INSTITUTE *for* ASIAN AMERICAN STUDIES

Information on Small Populations with Significant Health Disparities

A Report on Data Collected on the Health of Asian Americans in Massachusetts

Carolyn Wong, Hannah Hosotani, and John Her

November 2012

UMASS BOSTON

Information on Small Populations with Significant Health Disparities: A Report on Data Collected on the Health of Asian Americans in Massachusetts

Carolyn Wong, Ph.D.
Institute for Asian American Studies
University of Massachusetts Boston

Hannah Hosotani, B.A.
Mayo School of Health Sciences

John Her
Institute for Asian American Studies
University of Massachusetts Boston

UMASS
BOSTON

About the Institute for Asian American Studies

The Institute for Asian American Studies at the University of Massachusetts Boston conducts community-based research on Asian American issues, provides resources to Asian American communities in Massachusetts, and expands opportunities on campus for the study of Asian American experiences.

Institute for Asian American Studies
University of Massachusetts Boston
100 Morrissey Boulevard
Boston, MA 02125-3393
(617) 287-5650
www.iaas.umb.edu
asianaminst@umb.edu

Executive Summary

In Massachusetts, public health data collected at both the state and local levels remain largely inadequate for the systematic study of health disparities faced by Asian Americans (not including Native Hawaiians or other Pacific Islanders).[1] The annual health survey conducted statewide by the Massachusetts Public Health Department, the Behavior Risk Factor Surveillance System (BRFSS), does not obtain large enough samples of Asian American respondents to draw conclusions about many health disparities (see p. 8). Fortunately, there is a growing scholarly literature on how to improve the collection and interpretation of data on the very diverse and geographically dispersed Asian American ethnic subgroups in the U.S. This paper draws from this literature to evaluate the suitability of several publicly accessible sources of data on Asian Americans in Massachusetts. We also recommend measures to improve the store of relevant data in the state.

Public health agencies in Massachusetts have made important strides by creating standards that disaggregate health data, breaking out specific ethnicity and country of origin categories. In the annual BRFSS health survey named above, the small numbers of Asians contacted by randomly selecting households have made it hard for health agencies to justify translating the survey to Asian languages. They have focused their efforts on increasing the sample of Asian Americans by oversampling geographic areas, particularly small cities and towns, where relatively more Asian Americans reside.

But beyond oversampling in large surveys of randomly selected households in cities and towns where Asian Americans are concentrated, more innovative survey approaches are needed. In particular, new survey designs should account for geographic variability, including neighborhood-level concentration of some parts and wide dispersion of other parts of the ethnic populations. Studies on health disparities clearly demonstrate that residential contexts are an important factor in explaining health disparities (Do, Finch et al. 2008), and the theme is now familiar in calls for research on health disparities and inequities issued by nonprofit policy institutes, such as the "Place Matters 2012 National Health Equity Conference" sponsored by the Health Policy Institute of the Joint Center for Political and Economic Studies. Certain place-based factors affect the health of residents, including degrees of social isolation, the quality of schools, the suitability and accessibility of neighborhood-based services, viability of local markets for consumer needs (including for food and other daily supplies), and physical environmental quality.

In states like Massachusetts where Asian Americans are not numerous but where the population grew at a remarkable rate of 46% between the years 2000 and 2010, dedicating

---

[1] In this paper we refer to "Asian Americans" as a category that does not include Native Hawaiians or other Pacific Islanders, since Asian Americans and Pacific Islanders are classified separately in data standards of the Massachusetts Public Health Department and the U.S. Census. The Office of Minority Health of the Department of Health and Human Services, on the other hand, has recently issued reports on "Asian Americans, Native Hawaiians, and Pacific Islanders" (AANHPI) as in Appendix A-1 and A-2.

resources to surveys of specific communities where ethnic populations are concentrated would be cost-effective, since a random household survey within such a community can obtain sizeable samples of persons of those ethnicities. It is also more effective to conduct the survey in languages spoken by respondents in the area, and to tailor questions and styles of contacting (whether by phone, mail, computer-assisted surveying at neighborhood sites, or other means) to factors specific to the community.

Our recommendations include the following (the complete list appears at the end of the report):

- Public policy makers and health organizations should conduct health surveys targeting and tailored to the specific local communities where many Asian Americans live. These community-targeted surveys should collect data on the array of socioeconomic and medical issues, as well as cultural understandings that are pertinent to health disparities. To pool the resources necessary for these enhanced data collection efforts, public agencies, private organizations, nonprofit groups, and academic research institutions should forge partnerships to design and implement new and better community-based health surveys.

- Placing more attention on design and implementation of community based surveys does not eliminate the need to get more accurate information on Asian Americans in the annual health surveys of randomly selected households already conducted in the state – namely the BRFSS. It will be important to continue and expand the positive efforts of the Massachusetts Public Health Department to oversample cities with concentrations of Asian American residents. As the number of Asian American persons who are contacted increases, the survey should include the major Asian language spoken by Asian American residents in Massachusetts – namely, Chinese, Korean, Vietnamese, and Cambodian.

- Community-based organizations should be enlisted in data collection efforts. These community groups can encourage the participation of community members in the design of health research methodologies such as survey instruments and in recruiting respondents participation in the research.

- In collection of public health data, it will be important continue to implement the Massachusetts standard of using finely grained ethnic categories that are distinct and recorded separately from racial and ancestry categories. In hospital datasets, when data on race, ethnicity, or country or birth consists of fill-in text, extractive tools such as MassCHIP should make such information accessible to users. Since many residents of neighborhoods with concentrations of low-wage immigrants do not regularly use acute care hospitals for routine screening and services, the collection mechanism currently mandated for acute care hospitals should be extended to cover data collection at community-based clinics.

- Public health data should regularly include standard information on socioeconomic status, including types of income and housing available to individuals and families, poverty levels, and education.

- Hospital and clinic staff responsible for intake of patient information on data collection forms require careful training about how to ask about and code racial-ethnic categories, as well as names of languages spoken. A variety of cultural factors that influence communication about health topics which may be sensitive or health terms often misunderstood or mistranslated cross-culturally.

Our report begins with a discussion of the important issues of data collection and reporting and then discusses the particular challenges of collecting and reporting on data in Massachusetts. Profiles of major datasets based on records for administrative entities are presented such as the Massachusetts Cancer Registry, hospital discharges, MassHealth, and Medicare, and mortality and natality records. This is followed by a description of major datasets based on population surveys such as the Behavioral Risk Factor Surveillance System (BRFSS) and the Youth Risk Behavior Survey (YRBS). We discuss how MassCHIP may be used as an extractive tool for researchers. And we conclude with recommendations for the Commonwealth of Massachusetts to better address the health disparities of Asian Americans.

Issues of Data Collection and Reporting

In recent years, a growing body of research has revealed significant disparities between the incidence of certain diseases and risk for illness among persons who identify as Asian American.

One telling statistic from a 2009-2010 study indicates, for example, that

> TB (tuberculosis) rates among Hispanics, blacks, and Asians were seven, eight, and 25 times greater, respectively, than among whites. Among persons with TB, approximately 95% of Asians, 75% of Hispanics, 34% of blacks, and 20% of whites were foreign-born (Center for Disease Control and Prevention 2011a).

On various health indicators, however, the picture greatly varies for different Asian American ethnic sub-populations. *When data collection or reporting aggregates all the subgroups under one umbrella, this method tends to mask differences between them.* Findings such as the following all reported by Barnes, Adams et al. (2008), do indicate the need for finely grained data collected from and reported for subgroups such as Filipinos, Japanese, or Asian Indians.

> Filipino adults (27%) and Japanese adults (25%) were more likely than Chinese (17%) or Korean adults (17%) to have ever been told they had hypertension.

> Asian Indian adults (9%) were about two times as likely as Korean adults (4%) to have ever been told they had heart disease.

> Vietnamese adults (13%) and Filipino adults (11%) were more likely to suffer from migraines or severe headaches than Chinese adults (7%). (Barnes, Adams et al. 2008)

Not all health surveys gather information on immigrant status, education and income, measure of poverty, primary language, and other factors known to affect health status. The interactive relationship between these and other variables as they affect different measures of health outcomes are hardly straightforward, and a long-term agenda for research is needed. The Office of Minority Health reports, for example, that on one indicator, "asthma prevalence," Asian Americans have lower rates of asthma than the white population in the U.S. in general, but on another "death rates from asthma," they had a 30% greater death rate than whites in 2007 (Pleis, Ward et al. 2010; Xu, Kochanek et al. 2010). In one survey, researchers found that in Boston Chinatown, U.S.-born Chinese American children had asthma at a greater rate than foreign-born children (Brugge, Lee et al. 2006). The mixed results points to unsolved puzzles. For example, the higher incidence of asthma among the U.S. born suggests that genetic markers are not the only factors leading to high rates of asthma; rather, environmental conditions in communities of residence are likely to matter. If community-specific factors are important, then more resources for collecting data at the community level are warranted.

HHS Issues New Guidelines

In November 2011 the U.S. Department of Health and Human Services (HHS) published new guidelines for collection of health data on race, ethnicity, sex, and primary language gathered under its auspices or funding resources (U.S. Department of Health and Human Services 2011b). This welcome step was influenced in part by the publication of the Institute of Medicine's 2002 landmark report, *Unequal Treatment* in 2002, after which the challenge of reducing racial and ethnic disparities in the quality of health care rose higher on the policy agenda of health care reformers (Smedley, Stith et al. 2002). The 2008 Affordable Health Care Act called for improvements in systems of collecting health data for underserved populations that face what are called "health disparities," including high incidence of illness, less access to care, or inequities in treatment options compared to the majority of Americans. In addressing Asian Americans particularly, Health and Human Services affirmed an important commitment as follows:

> This plan contains the measurable objectives that the Department of Health and Human Services will pursue to raise the visibility of Asian American, Native Hawaiian, and Pacific Islander (AANHPI) health issues, health care, and human services disparities. This plan is meant to be a first step in elevating AANHPI issues across the Department under the leadership of the Assistant Secretary for Health, Dr. Howard Koh. (White House Initiative on Asian Americans and Pacific Islanders 2011a).

At the federal agency level, the guidelines for improving health data on Asian Americans call for, first, Health and Human Services sponsored surveys to report on a set of self-identified racial sub-categories that are disaggregated further than in the standard formerly in use as promulgated by the Office of Management and Budget in 1997. Rather than aggregating Asian ancestry groups under the "Asian" category, an individual can now select one or more of several racial categories that roll up to the Asian umbrella: Asian Indian, Chinese, Filipino, Japanese, Korean, Vietnamese, or Other Asian. The following racial categories roll up to the Native Hawaiian or Other Pacific Islander category: Native Hawaiian, Guamanian or Chamorro, Samoan, and Other Pacific Islander (U.S. Department of Health and Human Services 2011d). Second, the guidelines called for data collection instruments (such as surveys or questionnaires used for health administrative purposes by hospitals or insurers) to ask for the individual's "primary language," which they identify on their own. Several longer-range goals spelled out by Health and Human Services address oversampling, language use in data collection, and reporting/viewing tools. The full statement of objectives, both immediate and longer-term is included in Appendix A-1, and specifications for disaggregated racial and ethnic categories are in Appendix A-2.

