

## [NHS Data Opt-outs: the missing pieces of the puzzle | Understanding patient data](#)

**10 Dec 2024**

NHS Data Opt-outs: the missing pieces of the puzzle

*In this guest blog, John Tazare (Assistant Professor) and Helen Strongman (Associate Professor) from the [Electronic Health Records Research Group](#) at the London School of Hygiene & Tropical Medicine discuss their perspectives on potential consequences of data opt-outs for health data research.*

When you interact with the health service – for instance, visiting your local general practitioner (GP) practice or hospital – relevant information about you and your health arising from these interactions is recorded electronically by healthcare professionals and stored in electronic health records.

Information from your records and the records of others, when brought together at a large scale, provide a source of data that is really useful for research and service improvement. In the vast majority of cases, [this data is de-identified, meaning details that could be used to identify you have been disguised or removed.](#)

For over 20 years, health record data has meaningfully contributed to important public health priorities. For example, key evidence generated from this data has helped establish the [safety of measles mumps rubella \(MMR\) vaccination](#), [measured long-term risks of heart disease in cancer survivors](#) and more recently, [helped identify groups of people at highest-risk of severe COVID-19.](#)

In our roles at the London School of Hygiene & Tropical Medicine, we use this data to study patterns of disease and the use of medications, with the aim of improving current and future patient care.

In order to achieve the benefits of analysing patient data, building and maintaining public trust is key. One mechanism designed to increase trust is the ability to ‘opt out’, preventing use of your confidential patient information for research or planning if you haven’t consented. In England, this is done under a scheme, called the National Data Opt-Out, wherein patients can opt out of their data being used for research and planning purposes, with some exceptions. This does not apply when anonymised data or high-level statistics are used.

In 2021, following concerns surrounding a government initiative – General Practice Data for Planning and Research (GPDPR) - to create a centralised national database for GP data, there was a sharp increase in the level of data opt-outs. In the following sections, we describe the patterns of opt-outs across this period by key demographics and highlight potential consequences for research using this data.

### **Patterns of people registering opt-outs**

The GPDPR initiative was announced in May 2021, with a deadline to opt out by 23<sup>rd</sup> June for those who did not want their data to be shared. In the month leading up to the national opt-out deadline, over 1.3 million people in England opted out, increasing the overall opt-out rate from 2.77% to 4.97% of the population. However, this rise was not constant by age, sex and region:

- Among females, the opt-out rate grew by 83%, from 3.02% to 5.53%, while for males, it increased by 76%, from 2.51% to 4.41%.

- Across ages, the most significant rise occurred in people aged 40–49, where the rate more than doubled from 2.89% to 6.04%. Rises of a similar size were seen across people aged 30-60, with lower increases amongst those aged <30 and >70.
- By the end of June 2021, people aged 30–39 had the highest opt-out rate at 6.37%. Younger children (1.38%) and adults aged 90 and over (2.77%) had much lower rates.
- Opt-out rates varied across England, with some regional grouping of high opt-out rates around the Greater London and North West areas, but no strong patterns across the country.
- There was no clear correlation with living in less or more deprived areas and opt-out rates.

### **Consequences of opt-outs for health research**

This increase in the level of opt-outs has two key consequences for health research using this data.

Firstly, opt-outs reduce the number of people we can study for a given research question. In health research, researchers are often trying to understand patterns that apply to a large group of people, for example sometimes an entire country or an entire group of people with the same condition. Depending on the type of research, it may not be logistically possible to study the entire population, so researchers focus on a sample - a group of people taken from a larger population meant to be representative of this underlying large group of interest. Generally, the bigger the sample, the higher the confidence we have in our research results. This applies to health data research too. If only 5% of people opt out, the sample may still be very large (if the underlying group of interest is large), but the larger the amount of people that opt out, the bigger the reduction in the size of the sample. Using smaller sample sizes reduces our confidence in the results and can ultimately limit our ability to draw meaningful conclusions that improve the management of medications and disease.

Secondly, opt-outs can ‘distort’ the findings of a given study. As we have seen, levels of opt-outs are not consistent across demographic groups. In fact, there are likely further key differences between those who do and don’t opt-out, but the limited amount of demographic information available in opt-out statistics means we cannot be sure what those other differences are. For instance, some research indicates that some ethnic minorities are likely to opt out at a higher rate, but it is hard to know to what extent this is the case without including this information in official statistics. As researchers, we have no way of knowing how the available data (on people who have not opted out) compares to the whole group of people we wish to study. Importantly, this can lead us to generate answers that are not necessarily relevant to the whole group. Because important policy decisions about the health and care system are often based on academic research, this means policymakers may potentially make inappropriate decisions based on these findings.

To illustrate how these issues play out in a given study, we can consider the analogy of a large, say 500-piece, jigsaw puzzle of some complex landscape. The pieces represent people included in the study and the completed puzzle represents the answer to a particular research question.

First, let us imagine we randomly take out a number of pieces from the completed puzzle. If we only take a small number of pieces out, the picture is likely to still be clear. However, the more pieces we remove the more obscured (or uncertain) the picture is.

Second, instead of randomly removing pieces, let's suppose we take out all the pieces from the top-right section of the puzzle. We can describe the picture as best we can from the pieces that remain and maybe, given what we see, we can assume something about the missing section. However, without these missing pieces, we cannot be sure this is sensible or what the complete picture looks like. In reality, we likely do not even know what part of the puzzle is missing - instead of the top right corner, it might be the bottom left - because we don't know who has opted out of the dataset we're studying.

## **Conclusion**

The level of opt-outs rose substantially across the consultation period in 2021. This rise was not consistent across all demographic groups and regions, and its full impact on health research is hard to measure. Importantly, it is not possible to measure or reduce this impact adequately using data available to researchers. Careful consideration of the potential effects of opt-outs on research and health policies is a key piece of the puzzle when creating future national data-sharing initiatives.

To learn more about the opt-out and your choices, see Understanding Patient Data's explainer [here](#).

*This guest blog is based on an Open Access article, available via the link below:*

*NHS national data opt-outs: trends and potential consequences for health data research. John Tazare, Alasdair D Henderson, Jessica Morley, Helen A Blake, Helen I McDonald, Elizabeth J Williamson, Helen Strongman. BJGP Open 9 July 2024; BJGPO.2024.0020. DOI: [doi.org/10.3399/BJGPO.2024.0020](https://doi.org/10.3399/BJGPO.2024.0020)*