

MAY 2004

Epidemiologic Profile Development Guidelines

Based on Lessons Learned
From Development of the
Visakhapatnam District
Epidemiologic Profile

ANDHRA PRADESH, INDIA



Epidemiologic Profile Development Guidelines

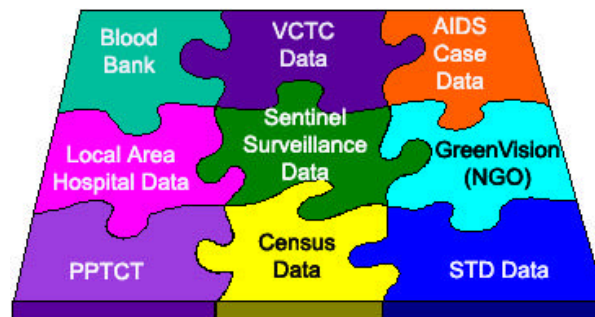
Based on lessons learned from development of the
Visakhapatnam District Epidemiologic Profile
Andhra Pradesh, India
2003/2004

Part 1 – Overview	3
Definition of an epidemiologic profile	3
Reasons to develop a profile	3
Purpose of this document	4
Part 2 – Starting the process	5
Determining the scope	5
Determining skills, time, and other resources needed	5
Determining the content of the profile	8
Part 3 – Collecting the data	9
Identifying data available to use in the profile	9
Learning more about the data	10
Collecting, cleaning, and storing profile data	11
Sampling	12
Issues of confidentiality	13
Part 4 – Using the data to describe the epidemic	15
Basic analytic concepts	15
Core epidemiologic questions	17
Answering the core questions	18
Part 5 – Completing the profile	24
Writing the remaining sections	24
Creating tables and graphs	24
Developing maps	25
Issues of confidentiality	26
Describing data strengths and limitations	26
Review	27
Part 6 – Lessons learned from profile development in Visakhapatnam	28
Coordination needs	28
Resource needs	28
Strengthening data sources	29
Identifying gaps in information	29
Part 7 – Using the profile for public health action	31
Comparing populations	31
Identifying additional areas for data collections	31
Evaluating HIV prevention and care services	32
Appendices –	
A. Putting Together the Puzzle (Abridged)	
B. Data Collection Guidelines	
C. Visakhapatnam – An Epidemiologic Profile of HIV/AIDS	

Part 1 - Overview

Definition of an epidemiologic profile

An HIV/AIDS epidemiologic profile is a document that describes the HIV/AIDS epidemic in various populations in defined geographic areas. It identifies characteristics of the general population, HIV-infected populations, and uninfected (and untested) persons whose behavior places them at risk for HIV. It consists of information gathered to describe the effect of the HIV/AIDS epidemic on an area in terms of sociodemographic, geographic, behavioral, and clinical characteristics. It combines information from a variety of sources to create a “picture” of the HIV/AIDS epidemic, as shown below.



Reasons to develop a profile

The profile is a valuable tool that can be used in a number of ways. These include:

- Sharing information – Informing policy-makers, health providers, program staff, and the general public about the HIV/AIDS epidemic, increasing community awareness and providing evidence to mobilize resources.
- Setting priorities – Providing one basis on which decision-makers may set priorities among competing HIV prevention and care needs to effectively target scarce resources.
- Monitoring and evaluation – Providing a baseline for monitoring the HIV/AIDS epidemic and evaluating the effects of interventions to control it.
- Identifying areas where additional information is needed – Identifying questions the data do not address and guiding future data collection activities.

Purpose of this document

The Andhra Pradesh State AIDS Control Society (AP SACS), in collaboration with the National Alliance for State and Territorial AIDS Directors (NASTAD) and the Centers for Disease Control and Prevention-Global AIDS Program/India (CDC-GAP India), jointly developed an HIV/AIDS epidemiologic profile for the district of Visakhapatnam. The purpose of the profile was to bring existing data sources together in a systematic way to ascertain what type of insight they provided into the local HIV/AIDS epidemic. The development process provided an understanding of the characteristics and quality of these data, the ease with which the data from each source could be accessed, the comparability of the data across sources, and the kinds of resources necessary to continue and/or replicate an activity of this type in other regions. The process also offered insight into the kinds of data currently unavailable through these existing sources that may be useful in addressing questions about the extent and scope of the epidemic and local activities to prevent and contain HIV.

This document describes some general steps to develop an epidemiologic profile, the specific steps taken to develop the epidemiologic profile in Visakhapatnam, and the lessons learned during this process. It includes everything from determining the scope of the profile to identifying data appropriate to describe the HIV/AIDS epidemic in the district to steps taken to obtain, analyze, and present the data. This document is intended to serve as a technical assistance tool to help other states and districts develop profiles to describe their local HIV/AIDS epidemics. For more information, please see “Integrated Guidelines for Developing Epidemiologic Profiles: HIV Prevention and Ryan White CARE Act Community Planning” by the Centers for Disease Control and Prevention and Health Resources and Services Administration. The document can be found at <http://www.cdc.gov/hiv/pubs/guidelines.htm> and was used as a resource for these guidelines.

Part 2 – Starting the Process

Determining the scope

The first step in preparing an HIV/AIDS epidemiologic profile is to determine its scope. Scope refers to the boundaries, such as time frame and geographic area, which define the extent of information in the profile. The scope of the profile has impact on the resources needed for developing the profile.

A number of decisions were made in regards to scope when starting the process of epidemiologic profile development in Visakhapatnam. These decisions were driven by a number of factors, including the pilot nature of this project as well as the original abbreviated time frame in which it was to occur (3 months). Scope was limited in the following ways:

- The epidemiologic profile was developed for only one district. Visakhapatnam was chosen by AP SACS based on the presence of Andhra Medical College in the district, established working relationships with personnel at the medical college, and accessibility to the district by plane.
- Data used for the profile were limited to existing data. No new data were generated for the Visakhapatnam profile.
- Data were extracted from records covering fiscal year April 2002 to March 2003.

Determining skills, time, and computer resources needed

Skills

A multidisciplinary approach is recommended for writing an epidemiologic profile, even if one person takes the lead in writing the profile. This approach allows people with a variety of skills and experiences to contribute to the document. A team should include people, or have access to people, who have:

- familiarity with the local HIV/AIDS epidemic
- familiarity with clinical aspects of HIV/AIDS and its treatment
- understanding of how data are collected and the strengths and limitations of available data sources
- skills in data analysis and interpretation
- knowledge of HIV/AIDS surveillance systems
- knowledge of basic principles of epidemiology and statistics
- knowledge of, and experience with, programs and research
- knowledge of the needs and duties of those planning prevention and care services
- knowledge of the data needs of the people providing prevention and care services
- knowledge of policy issues
- understanding of the importance of protecting confidentiality of individuals
- ability to communicate, both in writing and orally, with diverse audiences in user-friendly language

- ability to interpret data from HIV/AIDS and other surveillance systems to make inferences for HIV prevention and care planning

Other specific skills are needed for profile development, and these may be provided by team members or, if that is not possible, accessed from other individuals. These include:

- statistical skills
 - using descriptive statistics (e.g., mean, median, frequency, percentage, statistical relationships)
 - calculating rates
 - assessing trends over time
 - sampling
- computer skills
 - word processing
 - use of basic statistical and graphics software (e.g., Microsoft Excel, Microsoft PowerPoint, EpiInfo, Statistical Package for Social Sciences, Statistical Analysis Software, mapping software)
- writing and speaking skills, including the ability to communicate difficult concepts clearly to a variety of audiences
- interpersonal skills and ability to work with persons from diverse backgrounds and disciplines

Additional skills needed for data collection (particularly existing data):

- medical record abstraction
- data entry
- quality assurance for these processes

For the Visakhapatnam profile, the task of profile development was assigned to a NASTAD consultant epidemiologist who had all of the specific skills listed above in the areas of statistics, using computers, communicating data, and working with others. She initially worked on the profile for three months, full-time, while located in Andhra Pradesh. Work on the profile continued in the months after she returned to the U.S.

While the consultant took the lead in writing the profile, a much larger group of people with a variety of skills and backgrounds contributed to the development of the profile. Five other NASTAD India team members (two HIV epidemiologists and three HIV/AIDS program directors with varied backgrounds, one of whom was also an infectious disease physician) worked on the profile on a part-time basis, while in India and while in the U.S. Other staff in U.S. health departments were occasionally involved in dealing with technical issues. Staff from AP SACS (including program directors, a physician, monitoring and evaluation, and data management staff) contributed to data collection, data entry and profile development. Staff from the Community Medicine Department at Andhra Medical College, including program directors and two house surgeons, contributed to data collection and profile development. Overall, this group of individuals provided the multidisciplinary approach needed for the profile.

Time

Activities to complete the profile were conducted in a relatively short period of time. After initial planning discussions, the timeline proposed was as follows:

- 1 week to select indicators
- 1 week to develop databases for each source, using EpiInfo
- 4 weeks for data collection with direct, onsite entry into EpiInfo databases
- 3 weeks for data cleaning and analysis
- 4 weeks for sections to be shared with respective sites/departments for initial feedback
- 3 weeks to incorporate feedback and share draft with AP SACS

The project required more time than was initially projected, primarily because data collection efforts were more extensive than anticipated. Factors that affected the timeline included:

- important need to develop relationships with people in the field who were to be involved in order to gain their support
- need to establish registers to consolidate data and indicators that were collected in multiple registers at most sites to protect confidentiality;
- issues related to access to, functionality of, and software available on computers;
- provision of feedback to field personnel (for example, counselors) about the prospective programmatic uses for certain data elements
- limited staff time (e.g. house surgeons were only available for about 4 hours each day)

Computer resources

In order to conduct analyses for data to be used in the epidemiologic profile, it is necessary to have certain capacity in the area of computers and available software. For development of the Visakhapatnam profile, abstracted data were entered into databases created in EpiInfo 2002. The most recent version of EpiInfo (as of February 2004) is EpiInfo Version 3.2 (information available at <http://www.cdc.gov/epiinfo/>). System requirements include:

- Windows 95, 98, ME, NT 4.0, 2000, or XP
- 32 MB of Random Access Memory (64 MB for Windows 4.0 and 2000, 128 MB for Windows XP)
- 200 megahertz processor recommended (300 for Windows XP)
- At least 260 megabytes of free hard disk space (Drive C) to install; 130 megabytes after installation.

In addition to EpiInfo, which provides tools for database development, data entry, data analysis, and mapping, software is needed for producing the text and other graphics for

the profile. For the Visakhapatnam profile, Microsoft Word and Microsoft Excel were used. Tools are also available in EpiInfo for these purposes.

For development of the Visakhapatnam profile, access to computers in the clinics where data were abstracted from charts was limited or non-existent. Access was limited in terms of numbers of computers, particularly computers that were able to run the necessary software. Additionally, problems were encountered with electrical power and computer viruses. These situations had some impact on the amount of time required to collect the data. The consultant and data collection team had access to a laptop computer and one desktop computer for 6 weeks. Data entry also occurred at AP SACS over about 3 weeks with ample computer support.

Determining the content of the profile

To be useful for prevention and care planning, a full epidemiologic profile should answer several core epidemiologic questions:

- What are the sociodemographic characteristics of the general population in the geographic area being described?
- What is the scope of the HIV/AIDS epidemic in this geographic area?
- What are the indicators of risk for HIV infection and AIDS in the population covered in this geographic area?

The profile should also contain the following sections:

- Front matter – List of contributors, executive summary, table of contents
- Introduction – Background about the history and purpose of the profile, general description of the data sources and their strengths and limitations, preparation information
- Body – The epidemiologic questions and analyses of the data that answer these questions
- Conclusion, appendices, and other back matter – A summary of the data, trends, and highlights, information on data sources, glossaries, lists of acronyms, etc..

The epidemiologic profile for Visakhapatnam includes all of the sections listed above. Information on development of the content of these sections is described in the following pages.

Part 3 – Collecting the data

Identifying data available to use in the profile

Data Sources

Before Visakhapatnam was identified as the pilot site for development of an epidemiologic profile, staff at AP SACS, NASTAD India Team personnel, and CDC-GAP/India personnel developed a list of potential sources of relevant data to describe the HIV/AIDS epidemic in Andhra Pradesh. These were:

- AIDS case data
- Blood bank data
- Behavioral Surveillance System (BSS) data
- Care and support center data
- Census data
- Community-based Sexually Transmitted Infection (STI) Study data
- Department for International Development (UK) Impact Assessment Study data
- National Female Health Survey (NFHS) data
- Prevention of Parent to Child Transmission data (PPTCT sites)
- Sentinel Surveillance data
- STD clinic data
- TB-HIV coinfection data
- Voluntary Counseling and Testing data (VCTC sites)

These data sources, including strengths, limitations, and data collection forms, were described in a “Putting Together the Puzzle: First Steps in the Creation of an HIV/AIDS Epidemiologic Profile in Andhra Pradesh.” (Abridged version, Appendix A) These data did not include all possible sources of information (for instance, vital statistics data on HIV/AIDS-related deaths and hospital discharge data) because they were not readily available for review at the time the document was written.

After Visakhapatnam was chosen as the district for development of the pilot epidemiologic profile, this list of data sources was reviewed again to identify which sources would be available that would specifically describe HIV/AIDS in that district. The majority of the data was collected from various departments at the Andhra Medical College and four government hospitals in Visakhapatnam. These included:

- AIDS case report data – King George Hospital and CHEST Hospital
- Blood bank data – King George Hospital Blood Bank
- Behavioral Surveillance System (BSS) data – General population BSS and Targeted High Risk Population BSS, Andhra Pradesh
- Care and support center data – Green Vision (a non-governmental organization)
- Census data – 2001 Indian census
- Prevention of Parent to Child Transmission (PPTCT) data – King George Hospital, Victoria General Hospital, Anakapalle Area Hospital

- Sentinel surveillance data – 2003 sentinel surveillance data – antenatal clinic data from Anakapalle Area Hospital, STD data from Andhra Medical College, STD and AIDS Department, and high risk population data from NGO
- Voluntary Counseling and Testing Center (VCTC) data – Andhra Medical College Microbiology Department, Anakapalle Area Hospital, Narsipatnam Area Hospital

Clinic and hospital data were collected at the different sites. Sentinel Surveillance data (Antenatal clinic and Sexually Transmitted Disease clinic data and high risk data) were available from AP SACS, as were 2001 Behavioral Surveillance System data.

