

Development and Validation of an Automated Emergency Department-Based Syndromic Surveillance System for Use at a Mass Gathering Event (2020 Arctic Winter Games, Yukon Canada) During a Global Pandemic: Implications for Lower Resourced and Remote Settings

Etran Bouchouar (✉ ebouchouar@gmail.com)

Government of Yukon <https://orcid.org/0000-0002-8825-7286>

Benjamin M. Hetman

Public Health Agency of Canada

Brendan Hanley

Government of Yukon

Technical advance

Keywords: syndromic surveillance, mass gathering, detection algorithm, COVID-19, case definitions

Posted Date: September 10th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-54086/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background: Automated syndromic surveillance systems are useful tools for rapidly identifying health risks during times when routine surveillance and follow-up cannot meet the demands of the population. In Yukon, Canada, the Arctic Winter Games were scheduled in March 2020, and were expected to increase the local population beyond the capacity of local public health surveillance. An emergency department-based automated syndromic surveillance system was therefore developed and validated using local hospitalization records for use during the event.

Methods: Syndromes of interest were identified in consultation with the local public health authorities. For each syndrome, case definitions were developed using published resources and expert elicitation. Natural language processing algorithms were then written to detect syndromic cases from three different fields (triage notes; chief complaint; discharge diagnosis) using Yukon emergency department case data containing information from 19,082 visits over the period of October 1, 2018 to April 30, 2019. The automatic identification of cases was then manually validated by two raters and results were used to calculate positive predicted values for each syndrome and identify improvements to the detection algorithms.

Results: A total of six syndromes were originally identified for the syndromic surveillance system (e.g., Gastrointestinal, Influenza-like-illness, Mumps, Neurological Infections, Rash, Respiratory), with an additional syndrome added to assist in detecting potential cases of COVID-19. The positive predictive value for the automated detection of each syndrome ranged from 48.8%-89.5% to 62.5%-94.1% after implementing improvements identified during validation. As expected, no records were flagged for COVID-19 from our validation dataset. However, the system was rapidly adapted into an additional surveillance tool for use in the COVID-19 pandemic.

Conclusions: Validation is an important step for measuring the accuracy of syndromic surveillance, and ensuring it performs adequately in a local context. Ultimately, the 2020 Arctic Winter Games were cancelled due to the risks associated with mass gatherings during the global pandemic of COVID-19 and could not therefore be tested under a mass gathering scenario. However, the results from our validation study suggest that our surveillance system may be useful for future mass gathering events and proved a timely development for integration into Yukon's COVID-19 surveillance infrastructure.

Introduction

The Arctic Winter Games (AWG) is an international sporting and cultural event that brings together athletes from countries in the Circumpolar North, including Canada; U.S.A.; Greenland; Norway; Sweden; Finland; and Russia. This mass gathering of athletes, coaches, and media was scheduled to occur in Whitehorse, Yukon in March 2020, and was anticipated to impose considerable pressure on the territory's public health infrastructure. In other mass-gathering sporting events, vaccine-preventable diseases such as measles, influenza, and meningococcal and gastrointestinal illnesses have been reported [1, 2]. Moreover, the pandemic of novel coronavirus disease (COVID-19) and an ongoing outbreak of pertussis in northern Canada underscored the need for enhanced surveillance during the AWG, especially in a lower-resourced rural/remote setting such as Yukon, where meeting surge demand in the event of an outbreak represents a significant challenge. This paper describes a novel, near real-time automated emergency department (ED) syndromic surveillance system (SyS) that was developed and validated in preparation for the 2020 AWG and integrated into Yukon's COVID surveillance infrastructure.

Emergency department-based SyS are recognized tools for enhanced surveillance that complement traditional laboratory-based surveillance methods [3–6]. These systems automate the use of existing (near-)real-time pre-

diagnostic data that is routinely collected in hospitals and apply statistical algorithms to detect aberrations for immediate public health investigation. While it is recommended that investment in mass gathering surveillance should provide a system that is sustainable for long-term use [2], this historically has not been the case in Yukon. Previous mass gatherings in the territory used daily “drop-in” paper-based data collection methods that, while effective at the time, were resource- and labor-intensive to be sustainable long-term. Although automated SyS are often perceived to require complex technology or significant human resources, they do not need to be highly technical or costly to provide lasting benefit once established [7, 8].

