

**Exploring Survey Data on Health Care**  
**Prof. Pratap C. Mohanty**  
**Department of Humanities and Social Sciences**  
**Indian Institute of Technology, Roorkee**

**Lecture - 36**  
**Needs for Evaluation**

Welcome friends. Once again to the NPTEL MOOC module on Exploring Health Survey Data. We are on the verge of the last week of our module. Here we have completely targeted to focus on the direct application of the recent, research in healthcare by which there are several awards, several policy makings taking place.

So, the recent example that I can cite is the Nobel Prize given in a very particular year, also two years back to Professor Abhijit Banerjee and their team.

So, policy valuation has taken the centre place, centre stage in the present-day research and nonetheless healthcare has taken the most important dimension in social science, where there are a large number of research researchers trying to give certain trying to take certain attempts to evaluate different policies at different levels.

That is why we have titled the lecture on the very first lecture of the last week is Need for Evaluation. We know that need for evaluation is understood just by its name, but there are certain techniques we are going to introduce to you, and how the need is actually defined.

On the first introduction, we have kept the title of this particular page as an evaluation. We are going to cite one important definition given by u N evaluation group. United Nations there are a number of projects going on.

(Refer Slide Time: 02:23)

## Evaluation

□ The UN Evaluation Group defines evaluation as follows:

- "...an assessment, as systematic and impartial as possible of an activity, project, programme, strategy, policy, theme, sector, operational area or institutional performance.
- It focuses on expected and achieved accomplishments examining the results chain, processes, contextual factors and causality, in order to understand achievements or the lack thereof.
- It aims at determining the relevance, impact, effectiveness, efficiency and sustainability of the interventions"

And, they are sponsoring various program evaluation projects, and a large number are at present being evaluated in India, in developing countries, in African countries as well. As per the definition, this is defined as an assessment, as systematic and impartial as possible of an activity, project, program, strategy, policy, theme, sector, operational area or institutional performance.

In continuation to this definition we are also adding other important features that is this definition also focuses on expected and achieved accomplishment, examining the results chain processes contextual factors, and causality in order to understand achievements or the lack thereof.

It aims at determining the relevance, impact, effectiveness, efficiency and sustainability of the interventions. So, basically, there are so many things. First of all, it has emphasized the activities and systematic and impartial approach to strategy, policy, sectors, operational area, institutional role.

It has given importance to examining the results with a chain of strategies than their processes, contextual factor, and causality also. To link with the treatment and the control group and accordingly, we find the impact, effectiveness and efficiency or the sustainability of a particular project.

(Refer Slide Time: 04:10)

❑ Evaluations are **periodic, objective assessments of a planned, ongoing, or completed project, program, or policy.**

❑ Evaluations are used to answer specific questions, often related to design, implementation, or results.

❑ Furthermore, evaluation should influence policy and operational decisions.

swajani 3

So, the evaluations are periodic, objective assessments of a planned, ongoing, or completed project, program, or policy. Evaluations are used to answer specific questions, often related to design, implementation, or results. Further evaluations should influence policy and operational decisions.

So, the ultimate target of this evaluation is to go for policy and to find out the loopholes of the policies to get a better implemented one so that the project is going to be more sustainable.

(Refer Slide Time: 04:59)

❑ *Summative evaluation* can also be subdivided:

- *outcome evaluations* investigate whether the program or technology caused demonstrable effects on specifically defined target outcomes
- *impact evaluation* is broader and assesses the overall or net effects – intended or unintended – of the program or technology as a whole
- *cost-effectiveness and cost-benefit analysis* address questions of efficiency by standardizing outcomes in terms of their dollar costs and values
- *secondary analysis* reexamines existing data to address new questions or use methods not previously employed
- *meta-analysis* integrates the outcome estimates from multiple studies to arrive at an overall or summary judgment on an evaluation question

swajani 4

Here is the presentation of summative evaluation that can be subdivided into outcome evaluations, impact evaluations, cost-effectiveness or cost-benefit analysis and some possible of meta-analysis then secondary analysis etcetera to examine the evaluation.

The first one is called outcome evaluation which investigates whether the program or tech technology caused demonstration effects on demonstratable effects on specifically defined target outcomes. So, the target outcomes are evaluated largely through the outcome evaluation approach.

Second one is called impact evaluation is this is broader and assesses the overall or net effects, intended or unintended of the program or technology as a whole. So, this also deals with the direct outcomes, but the net effects are mostly emphasized.

The third one is more important that is called cost-effectiveness or the cost-benefit analysis addresses the questions of efficiency by on standardizing outcomes in terms of their cost or the values incurred for running the project and it is till the final outcome.

So, it is the poly the evaluation methods are also simultaneously discussed with their budget, the budget is very less. Usually, these evaluation methods are not going to be very effective. It is too less then it is not effective if it is in within a range generally it is considered to be good.

The second analysis examines existing data to address new questions or use methods not previously employed. So, this secondary analysis actually gives the gaps in the existing study and accordingly new variables or methods can be adopted. Meta-analysis as we know that the meta-analysis is the analysis of analysis studies of studies.

This is gives multiple studies to concluded and to give a judgment about certain evaluation question if we get large number of studies if we review all of them, we can get certain direction. Usually, these days we apply some meta-analysis approaches like Prisma, that is through literature review, but some conclusion can be also derived for certain policies.

(Refer Slide Time: 07:29)

□ *Formative evaluation* includes several evaluation types:

- *needs assessment* determines who needs the program, how great the need is, and what might work to meet the need
- *evaluability assessment* determines whether an evaluation is feasible and how stakeholders can help shape its usefulness
- *structured conceptualization* helps stakeholders define the program or technology, the target population, and the possible outcomes
- *implementation evaluation* monitors the fidelity of the program or technology delivery
- *process evaluation* investigates the process of delivering the program or technology, including alternative delivery procedures

So, formative evaluation includes several evaluation types that needs that is one needs assessment, then evaluate evaluability assessment, implementation evaluation, then process evaluation. Needs assessment determines who needs the program, how great the need is, and what might work to meet the need; evaluability assessment determines whether an evaluation is feasible and how stakeholders can help shape it is usefulness.

