Vaccine 37 (2019) 49–55

Contents lists available at ScienceDirect

Vaccine

journal homepage: www.elsevier.com/locate/vaccine

Tree-based scan statistic – Application in manufacturing-related safety signal detection

Olivia Mahaux *, Vincent Bauchau, Ziad Zeinoun, Lionel Van Holle

Vaccine Clinical Safety and Pharmacovigilance, GSK, Waver, Belgium

ARTICLE INFO

Article history:
Received 16 August 2018
Received in revised form 23 October 2018
Accepted 14 November 2018
Available online 22 November 2018

Keywords:
Tree-based scan statistic
Safety signal detection
Manufacturing
Data-mining

ABSTRACT

Background and objectives: Over the last decades, medicinal regulations have been put into place and have considerably improved manufacturing practices. Nevertheless, safety issues may still arise. Using the simulation described in this manuscript, our aim is to develop adequate detection methods for manufacturing-related safety signals, especially in the context of biological products.
Methods: Pharmaceutical companies record the entire batch genealogies, from seed batches over intermediates to final product (FP) batches. We constructed a hierarchical tree based on this genealogy information and linked it to the spontaneous safety data available for the FP batch numbers. The tree-based scan statistic (TBSS) was used on simulated data as a proof of concept to locate the source that may have subsequently generated an excess of specific adverse events (AEs) within the manufacturing steps, and to evaluate the method's adjustment for multiple testing. All calculations were performed with a customized program in SAS v9.2.
Results: The TBSS generated a close to expected number of false positive signals, demonstrating that it adjusted for multiple testing. Overall, the method detected 71% of the simulated signals at the correct production step when a 6-fold increase in reports with AEs of interest (AEOI) was applied, and 31% when a 2-fold increase was applied. The relatively low detection performance may be attributed to the higher granularity associated with the lower levels of the hierarchy, leading to a lack of power and the stringent definition criteria that were applied for a true positive result.
Conclusion: As a data-mining method for manufacturing-related safety signal detection, the TBSS may provide advantages over other disproportionality analyses (using batch information) but may benefit from complementary methods (not relaying on batch information). While the method warrants further refinement, it may improve safety signal detection and contribute to improvements in the quality of manufacturing processes.

1. Introduction

Post-marketing safety signal detection methods, particularly on spontaneous data, have been widely discussed [1–6]. While most of the detection methods focus on the product or active ingredients as being the source of variability, the origin of the safety issue may lie in the product manufacturing. Unlike chemical products, complex biological preparations such as vaccines are more prone to variability during their manufacturing process. Thus, strict quality control monitoring is necessary to ensure that appropriate measures are rapidly taken when an unexpected incident impacts the final product's quality and safety.

While recalls of unsafe vaccine lots due to defective manufacturing processes are rare events [7], the wide-ranging consequences of production failures have been well documented. In 1942, human serum, which was used to stabilize yellow fever vaccine, proved to be contaminated with hepatitis B in a limited number of lots. The vaccine was distributed among military recruits and resulted in cases of delayed jaundice, and a mortality rate of 3 per 1000 individuals [8]. In 1955, the "Cutter incident", caused by vaccination with lots containing improperly inactivated poliomyelitis virus (Cutter Laboratories), led to more severe

Abbreviations: AE, adverse event; AEOI, adverse event of interest; CI, confidence interval; DPA, disproportionality analyses; FP, final product; MedDRA, Medical Dictionary for Regulatory Activities; FPR, false positive rate; ROR, reporting odds ratio; H0, null hypothesis; HLGT, high level group of terms; LR, likelihood ratio; TBSS, tree-based scan statistic.

* Corresponding author at: Vaccine Clinical Safety and Pharmacovigilance, GSK, Avenue Fleming 20, 1300 Waver, Belgium.
E-mail address: Olivia.x.mahaux@gsk.com (O. Mahaux).

https://doi.org/10.1016/j.vaccine.2018.11.044
0264-410X/

disease symptoms than those caused by the natural poliomyelitis virus, resulting in the paralysis of 200 children and 10 deaths [9]. Changes in medicinal regulations have considerably improved manufacturing practices, and no event of that dramatic scope has since occurred for vaccines. However, safety issues continue to arise as exemplified by 2 incidences on biologicals. For treatment of chronic kidney disease, a specific epoetin alfa formulation in uncoated rubber stopper syringes increased the incidence of pure red cell aplasia with a peak during 2002–2003 [10,11]. Moreover, multi-use vial formulation of the erythropoietin analog peginesatide, which included phenol, led to severe anaphylactoid reactions with associated fatalities in 2012 [12,13]. Therefore, it remains essential to further improve passive reporting systems, and to develop adequate manufacturing-related detection methods for safety signals, especially for biological products [14,15].