To properly support an ED-SyS, operational case definitions must be developed and validated. These influence the balance between identifying all possible cases (i.e., sensitivity) and excluding those without the disease of interest (i.e., specificity) [1, 2]. There is no gold standard approach for developing or validating ED-SyS case definitions; methods described in literature suggest developing definitions using expert-based consensus followed by an ongoing refinement process, with validation based on chart review or comparison with a gold-standard dataset [9–17]. Most case definitions rely on algorithms that identify keywords associated with a syndrome of interest in the chief complaint (CC) field and discharge diagnosis (DD) fields [6, 16, 18–23], with the emerging use of clinical triage notes (CN) [16, 24–26]. Several studies have noted disagreement between syndrome definitions when using CC fields versus DD fields but, in general, case detection improves when leveraging information from both fields [11, 16, 22, 24, 25, 27, 28]. The use of CN fields also increases the sensitivity of the definition but may decrease specificity [16, 24–26].

There is currently a need to identify methods that lower-resourced communities can use to develop and operate automated ED-SyS that approach the efficacy of a fully validated real-time system, without the steep setup time and resource costs required by more sophisticated software. Moreover, these systems should be validated and optimized for use in the unique communities where they are applied. The objectives of the present study are therefore to describe the development and validation of Yukon’s first automated ED-SyS and evaluate the effectiveness of three different ED data fields. Taken together, these results may be used to inform the development of novel ED-SyS in communities similar to Yukon that lack the resources required for larger, more complex surveillance systems. While the AWG event was ultimately canceled in light of COVID-2019 physical distancing precautions, ED-based SyS described here may provide lasting benefit beyond the potential for use during a single mass-gathering event, in particular among similarly resourced rural community settings as Yukon.

Methods

The design and validation of Yukon’s ED-SyS occurred between November 2019 and February 2020, involving the following stages: (i) initial review of available data sources; (ii) development of initial case definitions; (iii) development of natural language processing algorithms and logic; (iv) validation of ED-SyS using validation dataset; (v) refinements to algorithms and logic.

Data Source: Emergency department records from Whitehorse General Hospital (WGH), the only hospital in Whitehorse and Yukon’s primary hospital, were used to perform this study. The Meditech ED-Tracker system at WGH is an electronic medical records database that captures demographic characteristics, date of visit, clinical notes (CN) (e.g., free-text describing a brief history of the stated complaint, the recorded temperature (in °C) at triage); the chief complaint (CC) containing the Canadian Emergency Department Information System (CEDIS) code [29, 30]; and the physician discharge diagnosis (DD) which is a free-text field providing the diagnosis at patient discharge.

ED-Tracker data from visits that occurred after January 1st, 2018 were provided for use. However, data before October 1st, 2018 were inconsistent (e.g. incomplete data and/or missing fields) and therefore excluded from the study. National Ambulatory Care Reporting System (NACRS) records between January 1st, 2018 and April 30th, 2019 were also used as a reference in our validation process for the inclusion of International Classification of Disease, 10th edition (ICD-10) codes. All analyses were performed using Stata LP 15.1 (Texas, USA).

Development of initial case definitions: Syndromes of interest were identified via consensus between Yukon's Communicable Disease Control Program (YCDC) Manager, the office of the Chief Medical Officer of Health (OCMOH), and the Territorial Epidemiologist. A review of existing syndromic surveillance platforms (e.g., NC-Detect, ESSENCE-II) [31, 32] along with literature searches were used to inform the terms used in the initial case definitions for each syndrome [6, 10, 12, 13, 15–19, 21, 22, 24, 33–40]; these were further refined for local context by review from stakeholders at YCDC and the OCMOH. In addition to key terms, CEDIS codes related to each syndrome were also identified and included in case definitions. Additionally, a syndrome related to COVID-19 was developed following recommendations published by the World Health Organization (WHO) [41] and local consensus with the OCMOH. This case-finding definition was further developed through extensive internal review, a review of WHO's COVID-19 weekly Situation Reports [42], and recommendations made available online by the Public Health Agency of Canada [43]. In brief, individual terms and codes related to each syndrome were organized into more broadly defined concepts that were then used to determine whether a patient record met the case definition for the syndrome of interest (see Supplement A1). Using these key terms and concepts we developed algorithms to query ED-Tracker records from October 1, 2018 to April 30, 2019 as this dataset coincided with available NACRS data. Algorithms were designed to query each data field using a forward-inclusion strategy, typically starting with DD, then querying the CC and CN fields (see Table 1).

