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

Lecture - 05 Understanding Healthcare Database: LASI

Welcome students to this module of NPTEL on health care research and database. We are on the fifth lecture to Understand Health Database from LASI. LASI is one of the finest dataset available in India and along with many collaborators and we are going to use it for our analysis. So, I thought it is better to include in the model that is this database is explaining about our old age longitudinal aging survey of India.

So, let us move forward and understand the details of this dataset. Myself Dr. Pratap Mohanty, I have been teaching over 14 plus years and the research methodology class I have been teaching over 6 years and regarding healthcare research I have been working on this area and trying to specialize in this area and I have already handled a previous course on handling large scale dataset with Stata.

So, this time it is purely on health care research we will be providing hands on experience on datasets and its analysis. In the introduction, I will give you the complete backdrop of the requirement of LASI dataset, why longitudinal aging survey is necessary and why should we focus on the research and aging population?

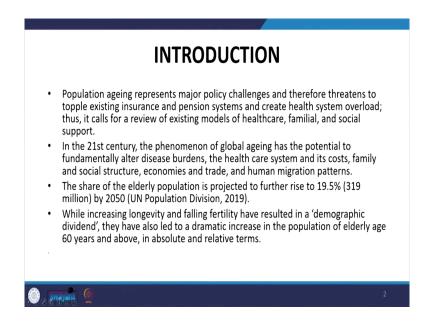
As we know that India is heading towards the most populated figures in the world, this is at present the ranked second, it is expected that in another 10-15 years, India is going to be the most populous country in the world.

And it's not just the total population that matters it also matters the composition of population area, structure of its population. So, population aging is important within the composition of population. Why it is important? Because it represents major policy challenges and threatens to topple existing insurance pension systems and create health system overload.

Thus, it calls for a review of existing models of health care family and social support. So, the existing health care model has to be reviewed since our population is aging. Though, at this moment India has huge young population, much better than that of for the China and that is why often it has been discussed within a population dividend in the literature.

So, in the 21st century the phenomenon of globalizing has the potential to fundamentally change the burden of disease. The health care system and its cost, family and social structure, economies and trade and human migration really matters in the present context. The share of the elderly population is projected to rise to 19.5 percent by 2050 as per the UN Population Division 2019.

(Refer Slide Time: 04:12)



So, by 2050 we will have a population of 319 million rise in the population increasing longevity and falling fertility have also resulted in a demographic dividend the present phase of population transition is that we have increasing longevity and falling fertility.

The death rate has been reduced as well as the crude birth rate has also been in a control, in that case we are left with young population and that dividend is expected to be very high. So, in this context it is important to highlight the dividend and that was possessed with the elderly population specially 60 years and above.

If they are not guided from the pre elderly to the elderly (pre elderly we mean 45 years onwards), that may create some form of disasters. So, the dramatic increase in population elderly and 60 year above in an absolute and relative terms are important dimension of research in the present days.

(Refer Slide Time: 05:37)



So, older age here we refer to 45 years and above form this form an age group that represents pre-aging and pre-retirement phases. This 45 years and above this is basically called pre-aging, 60 years and above is called elderly population. Oldest population is defined as 75 years and above as per the elevation and WHO definition.

Even this report has also clearly mentioned that these are different brackets of old age population i.e., pre-elderly, elderly then old age and oldest old (75 years and above). And, why this study is essential because when the population are aging it has social economic and health dimension primarily.

Those social dimension include social structure, there is perseverance of the society, then social networking, the societal adjustment within the family structure how old age are being put, how they are being treated within the social norm these are becoming challenging in the present days context.

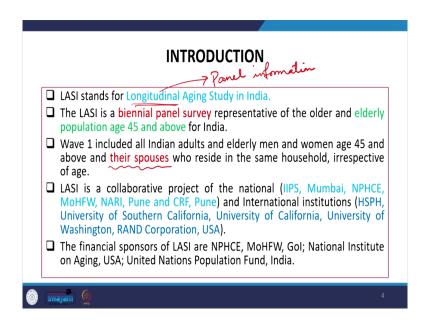
They are often understood in different newspaper article of their desertion they have been highly deserted in the old age and most importantly the old age care units are not well built in Indian context. So, this segment is upcoming therefore, studies is very essential.

