# An introduction to surveys and longitudinal data: cross sectional surveys

## Transcript

I am going to be talking today about cross-sectional surveys at the UK Data Service which tell us something about mental health. I’m going to talk a little bit about the variety of services that are available, talk about some of the analysis issues, present some examples of how they've been used and then say a little bit at the end about how you can access them.

Okay, so in the UK Data Service, which we love, there are many, many different cross-sectional surveys. Now there are a series of general health surveys which are really useful if you want to look at sort of nonspecific general mental health problems.

So, for example, the Health Survey for England, which been running since the early 1990s, includes the General Health Questionnaire, which is a very useful measure.

Then it is really worth having a look at some of the attitudinal surveys. So, for example, British Social Attitudes and also the Scottish Social Attitudes Survey have included modules in various years that ask people questions about their attitudes towards people with different mental health conditions. They've also included modules which ask people about their attitudes towards therapy and treatment and things like that. Really interesting and worth a look.

Also, worth noting that many different surveys now include subjective wellbeing measures like the Warwick Edinburgh Measure and the ONS4. So, that opens up a whole range of different studies if you're interested in more general wellbeing.

Also have a look at something like the Crime Survey because they ask questions of victims, for example, victims of violence, about how their experiences have affected their mental health. So, there's also different ways in like that.

Now my focus today is particularly on the specialist mental health surveys. Those are surveys that are focused primarily on mental health. Now I'm particularly interested in the Psychiatric Morbidity Survey series, but there's also another series that's available in UKDS called the NHS Community Mental Health Survey and that has been running for 20 years or so now. That's another resource.

Okay, so what this chart shows us is some of the different surveys that form part of the Psychiatric Morbidity Survey series. It's a terrible name. It really should just be called the National Mental Health Surveys. The ones in green at the top refer to the adult surveys, APMS, the Adult Psychiatric Morbidity Survey, and the ones in the bottom refer to the child surveys. That's the Mental Health of Children and Young People Survey.

Now these are general population surveys. That is, they include people with mental health problems and those without. So, we can look at prevalence and we can make comparisons. They also are carried out in home and they span different ages. So, the surveys at the top are from age 16 upwards. The earlier surveys were a younger age range, but from 2007 onwards the whole adult age range and no upper age limit.

I’m going to say a little bit about the child surveys, but I'm not going to focus on them so much. Just to say that if you are interested in child mental health, the website to go to is NHS Digital have a really good page on the Mental Health of Children and Young People Survey. It's a survey series that's run by ONS and the National Centre for Social Research. It's funded by the Department of Health and Social Care. And it's really well managed and overseen by NHS Digital. So, it's a very close collaboration of a number of different organisations.

Academic leadership has throughout the whole survey series really been provided by Tamsin Ford and also Tamsin Newlove-Delgado. So, they know absolutely everything about the measures that are used and there's very detailed measures that are used. And I recommend you have a read of this data resource profile in the International Journal of Epidemiology.

Also, to flag about the child surveys is that although the main face to face surveys were carried out 99, 2004, 2017, there have been a series of COVID follow-up surveys have been carried out online and these have provided really, really useful insight into the pandemic and school closures in relation to child mental health.

Now my focus today is on the Adult Psychiatric Morbidity Survey series. Again, this is also funded by the Department of Health and Social Care and is commissioned, overseen by NHS Digital, now NHS England.

The first one was carried out in 1993 when the government realised we know very little about the prevalent inequalities treatment gap in relation to mental health, and the design of this study was set up and the methods have remained remarkably consistent over time. Because the survey’s been carried out every seven years or so, we can really look at change over time. It's repeated cross-sectional surveys, so whilst the longitudinal studies you'll be hearing about in the other sessions today are really good for looking at causal inference, at individual trajectories, these are the surveys that are really good if you want to look at change at the population level and at temporal trends.

It's a household survey. Interview takes place in people's homes. Medium sort of size around 7,000 to 8,000. The interview is really, really detailed. It takes on average about an hour and a half, but it can go up to three hours. It's a two phase interview, so everybody's interviewed in the first wave but a subsample are followed up with very detailed different sorts of questions where the trained research interviewers are able to make judgements that are more clinically informed.

It also assesses certain mental health conditions to diagnostic criteria. Diagnostic criteria are used by clinicians if they need to decide whether or not somebody meets the threshold for a particular mental health condition. That takes account of the range of symptoms that they're presenting with, the severity, the frequency, the impact that it has on their life.

Now this means that the assessments used are far more detailed. So, whereas on a lot of surveys will have perhaps a 12 or 10 item screening tool, the number of the conditions that are assessed on the APMS series go into much more detail. So, for example, when we're looking at anxiety disorders and depression, we use information collecting the clinical interview schedule revised. It comprises well over 130 questions. So, it's much more detailed.

