

BMJ Open Assessing the performance of the Asian/Pacific islander identification algorithm to infer Hmong ethnicity from electronic health records in California

May Ying N Ly ¹, Katherine K Kim ², Susan L Stewart ³

To cite: Ly MYN, Kim KK, Stewart SL. Assessing the performance of the Asian/Pacific islander identification algorithm to infer Hmong ethnicity from electronic health records in California. *BMJ Open* 2019;9:e031646. doi:10.1136/bmjopen-2019-031646

► Prepublication history and additional material for this paper are available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2019-031646>).

Received 13 May 2019

Revised 14 November 2019

Accepted 19 November 2019



© Author(s) (or their employer(s)) 2019. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

¹Metropolitan Studies, University of North Carolina at Charlotte, Charlotte, North Carolina, USA

²Betty Irene Moore School of Nursing, University of California Davis, Davis, California, USA

³Department of Public Health Sciences, Division of Biostatistics, University of California Davis, Sacramento, California, USA

Correspondence to

Dr May Ying N Ly;
mly6@uncc.edu

ABSTRACT

Objective This study assesses the performance of the North American Association of Central Cancer Registries Asian/Pacific Islander Identification Algorithm (NAPIIA) to infer Hmong ethnicity.

Design and setting Analyses of electronic health records (EHRs) from 1 January 2011 to 1 October 2015. The NAPIIA was applied to the EHR data, and self-reported Hmong ethnicity from a questionnaire was used as the gold standard. Sensitivity, specificity, positive (PPV) and negative predictive values (NPVs) were calculated comparing the source data ethnicity inferred by the algorithm with the self-reported ethnicity from the questionnaire.

Participants EHRs indicating Hmong, Chinese, Vietnamese and Korean ethnicity who met the original study inclusion criteria were analysed.

Results The NAPIIA had a sensitivity of 78%, a specificity of 99.9%, a PPV of 96% and an NPV of 99%. The prevalence of Hmong population in the sample was 3.9%.

Conclusion The high sensitivity of the NAPIIA indicates its effectiveness in detecting Hmong ethnicity. The applicability of the NAPIIA to a multitude of Asian subgroups can advance Asian health disparity research by enabling researchers to disaggregate Asian data and unmask health challenges of different Asian subgroups.

BACKGROUND

Numerous studies have pointed to substantial racial/ethnic health disparities in the USA.^{1–3} The US racial and ethnic minority population, including immigrants, has doubled in the last decade.⁴ Combined with the increasing trend of chronic diseases such as diabetes, obesity, cardiovascular disease (CVD) and cancer, racial/ethnic health disparities have become an increasingly important and difficult area of public health research and practice. Yet, there are few baseline data that indicate and highlight specific health challenges faced by many racial and ethnic groups. Newer racial/ethnic groups that were immigrants and refugees to the USA are the most disadvantaged because of differences in migratory patterns (many coming from war-torn countries),

Strengths and limitations of this study

- The North American Association of Central Cancer Registries Asian/Pacific Islander Identification Algorithm (NAPIIA) is highly sensitive (78.0%) and specific (99.9%) as a tool for researchers to use to infer Hmong ethnicity from electronic health record data.
- Demonstration of the application of the NAPIIA algorithm using electronic health records data is relevant to researchers across the nation as these systems are in widespread use at healthcare institutions throughout the country.
- The NAPIIA did not identify some Hmong patients with surnames that are common in other Asian subgroups. This limitation can exclude a substantial number of ethnic Hmong in any application of the algorithm.
- Married Hmong women without Hmong surnames maybe excluded by the algorithm.
- The use of data from one institution, although representing a substantial proportion of the target Hmong community within California, may not be representative of the broader population of Hmong in the USA, thus limiting generalisability of the findings to other regions.

dietary habits, cultural practices and socioeconomic status. There are major differences between those who immigrate to the USA normally and those who come to the USA as refugees. Immigrants typically are more educated, have resources at their disposal and are prepared to adjust to life in the USA, versus refugees who have fewer resources and have less knowledge to help them adjust to life in the USA. The Hmong are among the Southeast Asian refugees entering the USA beginning in 1976. Today, many health challenges faced by Hmong Americans are still not known because they are often aggregated with other Asians.^{5 6}

