INTEGRATED DISEASE SURVEILLANCE PROJECT

TRAINING MANUAL FOR STATE & DISTRICT SURVEILLANCE OFFICERS

ANALYSIS AND INTERPRETATION OF DATA

Module -9

CONTENTS

1. INTRODUCTION TO THE UNIT

While collection of good quality data is important for a surveillance Programme, analysis and interpretation of this data is of equal significance. Without this, all the hard work put in by the workers becomes meaningless. Data Analysis serves six important purposes.

- ☞ Analysis of routine data helps in identifying outbreaks or potential outbreaks e.g. a case of measles identified should alert the health services about a potential outbreak.
- ☞ Analysis also helps to identify high-risk population groups, so that targeted interventions can be provided to them and scarce resources utilized appropriately.
- ☞ During an outbreak, analysis of the data identifies the most appropriate and timely control measures. Analysis in terms of person, time and place will be able to focus the intervention; e.g. analysis of a diarrhoeal outbreak will be able to identify the affected families and the cause of the outbreak so that corrective action can be targeted at this cause.
- ☞ Analysis of routine data provides information for predicting changes of disease rates over time and enables appropriate action. E.g., the increasing trends in Road Traffic injuries should help the public health manager to inform other sectors, redirect resources and plan interventions to reduce the same
- ☞ Analysis enables identifying problems in the health system, so that gaps can be effectively plugged; e.g., an outbreak of measles should alert the public health manager about the possibility of low vaccination coverage in that region.
- ☞ Comparison of analyzed data between regions or between sectors (public and private) helps the public health manager in identifying regional differences, improving the quality of the surveillance system and greater cooperation of all partners.

Please remember that-

- ☞ Regular and timely analysis is as important as regular and timely collection of data. The very purpose of collection of data is defeated if data are not analyzed properly and timely and used effectively.
- ☞ The effectiveness of the follow-up action depends on the promptness with which data is analyzed, especially in the event of an outbreak.
- ☞ The data should be checked for regularity and completeness.

The analysis of the data should include assessment of the quality of information.

A simple way of doing this is to maintain a monthly record of data received from the reporting sites, as per the following illustration month wise for the whole year.

2. SPECIFIC LEARNING OBJECTIVES

On completion of this module you will be able to

1. Identify the roley, importance and brief techniques of data analysis and value of interpreting the same in a scientific manner.
2. Outline the various sources of data in the district for disease surveillance and management of this data to draw valid conclusions.
3. Describe the reasons for separate analysis and reporting of data from various sources (Syndromic/Presumptive/Confirmed diagnosis; Sentinel/Regular surveillance) for making valid conclusion.
4. To choose descriptive and appropriate simple analytical methods to describe, present and interpret the outcome measures in reports and feed back
5. List outcome measures that will be used for developing periodic reports and feed back from DSO office (Weekly, Monthly, yearly)
6. Make valid conclusions from a given set of Data provided in
   - ❖ Raw Data from the District
   - ❖ Tables
   - ❖ Graphs
   - ❖ Feed back and reports generated

3. MODULE STRUCTURE AT A GLANCE

Duration of Session **3 Hours**

Unit No	CONTENT	METHODOLOGY	TENTATIVE DURATION	TEACHING AIDS
1.	Analysis approach and methods along with interpretation	Lecture followed by Module reading and Discussion	30 Minutes 60 Minutes	Training modules/ Slide Projector/ overhead projector
2.	Analysis Techniques	Group Activity	30 Minutes 60 Minutes	Group discussion and interpretation Exercises

Instructions for Trainees

- ☛ These modules have been designed as a self-contained unit of learning. You are expected to learn in small groups of 6-8 persons under the guidance of a facilitator after initial presentation
- ☛ You are expected to share your experiences with other trainees.
- ☛ The module has been divided into various units to cover various aspects of the topic.
- ☛ The unit structure given at the beginning of each unit will give you an overall view of the unit and the sequence in which you should go through.

- ☛ The learning objectives of the unit will inform you what is expected from you once you complete the unit.
- ☛ Please go through each learning activity and complete the exercise(s) given at the end of the unit.

4. KEY POINTS TO REMEMBER

- ☞ Surveillance data is useful for identifying outbreaks, tracking epidemics and measuring changes and describing various disease conditions by place, time and person.
- ☞ Surveillance helps in monitoring disease control activities, assessing the impact of and quality of services and timely intervention in high-risk pockets of population for special action.
- ☞ The various sites of data collection are sub-centre, primary health center, community health center, private practitioners and hospitals, medical colleges, taluka hospitals and district hospitals. Data collation should be done at district and state level.
- ☞ Data analysis should ideally be done at each level
- ☞ The standard definitions of case and suspect should be followed for surveillance of all diseases.
- ☞ The quality of surveillance data depends upon regularity of reports, timeliness, completeness and correct diagnostic criteria.
- ☞ A report (from a reporting unit) is said to be on time, if it reaches the designated level within the prescribed time period. If it reaches later, then the report is considered to be late (and of lesser public health use). The timeliness of a reporting unit can be calculated by assessing how many of its expected reports have come on time.
- ☞ The periodicity of data collection varies from weeks to months depending upon the disease.
- ☞ The analyzed data should be presented in terms of simple tables, graphs and charts and comparisons made with previous data
- ☞ Line listing of cases, spot maps, high risk mapping can be done and is very important.

5. GROUP ACTIVITIES

Group Activities and Exercises

Exercise 1

The weekly report of the number of diarrhoeal diseases from one district is provided in the table. Please comment on the data and specify what action will be taken in this problem.

