University of New Mexico
UNM Digital Repository

Native Health Database Full Text

Health Sciences Center Archives and Special Collections

1999

Northwest Tribal Registry--Seattle Indian Health Board Record Linkage Project.

J Bettles
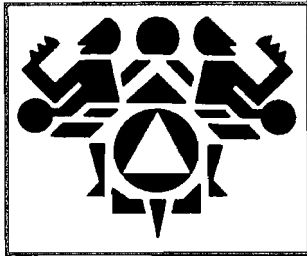
D. Robertson

FC. Romero

T. Becker

P. Stehr-Green

Follow this and additional works at: https://digitalrepository.unm.edu/nhd

Northwest Tribal Registry – Seattle Indian Health Board Record Linkage Project

A report by

James Bettles, B.S., Registry Manager
Dee Robertson, M.D., M.P.H., Director
Francine C. Romero, Ph.D., M.P.H., Epidemiologist

with contributions from

Tom Becker, M.D., Ph.D., Consulting Epidemiologist
Paul Stehr-Green, Dr.P.H., M.P.H., Consulting Epidemiologist

The Northwest Tribal Epidemiology Center (*The EpiCenter*)

Northwest Portland Area Indian Health Board
Portland, Oregon

November 15, 1999

Northwest Tribal Registry–Seattle Indian Health Board Record Linkage Project

SUMMARY

In the last ten years several reports using a variety of health data sets have shown a high proportion of racial misclassification for Northwest American Indians and Alaskan Natives (AI/ANs). The result of this has been an underestimation of the burden of various diseases among Northwest AI/ANs. To address health issues related to misclassification, the Northwest Tribal Epidemiology Center has developed the Northwest Tribal Registry (NTR) which contains a demographic data set of 135,060 Northwest AI/ANs, and which can be used for record linkage studies to ascertain accurate disease rates for Northwest AI/ANs. However, the Seattle Indian Health Board Registration contains a large number of Northwest AI/ANs, and it has not yet been incorporated into the NTR. To ascertain the degree of concordance between the NTR and SIHB Registration, we conducted a record linkage study linking these two files. **Results:** Of the 23,337 entries in the SIHB Registration, only 7,005 (30%) linked or matched with the NTR, leaving 16,332 who were not matched. Among the SIHB-NTR matches, 1,597 (22.8%) were from Northwest tribes; among the SIHB-NTR non-matches, only 3,413 (20.9%) were from Northwest tribes. **Conclusion:** The NTR is missing information on a substantial number of AI/ANs who have utilized the Seattle Indian Health Board for health care services. However, less than a fourth of these individuals are from Northwest tribes served by the Northwest Portland Area Indian Health Board.

INTRODUCTION

The last few years have brought increasing recognition to the vast disparities in health status among the different racial and ethnic populations in the United States. As a result of this awareness, the major theme of the Healthy People 2010 public health agenda for the nation is to eliminate racial health disparities. The Northwest Tribal Epidemiology Center was established in 1997 with the mission of providing Northwest tribes with timely, accurate, and useful health status information. A key part of this effort is the Northwest Tribal Registry (NTR) Project that was started in January 1999. The goal of the NTR Project is to improve the accuracy of health data for Northwest American Indians and Alaskan Natives (AI/ANs) through record linkage with a variety of health data sets. These data sets contain health data for Northwest AI/ANs, but either do not have racial identifiers or have inaccurate racial identifiers.

The validity of the information obtained from our record linkage studies depends on the accuracy, completeness, and representativeness of the data in the NTR. If the Registry is not complete or not representative, the record linkage studies may not yield valid conclusions. A potential concern in this regard has been that the Registry may not contain a significant proportion of the AI/ANs in the Seattle area who utilize the Seattle Indian Health Board (SIHB) for their health care services. SIHB is the only large Indian health care program in the Northwest whose patient registration list is not included in the source file from which the Northwest Tribal Registry was developed. The proportion of AI/ANs who live in urban settings has been increasing for decades and is now estimated to be about 50%.