Then coming to the economic front of aging, since persons are aging the morbidity percentage is also expected to be higher and so, the aged population and their expenditure, their health care unit, their inpatient care expenditure, their nutritional requirement, their settlement, their protein diet their all sort of aspects require management through its economic front.

How the budget should be allocated on them, how their pension should be dealt, how their insurance schemes are dealt, whether those are settled correctly or not since old age population rising. So, it is essential for a detailed study.

Now, similarly health dimensions of aging. So, as we already mentioned there are various indicators of health dimensions and those are included in this survey. So, these all suggest that there should be an exclusive study and the LASI has already considered it. Now coming to the LASI datasets this stands for Longitudinal Aging Study in India.

(Refer Slide Time: 08:56)



The LASI is a biennial panel survey and is quite representative one of the older and elderly population with an age of 45 and above. This 45 and above population as I have already said that these includes you know pre elderly, elderly and late elderly or oldest old population as well.

Why this is a longitudinal survey? Longitudinal survey, it is going to give the portion and a study over time period, this is going to give us panel information. Panel information means repeated cross sectional. The cross sectional units are repeated in different time period.

And this is going to be very helpful for policy making for policy evaluation or evaluation of the elderly programs and so, far only one round has been released and though the you know pilot study is also there that was conducted in 2010 we are going to discuss this in detail. So, wave 1 included all Indian adults and all elderly men and women with age 45 and above and their spouses.

So, this is important to note their spouses who reside in the same household are also covered. Now one critical aspects of this dataset is that if the spouses who are also covered are also asked even if they may not be of 45 years age or above.

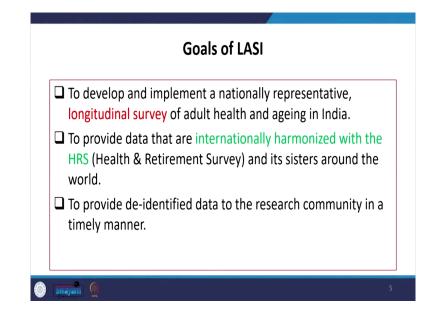
So, that is a concern you need to thoroughly understand. So, the kind of well being of the elders the kind of welfare analysis we wanted to do support in certain behavioral analysis we will do with the dataset, but the spouse age is not in that bracket then it is going to be problematic, but what sort of research question lies in that section we are going to understand.

LASI is a collaborative project of the national and international institutions, national institutions like Indian Institute of International Institute of Population Sciences located in Mumbai, NPHCE, MoHFW, NARI and CRF in Pune, international institutions like HSPH that is in University of Southern California, University of California, University of Washington Rand Corporation of USA.

So, these are the institutions attached with the study the financial sponsors of this study are NPHCE, MOHFW Government of India, National Institute of institute of Aging USA and then United Nations Population fund UNPF of India. So, in the introduction we have discussed all those things.

Coming to the certain goals of this LASI dataset, this helps in developing and implementing a Nationally Representative Longitudinal Survey of Adult Health and Aging in India. This provides data that are internationally harmonized with the HRS that is Health and Retirement Survey.

(Refer Slide Time: 13:22)

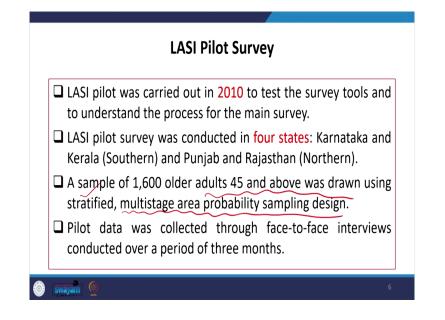


International there are datasets and this has also tried to go as per the standardize or course harmonize variables. So, international dataset and its sisters around the world to provide de-identified data to the research community in a timely manner.

So, though data, but not exactly identified, but certain codes are given, certain identifiers are given for each research community, research respondent in different time period. So, this is broadly speaking this gives representative data about adult age especially aging and that too if you are interested to work for so many countries this dataset is also comparable. So, that is one of the important aspects of LASI dataset.