Month wise and centre wise diarrhoeal diseases reports in the district

Reporting Site / months	A	M	J	J	A	S	O	N	D	J	F	M
Hospital 1…….	**30**	**25**	**22**		**17**	**15**	**11**	**10**	**10**		**9**	**9**
Hospital 2……	**42**	**37**	**30**	**17**	**17**	**18**	**14**	**11**	**9**	**10**	**8**	**7**
CHC 1	**15**	**13**	**14**	**12**	**15**		**13**	**12**	**10**	**9**	**11**	
CHC 2……	**34**	**23**	**28**	**18**	**17**	**16**	**18**	**17**	**15**	**12**	**11**	**13**
PHC 1……	**10**		**9**	**6**	**7**	**5**		**8**	**7**	**7**	**6**	**9**
Total	**131**	**106**	**95**	**20**	**73**	**68**	**65**	**58**	**51**	**47**	**45**	**50**

Exercise 2

Malaria cases from the District A (hypothetical) is provided below:

Reporting Site	Jan	Feb	Mar	*Cum 1999*	Jan	Feb	Mar	*Cum 2000*
Hospital 1…….	12	14	10	36	13	16	12	41
Hospital 2……	5	8	2	15	4	6	3	13
CHC 1	0	1	0	1	0	0	2	2
CHC 2……	1	0		1	0	10	0	10
CHC 3……	0	5		5		0	0	0
PHC 1……..			0	0	0		1	1
PHC 2……..			15	15	0	0	0	0
PHC 3……								
Total	18	28	27	73	17	33	18	68

- ☞ What is your comment? What action will be proposed to handle the problem?
- ☞ What do you think are the possible reasons for increase in the reported number of case?
- ☞ What do you think are the possible reasons for decrease in the reported number of cases?

Exercise 3. Calculating disease incidence and trend analysis

Calculate the incidence of the disease and interpret. Comment on the disease trend of this Vector Borne Disease in these four villages.

Table of reported cases during 1995 - 2002

		Population at risk	1995	1996	1997	1998	1999	2000	2001	2002
1	Village A	1750	61	10	6	2	0	0	41	84
2	Village B	2300	0	0	53	23	0	0	0	0
3	Village C	1800	62	4	1	0	0	0	13	0
4	Village D	2100	51	26	15	22	10	30	5	14
	TOTAL	7950	174	40	75	47	10	30	59	98

Exercise 4. Understanding cyclic pattern of the disease

The following table represents the number of cases of diarrhoeal diseases being reported from one large district.

Comment on the time trends of the disease over the years

Year	Number of cases
1974	9,714
1975	3,561
1976	7,047
1977	3,654
1978	1,313
1979	3,338
1980	36
1981	2,054
1982	6,196
1983	9,639
1984	881
1985	1,216
1986	5,076
1987	2,940
1988	1,379
1989	5,771

Exercise 5: Exercise on understanding age and sex distribution

The following table gives the age sex distribution of number of cases of STDs reported from 23 hospitals.

Comment on the data provided in the table

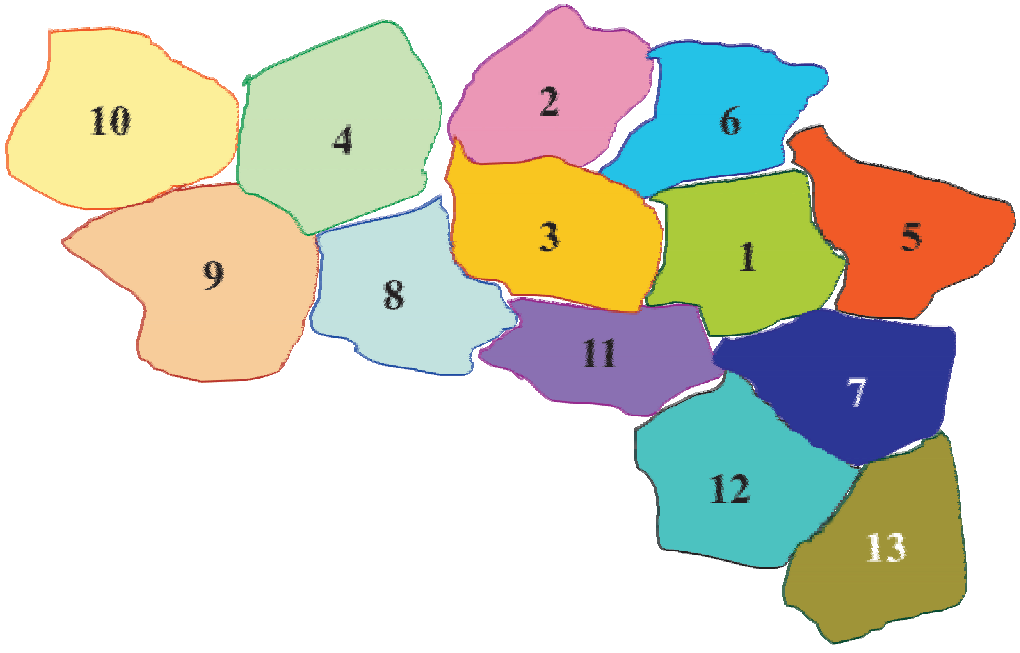
Age group	Male	Female	Total
0 - 14 yrs	1326	837	2163
15 - 29 yrs.	13073	7023	20096
30 - 44 yrs.	24887	6545	31432
> 45 yrs.	3276	814	4090
Total	42562	15219	57781

Exercise 6: Interpreting data from a map on place distribution.

The table given below depicts the occurrence of JE cases in 13 talukas of a district. What changes do you observe a period of time and interpret the same? Use the map given below to plot the taluka wise JE cases for any one year and interpret regarding the place distribution of JE

Table of reported cases select talukas during 1997 to 2002

Sl.No.	States/UTs	1997	1998	1999	2000	2001	2002
1	Taluka 1	247	192	203	72	4	3
2	Taluka 2	26	6	2	69	200	150
3	Taluka 3	0	0	0	19	11	1
4	Taluka 4	8	0	2	3	2	0
5	Taluka 5	0	16	56	43	22	40
6	Taluka 6	19	24	4	2	5	0
7	Taluka 7	87	50	98	45	14	15
8	Taluka 8	3	0	0	0	1	16
9	Taluka 9	0	1	2	0	0	1
10	Taluka 10	0	0	6	0	0	2
11	Taluka 11	42	14	5	0	0	0
12	Taluka 12	76	195	275	253	199	133
13	Taluka 13	124	9	27	50	21	0



Exercise-7 As part of your strategies to reduce Measles you have ensured immunization coverage of all eligible children.