THE NORTHWEST TRIBAL REGISTRY

The Northwest Tribal Registry is an enumeration of AI/ANs primarily from Idaho, Oregon, and Washington. The source data for the Registry came from the Portland Area Indian Health Service *Area Patient File*—a compilation of patient demographic data from Indian health care facilities that use the Resource and Patient Management System (RPMS) and export patient data to the Portland Area IHS Office. It includes individuals who received services from Northwest Indian health care facilities (except SIHB and Nisqually Tribal Clinic) from the mid-1980s to the present—a total of 135,060 records at the time of this linkage with SIHB in July 1999. However, because of limitations of the data in the source file, the number and percent from each of the Northwest tribes could not be ascertained. The Registry has been rigorously cleaned and unduplicated for use in all record linkage projects (Appendix A).

SEATTLE INDIAN HEALTH BOARD (SIHB) DATA SET

The SIHB data set included 23,337 AI/AN individuals seen at their health care facility since 1991. Prior to the linkage study this data set was standardized (Appendix A). Table 1 summarizes the missing data in the SIHB data set after standardization.

Table 1. Data missing from the SIHB Registration file

Field Name	Number of records with missing data	Percentage of records with missing data
Social security number	4,579	19.6
First name	29	0.1
Middle name	2,967	12.7
Last name	8	0.03
Sex	26	0.1
Date of birth (y/m/d)	38	0.2

PURPOSE

To better define the representativeness of the Northwest Tribal Registry, we conducted a record linkage study between the NTR and the Seattle Indian Health Board Registration file in July of 1999.

METHODS

The SIHB-NTR Registration record linkage study was conducted on-site at SIHB using the probabilistic record linkage software *Automatch©* (Matchware, Inc. v. 4.2). *Automatch* is able to link (or match) entries in two different data files when you have the same individual listed in each of the two files and the same types of characteristics (e.g., name, date of birth, social security number) for the individual are identified in both files. Uncertainty arises when there are differences between the data in the two files for the same characteristic of the same individual. For example:

	NAME	DOB	SSN	BIRTH PLACE
FILE A:	Jon Q. Doe	01/01/1951	999-88-7667	Humptulips, WA
FILE B:	John Q. Doe	01/01/1951	999-88-7676	Humptulips, WA

If a deterministic linkage were done using the above information, the entries would <u>incorrectly</u> be considered to represent two different individuals because of the small variances in the NAME and SSN data. Using probabilistic matching software, these two entries would <u>correctly</u> be considered to be the same individual. An extended explanation of probabilistic matching methods can be found in Appendix B.

The NTR and SIHB Registration were linked using the following characteristics (data elements):

- Name: first, middle, last
- Date of birth (DOB): year, month, day
- Social security number (SSN)
- Sex

The record linkage was conducted in a five-pass run, with each pass allowing for errors on some fields but not others. In succeeding passes, the matching combinations were rotated so that after five rounds all possible matches were obtained using the given fields. The initial error rates for each pass were given wide margins, and then recalibrated as necessary to minimize false-positive

matches. Clerical review sessions (case by case reviews of questionable matches) helped further clarify and separate the true matches from false matches. Upon completion of each linkage the results were extracted for data analysis, and all identifiers permanently removed from the new files created by the linkage.

RESULTS

On the first pass, social security number was used as the blocking field (the primary linkage field), and 5,433 individuals were identified. From a subsequent pass using last name as the blocking field, 1,406 additional cases were identified. An additional 166 individuals were identified during the final three passes utilizing sex and date of birth as blocking fields.

Thus, of the 23,337 AI/ANs in the SIHB Registration, the record linkage identified 7,005 (30.0%) who were also listed in the NTR, called "SIHB-NTR matches". A total of 16,332 AI/ANs (70.0%) in the SIHB Registration did not match, called "SIHB-NTR non-matches" (Table 2 and Figure 1).