Structure conceptualizations help stakeholders define the program or technology, the target population and the possible outcomes of implementation. Evaluation usually deals with the monitoring aspect of it, monitors the fidelity of the program or technology delivery that is precisely called implementation evaluation.

Then, process evaluation basically investigates the process of delivering the programs to the larger masses especially the technology how it is reaching to the larger population all it is steps processes are actually evaluated.

So, these are all important needs (Refer Time: 08:49) assessment then evaluation or evaluability assessment, whether that is feasible or not then structure conceptualization-based approach is followed, then implementation how it is implemented that evaluation can be also done, then how the process are rightly followed or not can also be checked.

Here the schematic details that is going to guide you how and which way it follows. Evaluation types then when to use then what it shows why it is useful these all details are explained in this particular chart. Like formative evaluation suppose we just take as an

example within that we can think of about evaluability assessment or needs assessment that has when to use that.

(Refer Slide Time: 09:46)

Evaluation Types	When to use	What it shows	Why it is useful
Formative Evaluation Evaluability Assessment Needs Assessment	<ul style="list-style-type: none"> <li>During the development of a new program.</li> <li>When an existing program is being modified or is being used in a new setting or with a new population.</li> </ul>	<ul style="list-style-type: none"> <li>Whether the proposed program elements are likely to be needed, understood, and accepted by the population you want to reach.</li> <li>The extent to which an evaluation is possible, based on the goals and objectives.</li> </ul>	<ul style="list-style-type: none"> <li>It allows for modifications to be made to the plan before full implementation begins.</li> <li>Maximizes the likelihood that the program will succeed.</li> </ul>
Process Evaluation Program Monitoring	<ul style="list-style-type: none"> <li>As soon as program implementation begins.</li> <li>During operation of an existing program.</li> </ul>	<ul style="list-style-type: none"> <li>How well the program is working.</li> <li>The extent to which the program is being implemented as designed.</li> <li>Whether the program is accessible and acceptable to its target population.</li> </ul>	<ul style="list-style-type: none"> <li>Provides an early warning for any problems that may occur.</li> <li>Allows programs to monitor how well their program plans and activities are working.</li> </ul>
Outcome Evaluation Objectives-Based Evaluation	<ul style="list-style-type: none"> <li>After the program has made contact with at least one person or group in the target population.</li> </ul>	<ul style="list-style-type: none"> <li>The degree to which the program is having an effect on the target population's behaviors.</li> </ul>	<ul style="list-style-type: none"> <li>Tells whether the program is being effective in meeting its objectives.</li> </ul>
Economic Evaluation: Cost Analysis, Cost-Effectiveness Evaluation, Cost-Benefit Analysis, Cost-Utility Analysis	<ul style="list-style-type: none"> <li>At the beginning of a program.</li> <li>During the operation of an existing program.</li> </ul>	<ul style="list-style-type: none"> <li>What resources are being used in a program and their costs (direct and indirect) compared to outcomes.</li> </ul>	<ul style="list-style-type: none"> <li>Provides program managers and funders a way to assess cost relative to effects. "How much bang for your buck."</li> </ul>
Impact Evaluation	<ul style="list-style-type: none"> <li>During the operation of an existing program at appropriate intervals.</li> <li>At the end of a program.</li> </ul>	<ul style="list-style-type: none"> <li>The degree to which the program meets its ultimate goal on an overall rate of STD transmission (how much has program X decreased the morbidity of an STD beyond the study population).</li> </ul>	<ul style="list-style-type: none"> <li>Provides evidence for use in policy and funding decisions.</li> </ul>

We use it during the development of a new program then whether the proposed program elements are likely to be needed understood or accepted by the population you wanted to reach and why it is useful because it allows for modifications to be made to the plan before full implementation begins.

Similarly, the evaluability assessment and needs assessment has the same kind of information like when to use, what it shows and why it is useful. Like, if you go for an example about process evaluation or program monitoring, these features, these points are important like when to use in these two cases for process and program monitoring.

As soon as program implementation begins the process evaluation is going to be active or program monitoring is also active. So, these also gives information about during operation of an existing program, existing program. What does it shows? It shows that how well the program is working the extent to which the program is being implemented as design whether the program is accessible and an acceptable to it is target population or not?

Why it is so useful, because it checks on the process it is guarantees the right delivery of the program. This provides an early warning for any problems that may occur. This also allows programs to monitor, how well their program plans and activities are going to be

applied effectively. Outcome evaluation or objective-based evaluation where when this is useful. This is useful after the program has contacted at least one person or group in the target population.

So, basically, once persons have been contacted for their, basically for the treatment group. How these treatment groups are actually now evaluated? So, they might have understood about the objectives whether the objective have been reached to the right person or not. Why this is useful, these tells whether the program is being effective in meeting it is right objective.

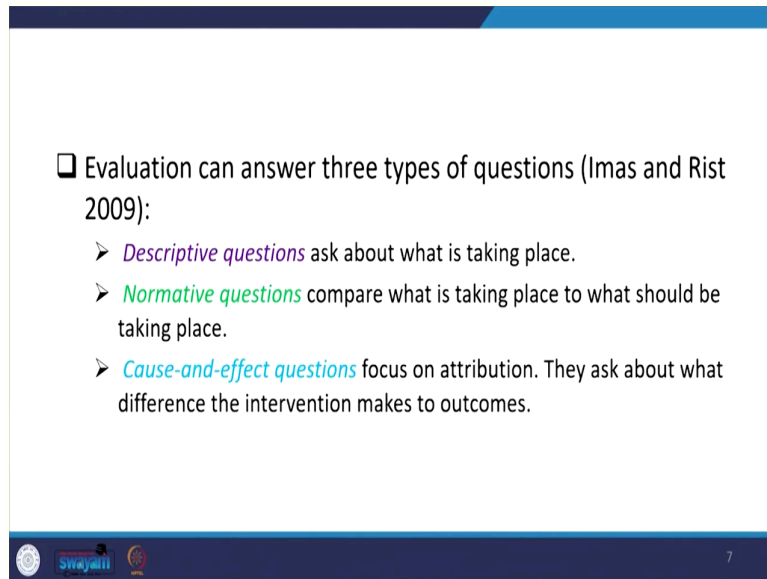
That is convinced to the persons who are the stakeholders can also be evaluated. There are various forms of economic evaluation that is very important economic evaluation you might have heard that is called cost-benefit analysis or cost analysis, cost-effectiveness evaluation, cost utility evaluation etcetera.