AP SACS and NASTAD staff initially examined the data collection forms being used in the different clinic and hospital settings and selected indicators from each form that would be used for analyses in the epidemiologic profile. The lists of indicators were then reviewed with representatives from each site (e.g., the department head, chief medical officer, counselor, and/or medical social worker) for appropriateness as well as completeness and relevance (see next section), and the lists were subsequently modified as needed. Although not all possible data elements from each source were selected for inclusion in the profile, those selected were thought to be most relevant to program planning and monitoring. Additionally, the indicators selected were, in most instances, also those for which the data were most complete. Data collection guidelines for the different sites are located in Appendix B.

Learning more about the data

The following are considerations for reviewing data and data sources that may be used in the epidemiologic profile:

- Completeness of the data – How well does the number of people reported with HIV/AIDS represent the true number of people with HIV/AIDS?
- Representativeness of the data – How well do the characteristics from a data source correspond to the characteristics of the overall population? For example, data from government-supported hospitals and clinics may not represent all HIV-infected individuals.
- Age of the data – How old are the data that will be used for analysis? It is best to use data, if available, that describe the recent characteristics of the epidemic.
- Timeliness – How long is the reporting delay between diagnosis of HIV/AIDS and report to the health department (if applicable?)
- Limitations of the data source or variable of interest: Consider the limitations of the data source or variable. For example, answers to behavioral questions on data collection instruments may not be consistently available.

- Surrogate, or proxy, markers – A proxy variable is used as a marker for other variables when what we want to measure is too difficult to measure. For example, some areas may use sexually transmitted infection (STI) data as a proxy when data on sexual behaviors are not available.
- Validity of the data – How well does a variable measure what it is intended to measure?
- Small numbers – If the data represent small numbers, they may be difficult to interpret because small absolute changes in the number of cases can produce large relative or proportionate changes in rates that may be misinterpreted by audiences. It is important to be prepared to interpret these data correctly, and this may happen through having access to people who know how to use advanced statistical tests.

In addition to taking all of these factors into consideration, representatives for each data source were asked questions pertaining to data quality, including information on gaps in data collection and any changes in data collection, forms, and procedures. These questions included the following:

- Date that site started?
- When did reporting in current format start?
- Any incomplete periods? e.g, missing forms for a certain period?
- Are all selected indicators usually completed?
- How are data collected for each individual?
- Does staff ask all form questions?
- How are data stored?
- Examples of additional site-specific questions:
 - Return rate for HIV test report?
 - Ratio of counselor to client?
 - Staff training?
 - Reason for incomplete data?

A thorough understanding of all data and data sources is necessary in order to more accurately interpret the data that are collected for the profile and to describe the strengths and limitations of the data. Additionally, it is important to understand the number and characteristics of individuals using services at a particular site. For instance, in order to determine what percentage of individuals at a particular clinic are HIV positive, it is necessary to know the total number of individuals tested at the clinic.

Collecting, cleaning and storing profile data

Individual-level data were collected from records in each of the various sites by a 4-person data abstraction team, which included two house surgeons from the AMC Community Medicine Department, and representatives from AP SACS and NASTAD. Access to different data sources was granted by appropriate authorities including,

individual department heads and hospital superintendents. In most settings where data were collected, clinic data on each individual client were located in multiple registers. In a limited number of settings (KGH STD clinic, hospitals collecting information on AIDS cases), data on each individual client were located in only one record (e.g. the STD card, the hospital chart).

Data were abstracted from each of the sites based on the final list of indicators (see Appendix B). Some indicators, such as demographic characteristics, were common to all data sources, while other indicators were unique to certain data sources. Records were reviewed by hand and manually entered into registers over a one-month period. In some cases, counselors and medical social workers assisted with data collection. Emphasis was placed on making sure that definitions of indicators were understood to assure they were consistent across data sets and that data were abstracted from information that was actually available, to avoid making assumptions about indicators in the absence of data.

Computerized databases were developed for each data source using EpiInfo 2002 software. Two individuals from the data abstraction team entered data from the registers into the EpiInfo databases. The majority of data entry occurred at the AP SACS office in Hyderabad over a three-week period. Data were then cleaned and analyzed for presentation in the epidemiologic profile. Cleaning the data primarily involved organizing and recoding the collected indicators and/or creating new variables. For instance, individual occupations were combined into general occupational categories.

Sampling

It is often impractical to collect large amounts of information on everyone who has HIV or is a risk for HIV because of resource constraints. As a result, either the amount of information is limited, or the quality of the information is sacrificed, or both. One solution to this problem is to collect information from a subset of people from the population of interest. This subset is called a sample. Information is collected from each person in the sample and this information is used to estimate the total number of people with some characteristic in the larger population. Statistical techniques may be used to make sure that sample estimates don't differ too much from one sample to the next.

In determining the sample size, it is important to be able to allow for estimates that are as precise as possible while minimizing wasted resources. Sample size depends on the characteristics of the population and the specific goals of the data collection activity. In general, the more rare the characteristic, the larger the sample needed, and the more common the characteristic, the smaller the sample needed.

There are a number of methods that can be used for sampling, and often when choosing a method, you're trying to strike a balance between what you can practically achieve and what you will gain from your data. Random samples, in which everyone in the population has the same chance of being picked, are preferred, but not always practical or possible. Different sampling strategies were used when abstracting different types of

data for the pilot epidemiologic profile. Simple random sampling and systematic sampling are described in the examples that follow.

Blood bank

Among blood bank donors, only data from those individuals who were initially accepted as potential donors based on prescreening and who then subsequently tested HIV positive were abstracted. No risk behaviors would have been identified among persons testing HIV negative and information on their sociodemographic characteristics would likely have contributed little additional information.

STD clinic

For STD clinic records, a simple random sample was attempted. When choosing a simple random sample, a random technique is used to select individuals. Either (a) everyone in the population is assigned a unique number and then random numbers are selected using a table of random numbers or a random number generator on a computer or (b) a single number is picked (such as the number 10) and then every 10th individual is selected from the larger population to be part of the sample. The idea is that everyone in the population has the same chance of being picked.

Due to a large volume of records in the one STD clinic site, data were collected from only one-third of STD patient records. Within the STD clinic, data were stored separately for male and female STD patients. An attempt to sample every third record was made; however, sampling was not uniform for records on males. At times, every second record was sampled, particularly when the record belonged to an HIV+ individual; consequently, men who were co-infected with STDs and HIV were oversampled.

PPTCT and VCTC

Another type of sampling that can be used is systematic sampling; this type of sampling is not random since not everyone has the same chance of being picked, but it is conducted in a uniform way. For PPTCT and VCTC records, all HIV positive cases and a sample of negative cases were abstracted for the 2002 fiscal year. Limited resources precluded collecting information from all of the HIV negative records. The sample drawn of HIV negative patients was calculated as ten percent of the total number of individuals tested per month. For example, if 100 individuals were tested for HIV in a given month, then information for the first ten HIV negative records was collected.

Issues of confidentiality

The activity of abstracting HIV-related data from clinic records and medical charts necessitates access to sensitive information on individuals. Institutional policies generally protect the confidentiality of individually-identifiable information. Some of the clinics where data were abstracted for the pilot project stored information about individuals in a number of different registers to protect the privacy of clients. This is important because people are likely to avoid service settings where personal and sensitive information about them is not safeguarded from individuals who do not need that

information to take care of them. Data that are collected manually into registers, combined data on individuals that are not regularly combined, and data entered into databases must not, when possible, include patient identifiers such as name so that individual records cannot be associated with individual clients.

Additionally, it is important to take into consideration the issue of sharing data. Institutions and state health departments should be consulted before sharing their data with outside parties, such as other state health departments, institutions, or researchers.

Part 4 – Using the data to describe the epidemic

Basic analytic concepts

Following are some basic terms, definitions, and calculations that may be used in an epidemiologic profile.

Case: A condition, such AIDS (e.g., an AIDS case), according to a standard case definition.

Confidence interval (CI): A range of values for a measure that is believed to contain the true value at a specified level of statistical certainty (e.g. 95%)

Convenience sampling: A technique that relies upon selecting people who are easily accessible at the time of a survey (e.g., a survey of clients who attend a group meeting or are in a clinic when a researcher happens to be there).

The advantage of convenience sampling is that it is easy to carry out. The weakness is that the findings may not represent the group you are trying to study.

Cumulative cases: The total number of cases of disease reported or diagnosed during a specified time. Cumulative cases can include people who have already died.

Cumulative incidence rate: The total number of cases during a specified time period, among people at risk for the disease.

Estimate: When accurate data are not available, an estimate may be based on the data that are available and an understanding of how they can be generalized to larger populations. In some instances, national or state data may be statistically adjusted to estimate local conditions. Estimates should be accompanied by statistical estimates of error (a confidence interval), which describes the uncertainty associated with the estimate.

Incidence: The number of new cases in a defined population in a certain time period, often one year, which can be used to measure disease frequency. It is important to understand the difference between HIV incidence and reported HIV diagnoses. HIV incidence refers to all persons infected with HIV during a specified period of time, usually one year. This is difficult to measure because it relies on people getting tested for HIV and found to be positive close to the time of their infection. New diagnoses describe when people get tested, which may or may not be close to the time they were infected; consequently, HIV diagnosed cases include cases in persons who have been infected for longer periods.

Incidence rate: The number of new cases in a specific area during a specific time period among those at risk in the same area and time period. It provides a measure of the effect of illness relative to the size of a population.

Interpretation: The explanation of the meaning of available data and associated analyses.

Mean: The sum of individual scores in a data set divided by the total number of observations. The mean is what many people refer to as an average.

Median: The middle value in a data set. Usually, approximately half the values will be higher and half will be lower. The median is useful when a data set contains a few unusually high or unusually low values, which can affect the mean. It is also useful when data are skewed, meaning that most of the values are at one extreme or the other.

Percentage: A proportion of the whole, in which the whole is 100.

Prevalence: The proportion of a given population who have a disease, and are not known to have died, at a particular time. Prevalence does not indicate how long a person has had a disease and cannot be used to calculate rates of disease. It can provide an estimate of probability that an individual in a population will have a disease at a point in time.

Probability sampling: A technique that relies upon random selection to choose individuals from a defined population; all individuals have a known chance of selection. Types of probability samples include simple random sample, systematic random sample, stratified sample, and cluster sample.

Probability (p) value: The probability that a statistical result (an observed difference or relationship) could have occurred by chance alone. Statistical results usually are regarded as significant if there is less than five percent probability that the observed difference or relationship was due to chance alone. In such situations, the p value is said to be less than .05 ($p < .05$).

Proportion: A portion of a complete population or data set, usually expressed as a fraction or percentage of the population or data set.

Range: The values of the largest and smallest values in a data set.

Rate: A measure of the frequency of an event or a disease compared to the number or persons at risk for the event or disease. Usually, when calculating rates to include in the epidemiologic profile, the general population, rather than the population potentially exposed to HIV infection by various high-risk behaviors, is used as the denominator. The size of the general population is known from census data, whereas the size of the high-risk population is usually not known. For ease of comparison, the resulting fraction is usually converted to number of cases per 100,000 population. The choice of 100,000 is standard practice.

Sample: A group selected from a total population with the expectation that studying this group will provide relevant information about the total population.

Seroprevalence: The number of persons in a defined population who test positive for HIV infection based.

Statistical significance - A finding (for example, the observed difference between two samples) is described as statistically significant when it can be demonstrated that the probability of obtaining such a difference by chance only is relatively low.

Stratification: The separation of a sample into subsamples according to predetermined criteria, such as age group, gender, or literacy level.

Trend: A long-term change in frequency. A simple linear trend can be described by calculating how much the quantity being measured increased (or decreased) from the beginning value (at the beginning of the period) to the ending value (at the end of the period). A key factor is the statistical significance of the trend, which could be a problem if the annual values fluctuate widely from year to year, which would be likely for small numbers.

Core epidemiologic questions

To be effective and useful, the profile should describe the epidemic from various perspectives, including:

- Characteristics of the general population in the geographic area covered by the profile
- Characteristics of HIV-infected persons and persons engaged in high-risk behaviors
- Indicators of risk
- Distribution of disease (geographically and by population)
- Trends, if any

For the pilot epidemiologic profile, epidemiologic analyses were limited in several ways. Since data were abstracted that covered a limited period of time, in most cases it was not possible to examine trends over time. For some data sources, while it was possible to examine some general characteristics individually, it was not possible to look at cross-tabulations of characteristics because numbers were too small. In some cases, raw data were not available on which to conduct analyses, so aggregate data were presented as they are included in other reports. The remedy for some of these limitations will be to continue to collect data over a longer period of time from the same data sources.

It should also be noted that, in most cases, various cross-tabulations of characteristics were examined for all data sources. Only data that yielded items of interest or demonstrated differences between groups were included in the pilot profile.

Answering the core questions

Question 1 – What are the sociodemographic characteristics of the general population in the geographic area being described?

The general characteristics of the population in the geographic area being described provide background information against which the characteristics of the population of interest (in this case, person with or at risk for HIV) may be compared. Data will help to examine characteristics from two perspectives – demographics and socioeconomic status (for which indicators such as literacy level may serve as a proxy).

Suggested analyses (when possible):

Number and % distribution by:

- Age group and sex
- Literacy level and sex
- Urban vs. rural
- Literacy level and urban/rural
- Trends in these indicators over time, if possible

Question 2 – What is the scope of the HIV/AIDS epidemic in this geographic area?

A description of the extent and effect of the HIV epidemic in broad population groups in the geographic area allows a closer examination of the effect on specific groups in order to better focus prevention and care services. Analyses used in the Visakhapatnam profile are described below; additional site-specific data might also be relevant.

Suggested analyses (when possible):

AIDS surveillance data

For most recently available calendar year:

- Number of AIDS cases diagnosed in that year
- Number, percentage distribution, and rates of cases by sex
- Number and percentage distribution of cases by
 - Age group and sex
 - Risk category and sex
 - Marital status and sex
 - Job category and sex
 - Job category and risk
 - Case numbers/rates by geographic subunits

Other things to look at – dual infection with HIV and TB, HIV/AIDS mortality, trend data after data collection has happened for an adequate amount of time (annual number of cases, annual number of deaths, number of new perinatal cases).

