

NORTHERN PLAINS CANCER DATA GAP ANALYSIS

Summary of Findings and Recommendations

BACKGROUND

Importance of Data

The problem of cancer disparities in Northern Plains American Communities is well known. In our region, American Indian men and women are at higher risk of developing and dying from cancer. Also, Northern Plains American Indians have one of the highest cancer incidence and mortality rates compared to American Indians and Alaska Natives in other regions.

There are many wonderful cancer prevention and control activities being done at local and regional levels. But we often forget about the importance of having accurate and relevant cancer information.

We need cancer data in order to:

- (1) Understand why we have such high rates of cancer incidence (new cancer cases) and cancer deaths
- (2) Decide on priorities - which cancers to focus
- (3) Develop programs to reduce cancer incidence and deaths
- (4) Evaluate the programs and track progress over time
- (5) Show the funding agency the need for cancer programs (i.e., to get more resources)

Data Issues Identified by Northern Plains American Indian Communities

Northern Plains Comprehensive Cancer Control Program is a regional cancer program administered through Great Plains Tribal Chairmen's Health Board (GPTCHB). Based on input from tribal health programs, the program identified the following data issues for Northern Plains American Indians:

- (1) An overall lack of or limited cancer data specific for Northern Plains American Indians
- (2) Barriers to accessing and utilizing existing cancer data
- (3) Fragmented cancer information systems
- (4) A lack of or limited cancer data available at a tribal community level
- (5) Non-dissemination of data or reports back to tribal communities for their use
- (6) Limited understanding and under-use of cancer data

NORTHERN PLAINS CANCER GAP ANALYSIS STUDY

In order to address some of these data issues, GPTCHB obtained a grant from Office of Minority Health to establish the “Northern Plains Tribal Cancer Data Improvement Initiative” (NPTCDI) in 2007. As part of this initiative, we conducted a study to document the extent of cancer data problems. We gathered information through literature review, internet search and interview of agencies that conduct national surveys. In the following sections we are showing the highlights of the findings.

The complete report is available from the NPTCDI website: <insert the web address here>

KEY FINDINGS

What kind of population-based data are available?

There are at least 24 national data sources that contain some kind of cancer data. Centers for Disease Control and Prevention (CDC) sponsors many of them. There are different types of data – surveys, administrative data, registries and vital statistics.

Table 1. Data type and sponsoring organization

| | Data Source | Sponsoring Organization |
|----------------------------|------------------------------------------------------------------------|-----------------------------------------|
| Survey | | |
| 1 | American Indian Adult Tobacco Survey (AI-ATS)(5) | CDC |
| 2 | Behavioral Risk Factor Surveillance System (BRFSS)(6) | CDC |
| 3 | Consumer Assessment of Healthcare Providers and Systems (CAHPS)(7) | AHRQ |
| 4 | Health and Retirement Study (HRS)(8) | NIA |
| 5 | Health Behavior in School-aged Children (HBSC) (9) | NICHD / HRSA |
| 6 | Health Information National Trends Survey (HINTS) (10) | NCI |
| 7 | Medical Expenditure Panel Survey (MEPS), Household component(11) | AHRQ / CDC's NCHS |
| 8 | National Ambulatory Medical Care Survey (NAMCS)(12) | NCHS |
| 9 | National Health & Nutrition Examination Survey (NHANES)(13) | CDC |
| 10 | National Health Interview Survey (NHIS) (14) | CDC's NCHS |
| 11 | National Hospital Ambulatory Medical Care Survey (NHAMCS) (12) | CDC's NCHS |
| 12 | National Immunization Survey (NIS)(15) | CDC |
| 13 | National Mortality Followback Survey (NMFS) (16) | CDC's NCHS |
| 14 | Pregnancy Risk Assessment Monitoring System (PRAMS)(17) | CDC |
| 15 | Tobacco Use Supplement to the Current Population Survey (TUS-CPS) (18) | NCI / CDC |
| 16 | Youth Risk Behavior Surveillance Systems (YRBS) (19) | CDC |
| Administrative Data | | |
| 17 | Medicaid Analytic Extract (MAX) (20) | CMS |
| 18 | Medicare Utilization - Standard Analytic Files (21) | CMS |
| 19 | National Longitudinal Mortality Study (NLMS) (22) | NCI / NHLBI / NIA / CDC's NCHS / Census |
| 20 | Resource and Patient Management System (RPMS) (23) | IHS |
| 21 | State Inpatient Database (SID)(24) | AHRQ / CDC's NCHS |
| Registries | | |
| 22 | National Program of Cancer Registries (NPCR)(25) | CDC |
| 23 | Surveillance, Epidemiology, and End Results (SEER) (26) | NCI |
| Vital Statistics | | |
| 24 | National Vital Statistics: Mortality (27) | CDC's NCHS |

