Results of Survey of PBS Schedule Outputs

Executive Summary

The Pharmaceutical Benefits Division produces the PBS Schedule every month in a variety of outputs and distributes these to Software Developers, under embargo, four weeks before the effective date. Developers incorporate these outputs into their software packages which they distribute to their clients throughout Australia.

In March 2015, the PBS Information Management Section created an online survey to investigate and measure the usage, suitability and the perceived quality of the monthly PBS outputs. The aim was to discover which outputs were working well and which areas needed improvement.

The survey was emailed to all external Software Developers who have approved secure access to the monthly PBS embargo data. The survey covered a range of issues relating to the resources available to developers and provided an opportunity for them to provide the Department with feedback.

What was measured

The survey measured the level of use and the perceived quality of the current PBS data outputs, which predominately includes the range of PDF's and XML related materials.

Respondents to the Survey

The survey was open for one week. 29 responses were received from 26 organisations. The responses included approximately 15 pages of feedback.

Summary of results

There was a high degree of participation in the survey indicating the readiness of developers to engage in a constructive discussion with the Department.

Overwhelmingly, the responses indicated that the PDF Summary of Changes and the PDF Schedule were the most frequently used, essential and positively rated of all outputs.

Despite being available in various iterations for nine years, the XML related outputs were used by a minority of developers, rated to be not useful and of poor quality and presented a significant barrier to entry for both existing and new users.

The average time taken to process each release of data was between two to three days with a maximum of ten days. The majority of responses indicated that the release of multiple versions of PBS data each month requires recipients of the data to re-process and re-do their work for each and every release. This can significantly impact on the time required to produce a final product and wastes significant resources on duplication whilst working within tight timeframes.

This is often compounded by the late delivery of the data. Many developers expressed considerable frustration at this burden; some commenting that they usually wait as late in the month as possible for the latest release as they have little faith the first version will be correct or final. Releasing data late in the month can also result in developers omitting last minute changes from that months software release – meaning prescribing/dispensing software may potentially not contain the same data as contained in the Schedule.

The majority of responses indicated they are aware that the TXT extracts (csv files) may be phased out. However, there was a strong desire for the retention of the TXT extracts, which, along with the

SQL version, were the most requested formats for the PBS data. Responses described the TXT files as being the most useful, reliable and easiest format to work with. Even some users of the XML commented they still require the TXT files.

Over half of respondents indicated they were aware of the approaching changes with XML Schema v3 and AMT v3. Several responses commented on Complex Authority (CAR) Drugs and requested they be better identified in the XML (this is being undertaken).

Over half of responses indicated they did not use the XML. One responder indicated they developed their systems to use the XML only because of the anticipated removal of the TXT files. Several comments were critical of the XML, citing it as being too big, too complex, volatile and from their business perspective, unusable.

The department's implementation of system changes was ranked moderately. 51% said it was average, 22% very good and 3% very good. However 17% rated the department's implementation performance as poor.

61% of respondents indicated that changes made to the PBS outputs required changes to their processes, software and resourcing. Several comments noted some changes can be costly, labor intensive and can take 4-6 months to implement. They requested more engagement, better communication and more lead time to implement changes.

Conclusion

Overall, the responses provided were comprehensive and informative. Developers and users of the data were very interested in providing feedback. The survey indicated several discrepancies with the current assumptions that the Department is making in regards to the PBS outputs. Notably the TXT files remain extremely popular and their future availability should be reviewed and confirmed at least until such time as a viable alternative is available. The complexity of the XML presents a barrier to entry of new developers and significant impost on current software providers. Multiple versions and the lateness of data has a significant resource impact on developers and upon downstream recipients. The popularity of the PDF versions of the Schedule was significantly higher than expected. The significant low use of the XML outputs and their low quality ranking was unexpected.



Question 1: How often do you use the following?

Frequently used PBS outputs	Rarely used outputs	
1. PDF Summary of Changes - 93%	1. XML Schema 72%	
2. PDF Schedules 86%	2. SQLite 72%	
3. PDF Errata 82%	3. XSL Stylesheets 68%	
4. Developers email alerts 79%	4. XML Schema doco 68%	
5. Developers website 75%	5. XML release notes 65%	



Question 2: How useful / essential do you find the following?

