



A Critical Evaluation of Safety Signal Analysis Using Algorithmic Standardised MedDRA Queries

Carolyn Tieu^{1,2} · Christopher D. Breder^{2,3,4,5}

Published online: 16 August 2018

© This is a U.S. government work and its text is not subject to copyright protection in the United States; however, its text may be subject to foreign copyright protection 2018

Abstract

Introduction Algorithmic Standardised MedDRA[®] Queries (aSMQs) are increasingly used to enhance the efficiency of safety signal detection. The manner that aSMQs affect capture of potential safety cases is unclear.

Objectives Our objective was to characterise the performance of aSMQs with respect to their potential for double counting, the likelihood of events in aSMQ positive cases being clinically related, how frequently terms are used for algorithmically positive cases, and the face validity of positive cases based on the drug inducing events. We were also interested in what effect requiring symptoms to overlap temporally would have on performance.

Methods We reviewed adverse event (AE) datasets of New Drug Applications and Biological License Applications and compiled a database including preferred terms and corresponding SMQs, SMQ term categories, AE start day, AE duration, drug name, and Anatomical Therapeutic Chemical class. Two reviewers independently determined if the algorithm was met and, if so, whether the broad terms overlapped temporally.

Results A total of 107 marketing applications were reviewed, including 103,928 patients and 277,430 AEs. Use of algorithms condensed the number of AEs to between 5 and 8% and the incidence to about 1.5% relative to when the SMQs are used without the algorithm. Certain aSMQs exhibited a potential for overcounting. Requiring symptoms to temporally overlap helped to eliminate irrelevant cases.

Conclusions Our findings demonstrate that algorithmic and temporal assessment increased specificity of case retrieval, though the reduction in the number of terms or incidence seemed excessive for certain aSMQs. Evaluating the day of AE onset and duration improve specificity through identification of outlying events. Identification of drug classes known to cause the aSMQ's clinical condition provides face validity for this tool, yet detection of cases associated with novel classes may provide new understanding of these disorders. Improvements in some of the SMQ term lists may improve the performance of SMQs in general.

✉ Christopher D. Breder
cbreder1@jhu.edu

¹ FDA Fellow in the Oak Ridge Institute for Science and Education (ORISE) Program, Silver Spring, MD, USA

² Division of Neurology Products, Office of New Drugs, Center for Drug Evaluation and Research, Food and Drug Administration, Silver Spring, MD, USA

³ Regulatory Science Program/Advanced Academic Programs, Johns Hopkins University, Rockville, USA

⁴ Center for Drug Safety and Effectiveness, Bloomberg School of Public Health, Johns Hopkins University, Washington, DC, USA

⁵ US Food and Drug Administration, 10903 New Hampshire Avenue, WO 22 RM 4218, Silver Spring, MD 20903-1058, USA

Key Points

Our findings demonstrate strengths and deficiencies in the algorithmic Standardised MedDRA Queries (aSMQs) in detecting potential cases of interest in adverse event safety databases.

Suggestions are made for some modifications to the current aSMQs that may increase the efficiency and productivity of investigators and regulators who use these for data safety mining.

Table 1 Algorithm for broad term categories of seven SMQs

SMQ	Class of broad term categories	Algorithm for broad term categories
Acute pancreatitis	B: Laboratory values C: Signs and symptoms	One term from each category: B + C
Anaphylactic reaction	B: Upper airway/respiratory C: Angioedema/urticaria/pruritus/flush D: Cardiovascular/hypotension	One term from two of the categories: B + C D + B D + C
Anticholinergic syndrome	B: Nervous system-related PTs C: Psychiatric-related PTs D: Other anticholinergic syndrome-related PTs	One term from two of the categories: B + C D + B D + C
Eosinophilic pneumonia	B: Eosinophilia terms C: Terms possibly indicative of pneumonia	One term from each category: B + C
Neuroleptic malignant syndrome	B: Fever-related PTs C: Muscle rigidity or injury-related PTs D: Other NMS-related PTs	One term from all the categories: B + C + D
Systemic lupus erythematosus	B: Photosensitivity (1) C: Oral ulcers (2) D: Arthritis (3) E: Serositis (3) F: Renal disorder (1) G: Neurological disorder (2) H: Haematologic disorder (3) I: Immunologic disorder (3)	Sum of categories weights > 6
Tumour lysis syndrome	B: Lab-related C: Renal injury/cytokine release	One term from each category: B + C

NMS Neuroleptic Malignant Syndrome, PT preferred term, SMQ Standardised MedDRA Query

1 Introduction

Safety signal detection in clinical research data requires screening tools with great sensitivity but with sufficient specificity to allow for appropriate cases' retrieval. To facilitate signal detection, adverse events (AEs) are coded in a standardised terminology with terms from the Medical Dictionary for Regulatory Activities (MedDRA[®]), which is required for premarket and postmarketing safety data submissions to regulatory authorities in numerous countries, including the United States (US), Japan, and the European Union [1]. MedDRA offers a highly granular system of describing AEs, with over 20,000 preferred terms (PTs). A MedDRA tool that is increasingly used to help identify and retrieve potentially relevant safety cases is the Standardised MedDRA Query (SMQ). SMQs such as SMQ *Hepatic disorders*, SMQ *Haematopoietic cytopenias*, and SMQ *Anaphylactic reaction* are lists of terms at the PT level related to defined medical conditions. Since

their inception, use of SMQs for safety signal detection has been strongly embraced, as demonstrated by a 40-fold increase in use based on total documented SMQ investigations within safety analyses of US marketing applications and their reviews by the US Food and Drug Administration (US FDA) [2].

