

DATA SOURCE ASSESSMENT WORKSHEET

Inter Tribal Council of Arizona Epidemiology Center

General Information:

- (1) Name of Agency that owns data [CHECK ONLY ONE]
- Arizona Department of Health Services
 - County Health Department _____ (SPECIFY)
 - University of Arizona
 - Tribe __ __ __ [ENTER TRIBAL CODE NUMBER]
 - IHS _____ (SPECIFY SERVICE UNIT)
 - Hospital _____ (SPECIFY)
 - Other State/County _____ (SPECIFY)
 - Other _____ (SPECIFY)
- (2) Name of Department/Program _____
- (3) Name of Data Set _____
- (4) Local Area Contact: (Check the names of persons interviewed)

POSITION TYPE	NAME	TITLE	PHONE
Administrator			
Data Manger/Analyst			
Other			

- (5) Agency Fax Number: () -

Background:

- (6) What other data sets are maintained on a regular basis?
- [1] _____
- [2] _____
- [3] _____

[ALL OF THE FOLLOWING QUESTIONS PERTAIN TO THE DATA SET IN QUESTION 3]

- (7) What is the primary reason for maintaining this data set?
- To identify trends & outbreaks
 - To monitor costs/services/resources
 - To keep legal/enforcement records
 - To monitor occurrence in the population
 - Other _____ (SPECIFY)

- (8) Are there any laws, statutes, or regulations that mandate the data system?
- Yes
 - No (SKIP TO QUESTION 10)

- (9) If 'YES', please list all that apply:

_____	_____
_____	_____

- (10) List personnel involved in the department noted in Question 2 involved with the data set by job titles and percent of effort devoted to data. FOR EXAMPLE: WHO DOES ABSTRACTION, DATA ENTRY, COMPUTER PROGRAMMING WITHIN THE AGENCY, QUALITY CONTROL, EDITING ANALYSES, SPECIAL PROJECTS, ROUTINE REPORTS AND SPECIAL REPORTS.

Job Title	Percent (%) Effort
_____	_____
_____	_____
_____	_____
_____	_____

(11) What is the primary type of data collection system?
PLEASE CIRCLE THE APPROPRIATE BOX FOR EACH DATA COLLECTION SYSTEM.

a. Surveillance (ongoing data collection) system? Yes No

b. Periodic survey (periodic point-in-time survey) system? Yes No

IF 'YES' to Question 11a. And 'NO' to Question 11b., THEN please go to Question 13.

(12) How often is this survey conducted?

PLEASE CHECK ONLY ONE: Biannually
 Annually
 Monthly
 Other

IF 'OTHER', [PLEASE SPECIFY]:

(13) Is this data set designed to capture information about persons who become ill and/or use health services?

- Yes
 No (SKIP TO QUESTION 17)

(14) What are the criteria for the case definition? (CHECK ALL THAT APPLY)

Clinical signs and symptoms

- | | |
|----------|----------|
| 1. _____ | 4. _____ |
| 2. _____ | 5. _____ |
| 3. _____ | 6. _____ |

Laboratory findings

- | | |
|----------|----------|
| 1. _____ | 4. _____ |
| 2. _____ | 5. _____ |
| 3. _____ | 6. _____ |

Utilization of specific health services

- | | |
|----------|----------|
| 1. _____ | 4. _____ |
| 2. _____ | 5. _____ |
| 3. _____ | 6. _____ |

Other: _____ (SPECIFY)

- | | |
|----------|----------|
| 1. _____ | 4. _____ |
|----------|----------|

2. _____ 5. _____
 3. _____ 6. _____

(15) What are the combinations of the criteria identified in Question 14 required to meet the case definition. (*Write out or use brackets, parenthesis, and NOT/OR/AND operators.*)

(16) What factors are known to determine whether a case is included in the data set?

- Insurance/eligibility _____ (SPECIFY)
- Use of a particular facility _____ (SPECIFY)
- Enrollment (TRIBAL ONLY) _____ (SPECIFY)
- IHS Eligible _____ (SPECIFY)
- Dependent on physician reporting from outside agency
- Dependent on self reporting/referral
- Use of a certain level of care (e.g., hospital admission)
- Other _____ (SPECIFY)

(SKIP TO QUESTION 18)

(17) The records in this data set pertain to what types of observations?

- Environmental Events/Accidents _____ (SPECIFY)
- Monitoring activities _____ (SPECIFY)
- Incident Investigations _____ (SPECIFY)
- Routine Inspections _____ (SPECIFY)
- Environmental surveys _____ (SPECIFY)
- Surveys of People _____ (SPECIFY)
- Other _____ (SPECIFY)

(18) Do all cases that occur in the population come to the attention of the persons responsible for reporting?

- Yes (*Go to question 19a*)
- No (*Go to question 19b*)

(19) a. If 'YES', what are the most important reasons for this under-reporting?

b. If 'NO', Specify any reasons for over-reporting.

