



---

# **Evaluation of Monthly Prescribing Open Data publication**

**Feedback, impact and  
webstats**

**ISD Prescribing**

---

## DOCUMENT CONTROL SHEET:

### Key Information:

<b>Title:</b>	<i>Evaluation of Monthly Prescribing Open Data publication</i>
<b>Date Published/Issued:</b>	<i>14 March 2017</i>
<b>Version/Issue Number:</b>	<i>1.0</i>
<b>Document Type:</b>	<i>Evaluation</i>
<b>Document status:</b>	<i>Pre-publication</i>
<b>Author:</b>	<i>Toby Stead</i>
<b>Owner:</b>	<i>ISD Prescribing</i>
<b>Approver:</b>	<i>Kathryn Neill</i>
<b>Approved by and Date:</b>	<i>Kathryn Neill, 14 February 2017</i>
<b>Contact:</b>	<i>Toby Stead, ISD Prescribing, 0131 314 1709</i>

## **About this document:**

Following the publication of ISD Prescribing's Monthly Prescribing Activity Files (our Open Data) on 12 April 2016, we have been collating and analysing information from a wide range of sources with the aim of improving the public value of the data release.

This document assesses the project's success at a six-month interval against 'public value'. Using customer and public feedback to guide our inquiry, we conducted further analysis on traffic to our website and demand for our services. Within the confines of the feedback we received, we were able to use these sources to assess the following domains of public value which are broadly derived from the properties by which data are considered 'open':

- Are the public able to find the data?
- Are the public able to use the data?

## **About Open Data**

Data are considered 'open' when they 'can be freely used, modified, and shared by anyone for any purpose' (see <http://opendefinition.org/>).

All data released by ISD into the public domain are technically 'open', but some data are more open than others, as described by a 5\* deployment scheme (see <http://5stardata.info/en/>). ISD Prescribing aimed to reach 3\* with its Monthly Prescribing Activity File, meaning that it is available in a structured, machine-readable, non-proprietary format.

The Monthly Prescribing Activity File itself presents volume, quantity and costs of prescriptions at drug and GP practice level. It is released every month, with the first release of three months' worth of data taking place on 12 April 2016.

## **About Information Services Division**

The Information Services Division (ISD) is a division of National Services Scotland, part of NHS Scotland. ISD provides health information, health intelligence, statistical services and advice that support the NHS in progressing quality improvement in health and care and facilitates robust planning and decision making.

As part of ISD's 'Transforming Information Programme', we have been re-examining the way we publish and present data and intelligence.

## Open Data feedback

### The feedback:

Feedback from the following three distinct sources was collected:

1. existing external users of ISD Prescribing data
2. internal ISD colleagues from a multitude of disciplines
3. external members of the Open Data community.

These audiences have distinct interests in and uses for the data and present an interesting triple perspective as we move from traditional to modern publication methods.

Some of the main themes include:

- Requests for additional fields
- The suppression of Quantity from the initial release
- Access to the file
- Discoverability of the file
- Reference files
- Backdated files.

The feedback on each will be summarised below, followed by a discussion of what this means in terms of this evaluation and any next steps for ISD Prescribing. Supplementary data, such as website traffic and social media engagement metrics, are used to provide extra context to the points raised.

### ‘The Quantity issue’

- Most common form of feedback
- Field described as ‘extremely useful’
- Omission came as a surprise to many customers.

‘Quantity’ is a measure describing, for example, the number of tablets of a drug prescribed (as opposed to the number of prescriptions for the drug). Although originally intended to be included in the Open Data, the field was suppressed late in the day due to a newly-discovered data quality issue. The high volume of feedback on this issue attests to the value of the Quantity field. Indeed, the field was described as ‘useful’ by three separate customers. These customers had been purchasing extracts containing the ‘Quantity’ field for varying lengths of time, usually several years, and reactions to the suppression included curiosity, confusion and surprise. The issue is described in detail on page 8 of the [FAQ document](#) published alongside the data files.

**Conclusion:** This feedback suggests that the ‘Quantity’ field, with eventual publication, is widely-used, but also that ISD Prescribing could have been better at communicating with its existing customers regarding the suppression of what is clearly a crucial field. With eventual release of the field in June 2016, we can be assured that our customers are using it.

### Access to data

- Feedback addressed the layout of the data, the size of the file, and the supplementary information provided
- Mixed feedback - both positive and negative
- Blank fields (eg missing data) requires careful explanation.

In roughly equal proportions, we received positive (eg “*the feedback I have got from our IT department is that the data is very clear*”) and negative (eg “*It seems that the data on the website is not user friendly (too large for our systems to handle)*”) feedback on access to the files.

We had anticipated that a large section of our current customers have been downloading and using corresponding HSCIC (now NHS Digital) data files for England which are much larger. Our files also meet the 3\* OGL (Open Government License) criteria.

We have taken steps to improve the accessibility of our files by producing an interactive visualisation which allows users to ask set questions of the data. This represents a two-tier approach to accessing the data – a pre-built interactive tool presenting a high-level summary and an in-depth, granular dataset presenting the full detail. These tiers are separated by a skills barrier.

With regards to the size of the files, we are investigating solutions that enable users to pre-select only a portion of the entire file. These include online portals dedicated to Open Data. Publishing the file in portions rather than whole has been discussed but rejected on the grounds that no consistent and future-proof approach could be determined.

We also continue to update the holding website with supplementary information where relevant. In particular, we should take care to explain why blank fields occur. Although blank fields represent a very tiny proportion of the data in the release, such is the amount of data published that they can add up to a large absolute volume of prescriptions and could facilitate misinterpretation.