According to the surveillance data, 20 cases of fever with rash were admitted in the district hospital last year. The surveillance data for this year shows no change in the hospital admissions.

1. List possible reasons why the surveillance data does not reflect any impact of the measures taken by you to reduce cases of Measles?

2. How would you check the validity of the reasons listed above?

3. What corrective measures would you take?

Exercise 8

How would you assess the quality of the surveillance system in your area from the given Data Set?

List the types of analysis that will provide information on quality

Tabulate the data to demonstrate validity of your observations

6. FREQUENTLY ASKED QUESTIONS

- ☞ Analysis cannot be done without computers.
- ☞ Analysis requires highly qualified scientists and we cannot do.
- ☞ Generating and interpreting tables is a complex task.
- ☞ Due to many constraints, we cannot act in a timely manner.
- ☞ We are not aware of statistical methods and cannot make analysis.
- ☞ We do not have data for previous years and cannot compare

7. HANDOUT ON 'ANALYSIS AND INTERPRETATION OF DATA'

7.1 As District Surveillance officer your major tasks are to collect, analyze, interpret, disseminate and initiate action on diseases identified under IDSP. In this direction you will receive data from the following sources at district level.

- ☛ Sub-Centre
- ☛ PHC
- ☛ CHC
- ☛ District
- ☞ Private practitioners
- ☞ Private nursing homes
- ☞ Identified laboratories
- ☞ Medical colleges (in some instances)

- ☞ Police departments
- ☞ State

The following types of Data will be available for analysis:-

- ☛ Syndromic case data
- ☛ Presumptive case data
- ☛ Confirmed case data
- ☛ Sentinel site data
- ☛ Regular surveillance data
- ☛ Urban data
- ☛ Rural data

The type and extent of data available will be based on reporting formats from different levels at periodic intervals. The reporting formats have already been discussed in earlier modules.

7.2 Periodicity of Data Collection and Reporting:

- ➢ Weekly
- ➢ In case of high priority, such as AFP Surveillance, as soon as the case is detected, and weekly report.

For example in Poliomyelitis, information about all cases of Acute Flaccid Paralysis (AFP) is collected. By collecting data on AFP, it is ensured that not a single case of Poliomyelitis is missed. The data should be collected in following age groups:

- ➢ Within one week of birth
- ➢ Between one week to 1 month of age
- ➢ Between 1 month to one year of age
- ➢ Between 1 to five years of age

7.3 What Information will be available for Analysis

Soon after surveillance Programme is in place, data will be / should be available on all conditions outlined in module 2 viz., Cholera, Other Diarrhoeal diseases, Typhoid, Malaria, Tuberculosis, Dengue fever, hepatitis, Japanese encephalitis, Vaccine Preventable Diseases, and some other diseases specified in individual states. The periodicity of data collection and availability will vary depending on situation and condition. The information will reach the DSO in regular formats as outlined in module 3 or sometimes by telephone, personal message in case of emergencies.

- ☞ Data on outbreaks are collected and analyzed separately to find out factors responsible for outbreaks and to plan control measures.

- ☞ For diseases, which are planned to be eradicated or eliminated, detailed information on each reported case is essential. Such diseases are poliomyelitis, neonatal tetanus, and measles. Information on all VPDs should continue to flow regularly maintaining the priority of Programmes for eradication and elimination.

7.4 Analysis

Surveillance is data collection for action. By collecting and analyzing data on cases and deaths due to selected diseases it is possible to pin point the Programmematic gaps for corrective actions or interventions.

- ☞ At both district and MO level, technical teams comprising data manager and data entry operators or other designated persons need to be constituted and oriented for their activities. At MO level, 1 or 2 interested (usually the person dealing with reports and statistics) persons may be involved in this activity.

7.4.1 Analysis - at which Level?

The surveillance system is most effective if the data are analyzed and used at the level at which they are collected.

Data analysis should ideally be done at each level

While, analysis should ideally be done at all level from the periphery upwards, the **officials responsible for analysis are the District surveillance officer (for rural areas) and the Corporation Health officer (for urban areas) and Medical officer at CHC and PHC levels.** The degree of analysis would depend on the capacity of the persons involved. For example, the community informants would be alert for any unusual increase in the number of fever cases occurring in the community. He should then be able to inform the MPW with details of how many people, what are the symptoms, where are they located etc. The PHC MO will limit his analysis to detection of outbreaks and anticipating seasonal trends. While the District Surveillance officer would be doing all of the above including recognition of outbreaks, informing all concerned to initiate action, responding to these outbreaks, monitoring trends, compilation of total information etc., as well as monitoring the effectiveness and efficiency of the health service. It is important for the Medical College staff to get involved in depth analysis of surveillance data.

7.4.2 When should analysis be done?

Analysis is done at various frequencies – daily, weekly, monthly, annually (see Table). Reports should be generated, either manually or computerized according to this frequency.

Frequency of reports and analysis

No	Reports	Daily[1]	Weekly	Monthly	Yearly
1	Timeliness and completeness of reports		✓	✓	✓
2	Description by time, place and person	✓	✓	✓	✓
3	Trends over time	✓	✓	✓	✓
4	Checking for crossing of threshold levels		✓	✓	
5	Comparison between reporting units			✓	
6	Comparison between public and private			✓	
7	Comparison between disease and lab data			✓	

7.4.3 How to analyse data?

Computer software Programmes have been designed to provide ready outputs at district level and Community health centre levels. At lower levels, manual handling of data would continue. As can be seen from table above, most of the analysis has to be done on a weekly basis. This will not be a major burden as number of reporting cases might be small except in epidemic or outbreak situations. In the event of an outbreak, the analysis has to be done on a daily basis. This will be dealt in the chapter on the investigation of and response to an outbreak. In this chapter, only the analysis of routine data has been dealt with. When analysing the data, the technical committee needs to keep some key points in mind.

- ☞ The strength and weaknesses of the data collection method
- ☞ The nature of reporting process.
- ☞ Examine each disease separately.
- ☞ Finally think whether the data generated is reliable and valid?