A commonly used safety signal detection method for manufacturing quality issues focuses on the time trending of adverse events (AEs) with the drawback of generating many false positive signals [16]. As another option, final product (FP) batch numbers, as reported in the spontaneous data, are used to perform disproportionality analyses (DPA) [17,18]. However, the FP batch may not be the right level on which to perform the signal detection as it marks the end of the manufacturing cascade, while variability may be introduced during all processes over the course of its complex production. In the DPA, different FP batches of the same product are compared. For instance, if several FP batches are affected by the same manufacturing defect, introduced at a shared bulk batch, the signal is less likely to be detected with the DPA. To ensure that these defects are spotted, one could apply the DPA on all level batches, although many false positives would be generated as a result of multiple testing.

The tree-based scan statistic (TBSS) may offer a possible solution to the challenges illustrated above [19]. With the TBSS, the independent variable needs to be defined as a hierarchical tree. The method adjusts for multiple testing which is inherent in the large number of potential combinations. Surveillance may thus be performed without requiring preconceived ideas about the causative relationship between a specific element or group of elements of the tree, and the increased risk. Pharmaceutical companies track and record the entire batch genealogies, from seed batches over intermediates to FP batches. This batch genealogy information can then be displayed in a hierarchical tree structure and be linked to spontaneous safety data that is available for the FP batch numbers. In this manuscript, we use the TBSS as a proof of concept based on simulated data to identify the manufacturing step at which a defect in the batch was most likely introduced, which subsequently generated an excess of AEs. Simulation was required to better control the size and location (within the hierarchy) of excess AEs, in the absence of actual 'true positives' in our records.

2. Methodology

Up to now, no proven link exists between manufacturing incidents of GSK vaccines and patient safety data; manufacturing-related safety signals detected using the DPA at the FP batch level could so far not be traced back to any manufacturing incident. Therefore, as a true positive is lacking, the test could not be performed on "real data". Test scenario simulations were thus performed using real GSK vaccine batch genealogies and spontaneous safety data as part of a realistic framework. The safety information was linked through the FP batch numbers (provided in the spontaneous reports) to manufacturing genealogies as illustrated in Fig. 1.

2.1. The hierarchical tree variable

Batch genealogies of 1 actual vaccine were used to construct a realistic hierarchical tree (as illustrated in Fig. 1) on which the TBSS was applied. The vaccine was selected as it possesses a relatively small amount of valid FP batch numbers reported in the spontaneous data, resulting in relatively simple batch genealogies. "Administration site reactions", as per the Medical Dictionary for Regulatory Activities (MedDRA) high level group of terms (HLGT), was chosen as AEs of interest (AEOI), because this group of terms is well represented in the database, and represents one of the groups of AEOI for manufacturing-related safety signal detection. For each leaf, the FP batch in our case, information regarding the number of cases with the AEOI and the total number of reports were retrieved from the company's vaccine spontaneous safety database (restricted to case reports with valid FP batch numbers). Branches connect related leaves at the same node (final container batch), and each node in turn stems from an above branch and thus connects to a higher-level node, until it ultimately connects to the uppermost node (seed batch). When defining 'batch', we refer to all batches within the hierarchical tree (Fig. 1).

2.2. Tree-based scan statistic

For the TBSS, the scanning windows were defined by all possible cuts on all branches of the tree. Each cut defined a larger or smaller group of related batches that was evaluated with respect to the AEOI reporting. The likelihood ratio (LR) was calculated for each cut under both the null hypothesis and the alternative hypothesis, and maximized over all possible cuts [19].

The analysis was performed using the Bernoulli probability model and conditioned on the total number of cases reporting the AEOI. Details about the likelihood formula can be found in the Supplementary material S1.

The p-values were estimated by Monte-Carlo simulations, for which the number of cases with the AEOI for each FP batch of the random datasets were obtained from a multivariate hypergeometric distribution (Supplementary material S1) [20,21].

The TreeScan software v1.1, available at that time, could not be used to perform the analyses due to the 'many-to-many' relationships between higher-level and lower-level nodes of the hierarchical tree described in the previous paragraph. Therefore, all calculations were performed using a customized program in SAS v9.2. In order to perform independent validation of this SAS code with the TreeScan software v1.1, a simplified version of the manufacturing hierarchical tree was developed to maintain only 'one-to-many' relationships between higher-level and lower-level nodes.