Table 2. Totals for the files used in the record linkage

| Data set	Count
NTR	135,060
SIHB Registration	23,337
Individuals in <u>both</u> SIHB Registration & NTR (SIHB-NTR match<u>es</u>)	7,005
Individuals in SIHB Registration <u>but</u> <u>not</u> in NTR (SIHB-NTR non-matches)	16,332
Individuals in NTR <u>but not</u> in SIHB Registration	128,055

Figure 1. Diagrammatic Representation of the Results of the SIHB-NTR Linkage Study
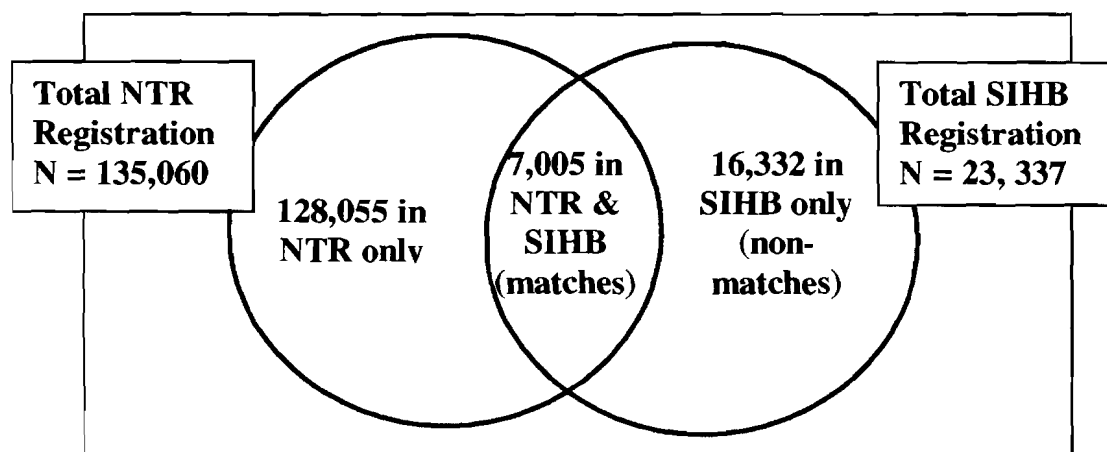
CHARACTERISTICS OF THE SIHB-NTR MATCHES AND NON-MATCHES

Figure 2 shows that the age distribution of SIHB-NTR matches and SIHB-NTR non-matches was similar, but showed a slightly higher proportion of the matched versus non-matched individuals were age 0-24 years, and a slightly higher percentage of the non-matched individuals compared to those who matched were age 25-64 years.

Figure 2. Age distribution of SIHB-NTR matches and non-matches
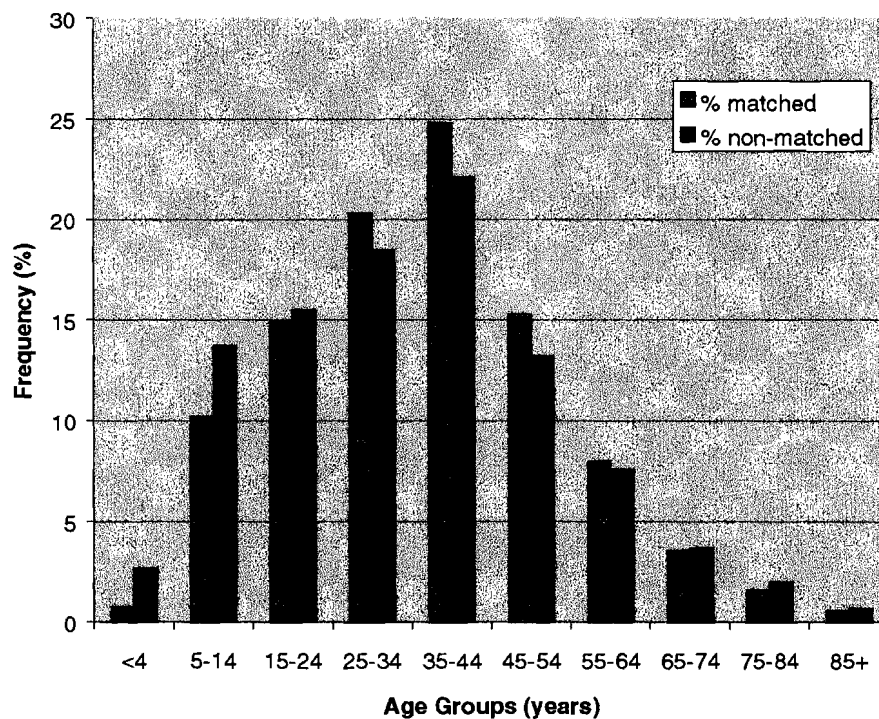
The majority of individuals in the linked data set reported Washington (82.4%) as the state of their residence while (5.8%) reported Idaho, (7.6%) reported Oregon and (4.2%) reported other state of residence.