LASI pilot survey was conducted in 2010 was carried out in 2010 to test the survey tools and to understand the process for the final survey and that was conducted in 2019-20. So, I just wanted to go back to check once again that is just fine. So, just a minute I will come back to it this was alright. So, goals I have mentioned.

(Refer Slide Time: 15:19)

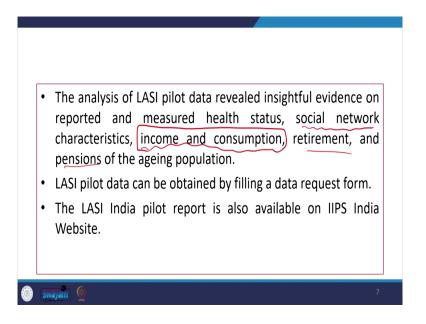


So, I am coming to the 2010 pilot survey. This has given survey tools to understand its main survey that was conducted in 2017 to 19. The LASI pilot survey was conducted in four states. So, Karnataka, Kerala, Punjab and Rajasthan, but it has not given whether these are going to be representative, but in small scale those data could be compared. Punjab and Rajasthan from the northern part and Karnataka and Kerala are from the Southern part of India were taken.

A sample of 1600 older adults with age 45 years and above was drawn using stratified multi-stage area probability sampling design. So, it has again followed multi-stage area probability sampling. So, in any case probability sampling is there and this is stratified to different strata are defined. So, the pilot data was collected through face to face interviews, conducted over a period of three months in the LASI pilot survey.

So, it was a face to face interview and that too spanned over 3 months period. So, this is the background of the LASI pilot study. The analysis of the LASI pilot data revealed insightful evidence on reported and measured health status, social network characteristics, income and consumption retirement and pensions of the aging population.

(Refer Slide Time: 17:12)



The LASI pilot data can be obtained by filling the data request form and through the process will be guiding the LASI data pilot report is also available on IIPS India website. So, this pilot survey gives information or report on major health status, social network characteristics income consumption, retirement and pensions of the aging populations.

So, you can take note social network is there those who are working on these issue they can compare, similarly pensions issues are there, retirement and pensions, income and interesting part of this is that it gives better structure about income and consumption unlike the NFHS dataset.

(Refer Slide Time: 18:08)

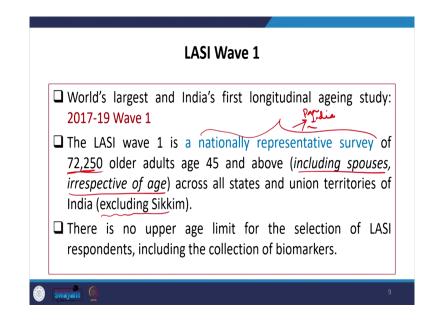
Results	Urban	Rural	Total
Household interviews			
Household selected	300	725	1025
Household interviewed	280	669	950
Household response rate1	93.33	92.28	92.68
Individual interviews			
Number of age-eligible individuals	495	1266	1761
Number of age-eligible individuals interviewed	472	1211	(1683
Individual response rate2	95.35	95.66	95.57
Biomarker interviews			
Number of eligible persons for biomarkers	472	1211	1683
Number of eligible persons who completed biomarkers	423	1107	(1530
Biomarker response rate 3	89.62	91.41	90.91
Dried blood samples (DBS) collection			
Number of persons who consented for DBS	366	1053	1419
Number of DBS collected	348	1008	1356
DBS response rate4	95.1	<u>95.7</u>	<u>95.6</u>
Note			
This table is based on an unweighted sample. Eligible individuals were persons ag	ed 45+ years and their spouses	less than 4	5 years.
 Indicates the number of households Interviewed/households selected. 			
2. Indicates the number of eligible individuals interviewed/number of age-eligible	individuals.		
 Indicates the number of eligible biomarkers completed/number of eligible biomarkers. 			