2.3. Simulated test scenarios

To obtain simulated datasets with the lowest possible pre-existing variability in the reporting of AEOI among FP batches, the link between safety data and manufacturing genealogy was broken, and the spontaneous reports were uniformly redistributed among the FP batches. Two different test scenarios were applied on this simulated safety batch hierarchical data. In the first test scenario, 100 different versions of these datasets were included to evaluate the false positive rate (FPR) using either the TBSS or the reporting odds ratio (ROR), with the latter method allowing the estimation of the impact of multiple testing (Fig. 2**A**). The ROR (a measure of association between an exposure and outcome [22]) and associated confidence intervals (CI) were indeed calculated one by one for all the batches in the hierarchy. False positives were defined as at least one signal (i.e. lower limit of the 95% CI > 1 for

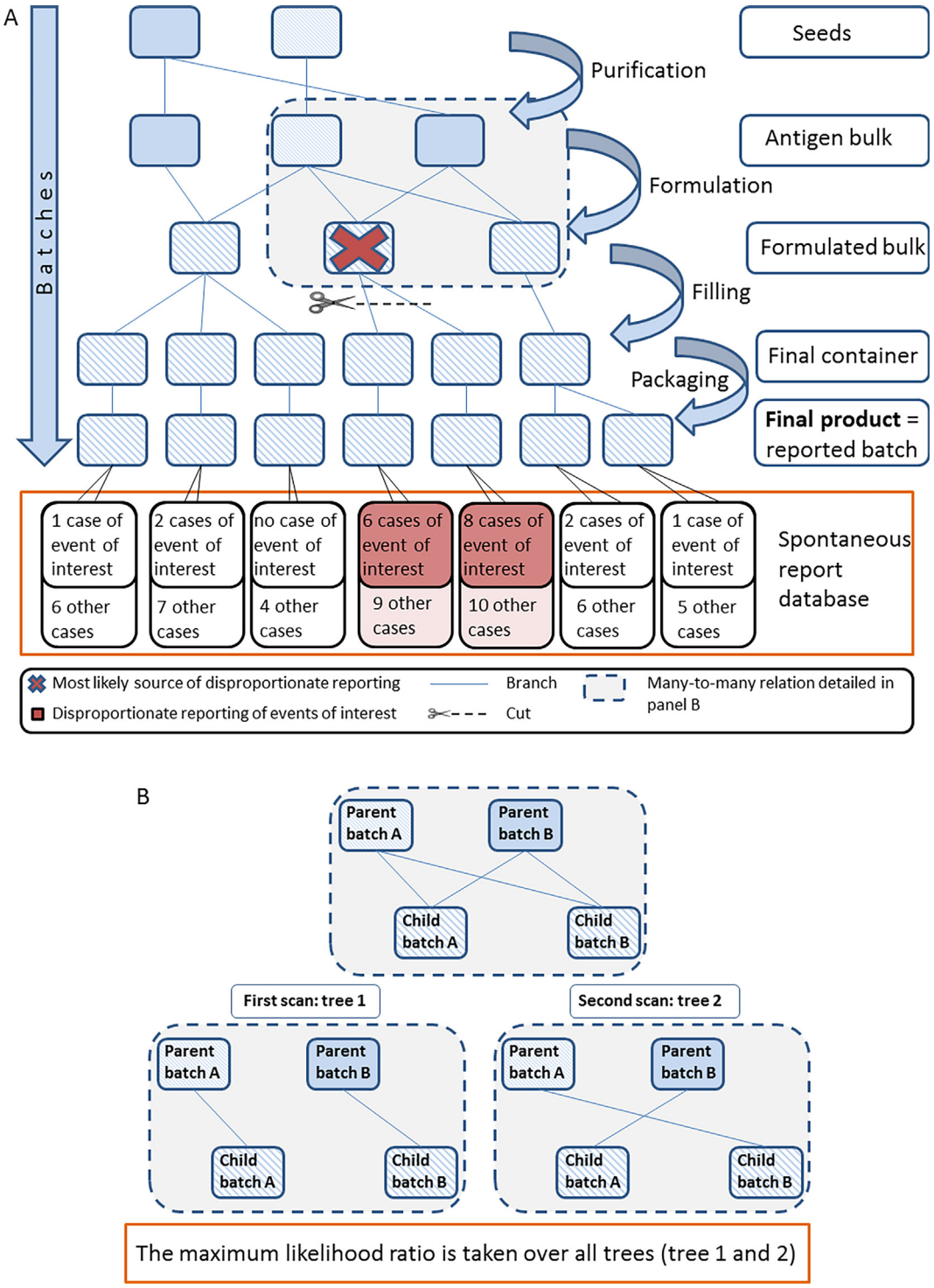

Fig. 1. Simplified scheme of manufacturing batches' genealogies represented as a hierarchical variable and tree-based scan.