A systematic approach to analysis will help the public health manager in getting a clear picture of the situation. The steps given below are the same whether the analysis is done on a weekly basis or on a monthly or annual basis.

7.4.4 Steps in Analysis

1. The technical committee should meet on a fixed day of every week / month to review all aspects of data collection – analysis – interpretation and feedback. This might change in outbreak situations depending on the need.
2. A minimum of 4 reports on a weekly basis and 7 on a monthly basis has to be developed as specified under IDSP.
3. Review the reports disease wise and look for any missing values.
4. Check the validity of the data.

5. Prepare summary reports, which are to be shared with colleagues at the same level as well as with the concerned officers at the higher level. This summary report, especially the monthly report should also be used as a tool for feedback.

6. Interpret the results keeping ground realities and time to act.

7. Recommend or Take action wherever necessary

The type, extent and depth of analysis vary at each level. It could be simple compilation of numbers at health workers level. Developing consolidated statements will be made at the PHC level. More work should be done at Community health centre and district levels.

7.5 Data presentation methods

Information can always be presented in different ways. Some basic steps are:

- ☞ Start with presentation of simple and crude numbers in the form of line listing or simple tables and then proceed to summarized data
- ☞ Tables are necessary, but graphs are easier to see and understand.
- ☞ When comparing between institutions / areas, use rates and ratios (like incidence rate / case fatality ratio) rather than actual numbers. This takes care of the effect of different populations in different regions. For example, if Block A has 50 cases of malaria and Block B has 75, it does not naturally imply that the situation in Block B is worse. If Block B has a larger population, then it could account for the higher caseload.

Some examples of types of reports that can be generated are as follows.

Report 1 – Completeness and Timeliness of data

This is one of the first reports that have to be generated. It is a reflection on the performance of the reporting units. For this, one needs to have a list of all cases reported against each reporting unit. The MO can then monitor which of the reporting units are sending complete reports on time. A simple tool to monitor the completeness and timeliness of the reporting units is provided.

A report (from a reporting unit) is said to be on time, if it reaches the designated level within the prescribed time period. If it reaches later, then the report is considered to be late (and of lesser public health use). The timeliness of a reporting unit can be calculated by assessing how many of its expected reports have come on time.

A report is said to be complete if all the reporting units within its catchment area has submitted the reports on time. If 8 out of 10 only have been submitted, then the report is said to be incomplete (or 80% complete)

Timeliness and completeness of reporting units is a proxy indicator of the alertness of the surveillance system. An alert system will have timeliness and completeness approaching 100%. Also, completeness of reporting units gives an idea about the reliability of the data; for example, if completeness of reports were only 50%, then the incidence of disease would be under reported by 50%. So the incidence rates and CFRs need to be read in conjunction with the Completeness reports.

Interpretation of the report

Scenario	Interpretation
Reporting unit A is timely and complete	An ideal scenario, everything is working well
Reporting unit B is timely, but regularly incomplete	The MO of B has understood the importance of reporting on time. But there are some reporting units under the jurisdiction of B who are not reporting on time. B's MO has to find out what the problem is.
Reporting unit C is late, but reports are complete	The MO of C has not understood the importance of reporting on time. He is probably waiting for all the reporting units under his jurisdiction to report before submitting his report. He needs to be impressed about the significance of timely reporting.
Reporting unit D is late and the reports are incomplete.	Major problem in this reporting unit – neither the MO of D nor the MOs of the reporting units under D have understood the importance of surveillance and timely data.

7.6 When is the disease occuring – analysis by time

Report 2 – Weekly / Monthly summary report

This is the second set of reports that need to be generated and consists of a subset of reports in the form of tables, graphs and maps. It is based on the compiled data of all the reporting units. Some samples are shown in here and reporting formats are discussed in module 3.

This information can be presented through various tables starting from the primary one giving the number of cases and deaths to tables with summarized data and rates etc., OR in the form of graphs like, a simple bar graph to identify the incidence of diseases and deaths; pie graphs to show the load of diseases OR even maps informing the place of disease occurrence. As can be seen, the tables are cumbersome to read and interpret. However, it is necessary for the sake of records. Also a cursory scanning should identify missing values, transposition of rows or columns and biases. In the event of computerization, once the data is entered, various tables to suit the need of the individual surveillance officer can be obtained. Preferably data should be presented in a graphical manner so that the MO can review the data easily.

When looking at the data of a single region / reporting unit, primary measures like cases and deaths would suffice, incidence rates and case fatality ratios are necessary for comparing data between reporting units and region. Some of the concepts with

regard to measuring the burden of diseases (like incidence, prevalence, case fatality rates; epidemic – endemic etc., trends of diseases) and representation of diseases in simple ways by various graphs and tables are discussed in module 5.

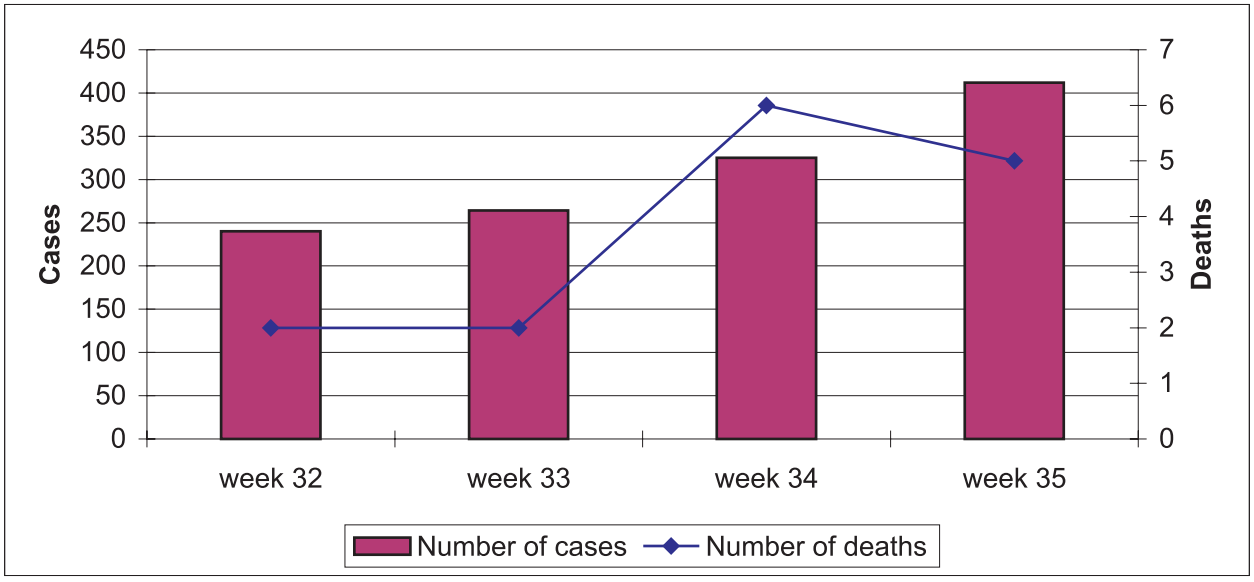
For the sake of clarity initially, separate tables and graphs should be made for data from the public and private sector.