Practices that Lead to Wrong Conclusions: Omission, Aggregation, and Extrapolation

A recently published study usefully distinguishes three practices that are likely to give rise to erroneous conclusions about Asian American health: omitting, aggregating, and extrapolating from data pertaining to ethnic subgroups (Holland and Palaniappan 2012). It is not

uncommon for surveys on health or access to medical services to omit Asian Americans as a relevant category all together. For example, the Urban Institute collaborates with researchers at the University of Minnesota to conduct the Massachusetts Health Insurance Survey, which tracks the effects over time of near universal health care coverage on residents of the state. However, the categories reported for race-ethnicity include only "white, non-Hispanic, Non-white, non-Hispanic, and Hispanic," omitting Asian race-ethnicities (Long and Stockley 2010). In the reports of the Behavioral Risk Factor Surveillance System (BRFSS), which is conducted by the Massachusetts Department of Public Health, some important information is included on Asian Americans, including ethnic subgroups, such as information on hepatitis b vaccination, asthma, dental care, overweight or obesity status, and alcohol use. However, other information is not available for Asian Americans as it is for other racial-ethnic groups, such as data on an individual's self-rated assessment of perceived health, diabetes, heart disease, cancer diagnosis, smoking cessation, flu vaccination, human papilloma virus (this list only gives illustrations and is not exhaustive). Sometimes data is omitted to protect the privacy of respondents when the sample is small (Massachusetts Department of Public Health 2011a). In a strong cautionary note, the survey administrators correctly point out that many estimates they provide for the small ethnic subgroups have large relative standard errors. This raises doubts about statistical reliability (Barnes, Adams et al. 2008).

The reasons for seeking finer grained health data on Asian Americans are well known among professional health researchers. As the Office of Minority Health of the U.S. Department of Health and Human Services reports "Chronic hepatitis B and liver cancer caused by hepatitis B in Asian Americans, Native Hawaiians and other Pacific Islanders (AA/NHOPI) comprise one of the most serious but frequently neglected racial and ethnic health disparities in the U.S." (U.S. Department of Health and Human Services 2011c). However, it is important to avoid extrapolating information that would lead to misleading conclusions, such as one that would attribute the same incidence of liver cancer, whether related to hepatitis B or not, to all Asian Americans. One study examined the incidence of liver cancer (without identifying a relationship between the liver cancer reported and hepatitis B) among Chinese, Filipino, Japanese, Korean, and Vietnamese men and women in the Greater San Francisco Bay Area between 2000 and 2004. The authors reported: "We found that incidence rates and trends varied among Asian ethnic subgroups, with non-significant declines in Chinese men, Japanese men, and Japanese women, but consistently high rates in Vietnamese, Korean, and Filipino men and women. Thus, liver cancer continues to affect Asian/Pacific Islander Americans disproportionately, with consistently high incidence rates in most ethnic subgroups" (Chang, Keegan et al. 2007).

When data from all ethnic subgroups are aggregated under the Asian American umbrella, chronic and even life-threatening health problems may be masked. Holland and Palaniappan (2012) note, for example, that the 2009 National Health Interview Survey (NHIS) reported that among Asian Americans as an aggregate group, there is a lower prevalence of heart disease than among non-Asian Americans; however, the authors cite

epidemiologic studies that demonstrate "a greater prevalence of coronary heart disease for Asian-Indian and lower prevalence for Chinese subjects compared with NHW's (non-Hispanic whites). There are similar differences in the prevalence of diabetes between Asian-Indian and Filipino research subjects.

Another pitfall lies in extrapolating findings about one ethnic subgroup to the entire Asian American population. To give on example, Holland and Palaniappan (2012) point out that reports and media took a study's findings that environmental factors affect the prevalence of cardiovascular disease, comparing rates in Japan, Honolulu, and San Francisco. In the latter two U.S. cities, Japanese men had a lower rate of heart disease than the non-Hispanic white population. But one media headline read "Asians have lower rates of heart disease," when as noted in the preceding paragraph, Asian Indians have a higher rate.

Capacity = More Data, Detail, Accuracy + Tools, Training, Partnerships

Public and private entities dedicated to health disparities research face high hurdles in their quest to attain better health data on Asian American populations. Appropriately, the new 2011 Health and Human Services guidelines begin by underscoring the need to "Increase the capacity to conduct more reliable health data and research throughout the U.S. and U.S. affiliated jurisdictions for Asian American populations to better describe and understand the need of the Asian American growing population as part of the Affordable Care Act." (See Appendix A-1.) Our investigation of the primary datasets available for researchers' use in Massachusetts, as reported below, has led us to emphasize a broader definition of "capacity" than the mere collection of more data and more detailed and accurate data. Certainly the collection stage is of prime importance, and the current capacity of public and private entities to collect data on Asian Americans lags behind the need. The data must be disaggregated according to ethnicity and in sample sizes sufficient to draw statistically reliable inferences. But capacity to conduct more reliable research encompasses more. The notion of sufficient capacity includes the availability of computer tools to extract the data in useful ways once it is collected. Building capacity includes the training of health researchers and providers in the use of this data in ways sensitive to the culture and social milieu of diverse Asian American populations living across the vast territory of the United States, in urban, suburban, exurban, and rural settings. And to achieve sufficient capacity for reliable research on health disparities faced by Asian Americans, ongoing partnerships between public health agencies, academic researchers, community-based organizations, health care providers and insurers are needed.

One of the hardest challenges faced by health care advocates concerned about Asian Americans is gaining recognition of the distinctive and pressing health needs of these populations so that real capacity for reliable research can be expeditiously built. The Asian American populations are often perceived by policymakers as marginal in most U.S. states and localities because of their relatively small size. In Massachusetts, among a population of 6,547,929 persons, individuals identifying as Asian American alone or in combination with

other races was 6% of the total number. The comparable percentages for other groups were: whites: 82.5%; blacks: 7.8%; Hispanics: 9.6% (Hispanics are classified as an ethnic group, and individuals may identify with any racial category), American Indian/Alaska Natives: 0.8%; Native Hawaiian/Pacific Islanders: 0.2 percent.[2]

Growth of the Asian American Population

Yet these 2010 population counts do not tell the whole story. Asian Americans comprise the fastest growing racial minority group in the U.S. as a whole, with numbers increasing by nearly 46% between 2000 and 2010 for those who identify as Asians alone or in combination with another race, and about 43% for those identifying as Asian alone. In Massachusetts over that time frame, the increase of Asian Americans identifying as Asian alone or in combination with other races was 48.9%, which was slightly faster than the comparable 46.4% rate for Hispanics or Latinos, who may identify with any race.[3] The expanding presence of Asian Americans at both the national and state population levels should draw greater attention to their needs by health policy and medical scientific researchers concerned not only about minorities but the health of all U.S. citizens and residents.

In the example of hepatitis B incidence mentioned above, it is worth noting that at least one of the most noteworthy health disparities is associated with some of the smaller groups of Asian Americans, the Vietnamese, Cambodians, and Hmong. In this case hepatitis B in this population is not limited to the U.S., but it is relatively widespread in Southeast Asia, and it affects groups that immigrated recently and have lived in the U.S. for a shorter duration of time than others (such as the Chinese or Japanese); hence, it is not surprising that since these comprise some of the smaller Asian American populations, a casual glance at the Asian American umbrella through the lens of the larger group will obscure the health problems of smaller groups.

Changing Concepts of Who is Asian American

 There is a further reason to gather more finely grained data on race and ethnicity, including information on how individuals may identify with more than one race. The Pew Research Center used data from the 2010 American Community Survey to analyze intermarriage rates, finding that the percentage of new marriages between partners of a different race or ethnicity increased to 15.1% in 2010, and the percentage of current marriages that are interracial or interethnic attained a high of 8.4%, greater than at any other time. Notably, the Pew Research Center reports that in 2010 among newlyweds married to persons of another race or ethnicity, Asian Americans comprised 27.7 percent, while the corresponding figure for Hispanics was 25.7%; for blacks it was 17.1%, and for whites it was 9.1% (Pew Research Center 2012). The relatively large rate of intermarriage among Asian Americans may have

---

[2] These figures (as are those in the next paragraph) are calculated from U.S. Census statistics and rounded to the first decimal place.

[3] These figures are calculated from counts reported in the 2010 U.S. Census Summary File 1.

medical implications for the descendents of intermarried couples, and the long-term trends can only be understood if baseline and trend data are carefully gathered.

For Asian Americans, a core concern addressed by the guidelines announced by the Department of Health and Human Services in November 2011 is the validity of racial-ethnic categories – that is, whether the names of categories (such as Korean, Japanese, or Burmese) match how respondents view themselves when categorizing the world in their own cultural milieu. The "Asian American" category was first constructed in the late 1960s by academics and advocates to increase the political leverage of Asian American constituent groups. However, the term was neither in popular usage at that time, nor is it fully accepted today. Most Vietnamese, for example, when asked their ethnic identification will say it is "Vietnamese," not "Asian American." Interestingly, researchers found that people who are classified by the census as Asian American, Native Hawaiian, and Pacific American Islanders will answer a survey question about their racial-ethnic identity by naming an ethnic-specific category, indicating it is their primary identity. But if asked in a follow-up question whether "in addition" they identify as Asian Americans, they will answer "yes"[4] (Lien, Conway et al. 2004). It is well documented that Medicare and Medicaid datasets underestimate the number of Asian Americans participating in these programs because many persons who do not self-identify as Asian, better identify with "white" or "other"[5] (Agency for Healthcare Research and Quality 2008). Thus, one consideration in constructing valid racial-ethnic categories is sensitivity to ordinary people's understandings of these terms.[6] As mentioned above, standardization is important so that meaningful comparisons can be drawn about ethnic groups named in the same manner across various data-gathering projects.

---

[4] In the 2008 National Asian American Survey of 5,129 respondents, researchers found that 47% think of themselves as Ethnic American; 40% as part of their Ethnic Group; 21% identify as Asian Americans; and 19% as Asian. Respondents could select more than one label, so the percentages sum to more than 100 (Wong, Ramakrishnan et al. 2011, pp. 161-2).

[5] In this report, researchers at Research Triangle, International, have investigated the accuracy of Medicare Enrollment Database (EDB) for Medicare beneficiaries, comparing the algorithm based on their unique surname to improve their coding on the EDB against self-reported race/ethnicity obtained from respondents to surveys conducted by the Centers for Medicare and Medicaid Services. They found, "the sensitivity for Hispanic coding on the EDB was a low 30 percent and for Asian Islander it was 55 percent. . . . Those low sensitivities largely reflected self-identified Hispanics coded as White on the EDB, and self-identified Asian/ Pacific Islander coded as Other on the EDB." Sensitivity is a statistical measure of the proportion of "actual positives."

[6] Agency for Health Research and Quality has documented the distinctive challenges involved with classifying Filipinos. According to the Office of Management and Budget (OMB) standards, people of Filipino descent are classified as Asian. In the National Health Interview Survey, for example, there are response categories for "Asian Indian, Chinese, Filipino, Japanese, Korean, Vietnamese, and Other Asian." In recent samples, 89 percent of single-heritage Filipinos selected the Asian category when presented with the OMB-minimum categorization. The balance of 11 percent chose primarily Native Hawaiian or Pacific Island. In various other surveys, Filipinos may self identify as Spanish, Pacific Islander, Asian American, or, if multiracial, white (Agency for Healthcare Research and Quality 2010, Box 3-3).

Data Collection Standards—Race, Ethnicity, and Language

One helpful set of recommendations that has guided reform of data collection standards in recent years is summarized in *Race, Ethnicity, and Language Data: Standardization for Health Care Quality Improvement*, a report by the Subcommittee on Standardized Collection of Race/Ethnicity Data for Healthcare Quality Improvement of the Institute of Medicine. Recommendation (3-1) states:

> An entity collecting data from individuals for purposes related to health and health care should:
>
> • Collect data on granulate ethnicity using categories that are applicable to the populations it serves or studies. Categories should be selected from a national standard list (see Recommendation 6-1) on the basis of health and health care quality issues, evidence or likelihood of disparities, or size of subgroups within the population. The selection of categories should also be informed by analysis of relevant data (e.g., Census data) on the service or study population. In addition, an open-ended option of "Other, please specify:____" should be provided for persons whose granular ethnicity is not listed as a response option.
>
> • Elicit categorical responses consistent with the current OMB standard race and Hispanic ethnicity categories, with the addition of a response of "Some other race" for persons who do not identify the OMB race categories (Institute of Medicine 2009).