the ROR and p-value < 0.05 for the TBSS) per simulated safety-batch hierarchy at any location in the tree. In the second test scenario, the sensitivity and positive predictive value of the TBSS method were assessed (Fig. 2**B**). For this purpose, 100 batches per hierarchical level were randomly selected. Only distinct batches were kept, and for each, the baseline safety data was increased by adding 100–500% (using increments of 100%) of the existing redistributed reports with the AEOI. When the increase had to be performed on higher levels of the hierarchy than the FP, the reports containing the AEOI for the FP batches produced from each randomly selected batch were increased by the same percentage, thus ensuring a uniform increase for the selected higher-level batch. Batches without AEOI cases were excluded from the simulations. A simulated batch detected by the method with a maximum LR and a p-value of <0.05 was considered to be a true positive, while a simulated TBSS-detected batch with a p-value < 0.05 that ranked second in terms of LR was considered to be a false negative.

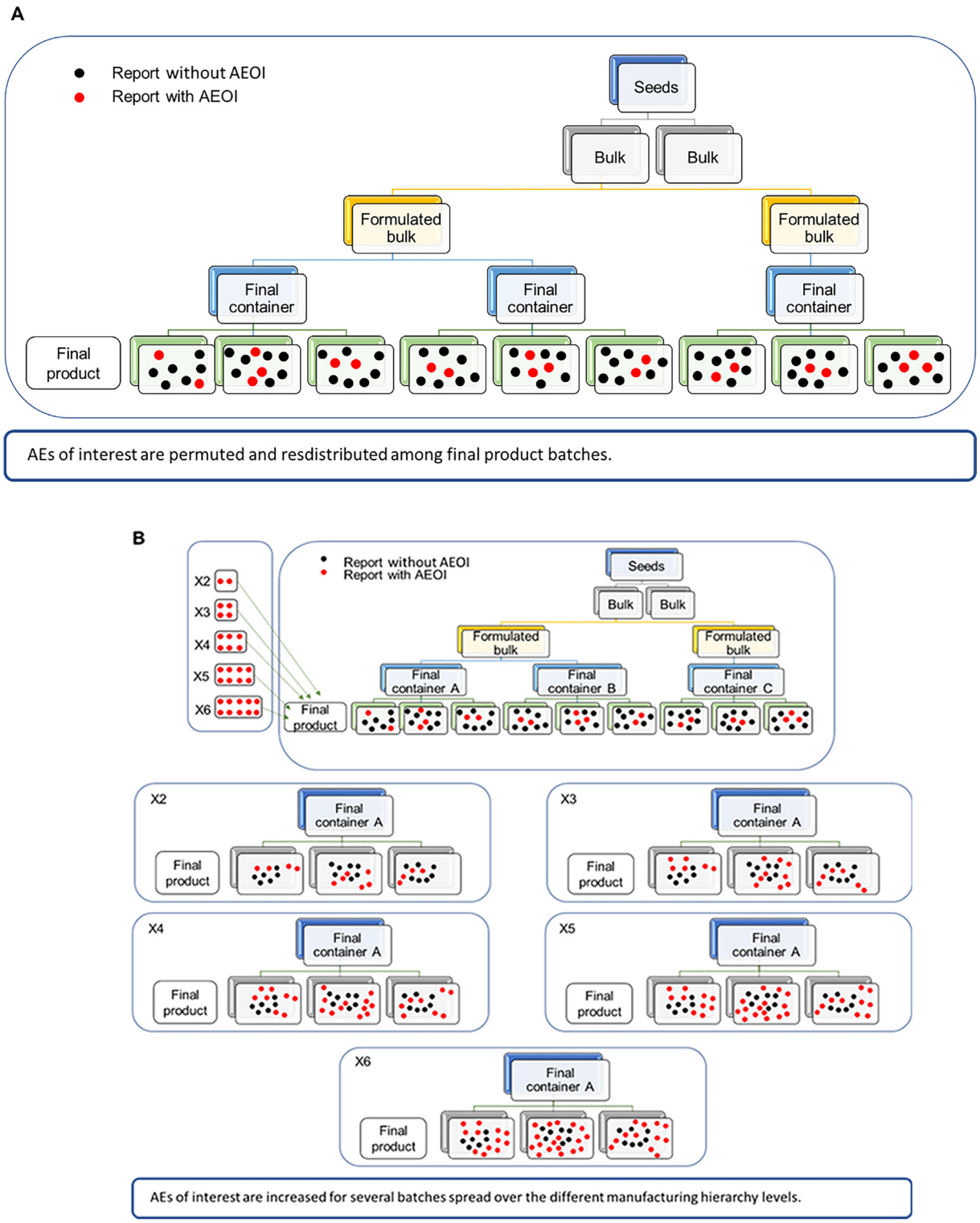

Fig. 2. Two simulated versions of hierarchical data. Abbreviations: AE, adverse event; AEOI, adverse event of interest; X2–X6, X2 to X6 increases.

