

Available next steps in implementing and delivering data usage reports: NHS Digital *can* offer a customised PDF to every patient, using only existing tools

Advances in 2021 make personalised data usage reports increasingly *technically* simple to deliver. As ever, the blockers remain largely political – grounded in an assumption that NHS England has an implacable objection to patients knowing how data about them is used.

Data is already in PowerBI, and PowerBI can output a PDF

NHS Digital has now launched the [web interface to its data use register](#); a clickable version of the big spreadsheet and internal database, all driven by the tool PowerBI. As well as the current spreadsheet and the web-based interface, PowerBI is equally capable of outputting a PDF.

NHS Digital could *today* provide individuals a human-readable, well formatted PDF of the information it already publishes, which says where *any* data went in the last period, showing where *their* data went. Something akin to the demo we produced in 2012 and updated for 2021, using the existing PowerBI system.

NHS Login is now widely used

The NHS Login, which underlies many NHS.UK services and the NHS App, has come into its own in 2021. This being the case, there is now a way for a verified user to log in to any digital NHS service with relative ease.

Integrating NHS Login with the existing PowerBI capability would allow a patient to be able to see where *their* data went, and how decisions were made by NHS Digital.

NHS Digital already has these information systems, it simply needs to join two services together – this should be straightforward as both services are *designed* to be glued together with others.

Include non-NHS Digital-owned Data Environments

NHS Digital's data use register includes all projects that went through the DARS process. Similarly, ONS only [publishes a list](#) of those projects that entered ONS SRS (its TRE) via the DEA accreditation process. Projects using NHS data in the ONS SRS on any basis other than Digital Economy Act powers are therefore invisible to the public.

ONS has a list of all projects that use NHS data in the ONS SRS,¹ and NHS Digital has contractual powers to require ONS to provide a list of projects which accessed NHS data. NHS Digital should therefore use these, and integrate the information into its existing records and infrastructure.

The use of multiple TREs is going to require this work anyway. Information from the long-running TRE at Genomics England should also be integrated, as part of the integration of GeL into mainstream NHS practice. And there is another data platform in use in the NHS whose providers *advertise* how well it interoperates: **Palantir**.

¹ Separately, ONS *must* commence a full accounting of *all* projects in its SRS to avoid secrecy undermining public confidence, especially as the IDP / IDS progresses.

Use of data in Palantir *can* and should be included

NHS England is very proud of its installation of Palantir's Foundry. [Palantir's own PR](#) claims Foundry has "full stack interoperability", offering "unfettered ownership of your data". On [GOV.UK's Digital Marketplace](#) Palantir states its platform can "integrate data of any scale, format or structure", tracing and auditing every access and analysis – unless configured not to.

While Palantir says it keeps detailed records, NHS England is much more reluctant to say what it uses the system for. In fact, in its own data use register, NHSE has avowed using it for precisely... nothing.

We would have happily included the Palantir activity feed into our new mock-up of data usage reports – and Palantir's APIs would make it trivial to do so – but the barriers here are not technical. The only difficult step is NHS England being willing to let go of its comfort blanket of secrecy. (Ideally, before the takeover of NHS Digital takes it away.)

Palantir claims that "by tightly integrating the data governance process into the underlying [access control](#) system", its platform promotes integrity of decision making – which will become a boon for the steady flow of legal cases and Judicial Reviews that NHS England faces as officials take decisions which are not obviously in the best interests of some patients. It could prove additionally challenging if their sensitive patient records – which NHS England wants to hold after having taken over NHS Digital – were to be used against them, say, by the part of NHS England that makes decisions about locations of services.

An expectation that research may use all health data might also be a catalyst for the creation of PROMS for diseases which lack widespread research infrastructure, e.g. post-viral fatigue (both COVID and other).

NHSE likes dashboards and the ability to cut and paste pretty graphs into reports – hardly demanding use cases for such technology. And the migration of councils into NHS Digital's "system access" for HES shows such uses don't require Palantir, or the more controversial activities of NHS England.

Include CPRD to begin some communications around GP data

When NHS E/x/D made claims that some GP data was already used the way they wished GDPR to operate, they were not entirely wrong. Patients knew nothing about it, but CPRD has operated in the shadows for decades. It is time for CPRD to come out of those shadows, ideally by choice.²

The introduction of personalised data usage reports is an opportunity for both CPRD and the MHRA (which runs it) to make the case to the patients whose data they make available why that data use is valuable – a case they have chosen never to make. Inaction is likely to become increasingly untenable. The pandemic GP dataset, for pandemic purposes only, offers a similar opportunity – and questions will and must be raised if one is included and the other not.

² As an aside, CPRD should become the centre of expertise in GP data within the statutory safe haven.

Should a future 'care.data 3' wish to make claims about data use, the best way to evidence those claims is to show *what GP data is used for today*. In practice, that largely means CPRD. The MHRA says it keeps records, and it publishes some of them to the CPRD website – so those are available in structured form, if the MHRA and NHS Digital can come to an agreement about process.

Of course, the MHRA does not itself know all of the data projects of its most high-paying recipients, who are allowed to pay more to avoid scrutiny.

Include what we learn from data driven research

Research can benefit a patient whether or not they were involved in it. And it is to the advantage of all research that patients have a better understanding of all research.

While patients should be able to read the outputs and lay summaries of research projects in which their data was included, they should also be able to read the outputs of all *other* projects that used NHS patients' data. The cancer registry has done some of this work for years, and this knowledge can and should be included in both the PowerBI interface and the PDF for patients.

Include Surveys and Longitudinal Studies too

There are some administrative longitudinal studies in which participation is secret, i.e. you have no way of knowing if you are included or not. While this raises ethical questions in several areas, some of those questions can be mitigated by replicating the approach taken to research outputs – offering full transparency on those surveys and studies to *everyone*, irrespective of whether that person was included in them.

Available shortly: next steps for civil society

If NHS England is successful in taking over NHS Digital and abolishing the "statutory" and "safe" parts of the statutory safe haven, some of these steps will become easier. And some will become harder.

Much of the structured information we refer to above is already published online in some form, and could be scraped into structured form without requiring the approval of gatekeepers who do not wish patients to understand how data about them is used.

With access to Palantir's logs, whether via its dashboards or an API, this becomes even more trivial to commence. Given its claims about governance, Palantir would probably even market its ability to do this, incorporating all of the above data sources, as a feature it will sell to NHS England.