**Conclusion:** Attempts to access the files have met with mixed results, although we continue our attempts to address the areas where access could be made easier thus enabling more of the public to use the data. We intend to review the supplementary material and metadata to improve understanding of the data by audiences with less experience of prescribing data.

### Discoverability

- Proactive use of social media and other communications tools to engage new audiences
- Large volume of traffic still arrives via conventional modes of publication
- Key issues faced by new audiences raised.

Downloads of the files are monitored using webstats. At least 100 attempts to download a fresh file are made each month. Downloads of every file continue to occur (albeit at a slower rate) even six months after publication of the first files. Since the average time spent on the

page is 4min 47secs, we can reasonably assume that perhaps half of attempts are long enough to complete the download.

Of 441 download attempts (of any file) made from the day of the July publication to the day before the September publication, 104 (23.6%) were from direct traffic. Direct traffic lands straight on the page supplying the Open Data and could be from bookmarks, from links sent out via social media or email or other direct routes. Direct traffic could therefore be a mix of new and established customers. Greater use of campaign tracking, in particular via social media, is likely to provide better insight into our ability to connect with new audiences.

There were 77 download attempts (17.5%) from traffic that reached the Open Data via other areas of the Prescribing and Medicines area of the ISD website and a further 72 (16.3%) from the remainder of the ISD website. This traffic, representing roughly a third of download attempts, reaches our Open Data via our more conventional models of publication such as listings on the ISD homepage and may more closely represent established users of ISD data. Traffic for the remaining 188 download attempts (42%) is not able to be classified in any meaningful way.

Thanks to a proactive social media campaign, we were able to obtain feedback relating to how the burgeoning Open Data community discovers data. The manner in which the files are presented and situated on the website, the 'predictability' of the URL (enabling automated algorithms to download new publications at regular intervals) and the content of the metadata are all issues which we could improve on as we engage more with this community. It has also been suggested that we apply for an Open Data Institute certificate which will help us to adopt the required conventions.

**Conclusion:** Interest in the files is sustained after six months and traffic reaches the site from a variety of sources. Better tracking of traffic and more sophisticated use of webstats will enable ISD to better measure its engagement outside of established audiences.

### Reference files

- Several sources indicated that reference files should be either provided within the data file or signposted more visibly
- Some integrity issues between open data and reference files
- 4% to 7% of website traffic known to navigate to signposted reference files.

The Open Data file is coded to save space and reference files are used to translate these codes to something intelligible. For example, the 14 codes for each Scottish Health Board can be translated to the name of each Health Board. We had decided early on to signpost users to reference files contained elsewhere on ISD and other NHS websites after reviewing the options available. Importantly, this means we decided not to provide reference files ourselves which is something that several customers have since asked for.

Other queries related to the integrity of the GP Practice reference file. This is a known issue; a small but significant volume of prescribing in the community is conducted outwith the

conventional GP practice model. Where this is the case, the Open Data prescribing will remain unassigned when matched to the GP practice reference file provided by ISD.

In the early months of publication, we found that roughly 4% to 7% of website traffic to the open data page (around 50 views) exits to the pages on which the reference files are stored. This suggests that our signposting approach is to some extent directing traffic.

**Conclusion:** We propose to review our approach to reference files, potentially making the signposting clearer (the links are currently stored within guidance document PDFs, which impacts on discoverability) and providing briefing on the reference files themselves as they relate to our own data. Further developments would involve closer collaboration with teams responsible for the files to enhance compatibility.

### **Extra fields**

- Sources requested even greater granularity almost immediately
- Customers know our data and have specific uses
- Must be balanced against broad appeal of Open Data.

On publication, several customers were interested in enhancing the detail of the data by including extra fields. As analysts ourselves, we understand that greater detail is almost always more useful. However, we have to balance this with issues around the size of the files and the potential disclosure risk.

**Conclusion:** ISD Prescribing is considering all requests for the inclusion of additional fields against a number of criteria. These include the added value of any single or combination of fields to a wide audience, the potential disclosure risk and the size of the expanded file. Customers remain able to request smaller extracts of more detailed data and, as at the beginning of the project, we continue to monitor these requests to inform the content of the Open Data.

### **Backdated files**

- Several requests for files to be backdated or purchased separately
- Indicates real demand.

Such is the popularity of the files that several customers have asked to receive files backdated to as early as February 2014.

**Conclusion:** ISD Prescribing is encouraged by such requests as it attests to the value of the data, and will give serious consideration to publishing backdated files if the level of interest in this is sustained.

## **Conclusion**

The ISD Prescribing Open Data project has been an intensive learning experience. In publishing this report, we hope to share what we learned with other organisations, NHS or otherwise, who are contemplating embarking on this journey.

As can be seen from the large amount of feedback and website traffic, interest in the data held by ISD Prescribing is extensive.

Much of what has been fed back to us will inform our efforts in continuous improvement or shared with the wider teams at ISD to help inform an ongoing Open Data strategy. Although we did not gain direct answers to the specific questions we asked of this evaluation, we got some way towards them while providing a proof-of-concept pilot to ISD to pave the way for other teams to improve on our work.

## **Further feedback and comment**

We encourage further comment and discussion of our Open Data. You can get in touch with a member of the ISD Prescribing team on [NSS.isdprescribing@nhs.uk](mailto:NSS.isdprescribing@nhs.uk) or tweet us [@NHSNSS](https://twitter.com/NHSNSS).