Example 1. Transmission Route Among Visakhapatnam AIDS Cases (Jan-Dec 2002)*

Transmission Route	Males		Females		Total	
	Number	%	Number	%	Number	%
Adults						
Sexual	74	89.2	47	82.5	121	86.4
Blood/Blood Products	0	0.0	1	1.8	1	0.7
Infected Syringe	0	0.0	0	0.0	0	0
Other	7	8.4	4	7.0	11	7.8
Not Specified	1	1.2	3	5.3	4	2.9
Children						
Perinatal	1	1.2	1	1.8	2	1.4
Other	0	0.0	1	1.8	1	0.7
Total	83	100	57	100	140	100

*Source: Andhra Pradesh State AIDS Control Society

Blood bank data

For most recently available calendar year:

- Number and proportion of donors who test HIV positive of all donors
- Description of characteristics of HIV positive donors (sex, education, age, job category)
- Geographic distribution of HIV positive donors
- Number of individuals coinfecting with Hepatitis B, Hepatitis C, and syphilis
- Donation status of donors (voluntary vs. replacement)
- When available, comparison of data from year to year to determine trends

Care and support data

For most recently available calendar year:

- Characteristics of individuals receiving services (sex, age, occupation)
- Infection status of spouse
- Risk information (if available), cross-tabbed with demographic variables
- Proportion receiving medical treatment (and description of medical treatment, e.g., antiretrovirals vs. prophylactics)
- Coinfection with TB

PPTCT data

For most recently available calendar year:

- Number of new clients registered, number and proportion of new clients counseled about HIV, number and proportion of new clients tested for HIV, and number and proportion of clients testing positive
- Comparison of characteristics of HIV positive and sample of HIV negative clients (age, education, occupation, spouses occupation)
- Comparison of risk behaviors of HIV positive and sample of HIV negative clients
- Comparison of geographic distribution of HIV positive and sample of HIV negative clients
- Use of therapies to prevent HIV transmission, if available
- Pregnancy outcomes, if available
- When available, comparison of data from year to year to determine trends

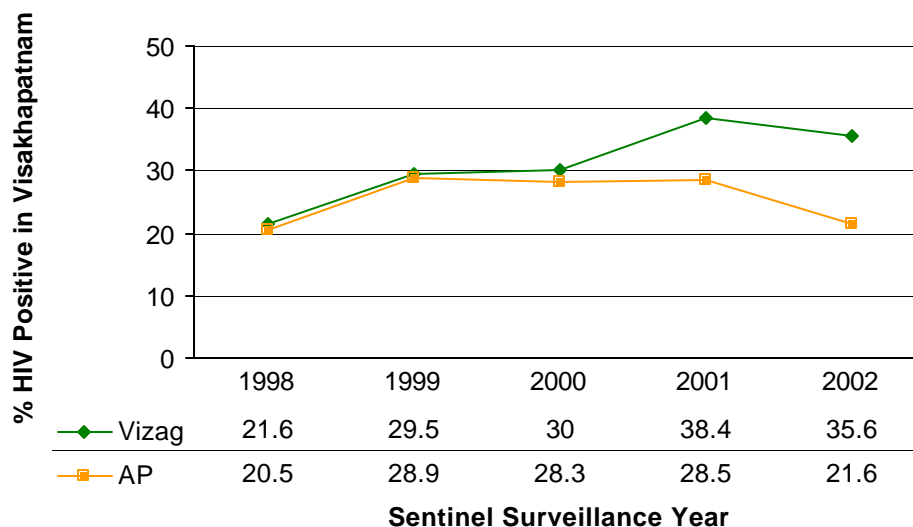
Sentinel surveillance data

For most recently available study data and for all populations studied:

- Number and percent of clinic clients testing positive for HIV at the different sites
- Comparison of data from year to year to determine trends for study populations with available data
- Distribution, by age and sex, of HIV positive clients
- Distribution, by locale and sex, of HIV positive clients
- Distribution, by literacy status and sex, of HIV positive clients
- Distribution, by occupation and sex, of HIV positive clients
- For STD clinic surveys, distribution, by STD and sex, of HIV positive clients

For studies that are conducted in clinics where only female patients are seen, for example, ANC clinics, cross-tabulations do not need to be done by sex.

Example 2. Trends in Visakhapatnam District STD Sentinel Surveillance Data (Andhra Medical College, Visakhapatnam) Compared to Statewide Trends, 1998-2002



VCTC data

For most recently available calendar year:

- Number and proportion of HIV positive clients among all clinic clients
- Comparison of characteristics of HIV positive and sample of HIV negative clients (sex, age, education, occupation, marital status)
- Geographic distribution of HIV positive clients
- Reason for VCTC visit, by sex
- Risk behaviors by sex for HIV positive and sample of HIV negative clients
- Comparison of data from year to year when available

Question 3 - What are the indicators of risk for HIV infection and AIDS in the population covered in this geographic area?

In this section of the profile, examine data on risk behaviors and markers of risk behavior (direct and indirect measures of risk for HIV infection and AIDS). Direct measures provide information about risk behavior that is directly associated with HIV transmission. Indirect measures describe other behaviors or conditions (for example, having a sexually transmitted disease) that indicate increased risk for HIV infection.

Suggested analyses:

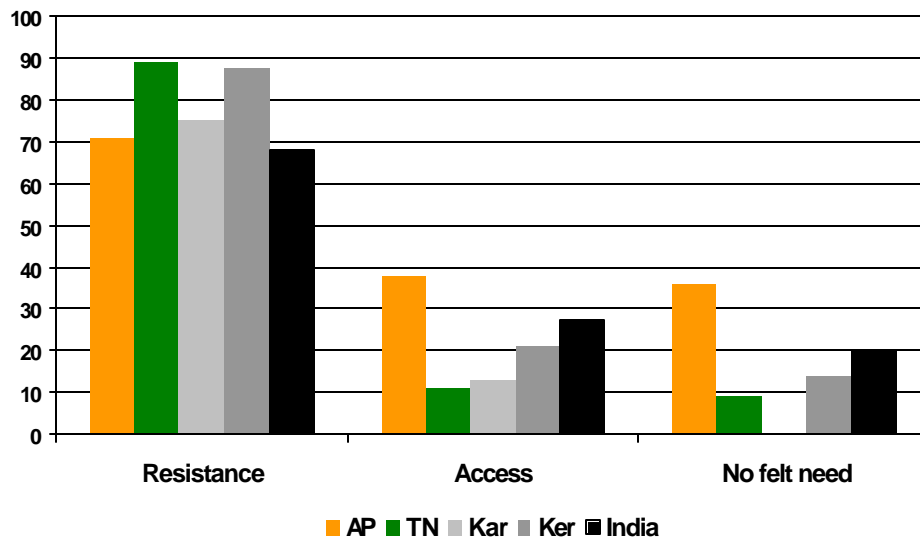
Behavioral Surveillance System (BSS) data

Results from the 2001 survey were available for inclusion in the Visakhapatnam profile. Aggregate statewide figures and tables created by NACO were available for Kerala,

Andhra Pradesh, Tamil Nadu, and Karnataka for both the general population and high-risk populations, including non-brothel based female sex workers and clients of sex workers. Analyses included in subsequent profiles will depend on which populations are included in future BSS surveys and what questions are asked. It is important to include both knowledge and behavior information for both the general population and any risk populations that may be surveyed. While the data may not specifically reflect responses from the district being studied, state data provide important context and information for the profile.

- Percent of general and risk populations that have heard of HIV
- Knowledge indicators among general population (e.g., HIV transmission routes, prevention strategies, linkage between STD and HIV)
- Behavioral indicators among general populations (e.g., non-regular sex partners, condom use)
- Knowledge indicators among risk populations
- Behavioral indicators among risk populations that may be specific to their risk (e.g., for sex workers, reasons for not using a condom at last sex with paying clients)
- Perception of risk
- STD treatment seeking behaviors

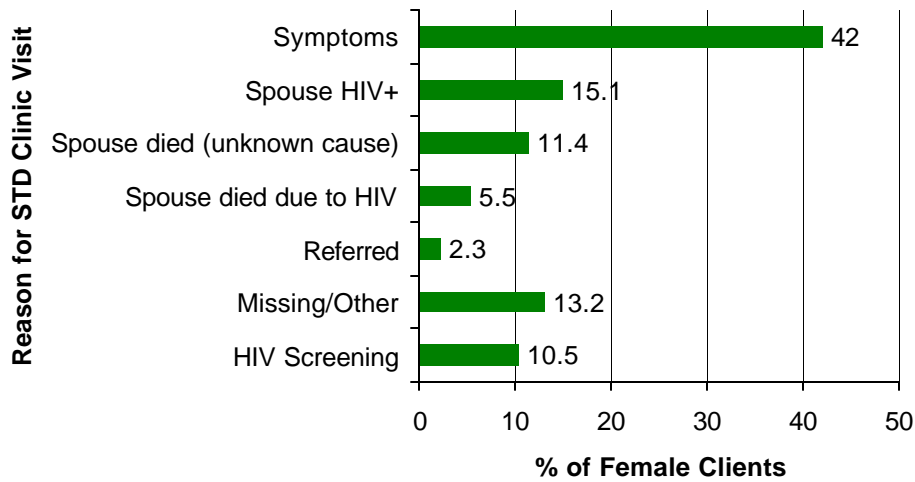
Example 3. Reasons for Not Using a Condom at Last Sex with Paying Clients, 2001 BSS Among High Risk Population*



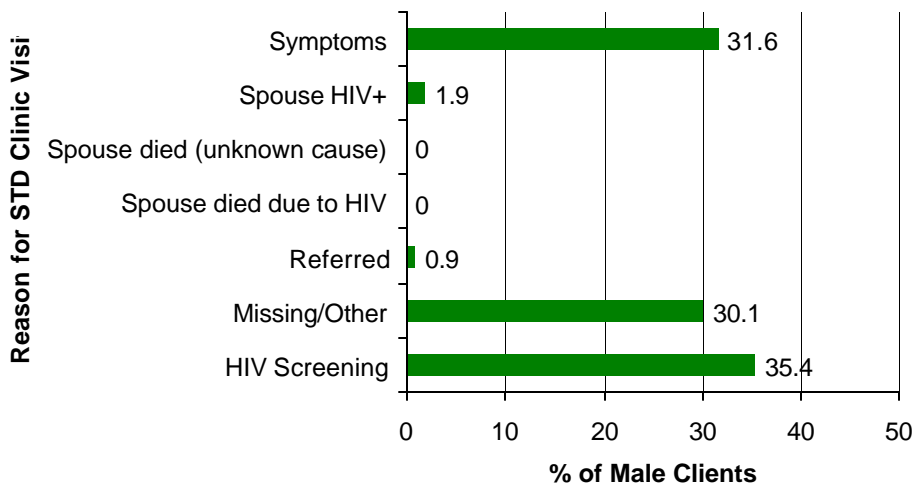
STD clinic data

- Description of characteristics of STD clinic clients (sex, education, age, occupation, marital status)
- Reason for visit by gender
- Risk behavior (for those with available information)
- HIV status of STD clinic clients (if available)
- STD diagnoses by HIV status

Example 4. Reason for Visit to KGH STD Clinic Among Female Clients (n= 219)



Example 5. Reason for Visit to KGH STD Clinic Among Male Clients (n=316)



Part 5 – Completing the profile (tables and graphs, maps, strengths and limitations)

Writing the remaining sections

Most of the data and interpretation of the data will be presented in the body of the epidemiologic profile using figures, tables, and narrative. It is important to include other sections in the profile to make the process easier to understand and to make the document easier to navigate and use. The remaining sections should include:

- Front matter
 - List of contributors
 - List of abbreviations
 - Executive summary
 - Table of contents
 - List of tables and figures

The executive summary is particularly important, because readers who do not have time or need to read the entire profile will use it. The executive summary presents the highlights of the document and conveys key points about the epidemic in the geographic area.

- Introduction
 - Background
 - Data sources
 - Strengths and limitations of the profile
 - Process followed in preparing the profile
- Conclusion
- Appendices
- Glossary of terms
- References

Creating tables and graphs

Summarizing data and presenting them in tables or figures is critical for an effective profile because raw data are difficult to understand, compare, and use to detect trends. Tips for presenting data in tables or figures include the following:

- Tables and figures explain the who, what, when, and where of the data. All relevant information needed to interpret the table or figure should be part of the table or figure so that the reader can understand without reference to the text.
- Figures are used to illustrate trends, relationships, or patterns. Tables are used to provide specific numeric values.
- Do not try to communicate too many ideas at once.
- Write clear, explanatory titles.
- Keep the table or figures uncluttered and free of unnecessary words.

- The labels on the axes of figures and the column headings and row entries in tables should be worded clearly and formatted consistently.
- Label all elements of a figure. If space doesn't allow you to label each element, include a legend.
- Three-dimensional graphs are harder to read and more likely to be misinterpreted than two-dimensional graphs.
- Make the scale appropriate for the findings you want to convey.
- Use the same scale for the y-axis when figures are meant to be compared.
- Use no more than 8 slices in a pie chart, and label all slices.
- When you present only percentages, include the total number (N). Do not chart percentages and numbers in the same graph.
- Name the sources of the data.
- In the accompanying text, refer to the key points of the table or figure; do not simply duplicate in words the content of the table or figure.

For an example of a table, see page 19. For an example of a figure, see page 21.

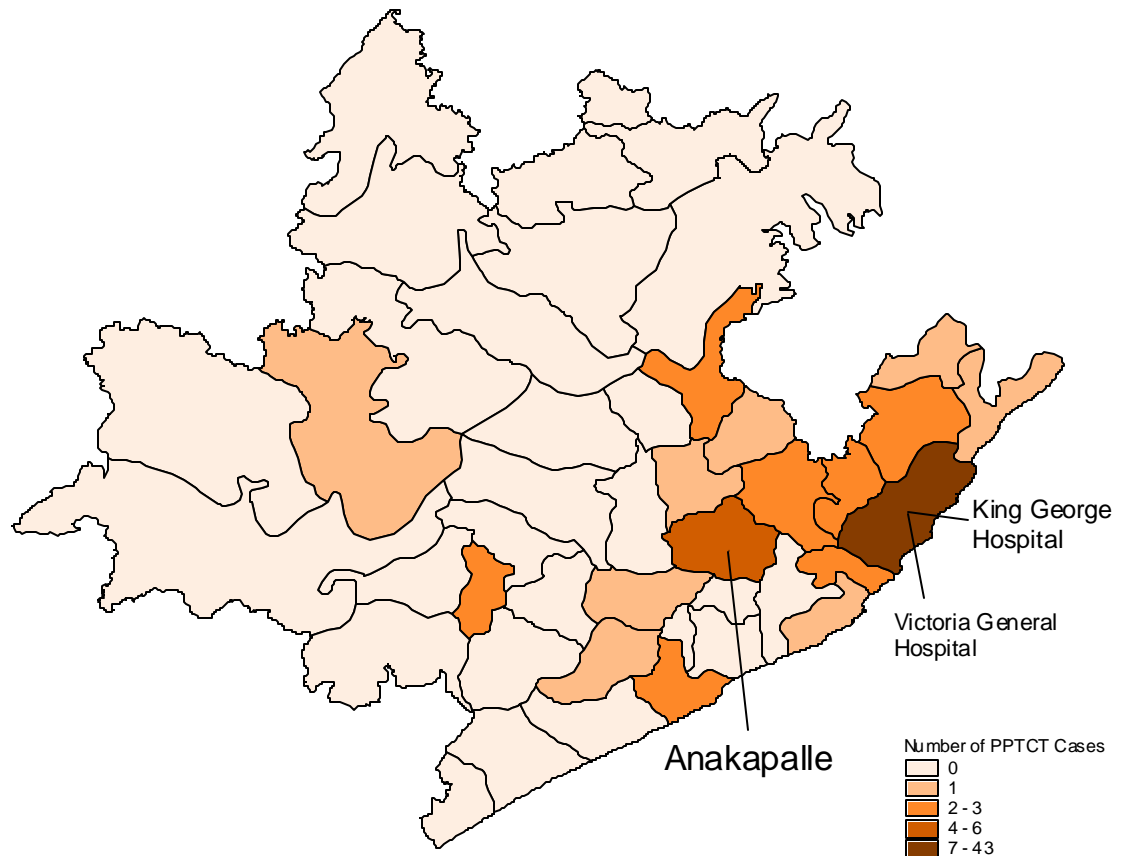
Developing maps

Maps are useful for showing the geographic location of events or attributes. Spot maps show where a disease or event occurred and area maps show either the incidence of an event in an area or the distribution of some condition throughout the geographic areas. Maps can help identify where HIV prevention and care services may best reach those who are at risk or infected.