Then how you should go for it? You should actually also try to understand this pilot of 2010. Let me read number of household number of interviews with LASI is eligible individuals, number of biomarker responses and response rates according to the place of residence covered in the LASI pilot 2010 is given here. Like in the pilot the household interviewed were 950, I will check once again the figure which I have already said 1600 older respondent adults were there with 45 age and above.

Here we are also going to mention around 1600 individual interviews. So, it is of number of age eligible individuals and number of age eligible individuals interviewed. So, when finally, interviewed was of 1600 plus 1683 and the biomarker interviews were also taken of those persons out of them only around 1530 respondents were there.

Dried blood samples were collected on 1356 persons. So, the response rate here in the household level was 92.68 whereas, at the individual level it is higher i.e., 95.57, let us compare a bit for urban and rural. In NFHS rural representation were better, but here household representations were better in urban areas and specially for the individual responses it is relatively better in rural area.

Biomarkers information's were taken better in rural areas as well. This table is also un-weighted, we represented in case of any purchase. Similarly, if it is un-weighted then what weight should be taken we will guide you through the process. So, I think I have already mentioned everything. (Refer Slide Time: 20:44)



Let us make a move to the next slide. We are now trying to understand LASI wave 1, this is in fact, the world's largest and the India's first longitudinal aging study conducted in the year 2017-19. This is also called wave 1 in every two years I think the other rounds are going to be conducted till 25 years study will be there. The LASI wave 1 is a national representative survey.

Those who are not familiar with what do you mean by national representative that they indicates in simple term at this moment I am guiding this indicates results the data and its results are going to represent the Indian. So, like in any kind of estimation you do you can claim that this is what happening in all India level.

So, national representative here in Indian context we are saying. So, there are of 72,250 older adults with the age 45 and above including spouses irrespective of their age.

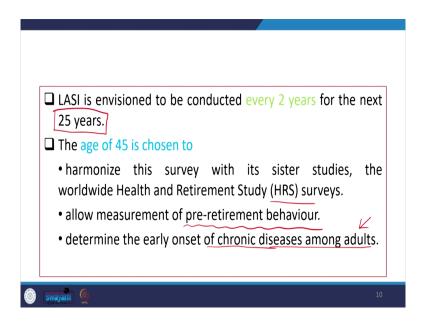
Now, this is what I said including spouses irrespective of their age. This is one of the catch we have to be careful while analyzing the dataset. So, spouse is usually in Indian context the wife between husband and wife, wife age is less, on the average, it is lesser.

So, that means, whoever are the respondent of 45 we may get very lower age group, but if you are simply getting the responses on the average those persons who are in the young age they might have also given their responses that may not be giving right result.

Now, if you are getting the average of those old age responses. So, we may need to carefully check their age. So, all may not be of 45 years and above. So, that is one of the aspects we need to be careful while checking or testing with the result.

This dataset wave 1 covers all the states and union territories except Sikkim. There is no upper age limit of the respondents and it also includes the collection of biomarkers as well bio biomarkers by which markers are also attached.

(Refer Slide Time: 23:56)



LASI dataset is envisioned to be conducted in every 2 years for the next 25 years. I have already mentioned this for another 25 years this data is going to be available so; that means, in every 2 years it will be conducted. So, at least we will have 12 waves.

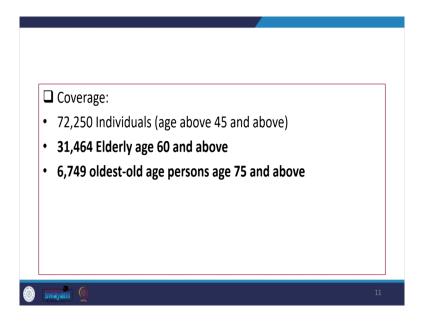
And this is going to be very interesting dimension for research those are the researcher at this moment walking using the old age dataset that is LASI dataset. They are going to exercise going to an experiment for another 12 years for their policy making. Their results will be quite interesting since these are longitudinal data.

The age of 45 is chosen to harmonize the survey with the sister studies the worldwide health and retirement study. So, this has attempted to also sync in with the existing sister studies. So, why 45 years year is the age considered as the benchmark is, because of this reason. This has also allowed measurement of pre-retirement behaviour. So, pre-retirement behavior, there are certain understanding with the person like you might have heard in the news or from different persons, those who are at the pre- retirement they might not be serious in their parent organizations where they are working, they use to say that I am going to get retired on another one year or so, so I am not putting my efforts.