(20) Are you aware of any reasons why a true case that comes to the attention of a person responsible for reporting would not actually be included in the data set? (*Refer to Question 13 and Question 14, IF APPLICABLE*)

- Yes
- No (SKIP TO QUESTION 22)

(21) If 'YES', PLEASE MARK ALL OF THE FOLLOWING THAT APPLY.

- Person responsible for reporting fails to report.
 - Errors in data entry or abstraction.
 - Errors in observation of people or environments.
 - Errors in instrumentation.
 - Errors in interpreting results.
 - Observation or testing done at inappropriate times.
 - Case-definition based on Signs and Symptoms alone not accurate.
 - Case-definition based on both signs/symptoms and laboratory results not accurate.
 - No case-definition used, clinical judgment only.
 - Other (SPECIFY) _____
-

Population Characteristics

(22) What primary population or environment is the data intended to represent? (If you were going to calculate a rate, what is the most appropriate denominator?)

- U.S. Populations
- Arizona Resident Population
- Enrolled Tribal Population
- Resident and members of Tribal Community
- Other _____ (SPECIFY)

(23) In your opinion, does your data set accurately represent the population or environment you are trying to study?

- Yes (SKIP TO QUESTION 25)
- No

(24) If 'NO', please describe:

Flow Chart

(25) Draw a flow chart of the data system and attach to this form.

[SEE ATTACHED FLOW CHART AS AN EXAMPLE]

(26) Which events trigger data collection? [SELECT ALL THAT APPLY IN COLUMN A]

(A) Data Collection Event	(B) Time to Completion of Data Collection (minutes, hours, days, months, etc.)
<input type="checkbox"/> Provider diagnosis of disease	
<input type="checkbox"/> Provider suspicion or clinical presumption of disease	
<input type="checkbox"/> Death	
<input type="checkbox"/> Delivery of Service [SPECIFY]	
<input type="checkbox"/> Discharge or completion of service [SPECIFY]	
<input type="checkbox"/> Laboratory/Pathology Report [SPECIFY]	
<input type="checkbox"/> Other [PLEASE SPECIFY]	
<input type="checkbox"/> Unknown	

Information Collected: Methods and Content

(27) What information is collected? [ATTACH A PRINTOUT OF VARIABLES AND THEIR RANGE OF POSSIBLE RESPONSES TO THIS FORM.] *[If not available ask for print out of data entry screen. If not available then ask for a questionnaire, and clarify which variables are not entered into the data set.]*

(28) Do documents or contacts report that there are variables of primary importance to the purpose of the surveillance system that have high rates of invalid missing values (i.e., missing values that are not supposed to be missing)? How high are these rates?

(29) How is the data routinely collected for the surveillance system? [CHECK ALL THAT APPLY]

- Patient Interviews
- Interview Healthcare Provider
- Parent Interviews
- Interview with other proxy _____ [PLEASE SPECIFY]
- Abstraction from Medical charts
- Abstraction from Discharge data
- Abstraction from ER logs
- Abstraction from other Administrative Records _____ [PLEASE SPECIFY]
- Copy of Medical charts
- Copy of Discharge data
- Copy of ER logs
- Copy of other Administrative Records _____ [PLEASE SPECIFY]
- Field notes
- Inspection/Audit Checklists
- Sampling/monitoring/lab results
- Other _____ [PLEASE SPECIFY]

Storage and Transfer of Data

(30) What is the form of the data when transferred from reporting sources to the agency that maintains the data set?

- Reporting source makes patients and records available to surveillance staff
- Records of individual data (e.g., case reports) sent to surveillance staff
- Tables of aggregated data sent to surveillance staff
- Other _____ [PLEASE SPECIFY]

(31) How is the information stored within the Department/Program? [CHECK ALL THAT APPLY]

- Hardcopy
- Computer File
- Other _____ [Please Specify]

(32) Is the information confidential?

- Yes
- No (SKIP TO QUESTION 36)

(33) If 'YES', please describe which information is confidential:

(34) How is confidentiality assured? (PLEASE SPECIFY)

(35) Cite, if any, legal confidentiality requirements that apply.

(36) Is a computer used to **store** data?

- Yes
- No (SKIP TO QUESTION 39)

(37) What software is used to store and transfer data. Provide an answer for each of the following:

- 1) Software (e.g. Paradox, Dbase, etc.) _____
Specify version(s) _____
- 2) Operating System (DOS or MAC) _____
- 3) Software used to transfer data (e.g. data compression software)

(38) What hardware requirements are needed? (e.g., Mainframe/PC, size of memory, RAM, etc.)

- 1) PC, LAN, WAN, OR Mainframe? (SPECIFY) _____

- 2) Model Names _____

- 3) Amount of Memory for Data Storage _____

- 4) Size of Database (Number of Records) _____

Analysis

(39) Are the data analyzed?