Last one is about impact evaluation. Impact evaluation is applied when the operation of an existing program at appropriate intervals are taken. If intervals are taken and two are compared then in that case it is called impact evaluation or generally at the end of the program, and why this is effective because this is useful because it provides evidence for use in policy and funding region.

Once this is done successfully an impact evaluation guarantees the right model was followed then that can directly be applied for policy or funding decisions. The degree to which the program meets it is ultimate goal on an overall rate of maybe STD transact transmission.

How much has the program X decreased the morbidity of an STD beyond the study sexually transmitted diseases etcetera, there are maybe many other such examples where this really going to be effective or not. So, that can be evaluated at the end of the program.

(Refer Slide Time: 13:35)



□ Evaluation can answer three types of questions (Imas and Rist 2009):

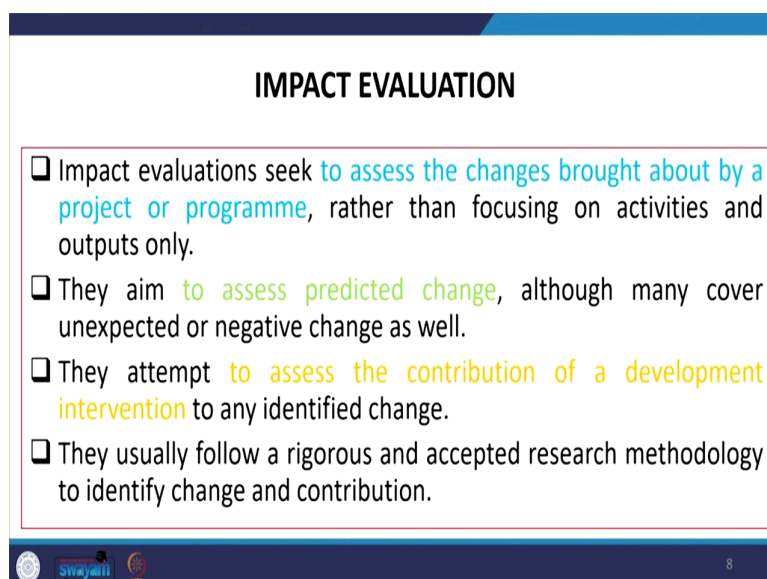
- *Descriptive questions* ask about what is taking place.
- *Normative questions* compare what is taking place to what should be taking place.
- *Cause-and-effect questions* focus on attribution. They ask about what difference the intervention makes to outcomes.

7

So, evaluation can answer three types of questions: one is called descriptive questions, then normative questions, then cause and effect related questions emphasized by Imas and Rist 2009 paper descriptive questions ask about what is taking place all about. Normative is which what compares what is taking place to what should be taking place.

That is basically comparing what should be, what is to be done. As compared to what is being what is taking place cause and effect questions focus on attribution. They ask about what difference the intervention makes to the particular outcome.

(Refer Slide Time: 14:20)



### IMPACT EVALUATION

□ Impact evaluations seek to assess the changes brought about by a project or programme, rather than focusing on activities and outputs only.

□ They aim to assess predicted change, although many cover unexpected or negative change as well.

□ They attempt to assess the contribution of a development intervention to any identified change.

□ They usually follow a rigorous and accepted research methodology to identify change and contribution.

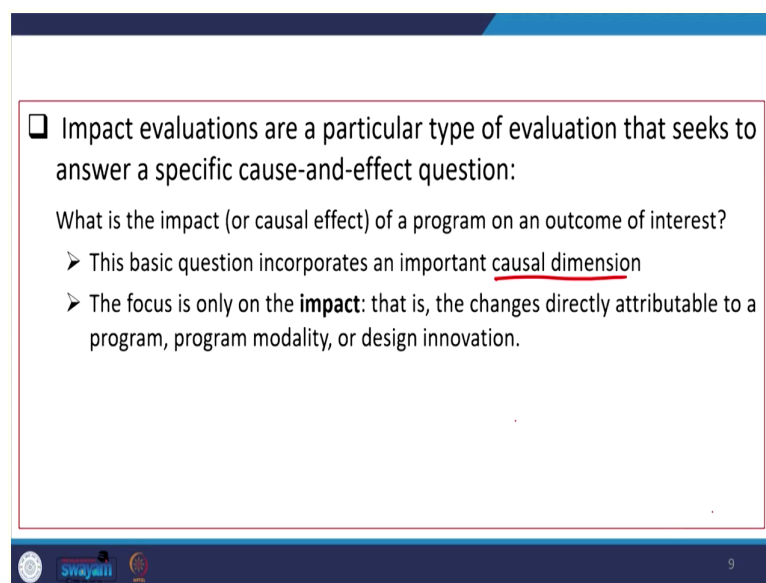
8



Now, we are explaining impact evaluation which I have already explained in the previous slides. Impact evaluation seek to assess the changes brought by a project or program, rather than focusing on activities and outputs only.

They aim to assess predicted change, although many covers unexpected or negative change as well. They attempt to assess the contribution of development intervention to any identified change. They usually follow a rigorous and accepted research methodology to identify their change and contribution.

(Refer Slide Time: 15:05)



□ Impact evaluations are a particular type of evaluation that seeks to answer a specific cause-and-effect question:

What is the impact (or causal effect) of a program on an outcome of interest?

- This basic question incorporates an important causal dimension
- The focus is only on the **impact**: that is, the changes directly attributable to a program, program modality, or design innovation.