It is important to recognize limitations to data presented on maps before drawing conclusions. This is particularly important when explaining potential associations between data. Residence at diagnosis may not be the location of risk-taking behavior.

For the Visakhapatnam profile, maps were generated using ArcView GIS 3.2. Maps can also be generated using EpiMap within Epi Info 2000. Even when mapping software is not available, maps can be generated manually.

Example 6. Distribution of PPTCT Cases by Mandal During 2002 Fiscal Year in Visakhapatnam District



*PPTCT sites include: King George Hospital, Victoria General Hospital, and Anakapalle Area Hospital

Issues of confidentiality

As is true of other methods of data presentation, with maps disclosure of information about individuals is a potential risk. Caution should be used when presenting geographic information in terms of protecting individual confidentiality and avoiding creation of stigma for certain groups or related to small geographic areas.

Describing data strengths and limitations

All data used to develop an epidemiologic profile will have strengths and limitations. It is important to describe the strengths, limitations, and data quality issues for every source of information used in the profile. Examples can be seen in Appendix A. It is also important to describe the more general strengths and limitations of the profile itself in the introductory section of the document.

Review

For development of the Visakhapatnam epidemiologic profile, it was important to have the document reviewed by a number of individuals, especially those at facilities from which data were abstracted. In preparing the Visakhapatnam profile, guidance and feedback on analytic output and narrative interpretation were provided primarily by AP SACS and NASTAD, and to a lesser degree by CDC-GAP/India. Feedback was also requested from the data collection sites themselves. All parties were asked to review the profile for accuracy of data presented, accuracy of interpretation, appropriate uses of data, and confidentiality concerns. Allowing people to provide feedback, especially the sites where the data are collected, increases the usefulness of the data as well as people's willingness to continue to work to improve the data, and to trust the data in the profile to guide decision-making.

Part 6 – Lessons learned from profile development in Visakhapatnam

Coordination needs

Development of the epidemiologic profile in Visakhapatnam required the efforts of a number of individuals in order to gain access to the data, abstract them, and analyze them – a full-time NASTAD consultant, other NASTAD team members, staff from AP SACS, the District Leprosy Officer, staff at the clinics and hospitals where the data were abstracted, and Community Medicine faculty at Andhra Medical College. Data abstraction and data entry activities, as well as examination of the data quality, were all coordinated by the consultant, who also did most of the work associated with writing the profile. In the future, in the absence of the consultant, it will be important to establish who the appropriate person will be to arrange access to data, coordinate data abstraction activities, and work with partners to maintain data quality. Ideally, this person will be someone with ties to the AP SACS office but who works in the field and has relationships with community medicine departments, government hospitals, clinics, NGOs, and other relevant organizations. Some discussion has indicated that the District Leprosy Officer may be the appropriate individual to coordinate profile development activities in the field.

It was also clear from profile development activities that it is helpful to have, in writing, agreements between parties involved in profile development to outline what the expectations are for parties' participation. Written authorizations for access to data and documents outlining objectives for the project and assigned responsibilities should be developed and agreed upon early in the process of profile development.

Resource needs

It will be important to build efficiencies into the process of profile development in order to reduce the use of time. The pilot project demonstrated that there is limited computer capacity in the field for data entry. Ideally, all clinic and hospital sites would be equipped to enter data on their clients into computerized databases on a regular basis, making data collection, overall, less burdensome because it wouldn't have to be done at one time. Since this is not an option at this time, it may be more efficient to centralize data entry to the AP SACS office, where there is computer capacity. Additionally, the more that pre-developed databases and previously planned analysis programs can be provided to staff, the less time will be required to prepare a profile.

EpiInfo was relied on heavily for use in profile development, for creating databases, data entry, cleaning data, and conducting analyses. Those who will be involved in these aspects of profile development in the future will need training in the use of EpiInfo.

Strengthening data sources

Since development of the Visakhapatnam epidemiologic profile was a pilot project and since individual-level data had never been extracted from clinic and hospital records for analyses for some data sources, much was learned about how data sources could be strengthened. Some of the areas for improvement included:

- Inconsistent use of forms across similar sites
- Lack of access to correct forms in some sites
- Use of forms that were not specific for the service being provided in some sites
- Patient responses did not always meet the intent of the questions asked
- Questions on instruments and responses to them were not being interpreted in a consistent way
- Incomplete data related to some variables, including those related to risk behavior (which are very important for prevention planning)

Recognition of these areas for improvement allows for development of solutions. Quality of the information collected from individuals should be checked regularly. Staff training on the intended use of the data collected and techniques for obtaining accurate data from clients would contribute to data quality.

Since data collection can be a resource-intensive activity, it will be important to conduct complete and accurate data collection activities on a limited number of variables that are most useful for program planning. It will also help to maintain electronic databases and conduct ongoing data entry over time in order to make data collection a more manageable process.

Identifying gaps in information

Due to the nature of the pilot project, data were abstracted for a limited time period and from limited populations (those receiving government services). Depending on what data would be most useful for HIV prevention and care planning, future data collection efforts could include:

- Continuing to collect information from the original sites in order to establish trends over time
- Collecting data on those who receive their services through the private sector or do not receive services at all
- Incorporating data from non-governmental organizations dealing with high risk populations
- Collecting data through focused, time-limited studies on specific risk populations to get more details about HIV-related behaviors
- Getting more complete AIDS surveillance information

The more that the quality and diversity of the information included in the epidemiologic profile increase, the more useful the document will be in planning HIV prevention and care services. However, it should be recognized that data will always be limited in some way, and that planning and implementation of services need to proceed with data that are available.

Part 7 – Using the profile for public health action

An epidemiologic profile serves as a starting point in responding to the epidemic. In addition to the data and analyses included in the profile, knowledge of HIV prevention activities and their effectiveness, an understanding of the various communities, and knowledge of local values and beliefs are needed in formulating policy decisions and an action plan.

Comparing infected and at-risk populations with those receiving services

The Visakhapatnam epidemiologic profile describes those who are infected with or at risk for HIV infection who received services in a limited number of government-funded facilities. These individuals had engaged in risky behavior that had resulted in or could have resulted in HIV infection. By describing the characteristics of these individuals, we can develop a clearer idea of those who need HIV prevention services that could contribute to decreases in risk-taking behaviors. These characteristics can be compared with the characteristics of people receiving prevention services to see where gaps may exist. These gaps may be described in terms of geographic location, literacy levels, certain occupations, risk behaviors, or certain demographic characteristics.

Abstracting data may also identify needs associated with delivery of clinic services. Collecting data for the pilot project revealed that many attendees of the STD clinic were actually coming to the clinic to be tested for HIV, and had to be referred to the VCTC. A possible outcome would be to conduct HIV counseling and testing in the STD clinic for this high-risk population.

Identifying additional areas for data collection when answers to specific, important questions are not available in the existing data

Abstracting data from existing clinic records will provide answers to some questions; the activity may also generate additional questions that cannot be answered by the existing data but may be important for planning prevention services. After characterizing populations most likely to be infected with or at risk for HIV, it may be important to collect more detailed information from certain populations on specific risk behaviors such as types of partners, practices, and condom use. It may be useful to ascertain the specific needs of these populations – what is the knowledge of this population related to the issue of HIV? Do members of this population have access to prevention services aimed at reducing risky behaviors? What are the barriers to them seeking out or receiving services – fear of stigmatization? Lack of reasonable availability? Beliefs about condoms? Attitudes about sex outside of marriage? What are community norms related to issues such as acceptance of HIV positive people or appropriateness of married people utilizing VCTC services or condoms, and how do they impact people receiving prevention services? Because data collection is a resource-intensive undertaking, it will be important to narrow additional questions down to the items of most importance.

Another way to get insight into some of these issues is to involve members of the risk populations in the planning of HIV prevention services. After characterizing the most at-risk populations in the epidemiologic profile, it may be useful to devise ways to get their input to allow for design of prevention programs that most effectively and appropriately reach these populations.

Evaluating HIV prevention and care services (process)

Ideally, the measure we would like to impact with prevention services is a reduction in HIV incidence. While it is difficult to measure new HIV infections, it is possible to collect information about HIV diagnoses in individuals attending clinics for services, as demonstrated in the pilot project. Collecting this information over time will contribute to knowledge about the continuing impact of HIV on these populations. It is also difficult to establish a causal relationship between prevention programs and trends in HIV; however, it is important to look at what prevention services are directed to these populations and then ascertain whether trends in HIV infection in these populations are changing, which may indicate association.

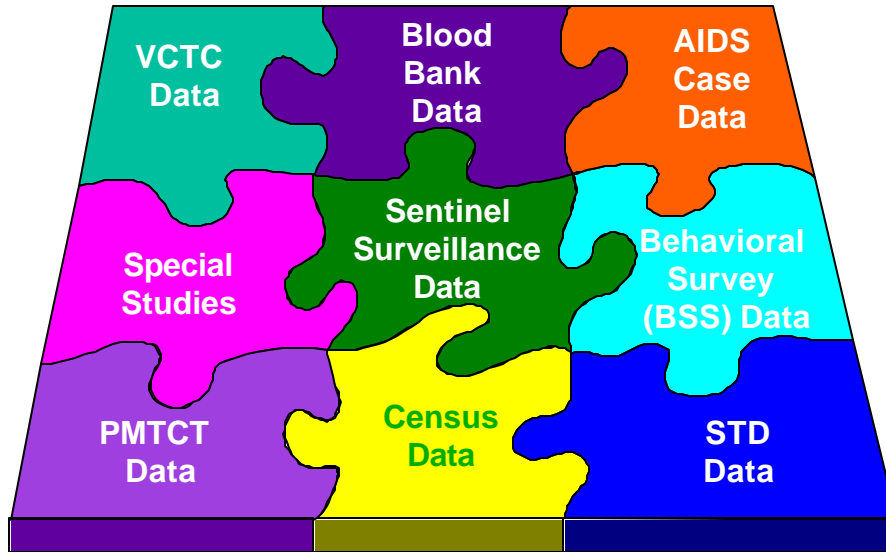
It is also important to measure changes in HIV-related knowledge and behaviors and to observe how prevention efforts may be impacting these indicators, in both general and high-risk populations. Some of this information is available from studies such as the Behavioral Surveillance System survey, as demonstrated in the pilot profile. Looking at these data over time, and looking at HIV prevention efforts directed at general and high-risk populations, may illustrate whether populations are being reached. Since it has been shown that increased knowledge does not always lead to behavior change, you will want to know if behaviors are changing over time. Again, the BSS includes questions about HIV-related risk behaviors, which if asked over time, could be used to demonstrate general trends in populations. Special, time-limited studies could be done with certain high-risk populations to observe whether individuals change their risk behaviors (or demonstrate intention to change these behaviors) as a result of receiving prevention services.

The delivery of ARV therapy at government clinics will present challenges in terms of resource utilization. It will also provide opportunities for collecting information on a greater number of individuals who have AIDS to characterize those who are receiving care services and perhaps describe improvement in health outcomes for those receiving the services. This will strengthen the profile in regards to care and treatment information.

To see an example of an epidemiologic profile, see the Visakhapatnam profile located in Appendix C.

Appendix A

Putting together the puzzle



First steps in the creation of
an HIV/AIDS epidemiologic profile
in Andhra Pradesh
(Abridged version)

A collaborative effort of NASTAD and APSACS
May, 2003

INTRODUCTION

In discussions with NASTAD, development of an epidemiologic profile of HIV/AIDS in Andhra Pradesh was identified as an area of interest by the APSACS office and a request for technical assistance was made. As a result, two epidemiologists with experience in HIV/AIDS surveillance and epidemiologic profile development were added to the technical assistance team. It is believed by all parties that development of this profile will aid in describing the characteristics and magnitude of populations at high risk for HIV infection, which will allow for better targeting of HIV prevention and care services.

Why develop an epidemiologic profile?

- To provide a thorough description of the HIV/AIDS epidemic in various populations
- To describe the current status of HIV/AIDS cases in the service area and provide some understanding of how the epidemic might look in the future
- To identify characteristics of the general population and of populations living with or at high risk for HIV/AIDS in defined geographic areas who need primary and secondary prevention or care services
- To provide information required to conduct needs assessment and gap analyses

How can an epidemiologic profile be used?

- To help develop a comprehensive HIV prevention or care plan
- To set priorities among populations who need prevention and care services
- To set priorities among prevention and care interventions
- To provide a basis for determining and/or projecting future needs
- To develop requests for proposals for providers and subcontracts
- To increase general community awareness of HIV/AIDS
- To disseminate data to providers
- To frame research and evaluation questions
- To apply for and receive funding
- To respond to public needs (educators, grant writers, media, policy makers)

What are the seven steps involved in developing an epidemiologic profile?

Step 1 – Determine profile scope and parameters and establish collaboration.

This step involves consulting with planning groups and potential stakeholders, identifying what planning jurisdictions and service areas should be included, figuring out what specific questions for prevention/care planning need to be addressed, and determining what resources are available.

Step 2 - Obtain core data and available supplemental data.

This step involves identifying what data are needed for development of the epidemiologic profile and what data sources are available to meet these needs. All data should be evaluated for completeness, representativeness, age, timeliness, and validity and their strengths and limitations should be described.

Step 3 – Analyze and interpret the data.

