Person-centred intelligence: Interpreting and using data

The Strategy Unit and Ipsos MORI

About this guide

- This guide forms part of the Strategy Unit and Ipsos MORI's series about person-centred intelligence.
- It focuses on the steps from once the data have been compiled, through to interpreting and using it to inform decisions.
- Having worked through this guide, readers should be able to judge the quality of their data, analyse it by referring back to the key purpose of the project, interpret it to inform decisions and present it to a range of stakeholders.

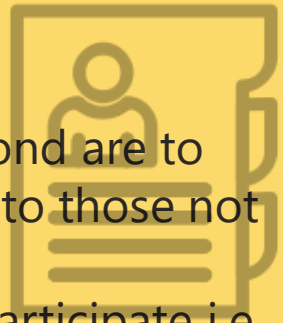
Contents

Assessing the quality of the data

How to assess data quality (Part I)

Assessing the quality of data helps to determine **how reliable it is as the basis of decision-making**. Consider the following elements:

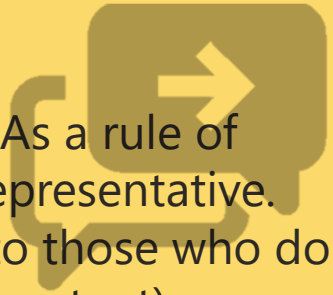
Profile of participants



- A key element is understanding how similar those who did not respond are to those who did respond. Where those responding are quite different to those not responding, there is bias within the results.
- To analyse this, compare the profiles of those who did and did not participate i.e. compare the age and gender profile from your participants with the age and gender profile of the entire population.
- Data permitting, consider the factors that are more likely to impact on participants' responses and compare on those (e.g. if you are measuring use of online services this will have a strong correlation with age).
- There is an option to weight the survey data (i.e. adjust it to match the entire population).

How to assess data quality (Part II)

Response rate



- Response rates are generally used as an indicator of survey quality. As a rule of thumb, surveys with higher response rates often tend to be more representative.
- However, this does depend on how similar those who respond are to those who do not respond (i.e. if they are similar, then the response rate is less important).
- Calculate the response rate by dividing the number of responses by the eligible population invited to provide data.
- A rate of around 50% or higher is typically regarded as a strong response rate, while 20% or lower is a weak response rate.

How to assess data quality (Part III)

Number of responses

- The absolute number of responses is also important in determining how much to trust the data.
- The more people from a population who are included, the more accurate the results.
- Each data point is surrounded by confidence intervals – e.g. a result of 50% could be anything from 45% to 55% if a census had been conducted – and the size of this confidence interval is determined by the number of responses.

Ability to respond

- Check answers for each question – if a large proportion (generally over ~10%) answer with 'don't know' or 'prefer not to say', it may be they did not understand the question and results should be treated with caution.

Tips on analysing data

The Strategy Unit.

Ipsos

Ipsos MORI

Overall approach to analysing the data

After assessing the data, it can be analysed and interpreted. Throughout analysis and interpretation, the golden rule is to **remain focused on answering the key questions**, and being able to use the data in the ways intended.

If you have developed a logic model, then this can form the basis of your analysis (see . This might mean not looking at every piece of data but being more focused.

Develop an analysis plan – the key points of analysis can be identified before even receiving the data, using the ideas on the following slides.

It can be daunting to have a dataset in front of you, but if you focus on your key questions and use the data to explore these (rather than looking at everything), you can be more streamlined in your approach.

Tips for analysing the data (Part I)

Look at scaled questions

- Scaled questions (e.g. 1-10 or very poor to very good) can be aggregated to illustrate an overall picture.
- However, in cases where the overwhelming majority of responses are either overwhelming positive or negative, it may be necessary to look at each point on the scale to differentiate between them.
- Looking at both ends of the scale is also important.

Carefully select which areas to further explore

- When choosing which areas needs additional research, note that small percentage differences (i.e. of 5 percentage points) between results does not always suggest distinct differences – as statistical reliability may mean that they are effectively equal.

Compare the responses of two different questions

- It may be useful to compare responses to different questions to understand why they have answered this way.
- For example, in a hospital staff survey comparing a question about how involved they feel they are in decision making and how positive/negative they feel about the quality of patient care provided may provide a better overall picture.

Tips for analysing the data (Part II)

Sub-group analysis

- Sub-groups are groups within the overall population such as specific age groups for patients, caring responsibilities for carers, or department for staff.
- Sub-group analysis will help to see whether certain groups of participants have different experiences or outcomes than others.

Tracking change

- If there are multiple data points over time, looking into how the findings have changed over time will help to analyse the direction of travel.

Be cautious of statistical reliability

- When analysing by sub-group or tracking change between data collections, note that results have to differ by a certain amount to be statistically significant (see end of pack).

Look for common patterns in the data

- During the analysis, do not look at questions in isolation, but instead aim to build a picture about the data by looking for common patterns.
- For example, if analysing a series of statements for sub-group differences, see if certain sub-groups have consistently more positive or negative outcomes or experiences. This is more powerful than having a list of differences for each statement.

Tips for interpreting the data

Tips for interpreting the data

Triangulate the data	While defining your vision and purpose, you will hopefully have identified a range of different data available to answer your key questions . Once you have analysed the new data, triangulate it against what else is available to build a wider picture.
Possible hypotheses	Look for logical hypotheses that might explain the data. It may be worth generating multiple hypotheses that can then be further tested with this data, other data that exists, and then discussed and interrogated.
Internal and external discussions	Disseminate and discuss the data so you can build a narrative around what is happening, and the reasons underpinning that. Many different people can input to this process including: frontline staff; patient, carer and staff representatives; patients and carers; those from outside your organisation who you work closely with (for example, others within the ICS who work with this population). This also helps build ownership for the data.
Additional qualitative research	Linked to the above, another option is to undertake further qualitative research with the population. This could be done fairly easy by asking those providing data if it would be possible to re-contact them to discuss in more detail and collecting contact details. In-depth discursive interviews or discussion groups can unpick the quantitative findings and the reasons underlying them in more detail, and can be useful to further develop action plans.
Co-design	Similarly, but building upon this, patients, carers and staff can be involved in co-designing action plans to address the findings of the data collection.