9

Impact evaluations are very particular kind of evaluation which explains a specific cause and effect questions: what is the impact like or causal effect of an impact of a program of an outcome of interest? This basic question incorporates an important causal dimension.

So, every time we are mentioning about causal dimension the Nobel Prize is on causal inferences, you can just have a check how the persons have received, their explanations are given. Just put the search in Google causal inferences Nobel Prize you will get the information about it.

The focus is on the impact evaluation one is on is only on the impact that is how the outcome has been changed due to the treatment. The changes are directly attributable to a program, program modalities, design innovation etcetera.

(Refer Slide Time: 16:03)

❑ Impact evaluations can be carried out on any kind of work, including service delivery, capacity development, mobilisation and policy influencing, and any sector of work from health through to governance.

❑ An impact evaluation can be applied at any level of work from small projects through to complex programmes covering multiple organizations, sectors and geographic locations.

10

Impact evaluations can be carried out on any kind of work, including service delivery, capacity development, mobilization and policy influencing, and any sector of work from health through to governance. An impact evaluation can be applied at any level of work from small projects to complex programs covering multiple organizations, sectors and geographic locations. So, in many directions that can be made.

(Refer Slide Time: 16:33)

**IMPACT EVALUATION QUESTIONS**

❑ What is the causal effect of scholarships on school attendance and academic achievement?

❑ Does health insurance reduce the Out-of-Pocket expenditure of the households?

❑ To what extent were the nutritional outcomes of children affected by POSHAN Abhiyaan?

❑ What is the impact of contracting out primary care to private providers on access to health care?

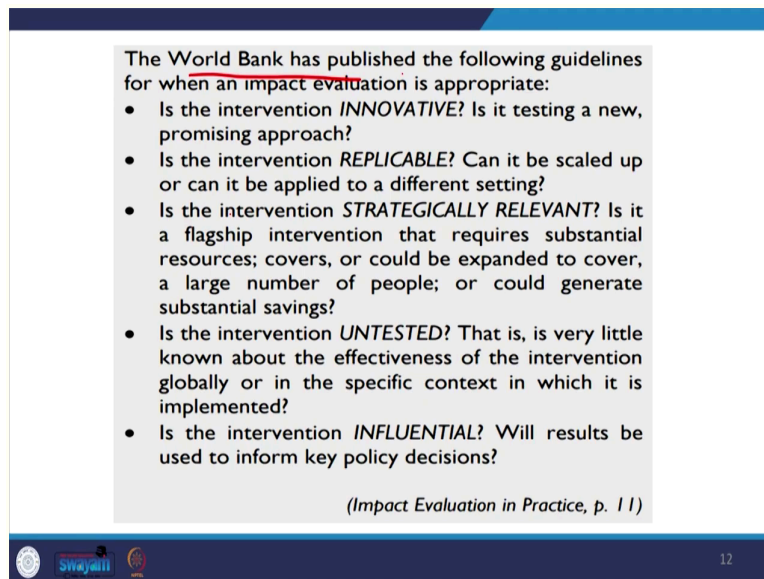
11

Impact evaluation questions some of the questions are important, like what is the causal effect of scholarships on school attendance and academic achievement. So, directly we can

ask about the causal effects of school scholarship or scholarship on them achieve performance does health insurance reduce the out of pocket expenditure of the household.

To what extent were the nutritional outcomes of children affected by POSHAN ABHIYAAN, which has been regulated at the present government initiated by the present government. What is the impact of contracting out primary care to private providers on access to health care?

(Refer Slide Time: 17:23)



The World Bank has published the following guidelines for when an impact evaluation is appropriate:

- Is the intervention *INNOVATIVE*? Is it testing a new, promising approach?
- Is the intervention *REPLICABLE*? Can it be scaled up or can it be applied to a different setting?
- Is the intervention *STRATEGICALLY RELEVANT*? Is it a flagship intervention that requires substantial resources; covers, or could be expanded to cover, a large number of people; or could generate substantial savings?
- Is the intervention *UNTESTED*? That is, is very little known about the effectiveness of the intervention globally or in the specific context in which it is implemented?
- Is the intervention *INFLUENTIAL*? Will results be used to inform key policy decisions?

(*Impact Evaluation in Practice*, p. 11)

Here is the detail when information about how World Bank has published some guidelines related to impact evaluation. How it is appropriate they talked about in different indicators like how innovative it is, how replicable it is, how strategically relevant the programme it is, whether it is tested or untested how influential this could be made. Is the intervention influential will results be used to inform key policy decisions or not etc., are mentioned by World Bank published paper?

(Refer Slide Time: 18:06)

**PROSPECTIVE VERSUS RETROSPECTIVE IMPACT EVALUATION**

- ❑ **Prospective evaluations** are developed at the same time as the program is being designed and are built into program implementation.
  - Baseline data are collected before the program is implemented for both the treatment and comparison group.
- ❑ **Retrospective evaluations** assess program impact after the program has been implemented, looking for treatment and comparison groups ex-post.

13

Coming to another difference called prospective versus retrospective impact evaluation. Prospective evaluations are developed at the same time as the program is being designed and are built into program implementation which is what is called perspective.

The example is that the baseline data are collected before the program is implemented for both the treatment as well as compare comparison group, that is why those are called baseline survey baseline surveys are made to have certain programs to be implemented later.

Retrospective evaluations assess program impact after the program has been implemented. We can retrospectively examine whether it had actually impacted the group or not. This is looking for the treatment and the comparison groups ex-post.

(Refer Slide Time: 19:04)

**EFFICACY STUDIES AND EFFECTIVENESS STUDIES**

- ❑ **Efficacy studies** assess whether a program can work under ideal conditions.
  - Efficacy studies explore proof of concept, often to test the viability of a new program or a specific theory of change.
  - If the program does not generate anticipated impacts under these carefully managed conditions, it is unlikely to work if rolled out under normal circumstances.
- ❑ **Effectiveness studies** assess whether a program does not work under normal condition.
  - When effectiveness evaluations are properly designed and implemented, the results may be generalizable to intended beneficiaries beyond the evaluation sample.

swayam 14

Others are like effect efficacy studies versus effectiveness studies. Efficacy studies assess whether a program can work under ideal conditions. Efficacy study explores proof of concept often to test the validity of a new program or a specific theory of change.