3. Results

The safety-batch hierarchical data consisted of 3660 vaccine batches spread over 7 hierarchy levels (Table 1). Seven false positives were obtained with the TBSS, while 5 were expected given the significance level of 0.05 used in the present analyses. With the ROR method, a FPR of 100% was obtained (expected 2.5%) (Table 1). A comparison between the TBSS and ROR methods demonstrated a mean of 71.53 batches per simulation that may be signaling with the ROR, and 0.13 with the TBSS method.

Sensitivity results using the simulated safety-batch hierarchical data are displayed in Table 2. Overall, 31% (157/499) of the simulated batches were correctly identified as most likely generating the excess AEOI when the number of AEOI was multiplied by 2, 52% (261/499) when multiplied by 3, 63% (312/499) when multiplied by 4, 67% (335/499) when multiplied by 5 and 71%

Table 1
False positive results obtained with the TBSS and ROR on the first version of simulated hierarchical data.

Hierarchical level	Batches N = 3670, n (%)	TBSS		ROR	
		False positives[a]	Number of batches included in false positives N′ = 13, n (%)	False positives[a]	Number of batches included in false positives N′ = 7153, n (%)
Seeds	25 (1)	0	0 (0)	40	54 (1)
Buffer/media	1763 (48)	1	4 (31)	94	3881 (54)
Bulk	143 (4)	0	0 (0)	94	356 (5)
Intermediate bulk	497 (14)	1	1 (8)	96	1274 (18)
Formulated bulk	156 (4)	0	0 (0)	94	313 (4)
Final container	415 (11)	3	3 (23)	99	754 (11)
Final product	671 (18)	5	5 (38)	100	521 (7)
Overall	3670 (100)	7[b]	13 (100)	100[b]	7153 (100)

Footnote: N, total number of batches; N', total number of false positive batches, n (%), number (percentage) of batches; ROR, reporting odds ratio; TBSS, tree-based scan statistic.
[a] Number of distinct versions 1 of simulated data with at least 1 false positive.
[b] Batches spread over several levels may share the same most likely position with a significant p-value, therefore, the numbers presented by level do not sum up to the Overall.

Table 2
Summary of the results obtained with the TBSS on the second version of simulated hierarchical data.

Hierarchy level	Total number of batches	Number of simulated batches	Number (%) of simulated batches identified as most likely to generate the excess of AEOI				
			X2 increase	X3 increase	X4 increase	X5 increase	X6 increase
Seeds	25	24	18 (75)	21 (88)	21 (88)	21 (88)	21 (88)
Buffer/ media	1763	96	45 (47)	59 (61)	69 (72)	75 (78)	76 (79)
Bulk	143	80	36 (45)	56 (70)	65 (81)	67 (84)	67 (84)
Intermediate bulk	497	87	32 (37)	57 (66)	66 (76)	70 (80)	71 (82)
Formulated bulk	156	69	11 (16)	26 (38)	36 (52)	41 (59)	44 (64)
Final container	415	85	11 (13)	29 (34)	41 (48)	46 (54)	52 (61)
Final product	671	58	4 (7)	13 (22)	14 (24)	15 (26)	21 (36)

Abbreviations: TBSS, tree-based scan statistic; AEOI, adverse event of interest.

(352/499) when multiplied by 6. Overall, the positive predictive value ranged from 69% (for the X2 increase scenario; 157 true positives out of 229 positive results and thus 72 false positives) to 87% (for the X6 increase scenario; 352 true positives out of 403 positive results and thus 51 false positives).

4. Discussion

We propose the TBSS as a tool to scan safety data linked to batches from all hierarchical levels of the manufacturing of a vaccine, and to ultimately identify the most likely batch and level at which an excess of AEOI will have been generated. To our knowledge, this is the first application of the TBSS to detect safety issues potentially originating from the manufacturing of a vaccine. From what has been reported so far, data-mining has only been performed at the FP batch level and not at earlier steps during the production [17,18]. If both the FP batch level and earlier steps are included, multiple testing will result in a higher number of false positives (as with the ROR). Applying the TBSS generated a close to expected number of false positive signals (7 versus 5 expected), demonstrating that it is indeed adjusted for multiple testing. Thus, the TBSS has major advantages over other conventional methods relying on batch numbers: it allows the evaluation of the complete manufacturing hierarchical data at once, the control for multiple testing and the approximate location of the excess AEOI reporting.

