

Senate Community Affairs Committee

ANSWERS TO ESTIMATES QUESTIONS ON NOTICE

HEALTH PORTFOLIO

Supplementary Budget Estimates 2016 - 2017, 19 October 2016

Ref No: SQ16-000661

OUTCOME: 0 - Whole of Portfolio

Topic: MBS/PBS Data Breach

Type of Question: Written Question on Notice

Senator: Watt, Murray

Question:

In relation to the breach of information in the MBS and PBS datasets published on the data.gov.au website:

- a) On what date and time was the Department informed that there was vulnerability in the MBS and PBS datasets published on the data.gov.au website?
- b) When was the data pulled off the website?
- c) When was the Office of the Australian Information Commission notified of this breach?
- d) How many line items of data were used in the datasets – both in terms of individuals, and health service providers?
- e) When did this information date back to? Is it correct it dated back to the 1980s?
- f) On how many occasions was this information accessed? Is the Department aware who accessed the information? If so please detail whether the information was accessed by individuals or other entities.
- g) The Department released a statement on September 29 confirming a data breach. When was this statement originally prepared?
- h) Did anyone in the Minister's Office give the Department advice or instruction on the timing of the announcement of the breach?
- i) Has the Department reviewed any other datasets, either on data.gov.au or on the Departmental websites, for similar vulnerabilities?
- j) Have any staff been counselled or disciplined over the breach?

Answer:

- a) On 8 September 2016 the Department of Health received a notification made in the public interest by the University of Melbourne indicating that it was possible to decrypt health service provider ID numbers in the 2014 MBS data in the 10 per cent linkable MBS and PBS datasets.
- b) 8 September 2016 at about four o'clock in the afternoon.
- c) On the morning of 9 September 2016.
- d) The dataset comprised MBS and PBS claims data for a random selection of 10% of Australian patients, representing approximately one billion lines of data relating to approximately 2.5 million Australians.

As the MBS/PBS data is a sample of patients, not a sample of claiming activity for the selected patients, it is not possible to quantify the number of individual service providers.

- e) Claims data for the MBS dataset covered the period 1984-2014. Data relating to the PBS covered the period 2003-2014.
- f) The MBS and PBS dataset consisted of 43 separate files. Between the period 1 August 2016 and 8 September 2016 (when the datasets were removed from data.gov.au) there were 1500 downloads of some or all of the 43 data files.

Of the 1500 downloads, approximately 500 were from academic or government domains, with the balance from a variety of other domains, including private sector firms and not for profit entities.

- g) The public statement of 29 September 2016 was finalised on the morning of 29 September 2016.
- h) The timing of the announcement was subject to decisions of Government on related matters, such as the intention to introduce legislation to protect de-identified data.
- i) The Department has not identified any other Health public source datasets that use the same de-identification methodology in relation to health service providers as the set removed.
- j) No. An independent review of the circumstances highlighted improvements the Department should make to its data governance and approval processes. A number of these were under way and have since been accelerated. No disciplinary matter was identified.