This preliminary analysis should give the MO an idea of the health problem under his/her jurisdiction in terms of basic epidemiological parameters (time, place and person). It thereby helps the MO to focus on problems that need further analysis.

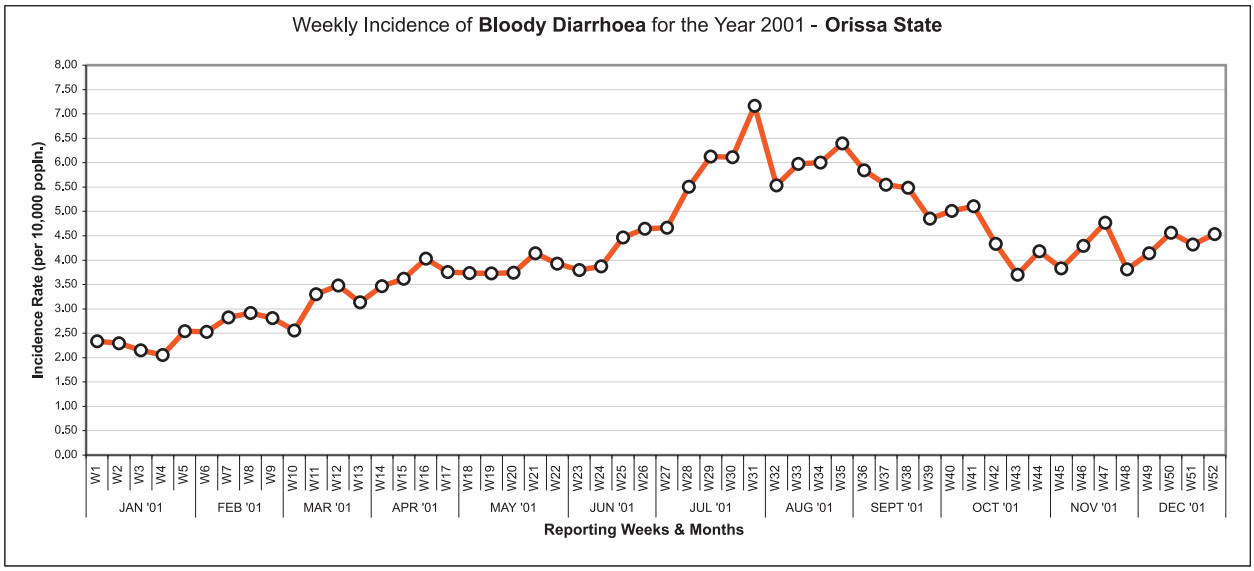
Report 3 - Comparison with previous weeks / months / years

This report helps the DSO and /or MO to detect the trend of the disease over time. It needs to be done for each disease and should be done on a weekly, monthly and annual basis depending on the situation. Weekly analysis should compare the current week's data with the data over the past three weeks. Here one takes the current week's cases and deaths and compares it with the cases and deaths for the same disease in the same region for the previous 3 weeks. An example is shown in Fig. below. As can be seen from the example, there seems to be an increasing trend in the number of cases of malaria. This should alert the district authorities to take the necessary preventive action.

Number of cases and deaths due to malaria in Keonjhar District – Orissa – 2001



Monthly and yearly analysis looks at the secular trends and tries to identify the months in the year when the disease tends to peak. This should alert the Public health manager about the possibility of intervention to prevent the peaks for investigation and a state of preparedness. An example is given below.



The main purpose of this report is to understand the trends over time.

Monthly Analysis of Data

Reporting Site / months	J	F	M	A	M	J	J	A	S	O	N	D
Hospital 1…….												
Hospital 2……												
CHC 1												
CHC 2……												
PHC 1……												
Total												

This record will show if all the reports from different sources have been received or not over a period of time. A single missing report from a large hospital during the month when the seasonal incidence is highest can lead to a false declining trend. All reporting units must be asked to submit a **"nil"** report even if no cases are seen to rule out low incidence due to incomplete reports. The example, given in the exercise shows incomplete reporting in the district.

Annual Analysis of Data

Trends in disease incidence

The annual number of cases is compared with the number of cases of previous years to know the trends in disease occurrence, track epidemics and to monitor changes. Improvement in Immunization Programme should reflect in declining trends of disease incidence and number of hospitalized cases. The declining trend of poliomyelitis in the country over a period of time is an example.

Cyclic pattern of the disease

Prior to the disease control Programmes an increase in incidence from year to year will be noticed. Following control activities in an area and decline in incidence the interval between the peaks will however increase and the intensity of the peaks will decline. For e.g., with control of water borne epidemics the cases and deaths due to diarrhoeal disease will decline; an increase in the incidence of poliomyelitis and measles was reported every two to four years – now with increasing immunization Programme this has changed. If the cyclic pattern of the disease in the area is known, an increase in the incidence can be anticipated and precautionary measures can be taken in the high-risk pockets. If the recorded data do not show any change in the cyclic pattern despite good interventions, the quality and other operational issues need to be examined.