SMQs consist of groups of PTs referred to as narrow and broad terms. For all SMQs, there is one group of narrow terms that is identical to or synonymous with the medical condition of interest. Every SMQ has at least one group of broad terms that describes symptoms, lab values, or other AEs associated with, but not necessarily specific to, the medical condition. Of the 101 SMQs available in MedDRA version 20.1, there are ten that have an algorithm with multiple categories of broad terms used to increase specificity. To be considered as a match or a positive algorithmic case, AEs coded at the PT level are matched to either a single, narrow term (category A) or a combination of broad terms from two or more term categories (e.g. category B and category C) (Table 1).

The use of algorithmic Standardised MedDRA Queries (aSMQs) is not without risk to the quality of the safety evaluation. Double counting of a positive algorithmic case can occur if the algorithm matches twice for a single patient, such as for a narrow term and also a combination

¹ "MedDRA[®] the Medical Dictionary for Regulatory Activities terminology is the international medical terminology developed under the auspices of the ICH. MedDRA[®] trademark is owned by the International Federation of Pharmaceutical Manufacturers and Associations (IFPMA) on behalf of ICH".

of broad term categories. This may lead to inflation of a safety signal, potentially affecting the user's interpretation. The investigation of Chang et al. revealed that temporal assessment of broad terms (i.e. whether the time frames of different AEs comprising a medical condition were near each other) in positive algorithmic cases was rarely reported, suggesting that those cases could potentially not be relevant [3]. An example of this would be a positive case for the algorithmic SMQ *Anaphylactic reaction* based on a brief episode of wheezing 3 days after study drug administration and urticaria 75 days later. The temporal separation of these events suggests that they might not be part of the same clinical event but is considered positive by the algorithm.

The degree to which aSMQs provide specificity in safety searches has not been explored, with the notable exception of the work of Botsis et al., who compared information retrieval approaches, including certain algorithmic SMQs, to classify cases from the Vaccine Adverse Event Reporting System [4]. Considering the widespread use of the algorithmic SMQ tool, we deemed a thorough evaluation of its performance and a recommendation of meaningful practices to be critically important.

In this study, we report on

- How the number of PTs and incidence are affected by the algorithm and temporal assessment.
- The overlap of narrow term and algorithmic broad terms that may lead to overestimation of event incidence.
- The average start day and duration of events in algorithmic cases to demonstrate how use of the algorithmic system without inspection of individual cases could lead to erroneous event rates.
- What terms are most commonly used in algorithmically positive cases.
- Which therapeutic drug classes are associated with positive signals.

2 Methods

We evaluated seven of the ten SMQs with algorithms: SMQ *Acute pancreatitis*, SMQ *Anticholinergic syndrome*, SMQ *Anaphylactic reaction*, SMQ *Eosinophilic pneumonia*, SMQ *Neuroleptic malignant syndrome*, SMQ *Systemic lupus erythematosus*, and SMQ *Tumour lysis syndrome*. We did not include SMQ *Generalised convulsive seizures following immunisation* and SMQ *Hypotonic-hyporesponsive episode* because of the limited situations and populations. We also did not include the SMQ *Drug reaction with eosinophilia and systemic symptoms*, which was produced in March 2016 (MedDRA version 19.0), because we used the MedDRA dictionary

version associated with the dataset for most of our analyses, and these predated this SMQ. First, we compiled a database of AEs from New Drug Application (NDA) and Biological License Application (BLA) electronic datasets that included AE terms at the PT level. AEs where the start or end day was missing or could not be calculated from relevant information were excluded from the database and were not part of the analysis. We used a program called MedDRA Adverse Event Diagnostics (MAED) to identify SMQ cases and obtain the AE term categories [5]. The dictionary version coded in each dataset was used in MAED. If the MedDRA version was missing from the datasets, the most current version (20.1) was used. In addition to the AEs and the variables of interest, a de-identified patient code, NDA or BLA tracking number, SMQ name, treatment type (test, active, or placebo), and the full seven-character Anatomical Therapeutic Chemical (ATC) class (https://www.whocc.no/atc_ddd_index/; last accessed 11/01/17) were recorded. This database provided SMQ cases for the seven medical conditions without applying the algorithm. A case is defined as any patient with any AE(s) of the SMQ. A case is considered algorithmically positive if it matches the algorithm of an aSMQ. A term is defined as an AE that is coded at the PT level in MedDRA.

Two reviewers independently analysed each case to determine whether algorithms were met or 'positive' for each SMQ. Any one term that matches from the list of terms in category A (narrow) in each SMQ is considered a positive case. Broad terms were evaluated using the algorithm outlined in the MedDRA version 20.1 (Table 1). A positive case occurred if the algorithms were met for the respective SMQ. This subset of data provided positive cases of SMQ after applying the algorithm.

Finally, for cases identified by the algorithm as positive, the same reviewers independently inspected each patient's AEs to determine whether the broad term categories were temporally overlapping (taSMQ+) or not (taSMQ-). Both reviewers adjudicated all discordant cases. Patients who were identified by meeting the algorithm through the narrow term, or category A, were not included in the temporal assessment procedure. This subset of data provided positive algorithmic cases of SMQ with the temporal assessment.

2.1 Statistical Methods

Descriptive statistics were generated for applications, patients, and AEs. Percentages of narrow terms that overlapped with positive algorithmic cases and temporal cases were calculated. Incidences of AEs for each SMQ were calculated. All data were entered in Microsoft Excel 2016 and analysed using JMP 13.1 (SAS Institute Inc.).