One of the problems that remains unsolved is how finely grained ethnic categories can be "rolled up" to larger classifications in a standard manner. Massachusetts established a set of standards that included granular ethnicities that are "locally relevant to the Commonwealth of Massachusetts" (Agency for Healthcare Research and Quality 2010).

The second type of measurement category that deserves close examination describes lifestyle or beliefs and attitudes that are related to health (such as on matters of diet; smoking; exercise; access to competent health services) or measurable facts about health status (such as whether a doctor informed patients that they had a specific illness).

Concepts of healthfulness or sources of illness differ across cultures, in general. Further, it is necessary to ensure cross-cultural validity of terminology about health-related concepts; for example, the meaning of the word "depression" in English is not always understood in the same way by persons embedded in different cultures (Tran and Ferullo 1997).

Making sure health-related terminology is cross-culturally validated is crucial when questionnaires are translated from one language to another. Researchers have devised some statistical tests to establish cross-cultural validation of questionnaires about health, and it is important that the cross-cultural validation process be undertaken competently.

When data collection methods fit standards of high reliability, the terms and choice of structures used to elicit answers produce consistent results. Such consistency depends in part on culturally appropriate and effective communication. There may be no greater obstacle to such communication than a language barrier: the inability of the respondent or questioner to understand or speak the language used to ask or answer questions. Thus, it is important to include questions about respondents' spoken languages so that research can discover associations between specific health outcomes and English language proficiency, assess language needs and ultimately improve language services. Accordingly, one recommendation (4.1) in *Race, Ethnicity, and Language Data* states:

- At a minimum, collect data on an individual's assessment of his/her level of English proficiency and on the preferred spoken language needed for effective communication with health care providers. For health care purposes, a rating of spoken English-language proficiency of less than very well is considered limited English proficiency.

- Where possible and applicable, additionally collect data on the language spoken by the individual at home and the language in which she/he prefers to receive written materials (Institute of Medicine 2009; p. 108).

As discussed below, in order to advance research on health disparities facing Asian Americans it will be important for public and private organizations to collaborate more fully in the creation of surveys that will be conducted in the Asian languages most commonly spoken, such as Chinese, Korean, and Vietnamese. As the capacity to do so in statewide surveys is built over time, it will be important to create surveys that focus on local Asian ethnic communities in the relevant languages. As noted above, in one national survey preparations were too far along at the time of this writing for National Health and Nutrition Examination Survey (NHANES) materials to be available in three Asian languages. The NHIS does not conduct its interviews in Asian languages.

The providers of public health data in Massachusetts can learn from the experiences of research capacity-building in other states. The California Health Interview Survey (CHIS**),** for example, is conducted in several Asian languages and collects sizeable samples of various Asian American ethnic subgroups. A strong partnership enabled the accumulation of resources to begin to meet the challenges of improving data collection on a California population that is very diverse in its racial-ethnic makeup. CHIS is conducted by the UCLA Center for Health Policy Research. Its principal partners are the California Department of Public Health and the Department of Health Care Services within the California Health and Human Services Agency. The funders include these two state agencies as well as several federal agencies, insurance companies, nonprofit foundations, and others (UCLA Center for Health Policy Research 2012).

Community Studies

Health disparities are correlated with high poverty rates, which are, in turn, correlated with reduced access to health care and insurance (National Cancer Institute 2008). The principle of affording equality of opportunity to all goes hand in hand with improving substandard health care and ending benign neglect of health problems of people living in poverty. According to standard poverty measures, Massachusetts and other states in the Northeast face particularly high rates of poverty for Asian Americans. And although, the umbrella category of Asian Americans (excluding Native Hawaiians and Pacific Islanders) nationally has a lower poverty rate than the national average, the rates are high among Hmong (37.8%), Cambodians (29.3%), Laotians (18.5%), and Vietnamese (16.6%) (White House Initiative on Asian Americans and Pacific Islanders 2011b).

Estimates of poverty rates from the 2005-2009 five-year American Community Survey provide striking news about poverty for Asian Americans in Boston: Persons living in Boston who identified racially as "Asian" in the American Community Survey Five-Year estimates for 2005-2009 had the highest poverty rate at 30.3% of any racial group (compared to "all," "white," "black," and "other"), although in Massachusetts as a whole the rate was only 13.8% (Albelda, Cadet et al. 2011).[7] By comparison, in Boston the poverty rate for the total population was 19.1%; for whites it was 13.0%; for blacks, it was 25.2%; and for the category designated "other" it was 27.4%.[8]

Some of the most revealing health disparities data comes from community studies, and it is instructive to highlight recent findings from a 2011 Vietnamese Adult Health Survey in Santa Clara, California. In the former category of economic disparities (measured along one dimension by percent of the population in poverty), some notable findings were:

> In 2011, Vietnamese adults in Santa Clara County cited finances and unemployment/ jobs (as well as health and health insurance) as top concerns facing their households. . . Roughly 1 in 10 Vietnamese families lived in poverty in 2007 to 2009, which was higher than for families in the county overall and for families of all other major racial/ethnic groups except Hispanics. Similar disparities were evident for educational attainment (Peddycord and Fenstersheib 2011).

With respect to health disparities,

> Vietnamese residents experience disparities in both chronic and infectious diseases

---

[7] First reported in Nov. 2011, the Supplemental Poverty Measure (SPM) is meant to capture variation in family costs of living and resources not measured in the standard measure of poverty, such as payroll taxes and in-kind benefits received from government program.

[8] In 2009, the poverty rate for Asian Americans was about 9%. Poverty rates for Asian Americans and Latinos have fluctuated in recent years, with the 2001 Asian American population in Boston possessing the highest rate living below the poverty line. Boston Public Health Commission (2010). Health of Boston 2010, Boston Public Health Commission.

relative to residents from other major racial/ethnic groups. For example, cancer was the leading cause of death among Vietnamese residents and accounted for a larger percentage of total Vietnamese deaths in 2011 than for all county residents or residents of all other major racial/ethnic groups.

Vietnamese adults had a higher incidence rate of (rate of new cases per 100,000 adults from 2007 to 2009) and mortality rate from several specific cancers than adults from other major racial/ethnic groups. Incidence and mortality rates for liver cancer were four times higher among Vietnamese adults than adults in the county as a whole. Vietnamese adults also had the second highest lung cancer incidence and mortality rates compared to other major racial/ ethnic groups. Vietnamese women had the second highest incidence rate of cervical cancer in 2007 to 2009 relative to women from other major racial/ethnic groups in the county. (Peddycord and Fenstersheib 2011).

The findings of the Santa Clara, California, survey of Vietnamese are relevant to communities of Vietnamese Americans in other cities, including the neighborhood of Dorchester, Boston, which is one of the largest urban concentrations of Vietnamese in the U.S. Through partnerships between government entities, medical researchers, and community providers, the relatively large communities of AANHPIs in California have successfully funded surveys that specifically focus on ethnic groups known to face social and health disparities.

In Massachusetts, on a much smaller scale, federal public health agencies and community organization have similarly collaborated to study the health of Cambodians. In 2010, a federal project sponsored by the Center for Disease Control, titled Racial and Ethnic Approaches to Community Health (REACH) funded a health survey in Lowell, Massachusetts, the second largest Cambodian community in the U.S. Researchers formed a coalition with community groups and employed survey sampling strategies where randomized lists of phone numbers or addresses located in neighborhoods with many Cambodian residents. From a sample of 381 Cambodian adults, 25 years and older, this survey found that 44% of respondents reported fair or poor health. Those most likely to report fair or poor health were "female, older, unable to work due to disability, to have spent a smaller proportion of their life in the U.S., and to have wanted to see a doctor in the past year, but not been able to" (Koch-Weser, Liang et al. 2006; p. 133).

Methods of Sampling

For surveys, the goal is to obtain, first, a representative sample (that is, a sample whose member characteristics are like those in the general population); and second, a sample large enough to offer statistical reliability to researchers seeking to infer information on central tendencies in the target population such as income level or percent of persons reporting they were told by a doctor they have tuberculosis. Various strategies are used to increase representativeness and sample size. To reduce costs, various multi-stage sampling strategies are well known, such as those that draw in the first stage on some fixed number of census tracts known to be the residential tracts where some minimum number of

persons with the target ethnicity live; in a second stage, some random selection of census blocks might be chosen; and finally, within the list of phone numbers of persons living on those blocks, only those whose surnames on a "Vietnamese" surname list would be contacted.

Two national surveys (as described below) the National Health Interview Survey (NHIS) and the National Health and Nutrition Examination Survey (NHANES), now oversample Asian American persons. Nevertheless, samples from these major federal datasets are not large enough for research that aims (1) to identify health disparities for specific ethnic subgroups or (2) to identify disparities in states with relatively small Asian American populations, like Massachusetts.

Islam, Kahn et al. illustrate the problem of small sample size for ethnic subgroups in a national survey with reference to the NHIS:

> In studies where there are sufficient samples of Asian Americans overall, it is often not possible to conduct meaningful subgroup analyses due to insufficient sample sizes by Asian ethnic group. For example, the sample size of individuals reporting Asian race in the 2005 NHIS sample was 3,748. The sample sizes for the subgroups available in NHIS, however, ranged from 261 to 704, which are sufficiently powered for bivariate analyses but are inadequate for more sophisticated multivariate models.[18] Even within the same survey, sample sizes can differ dramatically across years, with the trend of sample sizes decreasing over time. For example, NHIS interviewed 27% fewer households in 2007 than in 2005 (45,000 in year 2005; 33,000 in year 2007)[9] (Islam, Khan et al. 2010; p. 1369).

Islam, Kahn et al. also point out that unevenness of geographic representation further plagues research on health disparities (National Cancer Institute 2008; Islam, Khan et al. 2010; p. 1369). In many cases, the Asian American population groups are large enough to warrant focused research but do not attract enough attention from policymakers to attain funding in a competitive environment. Geographic areas with large ethnic concentrations are generally preferred opportunities for cost-effective research. In Massachusetts, oversampling of Asian Americans has occurred in Lowell, Massachusetts, but there is insufficient coverage of the Vietnamese, Chinese, Koreans, and others in the Boston metro area and neighboring towns and cities, such as Quincy and Lawrence.

In the conclusion to this report, we note of some successful strategies for surveying Asian American persons used by the California Health Information Survey at the state level, from which organizations dedicated to public health research in Massachusetts can learn.

Tradeoffs Between Finely Grained Ethnic Categories and Small Sample Size: There is an apparent tradeoff between the move to disaggregate data on Asian Americans into seven or more ethnic subgroups and the logical consequence, that is, sample sizes for the

---

[9] Islam, Khan et al. (2010) refer here to public use data release of the 2008 National Health Interview Survey.

narrower ethnic categories will tend to be smaller than if all Asian Americans are counted under the umbrella Asian American category as one racial-ethnic category alone. If data are collected for a relatively small group, such as the Cambodians, the costs may be prohibitive.

Native Hawaiians and Pacific Islanders are counted separately in the census, but for purposes of advocacy on health disparities, advocacy groups and researchers have included them with Asian Americans, while recognizing their distinctive situation. Now that ethnic groups will be identified separately within the umbrella, there is good reason for Native Hawaiians and Pacific Islanders to be included in discussions by governmental agency and nonprofit leaders of API concerns. Yet there are compelling reasons to collect data for Asian ethnic populations according to finely grained ethnic categories, even when the number of persons classified into these categories is small. First, when surveys are conducted fairly often, such as the yearly health surveys administered by state public health agencies and coordinated by the federal Health and Human Services Department, it is possible to pool samples across a few years and in this way increase the sample size. Second, there are good reasons to believe that health disparities are large for several of the smaller Asian American racial-ethnic populations. From an ethical standpoint, where health disparities are large for a population group that has established its residence in the U.S. and is making its contribution to economic and civic life in this country, public health agencies have an obligation to address these disparities. To understand and narrow the disparities requires precise health data accumulated in ways that will yield reliable statistical inferences (such as sufficiently large samples).