And regarding health care behaviour they might be over conscious since they are approaching towards you know 60. So, they might be over conscious about their health. So, we need to check the original data and the responses those are received and to identify whether this is correctly fitting to the analysis for the Indian context.

This dataset determines the early onset of chronic diseases among adult, this is what the health care aspects which I wanted to refer. Early onset of chronic disease among adults. So, if we are referring to before 45 age, early onset of some of the Non Communicable Diseases NCDs could be well tapped.

(Refer Slide Time: 26:57)



So, then the next aspect we will emphasize is, its coverage; coverage of LASI dataset that was of 72,250 individuals covered on this wave 1 and it includes all age above 45 and above.

So, out of these 72,250 individuals respondents, we have in this data 31,464 elderly we have defined the age group that is of 60 years till 74 and 60 years and above elderly and out of that 6,749 oldest old age persons with the age 75 above and so, they are also covered and if we want to stick to their individual perspective you can able to find out correctly.

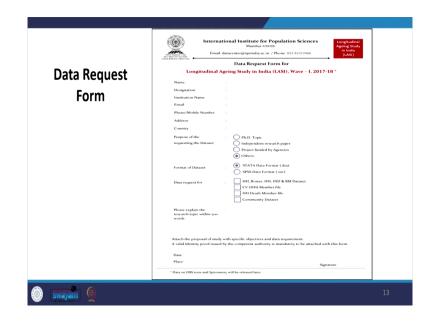
Now, we are guiding how to download this dataset if you are very new to IIPS dataset in the last lecture I have guided you about IIPS dataset it's a national family health survey data and that is considered to be one of the best dataset for the women questionnaire.

(Refer Slide Time: 28:25)

•	Visit IIPS India website at https://www.iipsindia.ac.in/lasi
•	Go to publications section on the website and download data request form.
•	Fill the form and send it to the Email Id: <u>datacenter@iipsindia.ac.in</u>
•	The IIPS will normally review all data requests within 24 to 48 hours on working days.

Here we are focusing on the old age, the old age covers information about all phases of their old age period. You can click on this website iipsindia.ac.in/lasi. LASI gives information about it correctly to go to publication section on the website and download the request form and fill the form and send it to the data center email is presented here. The IIPS usually again take one day to two days to respond and on its working days to make you available with the data.

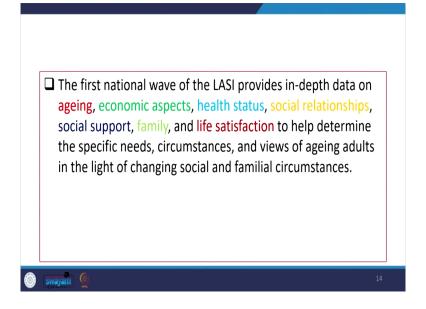
(Refer Slide Time: 29:27)



Now, the longitudinal aging study dataset, it is how you could able to request the form we have just given you for your clarity you might get confused what is that form which form how it looks. You can choose the purpose of the study for the PhD or others format of dataset, you may take Stata format or SPSS data format and which file you require in this case I think you require the community dataset household dataset and the member file as well.

Individual household and all datasets I will suggest you to download give that and in the end write down your reasons for downloading all your research topic for what you are working and with your signature you just send it to the email I have guided and your data will be sent to you or the link will be sent to you by the office IIPS office.

(Refer Slide Time: 30:17)



The first national wave of the LASI provides in depth data on ageing, economic aspects, health status, social relationship, social support, family, social satisfaction and these help in determining the specific needs circumstances views of aging adults in the light of changing social and familial circumstances.

Over the time I think we all know that the society is changing, society is getting reformed, society is adopting various practices specially in the digital days, in the ICT days, in the automation old age are the biggest prey to various difficulties and they are completely isolated specially in the young generation are more charged with the social networking sites. And much information they used to get it through that whereas, the old age are not trained and they usually are well connected with their peers and because of the economic transformation or the structural change in the Indian system these groups are highly isolated.