There are over 260 000 Hmong Americans among the 17.3 million Asian Americans in



the USA.⁴ Since their initial migration to the USA, the Hmong population has substantially grown. Between 2000 and 2010, the Hmong population grew over 50% as compared with the growth of the US population at 9.7%.⁷ The Hmong are among one of the most socioeconomically disadvantaged racial/ethnic groups. One in four Hmong Americans lives below the federal poverty line, a rate that exceeds that of Latinos and African Americans, and as with many racial/ethnic groups, language access is a major barrier to optimal health and services.⁷ According to Asian Americans Advancing Justice (2011), 43% of Hmong older than 5 years old are limited English proficient, and the percentage who speak a language other than English is 91%.

Hmong Americans are included as Asians under the Federal Office of Management and Budget's classification category.⁸ As a whole, Asians have reported better health as compared with other racial/ethnic populations¹⁹; however, closer examination at the subgroup level indicates wide disparities between subgroups.^{10–12} For example, in California, Vietnamese women have had the highest incidence and mortality rate from cervical cancer compared with other Asian ethnic groups and Koreans have a singularly high incidence of stomach cancer.¹³ Filipino men had the highest incidence and mortality rates for prostate cancer among Asian ethnic groups,¹³ but Korean men have the fastest growing incidence rate for prostate cancer.¹⁴ A review of national mortality records for the six largest Asian groups (Asian Indian, Chinese, Filipino, Japanese, Korean and Vietnamese) shows cancer to be the leading cause of death for all Asian American females collectively (as aggregated Asians) and all Asian female subgroups (disaggregated) except for Asian Indian women. The leading cause of death for Asian Indian women was heart disease.⁹ Aggregating Asians as a whole can mask the real health challenges of the diverse Asian subgroups. When Hmong are aggregated with other Asians, the health disparity faced by this subgroup may be masked.

The number of studies on Hmong health disparity is limited. No longitudinal studies in this population have been conducted focusing on disease trajectories. A literature search on research among Asians by Smalkoski *et al*⁶ resulted in 2681 articles that found only 4% were on Hmong and cancer, 3% were on Hmong and diabetes and 5% on Hmong health inequality and disparities. No studies have been conducted on strokes or CVD among the Hmong; however, one study found Hmong in California, the state with the largest Hmong population, to have greater prevalence of advanced disease.¹⁵ A study by Lee and Vang¹⁶ showed mortality rates among Hmong for stomach cancer to be 3.5 times higher than Asian Americans overall. Additionally, mortality rates for liver and cervical cancer were 3–4 times higher for Hmong women as compared with other Asian Americans and Pacific Islanders in California.¹⁵ Fifty-one per cent of Hmong women in the study also chose no treatment, compared with 5.8% for Asian/Pacific Islander women and 4.8%

for non-Hispanic whites.¹⁵ Obesity among Hmong Americans is on the rise. Hmong are shown to have higher prevalence of diabetes and higher proportion above the standard and Asian body mass index (BMI) cut-points for diabetes screening as compared with Korean, Vietnamese and Chinese.¹⁷ Franzen and Smith¹⁸ studied acculturation and environmental change impact on dietary habits and found that among Hmong men, the mean BMI was 32.4 kg/m² and among Hmong women, the mean BMI was 29.1 kg/m². A local study in Wisconsin reported that 65% of Hmong youths were overweight or obese.¹⁹

A primary methodological obstacle to quality health research about Asian subgroups, including Hmong Americans, is the inconsistency of data on race and ethnicity.¹¹ In order to overcome this barrier to identify Asian subgroups, researchers have begun using surnames to infer ethnicity, developing algorithms that can use names to infer ethnicity.²⁰ In the USA, several name algorithms have been developed with this methodology used to identify racial and ethnic groups in the absence of available data on race and ethnicity.^{21 22} The North American Association of Central Cancer Registries (NAACCR), a consortium of cancer registries, governmental agencies, professional associations and private groups in North America interested in enhancing the quality and use of cancer registry data, developed the NAACCR Hispanic/Latino Identification Algorithm (NHIA) to directly or indirectly classify cases as Hispanic/Latino for analytic purposes. The NHIA expanded to include Asian Pacific Islanders in the NAACCR Asian Pacific Islander Identification Algorithm (NAPIIA).²¹ The NAACCR algorithms are available in SAS and can be downloaded at <https://www.naacr.org> for use.²³ The process flow of the NAPIIA algorithm is described in online supplementary appendix A.²¹

