

## **Publication of GP practice level prescribing data and privacy risk**

### **Introduction**

In line with Scottish Government policy and strategy on Open Data, ISD Scotland will be publishing GP practice level prescribing data starting in April 2016. The data published will be similar to that released by the other countries of the United Kingdom and follows a consultation carried out in Autumn 2015 across a range of stakeholders. Two key messages from that consultation were that:

- Potential users of the data indicated that, to be useful, it needed to be at GP practice level and indicate the drug preparation
- NHS respondents in particular were concerned about the potential risk of disclosure and the ability to identify individual patients. This was considered to be greatest for single-handed practices.

This paper describes the principles and processes applied by ISD in assessing these concerns and the risks that might be associated with the data release.

Data to be released will be based upon reimbursed prescriptions and will consist of:

- Preparation paid (i.e. specifying the drug, formulation and strength)
- The number of prescription items paid
- The quantity paid (i.e. total number of tablets, capsules etc.)
- The gross ingredient cost (GIC) (i.e. the cost at NHS list price)

Information will be released on a monthly basis for each GP practice in Scotland. Information about paid prescriptions written by other groups e.g. hospital clinics, dentists etc, will also be released and will be aggregated at Health Board.

### **Addressing disclosure and privacy concerns**

The data release will contain no information about either the individuals being treated or about the number of patients being treated. However, at the level of granularity that the data will be released, approximately 75% of data rows will relate to four or fewer prescriptions and 40% to a single prescription. A single prescription for a preparation must clearly relate to a single individual. In recognition of this, the concerns expressed during the consultation, and to reassure both the public and wider healthcare community, ISD has given very careful consideration to this issue and this is described below.

As noted above, the other countries of the UK already routinely publish prescribing data. NHS England initially released highly aggregated data by British National Formulary (BNF) classification at Primary Care Trust level. In response to the open government agenda and following a consultation this was extended to chemical (drug) substance and at GP practice level: single handed practices were excluded initially. Following further discussions involving the Information Commissioner's Office it was concluded that although there might be a small risk of identification, this was outweighed by the potential public benefit and so information at preparation level for all practices started to be published in 2012. There have been no adverse incidents reported as a result of this.

In assessing the potential disclosure risks as a consequence of releasing Scottish practice level prescribing data, ISD did so with reference to our established disclosure risk assessment processes.

1. The files to be released are large, containing approximately 1.2 million rows of data each month. Furthermore, the data about prescribing and the details of the practice (name, location etc) will be published as separate files that will need to be linked by users of the data. We have therefore concluded that casual disclosure is extremely unlikely and so have approached our assessment from the assumption of a motivated intruder.
2. Although the data to be published includes large amounts of data relating to single prescribing events, this was considered to be low risk for a number of reasons:
  - a. The data is published monthly based on when the claim for payment was made. This may not be the same month that the prescription was written, which could be up to six months earlier.
  - b. Many drugs have multiple therapeutic uses therefore even if one could identify the individual being treated with a particular drug, this does not unequivocally identify the condition suffered by that individual.
  - c. The average GP practice serves a population of approximately 5,000 people and often there is considerable overlap between practices regarding where their patients live.

However, we further considered the possibility that there might be specific situations where the potential disclosure risk is increased. This could be where the condition being treated was rare, where a drug had a very specific use, or where practices were small (<1,000 patients) or served geographically well defined populations. We considered and assessed these issues in the context of treatments for conditions normally considered sensitive under existing ISD guidelines i.e. mental health, sexual health, blood-borne viruses, substance misuse and vulnerable groups such as children. In doing so we considered:

Would the release of practice level prescribing data allow someone using that and any other data that might be available to identify where individuals with a rare or potentially sensitive condition were located?

If the above was true and / or a GP practice was small (< 1,000 patients), is there anything about the drug treatment or condition that would make the individual distinguishable from any other individual?

In examining the data and the potential disclosure risk, ISD made use of additional information available to us but that will not be available within the prescribing data to be published. For most of the areas considered potentially sensitive, the overall volume of prescribing and distribution among GP practices meant that we do not believe that it would be possible to specifically identify small groups of individuals or attribute a prescription event to a specific individual. For therapy areas with small annual levels of prescribing, more detailed analysis revealed that the pattern of prescribing indicated that such prescriptions were often sporadic and so were more likely to represent an occasional supply to someone normally receiving treatment through a specialist hospital service but who had perhaps run out of their medicine. Many of these potentially sensitive drugs are normally provided through specialist hospital services. Such treatments might be provided directly and so not included in this published data but where it has been supplied via a prescription dispensed in the

community then it is reported in this data aggregated to Health Board level. This will be the health board of treatment and so not necessarily the Health Board in which a patient lives. There were no situations suggestive of regular treatment within a small and geographically well defined GP practice.

## **Conclusions**

ISD has therefore concluded that the volume and complexity of the data means that there is a negligible risk of casual disclosure.

A motivated intruder that considered that they had identified an individual from the published prescribing data would have made a number of assumptions:

- The location where treatment took place (GP practice versus alternative possibilities),
- The timing of the prescribing act in relation to the claim for payment
- That the drug had only a single possible use and that the condition being treated had no alternative treatments (including not receiving the drug treatment)
- There were no other patients within the GP practice who could have the condition

The risk of potential disclosure and consequences to the individual is greatest for small populations (e.g. rural practices) and sensitive conditions. Following detailed consideration and analysis we conclude that this would require additional knowledge or information that is not readily available and so the risk is small. Although we cannot say with certainty that the risk is nil, this does not prevent the publication of prescribing data at the level of the GP practice and the preparations prescribed.