That is a biggest concern and not just isolated this has long connection with their health issues. Specially during COVID-19, this is quite important, this is really important because in the COVID-19 over last two years they are completely paranoid and these groups require huge support their usual consultation with the doctors are delayed and there have been many instances where the old age sections be borne the huge burden of the disease.

Now, this pandemic many old age population could not able to continue sustain their life and that is not due to the virus itself that is really due to the present health care system and how our later ages and their age group populations are dealt and how they have been actually managed.

This is a bigger question at this moment and I am sure if you target yourself to do certain study on this light I am sure your work is going to be relevant for some years.

Demographics Work and employment Household economic status The LASI's key Biomarkers objective is to provide Chronic health conditions comprehensive data • Health insurance and healthcare utilization on following: Symptom-based health conditions • • Family and social networks Functional health • Welfare programmes Mental health retirement, satisfaction, and life expectations 💿 swayan 🛞

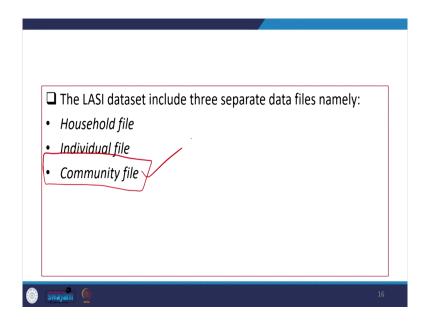
(Refer Slide Time: 33:07)

So, the LASI's key objective is to provide comprehensive information on demographics work and employment, household economic status, their biomarkers about their health conditions especially their chronic health conditions, insurance, health insurance health care utilization pattern through the insurance symptom based on health conditions their family and social networks.

And social networks like they are acquaintances with different layers in the society, how frequently they meet and how they are discussing especially in the COVID-19 period they are in trouble and though these data is not given in the wave 1, but the next wave would be maybe cover it.

The functional health care aspects are also mentioned the welfare programs which are relevant for this group is necessary and that those are also covered the mental health aspects is the emerging one in the section and they are also dealt. The retirement issue, satisfaction life expectations etc. are carefully covered in this dataset.

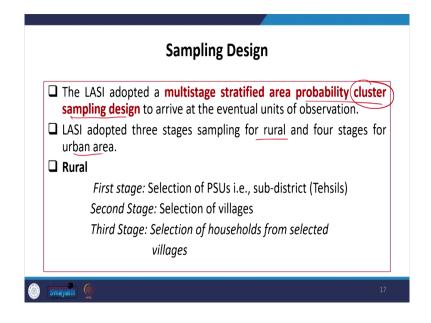
(Refer Slide Time: 34:34)



The LASI datasets includes three separate data files namely household file, individual file and community file. One of the important aspects of this dataset that it covers exclusive and comprehensive details and community file where the old age are highly connected. I will suggest all of you to follow this questionnaire and you will get lots of inferences about your other research maybe policy based research you are targeting.

This is what is in this dataset. We are now discussing about the sampling design that was followed: it is of the LASI adopted multistage stratified area probability cluster sampling design. So, some of the sampling probability sampling design we have discussed. So, here it is only not differentiating with clusters we have already discussed about cluster sampling in our previous module on handling large scale dataset.

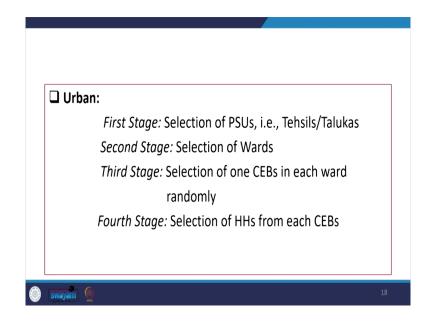
(Refer Slide Time: 35:37)



So, LASI also adopts three stages sampling for rural and four stages for urban areas. In the rural area first stage is primary sampling unit.

How to select the first PSU that is the sub district (Tehsil level) in the rural area. Once the Tehsil is determined then the second stage is to identify the village within the Tehsil then the third one is of course, your household within the village. So, the selection of household from the village is also important and that has been covered in three layers.