There are other compelling reasons to disaggregate ethnic data even if sample sizes will be small in many surveys of these groups. In contemporary times, public health research is decidedly global in scope and necessarily takes account of incidence and correlates of disease that crosses sovereign national borders. Within the U.S., metropolitan areas and rural towns are receiving a growing number of racial-ethnic minorities whose health status and access to medical services and insurance are becoming an increasingly important part of the overall public health picture of those communities.

Although small sample sizes generally reduce the statistical reliability of inferences, all else being equal, a little information about an understudied ethnic minority with large health disparities is better than none at all. If the data meet standards of validity and reliability as discussed above, we can acquire useful cues about an ethnic population's health status from data that do not yield statistically significant estimates. Also, exploratory research using initial findings can open up fruitful questions for more systematic research and justify the allocation of resources. For sure, data providers should clearly flag estimates of population parameters derived from small samples or with large variances in ways that can be easily understood by users. With these points as guideposts, we next discuss the current state of health data on Asian Americans in Massachusetts.

Massachusetts Collection and Reporting Practices

In Massachusetts as in all the American states, health data standards have evolved within the context of a complex federal system of health care research and delivery. When federal health agencies attempt to coordinate the gathering and archiving of health data, they often rely on state health agencies and their established data collection systems. Accordingly, improvements in national standards require cooperation between interlocking state and federal agencies.

In Massachusetts some of the Asian American populations with the largest health disparities are relatively small in number but have high poverty rates, such as the Cambodians and Vietnamese. Because these populations are small, they tend to fall under the radar screen of public health policymakers and are not among the research priorities of government-funded health projects, major research universities, insurers, and others. Much of the impetus to increase study of health disparities for racial and ethnic minorities has come from the demographic profile of Boston, where about half of the residents are members of racial and ethnic minority groups, compared to twenty percent in the state.

Administrative Health Data Collection

Massachusetts has taken positive steps to define a relatively flexible rubric for collecting health data from hospitals and vital records. Information on an individual's race, ethnicity, ancestry, and principal written and spoken language are collected. Options are provided for those who wish to self-identify. By contrast, the U.S. Census standards and the approach in other states subsume the concept of ethnicity under race; for example, the ethnic category of "Chinese," is considered a sub-racial category rather than an ethnicity or ancestry. In hospital data collected in Massachusetts, the conceptual independence of racial, ethnic, and ancestry categories enables individuals to self-report these dimensions of their identity in multiple ways, lending flexibility to self-categorization and richness to the data. For example, if an individual Cambodian person does not readily self-identify as "Asian" in racial terms (perhaps not recognizing the term), this individual has the opportunity to self-identify "Cambodian" as an ethnicity while perhaps leaving the racial category "unknown" or "other"; in addition, the person also has the opportunity to identify a place of birth or origin as "Cambodia." Researchers examining the relevant datasets can search for terms in multiple categories, as a result.

In addition, according to these standards, patients self-identify their race and ethnicity, rather than have hospital staff members attribute it, which often resulted in inaccuracies due to guesswork on the part of the staff. There are currently thirty-one ethnicity categories that reflect the diversity of the Massachusetts population. There are nonetheless substantial difficulties when it comes to implementing these standards, since hospital administrators have had to re-program computer tools that recorded admissions and discharges, train staff in new procedures, including in how to query patients about their self-identified race-ethnicity and then record it to fit into one of the thirty-one specified categories. With many

newly defined racial-ethnic categories, some hospitals inevitably have few numbers in some categories, and at the reporting and analysis stage, it is necessary to decide whether and how to amalgamate categories (Weinick, Caglia et al. 2007). The perception that the Asian American population is numerically less significant than others further slows implementation.

Reporting from Public Health Surveys

As noted, although some useful information on Asian American health is collected in Massachusetts health surveys, small sample sizes make it impossible to draw conclusions on important indicators. In 2011, the BRFSS sample size (nonweighted) of Asians was just over two percent (462 persons) of the total sample size compared to a population proportion of 6% of Asian Americans in Massachusetts (2010 U.S. Census). For a summary of responses given by respondents who identified under the umbrella racial category of "Asian American," the BRFSS reports some but not all data it collects, suppressing some information when sample sizes are so small as to potentially violate privacy of respondents. In light of these limitations, we make recommendations at the end of this paper that academic institutions, public agencies, and various private organizations all work together to oversample communities where many persons of Asian American ancestry live and to conduct local surveys in those communities in the principal Asian languages spoken by the residents of Asian ancestry. At this time, one limitation of the BRFSS is that it is not conducted in any Asian languages.[10]

The Boston Public Health Commission (BPHC) conducts an independent survey, the Boston Behavioral Risk Factor Surveillance System (BBRFSS), every other year. Surveying residents of Boston, the BBRFSS models its questionnaire and sampling methodology on the statewide (Massachusetts) Behavioral Risk Factor Surveillance System, although it is not part of the statewide BRFSS project. The reports of findings from Boston-based survey are especially useful for researchers interested in data disaggregated at the neighborhood level. The commission's 2011 report, *Health of Boston*, for example, usefully focuses on identifying groups of individuals and communities that face the greatest risk for poor health conditions, and it includes analysis of contextual factors for health outcomes, including income, education, employment, housing, exposure to racism and discrimination[11] (Boston Public Health Commission 2010).

The *Health of Boston* report provides some informative data on the health of Asian Americans.   For example, the 2011 report uses the 2010 BBRFSS to construct estimates of

---

[10] A spokesperson at the Massachusetts Department of Public Health informed the authors that early efforts to conduct this survey in Mandarin were not cost effective because four or fewer respondents among the small number of Asian respondents spoke Mandarin.

[11] In *Health of Boston* cited here, as shown in Figure 6.15a, p. 128; p.188, rates are not presented for Asians for 1999-2008 due to the small number of cases. Figure 12.6 Pap Test within the Past Year by Race/Ethnicity and Household Income, 2010; Figure 12.1 Mammogram within the Past Year by Selected Indicators, 2010, p. 201

health indicators for certain racial-ethnic groups, including the category of "Asians," such as the percentage of adults reporting high blood pressure, obesity, and certain mental health problems. The report also draws from the Communicable Disease Database held by the Boston Public Health Commission to report rates of salmonella infection and tuberculosis among Asians. In the annual reports, information on disaggregated ethnic subgroups is typically not provided. The sample size for Asian Americans is insufficient, moreover, to provide information on certain health indicators, such as diabetes among Asian American adults from 2004 to 2008. The health data related to cancer illustrates the strengths and weaknesses of the report. Because data on incidence of cancer and cancer mortality is recorded in cancer registries (not in population surveys, such as the BBRFSS), the authors of *Health of Boston* were able to take note of important trends in incidence of cancer and cancer mortality rates for Asian Americans. For example, the report states "all-cancer age-adjusted mortality rate for Boston's Asian population increased 43% from 2006 to 2008. Lung cancer was the leading type of cancer mortality among Boston's Asian residents each year from 2006 to 2008." In contrast, when reporting on cancer screening rates, the report relies on the survey data of the BBRFFS. In the 2010 survey, there was an insufficient sample size to determine rates of screening of Asian American women for cancer by yearly mammograms or Pap tests.

Among the agencies responsible for collecting public health data, the Boston Public Health Commission plays an important role. In addition to conducting the BBRFSS, for example, the commission also conducts surveillance of infectious diseases, lead screening, and sexually transmitted diseases. The BBRFSS is not included in our profile because details on the methodology for data collection are not available.

Profiles of Datasets

This section provides a profile of strengths and weaknesses of several widely used datasets insofar as they provide information on the health of Asian Americans in Massachusetts. Our selection does not exhaust by any means the universe of health-related data on Asian Americans as it is collected by various private or public entities, such as immigrant and refugee service agencies. But we selectively describe those datasets typically used to portray the health of Massachusetts residents. We classify the datasets into two general types: first, datasets created from patient records by administrative entities such as hospitals, other medical providers, insurers, organizations that create disease indices or registries, and local governments and second, datasets taken from population surveys. At the end of this section we present a table with detailed information on several datasets.

Among the features tabulated for each dataset are the (a) validity of racial-ethnic categories; (b) whether primary or preferred language is self-identified by the survey respondent or patient, as a facet of reliability; (c) whether information was ascertained about socioeconomic status, country of origin, immigrant status, and citizenship; and (d) for surveys, strategies used to sample small racial-ethnic populations.

Datasets Based on Records from Administrative Entities

Administrative Records: Medical

1) Massachusetts Cancer Registry (MCR**)** is a registry that records all newly diagnosed cases of cancer in the state. (Massachusetts Cancer Registry 2003)

2) Hospital Discharges is a database that consists of records submitted by acute care hospitals of all patients admitted. The Massachusetts Division of Health Care Finance and Policy has required specific fields to appear in each record, including demographic information, payer, diagnoses, caregivers' names, source of admission, discharge, or transfer, and other pertinent information. (Boston Public Health Commission 2007 ; Massachusetts Division of Health Care Finance and Policy 2007)

3) MassHealth is a public health insurance program for low and medium-income residents of Massachusetts. MassHealth is the name used for the program encompassing both Medicaid and the State Children's Health Insurance Plan (SCHIP). Collection of data on patients generally follows the standards set by the Uniform Hospital Discharge Dataset (UHDDS), a federal standard.

4) Medicare is public health insurance programs for individuals aged 65 or older, persons under 65 with certain disabilities, and people of any age with end-stage renal disease. Like Medicaid/MassHealth, collection of data on patients generally follows the standard set by the UHDDS.

Validity of Racial-ethnic Categories

Massachusetts Cancer Registry and MassHealth data collection evolved over the years to include disaggregated information on race and ethnicity (Massachusetts Cancer Registry 2003). The Emergency Hospital and the MassHealth datasets fulfill the HHS guidelines of specifying 11 Asian ethnic groups, not counting categories of "other." The Massachusetts Cancer Registry dataset follows the U.S. Census Bureau procedures, containing 28 racial subgroups, including 23 Asian racial subgroups. Standardized forms used for collecting data often include one field for entering a race according to standard codes (where the recorder must choose from predetermined racial terms), and also a second fill-in space that allows the persons entering information to write in answers using their own words. In the birth and death records held by the Massachusetts Registry of Vital Records and Statistics, and also in hospitalization records, certain variables contain text fields that capture what was written in the fill-in blanks. For example, in the original dataset for birth records there are both coded and text fields for both parents' race (as well as place of birth, ancestry, and language preference, which are discussed below). Hospital discharge records include not just a field for the individual's race using predetermined codes, but also fill-in text fields for "other race" and "other ethnicity."

Increasingly, people may select multiple races. The Massachusetts Cancer Registry and various other administrative records allow for multiple races to be entered (as indicated in Table 1.1), and MassHealth offers an "interracial" category[12] (MassHealth 2012).

MassHealth requires each recipient to self-identify his/her race and ethnicity,[13] and the Massachusetts Cancer Registry attempts to have all patients self-identify but allows for coders to make educated guesses about race based on related information that is available. In general, the cancer registry guidelines state "When coding race it is important to remember that race is defined by specific physical heredity or origin—NOT by birthplace, place of residence, language or citizenship." But an exception is made for Asian birthplaces, in this way: If the race of a patient is unknown, a default race is entered based on the patient's nationality or origin (Massachusetts Cancer Registry 2003; p. 61). Although finely grained racial-ethnic categories are specified on the enrollment form for MassHealth, our research revealed that many beneficiaries do not fill out these forms because they are automatically enrolled in MassHealth when they are eligible for other need-based aid programs, such as Social Security Supplemental Income.