It allows us to look at the presence of 14 different types of symptoms. So, when you're analysing this data, you might decide you're just interested in one of those symptoms and you could home in specifically on that.

We then apply algorithms to the data using those diagnostic manuals, and that enables us to identify up to six different types of common mental disorders. So, a generalised anxiety disorder being a very common one, as well as depression.

We are able to look at any common mental disorder being present, but we're also able to look at a general severity score. So, that is the CIS-R score. We can look at it as a dimensional measure. You can also look at it by applying thresholds as well.

So, there's many different ways that analytically you can use this measure.

Now APMS data, as I said, isn't so good for causality, but it is good if you want to look at some of these types of research questions.

So, first of all, population prevalence. What does it mean when we talk about the proportion of people with a mental health problem? So, this paper, I really like this paper because it's incredibly simple analysis, but it was kind of trying to highlight what it is that we're talking about when we talk about, for example, one in four of the population, which is a commonly used term, has a mental health problem.

Now on the survey, we look at some very specific severe mental illnesses. For example, using the survey, we're able to identify that about one in a hundred in the population experienced a psychotic episode in the past year. So, that would suggest a particular sort of rarity.

Now if we expand that definition to include those with a particularly high score in the CIS-R threshold, so that's scoring 18 or more, that's a group that we would really expect to need intervention of some sort, being talking therapy or possibly some sort of psychotropic medication. We would expect around one in 10 in the population to be having a mental health problem.

Now if we expand that further to include those with milder symptoms, symptoms that may well be troubling for them but perhaps wouldn't necessarily warrant intervention, then we start to approach one in five. If we add in, for example, those you've got dependency, substance dependence disorders, a bigger group still. And then the survey also allows us to identify people who may well have different sorts of neurodevelopmental disorders, so attention deficit hyperactivity disorder. We've got a short screening questionnaire for that as well. We can certainly reach one in four of the population.

Another approach that we've used with the data is to take a latent class analysis approach that allows us to produce a population segmentation that looks at how these different conditions might overlap, identifies those who have multiple, those who have a few, and that's another approach which you can apply using this particular data set.

Now this is possible because the survey covers such a wide range of different types of mental health problems, symptoms and conditions. So, you could look at different aspects of substance dependence, you can look at very specific and quite rare disorders such as eating disorders and things like problem gambling as well.

So, population prevalence is one thing you can look at. You can also look at, because we have this repeated cross-sectional survey series, you can look at temporal trends, a change over time. Now, in this particular paper we used data from a number of the surveys to look at trends over time in self-harming behaviours.

Now this is quite a useful example for us to talk about in terms of what you need to consider when you're planning an analysis of this type. First thing to do is when you've downloaded the data from the UK Data Service website is to download the questionnaire and to download the information that is provided with it. This documentation is really key because it allows you to look at the exact question wording that was used each time. So, you need to make sure that the question wording is consistent before you start to produce temporal trends.

Another thing you can do is make sure that you are looking at the same people each time. So, in 1993 the questions about self-harming was filtered and it was only asked of those reporting depressive ideas. Therefore, we had to exclude that data when we're looking at temporal trends because we didn't have a comparable sample for the analysis that we wanted to do.

The other thing is to look at the mode. So, whilst on APMS we've used a consistent methodology over time. Still there are subtle changes. So, when looking at temporal trends, I limited the analysis to where it was asked face to face because it had been asked face to face in every survey. But in more recent surveys, it was also asked about in the self-completion. But I had to exclude that data when looking at temporal trends. So, consider that you've got a consistent mode of data collection when you're constructing temporal trends like that.

Geographical coverage also had to be considered because the earlier waves in the survey series were all of Britain, whereas the later waves in the survey were England only. So, when doing temporal trends, I had to restrict the earlier waves just to the sample that was living in England at the time that the data was collected.

Likewise, age group has changed over time. The earliest survey only went up to age 64, so when looking at temporal trends, I had to limit it to a comparable age group over time.

So, using this approach, we're able to see that in fact the prevalence of self-harming behaviour has increased over time. Chart A here is men, Chart B here is women. We see there's been an increase in both groups, but the proportion of women reporting that they were using self-harming behaviours in order as a coping mechanism, that had changed particularly steeply over time.

Okay, another thing that the survey series is really, really good for is looking at the treatment gap. So, if we rely on data from those in contact with health services, then it doesn't tell us anything about those people not in contact with health services. Sorry, that's my cat coming in.