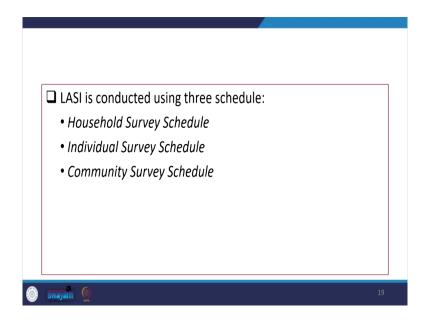
(Refer Slide Time: 36:27)



But in urban areas there are four layers (four stages): first is the PSU that is the urban talukas or tehsil urban areas then within the tehsils we all know that there are through the census enumeration blocks of census we can find out their selection of wards.

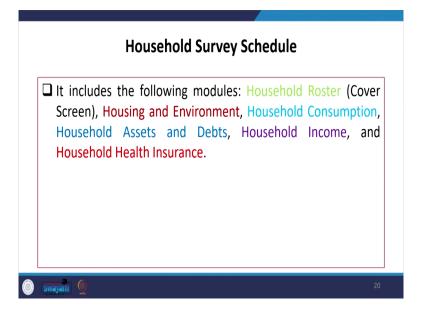
Then third stage is the CEB selection: one CEB census enumeration block in each ward randomly then the fourth stage is of household within the selected CEB census enumeration block.

(Refer Slide Time: 37:04)



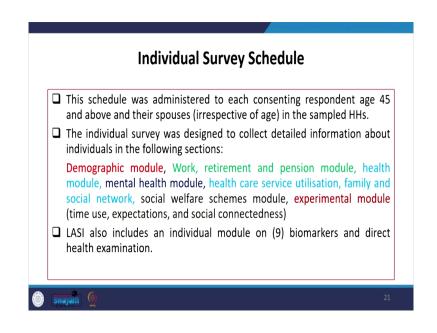
LASI is conducted using three schedule which I have already said first schedule on household survey, second one is individual survey and third one is on community survey schedule. Community survey schedule is very essential for further research.

(Refer Slide Time: 37:21)



Household survey schedule includes household roster that also gives information about the household and their environment, housing conditions, household consumption, household assets and debts, household income and household health insurance.

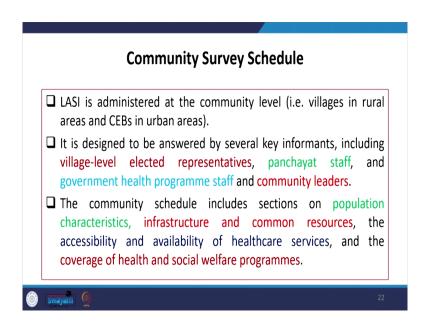
(Refer Slide Time: 37:42)



Then in the individual survey schedule they have the schedule which was administered to each consenting respondent of 45 years and above and their spouses irrespective of their age in the sampled household. The individual survey was designed to collect detailed information about individuals in the following section.

They covered in the individual schedule about demographic module, work retirement pension module, health module, mental health module, health care service utilization, family and social network, social welfare schemes module and experimental module also where they also included time use extent of time they put in and expectations and social connectedness. So, this section is very important. LASI also includes an individual module on biomarkers and their direct health examinations as well.

(Refer Slide Time: 38:41)



Last module it has included is on community health survey schedule this is administered at the community level of course, that is in the villages in rural areas and the census enumeration block in urban areas. This is in fact, designed to answer on several front like village level elected representatives, panchayat staff, government health program staffs and community leaders.

So, it's very finest aspects that is included for research. The community schedule includes sections on population characteristics, infrastructure and common resources, the accessibility and availability of health care services and the coverage of health and social welfare schemes are also important for analysis.

So, I think I have given you all the backdrop of the LASI data how to download and what are the important features, what are the coverage, who are included, who are not included, its age group different layers of research group, different variables included; I think these are quite important for you to start research on old age population. So, I will be guiding all other details in the next class with this I think it's time to close. So, I suggest you to please follow the details.

Thank you.