Table 2 Incidence and preferred term count for the algorithmic SMQs

	SMQ			SMQ+ algorithm			SMQ+ algorithm+ temporal		
	AEs, <i>n</i>	Patients w/ AEs, <i>n</i>	Incidence, %	AEs, <i>n</i>	Patients w/ AEs, <i>n</i>	Incidence, %	AEs, <i>n</i>	Patients w/ AEs, <i>n</i>	Incidence, %
Total	27,7430	103,928	29.65	22,496	5630	1.61	15,310	3489	1.00
Acute pancreatitis	81,874	46,249	13.19	3506	920	0.26	2341	605	0.17
Anaphylactic reaction	58,493	32,393	9.24	16,626	4309	1.23	10,937	2446	0.70
Anticholinergic syndrome	63,326	41,923	11.96	1009	208	0.06	967	201	0.06
Eosinophilic pneumonia	16,620	10,975	3.13	367	204	0.06	301	181	0.05
Neuroleptic malignant syndrome	34,805	24,896	7.10	167	41	0.01	52	18	0.01
Systemic lupus erythematosus	18,793	8415	2.40	33	20	0.01	33	20	0.01
Tumour lysis syndrome	3519	2060	0.59	788	166	0.05	679	128	0.04

AE adverse events, SMQ Standardised MedDRA Query, w/ with

3 Results

3.1 Database Characteristics

A total of 107 applications, including 96 NDAs and 11 BLAs, were used to gather AEs, which included 350,550 patients at risk (defined as those who received any treatment), and of those patients at risk, 233,318 patients experienced at least one AE, accounting for 1,359,736 AEs.

After using the MAED program to screen for AEs for each of the seven SMQs, the dataset with SMQ positive cases contained 103,928 patients with 277,430 AEs (Table 2). The median number of patients per application in our SMQ dataset was 681 [interquartile range (IQR) 1035], and the median number of AEs was 1256 (IQR 2783; range 23–33,055). Drug therapeutic class, represented by the ATC code, included the following (ATC1 code, *n*): alimentary tract and metabolism (A, *n* = 11), blood and blood forming organs (B, *n* = 2), cardiovascular system (C, *n* = 4), dermatologicals (D, *n* = 3), genito-urinary system and sex hormones (G, *n* = 3), systemic hormonal preparations (H, *n* = 1), anti-infectives for systemic use (J, *n* = 15), antineoplastic and immunomodulating agents (L, *n* = 13), musculoskeletal system (M, *n* = 3), nervous system (N, *n* = 40), respiratory system (R, *n* = 4), and various (V, *n* = 1). Seven applications had drugs that were not assigned ATC codes.

3.2 Preferred Term and Incidence Reduction in Algorithmic Positive and Temporally Overlapping Cases

The number of AEs in the dataset containing SMQ positive cases was condensed to 22,496 and then 15,310 using the algorithm and temporal overlap requirements, reflecting a reduction of about 92 and 95%, respectively (Table 2).

Similarly, the incidence rates decreased to 1.61 and 1.00% after applying the algorithm and temporal overlap, respectively. Figure 1 depicts the reduction in PTs for each SMQ.

3.3 Duplication of Narrow Terms and Algorithmically Matched Cases

SMQ *Acute pancreatitis* had the highest proportion of overlap of narrow term and positive algorithmic cases at 76.1% (Table 3). This proportion was further decreased to 54.2% after assessing the temporal relationship. For SMQ *Anaphylactic reaction*, the proportion of overlap between narrow term and positive algorithmic cases was 69%, which further decreased to 39.7% in temporally positive cases. There were no overlapping cases for SMQ *Eosinophilic pneumonia*, despite a relatively large number of narrow terms and positive algorithmic cases. For both SMQs *Tumour lysis syndrome* and *Neuroleptic malignant syndrome*, the proportion was 50 and 30.8%, respectively, in positive algorithmic cases. There were no changes after using the temporal assessment. For the SMQs *Anticholinergic syndrome* and *Systemic lupus erythematosus* there were no overlapping cases because there were no narrow term cases for SMQ *Anticholinergic Syndrome* and no positive algorithmic cases for SMQ *Systemic Lupus Erythematosus*.

3.4 Patterns of Adverse Event Onset and Duration

The average start day of AEs by broad term category for cases where the algorithm was positive, where the algorithm was not positive, and after assessing term category overlap in the cases where the algorithm was positive is demonstrated in Table 4. Compared to the average of all positive algorithmic cases, cases that also had temporal overlap of the term categories had an earlier average start