Data should be analyzed in such a way that questions of interest are answered to the extent possible in the epidemiologic profile (and gaps in data can be identified). These questions could include the following:

- What are the sociodemographic characteristics of the general population in your geographic area?
- What is the scope of the HIV/AIDS epidemic in this geographic area?
- What are the indicators of HIV infection risk in the population covered by the geographic area?
- What are the HIV services utilization patterns of the individuals in this population?
- What are the number and characteristics of individuals who know they are HIV-positive but who are not in care?

Step 4 – Present data in user-friendly formats.

Step 5 – Draw overall conclusions and write effective and useful narrative.

Step 6 – Write remaining sections (executive summary, appendices, glossary, references) and compile the complete profile.

Step 7 – Prepare clear presentations for appropriate audiences.

OBJECTIVE

In order to get started on the process of epidemiologic profile development for Andhra Pradesh, TA providers from NASTAD decided that it would be useful to work with APSACS staff to conduct **Step 2** of profile development. Together they worked to identify existing sources of data to be used in an epidemiologic profile and described, to the extent possible, the characteristics of these data sources.

METHODOLOGY

In an effort to describe existing sources of data to be used in an epidemiologic profile of HIV/AIDS in Andhra Pradesh, NASTAD TA providers, staff at APSACS, and Dr. Dora Warren first created a list. The list included the following sources:

- AIDS case data
- Blood bank data
- Behavioral Surveillance System (BSS) data
- Care and support center data
- Census data
- Community based STI study data
- DFID study data
- National Female Health Survey (NFHS) data
- Prevention of Mother to Child Transmission (PMTCT) data
- Sentinel surveillance data
- STD clinic data
- TB-HIV data
- Voluntary counseling and testing (VCTC) data

The TA providers then compiled a list of questions to ask about each of the listed data sources in order to ascertain how they would be useful for incorporation in an epidemiologic profile. It was not possible to answer every question about every data source; the list just provided structure for the conversations about the data sources. Over the course of several days, they met with APSACS staff in order to create a data inventory. The list of questions is as follows:

Data Review Worksheet

Name of the data source:

When was data collection initiated?

Are the data collected on an ongoing basis? If not, how often are they collected?

Where are the data collected (sites)?

Sample size?

Which clients are data collected from? Who is missing? (can clients refuse?)

Are there written protocols available for how the data are collected? (If so, attach if available)

Who is collecting the data?

How are they trained for data collection – is there standardized training?

When during the course of a client's care/visit are the data collected?

Is there a data collection form that gets filled out? (if available, attach)

How is the form filled out by the person collecting the data?

What data elements are collected?

Have these changed over time or is there intention of changing them?

Are data collection forms reviewed for completeness or accuracy by anyone?

Are duplicate data collected on clients?

Are lab data included?

Are standardized protocols followed for specimen processing?

Are lab results matched to other data? How?

Are data entered into an electronic system? Describe.

Is this system provided by NACO?

Is there any ability to make local changes to this system?

Who enters the data?

How are they trained?

Are the data analyzed after they are entered into the electronic system?

Who performs these analyses?

How are they trained?

When data come into APSACS and then sent to NACO, are they retained locally?

If not, is it possible to get a local database for a analysis from NACO?

What software is used for different types of analyses?

What types of reports are created from the data? (if available, attach an example)

Do these reports contain tables, graphics, maps, narrative interpretation?

How often are these reports created?

Who creates these reports?

Who reads these reports?

Aside from these reports, are other requests received for these data? From whom? What elements are requested?

Other notes:

The list of data sources is not exhaustive; other sources could be included in an epidemiologic profile, such as vital statistics data (to describe mortality) and hospital discharge data. The information collected about each of the data sources is also not exhaustive. However, this information should provide a good starting point for compiling an epidemiologic profile

Description of the Computerized Management Information System (CMIS)

Under direction from the National AIDS Control Organization (NACO), ORG Centre for Social Research developed the Computerized Information Management System (CMIS) to allow those working at the state and national level to:

- continuously assess project planning and implementation; and
- critically analyze different aspects of implementation, efficiency and effectiveness, time and cost overruns, and impact

CMIS was installed at each of the 38 state/municipal AIDS Control Societies and is used to monitor HIV -related activities at blood banks, STD clinics, voluntary counseling and testing centers, NGOs, and other sites. NACO provided training to the SACS/MACS on the use of the system, and placed a staff person in Andhra Pradesh to operate the system (while the assignment was to be for a year, this person has remained with APSACS). Data are received in aggregate form, on paper, from the various project sites. At some point in the future, the project sites will have the capacity to enter data and then transmit data to the state electronically. Currently, the aggregate data are received at the state office and manually entered into the CMIS system. The server for the system is located at NACO, and there is a provision for direct data entry on the web. Once data are entered into the system, they are available for analyses. The system has the ability to produce standard output reports in tabular and graphical form, including maps. There is also the ability to run ad hoc reports based on different variables.

RESULTS

AIDS Case Data

Background: Data collection was initiated in 1989 using the WHO case definition. In 1994, NACO distributed a clinical case definition to be used by providers (see attached). There is a requirement for physicians to report cases, but this has not been adhered to in a uniform way. Cases are reported from hospitals and also by a very limited number of physicians.

Data elements collected: The individual-level form contains the following data elements: sociodemographic characteristics, risk factors and practices (including sex with multiple partners and prostitutes and sex for incentives), history of STDs, drug use, syringe sharing, pregnancy information for women, AIDS symptoms and signs, lab tests performed, and HIV tests. On the aggregate form, information about mortality status, age by gender, mode by gender by adult/child, presenting signs and symptoms, and opportunistic infections is collected (see attached forms).

Data flow: The individual-level forms are completed by medical officers at 52 major hospitals. These forms are supposed to come to the state office but for the most part, the majority of clinicians are not aware of the form and are not filling it out. The aggregate forms come to APSACS, and reporting from hospitals has improved with encouragement from the state office – at this time, 30 to 35 of the 52 hospitals are sending these forms. This information is sent from the state office to NACO. All forms and data are received in paper format.

Strengths and limitations: AIDS case data describe those who are at advanced stages of HIV disease. Since many clinicians are not aware of the case report form and consequently don't fill it out, AIDS cases are very underreported. As with other data sources, individual level information is not received by the state; only aggregate level data are received, allowing for limited analyses.

Suggested analyses: It would be useful to describe AIDS cases by gender x mode x age category to the extent possible. Since cases are so underreported, it will be difficult to look at trends over time, although this would be important to look at if data are more complete in the future. It would be ideal to get the individual-level data since these provide more detail on risk behaviors.

Behavioral Surveillance System Data

Background: The BSS in the general population and high-risk groups (female sex workers and their clients, MSM, and IDUs) is proposed to be undertaken three times during the period 2001-2005. The first survey was conducted between October 2001 and March 2002 in 32 states and union territories. The target population for the BSS in the general population included men and women between the ages of 15 and 49 years from urban and rural areas. The target population for the BSS in high-risk groups in AP included non-brothel based female sex workers and clients of sex workers.

For the BSS in the general population, data were collected from April to December 2001. States were categorized into sampling units and for each sampling unit, respondents were drawn based on a cluster sampling format, with 3832 total respondents, equally divided between males and females, urban and rural. For the BSS with female sex workers and their clients, data were collected between October 2001 and March 2002. States were categorized into sampling units and respondents were selected via a cluster sampling format. For AP, 279 female sex workers and 277 clients were interviewed.

Data elements collected: The core indicators for BSS include: awareness of HIV/AIDS; awareness of STD, STD prevalence and treatment seeking behavior; sexual behavior and condom usage; and exposure to mass media. For the BSS in female sex workers and their clients, characteristics of each population (such as marital status, education, etc.) were collected. Interview instruments are located in the study reports at APSACS.

Data flow: Data were collected nationwide by an external agency (ORG-CSR) contracted by NACO for conducting the baseline, mid-term, and end evaluation of BSS. Data were entered in the Integrated System for Survey Analysis (ISSA) database and analyzed using SPSS (version 10) statistical package. Reports were distributed by NACO and include state-level information.

Strengths and limitations: The observations of the national baseline BSS survey provide basic information needed to strategize and prioritize prevention programs and interventions. Data are broadly collected via a standardized protocol. Sampling is done in such a way as to try to capture a representative sample, particularly in regards to the general population. It doesn't appear that survey data are weighted to the general population. Generalizability is more limited for the study done in female sex workers and their clients, but the data are still valuable.

Suggested analyses: Analyses are limited to those that are published in the reports. These analyses would be useful to include in the epidemiologic profile and for the general population include crosstabs of awareness, STDs, sexual behavior and condom use by gender and geographic location. For female sex

workers and their clients, there is useful information available to describe these individuals from limited areas in AP in regards to demographic and socioeconomic characteristics, awareness, STDs, sexual behavior, and condom use.

Blood Bank Data

Background: Data collection at blood banks started in 1998. At this time, data were consolidated manually. As of 4/2002, data entry into CMIS was fully implemented. There are 169 blood banks statewide. Of these, 137 report (81%). Of the 137, 46 (34%) are government blood banks and the remainder are private or voluntary/charitable.

Data elements collected: Staffing and training details, status of availability of equipment and consumables, stock position of testing kits and blood bags, status of testing of blood units (blood samples are tested for HIV, hepatitis B surface antigen, hepatitis C antibody, VDRL, and malaria). Samples can be described by positivity and whether the sample was a replacement sample or from someone who donated on a voluntary basis. There is also some information on gender and type of donation (see attached form).

Data flow: Data forms are completed by the medical officer or designee (lab technicians, other assignees). Training for completion of these forms is provided at the state level. Data are sent in paper format to APSACS monthly, entered into CMIS, and sent to NACO.

Strengths and limitations: This is one of the few sources of data that describes those receiving services from the private sector as well as the public sector. Some information on the forms is left blank and there is no follow-up, so all data elements are not complete. The state does not receive individual level data so information on gender and age and risk is not available at the state level for analyses, although it is available at the blood banks. Characteristics of those who get their blood tested for replacement reasons may differ from those of people who donate on a voluntary basis (this population may consist of more individuals of lower socioeconomic status). There is no way to eliminate duplicates, and geographically, people are described by where they donated blood.

Suggested analyses: With these data, there is the ability to describe the number of seropositives by geographic location of the blood bank (which may or may not reflect residence) and whether the samples are from volunteers or replacement donors. Eventually, it will be useful to look at trends over time. Information about the number of people testing positive for other diseases is useful, in combination with HIV and alone.

Care and Support Center Data

Background: Data collection for care and support centers started in April 2002. Data are collected from 13 NGOs and one TB/ID government hospital; two of these facilities are drop-in centers. The NGOs have very limited capacity (about 20 beds each). People come to the NGOs for both inpatient and outpatient care. Approximately 10,000 AIDS cases have been reported to date.

Data elements collected: Number of new PLWAs registered, admitted, and who have died, by gender; services provided, by gender; new cases, sex x age, deceased cases, sex x age; symptoms by status; age x sex x mode of transmission (and sexual behavior is broken out by type of sexual behavior, which is not always the case); disease described by those admitted vs. those receiving outpatient or home care; symptoms and opportunistic infections (see attached form).

Data flow: Data collection is being done by doctors at each of the NGOs. While individual-level information is collected by each NGO, data are recorded on an aggregate form and the form is sent in paper format to the state office and then on to NACO. While data are not yet being entered into CMIS, they will be entered soon.

Strengths and limitations: AIDS data are very useful for describing severe morbidity and mortality related to HIV disease. There is some useful information about age x sex x mode of transmission that describes those diagnosed with AIDS, and this is the only CMIS-related data that are cross-tabbed in this way. However, there is some question about the accuracy of this information since sometimes the totals are not consistent throughout the aggregate form. This is a source of data about opportunistic infections that may be useful. However, NGOs have a limited capability to diagnose OIs (with the exception of candidiasis, TB, and herpes zoster). The biggest limitation is that people who receive services are a very small representation of those with AIDS in AP, and AIDS is very underreported.

Suggested analyses: If accuracy could be maintained, the information on age x sex x mode of transmission would be very useful to describe those who have AIDS. Information about mortality and opportunistic infections could also be described.

Census Data

Background: Census data are collected every 10 years. The last census was conducted in 2000. For intercensal years, population estimates are calculated based on growth rates.

Data elements collected: The standard information describing demographic characteristics, socioeconomic characteristics, and geographic location is collected (see attached form).

Data flow: Data are available for AP via reports and the internet.

Strengths and limitations: Census data are the most complete and descriptive data that can be found to characterize the general population. However, they may underrepresent populations that are more disenfranchised or mobile.

Suggested analyses: Standard analyses of demographic and socioeconomic indicators should be conducted for the state as well as the districts in order to describe the general populations and ascertain if HIV has a disproportionate impact on certain populations.

CHARCA Project Data (Coordinated HIV/AIDS Response Through Capacity-Building and Awareness)

Background: This project was a multi-agency effort of UNAIDS agencies, state governments, civil society and women's groups. The project was implemented in six districts in India where young women (ages 12-30) were identified to be at heightened risk. The project was conducted from August - December, 2002, and the site that was randomly selected for AP was Guntur. The study was population based, with 20 sites in Guntur covering 3,200 individuals. Both urban and rural sites were included, and the choice of sites in Guntur was based on differing levels of development status (economic, educational, social). Components of the study included focus groups, in-depth interviews, and polling groups. Study subjects included young women, young men, guardians, NGOs, religious leaders, service providers, opinion leaders, and others.

Data elements collected: Information was collected on knowledge, awareness, skills, service support structure, environment, and factors related to vulnerability to disease, such as premarital sex, abuse, and alcohol use.

Data flow: UNAIDS collected the data and provided some consolidated cross-tabbed data back to the individual states.

Strengths and limitations: Stakeholders and researchers were involved in working out the research design and methodology and selecting sample sites. Information was collected from a broad variety of individuals, and since there was so much participation and buy-in in regards to data collection, it probably made the data more credible to members of the represented communities. The data collected were largely qualitative, so the size of the problem or the extent of its prevalence was difficult to estimate. Recall bias may affect data collected from individuals. All analyses were based on subjective perceptions and disclosure, and the study didn't include every known subgroup, so may not be representative of all views and ideas.

Suggested analyses: Analyses have been conducted in order to describe factors enhancing and reducing vulnerability of the population of interest. Topics examined include STI/HIV/AIDS-related protective knowledge and skills, life skills related to sexual activity, alcohol consumption, occupational related factors, support structures/individuals, STI/HIV/AIDS services, environmental factors, and gender/economic/social exclusions. These data would be valuable to include in a district epidemiologic profile for Guntur.

Community-Based STI Study Data

Background: Data for the most recent community-based study were collected September-November, 2002. The study was organized and conducted by NACO. Sites from all over the country were randomly selected. In AP, sites from four regions were selected and 300 people were sampled in each region, so the total sample size for AP was 1200 people. Fifty men and 50 women were sampled per district.