The performance results of our model indicate that the higher levels of the batches' hierarchy are associated with a higher detection power as expected from the lower granularity of the safety data when grouped at those levels. Of note, this does not apply for the false positives, which were relatively uniformly distributed among all hierarchy levels. Moreover, we achieved good sensitivity and positive predictive values using the TBSS, especially given the stringent definition criteria used for true positives.

Our simulation method assured the same proportional increase for all cases through the multiplication of existing cases, but this approach also implied that FP batches without AEOI cases were not eligible for multiplication, which resulted in a lower detection probability of the upper level batch producing those FP batches. Furthermore, batches with a higher number of initial case reports were advantaged with this simulation method. However, as an additional asset, the same intended strength of issue simulation was applied to batches from all hierarchical levels.

To ensure that all important safety concerns are captured, all significant non-overlapping batches (p-value < 0.05) that are not related to each other by their genealogy could be evaluated. However, as the identification of all non-overlapping batches, or in particular the ones with a link with their genealogy would represent further challenges, this approach was not taken in the current test scenarios.

It remains questionable whether the uppermost level of hierarchical batch data should be included in the analysis. While the uppermost level batch data account for very few different batch numbers throughout the complete lifecycle of a vaccine, they are used to produce several different lower-level batches spanning over long manufacturing periods, and thus may include more than 50% of all the AEOI. The calculated increased risk for this uppermost level batch may just highlight a lower than expected risk outside the cut as has been suggested by similar discussions regarding temporal scan tests [19].

The TBSS was run on spontaneous safety data known to be compromised by underreporting. As a result, not all simulated FP batches had reported AEs, and many different simulated batches in the hierarchical tree were incorrectly considered to be identical

by the method (even without a direct 'parent-child' link). This mainly affected the buffer/media level, at which 78% of the batches were duplicates. After removing the duplicates, the sensitivity increased on average by 17% for that level, without affecting the specificity.

Accordingly, an automated system should be developed to track the proportion of each batch used to produce subsequent batches, and ultimately the FP batches that are potentially associated with safety signals. This tracking information would allow for a better allocation of the batch in the manufacturing hierarchy, which will have most likely generated the potential signal. Indeed, in a scenario where 2 different batches produce the same FP batch, 5% of a given batch may have been used to produce the FP batch, while 25% of the other batch may have contributed to it.

The current application of the TBSS includes so called 'active-ingredient' batches, and thus already covers major aspects of manufacturing activities. Nevertheless, one may consider including chemicals and other material batches. On the other hand, manufacturing issues may be related to buildings, machines and equipment which are not tracked as a batch but could affect the safety of a product for a given period of time. Until now, other methods, not relying on batch information, such as a temporal scan statistic [23] on vaccination periods are still needed to complement the batch signal detection. The current application of the TBSS relies on a correct FP batch number, which is reported, on average, in 42% of spontaneous reports. Thus, in the GSK safety database, the spontaneous reporting system could still benefit from more awareness surrounding the importance of batch numbers' reporting.

While the TBSS was applied on the batch hierarchical data for only 1 vaccine in this work, the simultaneous analyses of data for more vaccines would include more safety information regarding batches shared across vaccines. This approach could provide more power in detecting potential manufacturing-related safety issues. However, this would be a computationally intensive task, and may introduce confounders due to the differences in the respective study populations who were administered with the vaccines of interest.

4.1. Future refinements of the TBSS

While the TBSS may serve as a powerful tool to locate the source batch in the manufacturing of a vaccine, which most likely causes an excess of AEOI cases, this method requires several refinements to allow for a broader application. For instance, one might consider assessing the impact of loosening the true positive definition criteria used for the performance of the TBSS. Furthermore, our performance results may also depend on the selected vaccine. Indeed, another vaccine may have a different number of FP and/or hierarchical levels and/or batches per hierarchical level. As a result, caution should be exercised when generalizing in this context, as the performance may also further depend on the respective chosen AEOI. In the present analysis, a common event was compared to the total number of reports which resulted in a relatively high probability of the event under the null hypothesis (H0). However, cases with AEOI may refer to a particular MedDRA preferred term, MedDRA HLGT or any other group of events with high variability in their H0 probability. In addition, the impact of a manufacturing issue on the spontaneous reporting is an unknown factor, and the simulations performed (X2 to X6 increases) may not represent a realistic situation.

