

ISDS 2018 Conference Abstracts

Leveraging the NSSP R Studio Server to Automate QA Monitoring and Reporting

Peter J. Rock* and Michael D. Singleton

College of Public Health, University of Kentucky, Lexington, KY, USA

Objective

The aim of this project was to develop a nimble system to both monitor and report on the quality of Kentucky emergency department syndromic surveillance (SyS) data at system-wide and facility levels.

Introduction

In 2016, the CDC funded 12 states, under the Enhanced State Opioid Overdose Surveillance (ESOOS) program, to utilize SyS to increase timeliness of state data on drug overdose events. In order to operationalize the objectives of the grant, there was a need to assess and monitor the quality of Kentucky's SyS data, with limited resources. We leveraged the NSSP's R Studio Server to automate quality assurance (QA) monitoring and reporting to meet these objectives.

Methods

Using the R Server, we pulled data from the process messages table, aggregating messages to single patient encounters. In addition to compiling the code on a powerful remote server, the server can access the process table messages relatively quickly. We developed an R Markdown report to produce a report that includes a variety of system- and facility-level metrics that highlight key indicators of system performance and data flows. By using R, we were able to create an auto-generating QA report that runs weekly and e-mails for analyst review. Quality metrics included: % completeness of chief complaint and discharge diagnosis codes (overall and by facility)[Fig 1 & Fig 2]; visit trend by day of visit (with interactive spark lines)[Fig 2]; maximum date of message created, date message arrived at NSSP server, date of visit, and total messages[Fig 3]; message arrived trend (interactive sparklines)[Fig 3]; volume and type of error messages failing to process[Fig 4]; message volume by ADT type[Fig 5]; and volume of patient class by type by day[not shown]. Our SyS analyst reviews the report and delivers it to stakeholders with general comments about ongoing and newly emerging data quality concerns.

Results

The report has proven to be beneficial in ongoing QA monitoring. The report is shared weekly with key stakeholders at the Kentucky Department for Public Health, Kentucky Health Information Exchange, NSSP, and regional ESSENCE users. Findings are reviewed at monthly SyS stakeholder meetings. The report has identified numerous errors, dead feeds, and other systems changes in near real-time; leading to corrective action and general data quality enhancement. Weekly monitoring of QA has improved data feed stability and communication of identified issue with key stakeholders.

Conclusions

The R Studio Server provides a nimble platform to develop, refine, and automate a QA reporting system that can lead to improved SyS data quality. In Kentucky, in addition to improving overall data quality, these weekly reports and subsequent communication have helped build relationships among key stakeholders and elevated the importance of syndromic surveillance data locally. Continual monitoring of data is critical to ensure quality and therefore the validity of the data.

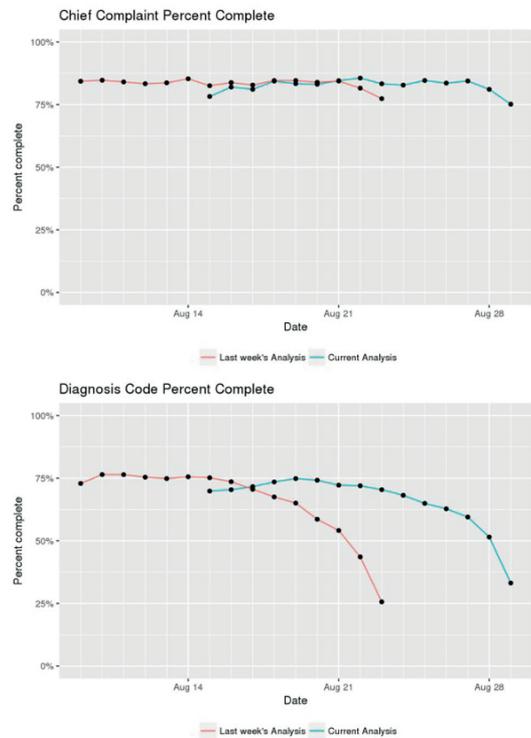


Figure 1

Chief complaint and diagnosis code completeness by Facility

The following table details the percent completeness of diagnosis code (DX), chief complaint (CC), both, and neither. The table also includes the total number of ED visits. The table covers the quality metrics of any E/Eme patients with a visit date on or after 2017-09-27

2017-09-27

Visit trends - The sparklines represent the volume of ED visits based on the visit date

Show 10 entries Search:

	Facility Name	%DX	%CC	%Both	% Neither	Total Visits	Visit Trend
56	Facility A	75.74%	33.20%	32.56%	23.62%	8960	
48	Facility B	32.44%	2.40%	2.40%	67.56%	7265	
29	Facility C	95.48%	100.00%	95.48%	0.00%	2721	
15	Facility D	0.82%	0.00%	0.00%	99.18%	2313	
75	Facility E	0.00%	6.43%	0.00%	93.57%	2131	
4	Facility F	79.45%	98.37%	79.45%	1.63%	2063	
70	Facility G	93.91%	99.95%	93.91%	0.05%	2052	
3	Facility H	86.06%	99.58%	86.06%	0.42%	1679	
63	Facility I	94.34%	100.00%	94.34%	0.00%	1644	
1	Facility J	85.49%	99.69%	85.49%	0.31%	1626	

Showing 1 to 10 of 78 entries Previous 1 2 3 4 5 ... 8 Next

Figure 2



Message Table Results

The following table shows counts of messages for emergency department visits.