Because of the question of the representativeness of the NTR, the tribal affiliation of both the matches and non-matches was of particular interest to us. Specifically, we wanted to know how many from each group were from the member tribes of the Northwest Portland Area Indian Health Board in the states of Idaho, Oregon, and Washington. Table 3 shows that only 22.8% of the SIHB-NTR matches were from Northwest tribes (including Federally recognized and non-Federally recognized tribes).

Table 3. Geographic distribution of SIHB-NTR matches, by region of tribe

Category	Frequency	Percent matches
Other Tribe	4,949	70.6
Northwest Tribe	1,595	22.8
Unspecified	372	5.3
Non-Indian	76	1.1
Non-Native Tribal Member	3	0.0
Tribal Total	6,995	99.9
Tribal designation missing	10	0.1
Grand Total	7,005	100.0

Table 4 shows the distribution of the SIHB-NTR matches by tribe, with Northwest tribes listed first. Only the top 15 are listed here. Appendix C contains a complete listing of the frequencies of all tribes represented in the SIHB Registration.

Table 4. Tribal affiliation of SIHB-NTR matches

Top Tribal Matches	Frequency of tribal affiliation, SIHB-NTR matches	Percent tribal affiliation, SIHB-NTR matches
Northwest Tribes		
Yakama	248	3.5
Colville	191	2.7
Lummi	166	2.4
Makah	103	1.5
Total Northwest Tribes	708	10.1
Non-Northwest Tribes		
Tlingit	704	10.0
Sioux, Dakota	520	7.4
Blackfeet	463	6.6
Chipewa, Ojibway	402	5.7
Canadian Native	342	4.9
Alaskan Native	318	4.5
Aleut	263	3.8
Eskimo	232	3.3
Cherokee	193	2.8
Haida	138	2.0
Navajo	130	1.9
Total Non-NW Tribes	3,702	53.0
All others	2,582	36.9
Total	6,995	100.0

Of even more interest was the geographic distribution of the SIHB-NTR non-matches. If a large proportion of the non-matches were from Northwest tribes, this could invalidate inferences we might draw from the results of our record linkage studies with external datasets.

Table 5 below shows that of the 16,332 SIHB-NTR non-matches, 71.8% were from tribes outside the Portland Area, and only 20.8% were from tribes within Portland Area.

Table 5. Geographic distribution of SIHB-NTR non-matches, by region

Category	Frequency	Percent non-matches
Other Tribe	11,731	71.8
Northwest Tribe	3,401	20.8
Unspecified	935	5.7
Non-Indian	174	1.1
Non-Native Tribal Member	3	0.0
Tribal Total:	16,244	99.5
Missing tribal designation	53	0.3
Potential SIHB duplicates	35	0.2
Grand Total	16332	100.0

Table 6 contains the tribal affiliation of the SIHB-NTR non-matches. Only the top 15 are listed here. A complete list can be found in Appendix D.

Table 6. Tribal affiliation of SIHB-NTR non-matches

Top Tribal Non-Matches	Frequency of tribal affiliation, SIHB-NTR non-matches	Percent tribal affiliation, SIHB-NTR non-matches
Northwest Tribes		
Yakama	487	3.0
Colville	390	2.4
Lummi	331	2.0
Total NW Tribes	1,208	7.4
Non-Northwest Tribes		
Tlingit	1,661	10.2
Sioux, Dakota	1,208	7.4
Blackfeet	1,058	6.5
Chipewa, Ojibway	934	5.7
Canadian Native	757	4.6
Alaskan Native	689	4.2