Tips for presenting the data

Tips for presenting the data (Part I)

Close the feedback loop

- People need to know what the findings and what is being done to address the findings.
- This is for those who participated and anyone involved in gathering the data, to demonstrate the impact of their effort.

Consider the audience

- Different audiences may require different things from the data and will have differing levels of data literacy. Therefore, ensure that each output produced is tailored to their needs. This could involve asking about what it is they require, or testing it with a few individuals before wider circulation.

Focus on the narrative

- Focus on the narrative and include data that supports it, instead of using all of the data.
- Make sure that the report highlights the narrative and talks people through it, including headings, rather than simply giving the data or question wording.
- Think about the best way of presenting the narrative – the Cambridge Centre for Health Services advocate using patient stories to illustrate the narrative, as people will engage with them better than data.

Tips for presenting the data (Part II)

Use the principles of good data visualisation

- Keep pages or slides 'clean' by not cluttering them up with too much information.
- Present it as simply as possible, select the best chart for each measure.
- Use headings or narrative to tell people what you want them to take from the data.
- Use colours and different size fonts to draw people's attention to where you want it to go.
- Where the data allows it, aggregate up to simplify the message – for example, aggregate a score of 9 and 10 rather than giving the results for 9 and the results for 10.

Provide methodological details:

- This helps illustrate the data's quality and reliability, and provides context when it is interpreted.
- Useful information includes: the purpose of data collection; target population; profile of participants; methodology; when the data were collected; the number of responses; and response rate.

Have the more detailed data available

- Some people might want to look at the data in more detail, so it is useful to have it available.

How sample size affects reporting of data

If reporting on a question asked of **fewer than 30 people**, report it as the number of people, rather than a percentage and include a footnote urging extreme caution in interpretation. In general, avoid reporting on these small sizes wherever possible.

If reporting on a question asked of **30 to 100 people**, include a footnote urging caution in interpretation. In general, if possible only report on these small sizes where you have confidence in the findings (e.g. through triangulation).

Ensure that any of the findings being reported on **do not identify individual participants**, unless they have agreed that they are happy to be identified. This means, for example, making sure that any sub-group analysis is only looked at where there are a minimum of 10 participants, and checking any quotes used from open ended questions.

Statistical reliability

Statistical reliability (Part I)

- When participants in a data collection exercise represent only a sample of the total population of interest, we cannot be certain that the results of a question are exactly the same as if everybody within that population had taken part ("true values").
- However, we can predict the variation between the results of a question and the true value by using the size of the sample on which results are based and the number of times a particular answer is given. The confidence with which we make this prediction is usually chosen to be 95% – that is, the chances are 95 in 100 that the true value will fall within a specified range (the "95% confidence interval").
- The table below gives examples of the confidence intervals for different numbers of responses.
- Strictly speaking, the tolerances applied here apply only to random samples.

Number of responses	Approximate confidence intervals for percentages at or near these levels (expressed in percentage points)		
	Level 1: ***10% or 90%***	***Level 2:*** ***30% or 70%***	***Level 3:*** ***50%***
	+/-	**+/-**	**+/-**
100	**6**	**9**	**10**
500	**3**	**4**	**4**
1,000	**2**	**3**	**3**

- For example, taking a sample where 500 responded and where 30% gave a specific answer, there is a 95% likelihood that the true value (which would have been obtained if the whole population had been interviewed) will fall within the range of +/-4 percentage points from that question's result (i.e. between 26% and 34%).

Statistical reliability (Part II)

- Different groups within a sample (e.g. different staff groups, or patients on different wards) may have different results for the same question. A difference has to be of a certain size in order to be statistically significant.
- To test if a difference in results between two sub-groups within a sample is statistically significant, at the 95% confidence level, the differences between the two results must be greater than the values provided in the table below.
- Similarly, if tracking changes between two different years of a survey the difference needs to be of a certain size to be statistically significant.
- Again, strictly speaking the sampling tolerances shown here apply only to random samples.

Number of responses	**Approximate confidence intervals for percentages at or near these levels (expressed in percentage points)**		
	Level 1: *10% or 90%*	*Level 2:* *30% or 70%*	*Level 3:* *50%*
	+/-	+/-	+/-
100	8	13	14
500	4	6	6
1,000	3	4	4

- This means, for example, that if surveying 100 patients in 2019 and 100 patients GPs in 2020, the difference in results between the two surveys for a question where c.50% give a particular answer must be 14 percentage points to be statistically significant.

Ready to move on?

When analysing, interpreting and presenting:

- Assess the quality of the data by looking at the response rate, profile of participants, number of responses and whether participants have been able to answer all questions.
- Moving on to analysing the findings, refer back to the original vision and purpose and the logic model (if relevant) to develop an analysis plan. This will help to focus the analysis.
- Analysis options include looking at each question individually, drawing findings across questions, considering how different groups in the sample compare, change over time and common patterns.
- It is important to also check statistical significance of these findings.
- To interpret the data, triangulate it with other relevant information, look for logical explanations for the findings, discuss it internally and externally and consider undertaking more detailed qualitative analysis.
- When reporting and presenting the data, consider the audience and their needs. Ensure methodology details are covered, focus on the narrative and the principles of good data visualisation. Also ensure that more detailed data is available for those who want it.