Essential PBS outputs		Not Useful outputs	
1.	PDF Summary of Changes 93%	1.	XML Schema documentation 58%
2.	PDF Schedule 79%	2.	XML Schema 51%
3.	PDF Errata 68%	3.	XSL Stylesheets 51%
4.	TXT Extracts 68%	4.	XML 44%
5.	Software Vendor Meetings 41%	5.	SQLite 44%





Positively rated outputs		Negatively rated outputs	
1.	PDF 86%	1.	XML 17%
2.	PDF Summary of Changes 79%	2.	XML Schema 17%
3.	PDF Errata 62%	3.	XML Schema Doco 17%
4.	TXT Extracts 58%	4.	XSL Stylesheets 14%
5.	Developers website 41%	5.	XML Release Notes 14%

Question 4: Please estimate the time, cost and resources you spend each month to process the PBS outputs. Please describe the impact of receiving multiple versions of data in one month:

Time taken (per release)	# of responses
5 hrs or less	8
2 – 3 days	10
3 – 5 days	4
10 days	1

Note: Before PharmCIS, PBS data was only ever released once a month and always on the same date. PBS data is currently routinely re-released at least twice each month, sometimes three or four times.

Key issues reflected in comments:

- 19 respondents (70%) advised that the re-release of PBS outputs each month has a considerable negative impact on their business. Multiple versions require developers to reprocess each release and can double, triple the time required – impacting on already very tight timeframes, compounded by the lateness of the data.
- 2. 12 respondents (44%) complained about the lateness, lack of a firm deadline and/or commented that they often wait as late as possible to process the data as they have little faith that the first release will be the final version.

Notable comments:

- a. "If late changes come through after we have released for that month, we have to rebuild, reverify and re-release, which has a heavy impact on our development team. Customer confusion around the extra release causes a large spike in support calls which has a heavy impact on our operations team."
- b. "Multiple versions result in rework from multiple teams and this impact is amplified the later in the month the change comes through. if the data is already out at pharmacy the rework is immense for rereleases"
- c. "Whenever there is an updated version, our processes require us to begin again to make sure everything is covered correctly. Consequently the costs for processing the files increases dramatically."
- d. "We now wait until the middle of each month before importing the PBS data as we assume that the first version will not be the final"
- e. "... if there are multiple releases, then depending on the time of these releases (early in the month or late) then this number of hours will be stretched, to sometimes double the number. The late re-release of PBS data, and often the lack of information about why it is being re-released, is a constant frustration and waste of resourcing from our perspective."
- f. "Multiple versions within a month has considerable impact with occasional rollback & reprocessing, or manual re-work of data that was processed using programs to load. it is very frustrating to NOT have a firm deadline to release that data for distribution to our clients."

Question 5: Are you aware that the TXT Extracts may eventually be phased out?



Question 6: Are you aware of the changes coming with XML Schema v3 & AMT v3?



Question 7: What features would you like to see included with the implementation of XML Schema v3?

Key issues reflected in comments:

- 1. 3 respondents (16%) commented on Complex Authority (CAR) Drugs and one respondent requested that CAR items be identified more easily. One respondent suggested added a drug type code 'CA'
- 2. 10 respondents (55%) to this question indicated they did not use the XML.
- 3. 11 respondents (37%) of the survey did not answer this question.

Notable Comments / requests:

- a. "Proper Use of AMT v3 NO More DHS Namespace encoding that is NOT AMT."
- b. "Complex Authority Required items to be extractable. Streamlined Authority Codes History for what SL Codes were valid for each month for the prior 12 months."
- c. "The current XML & Text Files ONLY represents what can be PRESCRIBED it does not properly represent what can be DISPENSED. This creates problems for items that can still be Dispensed"
- d. "1. we can identify PBS-items that have a TGP, we want also to be able to identify the PBS Therapeutic Group base-priced items. 2. we need more detail in Authority-Required items to enable machine interpretation of when an item is written only - there are many examples where the #in-writing v. #telephone options are conditional and those conditions are only spelt out in plain text"
- e. "Consistent Primary Keys in XML files from month to month (they change each month). -If a TPP is changed it would be good to see a previous version attribute for matching purposes."
- f. "I am hoping we will still be able to import the monthly data from the PBS XML file."
- g. "THE FEATURE I SEE AS MOST IMPORTANT WITH XML IS A FULL SET OF STYLE SHEETS AND THE TXT FILES BEING GENERATED I... The XML is too big and too complex to count on as accurate and it is just too hard to check the txt files are much easier and quicker to check. Even the experts who make the xml struggle to get it right and we have been let down by it a number of times as I am sure you know. The fact that the experts are reluctant to do style sheets indicates how complex the job is how can we be expected to produce critical data from the xml when even the experts struggle. There are major works required every time the xml version changes

that can be avoided if using the txt files (easier to implement changes) and probably also if using SQL. "

- h. "Reduction in verbosity and complexity."
- i. "Since we process and release data for Oracle databases, we prefer not to be forced to use XML files."
- j. "Some kind of formal governance over the development process and some transparency with the schema requirements. To help lift the quality of the final product."