Top Tribal Non-Matches	Frequency of tribal affiliation, SIHB-NTR non-matches	Percent tribal affiliation, SIHB-NTR non-matches
Aleut	673	4.1
Eskimo	546	3.4
Cherokee	488	3.0
Haida	349	2.1
Tsimpshian	301	1.8
Navajo	296	1.8
Total Non-NW Tribes	8,960	55.1
All others	6,111	37.5
Total	16,279	100

DISCUSSION

This study indicates that a substantial number of AI/ANs from Northwest tribes has received services from SIHB but have not accessed health care services at any other Northwest Indian health care facility in the previous ten years, and thus are not included in the NTR. However, the total number of these non-matches (3,401) represents less than 3% of the total listed in the NTR; further, these 3,401 are distributed among all the Northwest federally recognized tribes. Thus, it seems unlikely that their absence from the NTR will result in a substantial bias in the results of record linkage studies using the Registry.

Of the 23,337 AI/ANs in the SIHB Registration data set, even after a very thorough, multifaceted record linkage study, only 7,005 (30.0%) matched with the 135,060 individuals in the Northwest Tribal Registry. About one fifth (20.9%) of the 16,332 non-matched individuals were from tribes in Idaho, Oregon and Washington. Among those, the Washington tribes of Yakama, Colville, Lummi, and Makah were the most frequently represented. The number of non-matches for each tribe represents only a very small proportion of the total tribal population.

The goal of the Northwest Tribal Epidemiology Center is for the Northwest Tribal Registry to be as complete and accurate as possible. Inclusion of all the SIHB Registration would expand the total number of AI/ANs listed in the NTR from 135,060 to 151,393—a 12.1% increase. To the extent that our primary interest is in Northwest AI/AN-specific rates, our record linkage study with SIHB has shown that a relatively low proportion of AI/ANs from Northwest tribes would be added, and that this would not substantially affect Northwest AI/AN-specific morbidity and mortality rate estimates. However, adding the SIHB registration would increase the accuracy of identification of AI/ANs in our record linkage studies,

and would allow future epidemiologic reports to better reflect the burden of disease for AI/ANs in the Seattle metropolitan area.

We recognize that data linkage studies should be as accurate and complete as possible. The Northwest Tribal Epidemiology Center is open and ready to include other validated lists of Northwest AI/ANs into the NTR. In this way the Northwest Tribal Registry will be able to provide the tribe-specific and health care program-specific analyses which have been long requested by Northwest tribes and Indian health care programs.

Appendix A
Cleaning and deduplication steps for the Northwest Tribal Registry

The Northwest Tribal Registry (NTR) was prepared in three conditioning steps: standardization, cleaning, and unduplication.

Definitions

Standardization – Conditioning step necessary to ensure the precision of probabilistic record linkages. Standardization involves transforming nicknames into formal syntax so that blocking steps of the matching process are more likely to find true matches.

Cleaning – Conditioning step that involves removing dummy patient records and nonsense data from the data file to be linked.

Unduplication – Conditioning step that involves the identification of duplicate entries in the data file to be linked. Duplicates are retained within the file that is linked. After linkage the unduplicated registry counts are used for denominator calculations.

The standardization process invokes *AutoStan*, a component software utility within the record linkage software program *AutoMatch*© (Matchware Inc., v. 4.2). Name fields and field lengths were standardized using this program and written into a flat ASCII file. The standardized ASCII file was then imported into the Microsoft Access database program for cleaning, where records were sorted by field type (e.g. first name, last name, birth year, community, etc). Nonsense data and dummy patient records were removed within each sorting step. Using the blood quantum field 20,301 non-AI/AN patients were identified in the NTR and removed to a separate database. The NTR was then converted back into a cleaned, standardized ASCII format. A deduplication program was executed using *AutoMatch*, which identified 10,276 duplicate entries within the NTR. Since these duplicates contained varying levels of information about the same individual, they were retained within the NTR to increase the probability of a true match. However, note that all denominator calculations rely upon the unduplicated NTR counts.

Appendix B
Methods for Probabilistic Record Linkage Studies

I. Investigation of the data files includes evaluating the content (variables, values, codes, etc.) and the methods used to collect the data. Discussions with dataset managers (such as Cancer Registrars) will allow for development of error margins necessary for the matching step. Health status information contained within the data file is examined and preliminary research questions developed.