---

[12] Various national standards for health data information are currently in use, such as on Medicaid and Medicare enrollees. For data collected on patients, one important standard is the Uniform Hospital Discharge Dataset (UHDDS). Originally developed by the National Center for Health Statistics in 1969, the UHDDS was adopted as a standard for the Medicare and Medicaid programs in the 1960s.

[13] See the online version of the Member Beneficiary Request for MassHealth currently in use: http://www.mass.gov/eohhs/docs/masshealth/appforms/mbr.pdf

Researchers have found significant problems in the reliability of racial-ethnic coding in Medicare data. One study found that the self-identified races that people selected were accurate at the 90% level for only those who were white and black. Some Hispanics chose "white" as their race/ethnicity, and some Asian Americans chose "other." This finding is corroborated by above-cited research documenting that some Asian Americans do not readily self-identify with the term "Asian" (Agency for Healthcare Research and Quality 2008).

The definitions of race and ethnicity are not consistent between the above four datasets. The Massachusetts Cancer Registry, for example, defines ethnicity as either Hispanic or non-Hispanic, and race as physical heredity or origin (such as white, black, Chinese, Japanese, or Filipino). In contrast, the Uniform Hospital Discharge Dataset (setting the standard for Medicare and Medicaid data collection) defines race as "a social construct in which individuals are grouped together based on similar physical characteristics" (American Indian/Alaska Native, Asian, black/African American, Native Hawaiian or other Pacific Islander, white, Latino/Hispanic, and other race) and ethnicity as a subgroup of race, where people that "may share a geographic origin, language, history, and/or religious tradition" (such as Bangladeshi, Bhutanese, Burmese, or Hmong.).

Primary Preferred Language Self-Identified[14]

MassHealth records both spoken and written language preference of prospective enrollees, and provides forms both in Spanish and English, as well as certain other languages upon request.

For the Massachusetts Cancer Registry, no data are collected on the patient's preferred language, and no specification on what language was used during the collection process.

The Boston Public Health Commission and state of Massachusetts regulations stipulate that data collected from acute care hospitals should query patients for "Preferred Language." However, current guidelines for collecting hospital discharge data in Massachusetts, according to the 2011 Massachusetts Division of Health Care Finance and Policy Hospital Discharge Data Codebook, do not contain specifications regarding data collection on language proficiency or patient's preferred language (Massachusetts Division of Health Care Finance and Policy 2011).

---

[14] Nationally, hospitals collect somewhat less data on patients' preferred or primary language than on their race-ethnicity. An Institute of Medicine (2009, part 5) report found that "...more than 89 percent of hospitals report collecting race and ethnicity data, and 79 percent report collecting data on primary language." However, the data on language may be self-reported or attributed by an observer, and about half of the hospitals maintain a database of information they collect information on the primary language of their patients.

Socioeconomic Status, Country of Origin, Immigrant Status, and Citizenship

MassHealth datasets include minimal information about income levels of patients. On the other hand, information about expected payer for services in various of the datasets examined allow broad if roughly grained inferences to be drawn about income because certain public insurance payees have income eligibility requirements. For example, the Massachusetts database on Hospital Discharges followed the UHDDS standard by asking for the expected payer for services, and if this is Medicaid/MassHealth, then the researcher can infer that the patient was eligible for Medicaid/MassHealth by income criteria. Currently, for example, a family of four with children under 18 is eligible if their income is at or lower than 150% of the federal poverty level, or at an income of around $34,000 per year.

Administrative Datasets: Vital Statistics

The Registry of Vital Records and Statistics within the Executive Office of Health and Human Services "collects, processes, corrects and issues copies of birth, death and marriage records that occur in Massachusetts."

1) Mortality Records. The death certificate currently requests one's "Race (please specify)" and "Ethnicity" in terms of Hispanic Origin. The Registry of Vital Records and Statistics records verbatim all information from the Death Certificate filed by a funeral director, medical examiner, or family member. If no family member is present, examiner attributes race and ethnicity.

2) Natality Records. The creation of birth records presents administrators with choices about how to assign the race of newborns. Prior to 1989, the race/ethnicity of an infant was assigned by combining information on the race/ethnicity of the mother and the race/ethnicity of the father. Since 1989, Massachusetts has followed national conventions which tabulate births by the self-reported race/ethnicity of the mother.[15]

Datasets Based on Population Surveys

1) Behavioral Risk Factor Surveillance System (BRFSS) is a statewide population survey inquiring into the health of Massachusetts residents. It is conducted annually with a randomly selected sample of respondents in households. The survey is part of a national survey system coordinated by the Center for Disease Control and Prevention. It is administered in the states by the state-level health departments. The respondents self-report on public health issues, their own health status or condition, and factors in their lives or behavior that entail health risk. The survey is administered by telephone, and since 2009 both landline and cell phone numbers were called. In 2010, a mail survey was added. In 2011, the nonweighted sample size of persons who identified as "Asian" (which

---

[15] Reports from the state health department on live births typically count births by mother's race and ethnicity. See http://www.mass.gov/eohhs/docs/dph/research-epi/birth-report-2009.pdf

does not include Pacific Islanders) was 462, which was 2.07 percent of the total sample (Massachusetts Department of Public Health 2011b).

Following standards of the Massachusetts Department of Health, the survey separately asks respondents for their race, whether they are Hispanic or non-Hispanic, and ancestry. The question scheme allows the race variable in the dataset to distinguish, first, whether the respondent self-identified as white non-Hispanic; black non-Hispanics, Hispanic, Asian non-Hispanic, and other (which includes American Indians and Pacific Islanders). Second, since "ancestry" is queried, the data also allow the recording of certain Asian ancestry subgroups, though not asked in terms of "ethnicity" per se. Specifically, the ancestries listed on the questionnaire are: Puerto Rican, Dominican, Mexican, Salvadoran, Chinese, Filipino, Cambodian, Vietnamese, Japanese, Indian (Asian), and Other Central American (fill in text), Other South American (fill in text), or Other Asian (fill in). The options "don't know/ not sure" or "refused" are also available. The ancestry "Korean" is not included in the list of named choices, and we recommend that the list be revised to correct this omission.

Massachusetts adds some questions that are state-specific. Since 1996, the survey has included additional interviews in certain major cities, a design that in effect oversamples residents of those cities. This in some cases increases the representation of racial and ethnic minorities because some selected cities are known to have relatively large concentrations of minority residents.   Lowell, a city with one of the largest populations of Cambodian Americans, has traditionally been included in the list of such cities. In the 2011 survey cycle, Quincy, whose Asian population surged by over 50% between 2000 and 2010, has been added to the list of cities that will be oversampled.[16]

2) Youth Risk Behavior Survey. The Departments of Elementary and Secondary Education and Public Health collaborate with the Centers for Disease Control and Prevention (CDC) to conduct the Youth Risk Behavior Survey (YRBS, also know as the MAYRBS). It is administered in randomly selected public high schools in every odd-numbered year. Respondents anonymously answer questions about "tobacco use, alcohol and other drug use, sexual behaviors that might lead to unintended pregnancy or sexually transmitted disease, dietary behaviors, physical activity, and behaviors associated with intentional or unintentional injuries"[17] (Massachusetts Department of Elementary and Secondary Education and Health 2011).

---

[16] The list of cities where additional interviews will take place in 2011-12 includes Brockton, Lynn, Lowell, Quincy, Springfield, Worcester, Lawrence, and Fall River. Brockton, Lynn, and Quincy are rotated with Lawrence and Fall River.

[17] "For the high school surveys, MYRBS and MYHS, the CDC used a two-stage sampling method to produce representative samples of students in grades 9–12. In the first stage, schools were selected with a probability proportional to school enrollment size. In the second stage, classes of a required subject or required period were selected randomly.[1] Once classes were selected, half were randomly assigned to receive the MYRBS and half the MYHS. Including both surveys, usable data were collected from 5,655 high school students in 52 schools. The overall response rates (i.e., the school response rate multiplied by the student response rate) were 67% for the MYHS and 65% for the MYRBS.

3) REACH U.S. Risk Factor Survey. The CDC administers a survey through its Racial and Ethnic Approaches to Community Health Program (REACH). The CDC has considered the REACH program "the cornerstone of CDC's efforts to eliminate racial and ethnic health disparities among African Americans, American Indians, Hispanics/Latinos, Asian Americans, Alaska Natives, and Pacific Islanders. REACH grantees use community-based participatory approaches that address social determinants of health through policy, systems, and environmental changes." The first REACH U.S. Risk Factor Survey was conducted in 2009 in 28 communities. It collects health-related information on a yearly basis from selected communities with concentrated populations of the above-named racial ethnic groups. As noted above, one Massachusetts city, Lowell, was included in the 2010 survey, enabling important data to be collected on Southeast Asian Americans, but there are no plans to repeat the survey in this city (Institute of Medicine 2009).

Some noteworthy features of the population surveys pertain to racial-ethnic categorization, and language, and socioeconomic information as follows:

- In its national survey, the National Health Interview Survey (NHIS) currently gives respondents the option to self-identify only with one or more of the largest AANHPI categories: white, black/African American, Indian (American), Alaska Native, Native Hawaiian, Guamanian, Samoan, other Pacific Islander, Asian Indian, Chinese, Filipino, Japanese, Korean, Vietnamese, or other Asian.[18]

- In Massachusetts, the BRFSS in 2009 (and since what year) asks respondents to identify with these racial-ethnic categories, which are not separated into "race" and "ethnicity" (Hispanic/Latino or non Hispanic/Latino): Puerto Rican, Dominican, Mexican, Salvadoran, Chinese, Filipino, Cambodian, Vietnamese, Japanese, Indian (Asian), or Other Central American, Other South American, Other Asian (specify ___), Don't Know/Not sure, Refused.

- Apparently due to an oversight, Koreans (numbering 17,369) are not included in the Massachusetts BRFSS as a racial-ethnic category as late as 2010, although they constituted the fifth largest Asian ethnicity in Massachusetts in 2000, only slightly smaller than the fourth largest group, the Cambodian category (19,696), and much

---

For the middle school survey (MYHS), CSR used similar scientific procedures to select a representative random sample of middle schools and classrooms within those schools. CSR staff administered the MYHS in selected schools and classes. Data were collected from 2,859 middle school students from grades 6 through 8 in 69 schools. The overall response rate was 56%." Text from report below, page 2. The MYRBS is supplemented by the Massachusetts Youth Health Survey (MYHS), which are administered as described above.

[18] Respondents are also asked to choose whether they identify with these categories that would fall under the ethnicity category used to identify Hispanic or Latino persons in the census: Puerto Rican, Cuban/Cuban American, Dominican (Republic), Mexican, Mexican American, Central or South American, Other Latin American, Other Hispanic/Latino/Spanish.

more numerous than Japanese or Filipinos, which were included in the Survey (Massachusetts Department of Public Health 2010). See Appendix B for a breakdown of Asian ethnic populations in Massachusetts according to the 2010 census.

The languages used by interviewers are quite limited. In addition to English, the BRFSS is conducted only in Spanish and Portuguese (Massachusetts Department of Public Health 2011b).

- The Youth Risk Behavior Survey asks student respondents to self-identify with one these five racial categories: American Indian or Alaska Native, Asian, black or African American, Native Hawaiian or other Pacific Islander, or white. In addition, the respondent is asked if he or she is Latino or Hispanic (yes or no). No further questions are asked about ethnicity. In general the sample sizes for Asians and Pacific Islanders are so small that no meaningful analysis can be performed for the APIs (Center for Disease Control and Prevention 2011b). In 2009, for example, there were 2,859 middle school students, including only 116 APIs (3.9%). In surveying high school students in 2009, the total sample size for the MYHS was 2,948, including 98 APIs, and for the MYRBS was 2,707, with 131 APIs.