Seasonal pattern of diseases

The understanding of the seasonal pattern is a tool to plan or modify interventions. Many diseases have a typical seasonal pattern. In endemic areas, cases of poliomyelitis increase during the period of May to September, with the maximum cases being reported in July-August. Similarly, nearly half the cases of neonatal tetanus in endemic districts occur in the months of August and September. Recording the seasonal pattern is important for Programme interventions and for monitoring impact.

If a seasonal pattern is recorded, the disease is endemic. If high immunization coverage levels are reported in such areas, these should be verified. Cases should be spot-mapped to identify pockets of low coverage (if over-all immunization coverage levels are found to be high).

Control measures taken prior to the seasonal increase can flatten the peak and can significantly reduce the annual number of cases. Thus, if action is taken to protect all pregnant women with estimated date of delivery in August and September by TT immunization at least 4 weeks before the date and providing the women with information on hygienic and aseptic deliveries, so that they can ensure these practices by whoever assists them during delivery, number of cases of neonatal tetanus will drastically fall.

Similarly other interventions for some communicable diseases conducted prior to the known seasonal increase of cases in known areas will have a significant epidemiological impact.

IEC activities can be Programme according to needs. (See module on Planning & implementation of UIP, community participation etc.)

If no cases are reported or detected through active search in high-risk pockets during the seasonal peak period, it is very likely that there are few or no cases.

7.7 Where is the Disease Occuring – Analysis by Place

Analyzing data by place gives information about where the disease is occurring. Under IDSP this can be done with help of computers using GIS. It can also be done in simple ways manually using already available local maps. It allows the Medical officer to:

Detect any clustering of cases – e.g. if there is an increase in the number of diarrhoea cases in a given geographical place, the GIS will show this clustering by shaded colours and also indicate whether this is a sporadic increase or whether there is a real clustering in a particular village etc. The latter has more significance.

Understand some of the risk factors that may have contributed to the spread of disease – e.g. in the above instance, if the GIS map shows the clustering to be around a water source, then one can hypothesize that this may be the source of this outbreak.

Predict any potential outbreaks e.g. if the water quality in a particular area is low, then one can predict a potential outbreak of water borne disease.

Some of the possible interpretations from these report(s) are understanding

- Any increase or decrease in incidence of a disease for a particular reporting unit, (as compared to other reporting units).
- Any increase or decrease in deaths from a disease for a particular reporting unit (as compared to other reporting units).
- Place where the events are occurring
- Any clustering of cases (from the spot map).

Preparation of a map of an area

For the medical officers in a PHC or slightly higher levels where computer facilities are not available, a simple map of each area - (health worker wise) is a prima facie requirement. The DSO should highlight this point to all MOs in his jurisdiction. It should have the geographical details and identity marks demarcating different areas topographically to understand the problems and to plan out the type of strategy to be evolved accordingly. This will also help to locate the **high-risk** areas. It will also help to prepare **spot maps** showing the occurrence of the cases. Cases of a particular disease can be represented on the map in different colours.

Profiling high-risk areas

The high-risk areas are those places from where there is a greater likelihood or probability of more cases occurring and can be broadly categorized as given below for a wide variety of communicable disease and VPDS. Such areas should be shown in the map in different colours or in different shadings.

- Inaccessible areas - Hilly, tribal, desert etc.
- Areas of socio economic backwardness-refugee colonies
- Disadvantaged ethnic groups.
- Nomadic & floating population areas.
- Areas of poor immunization coverage due to any of the above.
- Areas of known endemicity.

- ☞ Peri-urban - urban slums, peri-rural - rural slums.
- ☞ Poor communication areas.
- ☞ Pockets of large vacancies of health workers.
- ☞ Poor literacy areas leading to lack of awareness and
- ☞ Unsatisfactory health seeking behaviour.
- ☞ Crowded localities with poor sanitation conditions.

Spot-map for cases by residential address

Having the general map as discussed above being available and the line list of the cases are received, these can be used for spot-mapping areas from where cases have been recorded. A colour 'dot' is put for each case in the village (or urban ward) of residence. Viewed even from a distance, it can reveal more number of cases occurring in these areas. If any area consistently reports cases, status of the services in the area should be checked and additional measures taken, if necessary.

Spot mapping of cases by residential status is important as often cases can come to a hospital from large distances. Cases may be from neighbouring districts or even neighbouring states / countries. Cases with residence in the concerned district may be clustered in selected PHC areas. If all cases are recorded by district in which the hospital is located, the analysis of the data will be skewed and will not be useful in monitoring the impact of the services in the district concerned or for initiating follow-up activities. This requires immediate notification to the concerned district (state) from where cases have come. While spot-mapping cases, it should be kept in mind that there might be more 'high-risk' areas than reflected by the reported number of cases. Cases from inaccessible areas with poor communication facilities may not come for treatment. In places where health infrastructure and IEC activities are inadequate, community awareness and health seeking behaviour may be less satisfactory than in other areas. Other parameters should also be used, in addition to surveillance data for identifying 'high-risk' pockets in each PHC area and in the district as a whole. All areas with long-term staff vacancy, difficult to approach villages with poor communication facilities, peri-urban areas with inadequate health facilities, crowded localities with poor sanitary conditions such as urban slums & rural slums and areas with recorded low coverage levels should also be considered as high risk areas. Spot maps help to identify:

- ☞ Pockets from where cases are consistently reported
- ☞ Pockets from where cases are expected but not reported
- ☞ To know the trend of the disease while comparing similar maps for the corresponding period.

The immunization status of cases of vaccine preventable diseases should be analyzed separately. As immunization coverage levels increase, the proportion of cases with history of immunization will also increase but the total number of cases will fall. A

rapid check of vaccine potency can be made. If the vaccine potency is less than expected, the reasons should be checked with the highest priority. The diagnosis of the cases should also be confirmed and second opinion taken if necessary.

7.8 Who are the Affected People - Analysis by Person

Analysis of data by some individual characteristics is important to identify affected people so that action can be initiated to protect them or to provide remedial measures.