Both the NHIA and NAPIIA follow common methodological process and research components for using names to infer ethnicities, typically including: (1) a name reference list that is independently built or sourced from another study or expert in the specific ethnic culture and language, (2) a separate target population that is manually or automatically classified into ethnic groups and (3) a verified or self-identified ethnicity in the target population as a 'gold standard'.²⁰ The reference lists in the NAACCR algorithms (NHIA and NAPIIA versions) are based on surname lists from: the 2000 US Census; the Lauderdale and Kestenbaum name list (known as the 'Lauderdale List') and the NAACCR name list.²¹ The ethnic/racial groups in all three lists include the following: white, black, American Indian/Alaska Native, Chinese, Japanese, Filipino, Korean, Asian Indian, Vietnamese, other Asian, Hawaiian, Guamanian/Chamorro, Samoan and other Pacific Islanders. A Hmong name list was supplied by Richard Yang of the Cancer Registry of Central California.²⁴ The purpose of the current study is to assess the performance of the NAPIIA to infer Hmong ethnicity from electronic health records (EHRs) provided by the UC Davis Healthcare System from 2011 to 2015.

METHOD

UC Davis Health has more than 31 000 patients admitted to the hospital and/or seen at satellite locations and clinics throughout the region annually. UC Davis Health serves patients from the entire geographical region spanning Sacramento County and four surrounding counties: Yolo, Placer, El Dorado and Yuba/Sutter. This project used a dataset from UC Davis Health that contained 83 090 EHRs for the time period from 1 January 2011 to 1 October 2015. The Affordable Care Act mandates that language assistance is provided to those who do not speak English.²⁵ UC Davis Health implemented the data collection on race, ethnicity and language spoken by patients as one of their objectives under the meaningful use objectives that is mandated by the Health Information Technology for Economic and Clinical Health Act.²⁶ This information is gathered by administering the RACE/ETHNICITY/LANGUAGE Questionnaire (online supplementary appendix B) to patients at the first point of contact; however, it is not mandatory for patients to complete the questionnaire. The questionnaire asked patients regarding their language preference, the racial and ethnicity categories the patient identifies with and whether or not the patient is Latino. The information gathered is entered into the EHR by clerical and admission staff. This self-reported ethnicity from the questionnaire is the ‘gold standard’ for this project. The inclusion criteria for the study included patients who were at least 30 years old and no older than 74, who had no current or prior CVD diagnoses, were not pregnant and had a lipid panel completed during the study period (the CVD study is reported elsewhere).

In Hmong society, Hmong clan names are used as Hmong surnames and there are 18 traditional Hmong clans.²⁷ Online supplementary appendix C lists the traditional Hmong 18 clan names. Hmong surnames typically do not deviate from this list of clan names due to religious and cultural practices; however, Hmong scholars, including this Hmong researcher, have found additional

Hmong surnames that were adopted through the Hmong diaspora.²⁸ Yang²⁸ found that many Hmong have adopted their grandfather’s name as their surname. We compiled a list of currently known adopted Hmong grandfather’s names as surnames and searched those names in our data set to include them. A limited list of adopted surnames using a grandfather’s name is included in online supplementary appendix C.

We ran the NAPIIA on all patient names in the data set containing the 83 090 EHRs to identify all Asian names, including Hmong, using the name classification SAS code in V.15.0 of the publicly available SAS program specified above. The total number of records identified by the algorithm as having Asian names was 11 851. We excluded 8601 Filipino, Japanese and Asian Indian or Pakistani names because the names of these ethnicities do not have any commonality with Hmong names. In addition, we excluded another 605 patient records that did not have a specific ethnicity in the questionnaire (ie, ‘decline to state’, ‘other’, ‘unavailable’ and ‘unknown’). The final data set with 2644 names includes Hmong, Vietnamese, Chinese, Korean and other Asian names for patients who self-reported ethnicity in the RACE/ETHNICITY/LANGUAGE questionnaire (our ‘gold standard’) (table 1). We compared the algorithm-identified Hmong to those who self-identified as Hmong from the questionnaire.