If the program does not generate an anticipated impact under these carefully managed conditions it is unlikely to work if rolled out under normal circumstances whereas, in case of effectiveness studies assess whether a program does not work under normal conditions.

Whether effectiveness evaluations are properly designed and implemented and the results may be generalizable to intended beneficiaries beyond the evaluation sample.

(Refer Slide Time: 19:58)

## EVALUATION DESIGN

- ❑ Key elements in evaluation design
  - Deciding whether to proceed with the evaluation
  - Identifying key evaluation questions
  - The evaluation design should be embedded in the program theory
  - The comparison group must serve as the basis for a credible counterfactual, addressing issues of selection bias and contagion
  - Findings should be triangulated
  - The evaluation must be well contextualized

Evaluation design: there are key elements in the evaluation design are like deciding whether to proceed with the evaluation or not identifying key evaluation questions.

The evaluation design should be embedded in the program theory. The comparison group must serve as the basis for a credible counterfactual we are going to explain just now, addressing issues of selection bias and contagion effects. Findings should be triangulated; the evaluation must be well contextualized to have a better implication.

(Refer Slide Time: 20:34)

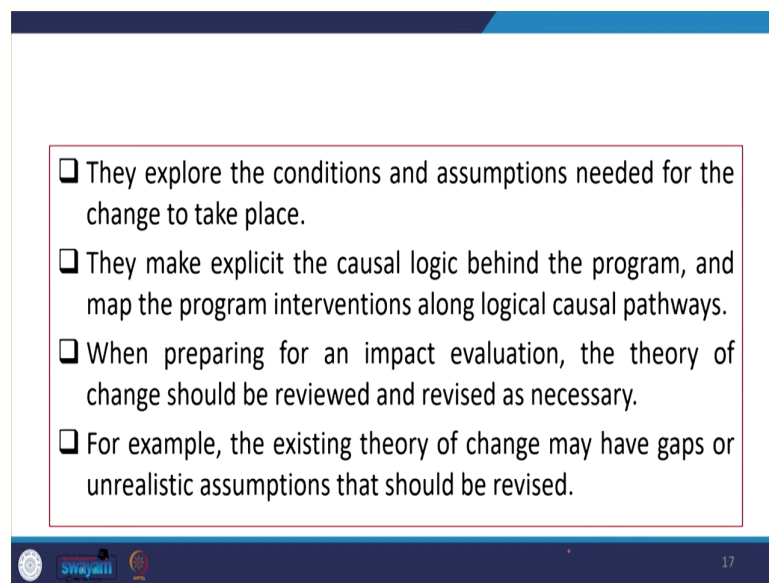
## THEORY OF CHANGE

- ❑ A theory of change explains how activities are understood to produce a series of results that contribute to achieving the intended or observed impacts.
- ❑ It is a description of how an intervention is supposed to deliver the desired results.
- ❑ It is a key underpinning of any impact evaluation, given the cause-and-effect focus of the research.

Theory of change is well referred in the context of evaluation. So, let us understand that a theory of change explains how activities are understood to produce a series of results. So, since the change is to be explained.

So, a chain of change has to be explained. This explains this contributes to achieving the intended or observed impacts. It is a description of how an intervention is supposed to deliver the desired results. It is a key underpinning of any impact evaluation given the cause and effect focus of the research.

(Refer Slide Time: 21:14)



- They explore the conditions and assumptions needed for the change to take place.
- They make explicit the causal logic behind the program, and map the program interventions along logical causal pathways.
- When preparing for an impact evaluation, the theory of change should be reviewed and revised as necessary.
- For example, the existing theory of change may have gaps or unrealistic assumptions that should be revised.

They explore the conditions and assumptions needed for the change to take place. They also make explicit the causal logic behind the program and map the program interventions along logical causal pathways, when preparing for an impact evaluation the theory of change should be reviewed and revised as per the necessity. For example, the existing theory of change may have several gaps or unrealistic assumptions that should have been revised on time.

(Refer Slide Time: 21:47)

❑ The impact evaluation can be improved by using theory of change to:

- identify relevant variables that should be included in data collection
- identify intermediate outcomes that can be used as markers of success, in situations where the impacts of interest will occur after the evaluation time frame.
- identify aspects of implementation that should be examined to see if the failure to achieve intended impacts is due to a failure to implement the intervention successfully.
- identify potentially relevant contextual factors that should be included in data collection
- guide data analysis
- provide a framework for reporting findings.

18

The impact evaluation can be improved by using the theory of change. To identify the relevant variable, the particular variable that should be included or not. Included identifying the immediate outcome of what could be the immediate success from the makers.

In a situation where the impact of interest will occur after the evaluation time frame, some basic some testing is required, and immediate outcome testing is required. This identifies aspects of implementation that should be examined to see if the failure of to achieve intended impacts is due to a failure to implement the intervention successfully.

This identifies potentially relevant contextual factors that should be included in data collection. Then the theory of change also includes some guides about data analysis. This provides a framework for reporting findings as well. Using the theory-based approach usually avoids black box impact evaluation this means that those which give a finding on impact but no indication is to why the intervention is or is not doing.

(Refer Slide Time: 22:58)



❑ Using the theory-based approach avoids 'black box' impact evaluations.

- Black box evaluations are those which give a finding on impact, but no indication as to why the intervention is or is not doing.

19

So, it is avoiding the black box impacts as well.

(Refer Slide Time: 23:08)

### DEVELOPING A RESULT CHAIN

❑ A results chain is one way of depicting a theory of change.

❑ Other approaches include theoretical models, logic models, logical frameworks, and outcome models.

❑ A results chain establishes the causal logic from the initiation of the program, beginning with resources available, to the end, looking at long-term goals.

20

So, developing a result chain is required in case of the evaluation method. A results chain is one way of depicting the theory of change as we already said. Other approaches include theoretical model, logic models, theoretical frameworks and outcome variables. This establishes the causal logic from the initiation of the program, beginning with resource able availability, or resources available to the end, looking at long-term goals etcetera.