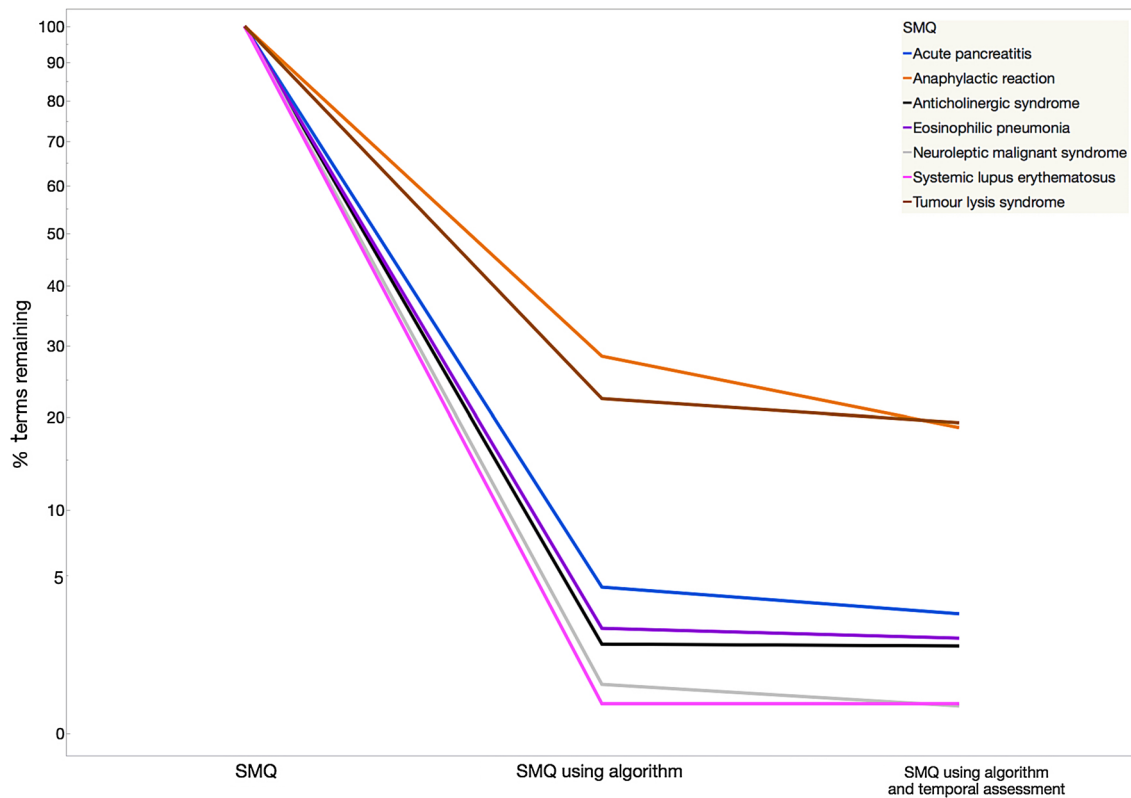


Fig. 1 Reduction of SMQ preferred terms through application of the algorithm and by requiring temporal overlap of term categories. The y axis is plotted on a $y^{0.5}$ power scale to improve visualisation of the changes at the lower end of the scale. *SMQ* Standardised MedDRA Query

Table 3 Percentages of overlap between narrow term and positive algorithmic case and temporally positive cases

SMQ	Narrow terms, <i>n</i>	Overlap of narrow term and positive algorithmic cases, <i>n</i> (%)	Overlap of narrow term and temporally positive cases, <i>n</i> (%)
Acute pancreatitis	310	236 (76.1)	168 (54.2)
Anaphylactic reaction	126	87 (69)	50 (39.7)
Anticholinergic syndrome	0	–	–
Eosinophilic pneumonia	198	0 (0)	0 (0)
Neuroleptic malignant syndrome	13	4 (30.8)	4 (30.8)
Systemic lupus erythematosus	22	0 (0)	0 (0)
Tumour lysis syndrome	10	5 (50)	5 (50)

SMQ Standardised MedDRA Query

date of AEs. This was especially notable for the SMQ *Neuroleptic malignant syndrome*. The term categories B, C, and D in positive algorithmic cases had an average start day of 127, 109, and 117, respectively, and 80, 56, and 27 in positive algorithmic cases with temporal overlap of broad term categories. Conversely, compared to the average of all positive algorithmic cases, term categories in the positive algorithmic cases where there was not temporal overlap of the broad term categories had a later average start day by at least 25 days.

The average duration of AEs by broad term category for the respective SMQ based on whether the algorithm was positive and, if so, whether the term categories were overlapping is demonstrated in Table 5. The AEs in most term categories in the positive algorithmic cases with temporal overlap had an average AE duration that was longer than that for those that were algorithmically positive, alone. One exception to this was the term category C (rigidity terms) in SMQ *Neuroleptic malignant syndrome*, where the average duration was 19 days in the positive algorithmic cases with

Table 4 The average start date for adverse events by broad term categories

	Broad term categories		
	B	C	D
SMQ with algorithm			
Acute pancreatitis	187 (287)	128 (212)	–
Anaphylactic reaction	176 (229)	139 (212)	147 (258)
Anticholinergic syndrome	101 (172)	167 (289)	131 (217)
Eosinophilic pneumonia	164 (222)	231 (279)	–
Neuroleptic malignant syndrome	127 (179)	109 (149)	117 (160)
Tumour lysis syndrome	166 (218)	165 (234)	–
SMQ with algorithm + temporal assessment (positive cases)			
Acute pancreatitis	136 (202)	121 (179)	–
Anaphylactic reaction	151 (203)	130 (204)	124 (187)
Anticholinergic syndrome	102 (175)	161 (279)	128 (202)
Eosinophilic pneumonia	56.3 (43.8)	218 (283)	–
Neuroleptic malignant syndrome	80.4 (111)	55.6 (50.7)	27.2 (30.6)
Tumour lysis syndrome	160 (203)	154 (218)	–
SMQ with algorithm + temporal assessment (non-positive cases)			
Acute pancreatitis	261 (368)	142 (267)	–
Anaphylactic reaction	223 (265)	161 (229)	172 (318)
Anticholinergic syndrome	90.3 (119)	276 (442)	187 (430)
Eosinophilic pneumonia	236 (263)	257 (274)	–
Neuroleptic malignant syndrome	147 (201)	132 (170)	152 (176)
Tumour lysis syndrome	234 (339)	259 (326)	–

All values are expressed as day (standard deviation)

SMQ Standardised MedDRA Query

temporal overlap and 28 days in the positive algorithmic cases. In most of the positive algorithmic cases where there was no overlap of the term categories, the term categories had an average AE duration that was shorter than that for the positive algorithmic group as a whole. The most notable example was SMQ *Anaphylactic reaction* where the categories of B (upper airway/respiratory) and C (angioedema/urticaria/pruritus/flush) had a much shorter average duration than the term categories in the positive algorithmic cases.