In the currently developed version of the TBSS, adjustment for confounders, such as country or time, has not been implemented. Variations in the passive reporting system in different countries may result in an artificial excess of AEOI when data derived from several countries are pooled together, and no adjustment is performed. Proper evaluation is also warranted with respect to the different batches that are detected with equal likelihood, irrespective of the presence of a 'parent-child' link.

The TBSS is adjusting for the multiple testing resulting from the evaluation of all possible cuts of the tree. For a signal detection purpose, those analyses may be repeated over time and adjustment for the multiple testing during the repeated analyses should be implemented in future TBSS as suggested by Kleinman et al. [24]. It has not been assessed how fast an emerging signal can be detected using the TBSS. It is presumed that consumption of some batches is initially required which might consequently result in a detection delay.

As a further refinement, the TBSS may be performed on MedDRA hierarchical data and batch hierarchical data simultaneously. In this scenario, more combinations will result in a computationally intense approach. In a more complex application and with proper adjustment for confounders, the TBSS may be applied for signal detection purpose on a hierarchical tree built from vaccines and their different components (i.e. adjuvant, excipients etc.) and may simultaneously scan the batch hierarchical vaccine data or even all 3, MedDRA, vaccine and batch information.

5. Conclusion

Using the TBSS, we have implemented a data-mining method for manufacturing-related safety signal detection that is specific and may thus provide major advantages over all other statistical methods relying on batch information. This method has the ability to locate the most probable source that might lead to an excess of AEOI in a hierarchical manufacturing structure, and to reduce the number of false positives by adjustment for multiple testing. The implementation of this method is in its early stages and therefore warrants further refinement.

Conflict of interest

All authors are employed by the GSK group of companies. Olivia Mahaux, Vincent Bauchau, Lionel Van Holle and Ziad Zeinoun hold shares in the GSK group of companies.

Acknowledgements

The authors would like to acknowledge Mathieu Vasselle for his valuable contribution through conceptual discussions regarding statistical aspects of the analysis, and Laëtitia Lastrayoli for her thorough review of the manuscript. The authors would also like to thank Anne-Theres Henze (XPE Pharma & Science c/o GSK) for medical writing support and Sophie Timmery (XPE Pharma & Science c/o GSK) for manuscript coordination and editorial support.

Funding

GlaxoSmithKline Biologicals SA was the funding source and was involved in all stages of the study conduct and analysis. GlaxoSmithKline Biologicals SA also took responsibility for all costs associated with the development and publishing of the present manuscript.

Contributorship

Olivia Mahaux contributed to the collection and generation of the data. All authors were involved in the conception and design of the work, and in the analysis and interpretation of the results. All authors participated in the development of the manuscript and approved its final version.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.vaccine.2018.11.044.

References

[1] Van Holle L, Zeinoun Z, Bauchau V, Verstraeten T. Using time-to-onset for detecting safety signals in spontaneous reports of adverse events following immunization: a proof of concept study. Pharmacoepidemiol Drug Saf 2012;21:603–10. https://doi.org/10.1002/pds.3226.

[2] Caster O, Juhlin K, Watson S, Noren GN. Improved statistical signal detection in pharmacovigilance by combining multiple strength-of-evidence aspects in vigiRank. Drug Saf 2014;37:617–28. https://doi.org/10.1007/s40264-014-0204-5.

[3] Harpaz R, DuMouchel W, LePendu P, Bauer-Mehren A, Ryan P, Shah NH. Performance of pharmacovigilance signal-detection algorithms for the FDA adverse event reporting system. Clin Pharmacol Ther 2013;93:539–46. https://doi.org/10.1038/clpt.2013.24.

[4] Van Holle L, Bauchau V. Use of logistic regression to combine two causality criteria for signal detection in vaccine spontaneous report data. Drug Saf 2014;37:1047–57. https://doi.org/10.1007/s40264-014-0237-9.

[5] Candore G, Juhlin K, Manlik K, Thakrar B, Quarcoo N, Seabroke S, et al. Comparison of statistical signal detection methods within and across spontaneous reporting databases. Drug Saf 2015;38:577–87. https://doi.org/10.1007/s40264-015-0289-5.

[6] Yih WK, Maro JC, Nguyen M, Baker MA, Balsbaugh C, Cole DV, et al. Assessment of quadrivalent human papillomavirus vaccine safety using the self-controlled tree-temporal scan statistic signal-detection method in the sentinel system. Am J Epidemiol 2018;187:1269–76. https://doi.org/10.1093/aje/kwy023.