To evaluate the accuracy of the algorithm, we use the epidemiological measures of sensitivity (the proportion of those who reported Hmong ethnicity who were correctly identified as Hmong by the algorithm), specificity (the proportion of those who did not report Hmong ethnicity who were correctly labelled as non-Hmong by the algorithm), positive predictive value (PPV) (the proportion of those who were labelled as Hmong by the algorithm who reported Hmong ethnicity), and negative predictive value (NPV) (the proportion of those who were labelled as non-Hmong by the algorithm who did not report Hmong ethnicity) among all records with Asian names in the data

Table 1 Self-identified ethnicity demographic

	Ethnicity				
	Chinese	Korean	Vietnamese	Hmong	Other
N (%)	1363 (51.6)	285 (10.8)	628 (23.8)	100 (3.8)	268 (10.1)
Age					
≤45	410	80	260	57	81
≥46	953	205	368	43	187
Sex					
Male	631	171	277	53	81
Female	732	114	351	47	187
Language spoken (N, %)					
Non-English	212 (15.6)	81 (28.4)	194 (31.0)	38 (38.0)	24 (9.0)
English	1145 (84.4)	204 (71.6)	431 (69.0)	62 (62.0)	244 (91.0)

‘Other’ column aggregated all other ethnicities not listed.

**Table 2** Self-reported ethnicity compared with algorithm identified Hmong

	Gold star Hmong		
	Yes	No	Total
NAPIIA			
Hmong	78	3	81
Not Hmong	22	2541	2563
Total	100	2544	2644

NAPIIA, North American Association of Central Cancer Registries Asian/Pacific Islander Identification Algorithm.

set. We stratified the data by sex (male/female) and age (≤ 45 and ≥ 46) for time in the USA. Language was stratified into English and non-English and used as a proxy for recent immigrant status. We excluded nine records that did not have specific data on language. We computed exact binomial 95% CIs to measure the precision of these metrics in all strata using an online binomial calculator at <https://stattools.crab.org/Calculators/binomialConfidence.htm>. Numerators and denominators were entered for each measure to calculate the upper and lower confidence limits.

Patient and public involvement

This study did not involve any patients or the public in the design or planning of the study. This is a secondary data analysis using EHRs from a health system.

RESULTS

The demographic from the self-identified ethnicity is display in [table 1](#) for the sample used in our analyses. We included a breakdown by age, sex and language.

[Table 2](#) shows the two-by-two cross-tabulation of Hmong ethnicity (gold standard) and Hmong identified by the NAPIIA. The final data set contained 81 names identified as Hmong by the algorithm versus 100 patients who self-identified as Hmong. Both methods identified 78 Hmong ethnicity patients. The algorithm identified an additional four Hmong names that self-identified as Laotian from the questionnaire; however, we counted these as Hmong. This determination was based on the expertise of this researcher (MYNL) who is Hmong, recognising the Hmong spelling in the name and the historical context that labelled all ethnic minorities from the country of Laos as Laotians, which included Hmong. Among those identified as Hmong by the algorithm, only three records self-identified as Chinese, Vietnamese and Mien. Among those who self-reported Hmong ethnicity (gold standard) from the questionnaire, the algorithm classified 22 records as belonging to other ethnicities: Korean (3), Chinese (16) and Vietnamese (3).

[Table 3](#) shows the sensitivity, specificity, PPV, NPV and prevalence of Hmong reporting Hmong ethnicity in the sample.

[Table 4](#) lists the Hmong surnames the algorithm did not identify as Hmong but which were self-identified

Table 3 Evaluation using sensitivity, specificity, PPV, NPV results

Evaluation method	Result, %	(95% CI)
Sensitivity	78.0	(0.69 to 0.86)
Specificity	99.9	(0.997 to 1.00)
PPV	96.3	(0.90 to 0.99)
NPV	99.1	(0.987 to 0.994)
Prevalence of Hmong	3.9%	

NPV, negative predictive value; PPV, positive predictive value.

as Hmong from the questionnaire. One record had a common Hmong surname (Xiong) and was identified by the algorithm as Hmong; however, the patient did not self-report as Hmong on the questionnaire (self-reported as Chinese).