3.5 Observations from Individual SMQs

3.5.1 SMQ Acute Pancreatitis

Application of the algorithm reduced the PTs to 4.3% of the SMQ terms and then to 1.9% after use of the temporal requirement (Fig. 1). The incidence rates were 0.26 and 0.17%, respectively (Table 2). Of the 605 patients who were positive for temporal overlap, the most frequent pattern was for terms related to laboratory values and nausea. A third of

the positive algorithmic cases were driven by narrow terms. The most common ATC 3 classes associated with these cases were from L01C *Plant alkaloids and other natural products*, which contains naturally derived chemotherapeutic agents, and L01X *Other antineoplastic agents*, which contains monoclonal antibodies and protein kinase inhibitors (Table 6).

3.5.2 SMQ Anaphylactic Reaction

Application of the algorithm reduced the PTs to 28% of the SMQ cases, with an additional reduction to 18.3% with the requirement for temporal overlap of term categories (Table 2). The incidence rates were 1.23 and 0.7%, respectively. The most common positive algorithmic cases were a combination of broad terms, typically from the B (upper airway/respiratory) and D (cardiovascular/hypotension) term categories. The most common cases were from the ATC 3 classes L01X and A16A *Other alimentary tract and metabolism products*. The latter class includes enzymes involved in metabolism.

3.5.3 SMQ Anticholinergic Syndrome

The use of the algorithm decreased the PTs to 1.59% of the number of SMQ positive terms and then to 1.53% after application of the temporal overlap analysis (Table 2). The incidence rates were similar between the positive algorithmic cases and temporally overlapping cases, at about 0.06% for both. The most common drug classes associated with aSMQ positive cases were in patients taking drugs from the ATC 3 classes N02B *Other analgesics and antipyretics*, which includes non-opioid analgesics, and A16A, where the signal emanated primarily from recombinant enzymes used to treat inborn errors of metabolism.

3.5.4 SMQ Eosinophilic Pneumonia

The algorithm reduced the PTs to 2.2% and then to 0.2% after requiring temporal overlap of the AEs (Table 2). The incidence rates were similar between the positive algorithmic cases and temporally overlapping cases at about 0.05%. Positive algorithmic cases were primarily driven by narrow terms, specifically *pneumonitis*, which appeared 111 times out of the 181 positive cases. The L01X and L02B *Hormone antagonists and related [chemotherapeutic] agents* ATC 3 classes were the most prevalent.

3.5.5 SMQ Neuroleptic Malignant Syndrome

The algorithms reduced the PTs to 0.5% and then to 0.07% after requiring temporal overlap of the events (Table 2). The incidence rates were both approximately 0.01%. There were

Table 5 The average duration for adverse events by broad terms categories

	Broad term categories		
	B	C	D
SMQ with algorithm			
Acute pancreatitis	38.6 (75.7)	22.4 (52.6)	–
Anaphylactic reaction	75.5 (160)	89.9 (178)	17.2 (81.8)
Anticholinergic syndrome	40.5 (130)	32.3 (71.1)	35.5 (98.6)
Eosinophilic pneumonia	90.9 (191)	35 (104)	–
Neuroleptic malignant syndrome	3.55 (4.36)	27.6 (41.9)	33.3 (60.9)
Tumour lysis syndrome	31.6 (74.8)	41.4 (92.1)	–
SMQ with algorithm + temporal assessment (positive cases)			
Acute pancreatitis	36.6 (71.3)	25.8 (56.2)	–
Anaphylactic reaction	94.4 (181.1)	111.9 (201.7)	24.6 (108.7)
Anticholinergic syndrome	41.2 (133.2)	30.6 (70.7)	35.3 (100.2)
Eosinophilic pneumonia	165.9 (284.6)	44.52 (125.4)	–
Neuroleptic malignant syndrome	3.7 (3.7)	19 (35.9)	41.4 (61.9)
Tumour lysis syndrome	32.7 (78)	43.2 (96.9)	–
SMQ with algorithm + temporal assessment (non-positive cases)			
Acute pancreatitis	41.5 (81.9)	15.5 (43.6)	–
Anaphylactic reaction	38.6 (97.4)	34.3 (69.1)	8.8 (26.8)
Anticholinergic syndrome	28.6 (45.6)	64.4 (73.8)	39.9 (61.8)
Eosinophilic pneumonia	40.8 (50.2)	16.6 (27.1)	–
Neuroleptic malignant syndrome	3.48 (4.7)	31.2 (44.3)	30.2 (60.9)
Tumour lysis syndrome	19 (15.7)	27 (32.8)	–

All values are expressed as day (standard deviation)

SMQ Standardised MedDRA Query

only 18 positive algorithmic cases after the temporal assessment. Out of 13 narrow terms, *serotonin syndrome* was the most prevalent term. The drug class most frequently associated with aSMQ positive cases came from the ATC 3 classes N04B *Dopaminergic agents*, where the signal was driven by patients taking the monoamine oxidase B inhibitors, and N05A *Antipsychotics*, where the signal emanated predominantly from patients taking indole derivatives, as well as obesity, monoaminergic stimulatory drugs and direct-acting antivirals.