Validation and refinement of initial syndromic surveillance case definitions: Records with an exact match between ED-Tracker and NACRS datasets (i.e., visits occurring between January 1st, 2018 and April 30th, 2019) were used to validate the ED-SyS. For syndromes that returned a large number of records (e.g., > 600) identified as positive, a subset of the data containing records from November 1, 2018 to January 31, 2019, was instead used. Following the example set by [25], records flagged for meeting syndromic case definitions were assigned one of three validation classifications: potentially related to the disease of interest (true-positive) = "0"; unclear/conflicting information (false positive) = "1"; does not match case definition/not related (false positive) = "2" (Table 1). These assignments were carried out first by reviewing available ICD-10 codes; any flagged case with an ICD-10 code equivalent to the those outlined in [44] was assigned a zero ("0"). This step was accomplished using the US-CDC's database of equivalent code translation (available at: <https://icd.codes/convert/icd9-to-icd10-cm>). Remaining records were validated manually by two epidemiologists reviewing the ICD-10 code, DD, CC and CN fields, typically in that order. After an initial review, a random sample of 10 cases was taken from the "0", "1", and "2" sub-groups for all syndromes (30 cases total per syndrome) and reviewed by both epidemiologists to ensure agreement and consistency among classification.

Following validation, we manually reviewed the validated records from each syndrome to identify: (1) misspellings or shorthand among the free-text fields that were not considered during the development of the initial algorithms; (2) terms and CEDIS codes that could be used to further refine the inclusion or exclusion of true- and false-positives, respectively; and (3) adjustments to the algorithm structure to improve case classification.

Evaluating the performance of the ED-SyS: Each syndrome included in the ED-SyS was evaluated by measuring the positive predictive value (PPV) using the validated case records. The PPV was defined as the total number of true-

positive records (classification= "0") over the total number of records (classification="0", "1", or "2"), as a percent. To evaluate the contribution of each data field on the overall performance of the ED-SyS, we measured the PPV from each component individually as well as in combination (e.g., CN; CC; DD; CN + CC; CN + CC + DD). Finally, to evaluate which terms were most influential in identifying each syndrome, the frequency of terms present among true-positive case records were measured and visually assessed using word clouds (NB: select terms may have been double-counted since some concept terms may exist in both CN and DD algorithms).

Results

Selection of syndromes for the ED-SyS: An overview of the Yukon ED-SyS is presented in Fig. 1. In total, seven syndromes of interest were selected for daily monitoring; these included: gastrointestinal illness (GI), influenza-like-illness (ILI), mumps, neurological infections (Neuro), rash, respiratory illnesses (Resp), and COVID-19. For each syndrome, diseases/conditions of interest were identified, and case definitions were developed based on algorithms that query information from the WGH ED-Tracker database (see Table 1). A complete list of terms used to query each syndrome is available in supplemental material (Supplement A1).

ED-SyS validation and performance: The merged dataset of ED visit data from the ED Tracker and NACRS contained a total of 19,023 unique ED visits between October 1, 2018 and April 30, 2019. After applying the initial case definitions to this dataset, GI and Resp were the only two syndromes that greatly exceeded our target case sample of approximately 500 hits for validation. For these two syndromes, a subset of ED visits from November 1, 2018 to January 31, 2019 was used for validation (n = 8246 records). In total, our original case definitions flagged 3,707 ED visits as potential cases for all syndromes (1,793 for GI; 966 for Resp; 593 for ILI; 64 for Rash; 234 for Neuro; 57 for Mumps; and 0 for COVID-19). As no records were flagged for COVID-19, the syndrome was not included in validation procedures.

Results from our validation are presented in Table 2. Among our initial case definitions, DD consistently returned the highest proportion of true positive cases (PPV: 51.3–100%), compared to CN (PPV: 22.8–86.1%) and CC (PPV: 0–35.0%) when used individually. In general, the CN field produced the most sensitive results, flagging the highest number of visits for all syndromes, while CC and DD fields provided a lower number of hits with higher specificity. These trends were maintained after adjustment to the initial case definitions, with the largest improvements observed in the GI syndrome (e.g., PPV for DD improved from 51.3–86.9%).