In [table 5](#), the sex stratification (female vs male) to account for missing data on married women show that the algorithm performed only slightly better for men. The stratification for age (≤ 45 and ≥ 46) and for language (English and non-English speaking) use as a surrogate for US-born/long time resident versus more recent immigrant showed similar results with no substantive difference in how the algorithm performed.

DISCUSSION

To the best of our knowledge, this may be the first study that has attempted to validate the NAPIIA to infer Hmong American ethnicity in health research. The NAPIIA is highly sensitive and specific as a tool for researchers to use to infer Hmong ethnicity in the absence of racial/ethnic data. There are some limitations to the algorithm. Married Hmong women who adopt the non-Hmong surname of their spouse will not be identified by the algorithm. We stratified our sample by sex (female vs male), age (≤ 45 , ≥ 46) and language (proxy for time in the USA and recent immigrants) and found no substantive differences in the performance of the algorithm. We found that

Table 4 Hmong surnames identified as other ethnicity by algorithm

Hmong surname	Other ethnicity	No of EHR cases
Chang	Chinese	3
Fang	Chinese	1
Lo	Chinese	1
Lee	Chinese	1
Lee	Korean	1
Vang-Her	Chinese	1
Yang/Yah	Chinese	5
Cha	Korean	1
Heu	Korean	1

EHR, electronic health record.

Table 5 Stratification for sex, age and language

Frequencies and sensitivity by gender

Frequencies	Female gold star Hmong			Male gold star Hmong		
	Yes	No	Total	Yes	No	Total
Hmong	39	2	41	39	1	40
Not Hmong	8	1438	1446	14	1103	1117
Total	47	1440	1487	53	1104	1157

Evaluation metrics

	Measure	(95% CI)	Measure	(95% CI)
Sensitivity	0.83	0.69 to 0.92	0.74	0.60 to 0.85
Specificity	0.999	0.995 to 1.00	0.999	0.995 to 1.00
PPV	0.95	0.83 to 0.99	0.98	0.87 to 1.00
NPV	0.994	0.989 to 0.997	0.987	0.979 to 0.993

Frequencies and sensitivity by age group (≤ 45 , ≥ 46)

Frequencies	Gold star Hmong ≤ 45			Gold star Hmong ≥ 46		
	Yes	No	Total	Yes	No	Total
Hmong	46	1	47	32	2	34
Not Hmong	11	829	840	11	1712	1723
Total	57	830	887	43	1714	1757

Evaluation metrics

	Measure	(95% CI)	Measure	(95% CI)
Sensitivity	0.81	0.68 to 0.90	0.74	0.59 to 0.86
Specificity	0.999	0.993 to 1.00	0.999	0.996 to 1.00
PPV	0.98	0.89 to 1.00	0.94	0.80 to 0.99
NPV	0.987	0.977 to 0.993	0.994	0.989 to 0.996

Frequencies and sensitivity by language

Frequencies	Gold star Hmong English speaker			Gold star Hmong non-English speaker		
	Yes	No	Total	Yes	No	Total
Hmong	50	2	52	28	1	29
Not Hmong	12	2022	2034	10	510	520
Total	62	2024	2086	38	511	549

Evaluation metrics

	Measure	(95% CI)	Measure	(95% CI)
Sensitivity	0.81	0.69 to 0.90	0.74	0.57 to 0.87
Specificity	0.999	0.996 to 1.00	0.998	0.989 to 1.00
PPV	0.96	0.87 to 1.00	0.97	0.82 to 1.00
NPV	0.994	0.990 to 0.997	0.98	0.96 to 0.99

NAPIIA, North American Association of Central Cancer Registries Asian/Pacific Islander Identification Algorithm; NPV, negative predictive value; PPV, positive predictive value.

half of the traditional 18 clan names (surnames) were also used by other Asian subgroups (see [table 4](#)). The false negative was substantially higher than the false positive in our sample, resulting in an under count of Hmong patients. The algorithm did not identify any patients with the ‘Lee’ and ‘Yang’ surnames as Hmong. We expected this because of the commonality of both the Lee and Yang surnames in other ethnic groups; however, these surnames are also the most common surnames within the Hmong community. This is a serious limitation, guaranteeing that

a substantial number of ethnic Hmong will be excluded in any application of the algorithm. When the algorithm is applied in a different population with similar Hmong names but with a smaller prevalence of Hmong ethnicity in the sample, the PPV would be expected to decrease and NPV to increase. The spelling variation of names can also be a limitation.