3.5.6 SMQ Systemic Lupus Erythematosus

The weighted algorithm reduced PTs to about 0.01% relative to the number of SMQ positive terms (Table 2). All PTs were from the narrow term (category A); thus temporal assessment of broad terms was not evaluated. The most common narrow terms were *systemic lupus erythematosus*, followed by *cutaneous lupus erythematosus* and *lupus nephritis*. The cases were predominantly from the L01X, B01A, and L04A *Immunosuppressants* ATC 3 classes.

3.5.7 SMQ Tumour Lysis Syndrome

The algorithms reduced the PTs to 22.4% and then to 18.4% after requiring temporal overlap of the events (Table 2). The incidence rates were both approximately 0.05 and 0.04%, respectively. The cases were mainly from ATC 3 classes L01X, where the signal primarily emanated from patients taking the protein kinase inhibitors and monoclonal antibodies, and B01A *Antithrombotic agents*, with a signal that was from patients taking direct thrombin inhibitors.

4 Discussion

This study provides useful insight in a time of increasing aSMQ use as a tool for detection of safety signals in clinical trial data. A database of 350,550 patients at risk from 107 approved drugs, of which about 103,928 patients experienced 277,430 AEs, represents one of the largest collections of controlled trial data to be evaluated in this level of detail. This data represents between 5 and 10% of the FDA's drug approvals between 2000 and 2017. In

Table 6 Safety signal by drug ATC 3 class for each SMQ detected with temporal assessment

SMQ	ATC 3 class	Patients, <i>n</i>	Patients at risk, <i>n</i>	Rate per 1000 patients
Acute pancreatitis	L01C—plant alkaloids and other natural products	31	1978	15.7
	L01X—other antineoplastic agents	182	13,895	13.1
	J05A—direct acting antivirals	67	13,404	5
	L02B—hormone antagonists and related agents	8	1680	4.76
	A07D—antipropulsives	10	2473	4.04
Anaphylactic reaction	L01X—other antineoplastic agents	1161	13,895	83.6
	A16A—other alimentary tract and metabolism products	18	273	65.9
	L01C—plant alkaloids and other natural products	55	1978	27.8
	L03A—immunostimulants	1	54	18.5
	J05A—direct acting antivirals	245	13,404	18.3
Anticholinergic syndrome	N02B—other analgesics and antipyretics	6	1064	5.64
	A16A—other alimentary tract and metabolism products	1	273	3.66
	L01X—other antineoplastic agents	38	13,895	2.73
	L01C—plant alkaloids and other natural products	4	1978	2.02
	N05A—antipsychotics	43	23,100	1.86
Eosinophilic pneumonia	L01X—other antineoplastic agents	59	13,895	4.25
	L02B—hormone antagonists and related agents	5	1680	2.98
	L01C—plant alkaloids and other natural products	3	1978	1.52
	N04B—dopaminergic agents	4	4434	0.9
	L04A—immunosuppressant	6	7805	0.77
Neuroleptic malignant syndrome	N04B—dopaminergic agents	2	4434	0.45
	N05A—antipsychotics	7	23,100	0.3
	A08A—antiobesity preparations, excl. diet products	1	4000	0.25
	N06A—antidepressants	2	12,699	0.16
	J05A—direct acting antivirals	1	13,404	0.07
Tumour lysis syndrome	L01X—other antineoplastic agents	508	13,895	36.6
	B01A—antithrombotic agents	6	2977	2.02
	J05A—direct acting antivirals	20	13,404	1.49
	C10A—lipid modifying agents, plain	7	7192	0.97
	G03C—estrogens	3	4569	0.66

ATC Anatomical Therapeutic Chemical, SMQ Standardised MedDRA Query

keeping with the objectives of this study, we believe our results are adequate to make some generalised observations on the performance of this tool.

4.1 General Performance of Algorithmic SMQs to Detect Safety Signal

Evaluation of the seven algorithmic SMQs demonstrates strengths of this tool and prompts several suggestions to improve the detection of safety signals. Given their role in capturing potential relevant cases in a clinical trial, aSMQs, such as SMQ *Anaphylactic reaction* and SMQ *Acute pancreatitis* seem to strike a balance between improving specificity and reducing the likelihood of missing a signal. Others, where the PTs were reduced to 0.1%, seem overly restrictive given that

aSMQ is frequently used as a screening tool [3]. In this case, it may be beneficial to also check cases detected by SMQ alone. Botsis et al. also noted that the elevated specificity of certain algorithmic SMQs came at a cost of the sensitivity in their assessment of case retrieval and classification methods [4].

4.2 Overestimation of Safety Signal

Our result suggest that the narrow terms regularly overlapped with positive cases where broad terms satisfy the algorithm, indicating the possibility of double counting that could lead to inflation of AEs. Typically, sponsors will code known syndromes with narrow terms if the diagnosis is apparent, while other AEs will be coded at the broad term

level. This is an issue for which education of sponsors would have a profound effect.

4.3 Average Start Day and Duration of Adverse Events and the Temporal Overlap of Broad Terms

Our analysis of the average start date and average duration of AEs suggests that incorporation of a temporal assessment is an important step in using this screening tool. We noted that the longer the duration of AE, the more likely it is to overlap with other term categories, which may lead to more positive cases. There may be occasions where the need for assessing temporal overlap is not so clear or appropriate and a case definition should be defined before the evaluation of AEs.

4.4 Performance of Individual SMQ

In the following sections, observations germane to specific aSMQs are described that relate to the general commentary from the preceding section.