Three national surveys are included in our profile to illustrate how surveys that have been influential in evaluating public health issues or health reform in Massachusetts have not included any information on Asian Americans, Native Hawaiians, or Pacific Islanders. These are the National Survey of Children's Health CDC 2007; the Massachusetts Health Reform Survey, 2009 (Urban Institute); and the National Immigration Survey (Child) CDC. In contrast, we also include in the profile the national surveys mentioned above that have implemented oversampling of Asians: the National Health Interview Survey (NHIS) and the National Health and Nutrition Examination Survey (NHANES). For a comprehensive description of national surveys containing health information on persons of Asian American race-ethnicity, see Islam, Khan et al. (2010; footnote 16).

Extractive Tools for Researchers: MassCHIP

The Massachusetts public health agency, like many other state public health agencies, has created an online data extraction and analysis tool that allows users free access to some 36 health datasets that include information on the "health status, health outcome, program utilization, and demographic datasets" of Massachusetts residents. Provided by the Massachusetts Department of Public Health, the tool, called the Massachusetts Community Health Information Profile (MassCHIP), allows users to create standard and custom reports in health data in current years and some in past years. The reports are intended to provide community-level information, such as health data for cities and towns. Users should be aware of several features of MassCHIP that affect usability for research on Asian American health disparities.

- One positive feature of the data collection system for hospital and vital records is the independence of categorization by race, ethnicity, and ancestry. A consequence of this independence is that numerical counts of persons who are identified with one or more of the many ethnic categories conventionally understood as "Asian" may not be counted as "Asian" if the data recorder did not indicate that particular racial selection. Users of MassCHIP should thus be aware that individuals may not be identified as "Asian" in the datasets even though their ethnicity would be considered "Asian" sub-race in other systems, such as the U.S. Census categorization scheme. For example, some persons who self-identify on hospital admission forms as "Filipino" in the ethnic category may not self-identify as "Asian" in the racial category.

- MassCHIP extracts information from primary datasets and in some cases summarizes it. To obtain the most accurate information for detailed studies of health disparities, researchers should generally use the primary datasets and not MassCHIP. In addition, data on specific ethnic subgroups are often suppressed because the number of respondents in that ethnic category is small, and revealing the data jeopardizes privacy of the respondents.

- "Bridging methods" allow comparison of race-ethnicity data over successive years or decades when categorization rules changed over time. After 2000, the Massachusetts Department of Public Health adopted new standards for data collection, giving respondents the option of checking more than one race category to self-identify their racial-ethnic identity (according to the U.S. census standards), as well as ancestry. In forming population counts of Asians in 2000 and afterward, demographers must decide for cross-decennial bridging and other purposes whether or not to include as "Asian" only respondents who select "Asian" singly on the census questionnaire, or also those who check "Asian" in combination with one or more other racial categories. In MassCHIP reports, the count of Asians in a geographic area follows a fractional formula. A respondent who has selected more than one race, including Asian, in the census is counted as a fractional part

Asian, where the fraction is meant to mirror the proportion of the population in the relevant geographic area that has identified as singly Asian, roughly speaking (Executive Office of Health and Human Services - Department of Public Health - Commonwealth of Massachusetts 2001).

- Some anomalies exist in reporting on the wide range of specific ethnic subgroups of Asian Pacific Americans. These problems likely stem from the complexity and growing number of Asian American ethnic subgroups represented in the state's population, which presents increasing challenges to data collection agencies.[19] Collaborative partnerships between public health agencies and specialists in Asian American studies may help data collection standards keep abreast of changing demographic features of these populations.

---

[19] In MassCHIP (Vital Records) the mortality dataset spanning from 1999-08, there is no option to select data on deaths of Korean, Asian Indian, Filipino, Japanese, or Thai persons. Since the disaggregated death data is in fact recorded in the vital records available through the Registry of Vital Records and Statistics, this anomaly arises from limitations in the extractive capabilities of MassCHIP. Also, in MassCHIP (Vital Records) dataset spanning the period from 1996 to 2009, the user does not have the option to select data on Filipino and Japanese births, in a similar way, certain ethnic selectors are not available for death data.

Conclusion and Recommendations

Massachusetts' efforts to address health disparities by improving collection of public health data on racial ethnic minorities can potentially serve as an example for many other states. This potential can be realized, however, only if valid and reliable health data are regularly collected on the growing population of Asian Americans as well as other minority groups. Following are our recommendations:

Partner with Target Communities to Conduct Local Studies

Public agencies, private organizations, nonprofit groups, and academic research institutions should forge partnerships to design and implement local health surveys. These surveys should target geographically concentrated Asian American communities, such as the Vietnamese American communities in Boston (Dorchester neighborhoods), Cambodian communities in Lowell and Lawrence, and Chinese American communities in Boston (Chinatown) and Quincy. Community surveys can be designed to obtain information relevant to the health problems known to affect the ethnic groups living in these areas, and data collection strategies can be tailored to reach immigrants and low-income individuals who may not respond to standard phone or mail surveys. Community organizations can be enlisted in data collection efforts, providing advice on research methodologies, such as the construction of survey instruments and ways to recruit respondents.

The large Asian Pacific American population in California is favorable for the accumulation of resources and to partnerships such as the one that has created the California Health Information Survey. On a smaller scale in Massachusetts, it would be beneficial to initiate cooperation between public agencies, academic research institutions, medical providers and insurers to design sampling strategies that will better capture information on the diverse Asian American population in the state. It would be important to continue the positive efforts of the Massachusetts Public Health Department to oversample cities with concentrations of Asian Pacific American residents, such as Quincy and Lowell. As the number of Asian American persons who are contacted increases, the range of languages in which the interviewee can respond should be expanded. The additional languages should include the major Asian language spoken by Asian American residents in Massachusetts, i.e. Chinese, Korean, Vietnamese, and Cambodian.

Collect Data from Community-Based Clinics

For collecting data on patients receiving care in acute care hospitals, Massachusetts has established a useful standard of using finely grained ethnic categories that are distinct and recorded separately from racial, ethnic, and country of birth or ancestry categories. Since many residents of neighborhoods with concentrations of low-wage immigrants do not regularly use acute care hospitals for routine screening and services, the collection mechanism currently mandated for acute care hospitals should be extended to cover data collection at community-based clinics.

Promote Community Participation

Community-based organizations should be enlisted in data collection efforts. These community groups can encourage the participation of community members in the design of health research methodologies such as survey instruments and in recruiting respondent participation in the research.

Improve Sampling of Asian Americans in BRFSS

The Massachusetts Public Health Department should continue collection of information on socioeconomic status in the Behavior Risk Factor Surveillance System (BRFSS) and include better reporting and analysis of the relationship between poverty and health risk in regular reports issued by public health agencies.

It will be important to continue and expand the positive efforts of the Massachusetts Public Health Department to oversample cities with concentrations of Asian American residents. As the number of Asian American persons who are contacted increases, the survey should include the major Asian languages spoken by Asian American residents in Massachusetts – namely, Chinese, Korean, Vietnamese, and Cambodian.

Improve Intake Worker Training

It is also vitally important to improve training of hospital and clinic staff responsible for intake of patient information on data collection forms. This training should include education on ethnic categories and names in relationship to countries of origin, names of languages spoken, and cultural factors that influence communication about health topics that may be sensitive or health terms often misunderstood or mistranslated cross-culturally.

To improve the usefulness of MassCHIP data, those who collect data should be encouraged to fill in text fields describing an individual's race, ethnicity, ancestry, parent's place of birth (for natality records), and preferred language, where the text fields are available in the original dataset, such as the birth, death, and hospitalization records.

Promote Consistent Sharable Data

From data currently collected, the usability of records can be improved by systems of sharing and integrating data on race and ethnicity, including Asian Pacific Americans across hospitals, community health centers, physician practices, health plans, and surveys, as recommended by the Agency for Healthcare Research and Quality (Agency for Healthcare Research and Quality 2010).

Public agencies should continue to provide support and necessary resources to regularize sharing of claims data by health plans, including private insurers, and they should take steps to make this data available to researchers. Already underway in Massachusetts is an All-Payer Claims Database for use by researchers, health providers, and others.

The reporting of race-ethnicity on insurance records will need improvement to make the All-Payer Claims Database a useful source of information on health disparities. Massachusetts is one of the few states that has required health plans to submit self-identified race-ethnicity data, but the law is being implemented very gradually; only 5 percent of the records have to meet this requirement by 2012; thus, most race-ethnicity fields on these records "remain empty" (Weissman, Betancourt et al. 2011).

Provide Better Online Documentation

Public agencies and other organizations holding primary datasets should provide better online documentation of health datasets. It is often difficult to find templates showing what variables are present in the data. As a result, navigating these data systems to find information on Asian Americans or other racial-ethnic groups can be very hard.

Table 1.0 Datasets from Population Surveys

Dataset Name	Data Sources	Frequency	Targeted Sample Size	Over-sampling of API Overall	Racial, Ethnic, and Ancestry Categories			
					5 Categories (OMB Standard)	7 Asian Sub-racial Categories (New HHS standard)	>7 Ethnic groups	Multiple Races can be Identified
Massachusetts Surveys								
Behavioral Risk Factor Surveillance (BRFSS)	Interviews by mail and landline phone	Yearly	2008:20,581 2010:16,311	No		✔[1]		No
Youth Risk Behavioral Surveillance Survey (YRBSS)	Surveys in public schools	Odd numbered years	2009: 5,655 high schools 2,859 middle schools	No	✔			Yes
REACH 2010 Surveillance of Health Status in Minority Communities[2]	Address based sampling by phone, face-to-face in 28 communities	Annual (Start in 2009)	1,000 Average in each community, Minority residents ≥ to 18 years	Yes			✔	Yes
National Surveys								
National Health and Nutrition Examination Survey (NHANES)	Interview and physical exam in homes	Annual (released every 2 years)	2009: 10,537 Interviewed	Yes			✔	Yes
National Health Interview Survey (NHIS)	Household interviews	Annual	2010: 34,329 households, 89,976 persons Interviewed	Yes (Start in 2006)			✔	Yes
National Survey of Children's Health (NSCH)	Households with children 0-7 years Interviews by landline/cell phone	Every 4 years	2007: 91,642 child-level interviews	No			✔	Yes

---

[1] Seven Heritage or Ancestry categories include 6 Asian choices plus one "Other Asian-fill in" option. The total list includes Puerto Rican, Dominican, Mexican, Salvadoran, Chinese Filipino, Cambodian, Vietnamese, Japanese, Indian (Asian), Other Central American (specify by fill in), Other South American (specify by fill in), Other Asian (specify by fill in), Don't know/Not Sure, and Refused.

[2] Racial and Ethnic Approaches to Community Health (REACH) 2001-2 (not repeated for Lowell since)

Table 1.0 Datasets from Population Surveys

Socioeconomic and Language Data

Dataset Name	Socioeconomic information included in dataset							Annual Income	
	Highest Education	Telecomm Access	Home Owner	Renter	U.S. Citizen	Country of Birth	Years of U.S. Residency	Personal	Household
Massachusetts Surveys									
Behavioral Risk Factor Surveillance (BRFSS)	Yes	Yes	No	No	No	No	No	Yes	No[3]
Youth Risk Behavioral Surveillance Survey (YRBSS)	Yes	No	NA	NA	No	No	No	NA	NA
REACH 2010 Surveillance of Health Status in Minority Communities	Yes (Anywhere)	Yes (Cell, Phone, Internet)	Yes	Yes	No	U.S. or Not	No	No	Yes
National Surveys									
National Health and Nutrition Examination Survey (NHANES)	Yes (Anywhere)	Yes (Internet, Cell)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
National Health Interview Survey (NHIS)	Yes	Yes (Internet, Cell)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
National Survey of Children's Health (NSCH)	Yes (Parent)	Yes (Landline)	Yes (Parent)	Yes	No	Yes (Child and Parent)	Yes (Parent)	NA	Yes

---

NA = Not Applicable
[3] Not distinguished from Personal Annual Income

Table 1.0 Datasets from Population Surveys

Socioeconomic and Language Data

Health Insurance Status		Language			
Carries or Not	Type Specified	Primary Language Self-Identified	Preferred Language for Medical Care Self-Identified	English Proficiency Self-Identified, 4 Levels	Interviewer or Questionnaire in Primary or Preferred Language
Yes	Yes	No	No	No	No
No	No	No	No	No	No
Yes	No	Yes	No	Yes (Reading Only)	Yes [4]
Yes	Yes	Yes	No	No	No (English Speaking Only)
Yes	Yes	No	No	NA [5]	No (English Speaking Only)
Yes	Yes	Yes	Yes	NA [6]	Yes [7]

---

[4] Interviews in Vietnamese, Khmer, Mandarin Chinese, English, Spanish Language

[5] The interviews are conducted in multiple languages; 4 levels of English proficiency are not recorded. Some information on English proficiency may be inferred if respondent chose a non-English language to conduct interview.