Questions 8: Over the years, the Department has made changes to the PBS outputs as a result of policy or system changes. How would you rate the Department's implementation of these types of changes?



Question 9: Please describe the impact these changes to the PBS outputs have had on your business. Do you have any suggestions on how the Department could manage these changes differently in the future?

Key issues reflected in comments:

- 1. 11 respondents (61%) indicated that changes made to the PBS outputs required changes to their processes, software and resourcing. Several comments noted the changes can be costly, labour intensive and can take 4-6 months to implement a change.
- 2. 2 respondents noted that MSIA's involvement in changes has been greatly beneficial.

Suggestions included:

- a. "Before making any changes, the Department should engage with stakeholders and communicate relevant information such as the estimated timeframes for changes take place."
- b. "Sample files need to be provided in greater lead time, and they need to mirror a full file, not just a cut down copy of the XML."
- c. "Continue to deliver data valid against previous schemas for at least 12 months after any schema has been superseded."
- d. "STICK WITH TXT FILES. Much more efficient with time, resources and costs."
- e. "We need to avoid non-passive changes and the best way for us is to have an alternative to XML files (SQLLite works fine for us)."
- f. "Transformations that demonstrate the actual data needed to describe the change, e.g. with the nurse practitioner changes, we had to develop our own transformations to manage the change. It would be preferable if the transformations were provided."

Notable comments:

- a. "we paid an external contractor to help changeover to importing the data from the PBS XML file... but have sometimes had to use the PBS TXT Extracts."
- b. "Changes to the PBS data outputs particularly with major changes eg Chemotherapy over 2

years have a big impact. There are additional programming resources and support requirements that require funding. Prior to PBS online the changes where infrequent, but now it seems that there is some form of change on the go all the time. While some changes have been funded, some eg Streamline PBS/RPBS Claiming have not - this is stress at our expense. DHS working with the MSIA has been of great benefit to have the projects well defined and specified to help vendors implement the changes."

- c. "Each change have required timely alterations made to our own parser"
- d. "major development work to switch from consuming TXT to XML, we expended a lot of resources in this transition; now that we have made the switch, we hope there will be no changes in the future i.e. continue with the same method of data delivery (via the secure website), continue with XML"
- e. "Obviously change has increased the workload from time to time but the maintenance of the txt files has been great and assisted to make changes easier. Talking about the Department I have been generally very happy with the support, and the consultation has been great. Sometimes the follow up seems a bit slow. I cannot say the same about the XML team. Questions have seemed to be ignored and with great support from MSIA I have at times had an answer but after 6mths really! Without MSIA questions have been passed around and ignored. No they have not been helpful but the Dept has and thanks again for maintaining the txt files which have saved the day for all vendors when the XML team has failed."
- f. "PBS output changes can put us under pressure to deliverer the updated dispense package to our customers on time."
- g. "Requires alteration and validation of processing flows which is somewhat labour intensive."
- h. "Since we have to plan for engineering resources, it generally takes a minimum of 4-6 months to implement a change. We need to avoid non-passive changes and the best way for us is to have an alternative to XML files (SQLLite works fine for us)."
- "The number of data re-releases seems to have increased over the past few months March 2015 had 4 releases! This is a poor reflection on the success of the implementation of these changes."
- j. "The progressive introduction of AMT over time has created difficulties with names of medicines."

File Format	# of requests	%
Text/CSV	11	44%
SQL	10	40%
XML	8 (2 negative responses)	32%
PDF	7	25%
Excel	6 (1 negative response)	24%
Real time	1	4%

Question 10: If you could have the monthly PBS Schedule outputs delivered in any format, what would that be and why? (i.e. XML, PDF, SQL, Excel, other)

Key issues:

- 1. TXT/CSV files are the most requested file type. Common response was they are the easiest to process, search and maintain.
- 2. SQL has a lot of interest and developers seem eager to embrace it.
- 3. XML responses had the most significant negative comments. Feedback included 'difficult to process without errors' and 'XML is okay for transferring small amounts of data but I believe

it is not a good choice for a project as big the PBS. It has shown to have problems all along the way.'

