Data analysis principles

A guide for all our analyses

At Impact on Urban Health, we have grown as a team, and we frequently work with partners and temporary consultants. It is important to work with a common definition of what responsible use of data looks like for us and people who work with us.

This is not a legal document but rather a framework to inform decisions on data analysis projects, their design, and the use we make of the insights we produce. To this aim, we kept the principles short and relevant for the type of work we do at Impact on Urban Health.

11 principles to use data responsibly

Acknowledging that we are not the first team or discipline to raise ethical questions, we got inspiration from existing sets of principles from science and medicine. We selected the ones that were most relevant for our work and amended them to reflect what mattered most for us:

1. We clearly and accurately present our analyses and do not mislead or allow misinterpretation of our work. We also ensure the insights are used responsibly.
2. Acknowledging that a data product is only as good as the data underlying it, we clearly document the source and context of a data set, including the quality, representativeness, missingness, biases, timeframe and overall limitations or caveats of the data.
3. We develop projects and undertake analyses in line with our data ethics guidelines as well as require the same of our partners.
4. We understand that data are not just numbers but represents people and **individuals'** situations. We acknowledge that data science can have unintended consequences, including exacerbating existing inequalities, and do everything we can to avoid doing harm.
5. We ensure there is a clear benefit to our work, including adding value to our programmatic work, advancing diversity, equity and inclusion, and reducing economic, social and health inequalities.
6. We continue to advance our data skills and promote personal development of others. We are **not afraid to say we don't know and we ask for support when** needed.
7. We act with caution when using data including checking that we have the right to use, analyse the data and/or share the data.

8. We make sure that data collected is anonymised so no individual can be identified and we suppress figures that would allow reverse engineering or re-identification. We put privacy at the forefront of all work we do.

9. We are transparent with our community and partners about the data we use including making them aware and/or involving them in project design especially when data is collected on them.

10. We are responsible for our own work and the respect of quality assurance processes; and if mistakes are made, we report and correct them immediately.

11. We ensure the security of data where required, especially with the use of personal data or patient level data.