AHRQ - Agency for Healthcare Research & Quality
 CDC - Centers for Disease Control and Prevention
 CMS - Centers for Medicare & Medicaid Services
 HRSA - Health Resources and Service Administration
 NCHS - National Center for Health Statistics
 NCI – National Cancer Institute
 NHLBI - National Heart, Lung & Blood Institute
 NIA - National Institute of Aging
 NICHD - National Institute of Child Health & Human Development
 US HHS - US Department of Health and Human Services

What is the representation of Native Americans?

Some data sources have a good representation, at least at the national level (Table 2). BRFSS (“Behavioral Risk Factor Surveillance Survey”) is one of the most widely used national survey which has lots of cancer related information. In 2008, 6470 Native Americans participated in BRFSS.

Other data sources have much smaller number of Native American participating (highlighted with yellow). For example, NHANES (“National Health and Nutrition Examination Survey”) included only 86 Native Americans in its 2007-2008 survey and HINTS (“Health Information National Trend Survey”) included only 61 Native Americans in its 2007 survey.

Table 2. Sample size

| | Data Source | Total Population/Sample Size | AI/AN Population/Sample Size |
|---------------------------|-------------|-----------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------|
| Survey | | | |
| 1 | BRFSS | 2008: 414,509 | 2008: 6,470 |
| 2 | CAHPS | 2008 Adult Commercial: 162,656 | 2008 Adult Commercial: 642 |
| 3 | HBSC | 2005-06 U.S. Survey: 9,016 | 2005-06 U.S. Survey: 526 |
| 4 | HINTS | 2007: 7,674 | 2007: 61 |
| 5 | HRS | 1992 Core Interview: 12,521 | 1992 Core Interview: 162 |
| 6 | MEPS | 2006: 32,577 | 2006: 275 |
| 7 | NAMCS | 2004: 25,286 | 2004: 93 |
| 8 | NHIS | 2008: 291,893 | 2008: 3,192 |
| 9 | NHANES | 2007-2008: 10,149 | 2007-2008: 86 |
| 10 | NHAMCS | 2004 Outpatient Dpt data set: 31,783 | 2004 Outpatient Dpt data set: 149 |
| 11 | NIS | 2008 Household Survey: 30,681 | 2008 Household Survey: 396 |
| 12 | NMFS | 1993: 22,957 | 1993: 205 |
| 13 | TUS CPS | 2008-09: 237,199 | 2008-09: 1,851 |
| 14 | YRBS | 2007: 14,041 | 2007: 297 |
| Administrative | | | |
| 15 | Medicaid | 2003: 55,157,775 | 2003: 806,211 |
| 16 | Medicare | 2008: 44,263,111 | 2008: 185,666 |
| 17 | RPMS | | |
| 18 | SID | Varies across states | Varies across states |
| Registries | | | |
| 19 | NPCR | NPCR and SEER together coverscover the entire U.S. In 2009 estimated 1,479,350 new cancer cases were to be diagnosed. | According to Espey et al. (2008), a total of 28,874 AI/AN cancer cases were found in 1999-2004 registry data. |
| 20 | SEERS | Covers 26% of the U.S. population and | Covers 42% of the AI/AN population |
| Vital Statistics / Census | | | |
| 21 | Mortality | 2003: 2,452,154 | 2003: 13,160 |
| 22 | NLMS | March 1994 CPS (Cohort V): 66,173 | March 1994 CPS (Cohort V): 830 |

What are the gaps in cancer information?

There are many different data sources we can use to understand cancer prevention (commercial tobacco use, physical activity, obesity). Also, screening, diagnosis, and treatment are covered fairly well. But we have very limited data on survivorship.