Data elements collected: Elements included sociodemographic characteristics; STI/HIV/AIDS-related knowledge, attitudes, and beliefs; STI history and treatment; and limited sexual behavior. Biological specimens of blood, urine, and genital fluids were also collected from respondents and tested for HIV, syphilis, gonorrhea, herpes, and other STIs (see attached form).

Data flow: Data were collected by NACO, and states have not yet been provided with state-level information.

Strengths and limitations: This study was conducted using a standardized protocol and covered a broad distribution of sites in the country. Collecting both behavioral and biological data from the respondents should yield some useful information for prevention planning. It is unknown how representative these individuals are of the general population in AP. In regards to analyses of those who are HIV infected, the number of those testing positive in this study was small (34), so it may be difficult to really perform any analyses on these individuals.

Suggested analyses: The most useful data from this study are those that describe prevalence of STIs in a randomly selected sample of the general population. It would be interesting to compare these data to data from those who voluntarily attend STD clinics.

DFID Study Data (Department for International Development)

Background: The Impact Assessment Project implemented in the state under the overall guidance of APSACS with technical assistance of Family Health International and funding from Department of International Development included the following studies: 1) HIV risk behavior surveillance survey, including female sex workers, their clients, male industry workers, fishermen, male university students, and female workers in the tobacco industry (end 1999-early 2000); 2) STI prevalence survey among female sex workers (April 6- May 30, 2000); and 3) knowledge and practices of health care providers (November 1999-September 2000).

Sampling of the different populations was done via a cluster method, and sample sizes were chosen based on estimated levels of key risk behaviors for certain parts of the study, as well as the ability to measure changes in indicators. Sampling methods are described for each part of the study in the books available at APSACS.

Data elements collected: Knowledge of HIV/AIDS/STI; sexual behaviors; condom usage; clinical findings; prevalence of HIV, STIs and Hepatitis B; and gaps between knowledge and practice of health care providers in STI case management.

Data flow: Data entry and data analysis were conducted by the Impact Assessment Project. The health care provider survey was implemented by Social and Environmental Research Division of Blackstone Market Facts (a market research agency). The HIV/STI survey among female sex workers was implemented by the Community Health Awareness and Natural Green Environment Society (CHANGES – Non-Governmental Organization). The HIV risk behavior surveillance survey was implemented by Taylor Nelson Sofres Mode Ltd. (TNS Mode).

Results are available in a series of booklets available at APSACS.

Strengths and limitations: These surveys provide useful information about the prevalence of STIs among female sex workers, the behaviors of certain high risk groups in the state, and the manner in which STI patients are managed in health care settings. The surveys followed methods outlined by UNAIDS/WHO for evaluating and monitoring HIV/STI prevalence studies. Conducting the studies involved a strong collaborative effort involving governmental institutions, non-governmental organizations, laboratories, and the private sector. Sampling methodologies were carefully thought out. Not every analysis of interest is available in the results book, and it is unknown how representative the data are for the risk populations.

Suggested analyses: For the epidemiologic profile, results described in the study booklets, particularly in regards to the risk populations, would be useful.

NFHS Data (National Female Health Survey)

Background: The NFHS is a general health survey of young ever-married women between the ages of 15 and 49 that was conducted by a contracted survey company located in Mumbai in 1998-1999 across India.

Data elements collected: The NFHS collects information about a broad variety of health issues, as well as demographic and socioeconomic information about young married women. General information includes age, residence (urban vs. rural), education, religion, caste/tribe, standard of living index, and exposure to mass media. HIV/AIDS related questions address knowledge of AIDS, sources of information about AIDS, beliefs about whether AIDS is preventable, and what precautions individuals can take to avoid infection. HIV/AIDS – related questions for AP were not included in the first phase of the study.

Data flow: These data are collected at the national level and the contracted survey company compiles the results. Results are provided for the entire country as well as each of the states in the form of written reports. These reports are available via the internet.

Strengths and limitations: These data provide important information about HIV/AIDS related knowledge and beliefs of young ever-married women and the data are available at the state level. Women's knowledge and beliefs can be examined by the available demographic and socioeconomic characteristics. It is unclear how subjects were selected, how interviews were conducted, how many young women were interviewed in AP in each phase of the study, and how they were geographically distributed. Since it is unknown how sampling was done, it is difficult to say how generalizable these data are to the population of young, ever-married women. It doesn't appear that survey data are weighted to reflect this population. Data are not collected about behaviors in this population.

Suggested analyses: Analyses are provided in report format and results could be incorporated as they are into an epidemiologic profile. Although the data are from 1998-1999, they are still useful in providing some information that would be important for prevention planning.

Prevention of Mother to Child Transmission of HIV (PMTCT) Data

Background: Prior to 2002, there was only one PMTCT site established on an experimental basis in order to perform HIV counseling and testing with pregnant women. In 2002, 14 centers were added at medical colleges. Recently (4/2003), 23 more PMTCT sites were added at government hospitals.

There are approximately 1.6 million pregnancies every year in AP. The PMTCTs provide services to about 100,000 pregnant women (1/16 of the women) who go to these hospitals for prenatal care and delivery. Of the 100,000 counseled, about 75% were tested for HIV.

The data collection form was revised in December 2002 with input from UNICEF but essentially, the data elements are the same.

Data elements collected: Number of women who are counseled and tested for HIV, who are HIV+, who receive post-test counseling, who receive their test results, number of spouses/partners who receive counseling and testing and are HIV+, number of women who are counseled and tested at time of labor, receipt of nevirapine by mother-child pairs, information about delivery and follow-up care of babies (see attached form).

Data flow: These data are tallied at the PMTCT sites and the forms are sent to APSACS. Data are then sent in paper format to NACO. Data are being entered at the state level in Excel.

Strengths and limitations: Collection of data from this population acts as a proxy for HIV seroprevalence in the general population. Data collection from this population is ongoing and provides a large sample size. This is a good source of information about followup of infants who are born to HIV+ mothers who receive services at these sites. These data describe only women who receive services at the PMTCT sites, which may not represent all women who are pregnant. There is no descriptive information collected on this form about the women. These data overlap with the data collected for the sentinel surveillance studies.

Suggested analyses: For the purposes of the epidemiologic profile, these data could be used to describe seropositivity in pregnant women. These data also provide some information on seroconcordant/serodiscordant couples and the number of babies who seroconvert, although it may be difficult to actually do these calculations because only tallies are provided, not information on individual clients.

Sentinel Surveillance Data

Background: Sentinel surveillance started in AP at a number of sites in 1985/86. Over time, and under the guidance of NACO, collection of sentinel surveillance data became more systematic in 1998 and data are collected in sites all over India. Antenatal clinics (ANCs) were chosen for sentinel surveillance to use as a proxy to measure prevalence of the epidemic in the general adult population. STD clinics were chosen for sentinel surveillance to use as a proxy to measure prevalence of the epidemic in the high risk population. In a limited number of states, the study is also conducted in commercial sex workers, men who have sex with men, and injection drug users. All sentinel sites are located in teaching hospitals, district hospitals, and area hospitals. STD clinic sites that serve more than 100 people in a month were chosen, as were ANC sites that serve more than 500-600 people in a month. Data are collected during a three month period, usually August through October, although in 2003 data will be collected June through August.

In AP, initial data collection sites included three STD clinics and three ANCs. In every year following, data have been collected at an increasing number of sites. Currently, data for AP are collected in 11 STD clinics and 23 ANC clinics. These include the clinics that got sanctioned by NACO (9 STD and 15 ANC) and clinics added by APSACS.

At each of the ANCs, 400 samples are collected. These are consecutive clients who are getting blood drawn for other health care services. At each of the STD clinics, 250 samples are collected from consecutive clients who are getting blood drawn for other clinic services (such as syphilis serologies). All samples are blinded and tested anonymously for HIV.

NACO developed the methodology for conduct of the surveys and provides guidance and training to the SACS, which in turn provide training to medical officers and lab technicians to perform the survey.

Data elements collected: Data are collected on each client and include gender, age, occupation, literacy level, migration status, and geography based on residence (see attached forms).

Data flow: At the site, data are collected on individuals. These data are consolidated onto a single form at the site and all forms are sent to APSACS at the end of the study. These forms are then sent from APSACS to NACO. In 2002, the state retained the data from the aggregate forms, entered these data into Excel, and analyzed them locally. At the NACO level, the data are consolidated and reported out in an official document (2001 and 2002 reports have still not been released).

Strengths and limitations: Sentinel surveillance data have been collected on an annual basis since 1998 and data have been collected at an increasing number of sites, increasing the sample size for the state. Data are collected in a standardized way, and include information about some important demographic characteristics. Nonresponse/nonparticipation bias is diminished by the fact that it is a blinded study and people do not self-select to be included in the study. APSACS retains the consolidated data so has some ability to do local analyses. Sentinel surveillance data describe only those who come for services at government sites, so may not represent the entire population of those who are infected with HIV. Characteristics of the individual ANCs and STD clinics are not available to describe the clients and ascertain how representative they are of the population being studied. Risk information is not collected but assumed based on clinic site; data collection is limited to information routinely collected in that particular clinic setting. There is no way to know if duplicates are included in the data. ANC data probably have overlap with data collected from the PMTCT sites.

Suggested analyses: Local analyses have included examining the characteristics of those who are HIV infected by gender, age, migration status, literacy level, and residence. It may be useful to compare the characteristics of those who test HIV positive to those who do not to see if the populations have different characteristics. It may also be useful to stratify the data further to look at the characteristics of, for instance, men versus women. Stratification should probably be conducted at the state level, not district level, for those who are positive because of the relatively small sample size. When looking at trends over time, particularly for individual sites, it would be important to calculate confidence intervals to see if trends are significant since the annual sample sizes are relatively small.

STD Clinic Data

Background: STDs have been reported since 1947, but are not notifiable conditions in India. The STD program came under the control of the SACS in 1992.

Data are collected on an ongoing basis. Since 1999, there were 28 STD clinics, 10 in medical colleges and 18 in districts. Sixty clinics were added in districts that were not covered and area hospitals to make 88 total clinics in the 23 districts.

There are 40 million new STIs contracted every year in India, and only about 5% are reported from the public sector. There is a five to seven percent annual incidence of STIs in the sexually active population (defined as 15-59 year olds) in both India and AP. This works out to 2.6 million cases for AP every year.

Data are mainly from medical colleges, where lab support is better. Quality of data is not as high in district hospitals and area hospitals.

Data elements collected: Individual-level data are collected on STD cards at the clinic; these data include identifying information as well as treatment and diagnosis, contact information, history, exposure, counseling services. An aggregate form is also filled out at the STD clinic. This aggregate form contains information on details of facilities at the clinic; details of staff; status of staff, equipment, consumables, and essential medicines; details of patients/cases at the clinics (new and old patients, by gender and attending vs. referred clients); details of cases detected (type of disease by married and single males and females); laboratory investigations (types of lab test for different diseases); and details of condom distribution, partner notification, and counseling services (by gender) (see attached forms).

Data flow: Data from the individual-level STD clinics that are recorded on the STD cards are retained at the clinics. These are not in a computerized system. Data that are received at APSACS are the aggregate data. These aggregate data are entered into CMIS and sent on to NACO. Effort is being made to obtain the more detailed individual level data.

Strengths and limitations: STDs are an indirect indicator of unprotected sexual activity that could result in transmission of HIV. Due to the considerably shorter incubation period for STDs, these data may provide more timely information on behavioral trends. Risk information is not included on the consolidated form received by the state, and it would be useful to have the more detailed risk information that is included on the STD cards at the clinic. People diagnosed with STDs are described geographically by where they live, not where the clinic is located, which is a strength. STD clinics that are required to report to the state are located in general hospitals, district area hospitals, and area hospitals. Data

are not included from private providers or clinics, where the majority of people with STDs are being seen. While the consolidated format does not allow for the elimination of duplicates, duplicate data are unlikely.

Suggested analyses: For purposes of the epidemiologic profile, the information from the consolidated form that would be most useful is the description of people with various STDs by gender. Individual-level information would allow for more useful analyses, which would include sexual behavior of high risk persons; age x gender x disease prevalence; types of STIs associated with HIV; results of therapy; and changes in clinical presentation.

TB-HIV Activities Data from VCTCs

Background: Data collection started when VCTCs were established for the purpose of counting those who are suspected of having or diagnosed with TB who are receiving HIV tests at the VCTCs.

Data elements collected: Tallies are made of those who are HIV infected who are suspected of having pulmonary or extrapulmonary TB, those who have been referred, and those who have been diagnosed by a clinician somewhere else. There are no clinicians diagnosing TB at the VCTCs. Information is also collected about the number of people receiving information and counseling on TB (both HIV+ and HIV - individuals) (see attached form).

Data flow: Data are tallied by the counselors at the VCTC sites. Tally sheets are sent to APSACS, where data are consolidated and then sent to NACO on a monthly basis. A copy is retained at APSACS in paper format.

Strengths and limitations: This source of information provides a count of those who are coinfecting with HIV and TB who attend the VCTCs. There is no demographic information recorded about these individuals. These tallies only describe those who are coinfecting with TB and HIV who attend the VCTCs and do not describe all individuals who are coinfecting. Data from these centers should not be generalized to the broader population to describe HIV/TB coinfection rates.

Suggested analyses: These data are useful for documenting those who are suspected as having TB, those who have been referred from TB facilities, and those who have been diagnosed with TB by a clinician who come to the VCTC for HIV counseling and testing. They also provide documentation about whether VCTC clients are receiving information and counseling on TB. There is not enough information available through this source to epidemiologically describe those who are coinfecting with HIV and TB. This information may be used in combination with other HIV -TB coinfection data, such as study data from medical colleges.

Voluntary Counseling and Testing Center (VCTC) Data

Background: Prior to 2001, there were three VCTCs in AP, one per region. Twenty-six were established in 2001, and this increased to 88 total sites in 23 districts in April, 2002. There are, on average, more than three VCTCs per district. As with other components of the CMIS system, full implementation of data collection took place in 4/2002.

VCTCs are located only in government hospitals. Clients can get tested at private hospitals, but then they have to pay. VCTCs serve referred patients (such as pregnant women and TB patients) as well as walk-ins.

Data elements collected: The pretest form includes sociodemographic information, history of risk behaviors, knowledge about HIV/AIDS, and disease-related symptoms (see attached form). On the aggregate form, information is collected about status of staff positions at the VCTC; status of availability of equipment and consumables; details of tests undertaken for HIV/AIDS (walk-ins vs. referrals); composition of people who test (sex x age); routes of transmission for seropositive cases (mode x sex); information exclusively about walk-ins (sex x age and sex x mode for all and seropositives); and counseling and other support services provided, by gender (see attached form).