In our review of the literature, we found an increasing trend of Hmong adopting a grandfather’s names as new surnames, for example: Lyfoung, Lynoulu, Bliatout ... (see



online supplementary appendix C). The utilisation of Hmong grandfather names is new in the Hmong community. We suggest updating the reference list with the new grandfather surnames in the algorithm as they are being adopted. The use of a Hmong expert to verify this variation and other variables that can validate a Hmong name, for example, the distinct spelling of a Hmong name, enhanced our ability to identify Hmong names. Additionally, we found the EHR to be inconsistent in how data on race, ethnicity and language were collected despite policies encouraging the data collection including the administration of a race/ethnicity questionnaire.

The inconsistency of gathering race/ethnicity data in EHR is a continual challenge. The NAPIIA includes the utilisation of maiden name as an additional variable; however, the EHR we examined did not contain maiden names. The absence of maiden names information in the EHR examined could lead to underestimation of the sensitivity of NAPIIA to infer Hmong ethnicity. Furthermore, although the race/ethnicity questionnaire was administered to every patient, it was not mandatory for every patient to provide a response. Therefore, this information may not have been captured in the EHR or captured correctly. Adhering to completion of the questionnaire for every Hmong person and entering every questionnaire into the EHR could increase the sensitivity of the method. However, the resulting estimation could be biased because those who do not complete the questionnaire might identify with a different ethnicity other than Hmong.

CONCLUSION

The applicability of the NAPIIA to a multitude of Asian subgroups can advance Asian health disparity research by enabling researchers to disaggregate Asian data and unmask health challenges of different Asian subgroups. The use of name lists to infer ethnicity is a useful tool; however, it cannot be used exclusively to identify ethnicity. EHRs need to include complete data on self-reported patient ethnicity in order to identify demographic patterns of disease risk and treatment outcomes, and to facilitate the development of individual and culturally tailored healthcare.

Twitter Katherine K Kim @kimkater

Acknowledgements The authors would like to thank their respective schools and data support personnel for their contribution to the management of the overall large data for the project. In addition, the first author would like to thank Gloria Wheeler for her support and guidance.

Contributors All authors (MYNL, KKK and SLS) contributed to the conceptualisation and design of the work, and reviewed and approved the final manuscript. MYNL drafted the manuscript. SLS assisted in the acquisition and analysis of the data. KKK assisted in editing the manuscript. All authors contributed to editing and review the manuscript critically for important intellectual content; and final approval of the version to be published. All authors agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Funding This research was made possible through a doctoral scholarship provided by the University of California at Davis, Betty Irene Moore School of Nursing and

the Gordon and Betty Moore Foundation (MYNL). Partial support is provided through a postdoctoral fellowship for the main author with the Women + Girls Research Alliance at University of North Carolina at Charlotte (MYNL).

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval The UC Davis Institutional Review Board approved this research on 11 May 2016 (# 8 93 313–1).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data may be obtained from a third party and are not publicly available. All data relevant to the study are included in the article or uploaded as supplementary information.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