[7] Center for Disease Control and Prevention (CDC). History of Vaccine Safety. Available at: https://www.cdc.gov/vaccinesafety/ensuringsafety/history/index.html. Accessed in November 2017.

[8] Frierson JG. The yellow fever vaccine: a history. Yale J Biol Med 2010;83:77–85.

[9] Nathanson N, Langmuir AD. The cutter incident. Poliomyelitis following formaldehyde-inactivated poliovirus vaccination in the United States during the Spring of 1955. II. Relationship of poliomyelitis to Cutter vaccine. 1963. Am J Epidemiol 1995;142:109–40. discussion 7-8.

[10] Macdougall IC, Roger SD, de Francisco A, Goldsmith DJ, Schellekens H, Ebbers H, et al. Antibody-mediated pure red cell aplasia in chronic kidney disease patients receiving erythropoiesis-stimulating agents: new insights. Kidney Int 2012;81:727–32. https://doi.org/10.1038/ki.2011.500.

[11] Boven K, Stryker S, Knight J, Thomas A, van Regenmortel M, Kemeny DM, et al. The increased incidence of pure red cell aplasia with an Eprex formulation in uncoated rubber stopper syringes. Kidney Int 2005;67:2346–53. https://doi.org/10.1111/j.1523-1755.2005.00340.x.

[12] Weaver JL, Boyne M, Pang E, Chimalakonda K, Howard KE. Nonclinical evaluation of the potential for mast cell activation by an erythropoietin analog. Toxicol Appl Pharmacol 2015;287:246–52. https://doi.org/10.1016/j.taap.2015.06.009.

[13] Bennett CL, Jacob S, Hymes J, Usvyat LA, Maddux FW. Anaphylaxis and hypotension after administration of peginesatide. N Engl J Med 2014;370:2055–6. https://doi.org/10.1056/NEJMc1400883.

[14] Guideline on good pharmacovigilance practices (GVP) – Product- or Population-specific considerations I: Vaccines for prophylaxis against infectious diseases, 9 December 2013. EMA/488220/2012 Corr. Retrieved from https://www.ema.europa.eu/documents/scientific-guideline/guideline-good-pharmacovigilance-practices-gvp-product-population-specific-considerations-i-vaccines_en.pdf. Accessed September 2018.

[15] Guideline on good pharmacovigilance practices (GVP) – Product- or Population-specific considerations II: Biological medicinal products, 4 August 2016. EMA/168402/2014 Corr. Retrieved from https://www.ema.europa.eu/documents/scientific-guideline/guideline-good-pharmacovigilance-practices-gvp-product-population-specific-considerations-ii_en-0.pdf. Accessed September 2018.

[16] DuMouchel W, Yuen N, Payvandi N, Booth W, Rut A, Fram D. Automated method for detecting increases in frequency of spontaneous adverse event reports over time. J Biopharm Stat 2013;23:161–77. https://doi.org/10.1080/10543406.2013.736809.

[17] van Puijenbroek EP, Broos N, van Grootheest K. Monitoring adverse events of the vaccination campaign against influenza A (H1N1) in the Netherlands. Drug Saf 2010;33:1097–108. https://doi.org/10.2165/11539270-000000000-00000.

[18] Purcell P, Barty S. Statistical techniques for signal generation: the Australian experience. Drug Saf 2002;25:415–21.

[19] Kulldorff M, Fang Z, Walsh SJ. A tree-based scan statistic for database disease surveillance. Biometrics 2003;59:323–31.

[20] Son DS, Lee D, Lee K, Jung SH, Ahn T, Lee E, et al. Practical approach to determine sample size for building logistic prediction models using high-throughput data. J Biomed Inform 2015;53:355–62. https://doi.org/10.1016/j.jbi.2014.12.010.

[21] Gentle JE. Random number of generation and Monte Carlo methods. 2nd ed. New York, NY: Springer; 2003.

[22] Bate A, Evans SJ. Quantitative signal detection using spontaneous ADR reporting. Pharmacoepidemiol Drug Saf 2009;18:427–36. https://doi.org/10.1002/pds.1742.

[23] Kulldorff M. Prospective time periodic geographical disease surveillance using a scan statistic. J Roy Stat Soc: Ser A (Stat Soc) 2001;164:61–72. https://doi.org/10.1111/1467-985X.00186.

[24] Kleinman K, Lazarus R, Platt R. A generalized linear mixed models approach for detecting incident clusters of disease in small areas, with an application to biological terrorism. Am J Epidemiol 2004;159:217–24.