4.4.1 SMQ *Acute Pancreatitis*

This algorithm provides a notable reduction of PTs (Table 2). The requirement for temporal overlap helps to eliminate AEs with an extreme average start day. A sizable portion of positive algorithmic cases was identified based on narrow terms ($n = 310$), possibly because the diagnosis is intuitive once an objective result such as an abnormal amylase or lipase result is obtained. The list of drugs associated with drug-induced pancreatitis is expanding at a remarkable rate, so a positive safety signal in most drug classes should not be surprising [6, 7]. Clinical trials that study oncology and antiviral drugs may expect more cases of acute pancreatitis based on our ATC class evaluations.

4.4.2 SMQ *Anaphylactic Reaction*

This SMQ provides the least reduction of PTs and incidence of all those evaluated in this study (Table 2). One reason may be the surprisingly long duration of many events leading to a high rate of overlapping term categories (Table 5). This was also shown by our analysis on average duration of AEs where most of the term categories in the positive algorithmic cases with temporal overlap had a much longer average duration than those positive algorithmic cases where the broad term categories did not overlap. Most of the positive algorithmic cases were through combinations of broad terms without much contribution of narrow terms. This is surprising since the clinical presentations of drug-induced anaphylactic reaction often commences with a characteristic and concomitant expression of cardiac and pulmonary

events with dermal effects sometimes being delayed [8–10]. In some scenarios, such as the setting of surgery under general anaesthetic, the clinician may recognise the onset of anaphylaxis through the sudden onset of tachycardia, hypotension or elevated airway inspiratory pressures before all of the component events have manifested [11].

4.4.3 SMQ *Anticholinergic Syndrome*

The reduction of PTs when the algorithm is applied seems reasonable for this SMQ (Table 2). A difficult aspect with respect to this medical condition to reconcile with is the use of an algorithmic retrieval method because of its highly variable presentation. Two reviews noted the highly inconsistent presentation of patients and the unreliable combination of symptoms [12, 13]. The lack of cases identified through narrow terms seems to reflect the difficulty of diagnosing anticholinergic syndrome during the clinical event.

4.4.4 SMQ *Eosinophilic Pneumonia*

This SMQ seems to derive the greatest reduction in incidence rate in the positive algorithmic cases following application of the temporal overlap assessment (Table 2). The most frequently used term leading to positive algorithmic cases was the narrow term *pneumonitis* ($n = 111$). This term and its role in the algorithm seem like a candidate for reevaluation since it seems to be more of a broad term than a phrase specifically synonymous with eosinophilic pneumonia. This form of pneumonia is well understood such that a more specific algorithm could be generated [14–16].

4.4.5 SMQ *Neuroleptic Malignant Syndrome*

There were few positive algorithmic cases with temporal overlap of broad terms (Table 2). The most commonly used term leading to positive algorithmic cases was the narrow term *serotonergic syndrome*, which is considered clinically different from neuroleptic malignant syndrome [17, 18]. This narrow term, *serotonergic syndrome*, may be included in a medical differential diagnosis but is not synonymous with neuroleptic malignant syndrome, and so the terms in this SMQ should be re-evaluated.

4.4.6 SMQ *Systemic Lupus Erythematosus*

Application of the algorithm produced the most profound reduction of PTs and incidence of any of the SMQs evaluated (Table 2). This SMQ is different from others in that the algorithm determination is based on a weighted point system for different term categories rather than the algorithmic system used by the other six aSMQs evaluated in this study. Surprisingly, all of the positive algorithmic cases came from the

narrow terms and none were from the combination of broad terms in the weighted algorithm. This weighted algorithm may be too restrictive to be used as a screening tool and should be re-evaluated. Screening cases without employing the broad term component of the algorithm may be useful in this case.

4.4.7 SMQ Tumour Lysis Syndrome

This SMQ resulted in a notable reduction in cases after using the algorithm, yet the number of PTs was among the least reduced of all of the SMQs (Table 2). The most common ATC class associated with this medical condition in our analysis was the L01X *Other antineoplastic agents* class, which was reasonable given the nature of this therapeutic class. Other classes of drugs that are not commonly associated with tumour lysis syndrome can still have a safety signal through direct nephrotoxic effects [19]. This was shown in our results, where some of the drugs in the ATC class C10A *Lipid modifying agents*, which were associated with renal toxicity, had several positive algorithmic cases.

4.5 Limitations

Our study had several limitations. First, we relied on datasets from clinical trials submitted to the FDA and cannot account for missing or incomplete information. Because we sought to eliminate this bias by excluding any subjects with missing pertinent variables, our results are more likely to be an underestimation. Second, SMQ is a screening tool. Obtaining algorithmically positive cases, or even simply SMQ positive cases, does not necessarily indicate that one has found a true case of the medical condition of interest. Verification is needed to confirm the safety signal for those cases that were positively identified by SMQs, a point discussed by Chang et al. [3]. Because this was not performed for the cases in this study, we do not know the accuracy of the positive cases. Third, the number of AEs was not uniform across the different SMQs. For example, SMQ *Acute pancreatitis* had about 80,000 AEs in the SMQ positive group, while SMQ *Tumour lysis syndrome* only had 3500 AEs. Therefore, it may be more challenging to detect differences or observe trends, if any existed, in the performance of the aSMQs in cases where the number of AEs left after application of the SMQs is very low. Finally, our findings are derived from premarket data, so they may not all be generalisable to other areas where SMQs are used, such as the detection of safety signals from post-marketing reports.

5 Conclusion

Our results support the use of algorithm and temporal assessment to improve specificity when retrieving appropriate cases and to reduce overestimating of event incidence.