The computer-generated tables indicate in which age, sex and place cases and deaths are being reported. The age group affected is important to prioritize action for preventive measures and IEC activities. At the national level, for example, **the median age for poliomyelitis is 18 months and for measles 24 months** (that is, half the total cases occur before 18 months and 24 months of age, respectively) highlighting the urgency for the early completion of the immunization schedule. Similarly, analysis by sex is vital to know sex differences in incidence and deaths. Line lists can be used to document the sex ratio of cases. If there is a large differential, IEC efforts must be stepped up in the area to promote equal treatment opportunities for children of both sexes. Data can also be analysed on many other variables. The immunization status of all cases should be recorded, irrespective of diagnosis. A high proportion of un-immunized cases seen in the hospitals are a reflection of low immunization coverage levels in the community

Under and Over-reporting – What does it mean?

Underreporting is nil or less reporting of cases and deaths even when actual cases are occurring. Rough estimates of the expected number of cases and deaths should be made and these should be compared with the total numbers actually reported. For all diseases and conditions, we should aim at 100% reporting. Underreporting indicates that action needs to be initiated for improving reporting systems.

Some of the possible interpretations are:

Scenario	Interpretation
Increasing trends	○ Could be a potential outbreak
	○ Could be better reporting
	○ Could be a change in the detection and reporting
Decreasing trends	○ Could indicate improved control measures
	○ Could indicate under reporting because of incomplete reports
	○ Could indicate a change in the detection and reporting
Plateau of the graph	○ Could indicate stable situation
	○ Has to be corroborated with the completeness report.

Report 4 - Crossing threshold values

This report helps to identify outbreaks early enough. The weekly / monthly data should always be compared with established threshold levels. These can be obtained in the following manner:

1. Pre-existing National / Internationally developed thresholds- e.g. a single case of measles in a tribal area is considered as an outbreak and reason for action
2. Based on historical data- e.g. if data for a particular disease is available, then the monthly mean should be calculated for the previous three years (excluding months in which there was an outbreak).
3. Increasing trends of the disease over a short duration of time (e.g. in weeks)

Some interpretations and examples of thresholds are given below in table

Interpretations that are possible are:

Scenario	**Interpretation**
Number of cases much below the threshold	No reason to worry Check for under reporting Review the threshold value
Trends approaching threshold	Potential outbreak
Number of cases have crossed the threshold	Outbreak situation, to take necessary action

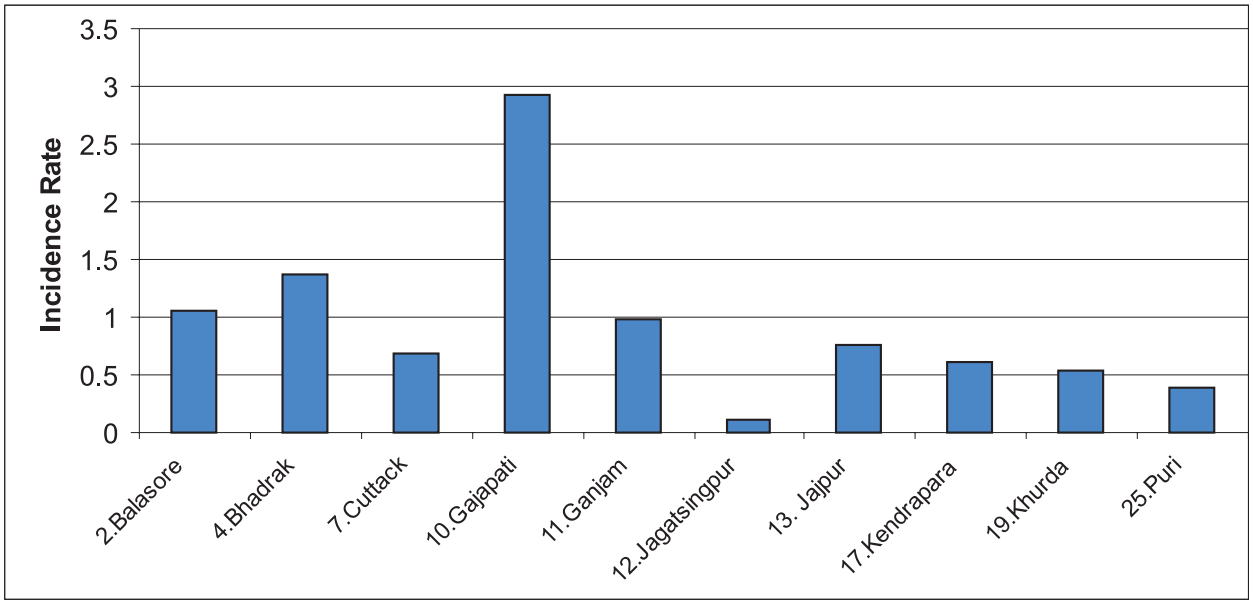
Table – Threshold levels for common epidemic prone diseases

Threshold level	**Disease / Condition**	**Action to be taken**
A single suspect case of	Cholera Dengue JE Measles Plague AFP	Immediate reporting to the next level to alert them. Investigation and confirmation of the existence of case Lab confirmation where possible Specific response if confirmed epidemiologically and/or by lab.
If the number of cases exceed the mean number of cases from the previous non-epidemic period	Diarrhoeal disease Typhoid Hepatitis Malaria Water pollution Air pollution	☞ Immediate reporting to the next level to alert them. ☞ Investigation and confirmation of existence of cases ☞ Check for epidemiological linkages ☞ Reviewing the past data ☞ Lab confirmation where possible ☞ Specific response if confirmed epidemiologically and/or by lab.
If the number of cases or deaths are increasing over a short period of time	Diarrhoeal disease Typhoid Hepatitis Malaria TB HIV Water pollution Air pollution	Immediate reporting to the next level to alert them. Investigation and confirmation of existence of cases Check for epidemiological linkages Reviewing the past data Lab confirmation where possible Specific response if confirmed epidemiologically and/or by lab.