- Yes
- No [SKIP TO QUESTION 44]

(40) Who analyzes the data and where?
[PLEASE SPECIFY... ASK TO SPEAK TO THAT PERSON]

(41) How are the data routinely analyzed?

- Hand generated
- Computer generated
- Other _____ [PLEASE SPECIFY]

(42) What type of analysis is conducted?

- Descriptive analysis of frequencies, rates, and proportions in graphs and tables.
- More advanced analysis _____ [PLEASE SPECIFY]

(43) What software, if applicable, is used to analyze the data? _____
_____ [PLEASE SPECIFY]

(44) Is a report generated from the data?

- Yes
- No [SKIP TO QUESTION 50]

(45) What software, if applicable, is used to generate the report? _____
[PLEASE SPECIFY]

Questions 46-49 are below the table. *Please complete the answers in the table provided.*

A Question (46) Types of Reports	B Question (47) Frequency of Reports (daily, weekly, monthly, etc.)	C Question (48) To whom circulated	D Question (49) How Distributed

(46) What types of reports are generated from this data?
[RECORD IN TABLE – COLUMN A]

(47) How often are these reports generated? [RECORD IN ABOVE TABLE – COLUMN B]

(48) To whom are they circulated? [RECORD IN ABOVE TABLE – COLUMN C]

(49) How are they distributed? [RECORD IN ABOVE TABLE – COLUMN D]

- 1) Hardcopy through mail
- 2) Hardcopy hand delivered
- 3) Computer/Electronic transfer
- 4) By verbal message over phone
- 5) Computer floppy disk by mail
- 6) Other _____ [PLEASE SPECIFY]

Availability

(50) On a scale of 1 to 5 (1 = LEAST LIKELY, and 5 = MOST LIKELY) what is the likelihood that raw data (without identifiers) would be available to:

i) Other Organizations/Departments/Programs within the Tribe

LEAST LIKELY	(1)	(2)	(3)	(4)	(5)	MOST LIKELY
-----------------	-----	-----	-----	-----	-----	----------------

ii) Federal Agencies

LEAST LIKELY	(1)	(2)	(3)	(4)	(5)	MOST LIKELY
-----------------	-----	-----	-----	-----	-----	----------------

iii) State/other governmental agencies

LEAST LIKELY	(1)	(2)	(3)	(4)	(5)	MOST LIKELY
-----------------	-----	-----	-----	-----	-----	----------------

iv) Non-Governmental Organizations (e.g., ITCA, Inc.)

LEAST LIKELY	(1)	(2)	(3)	(4)	(5)	MOST LIKELY
-----------------	-----	-----	-----	-----	-----	----------------

v) University/Colleges

LEAST LIKELY	(1)	(2)	(3)	(4)	(5)	MOST LIKELY
-----------------	-----	-----	-----	-----	-----	----------------

vi) General Public

LEAST LIKELY (1) (2) (3) (4) (5) MOST LIKELY

(51) Would the data be available to organizations outside your agency in aggregate form? (e.g., tables)

- Yes
- No [SKIP TO QUESTION 53]

(52) If 'YES', what (if any) requirements apply? [PLEASE SPECIFY]

(53) What variables can be used to identify tribal affiliation?

(54) What variables can be used to identify whether or not the occurrence of case resides on a particular Indian reservation?

(55) In order to locate residence, what are the geographic units that could be identified in available data?

- Census block group
- Census tract
- Zip code
- Town, city, village
- Reservation
- County
- State
- Other [PLEASE SPECIFY] _____

Computer files that store data can be arranged differently. In a rectangular file, columns represent different variables (e.g., name, age, sex) and row represent one type of observation such as patients or visits to the hospital. Sometimes more than one type of observation is of interest (e.g., patient and visits, households and individuals, mothers and children). When data are stored in multiple rectangular files, there is a file pertaining to each type of observation (e.g., a file on patients and a separate file on visits). These different files can be linked. Each

file, for example, has a variable that indicated which visits belong to which patients. Sometimes, due to inexperience in programming or due to advanced programming, different types of observations can be found in a single row. These are complex files with variable-length records. For example, one row in the file contains both fixed information on the patient and variable information on visits including diagnostic, treatment and laboratory services. There are also more complicated, three-dimensional, nonrectangular files.

(56) How would the computer files be structured?

- Single rectangular file
- Multiple, linked rectangular files
- Complex files with variable-length records.
- Other [PLEASE SPECIFY] _____

(57) How long does the agency store the data? _____ (years)

(58) If ITCA, Inc. could receive data from the system, what would be the average time between data entry and its transfer to ITCA, Inc.?

- < one week
- 1 week to 1 month
- 1 to 6 months
- 6 months to 1 year
- > 1 year
- Other _____ [PLEASE SPECIFY]

THANK YOU FOR YOUR TIME AND COOPERATION.