Okay. So, when we were looking at the data, we could see that of those people who reported having self-harmed, around half said that they had not received either medical or psychological support for that. And we saw no evidence of that gap improving over time. And in fact, certain groups were particularly unlikely to get either medical psychological support after self-harm, and that's people who are younger, it was men and it was those in debt. So, there's a real sense of socioeconomic and demographic inequalities in treatment access. And the APMS data is ideal for looking at that.

It's also really good if you want to do all sorts of subgroup trends, comparisons, looking at inequalities and looking at the context and circumstances for people experiencing difficulties of that kind. So, this is a really lovely paper I highly recommend. Produced by Gargie Ahmad, who is based at King’s. And she used data from the 2007 and the 2014 surveys to look at inequalities and treatment access by ethnicity. However, she was really held back by the fact that the surveys so far have had a relatively small number of participants from different ethnic groups, simply because they were representative samples. But in the next survey, which is currently in field, this is the 2023 survey, it's going to be a full year of data collection. Then there'll be some processing of the data, but that data should be available come 2025.

And it's going to be an absolute corker of a survey. It has for the first time in its series a really substantial boost of different ethnic groups. This has been funded by the Department of Health, it's been hard won. So, I'm very, very grateful to the Department of Health for making this happen.

There's going to be a translation into Urdu that's available for those participants who need it.

So, the next survey really will be much better placed to look at variation by ethnicity and you'll be able to use the data to explore variation by ethnicity and a whole range of different outcomes.

These are some of the characteristics that are studied on the survey. Note also that as well as an ethnic boost, there will also oversampling of areas in more deprived neighbourhoods, which means that we're really going to be much better placed for exploring things like the impact of debt, context of poverty, housing conditions and other socioeconomic circumstances.

Just to flag a few uses of the data, just for examples of different sorts of outputs that are possible, this is a report that looked at those with intellectual impairment. That's another group that can be identified on the survey with the measures used. And it really showed a very steep gradient whereby those who had the greatest intellectual impairments are also the most likely to be experiencing depression and anxiety.

This is another analysis done which really highlights the situation for people with recent experience of domestic violence, particularly flagging up the very high rates of suicidality and self-harm in this group.

And finally, this paper as well shows how just having one question can open up a whole analysis. So, there was one question that was added in the 2014 survey about whether or not people had ever been in prison. And then this paper used analysis that that was centred around that variable that just compared the experiences of this group and their childhood antecedents and their life subsequent to that.

So, you need to access this data. It's a wonderful data source. This paper by Tamsin Ford and many others highlights that accessing mental health data of all kinds is an opportunity, but also a challenge.

Just to say that data that was collected up to about 2010 is very, very straightforward to get hold of via the UK Data Service. So, what you need to do is you go to the UK Data Service website, you log in. If you don't yet have a registration, you can register. You need to have information about your university or your academic institution to do so. Then you can search for different, it's very useful if you know the name of the data set you're looking for, but there's also keywords you can search on. And you'll be able to find your data set in that way. There's different access conditions for different surveys, different classes of use. Increasingly, there will be other and additional criteria, conditions to you accessing that data, but once you've found the data set that you're interested in, you'll come to a page like this. Highly recommend that you look at the documentation page. There's reports and lists of variables and other key information, and you can download the data set from there.

For the more recent surveys, you need to have permission from what was NHS Digital, now NHS England, the data access request service there.

Now what you need to do first is go to this page, which is the data access request service site. You will need to register with them if you've not already done so. Then you open up a form. It's quite a detailed form. You need to have information about your organisation and how you can securely hold the data. You will need to have information about what you need to do with the data and very clearly be able to explain how your analyses can lead to patient benefit and improvement in services. There are increasingly new ways being offered by NHS Digital about being able to access the data that they hold remotely, so there may well be other information about that.

Sometimes the system looks like you're going to be charged to access survey data, but our understanding is that there are no plans to introduce charges for accessing survey data. So, let us know if you have any difficulties with that.

Be prepared that this process could take quite a long time. So, it can take at a minimum several months.

This is the last slide, really just to flag up the very, very wide range of different topics that are covered on the survey. It's ideal for really detailed exploration of different types of mental health conditions that allow you to look at severity in different thresholds, but also provides insight on a wide range of different subject areas known to be associated with poor mental health outcomes. So, you might be interested, for example, in carers or in parenting or in sensory impairment or any of a wide range of other topics.

National Centre for Research Methods (NCRM)  
Social Sciences  
Murray Building (Bldg 58)  
University of Southampton  
Southampton SO17 1BJ  
United Kingdom

**Web** www.ncrm.ac.uk   
**Email** info@ncrm.ac.uk  
**Tel** +44 23 8059 4539  
**Twitter** @NCRMUK