[6] Interviews were conducted in multiple languages; 4 levels of English proficiency are not recorded. Some information on English proficiency may be inferred if respondent i.e. parent chose a non-English language to conduct interview.

[7] Interviews in English, Spanish, Cantonese, Mandarin, Korean Vietnamese; 0.2% of all completed interviews

Table 1.1 Datasets from Administrative Medical Records

| Dataset Name | Racial-Ethnic Categories[1] | | | | | |
| | 5 Racial Categories (OMB) | Additional API Racial or Ethnic Sub-groups | Additional Ancestry or Birthplace Categories | Multiple Races Option | Data Input by Format | |
					Textual	Coded
Datasets that Designate Independent and Separate Categories for Race and Ethnicity						
Hospital Inpatient Discharge Data 2011	Yes	Yes (9 ethnic groups[2])	No	Yes	Other Race, Other Ethnicity	Race1, Race2, Ethnicity1, Ethnicty2, Hispanic Ind: (Y/N)
Birth (Natality Records) Information on Mother & Father (M/F)	Yes	Yes (12 ethnic groups including other API[3])	Yes	Yes	M/F Birthplace, Race, Ancestry, Language Pref.	M/F Birthplace, Race, Ancestry, Language Pref.
Mass Health Sept 2011	Yes	Yes (11+Other ethnic groups[4])	No	No (but Interracial is coded)	None	Race Ethnicity
Datasets that Use the U.S. Census Method of Classifying Race (and Sub-racial Categories)						
Massachusetts Cancer Registry 5th Ed. 2007	Yes	Yes (23+other sub-racial groups[5])	Yes	Yes	None	Race 1-5, Spanish/ Hispanic Origin
Datasets that Use Other Methods of Classifying Race and Ethnicity						
Medicare	Yes	No	No	No	None	Race
Death (Mortality Records)	No[6]	No[6]	Yes[6]	No[6]	Birthplace: city, town, country; M/F Birthplace	Nativity, Hispanic Origin, Race

---

[1] The columns for race-ethnic categories are organized to reflect separation of race, ethnicity, and ancestry (or birthplace) according to the standards of the Massachusetts Public Health Department. These standards differ from the 2011 Health and Human Services standards, which form the contrasting basis for describing race-ethnicity in population surveys in Table 1.2.
[2] Massachusetts Division of Health Care Finance and Policy Hospital Inpatient Discharge Data Electronic Records Submission specification May 2011: http://www.mass.gov/eohhs/docs/dhcfp/g/regs/114-1-17-inpatient-specs.pdf
[3] Home Birth Worksheet for Birth Certificates 2003
http://massmidwives.org%2Findex.php%2Fdownload_file%2Fview%2F50%2F110%2F&ei=KVkpUJGoCaiu0AGcVA&usg=AFQjCNEc65EMhI9kUAvJCa
uxzyxvxDnxRg&sig2=hxnu6IlzsGvVB7b3bBmLqQ
[4] MassHealth's MA21 Computer System that holds the data.
[5] MA Cancer Registry Code Manual 5th Inpatient http:// www.mass.gov%2Feohhs%2Fdocs%2Fdph%2Fcancer%2Fregistry-code-manual-5th-
patient.doc&ei=emspUKfvA6Xu0gGAoICoDA&usg=AFQjCNEDawu2vr89sAP9cbkplQ5bsTmV5g&sig2=ij_tJu6hg1-WVSu053qjmw
[6] Information is attributed; Death certificate asks to specify race, Hispanic origin, birthplace with fill-in text.

Table 1.1 Datasets from Administrative Medical Records

Language Used by Medical Provider		
Primary Language Self-Identified	Preferred Language for Medical Care Self-Identified	English Proficiency Self-Identified, 4 Levels
No[7]	No	No
No	Yes	No
Yes[8]	Yes	Self-identified or determined by agency[9]
No	No	No
No	No	No
No	No	No

[7] In order to meet federal and state legal requirements, Massachusetts hospitals are required to submit a Comprehensive Language Needs Assessment (LNA) every three years, and annual LNA reports highlighting relevant updates in consecutive years. http://www.mass.gov/eohhs/docs/dph/health-equity/lna-comp-template.pdf

[8] MassHealth's MA21 Computer System that holds the data: indicates yes, but primary and preferred not distinguished

[9] Informs applicants and members of the availability of interpreter services or he/she can choose to provide his/her own services. MassHealth agency can also determine when such services are necessary. Chapter 501,Page 501.009 http://www.mass.gov/eohhs/docs/masshealth/regs-member/regs-memb-501.txt

Table 1.1 Datasets from Administrative Medical Records

Dataset Name	Socioeconomic information included in dataset								
	Highest Education (In U.S.)		Highest Education (In Other Country)		Home Owner	Renter	U.S. Citizen	Country of Birth	Years of U.S. Residency
	High School	College	High School	College					
Datasets that Designate Independent and Separate Categories for Race and Ethnicity									
Hospital Inpatient Discharge Data 2011	No	No	No	No	No	No	Yes	No	No
Birth (Natality Records)	Yes [10]	Yes [10]	Yes [10]	Yes [10]	No	No	Yes	Yes	No
Mass Health Sept 2011	No	No	No	No	No	No	Yes	No	No
Datasets that Use the U.S. Census Method of Classifying Race (and Sub-racial Categories)									
Massachusetts Cancer Registry 5th Ed. 2007	No	No	No	No	No	No	No	Yes	No
Datasets that Use Other Methods of Classifying Race and Ethnicity									
Medicare	No	No	No	No	No	No	No	No	No
Death (Mortality Records)	Yes [10]	Yes [10]	Yes [10]	Yes [10]	No	No	No	No	No

[10] Does not specify where education was attained

Table 1.1 Datasets from Administrative Medical Records

Annual Income		Health Insurance Status	
Personal	Household	Insured?	Type specified
No	No	Yes	Yes
No	No	Yes	Yes
Yes	Yes	Yes	Yes
No	No	No	No
No	No	Yes	Yes
No	No	No	No

Appendix A

National Data Collection

Appendix A-1
Long-term plans to Improve Data Collection Nationally — Office of Minority Health, HHS

Improve data collection in AANHPI communities			
Data Collection Goals	**Strategies**	**Lead Agencies**	**Benchmarks/Measurable Outcomes**
Goal 1: Increase the capacity to conduct more reliable health data and research throughout the U.S. and U.S. affiliated jurisdictions for AANHPI populations to better describe and understand the need of the AANHPI growing population as part of the Affordable Care Act Provision: Understanding Health Disparities: Data Collection and Analysis (Sec. 4302).	1. Work with HHS 4302 Workgroup in the full implementation of section 4302 regarding data collection on race, ethnicity, sex, primary language and disability status.	OMH , AHRQ, CMS, HRSA	Work with Federal partners, AANHPI organizations and communities in implementing the HHS 4302 Workgroup recommendations. Work with ONC in disseminating available data on AANHPIs and their ethnic subgroups.
Goal 2: Improve the collection, reporting and disaggregation of race, ethnicity and primary language data on AA, NH and PIs within HHS Departments to reflect the revised OMB requirements for data collection, analysis, and reporting of racial and ethnic data in the Continental US, Hawaii and 6-Pacific Islands.	1. In the implementation of section 4302, work to improve AA, NH & PI data disaggregation.	HRSA, OMH, SAMHSA	1. By 2012, US DHHS reports will reflect separate NH and PI categories for data collection, analysis, and reporting.
	2. Enhance the quality of data collected within SAMHSA's National Survey on Drug Use & Health (NSDUH) for AAANHPI populations.	SAMHSA	2. Embed the enhanced indicators into the 2011 NSDUH.
	3. Continue oversampling of Asian Americans in NCHS's National Health Interview Survey (NHIS).	CDC	3. Prioritize Asian sample size in NHIS

	4. Include an oversampling of Asian Americans in 2011-14 National Health and Nutrition Examination Survey (NHANES).	CDC	4. By Fall 2013, estimates of prevalence among Asian Americans of undiagnosed conditions such as hypertension, high cholesterol, and diabetes will be available
	5. Develop improved tools for accessing and analyzing vital statistics and survey data for small populations.	CDC	5. Improved tools available for data access and analysis.
Goal 3: Partner with NHPI communities to identify additional strategies to collect data on NHPI population groups.		OMH, CDC, AHRQ, NIH	
Goal 4: Collaborate with leading organizations to deliver messages around the Affordable Care Act.	Support messaging of the Affordable Care Act and outreach to AAANHPI communities.	OMH	

Source: http://minorityhealth.hhs.gov/templates/content.aspx?ID=8804&lvl=3&lvlid=573

Appendix A-2
Disaggregated Categories for Race and Ethnicity, HHS Standards

Excerpts from "Implementation Guidance on Data Collection Standards for Race, Ethnicity, Sex, Primary Language, and Disability Status." October 2011 the U.S. Department of Health and Human Services:

The categories for HHS data standards for race and ethnicity are based on the disaggregation of the OMB standard used in the American Community Survey (ACS) and the 2000 and 2010 Decennial Census. The data standard for race and ethnicity is listed below. Race and ethnicity data collection applies to survey participants of all ages.

Ethnicity Data Standard *Are you Hispanic, Latino/a, or Spanish Origin?* *(One or more categories may be selected)*	
Categories	**Notes**
a. _____ *No, not of Hispanic, Latino/a, or Spanish origin*	These categories roll-up to the Hispanic or Latino category of the OMB standard
b. _____ *Yes, Mexican, Mexican American, Chicano/a*	
c. _____ *Yes, Puerto Rican*	
d. _____ *Yes, Cuban*	
e. _____ *Yes, Another Hispanic, Latino/a or Spanish origin*	

Race Data Standard *What is your race?* *(One or more categories may be selected)*	
Categories	**Notes**
a. _____ *White*	These categories are part of the current OMB standard
b. _____ *Black or African American*	
c. _____ *American Indian or Alaska Native*	
d. _____ *Asian Indian*	These categories roll-up to the Asian category of the OMB standard
e. _____ *Chinese*	
f. _____ *Filipino*	
g. _____ *Japanese*	
h. _____ *Korean*	
i. _____ *Vietnamese*	
j. _____ *Other Asian*	
k. _____ *Native Hawaiian*	These categories roll-up to the Native Hawaiian or Other Pacific Islander category of the OMB standard
l. _____ *Guamanian or Chamorro*	
m. _____ *Samoan*	
n. _____ *Other Pacific Islander*	

http://aspe.hhs.gov/datacncl/standards/ACA/4302/index.shtml
Accessed on October 1, 2012.