Table 3. Information Type based on Cancer Continuum of Care Model

| Prevention | Early Detection | | Treatment | Survivorship | Other |
|-------------------------------------------|-----------------------------------------|------------------------------------------------------|-----------------------------------------------|-------------------------------------------------|---------------------------------------------|
| | Screening | Diagnosis | | | |
| Commercial tobacco use (10) | Breast cancer screening receipt (7) | Self-report of diagnosis (5) | Treatment receipt (7) | Type of provider who gives majority of care (2) | Family history of cancer (2) |
| Secondary smoke exposure (5) | Cervical cancer screening receipt (7) | Diagnosis based on provider records (7) | Type of primary care provider (1) | Insurance coverage of cancer treatment (1) | Knowledge and perception of cancer risk (1) |
| Smoking cessation (9) | Prostate cancer screening receipt (7) | Availability of diagnostic tests at the facility (1) | Routine cancer check-up (1) | Limitations caused by cancer (2) | Cancer care giving experience (1) |
| Tobacco policies (3) | Colorectal cancer screening receipt (7) | | Participation in clinical trial (1) | Pain and pain control (1) | Death caused by cancer (4) |
| Youths access and exposure to tobacco (1) | | | Cost of treatment (3) | | |
| Alcohol use (7) | | | Availability of treatment at the facility (1) | | |
| Physical activity (6) | | | | | |
| Nutrition (6) | | | | | |
| Obesity (6) | | | | | |
| Sexual behavior (2) | | | | | |
| HPV infection (2) | | | | | |
| HPV vaccine receipt (4) | | | | | |
| Hep B vaccine receipt (4) | | | | | |
| UV exposure (4) | | | | | |
| Occupational exposure (2) | | | | | |
| Pesticide use (2) | | | | | |

The number in the parenthesis indicate the number of data sources which cover the particular topic

SUMMARY & RECOMMENDATIONS

1. Improve representation of AI/AN in key national surveys

If we do not have adequate sample size of AI/AN respondents, how can we draw a meaningful conclusion from the surveys? At least for major surveys on cancer such as National Health and Nutrition Examination Survey and Health Information National Trend Survey, we think more efforts should be invested to increase the sample size.

Considering the known geographic variations in cancer incidence, mortality and risk factors, the sample size should be large enough for at least regional-level analysis. To reach the adequate sample size for the trend analysis, here are two potential solutions or approaches:

- Conduct oversampling of the AI/AN population in a certain interval (every 3 to 4 years) to compare estimates from specific years of data collection.
- Increase the AI/AN sample size by a smaller proportion for each data collection point (every year or every other year). The increase needs to be large enough so that aggregated data from multiple years (3 to 4 years) can achieve the adequate sample size for each period for comparison over time.

2. Build data infrastructure within tribal communities

Cancer prevention and control activities are planned and implemented by individual tribes. That is why we advocate for local data.

The mechanism used for the American Indian Tobacco Survey and South Dakota Tribal Pregnancy Risk Assessment Monitoring System has worked well for the tribes in the Northern Plains to collect data. These surveys did not only produce the data tribes needed but also helped them to build and strengthen data and evaluation infrastructure for tribes and tribal organizations (e.g., tribal epidemiology centers).

By its design to involve tribal leaders and community members from the designing and dissemination stages, these surveys also helped increase awareness of health issues in the tribal communities. To address effectively disparities in cancer and other health problems, resources should be invested in the communities that suffer from disparities.

3. Balancing information across the cancer continuum

The majority of existing cancer data sources focus on cancer prevention and early detection. Treatment information is relatively scarce and/or difficult to access. The cancer registries provide the first course treatment information but details of treatment (e.g., dosage and specific drugs used in chemotherapy) is not available.

Combined with lack of detailed co-morbidity (one or more diseases) and clinical information in the registry makes it difficult to evaluate the use of guideline based treatment in a given population group. In addition, population-based palliative care and end-of-life care information is largely missing.

As for survivorship, with the exception of a few surveys (e.g., BRFSS), little data exist to understand the experience of survivors, families and care givers. An increased collaboration and coordination across agencies should happen to eliminate duplicating efforts and to fill in cancer data gaps.

4. Increasing data access and sharing

While protecting the confidentiality of data, it is important to make the data more accessible for program evaluation and research. There is a desire that national efforts to establish clinical and public health data standards will make data systems more inter-operational. Promotion of formal data exchanges and use agreements would lead to increased access to data. This is not only important from a cancer surveillance standpoint but from a clinical perspective. Increase in sharing of clinical data, perhaps through RPMS and similar system, may be helpful in improving communication and coordination of care across providers.

In summary, we recommend that resources be allocated for collection of cancer data be done based on a careful analysis of data gaps for the overall U.S. population and for specific population groups. Partnership between tribal communities and agencies that sponsor and/or administer data collection is a key to address cancer data gaps for the AI/AN population.