(Refer Slide Time: 23:39)

□ A basic results chain will map the following elements:

- **Inputs:** Resources at the disposal of the project, including staff and budget.
- **Activities:** Actions taken or work performed to convert inputs into outputs
- **Outputs:** The tangible goods and services that the project activities produce; these are directly under the control of the implementing agency.

21

The basic result chain will actually map the following elements those are inputs, activities, and outputs like inputs where resources at the disposal of the project, including staff and budget etcetera should be well examined.

Activities such as actions taken or work performed to convert inputs into outputs. Outputs are the tangible goods and services that the project activities produce, these are directly under the control of the implementing agency.

(Refer Slide Time: 24:12)

➤ **Outcomes:** Results likely to be achieved once the beneficiary population uses the project outputs.  
Usually not directly under the control of the implementing agency.

➤ **Final outcomes:** The final results achieved indicating whether project goals were met.

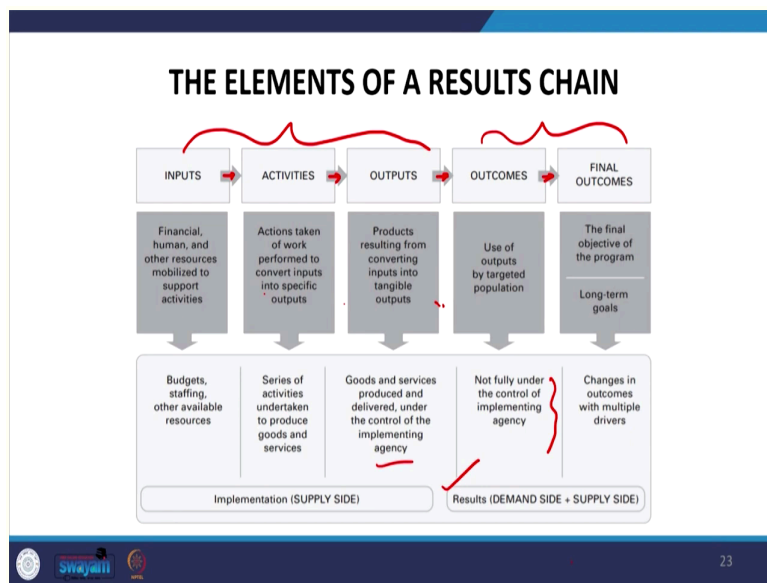
22

Similarly, besides three important direction inputs, activities outputs. The basic result chain also maps for the two other outcomes, not just output outcomes is the particular target

beneficiaries. The beneficiaries who have been targeted should have been well dealt with outputs the; of the project. Usually, this is not directly under the control of the implementing agency.

Final outcomes, the final results achieved indicating whether project goals are met or not. The implementation of a result chain follows inputs to then inputs to activities than to then to outputs.

(Refer Slide Time: 24:59)



Activities to output then output to outputs to your outcomes then output to final outcomes. The inputs we have already discussed it include budgets also, staffing etcetera activity series of activities taken in between outputs. Basically, whether the products to be delivered on time under the control of the implementing agency the here, the implementing agency control is there. In case of outcome implementing agencies is not having any control.

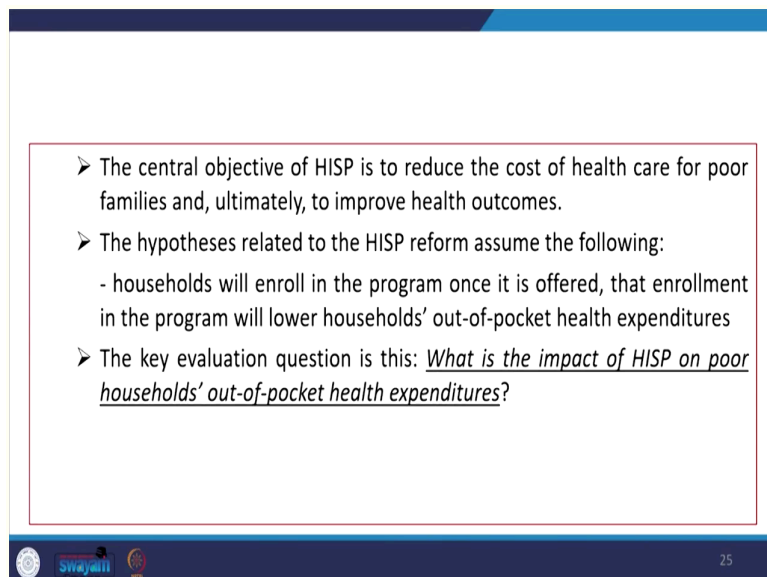
So, these first three sides that is inputs, activities and outputs is usually referred as supply-side implementation. Other two since we are saying the beneficiaries and the final reachability of the project we are linking with the demand side as well as the supply side. So, the, therefore, these are called the real results like in technical institutions do invent many things, but many of the inventions are not catering to society.

So, in the program implementation we do appreciate the stand of the technical institutions, but in the end of the result that those may not be sustainable because it is not going to deal

with the final beneficiaries. So, example like health insurance subsidy program, the ultimate objective of health insurance subsidy program is to improve the health of the country's population.

The government is concerned that poor rural households are unable to afford the cost of basic healthcare with detrimental consequences for their health. To address this issue HISP subsidizes health insurance for rural households that has covered their cost, relative there relative to their other healthcare and primary healthcare and medicinal cost. The central objective of HISP is to reduce the cost of healthcare for poor families and to reach the masses and improve their health outcomes.