Appendix B

Massachusetts Asian American Subgroups

Asian Americans by Subgroup in Massachusetts*

	2000	2010	% Change
Chinese (except Taiwanese)	82,028	118,164	44.1%
Indian	43,801	77,177	76.2%
Vietnamese	33,962	42,915	26.4%
Cambodian	19,696	25,387	28.9%
Korean	17,369	24,110	38.8%
Filipino	8,273	12,309	48.8%
Japanese	10,539	9,224	-12.5%
Pakistani	2,145	6,205	189.3%
Other Asian, not specified	4,019	5,729	42.5%
Taiwanese	2,364	4,502	90.4%
Laotian	3,797	3,632	-4.3%
Thai	2,141	3,529	64.8%
Nepalese		2,580	
Bangladeshi	573	2,109	268.1%
Sri Lankan	651	1,034	58.8%
Hmong	1,127	992	-12.0%
Burmese		923	
Indonesian	730	847	16.0%
Bhutanese		425	
Malaysian	218	357	63.8%
Other Asian, specified	614	275	-55.2%
TOTAL ASIAN (one ethnicity only)	234,047	342,425	46.3%
TOTAL ASIAN IN MA (one race)	238,124	349,768	46.9%

*Asian Americans alone who selected one Asian subgroup only.

Dataset: Census 2000 and 2010 Summary File 1 (SF 1) 100 Percent Data.

References

Agency for Healthcare Research and Quality (2008). *Creation of New Race-Ethnicity Codes and Socioeconomic Status (SES) Indicators for Medicare Beneficiaries. Final Report.* Rockville, Maryland.

Agency for Healthcare Research and Quality (2010). *Race, Ethnicity, and Language Data: Standardization for Health Care Quality Improvement . Part 5: Improving Data Collection Across the Health Care System.* Institute of Medicine Report. Rockville, Maryland.

Albelda, Randy, Ferry Cadet, Dinhong Mei (2011). *Poverty in Massachusetts by Race.* Center for Social Policy Publications. University of Massachusetts, Boston.

Barnes, Patricia M., Patricia F. Adams, Eve Powell-Griner (2008). "Health Characteristics of the Asian Adult Population: United States, 2004-2006." *Advance Data* (394): 1-22.

Boston Public Health Commission (2007). *Guidelines for the Implementation, Interpretation, and Enforcement of the Boston Public Health Commission's Standardized Data Collection Regulation.* Boston, Massachusetts. Retrieved March 5, 2012, from http://www.bphc.org/chesj/resources/Documents/Regs/Data%20Collection%20 GuidelinesData.pdf

Boston Public Health Commission (2010). *Health of Boston 2010.* Boston, Massachusetts.

Brugge, Doug, Angela C. Lee, Mark Woodin (2006). "Asthma Prevelance in Native and Foriegn Born Children in a Predominantly Asian Community Setting." *Epidemiology* 17(6).

Center for Disease Control and Prevention (2011a). "Trends in Tuberculosis -- United States, 2010 " *Morbidity and Mortality Weekly Report (MMWR)* 60, 333-337. Atlanta, Georgia.

Center for Disease Control and Prevention (2011b). *Youth Risk Behavior Surveillance System: Selected 2011 National Health Risk Behaviors and Health Outcomes by Race/Ethnicity.* Atlanta, Georgia/
Retrieved May 3, 2012, from http://www.cdc.gov/healthyyouth/yrbs/pdf/us_ disparityrace_yrbs.pdf.

Chang, Ellen T., Theresa H. M. Keegan, Scarlett L. Gomez, Gem M. Le, Christina A. Clarke, Samel K.S. So, Sally Glaser (2007). "The Burden of Liver Cancer in Asians and Pacific Islanders in the Greater San Francisco Bay Area, 1990 through 2004." *Cancer* 109(10): 2100-2108.

Do, D. Phuong, Brian Karl Finch, Ricardo Basurto-Davila, Chloe Bird, Jose Escarce, Nicole Lurie (2008). "Does Place Explain Racial Health Disparities? Quantifying the Contribution of Residential Context to the Black/white Health Gap in the United States." *Social Science and Medicine* 67(8): 1258-1268.

Executive Office of Health and Human Services - Department of Public Health - Commonwealth of Massachusetts (2001). *Bridging Population Data Using 2000 Census Race/Ethnicity. for the Massachusetts Department of Public Health.* Unpublished draft memo (December 31). Boston, Massachusetts.

Holland, Ariel T. and Latha P. Palaniappan (2012). "Problems With the Collection and Interpretation of Asian-American Health Data: Omisssion, Aggregation, and Extrapolation." *Annals of Epidemiology* 22(6): 397-405.

Institute of Medicine (2009). *Race, Ethnicity, and Language Data: Standardization for Health Care Quality Improvement.* National Academies Press. Washington, D.C.

Islam, Nadia Shilpi, Suhaila Khan, Simona Kwon, Deeana Jang (2010). "Methodological Issues in the Collection, Analysis, and Reporting of Granular Data in Asian American Populations: Historical Challenges and Potential Solutions." *Journal of Health Care for the Poor and Undeserved* 21(4): 1354-1381.

Koch-Weser, Susan, Sidney L. Liang, Dorcas C. Grigg-Saito (2006). "Self-reported Health Among Cambodians in Lowell, Massachusetts." *Journal of Health Care for the Poor and Underserved* 17(2 Suppl).

Lien, Pei-te, Margaret M. Conway, Janelle S. Wong (2004). *The Politics of Asian Americans: Diversity and Community.* New York, N.Y., Routledge.

Long, Sharon K. and Karen Stockley (2010). *Health Reform in Massachusetts: An Update As Of Fall 2009.* Urban Institute. Washington, D.C. Retrieved February 6, 2012, from http://bluecrossmafoundation.org/~/media/Files/Publications/Policy%20 Publications/060810MHRS2009FINAL.pdf

Massachusetts Cancer Registry (2003). *Abstracting and Coding Manual for Hospitals, 5th Edition.* Boston, Massachusetts. Retrieved June 2, 2012, from www.mass.gov/.../ cancer/registry-code-manual-5th-intro.doc

Massachusetts Department of Elementary and Secondary Education and Massachusetts Department of Public Health (2011). *2009 Health and Risk Behaviors of Massachusetts Youth.* Boston, Massachusetts. Retrieved May 1, 2012, from www.doe.mass.edu/cnp/ hprograms/yrbs; and www.mass.gov/dph.

Massachusetts Department of Public Health (2010). *Behavioral Risk Factor Surveillance System Questionnaire Massachusetts - 4869.* Boston, Massachuetts. Retrieved May 3, 2012, from http://www.mass.gov/eohhs/docs/dph/behavioral-risk/survey-10.pdf.

Massachusetts Department of Public Health (2011a). *A Profile of Health Among Massachusetts Adults, 2010.* Boston, Massachusetts.

Massachusetts Department of Public Health (2011b). *Methodology, Behavioral Risk Factor Surveillance System.* Boston, Massachusetts. Retrieved May 3, 2012, from http:// www.mass.gov/eohhs/consumer/community-health/brfss/methodology.html.

Massachusetts Division of Health Care Finance and Policy (2007). *Inpatient Hospital Discharge Database Documentation Manual - Fiscal Year 2006.* Massachusetts Division of Health Care Finance and Policy. Boston, Massachusetts. Retrieved March 5, 2012, from http://www.mass.gov/eohhs/docs/dhcfp/r/hdd/hdd-fy06.pdf

Massachusetts Division of Health Care Finance and Policy (2011). *Hospital Inpatient Discharge Data Electronic Records Submission Specification.* Boston, Massachusetts. Retrieved March 5, 2012, from http://www.mass.gov/eohhs/docs/dhcfp/r/hdd/hdd-fy06.pdf

MassHealth (2012). *Member Benefit Request Instruction Page.* Boston, Massachusetts. Retrieved February 25, 2012, from http://www.mass.gov/eohhs/docs/masshealth/ appforms/mbr.pdf.

National Cancer Institute (2008). *Cancer Health Disparitie*s. National Cancer Institute Factsheet. Retrieved May 15, 2012, from http://www.cancer.gov/cancertopics/factsheet/disparities/cancer-health-disparities

Peddycord, Dan, and Martin Fenstersheib (2011). *Status of Vietnamese Health, Santa Clara County, California 2011*. Santa Clara County Public Health Department. Santa Clara County, California.

Pew Research Center (2012). "The Rise of Intermarriage." *Social & Demographic Trends*. Pew Research Publications, Washington, D.C. Retrieved on September 10, 2012, from http://pewresearch.org/pubs/2197/intermarriage-race-ethnicity-asians-whites-hispanics-blacks

Pleis, John R., Bruce W. Ward, Jacqueline W. Lucas (2010). "Summary Health Statistics for U.S.: National Health Interview Survey, 2009." *Vital Health Statistics* 10(249).

Smedley, Brian D., Adrienne Y. Stith, Alan R. Nelson, Eds. (2002). *Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care*. National Academies Press. Washington, D.C.

Tran, Thanh V. and Donna L. Ferullo (1997). "Indochinese Mental Health in North America: Measures, status, and treatments." *Journal of Sociology and Social Welfare* 24(2): 3-20.

U.S. Department of Health and Human Services (2011b) *Improving Data Collection to Reduce Health Disparities*. Retrieved May 12, 2012, from http://www.healthcare.gov/news/factsheets/2011/06/disparities06292011a.html

U.S. Department of Health and Human Services (2011c). *National Hepatitis B Initiative for Asian Americans and Pacific Islanders*. Washington, D.C. Retrieved May 12, 2012, from http://minorityhealth.hhs.gov/templates/browse.aspx?lvl=2&lvlID=190.

U.S. Department of Health and Human Services (2011d). *Implementation Guidance on Data Collection Standards for Race, Ethnicity, Sex, Primary Language, and Disability Status*. Washington, D.C. Retrieved May 8, 2012, from http://www.healthcare.gov/news/factsheets/2011/06/disparities06292011a.html

UCLA Center for Health Policy Research (2012). *California Health Interview Survey*. Los Angeles, California. Retrieved February 2012, from http://www.askchis.com/main/default.asp.

Weinick, R. M., J. M. Caglia, et al. (2007). "Measuring Racial And Ethnic Health Care Disparities In Massachusetts." *Health Affairs* 26(5): 1293-1302

Weissman, Joel, Joseph R. Betancourt, Alexander R. Green, Gregg S. Meyer, Aswita Tan-McGrory, Jacob Nudel, Jessica A. Zeidman, J. Emilio Carrillo (2011). *Healthcare Disparities Measurement*. Commissioned Paper. Disparities Solutions Center. Massachusetts General Hospital. Boston, Massachusetts. Retrieved March, 2012, from http://www2.massgeneral.org/disparitiessolutions/z_files/Disparities%20Commissioned%20Paper.pdf

White House Initiative on Asian Americans and Pacific Islanders (2011a). *HHS Plan for Asian American, Native Hawaiian and Pacific Islander (AANHPI) Health*. Washington, D.C. Retrieved March 1, 2012, from http://www2.ed.gov/about/inits/list/asian-americans-initiative/criticalissues.html

White House Initiative on Asian Americans and Pacific Islanders (2011b). *Critical Issues Facing Asian Americans and Pacific Islanders*. Washington, D.C. Retrieved March

1, 2012, from http://www2.ed.gov/about/inits/list/asian-americans-initiative/
criticalissues.html.

Xu, Jiaquan, Kenneth D. Kochanek, Sherry L. Murphy, Betzaida Tejada-Vera (2010). "Deaths:
Final data for 2007". *National vital statistics reports*. National Center for Health
Statistics. Hyattsville, Maryland.

Acknowledgements

INSTITUTE FOR ASIAN AMERICAN STUDIES
University of Massachusetts Boston

The Institute for Asian American Studies at the University of Massachusetts Boston conducts community-based research on Asian American issues, provides resources to Asian American communities in Massachusetts, and expands opportunities on campus for the study of Asian American experiences.

100 Morrissey Boulevard
Boston, MA 02125-3393
(617) 287-5650
www.iaas.umb.edu
asianaminst@umb.edu