Data flow: A pretest counseling form is filled out by the counselor. The counselor doesn't complete the form while talking to the client, but afterwards. Based on the information provided on the pretest form, the counselor fills out an aggregate report. Information for the aggregate report is also provided by lab technicians. The aggregate report is sent in paper form to APSACS on a monthly basis, data are entered into CMIS, and then are sent to NACO.

Strengths and limitations: Data collected at VCTCs are more likely to include a higher risk population. There is some ability to describe those who are HIV infected. Data from VCTCs don't describe those who are tested for HIV outside of government hospitals. Much of the information collected on the pretest counseling form, including the detailed behavioral information, is not available at the state level but retained at the VCTCs. Accuracy of recall when counselors complete the pretest form appears to be in question. There is no way to eliminate duplicates. Information about geography reflects where a person got tested, not where they reside. Definitions of the other/not specified categories of transmission are not clear.

Suggested analyses: As is already being done, to the extent possible, describe those who test positive for HIV in regards to gender, age, and mode of transmission. Compare demographic characteristics to those who do not test positive. It will also be useful to describe walk-ins as compared to those who are referred. Care should be taken with analyses that include small numbers,

particularly when looking at results by location and over time. In these instances, it would be useful to calculate confidence intervals.

CONCLUSIONS/SUGGESTIONS

The amount of work that has been completed in a relatively short time in regards to scaling up the number of sites at which people in the state and the country can receive HIV-related services, as well as collecting data from these sites, is very impressive. It is also a strength that NACO has put into place a uniform electronic system for reporting of the data. Through this system, there is a great deal of information about the number of people who receive specific services as well as the amount of resource that goes into providing these services. These data are very useful for looking at present and future needs for program delivery at various sites.

Information about the number of people who are infected and the characteristics and behaviors of these individuals is also very important for planning and targeting HIV/AIDS-related services. Review of the data sources included in this inventory indicate that while there are data available that may be used for these purposes, there are some gaps in regards to knowledge of the representativeness and quality of some of the data and in particular types of data. Following are some suggestions about how to move forward to create an epidemiologic profile with the available information while at the same time working to resolving some of the issues related to the gaps.

E. Use the data that are available to create an epidemiologic profile

- In the short term, create an epidemiologic profile for one district in AP with existing available data sources as described in the inventory. This district should be one that has a large number of available data sources, so would likely be a district that has higher prevalence and has been more studied. It will be important to provide narrative interpretation of results in addition to charts, maps and graphs. It would be useful to format the profile to answer specific epidemiologic questions; this would highlight where there are gaps in the data. There appears to have been much attention focused on trying to determine the number of HIV positive individuals who live in the state. This is important information, but it is also important to be able to describe the characteristics and behaviors of these individuals as much as possible. Consequently, it would be advantageous to place more emphasis on these types of descriptions. With the available data, this may be difficult to do because of the lack of individual level data. The profile should also describe strengths and limitations of data sources (information from this inventory could be included). As data sources are improved, accessed, or added over time, the profile should be updated. One possibility would be to create a “before” profile for a single district using data that are available now, access other data sources as described in the following steps, add these data to the profile to create an “after” profile, and then use this example from the single district to create enthusiasm about going through the same exercise in other districts.

- It would be useful to map not only where the risk populations are, but also where interventions are taking place and where clinics and hospitals are located.

Collect individual-level data from existing sources

- As mentioned a number of times in the data inventory, data are often received at the state level in aggregate format, and there is a very minimal amount of behavioral information available. It would be advantageous to work on methodologies to obtain and record individual-level client data from STD clinics, blood banks, VCTCs, PMTCTs and NGOs. Having these data available at the state level will allow for more complex analyses to describe populations that are infected or at risk. An important part of doing this type of data collection will be to have the ability to enter these data into an electronic system for analyses. The system may be one such as EpiInfo, in which case it would be important to provide training and capacity in using this system. As mentioned in the section above, it may be a useful exercise to pursue these individual level data in one district to work out methodologies and systems development and then expand to other districts.

Improve understanding and quality of the existing data

- Collect information about the various blood banks, STD clinics, NGOs, ANCs, VCTCs and PMTCTs in order to better understand how representative the clients who attend these sites are of the population that is being described and the characteristics of the geographic areas where the sites are located. As suggested previously, it may be best to pilot data collection in one district.
- Data quality, accuracy, completeness, and validity are unknown at a number of the sites. It would be useful to examine what happens from the point of data collection from the client to the time that the data are received at the state to see what procedures are being followed and evaluate these procedures. Aggregate data received at the state level should be examined for elements that are not consistent and follow-up should be done with the sites. Again, it may be best to pilot in one district first.
- It may be worthwhile to conduct a similar exercise at sentinel surveillance sites while these studies are being conducted.

Collect new information

- AIDS surveillance data, collected in a way that is population-based, would provide good information about needs for resources, both for prevention and care. The process of initiating a surveillance system would need to be inclusive, so thought would need to be put into who the stakeholders are

- (ideally both public and private). There should be discussion about the case definition, reporting requirements, a case report form (using the one that already exists or revising it), and an electronic system for the data. Again, it may be easier to start in one district and build a model/protocol for implementation.
- There also appears to be interest in collecting more behavioral information from those who are infected. The model for this could be the U.S. based CDC project called SHAS (Supplement to HIV/AIDS Surveillance). Sampling is done of those who are reported to the surveillance system and these people are recruited for participation in the study. A face-to-face interview is conducted and an extensive questionnaire is filled out about demographic and socioeconomic characteristics, drug use and sexual behaviors, access to health services, medication regimens, etc. Since more behavioral information is needed from high-risk populations, this may be a place to start. The SHAS survey tool could be adapted (it may also be worthwhile looking at the survey instrument from another CDC project called HITS, the HIV Testing Survey), protocols written and interviewers trained.
 - There appears to be a lot of data that are being collected about pregnant women, and quite a bit of overlap between the ANC data and the PMTCT data. In the future, when adding sites outside of the ones mandated by NACO, APSACS may want to consider scaling back on the ANC sites and adding more sites to capture high risk populations (more STD clinics or perhaps NGOs?).

This is an ambitious and broad scope of work. It provides information to describe the skills needed for the NASTAD consultant (see appendix), who will spend 3 months working in AP. It also appears to fit well with the capacity that is present with the staff at the Institute of Health Systems, so it would be ideal to have that institution involved in doing some of this work.

Appendix 1 – Job description for consultant

Short-term consultant for NASTAD technical assistance to India

At the request of the Project Director of the Andhra Pradesh State AIDS Control Society (APSACS) and the Institute of Health Systems, NASTAD and the CDC Global AIDS Program's India office will place a short-term consultant in Hyderabad, India for three months beginning July 2003 through late September 2003. This consultant will be responsible for coordination of public and private sector partners, as well as NASTAD technical assistance, in the development of an epidemiological profile pilot for one district in the state of Andhra Pradesh. In doing so, the consultant will assist in a) the development of methods to evaluate existing data sources in regards to quality, and b) the development of systems for collecting additional information that could be used to better describe those at high risk for or infected with HIV in Andhra Pradesh.

The consultant will:

- Have lived and/or worked in cross-cultural settings outside of the United States
- Possess strong interpersonal communication skills
- Have the ability to sort and analyze a broad variety of data sources, as well as develop and train others in relevant databases (e.g. Epi Info), provide guidance and support in data quality assurance, and lead statistical analyses while building local capacity
- Have written or contributed to the development of U.S. state level HIV/AIDS epi profiles
- Have experience in HIV/AIDS surveillance and the evaluation of surveillance systems
- Understand the context of HIV/AIDS outside of the United States
- BA required, Masters degree preferred

Appendix B

Data Collection Guidelines

**Visakhapatnam HIV/AIDS Epidemiologic Profile
AIDS CASE DATA COLLECTION GUIDELINES**

Data Type	Individual-level AIDS Case Sheets
Service Area(s)	King George Hospital CHEST Hospital
Time Period	April 2003 – December 2003
Sampling Frame	All AIDS cases
Contact(s)	
Notes/Comments	Dr. TL Prasad completed data abstraction. NACO forms are not available. Records were abstracted from individual patient medical records from KGH and the CHEST hospital.

INDICATORS TO COLLECT	VALUES/FORMAT	NOTES/COMMENTS
1. Date		
	Month/Year	e.g. July 2003
	'blank' for missing values	
2. Age		Continuous variable
	<1 year	
	# years	
	'blank' for missing values	
3. Sex		
	Male	
	Female	
	Missing	
4. Marital Status		
	Married	
	Unmarried	
	Missing	
5. STD/O.I.		
	VRDL (syphilis)	
	Kaposi Sarcoma	
	Cryptosporidiosis	
	Toxoplasmosis	
	Tuberculosis	
	Other	
6. Occupation		Need to categorize values from individual records
	Agriculture	
	Business	
	CSW	

	Employed – private	
	Employed – public/government	
	Employed – unspecific	
	Housewife	
	Professional	
	Service Sector – skilled	
	Service Sector – unskilled	
	Student	
	Transportation	
	Unemployed	
	Other	
	Missing	
7. District		Need to determine from place of residence
	Adilabad Anantapur Chittoor Cuddapah East Godavari Guntur Hyderabad Karimnagar Khammam Krishna Kurnool Mahaboobnagar Medak Nalgonda Nellore Nizamabad Prakasam Rangareddi Srikakulam Visakhapatnam Vizianagaram Warangal West Godavari Missing	
8. Vizag Mandal		Need to determine from place of residence
	Anakapalle Anandapuram Ananthagiri Araku Valley Atchutapuram Bheemunipatnam Butchayyapeta Cheedikada	

	Chinagadili Chinthapalli Chodavaram Deverapalli Dumbriguda G. Madugula G.K. Veedhi Gajuwaka Golugonda Hukumpeta K. Kotapadu Kasimkota Kotauratla Koyyuru Madugula Makavarapalem Munagapaka Munchingput Nakkapalee Narsipatnam Nathavaram Paderu Padmanabham Paravada Payakaraopeta Pedabayalu Pedagantyada Pendurthi Rambilli Ravikamatham Rolugunta S. Rayavaram Sabbavaram Visakhapatnam Yellamanchili Missing	
--	---	--

**Visakhapatnam HIV/AIDS Epidemiologic Profile
BLOOD BANK DATA COLLECTION GUIDELINES**

Data Type	Individual-level Blood Bank Donor Data
Service Area(s)	King George Hospital
Time Period	April 2003 – December 2003
Sampling Frame	All HIV positive, accepted donors
Contact(s)	Dr. Manjula Vani
Notes/Comments	Demographic data and HIV status stored in separate registers

INDICATORS TO COLLECT	VALUES/FORMAT	NOTES/COMMENTS
1. Date		
	Month/Year	e.g. July 2003
	'blank' for missing values	
2. Age		Continuous variable
	<1 year	
	# years	
	'blank' for missing values	
3. Sex		
	Male	
	Female	
	Missing	
4. Education		Need to categorize values
	Illiterate	
	1 st -4 th	
	5 th -8 th	
	9 th -10 th	
	Intermediate	
	Diploma/Degree	
	Post Graduate	
	Missing	
5. Donor Type		
	Replacement	
	Voluntary	
	Missing	
6. Occupation		Need to categorize values from individual records
	Agriculture	
	Business	
	CSW	
	Employed – private	
	Employed – public/government	

	Employed – unspecific	
	Housewife	
	Professional	
	Service Sector – skilled	
	Service Sector – unskilled	
	Student	
	Transportation	
	Unemployed	
	Other	
	Missing	
7. District		Need to determine from place of residence
	Adilabad Anantapur Chittoor Cuddapah East Godavari Guntur Hyderabad Karimnagar Khammam Krishna Kurnool Mahaboobnagar Medak Nalgonda Nellore Nizamabad Prakasam Rangareddi Srikakulam Visakhapatnam Vizianagaram Warangal West Godavari Missing	
8. Vizag Mandal		Need to determine from place of residence
	Anakapalle Anandapuram Ananthagiri Araku Valley Atchutapuram Bheemunipatnam Butchayyapeta Cheedikada Chinagadili Chinthapalli	

	Chodavaram Deverapalli Dumbriguda G. Madugula G.K. Veedhi Gajuwaka Golugonda Hukumpeta K. Kotapadu Kasimkota Kotauratla Koyyuru Madugula Makavarapalem Munagapaka Munchingput Nakkapalee Narsipatnam Nathavaram Paderu Padmanabham Paravada Payakaraopeta Pedabayalu Pedagantyada Pendurthi Rambilli Ravikamatham Rolugunta S. Rayavaram Sabbavaram Visakhapatnam Yellamanchili Missing	
--	--	--

**Visakhapatnam HIV/AIDS Epidemiologic Profile
GREEN VISION COLLECTION GUIDELINES**

Data Type	Individual-level Care and Treatment Data
Service Area(s)	Green Vision – NGO providing O.I. care and treatment
Time Period	April 2003 – December 2003
Sampling Frame	All HIV positive clients
Contact(s)	Dr. G. Prabhakar
Notes/Comments	Individual-level records include information on client’s symptoms and in some case, diagnosis. Treatment provided is primarily for O.I. symptoms. A letter of intent was provided to Green Vision prior to data abstraction.