II. Conditioning increases the precision of linking two data files. The purpose of conditioning is twofold: to increase the probability of true matches and decrease the probability of false negatives. *Autostan*, a component of the matching software, helps accomplish this objective by standardizing key matching fields. There are four rule sets required for conditioning: classification, dictionary, pattern action, and optional conversion files. The data are parsed according to its character (e.g. numeric vs. string) and then classified (e.g. name vs. address). General patterns are then typed and the file rebuilt into a standardized form. Soundex and NYSIIS codes on any of the fields can also be generated to increase matching options.

III. Matching is the key step in record linkage and employs a computer software program called *AutoMatch (v. 4.2)*. Data dictionaries for each file to be linked are created. These dictionaries specify the matching fields and the epidemiological data fields of interest.

The matching strategies may differ according to specific data files and the validity of certain fields. Essentially AutoMatch works by minimizing the amount of matching combinations that have to be run. If one small file containing 1,000 records were to be matched sequentially against another file containing 1,000 records, there would be 1,000 x 1,000 = 1,000,000 possible combinations. Sequential matching with files of increasingly greater size easily exceeds the processing capacity and memory cache of most computers. This can be bypassed by reducing the two sets to be matched into "blocks," where one field (e.g. social security number) must match exactly and the other fields are allowed to contain errors (e.g. name, date of birth, etc.). In this fashion our example of a 1000 x 1000 pass could be reduced to say, 100 blocks each containing (10x10=100) combination possibilities. In the next pass the name field would have to match exactly and errors would be allowed on social security number in a new set of blocks. The user defines the matching strategies, tolerable error margins, probability weight scores, and upper and lower cut-off points.

The cut-off weights are used to define three sets of data. The first of these sets are the total number of true matches whose raw weight scores fall above the upper cut-off weight. The user reviews this data to ensure there are no false positives

and can reassign the upper cut-off weight if necessary. The second data set is composed of the total unmatched or unlinked cases for that pass called residual records or "residuals" and fall below the lower cut-off weight. The third set of data produced are those that fall within the lower and the upper cut-off weights called clerical review records or "clericals". Clerical review involves looking at each potential linked pair and assigning them a matched or residual status. Although unmatched records from a clerical review are reassigned to the residual pool, they are not considered for future match passes.

Running a series of matching passes reduces the processing time needed for matching to occur. Clerical review occurs after each pass, with about four or five passes sufficient to account for all possible matches. Incomplete data represent the major limiting factor in record linkages but with sound matching strategies the probability of false negative matches is greatly reduced.

An interesting feature of AutoMatch is the ability to recognize the commonality of different fields within a file and to reassign the probability weight scores on a case by case basis. For example the name "Smith" is more likely to appear in two files than the name "Flaverly," so the program will reset the probability weight for "Flaverly" higher than "Smith". Similarly it would recognize a year of birth of "1910" as being less common and less reliable than a year of birth of "1980". These applications of fuzzy logic are accomplished through the loading of a frequency index prior to matching and the creation of a tree index which factors scores into the matching algorithms.

IV. Survivorship and formatting occur after all passes are completed. A new file is extracted using the linked cases to import all of the epidemiological data fields contained in the dictionary file. This file is imported into a statistical analysis software such as SPSS© for analysis. Research questions include ascertaining rates of racial misclassification, overall morbidity and mortality profiles, and demographic information for both the regional and community level. A unique-identifier match key is generated to allow recall of key epidemiological data minus personal identifiers.

Ascertainment of racial misclassification involves two record linkages per project. The first linkage matches the Tribal Registry against the complete data file and the production of summary statistics. The second record linkage involves repeating the linkage against the same data file which contains only those patients identified as American Indian. The summary statistics are then analyzed and rates of misclassification derived.

Appendix C

Appendix C contains the tribal affiliation of the SIHB-NTR matches. All Portland Area tribes (located in Idaho, Oregon and Washington) are included. Non-Portland Area tribes with frequencies of 50 or greater are specified by name; all others are included in the "Other Tribe" category.