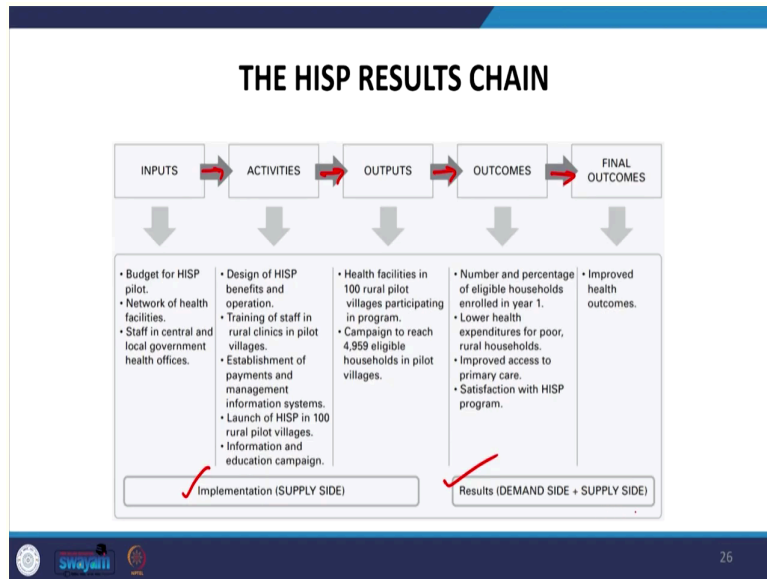
(Refer Slide Time: 27:16)



- The central objective of HISP is to reduce the cost of health care for poor families and, ultimately, to improve health outcomes.
- The hypotheses related to the HISP reform assume the following:
  - households will enroll in the program once it is offered, that enrollment in the program will lower households' out-of-pocket health expenditures
- The key evaluation question is this: What is the impact of HISP on poor households' out-of-pocket health expenditures?

The hypothesis related to the HISP reform assume the following that is the household will enroll in the program once it is offered, and enrolment in the program will lower households out of pocket healthcare expenses. The key evaluation question is that what is the impact of HISP on poor households out of pocket health expenditures.

(Refer Slide Time: 27:40)



The HISP results also follow inputs to inputs to these activities then outputs then outcomes then final outcomes. So, accordingly, you can think of their supply-side indicator and the demand and demand-side indicators as well along with the supply side indicators.

(Refer Slide Time: 28:03)

### CAUSAL INFERENCES

Many policy questions involve cause-and-effect relationships:

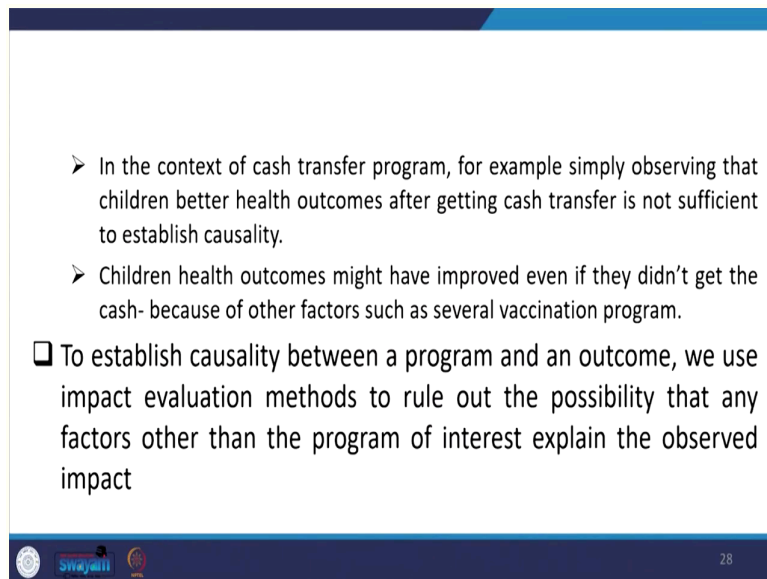
- Does teacher training improve students' test scores?
- Do conditional cash transfer programs cause better health outcomes in children?
- Do vocational training programs increase trainees' incomes?

Although cause-and-effect questions are common, answering them accurately can be challenging.

The causal inferences are very important in the case of evaluation method many policy questions involved cause and effect relationships for example, does teacher training improve students test scores? Do conditional cash transfer programs cause better health outcomes in children? Do vocational training programs increase trainees out in incomes etc.?

So, these are some of the questions usually asked we have referred here for your easy understanding. Although cause and effect questions are in common answering them accurately can be actually very challenging and should make it scientific even more challenging. In the context of cash transfer programs or direct cash benefit or direct transfer benefit programs for example, simply observing children.

(Refer Slide Time: 29:03)



➤ In the context of cash transfer program, for example simply observing that children better health outcomes after getting cash transfer is not sufficient to establish causality.

➤ Children health outcomes might have improved even if they didn't get the cash- because of other factors such as several vaccination program.

☐ To establish causality between a program and an outcome, we use impact evaluation methods to rule out the possibility that any factors other than the program of interest explain the observed impact

28

Better health outcomes after getting a cash transfer are not sufficient to establish causality. Children's health outcomes might have improved even if they did not get the cash because of other factors such as several vaccination programs. Usually, even if cash is not there by the other programs might have benefited.


So, policy evaluation should be very particular in establishing what exactly they are going to observe and how they can utilize the other effects. To establish causality between a program and an outcome we use impact evaluation methods to rule out the possibility that any factors other than the particular program, we are interested in evaluating really explaining the observed impact.

(Refer Slide Time: 29:49)

□ The answer to the basic impact evaluation question—what is the impact or causal effect of a program (P) on an outcome of interest (Y)? is given by following formula:

$$\Delta = (Y|P=1) - (Y|P=0) \rightarrow cf$$

➤ It states that the causal impact ( $\Delta$ ) of a program (P) on an outcome (Y) is the difference between the outcome (Y) with the program (in other words, when P = 1) and the same outcome (Y) without the program (that is, when P = 0).



The answer to the basic impact evaluation question is that what is the impact or causal effect of a program that is P-value or sorry or the P program success of our program effect of, or the causal effect of a program; that is on an outcome of interest Y.

Now, P here we are taking a program and the outcome we are taking with the notation as Y is given as a delta that changes in the outcome to that of the change in the P. Why? What is it?