May Ying N Ly <http://orcid.org/0000-0002-6733-2355>

Katherine K Kim <https://orcid.org/0000-0001-5766-3938>

Susan L Stewart <http://orcid.org/0000-0002-1593-440X>

REFERENCES

- Adler NE, Rehkopf DH. U.S. disparities in health: descriptions, causes, and mechanisms. *Annu Rev Public Health* 2008;29:235–52.
- Bach PB, Cramer LD, Warren JL, et al. Racial differences in the treatment of early-stage lung cancer. *N Engl J Med* 1999;341:1198–205.
- Barr DA. *Health disparities in the United States: social class, race, ethnicity, and health*. JHU Press, 2014.
- Bureau UC. *American factfinder*. US Census Bureau Washington, DC, 2010.
- Harrison GG, Kagawa-Singer M, Foerster SB, et al. Seizing the moment: California's opportunity to prevent nutrition-related health disparities in low-income Asian American population. *Cancer* 2005;104:2962–8.
- Smalkoski K, Herther NK, Ritsema K, et al. Health disparities research in the Hmong American community: implications for practice and policy. *Hmong Studies Journal* 2012;13.
- Justice AAA. *A community of contrasts: Asian Americans in the United States*. Washington, DC: Asian Americans Advancing Justice, 2011.
- Budget USOOma. Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity. In: *Budget OoMa*, 1997: 58782–90.
- Hastings KG, Jose PO, Kapphahn KI, et al. Leading causes of death among Asian American subgroups (2003–2011). *PLoS One* 2015;10:e0124341.
- Ghosh C, People H. *Healthy People 2010 and Asian Americans/Pacific Islanders: Defining a Baseline of Information*. *Am J Public Health* 2003;93:2093–8.
- Srinivasan S, Guillermo T. Toward improved health: disaggregating Asian American and native Hawaiian/Pacific Islander data. *Am J Public Health* 2000;90:1731.
- Tanjasiri SP, Kagawa-Singer M, Foo MA, et al. Designing culturally and linguistically appropriate health interventions: the "Life Is Precious" Hmong breast cancer study. *Health Educ Behav* 2007;34:140–53.
- McCracken M, Olsen M, Chen MS, et al. Cancer incidence, and associated risk factors among Asian Americans of Chinese, Filipino, Vietnamese, Korean, and Japanese ethnicities. *CA Cancer J Clin* 2007;57:190–205.
- Gomez SL, Noone A-M, Lichtensztajn DY, et al. Cancer incidence trends among Asian American populations in the United States, 1990–2008. *J Natl Cancer Inst* 2013;105:1096–110.
- Yang RC, Mills PK, Riordan DG. Cervical cancer among Hmong women in California, 1988 to 2000. *Am J Prev Med* 2004;27:132–8.
- Lee HY, Vang S. Barriers to cancer screening in Hmong Americans: the influence of health care accessibility, culture, and cancer literacy. *J Community Health* 2010;35:302–14.
- Stewart SL, Dang J, Chen MS. Diabetes prevalence and risk factors in four Asian American communities. *J Community Health* 2016;41:1264–73.

- 18 Franzen L, Smith C. Acculturation and environmental change impacts dietary habits among adult Hmong. *Appetite* 2009;52:173–83.
- 19 La Crosse County Health Department W. *Result of Hmong health screenings*. Department LCCH, 2001.
- 20 Mateos P. A review of name-based ethnicity classification methods and their potential in population studies. *Popul Space Place* 2007;13:243–63.
- 21 Boscoe FP. NAACCR Asian/Pacific Islander Identification Algorithm [NAPIIA v1. 2.1]: Enhancing the Specificity of Identification; 2009.
- 22 El-Sayed AM, Lauderdale DS, Galea S. Validation of an Arab name algorithm in the determination of Arab ancestry for use in health research. *Ethn Health* 2010;15:639–47.
- 23 NAACCR R. *Ethnicity Work Group: Naacccr guideline for enhancing hispanic/latino identification: Revised naacccr hispanic/latino identification algorithm [nhia v2. 2.1]*. Springfield IL, North American Association of Central Cancer Registries, 2011.
- 24 Yang RC, Mills PK, Nasser K. Patterns of mortality in California Hmong, 1988–2002. *J Immigr Minor Health* 2010;12:754–60.
- 25 Hellerstedt WL. The Affordable care act. *Healthy Generations* 2013;1.
- 26 Steinbrook R. Health care and the American recovery and Reinvestment act. *N Engl J Med* 2009;360:1057–60.
- 27 Duffy J. The Hmong: an introduction to their history and culture Center for Applied Linguistics, Cultural Orientation Resource Center; 2004.
- 28 Yang K-L. Problems in the interpretation of Hmong surnames. Hmong/Miao in Asia; 2004: 179–216.