Tribe	Frequency	Percent
Northwest Tribes		
Yakama	248	3.5
Colville	191	2.7
Lummi	166	2.4
Makah	103	1.5
Nez Perce	79	1.1
Nooksack	67	1.0
Quinault	63	0.9
Tulalip	60	0.9
Snoqualmie	54	0.8
Clallam	53	0.8
Shoshone	46	0.7
Muckleshoot	41	0.6
Spokane	39	0.6
Suquamish	38	0.5
Puyallup	35	0.5
Coeur D'Alene	34	0.5
Quilleute	26	0.4
Skagit, Upper	25	0.4
Swinomish	23	0.3
Paiute	21	0.3
Kootenai Band	19	0.3
Snohomish	19	0.3
Conf. Tribes of Umatilla	18	0.3
Shoalwater	18	0.3
Skokomish	16	0.2
Squaxin Island	15	0.2
Capalis	14	0.2
Klallam Lower Elwha	13	0.2
Duwamish	12	0.2
Nisqually	12	0.2
Sauk-Suiattle	9	0.1
Cowlitz	8	0.1
Kalispel	6	0.1
Klallam Port Gamble	5	0.1
Siletz	5	0.1
Chinook	< 5	0

Tribe	Frequency	Percent
Stilaquamish	< 5	0
Grand Ronde	< 5	0
Chinook, landless	< 5	0
Hoh Band	< 5	0
Samish	< 5	0
Shoshone-Bannock	< 5	0
Bannock	< 5	0
Cowlitz, landless	< 5	0
Satsop, Lower Chehalis Band	< 5	0
Modoc	< 5	0
Skykomish	< 5	0
Snohomish, landless	< 5	0
Conf. Tribes of Coos	< 5	0
Non-Northwest Tribes		
Tlingit	704	10.0
Sioux, Dakota	520	7.4
Blackfeet	463	6.6
Chipewa, Ojibway	402	5.7
Canadian Native	342	4.9
Alaskan Native	318	4.5
Aleut	263	3.8
Eskimo	232	3.3
Cherokee	193	2.8
Haida	138	2.0
Navajo	130	1.9
Tsimpshian	101	1.4
Cree	69	1.0
Gros Ventre	69	1.0
Assiniboin	65	0.9
Apache	60	0.9
Tonto Apache	52	0.7
Cheyenne	51	0.7
Crow	51	0.7
Flathead	50	0.7
Other Tribes	436	8.6
Unknown	23	.33
Unspecified	372	5.3
Non-Indian	76	1.1
Non-tribal Members	3	.0
Tribal Total	**6995**	**99.9**
Missing	10	.1
Grand Total	**7005**	**100.0**

Appendix D

Appendix D contains the tribal affiliation of the SIHB-NTR non-matches. All
Portland Area tribes (located in Idaho, Oregon and Washington) are included.
Non-Portland Area tribes with frequencies of 50 or greater are specified by name;
all others are included in the "Other Tribe" category.

Tribe	Frequency	Percent
Northwest Tribes		
Yakama	487	3.0
Colville	390	2.4
Lummi	331	2.0
Makah	254	1.6
Nez Perce	175	1.1
Nooksack	157	1.0
Tulalip	157	1.0
Quinault	141	0.9
Shoshone	126	0.8
Clallam	125	0.8
Snoqualmie	86	0.5
Muckleshoot	85	0.5
Spokane	85	0.5
Puyallup	77	0.5
Suquamish	75	0.5
Swinomish	67	0.4
Paiute	61	0.4
Coeur D'Alene	58	0.4
Skagit, Upper	58	0.4
Quilleute	44	0.3
Nisqually	36	0.2
Cowlitz	35	0.2
Skokomish	35	0.2
Conf. Tribes of Umatilla	34	0.2
Conf. Tribes of Warm Springs	33	0.2
Kootenai Band	32	0.2
Duwamish	28	0.2
Shoalwater	26	0.2
Snohomish	25	0.2
Sauk-Suiattle	24	0.1
Capalis, Lower Chehalis Band	20	0.1
Squaxin Island	17	0.1
Klallam Lower Elwha	14	0.1
Modoc	13	0.1
Shoshone-Bannock	13	0.1