Combining multiple levels of data from ED-Tracker produced an average of 6.7-fold more hits to each syndrome than querying each component individually (Table 2). The largest example of this was observed when querying for visits related to the Resp Syndrome; using CC alongside CN increased hits from our original case definitions over 33-fold. Using our final case definitions for the combined fields produced the largest improvements to the GI Syndrome, with the PPV for the CN/CC combination increasing to 44.5% from 22.3% and the CN/CC/DD combination increasing to 78.8% from 48.8%. Changes observed among the other syndromes were minimal, save for the PPV of detecting the Mumps Syndrome, which increased from 50.9–94.1%, although the total number of observed hits decreased from 57 to 17.

Evaluating terms and logic used to identify syndromes: After an initial review of the ED-SyS performance using our validation dataset, several adjustments were made to the terms and algorithms used in our ED-SyS to improve the system's performance within the local context. In general, three areas were useful in redefining ED-SyS queries to provide more accurate results in the Yukon setting

(1) *Misspellings or shorthand among the free-text fields that were not considered during the development of the initial algorithms.* For example, the acronym “LWBS” in the DD field was often used to indicate a patient had “left without being seen”, and additional acronyms were needed to describe the diarrhea, nausea, and vomit concepts within the GI syndrome case definition including “N&V” (nausea and vomiting), “N, V, D” (nausea, vomiting, diarrhea), and “V/D” (vomiting/diarrhea).

(2) *Terms and CEDIS code that could be used to further refine the inclusion or exclusion of true- and false-positives, respectively.* For example, additional negating terms were identified for several syndromes; the inclusion of “no”, “denies” and “-” were therefore added to indicate negation within the diarrhea, nausea, vomit, fever, and cough concepts. Several CEDIS codes were consistently present among true- and false-positive results for each syndrome; we used these codes to help provide additional specificity to our inclusion and exclusion criteria based on local context (Table 3). For example, a high proportion of visits flagged by our GI case definition presented with abdominal pain (CEDIS 251), but were validated as unclear/conflicting information (n = 91/571, classification = “1”) or unrelated (n = 365/571, classification = “2”). For this reason, we elected to remove CEDIS 251 from the final GI syndrome case definition in addition to several related terms that were present in the original syndromic case definition at the CN and DD level (“ascites”, “LUQ”, “LLQ”, “RLQ”, “RUQ”, “diverticulitis”, “appendicitis”, “abdominal bloating”, and “flatus”) and were also associated with false-positive records.

(3) *Adjustments to the algorithm logic that could improve the application of our algorithms.* For example, the original algorithm for detecting the Neuro syndrome queried the CN and CC fields only if the information from the DD field was blank. We modified this process to query CN and CC when the DD field contained at least one of the terms “febrile seizure,” “headache nyd,” “malaise nyd”, or “sepsis” without mention of the term “mening”. This improved both the sensitivity and specificity of the algorithm, increasing the number of hits from six to 19, while concomitantly increasing PPV from 83.3–89.5%.

After completing the above adjustments to improve case-finding, we visualized the most influential terms among the true-positive case records for each syndrome (Fig. 2). Each syndrome appeared to have several terms that were essential to identifying the syndrome of interest. In three syndromes (Rash, Resp, and ILI) CEDIS codes appeared among the most frequent terms.

Discussion

Our ED-SyS is the first automated system to be implemented in Yukon, and, to our knowledge in northern Canada. In preparation for the 2020 AWG, we developed an ED-SyS that included six syndromes, with a seventh syndrome created in response to an ongoing global pandemic. While the AWG was ultimately canceled, the preparations allowed us to later activate the ED-SyS for enhanced surveillance during the COVID-19 pandemic, thus highlighting the flexibility and utility of these types of systems in settings like Yukon where surveillance capacity may otherwise be limited.