Report 5 - Comparison between the reporting units in the region

This is a useful report and is a good proxy indicator for the quality of the data generated. One compares the *Incidence rates and Case Fatality Ratios* for the current month between the various reporting units. This should ideally be done from the Block and above level. If there are sharp rises or falls in the incidence rates (where one is not expecting one), then one should look more carefully at the veracity of the reports from that reporting unit. In a given region, one will not expect a major difference in incidence rates, unless there are some specific interventions there. This type of comparison also helps in improving the performance of the system. An example of this comparing 10 coastal districts of Orissa is given below :

Comparison of the Incidence rate for malaria for the week 27 (2001) for the 10 coastal districts of Orissa.



It can be observed that Jagatsinghpur has a very low incidence rate as compared to its neighbouring districts. As there are no particular intervention Programme in this district, one may need to look carefully at the data from this district.

Possible interpretations are:

Scenario	Interpretation
IR and CFR in the various reporting units are similar	May be indicative of good reporting mechanism
Markedly low IR / CFR in a reporting unit	Quality of data from this unit needs to be reviewed – possibility of under reporting
Markedly high IR / CFR in a reporting unit	Quality of data from this unit needs to be reviewed – possibility of an outbreak or a data entry error.

Report 6 - Comparison of reports between Public and Private sources

The data from these two independent sources is a good proxy indicator of the quality of data generated from the two sectors. The trends in the incidence of new cases / deaths in the public and private health sector may be analysed to see if they are following a similar pattern. If there is correlation between the two sources, then one can assume that the quality of data is good and it represents the events in the community. In case there is discordance between the two data sets, one has to do further operational research to identify which data source is more reliable and measures to correct the unreliable source. It should also be remembered that factors related to service utilization might influence disease occurrence patterns.

Possible interpretations are:

Scenario	Interpretation
Trends in public sources are similar to that in Private sources	Quality of data in the Public health services seems to be representative.
Trends in Public sources are not similar to that in Private sources	Quality of data in the Public health services may not be reliable (assumption being that the Private sector data is reliable).

Report 7 - Comparison of reports between public health and laboratory sources

It is important to correlate the findings of the data analysis with the availability of other data obtained from laboratories. Some examples of this comparison are:

- The cases diagnosed in the labs and the number of cases seen by the providers.
- The water quality reports and the cases of water borne diseases. For example contamination of a water source may be detected by the routine water testing and the resultant outbreak of jaundice may be well within the incubation period of the disease, thus pointing to a single source epidemic.
- The entomological data and the cases of vector borne diseases. For example a high vector density of aedes mosquitoes could clearly link to an outbreak of dengue fever in that area.

Once again this sort of comparisons should validate the data as well as identify potential areas of problems in data collection and generally in the surveillance systems.

7.9 Periodicity of Reports

The general ways of analysing data and generating reports 1 –7 have been discussed till now with some examples. Data from the other diseases like TB, HIV, Malaria, Road Traffic Accidents and others should be incorporated and analysed in a similar way. A summary should be prepared for the month which should be shared with the concerned officers at that level; e.g. at the district level, the summary should be shared at the monthly meeting of the MOs, with the Programme Managers and with the District Collector / Magistrate. This summary sheet should then be forwarded to the next level for information.

The first 4 reports need to be generated and reviewed on a weekly basis at community health centre levels. A technical committee comprising of the MO in charge of surveillance and some other medical and Para-medical workers can do this. The review should try and identify any lacuna in the system (through Report 1) and presence of any outbreak (through Reports 2 – 4)

On a monthly basis, at least 7 reports need to be generated and reviewed. The Reports 5 – 7 would help the DSO/MO to better review the performance of the surveillance system, by validating the data Based on the review, a summary should be prepared which should be sent to the next level on a weekly basis along with the compiled data. From the district level, only the weekly summary will reach the State level.

7.10 Conclusion

Analysis is the major component of any surveillance Programme linking data collection with Programme implementation. A combination of accurate data and reasonable analysis is a powerful tool to identify potential and real outbreaks and take focused action so that unnecessary morbidity and mortality are prevented. While it is important to analyze the data, it is also important that analyzed reports are sent to the appropriate authorities, both to higher and peripheral levels. The latter is very important as it gives the staff a tool to assess their own performance. This sort of feedback is also a good motivator.

However, while doing the analysis, one must also be aware of some limitations, which can always be improved upon.

- Sometimes, the quality of data may not be very high. There are various reasons starting from inconsistent use of case definitions to difficulty in confirming cases. In depth analysis on poor quality data is not of much use
- There is a time lag between detection, reporting and analysis. The ground situation may have changed by the time of analysis
- There could be under – reporting in surveillance data and it might not have captured all the health events that have occurred in the community. However surveillance data gives trends which is of importance
- The data is not representative and the only way to overcome this is to increase the sources of data, including the private sector and the NGO sector etc.
- Analysis of data for the risk factor surveillance is dealt with in the specific chapter.
- Data must be analyzed carefully and interpreted prudently
- Ability to effectively analyze, interpret and present surveillance data is an important skill for the Public Health Manager.

8. EVALUATION QUESTIONS

1. Being the DSO of district Pavagada, Please list all reports that should reach you in electronic format on Monday of every week
2. Please enumerate all steps you would undertake for obtaining and checking uniformity, timeliness and completeness of reports.
3. How would you identify an emerging outbreak in your district
4. How will you identify the seasonal pattern of diarrhoeal diseases in your district with the given set of data?
5. If data from 30 CHCs are available to you on 1st Jan 2004 on reported deaths due to malaria, how will you rank order CHCs for intervention based on prioritization and what follow-up actions you would undertake.
6. If Case fatality rate of measles in your district were reported as 10 %, what other additional information you would require for obtaining total picture of disease.
7. As DSO, list all periodical reports you need to send for higher authorities during first week of every month.
8. What measures you would undertake to check validity and reliability of data you receive from CHCs and PHCs.
9. If a local newspaper reports 52 sudden deaths due to fever and convulsions over 3 days, what data you would look for to initiate action.
10. With regard to IDSP, what issues of data quality you would discuss in the monthly meeting of MOs.