INDICATORS TO COLLECT	VALUES/FORMAT	NOTES/COMMENTS
1. Date		
	Month/Year	e.g. July 2003
	‘blank’ for missing values	
2. Age		Continuous variable
	<1 year	
	Number of years	
	‘blank’ for missing values	
3. Sex		
	Male	
	Female	
	Missing	
4. Occupation (client)		Need to categorize values from individual records
	Agriculture	
	Business	
	CSW	
	Employed – private	
	Employed – public/government	
	Employed – unspecific	
	Housewife	
	Professional	
	Service Sector – skilled	
	Service Sector – unskilled	
	Student	
	Transportation	
	Unemployed	
	Other	
	Missing	
5. Spouse HIV Status		
	HIV positive	
	HIV negative	

	Unmarried	
	Missing	
6. Spouse Tested for HIV		
	Yes	
	No	
	Unmarried	
	Missing	
7. Present Health Problem		
	General Symptoms (e.g fever, body ache)	
	Kaposi Sarcoma	
	Cryptosporidiosis	
	Toxoplasmosis	
	Tuberculosis	
	Other	
	Missing	
8. Treatment Provided		
	Yes	
	No	
	Missing	
9. District		Need to determine from place of residence
	Adilabad Anantapur Chittoor Cuddapah East Godavari Guntur Hyderabad Karimnagar Khammam Krishna Kurnool Mahaboobnagar Medak Nalgonda Nellore Nizamabad Prakasam Rangareddi Srikakulam Visakhapatnam Vizianagaram Warangal West Godavari Missing	

10. Vizag Mandal		Need to determine from place of residence
	Anakapalle Anandapuram Ananthagiri Araku Valley Atchutapuram Bheemunipatnam Butchayyapeta Cheedikada Chinagadili Chinthapalli Chodavaram Deverapalli Dumbriguda G. Madugula G.K. Veedhi Gajuwaka Golugonda Hukumpeta K. Kotapadu Kasimkota Kotauratla Koyyuru Madugula Makavarapalem Munagapaka Munchingput Nakkapalee Narsipatnam Nathavaram Paderu Padmanabham Paravada Payakaraopeta Pedabayalu Pedagantyada Pendurthi Rambilli Ravikamatham Rolugunta S. Rayavaram Sabbavaram Visakhapatnam Yellamanchili Missing	

**Visakhapatnam HIV/AIDS Epidemiologic Profile
PPTCT DATA COLLECTION GUIDELINES**

Data Type	Individual-level PPTCT Data
Service Area(s)	King George Hospital, OB/GYN Department Anakapalle Area Hospital Victoria General Hospital
Time Period	April 2003 – December 2003
Sampling Frame	All HIV positive clients and a sample of HIV negative clients <i>Sample of HIV negative clients:</i> For each month, calculate 10% of the total number of individuals tested for HIV. For example, if 100 clients were tested for HIV in July 2003, then you would need to abstract records from 10 HIV negative clients in July. The <i>first</i> ten HIV negative records should be abstracted beginning from the first of the month. Therefore, in July, records from the first 10 clients that tested HIV negative would be abstracted beginning from July 1 st . The sampling is not random, but systematic. <i>Note:</i> It was helpful to review the monthly aggregate reports to determine the number of HIV negative records needed.
Contact(s)	<i>KGH</i> – OB/GYN Head of Department, Mrs. Sailaja (counselor), and Mr. Venkatraman (counselor) <i>Anakapalle</i> - Dr Arvind (Hospital Superintendent), Dr. Joythi, Mrs. Sailaja (counselor) <i>VGH</i> – Dr. J. Gowri and Dr. Vani, Mrs. Sailaja and Mr. Umamahesh
Notes/Comments	Demographic data and HIV status stored in separate registers. DLO contacted the hospital superintendent at Anakapalle to notify him about the project and the requested data to have ready. Faculty at the Community Medicine Department introduced us to the chief medical officer and counselors at the KGH OB/GYN department and VGH.

INDICATORS TO COLLECT	VALUES/FORMAT	NOTES/COMMENTS
1. Date		
	Month/Year	e.g. July 2003
	'blank' for missing values	
2. Referral Information		Abstract only the coded response (e.g. '4')
	1. Self	
	2. Doctor	
	3. NGO/CBO	
	4. Spouse	
	5. Other family members	
3. Age		Continuous variable
	<1 year	
	Number of years	
	'blank' for missing values	

4. Sex		Abstract only the coded response
	1. Male	
	2. Female	
	5. Missing	
5. Education		Abstract only the coded response
	1. Illiterate	
	2. Up to 4 th Std.	
	3. Up to 8th Std.	
	4. SSC	
	5. Degree	
	6. Post Graduate degree	
	10. Missing	
6. Occupation (client)		Need to categorize values from individual records
	Agriculture	
	Business	
	CSW	
	Employed – private	
	Employed – public/government	
	Employed – unspecified	
	Housewife	
	Professional	
	Service Sector – skilled	
	Service Sector – unskilled	
	Student	
	Transportation	
	Unemployed	
	Other	
	Missing	
6b. Occupation (spouse)		Available for some records at KGH and VGH
	Same values as #6	
7. Marital Status		Abstract only the coded response
	1. Married	
	2. Unmarried	
	3. Divorced	
	4. Widower	
	10. Missing	
8. Reason for Visit		Abstract only the coded response
	1. Spouse death	

	2. H/o STD	
	3. H/o risk behavior	
	4. Prolonged illness	
	5. Spouse is infected	
	6. Baby of infected mother	
	7. AIDS-phobia	
	8. Any other (specify)	Specify in 8a.
	10. Missing	
9. Risk Behavior		Abstract only the coded response
	1. Heterosexual – MPS	
	2. MSM	
	3. IDU – Needle Sharing	
	4. H/o Blood Transfusion	
	5. Baby of HIV infected mother	
	6. Needle stick injury	
	7. Other	
	8. Unknown	
	10. Missing	
10. HIV Status		Abstract only the coded response
	1. HIV positive	
	2. HIV negative	
	3. Unknown	
	10. Missing	
11. District		Need to determine from place of residence
	Adilabad Anantapur Chittoor Cuddapah East Godavari Guntur Hyderabad Karimnagar Khammam Krishna Kurnool Mahaboobnagar Medak Nalgonda Nellore Nizamabad Prakasam Rangareddi Srikakulam	

	Visakhapatnam Vizianagaram Warangal West Godavari Missing	
12. Vizag Mandal		Need to determine from place of residence
	Anakapalle Anandapuram Ananthagiri Araku Valley Atchutapuram Bheemunipatnam Butchayyapeta Cheedikada Chinagadili Chinthapalli Chodavaram Deverapalli Dumbriguda G. Madugula G.K. Veedhi Gajuwaka Golugonda Hukumpeta K. Kotapadu Kasimkota Kotauratla Koyyuru Madugula Makavarapalem Munagapaka Munchingput Nakkapalee Narsipatnam Nathavaram Paderu Padmanabham Paravada Payakaraopeta Pedabayalu Pedagantyada Pendurthi Rambilli Ravikamatham Rolugunta S. Rayavaram Sabbavaram Visakhapatnam	

	Yellamanchili Missing	
Additional Variables for HIV positive women		Variables 14-16 will only be available for a limited number of HIV+ women.
13. Spouse's HIV Status		
	1. HIV positive	
	2. HIV negative	
	3. Unknown	
	4. Missing	
14. Pregnancy Outcome		
	1. Medical Termination of Pregnancy (MTP)	
	2. Miscarriage	
	3. HIV positive baby	
	4. HIV negative baby	
	5. Loss to Follow-Up	
	6. Unknown	
	7. Missing	
15. Nevaripine Use		
	1. Administered	
	2. Not administered	
	3. Unknown	
	4. Missing	

**Visakhapatnam HIV/AIDS Epidemiologic Profile
STD DATA COLLECTION GUIDELINES**

Data Type	Individual-level STD Clinic Data
Service Area(s)	Andhra Medical College STD Department
Time Period	April 2003 – December 2003
Sampling Frame	A sample of male STD clients and a sample of all female STD clinic clients. Abstract data from every third record. Final abstracted records should include 1/3 male clients and 1/3 of female clients. <i>Sample of STD clients</i> – data abstraction was conducted separately for male (yellow cards) and female (pink cards) records STD records were stored in bundled stacks by month and by client’s sex. Male records are pink and female records are yellow.
Contact(s)	Dr. Ashok Babu (HOD), Dr. Bose, Mrs. Sailaja (medical social worker)
Notes/Comments	Self-reported HIV status is not uniformly collected on the STD case sheet, but is available for many records. ‘Reason for visit’ and ‘risk behavior’ are also variables that are not part of the standard case sheet, but are often available.

INDICATORS TO COLLECT	VALUES/FORMAT	NOTES/COMMENTS
1. Date	Month/Year ‘blank’ for missing values	e.g. July 2003
2. Age	<1 year Number of years ‘blank’ for missing values	Continuous variable
3. Sex	Male Female	
4. Marital Status	Married Unmarried Missing	
5. Education	Illiterate 1 st - 4 th 5 th - 8 th 9 th - 10 th Diploma/Degree Intermediate Post Graduate degree Other	

6. Occupation	<p>Missing</p> <p>Agriculture Business CSW Employed – private Employed – public/government Employed – unspecific Housewife Professional Service Sector – skilled Service Sector – unskilled Student Transportation Unemployed Other Missing</p>	<p>Need to categorize values from individual records</p>
7. Reason for Visit	<p>Symptoms HIV Screening Spouse HIV positive Spouse died (cause not specified) Spouse died due to HIV Referred Other Missing</p>	<p>Complete this variable when data are available</p>
8. Risk Information	<p>Extra Marital Exposure Spouse Infected Unprotected Sex MSM CSW Needle Stick Other Missing</p>	<p>Complete this variable when data are available</p>
9. HIV Status	<p>HIV positive HIV negative HIV Status Unknown</p>	<p>Complete this variable when data are available</p>
10. STD Diagnosis (laboratory confirmed)	<p>Balanoposthitis (BP)</p>	<p>Mark all that apply (e.g. if client has more than one STD)</p>

Candidiasis
Cervicitis
Chlamydia
Donovanosis
Gonorrhea
Herpes
Lymphogranularum veneris
Molluscum contagiosum
Non-gonococcal Urethritis
Scabies
Syphilis
Tinea cruris
Trichomonal vaginitis
Ulcer
Warts (includes HPV and
Condylomata accuminata)

11. District

Need to determine from
place of residence

Adilabad
Anantapur
Chittoor
Cuddapah
East Godavari
Guntur
Hyderabad
Karimnagar
Khammam
Krishna
Kurnool
Mahaboobnagar
Medak
Nalgonda
Nellore
Nizamabad
Prakasam
Rangareddi
Srikakulam
Visakhapatnam
Vizianagaram
Warangal
West Godavari
Missing

12. Vizag Mandal

Need to determine from
place of residence

Anakapalle
Anandapuram
Ananthagiri
Araku Valley

Atchutapuram
Bheemunipatnam
Butchayyapeta
Cheedikada
Chinagadili
Chinthapalli
Chodavaram
Deverapalli
Dumbriguda
G. Madugula
G.K. Veedhi
Gajuwaka
Golugonda
Hukumpeta
K. Kotapadu
Kasimkota
Kotauratla
Koyyuru
Madugula
Makavarapalem
Munagapaka
Munchingput
Nakkapalee
Narsipatnam
Nathavaram
Paderu
Padmanabham
Paravada
Payakaraopeta
Pedabayalu
Pedagantyada
Pendurthi
Rambilli
Ravikamatham
Rolugunta
S. Rayavaram
Sabbavaram
Visakhapatnam
Yellamanchili
Missing

**Visakhapatnam HIV/AIDS Epidemiologic Profile
VCTC DATA COLLECTION GUIDELINES**

Data Type	Individual-level VCTC Data
Service Area(s)	Andhra Medical College, Microbiology Anakapalle Area Hospital Narsipatnam Area Hospital
Time Period	April 2003 – December 2003
Sampling Frame	All HIV positive clients and a sample of HIV negative clients <i>Sample of HIV negative clients:</i> For each month, calculate 10% of the total number of individuals tested for HIV. For example, if 100 clients were tested for HIV in July 2003, then you would need to abstract records from 10 HIV negative clients in July. The <i>first</i> ten HIV negative records should be abstracted beginning from the first of the month. Therefore, in July, records from the first 10 clients that tested HIV negative would be abstracted beginning from July 1 st . The sampling is not random, but systematic. <i>Note:</i> It was helpful to review the monthly aggregate reports to determine the number of HIV negative records needed.
Contact(s)	AMC – Dr. Ramani (HOD) & Dr. Joshi, Ms. Mary (counselor) Anakapalle - Dr Arvind (Hospital Superintendent), Dr. Jogan Mohan, Mr. Naidu (counselor) Narsipatnam - attn. Dr. R. Susheela (Hospital Superintendent), Dr. Gowri Devi, Ms. Sailaja (counselor)
Notes/Comments	Demographic data and HIV status stored in separate registers. DLO contacted hospital superintendents at Anakapalle and Narsipatnam to notify them about the project and the requested data to have ready

INDICATORS TO COLLECT	VALUES/FORMAT	NOTES/COMMENTS
1. Date	Month/Year 'blank' for missing values	e.g. July 2003
2. Referral Information	1. Self 2. Doctor 3. NGO/CBO 4. Spouse 5. Other family members 10. Missing	Abstract only the coded response (e.g. '4')
3. Age	<1 year Number of years 'blank' for missing values	Continuous variable

4. Sex

Abstract only the coded response

1. Male
2. Female
5. Missing

5. Education

Abstract only the coded response

1. Illiterate
2. Up to 4th Std.
3. Up to 8th Std.
4. SSC
5. Degree
6. Post Graduate degree
10. Missing

6. Occupation

Need to categorize values from individual records

- Agriculture
- Business
- CSW
- Employed – private
- Employed – public/government
- Employed – unspecified
- Housewife
- Professional
- Service Sector – skilled
- Service Sector – unskilled
- Student
- Transportation
- Unemployed
- Other
- Missing

7. Marital Status

Abstract only the coded response

1. Married
2. Unmarried
3. Divorced
4. Widower
10. Missing

8. Reason for Visit

Abstract only the coded response.

1. Spouse death
2. H/o STD
3. H/o risk behavior
4. Prolonged illness
5. Spouse is infected
6. Baby of infected mother
7. AIDS-phobia

	8. Any other (specify) 10. Missing	Specify in 8a.
9. Risk Behavior	1. Heterosexual – MPS 2. MSM 3. IDU – Needle Sharing 4. H/o Blood Transfusion 5. Baby of HIV infected mother 6. Needle stick injury 7. Other 8. Unknown 10. Missing	Abstract only the coded response
10. HIV Status	1. HIV positive 2. HIV negative 10. Missing	Abstract only the coded response
11. District	Adilabad Anantapur Chittoor Cuddapah East Godavari Guntur Hyderabad Karimnagar Khammam Krishna Kurnool Mahaboobnagar Medak Nalgonda Nellore Nizamabad Prakasam Rangareddi Srikakulam Visakhapatnam Vizianagaram Warangal West Godavari Missing	Need to determine from place of residence
12. Vizag Mandal	Anakapalle	Need to determine from place of residence

Anandapuram
Ananthagiri
Araku Valley
Atchutapuram
Bheemunipatnam
Butchayyapeta
Cheedikada
Chinagadili
Chinthapalli
Chodavaram
Deverapalli
Dumbriguda
G. Madugula
G.K. Veedhi
Gajuwaka
Golugonda
Hukumpeta
K. Kotapadu
Kasimkota
Kotauratla
Koyyuru
Madugula
Makavarapalem
Munagapaka
Munchingput
Nakkapalee
Narsipatnam
Nathavaram
Paderu
Padmanabham
Paravada
Payakaraopeta
Pedabayalu
Pedagantyada
Pendurthi
Rambilli
Ravikamatham
Rolugunta
S. Rayavaram
Sabbavaram
Visakhapatnam
Yellamanchili
Missing