4. Excel was the least preferred format. One comment was that it 'is easy to corrupt the data'

Notable comments:

- a. "PDF files are and always will be critical for data validation. It is the closest thing we have to an official document that can be used as a reference source. The file can also be saved monthly as a reference to past states of the PBS if requiring to confirm a streamline code in a particular month or the status of a benefit at a previous time eg Authority, repeat, quantity, restriction, price etc"
- b. "stick with XML, because we have now built our system around that... stick with PDF, as we continue to have a need for a human-readable form of the data during the embargo period"
- c. "we prefer CSV text file as we have the tools setup to handle them. Also I feel they are the easiest to process, search and maintain"
- d. "XML has been constantly changing & requires expertise with XSL style sheets. It also is susceptible to the xml parser you use & the data is now so big that it is difficult to process with out errors."
- e. "Since we have built a suite of utilities to import and process the text files, we would prefer to see them delivered in his format, but if that is impossible, we would probably prefer to have them either as SQL or Excel files that we could then process in a similar fashion."
- f. "We have now changed our process, with the assistance of external contractors (at not insignificant cost), to import the required PBS data from the XML file."
- g. "We hope that the changes to XML Schema v3 and AMT v3 do not impact on our existing process??"
- h. "TXT files as they are the most efficient to work with. Compact, robust, easily searched and can be loaded into almost anything. They have been around since the start and changes are easy to implement."
- i. "(Never) XML, XML is okay for transferring small amounts of data but I believe it is not a good choice for a project as big the PBS. It has shown to have problems all along the way. The file has become too big to handle with any standard tools which means writing specialized specific software that requires a major re-write with each new version. "
- j. "Our newer product delivery of drug updates is via consumption of the XML file which would be desirable moving forward.. again as stated previously it is high risk with the changes that are unknown into the future."
- k. "We strongly prefer the SQLite format ... this format has been trialled over the recent months, but we would welcome and encourage this format to be invested in further, promoted and expanded."

Question 11: Do you have any additional comments or feedback about the PBS Schedule outputs?

Key issues:

- 1. 3 respondents made suggestions for additions to the PDF Summary of Changes
- 2. 6 respondents (35%) made further complaints about the multiple releases every month, the lateness and lack of a regular delivery schedule.
- 3. Further comments were made about the complexity and volatility of the XML and several more requests were made to keep the TXT files.

Notable comments:

- a. "Adding of a table of contents with navigation links to the PDFs would be nice"
- b. "From a quality assurance point of view we use the summary of changes extensively for manual testing purposes can it please include the prescriber type... would also like to see the authority status and streamlined codes listed here against each drug"
- c. "XML is not the answer it is not the medium for such a large and complex task as the PBS it was developed as a data transfer medium not a data storage system. XML can be good in being humanly recognisible but the PBS file is too big and too complex for that now and additionally it is too big to load into any standard too, to look at. It is too big and we cannot work with a delta file can explain why if you do not know. Many rules written into the XML we encoded years ago and are unnecessarily being re-invented by people who should have consulted those that have been working with this data for decades before wasting that time, energy and creating file bloat. "
- d. "Multiple releases and late releases make data incorporation extremely difficult. Having errata that is ongoing (and thus not incorporated into the xml and text files) limit the usefulness of the data released. "
- e. "Our processing of PBS data is reliant upon the Avergae Rates text file ... this file is typically delivered ~17th of each month. Until this arrives, we cannot begin our processes. Could this file, or data contained within it, be delivered in a more timely manner in the future?"
- f. "Please add the following to the summary of changes: prescriber type flags (eg: NP), authority indicator, - applicable streamlined codes. At the moment we need to cross reference with the full schedule to verify this information which is very time consuming. Sometime the SOC becomes available without the full schedule, and these missing details impede our QA"
- g. "Programming interfaces that require a significant commitment to implement must be treated like contracts. The terms and conditions must be understandable and once they are agreed, they must be stable. If one party to the transaction continually, unilaterally alters the contract, it will become harder to convince parties to invest in the transaction. Programmers have looked at the mind-boggling complexity and volatility of the PBS XML and the simplicity and stability of the text files and have decided that an investment in the XML is too risky and offers too little return."
- h. "The hardest part of the distribution is the uncertainty of the finalisation of the data for distribution. With the multiple versions produced each month, to add items, fix items etc if one thing would make it easier, it would be to have a deadline for the month & run with that for production."