Generally, the scope of adjusting between sensitivity and specificity will differ between those who are using this tool to screen safety signals from new drugs and those using it for signal strengthening and evaluation of specific conditions. Currently, algorithmic SMQs do not assess for a temporal relationship, but this should be considered for some of them. The algorithm may not be useful as a screening tool in some cases if the results are too restrictive. Evaluating the onset and duration of AEs may be another approach for users to retrieve relevant cases for further verification. A better understanding of the pattern of onset AEs can help provide guidance for case retrieval for safety signal detection. As the popularity of SMQs used to screen for safety signals grows, we recommend revisiting and refining their term lists.

Acknowledgements We thank Ms. Joean James and Ms. Gretchen Kindle for their administrative support.

Compliance with Ethical Standards

Conflict of interest Carolyn Tieu has no conflicts of interest that are directly relevant to the content of this study. Christopher Breder is a medical officer at the US Food and Drug Administration (FDA); he is also the FDA Topic Expert in the M1 Points to Consider Group.

Funding Dr. Tieu's fellowship was funded by an Oak Ridge Institute Science and Education Grant # RSR-702. The funding source had no role in the design and conduct of the study, analysis or interpretation of the data, and preparation or final approval of the manuscript prior to publication.

Disclaimer The views expressed in this paper are the authors' views and do not necessarily represent those of the FDA or the ICH.

References

1. MedDRA Maintenance and Support Services Organization. Coding with MedDRA found at the URL https://www.meddra.org/sites/default/files/training/file/000133_coding_with_meddra_f2f_course_7.pdf. Accessed 15 Jun 2018.
2. MedDRA Maintenance and Support Services Organization. Introductory guide for Standardised MedDRA Queries (SMQs) Version 20.1. Accessed 15 Jun 2018. In: The International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use, editor. Geneva, Switzerland 2017.
3. Chang LC, Mahmood R, Qureshi S, Breder CD. Patterns of use and impact of standardised MedDRA query analyses on the safety evaluation and review of new drug and biologics license applications. *PLoS One*. 2017;12(6):e0178104.
4. Botsis T, Woo EJ, Ball R. Application of information retrieval approaches to case classification in the vaccine adverse event reporting system. *Drug Saf*. 2013;36(7):573–82.
5. Li XJ, Cooper C, Luo Z. MAED Service: FDA-developed tool for clinical AE data signal detection. *Computational Science Symposium*; March 18th–19th, 2013; Silver Spring, MD 2013.
6. Douros A, Bronder E, Andersohn F, Klimpel A, Thomae M, Ockenga J, et al. Drug-induced acute pancreatitis: results from the hospital-based Berlin case-control surveillance study of 102 cases. *Aliment Pharmacol Ther*. 2013;38(7):825–34.

7. Nitsche CJ, Jamieson N, Lerch MM, Mayerle JV. Drug induced pancreatitis. *Best Pract Res Clin Gastroenterol.* 2010;24(2):143–55.
8. Webb LM, Lieberman P. Anaphylaxis: a review of 601 cases. *Ann Allergy Asthma Immunol.* 2006;97(1):39–43.
9. Bousquet PJ, Kvedariene V, Co-Minh HB, Martins P, Rongier M, Arnoux B, et al. Clinical presentation and time course in hypersensitivity reactions to beta-lactams. *Allergy.* 2007;62(8):872–6.
10. Bircher AJ, Scherer K. Delayed cutaneous manifestations of drug hypersensitivity. *Med Clin N Am.* 2010;94(4):711–25.
11. Michavila Gomez AV, Belver Gonzalez MT, Alvarez NC, Giner Munoz MT, Hernando Sastre V, Porto Arceo JA, et al. Perioperative anaphylactic reactions: review and procedure protocol in paediatrics. *Allergol Immunopathol (Madr).* 2015;43(2):203–14.
12. Patel RJ, Saylor T, Williams SR, Clark RF. Prevalence of autonomic signs and symptoms in antimuscarinic drug poisonings. *J Emerg Med.* 2004;26(1):89–94.
13. Hall RCW, Fox J, Stickney SK, Gardner ER, Perl M. Anticholinergic delirium—etiology, presentation, diagnosis and management. *J Psychedelic Drugs.* 1978;10(3):237–41.
14. Tazelaar HD, Linz LJ, Colby TV, Myers JL, Limper AH. Acute eosinophilic pneumonia: histopathologic findings in nine patients. *Am J Respir Crit Care Med.* 1997;155(1):296–302.
15. Jhun BW, Kim SJ, Son RC, Yoo H, Jeong BH, Chung MP, et al. Clinical outcomes in patients with acute eosinophilic pneumonia not treated with corticosteroids. *Lung.* 2015;193(3):361–7.
16. Giovannini-Chami L, Blanc S, Hadchouel A, Baruchel A, Boukari R, Dubus JC, et al. Eosinophilic pneumonias in children: a review of the epidemiology, diagnosis, and treatment. *Pediatr Pulmonol.* 2016;51(2):203–16.
17. Musselman ME, Saely S. Diagnosis and treatment of drug-induced hyperthermia. *Am J Health Syst Pharm.* 2013;70(1):34–42.
18. Burkhard PR. Acute and subacute drug-induced movement disorders. *Parkinsonism Relat Disord.* 2014;20(Suppl 1):S108–12.
19. Criscuolo M, Fianchi L, Dragonetti G, Pagano L. Tumor lysis syndrome: review of pathogenesis, risk factors and management of a medical emergency. *Expert Rev Hematol.* 2016;9(2):197–208.