The following table details the:

- Facility name
- Facility type (repeated values for facilities with multiple types)
- Max message date - The maximum date of when a message was created (facility side)
- Max Arrived date - The maximum date of when a message hit the NSSP server (CDC side)
- Max visit date - The maximum patient visit date the message is referencing (facility side)
- Message count - The total number of messages since 2017-09-27
- Message trends - The sparklines represent the volume of messages based the time it arrived at NSSP/CDC

If a facility is functioning as expected we would expect to see arrive date ≈ message date ≈ visit date. Visit Date may be slightly behind based on server volume and when the report was generated.

Show 10 entries

	Facility Name	Max. Message Date	Max. Message Arrived Date	Max. Visit Date	Message Count	Message Trends
47	Hospital F	2017-08-16	2017-09-30	2017-08-16	1	
66	Hospital G	2017-10-01	2017-10-05	2017-09-27	10	
11	Hospital B	2017-10-10	2017-10-10	2017-10-09	48	
22	Hospital D	2017-10-09	2017-10-10	2017-10-02	176	
27	Hospital O	2017-10-09	2017-10-10	2017-10-09	1052	
29	Hospital W	2017-10-10	2017-10-10	2017-10-10	12636	
42	Hospital A	2017-10-09	2017-10-10	2017-10-05	215	
57	Hospital X	2017-10-10	2017-10-10	2017-10-10	2056	
68	Hospital N	2017-10-10	2017-10-10	2017-10-10	5118	
69	Hospital S	2017-10-10	2017-10-10	2017-10-10	7687	

Showing 1 to 10 of 79 entries

Figure 3

Exceptions Table

The following table shows counts of errored messages that do not make it to the processed table

The following table details the:

- Facility name
- Count of messages
- Message error type

This section of the report only looks at errored messages that occurred since 2017-09-26

NOTE: This is for any/all messages not just ED visits. Because these are errored messages that do not get processed there is no easy way to determine which messages would have been processed to ED visits.

Show 15 entries

	Facility Name	Message Count	Message Error Type
1	Hospital D	1427	Message did not include valid C_Unique_Patient_ID
2	Clinic S	10	C_Visit_Date_Time is set in the future
3	Hospital M	646	C_Visit_Date_Time is set in the future
4	Clinic Q	570	C_Visit_Date_Time is set in the future
5	Hospital U	469	C_Visit_Date_Time is set in the future
6	Hospital X	212	C_Visit_Date_Time is set in the future
7	Clinic T	6	C_Visit_Date_Time is set in the future
8	Hospital I	11	C_Visit_Date_Time is set in the future
9	Hospital 8	206	C_Visit_Date_Time is set in the future
10	Hospital 7	12822	C_Visit_Date_Time is set in the future
11	Hospital R	6	C_Visit_Date_Time is set in the future
12	Clinic W	1	Message did not include valid C_Unique_Patient_ID
13	Clinic Q	44	C_Visit_Date_Time is set in the future

Figure 4

Overall performance

This section attempts to indicate some of the overall status of the data flow.

The graph below details the number of A01, A03, and A04 messages as they hit the CDC, by day, for the last 2 weeks (since 2017-09-26)

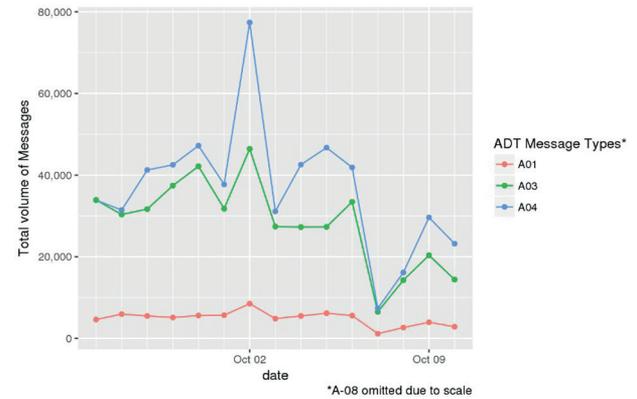


Figure 5

Keywords

Quality assurance; R Studio; NSSP; R Markdown

Acknowledgments

We acknowledge and thank the following agencies for their support of this work: The Kentucky Department for Public Health, the Kentucky Health Information Exchange, and the National Syndromic Surveillance Program.

***Peter J. Rock**
E-mail: pjrock2@uky.edu