Tribe	Frequency	Percent
Siletz	13	0.1
Grand Ronde	12	0.1
Hoh Band	10	0.1
Kalispel	10	0.1
Klallam Port Gamble	9	0.1
Chinook	7	0
Samish	7	0
Stilaquamish	7	0
Bannock	<5	0
Cowlitz, landless	<5	0
Snohomish, landless	<5	0
Chinook, landless	<5	0
Humtulips, Lower Chehalis Band	<5	0
Conf. Tribes of Chehalis	<5	0
Skykomish	<5	0
Steilacoom	<5	0
Conf. Tribes of Coos	<5	0
Celilo Wyam Board	<5	0
Skagit, Lower	<5	0
Non-Northwest Tribes		
Tlingit	1661	10.2
Sioux, Dakota	1208	7.4
Blackfeet	1058	6.5
Chipewa, Ojibway	934	5.7
Canadian Native	757	4.6
Alaskan Native	689	4.2
Aleut	673	4.1
Eskimo	546	3.4
Cherokee	488	3.0
Haida	349	2.1
Tsimpshian	301	1.8
Navajo	296	1.8
Gros Ventre, Hidatsa	174	1.1
Cheyenne	156	1.0
Apache	149	0.9
Assiniboin	140	0.9
Cree	138	0.8
Crow	137	0.8
Flathead	126	0.8
Tonto Apache	123	0.8
Ramona Band Billage Cahuilla	117	0.7
Cowichan	90	0.6
Choctaw	70	0.4
Cheyenne-Arapahoe	65	0.4

Tribe	Frequency	Percent
Other Tribe	927	5.7
Non-Indian	174	1.1
Unknown	237	1.4
Unspecified	935	5.7
Non-Native Tribal Member	3	0
Tribal Total	**16244**	**99.7**
Missing	53	0.3
Potential SIHB Duplicates	35	.2
Grand Total	**16332**	**100.0**

Accession Number	**9703851**

Entry Date	01/22/2002

Title	Northwest Tribal Registry--Seattle Indian Health Board Record Linkage Project.

Author(s)	Bettles J. Robertson D. Romero FC. Becker T. Stehr-Green P.

Institution	The Northwest Tribal Epidemiology Center.

Source	Northwest Portland Area Indian Health Board, Portland, Oregon.. November 15

Publication Date	1999

Document Type	Report

Tribe	Alaska Native, Aleut, Blackfeet, Cherokee, Chippewa, Eskimo, First Nations, Haida, Lummi, Makah, Navajo, Ojibway, Sioux, Tlingit, Yakama

Geographic Region / Country	Northwest, United States

U.S. State / Province	Idaho, Oregon, Washington

KeyWords (Medical Subject Heading Terms)	Database Management Systems/standards; Documentation/standards; Quality Assurance, Health Care; Population Surveillance; Terminology

Abstract

In the last ten years several reports using a variety of health data sets have shown a high proportion of racial misclassification for Northwest American Indians and Alaskan Natives (AI/ANs). The result of this has been an underestimation of the burden of various diseases among Northwest AI/ANs. To address health issues related to misclassification, the Northwest Tribal Epidemiology Center has developed the Northwest Tribal Registry (NTR), which contains a demographic data set of 135,060 Northwest AI/ANs, and which can be used for record linkage studies to ascertain

accurate disease rates for Northwest AI/ANs. However, the Seattle Indian Health Board Registration contains a large number of Northwest AI/ANs, and it has not yet been incorporated into the NTR. To ascertain the degree of concordance between the NTR and SIHB Registration, we conducted a record linkage study linking these two files. RESULTS: Of the 23,337 entries in the SIHB Registration, only 7,005 (30%) linked or matched with the NTR, leaving 16,332 who were not matched. Among the SIHB-NTR matches, 1,597 (22.8%) were from Northwest tribes; among the SIHB-NTR non-matches, only 3,413 (20.9%) were from Northwest tribes. CONCLUSION: The NTR is missing information on a substantial number of AI/ANs who have utilized the Seattle Indian Health Board for health care services. However, less than a fourth of these individuals are from Northwest tribes served by the Northwest Portland Area Indian Health Board.