Its outcome with a changed with the one that is the treatment group as compared to the outcome as 0, and the control group when there is no program P is equal to 0 no program and here we are actually subtracting the impact of change in the program and it is outcome value, because of the change as compared to no change in the P program.

So that means, the control value as compared to the treatment group if there are any changes that is captured and that is all about called evaluation. It states that the causal impact of a program of on an outcome Y is the difference between the outcome Y with the program. In order in other words P when P is equal to 1, the outcome without the program outcome that is equal to Y without the program when P is equal to 0.

(Refer Slide Time: 31:23)

## THE COUNTERFACTUAL

- ❑ The **counterfactual** is what would have happened—what the outcome (Y) would have been for a program participant— **in the absence of the program (P)**.
- ❑ In the basic impact evaluation formula, the term  $(Y|P=0)$  represents the counterfactual.
- ❑ Since we cannot directly observe the counterfactual, we must estimate it.
- ❑ Solving the counterfactual problem would be possible if the evaluator could find a “perfect clone” for a program participant



30

Now, a most important aspect is called counterfactuals. Counterfactuals are we discussed in program evaluation largely the counterfactual is what would have happened and what the outcome would have been for a program participant in the absence of the program.

In the absence of the program, if the program is not there then what would have happened to these in a particular group is basically called counterfactual, if we just go by no availability of the program then who are those people, who are those nonbeneficiaries and what has happened with them without the program.

They are actually called counterfactuals. In the basic impact evaluation formula, the term  $Y$  which is  $P$  is equal to  $0$  represents the counterfactuals. So, this is what is not though we are saying control, actually called counterfactuals. So, since we cannot directly observe the counterfactual, we must estimate the counterfactual.

Solving the counterfactual problem would be possible if the evaluator could find a perfect clone for a program participant. If a program participant and a replicated one without the program participant is represented, then the two groups can be compared.

So, comparing the treatment group with the counterfactuals is actually going to give us better results in terms of evaluation. The key to estimating the counterfactuals for program participants is to move from the individual or unit level to the group level.

(Refer Slide Time: 33:05)



Slide 31 contains three bullet points:

- ❑ The key to estimating the counterfactual for program participants is to move from the individual or unit level to the group level.
- ❑ The group that participates in the program is known as the **treatment group**, and its outcome is  $(Y/P=1)$  after it has participated in the program.
- ❑ The statistically identical comparison group is the group that remains unaffected by the program, and allows us to estimate the counterfactual outcome  $(Y/P=0)$ .

The slide footer includes logos for Swajail and a page number of 31.

The group that participates in the program is known as the treatment group and its outcome is  $Y$  given  $P$  is equal to 1 after it has participated in the program. The statistically identical comparison group is the group that remains unaffected by the program, and allows us to estimate the counterfactual outcome with  $P$  is equal to 0.

(Refer Slide Time: 33:32)

Slide 32 contains two main bullet points:

- ❑ The challenge of an impact evaluation is to identify a treatment group and a comparison group that are statistically identical, on average, in the absence of program.
- ❑ A valid comparison group
  - has the same characteristics, on average, as the treatment group in the absence of the program;
  - remains unaffected by the program;
  - would react to the program in the same way as the treatment group, if given the program.

The slide footer includes logos for Swajail and a page number of 32.

The challenge of an impact evaluation is to identify a treatment group and a comparison group that is statistically identical on average. So, this is what on average that should be statistically identified and statistically identical. So, these two groups in the absence of a particular program.

A valid comparison group has the same characteristics on average as the treatment group in the absence of the particular program. This remains unaffected by the program and would react to the program in the same way, as the treatment group if given the program.

(Refer Slide Time: 34:10)

❑ When the comparison group does not accurately estimate the true counterfactual, then the estimated impact of the program will be invalid.

❑ In statistical terms, it will be biased.

33

When the comparison group does not actually estimate the true counterfactual then the estimated impact of the program would be actually invalid. In statistical term, it will be called biased or giving biased results.

(Refer Slide Time: 34:25)

### TWO COUNTERFEIT ESTIMATES OF THE COUNTERFACTUAL

- ❑ **Before-and-after comparisons**
  - Also known as pre-post or reflexive comparisons
  - Compare the outcomes of the same group before and after participating in program.
- ❑ **Enrolled-and-nonenrolled comparisons**
  - also know as self-selected
  - compare the outcomes of a group that chooses to participate in a program with those of a group that chooses not to participate.

34

Two counterfeit estimates of the counterfactual that is called before and after comparison enrolled and non-enrolled comparisons. Before and after comparisons are also known as pre-post or reflexive comparisons this compares the outcomes of the same group before and after participating in the program.

Enrolled and non-enrolled comparisons are like, these are also known as self-selected comparisons. This compares the outcomes of a group that chooses to participate in a program with those of a group that chooses not to participate.

So, these are the two-comparison. There are some methodologies in impact evaluation mostly applied those are called experimental design and called non-experimental design.

(Refer Slide Time: 35:21)

The slide is titled "METHODOLOGIES IN IMPACT EVALUATION". It is divided into two main sections: "Experimental" and "Non-experimental". Under "Experimental", there is one bullet point: "Randomized Control Trials". Under "Non-experimental", there are three bullet points: "Propensity score matching", "Regression discontinuity approach", and "Difference-in-difference". The slide has a blue header and footer. The footer contains logos for Swajati and NPTEL, and the number 35.

**METHODOLOGIES IN IMPACT EVALUATION**

- Experimental**
  - Randomized Control Trials
- Non-experimental**
  - Propensity score matching
  - Regression discontinuity approach
  - Difference-in-difference

Experimental design we use randomized control trial and in non-experimental design there are largely three called propensity score matching, regression discontinuity approach and difference in difference approach.

So, these four aspects we are going to discuss in our respective lectures in this week, and we will also try to cite examples for your reference and, I am quite sure that, this is going to be useful, and will be stimulating your research further.

With this basic guidance about evaluation, I am quite hopeful that you have been recharged to carry forward evaluation-based methods any sort of doubts you have, do not hesitate and put these questions on the screen of the NPTEL query box. We will be happy to address it.

Thank you.