Results from our validation indicated that for each syndrome tested, the CN field generally provided the most sensitive results, whereas the DD field was most specific. These results were not surprising to us, as the information stored in each field tended towards specificity of a diagnosis moving from CN to CC and DD. Our findings also highlight the value of leveraging the three fields together to improve case detection. This finding is noteworthy as a significant number of cases could be missed without considering one or more fields. This was particularly evident with our respiratory and ILI syndromes, where a significant number of cases could be missed without using the CN

field. Our results for the performance of individual fields towards the overall surveillance system should be interpreted with caution, however, as many of the built algorithms were intended to pull information from CC and CN data only in the absence of a confirmed case or non-case from the DD field. For a more accurate assessment of the performance of each field, algorithms should be structured with the intent to use information uniquely from each data source, instead of in a combined fashion.

A useful step in designing our ED-SyS was ensuring that we allotted time for refining the system using local data. The original iteration of our ED-SyS was built using terms and case definitions informed by relevant literature, which did not capture the range of nuances inherent in our local data source. By reviewing and validating records on a case-by-case basis, we were able to identify additional terms and patterns of chart record-keeping that allowed us to make adjustments to the syndromic case definitions that improved the predictive ability of the ED-SyS. We found the CEDIS terms used in the CC field were especially useful for exploring false- positive cases flagged in our results. Having a standard terminology allowed us to group false positive cases from each syndrome by their CEDIS term and explore whether it was necessary to build additional exclusions into our syndromic case algorithms. This was not as feasible with either of the CN or DD fields, as they both contained free-text input, which proved much more variable than the CEDIS-codes and their standardized terms.

We were motivated to establish an ED-SyS to not only support surveillance activities during the 2020 AWG, but also enhance Yukon's surveillance infrastructure for detection and response capabilities for years to follow. Throughout all stages of development, we found it essential to leverage both local knowledge of contextual issues and disease risk and existing research to establish and refine our syndromic case definitions and algorithms. Partnerships between WGH, OCMOH and the Public Health Agency of Canada greatly contributed to the timely design and implementation of the ED-SyS. Future development of the ED-SyS will include expanding to the syndromic surveillance of other public health issues in Yukon, including opioids, cannabis, forest fires, and secondary health impacts of the COVID-19 pandemic.

Conclusions/ Practice Implications

Our study highlights the feasibility of implementing an automated ED-SyS and validating syndrome case definitions in a low-resourced/remote setting using simple tools, resources, and adapted gold standard methods. Our approach to developing an ED-SyS allowed us to move away from "drop-in" paper-based methods to create a "low-tech" sustainable system that can be leveraged for other mass-gathering events, other emerging health issues of concern, and general ongoing surveillance. Our study also reinforces the importance and value of validating syndrome case definitions using local data. Importantly, our study provides a path forward for other lower-resourced rural/remote settings on how to develop and validate syndrome case definitions.

List Of Abbreviations Used In The Manuscript

AWG Arctic Winter Games

ED Emergency department

SyS Syndromic surveillance system

CC Chief complaint

DD Discharge diagnosis

CN Clinical (triage) notes

WHG Whitehorse General Hospital

CEDIS Canadian emergency department information system

NACRS National ambulatory care reporting system

ICD International classification of disease

YCDC Yukon Center for Disease Control

OCMOH Office of the chief medical officer of health

NC Detect North Carolina disease event tracking and epidemiologic collection tool

ESSENCE-II Electronic surveillance system for the early notification of community-based epidemics

COVID-19 Coronavirus-disease 2019

US-CDC United States Centers for Disease Control and Prevention

GI Gastrointestinal illness

ILI Influenza-like illness

Neuro Neurological illness

Resp Respiratory illness

PPV Positive predictive value

LUQ Left upper quadrant

LLQ Left lower quadrant

RLQ Right lower quadrant

RUQ Right upper quadrant

nyd Not yet determined

Author Declarations

Ethics approval and consent to participate

As per Yukon's Health Information Privacy and Management Act (DIVISION 3 – USE (4))

(http://www.gov.yk.ca/legislation/acts/hipm_c.pdf), this study was recognized as “public health surveillance” work and as such it governed the collection, use, and disclosure of the associated data. Submission and approval by

Yukon's Health Research Review Committee was therefore not required. No formal IRB body operates in the Yukon. A privacy impact assessment (PIA) was required as part of routine Yukon Government processes to outline the relevant authorities under the HIPMA legislation and data safeguards. The PIA was reviewed and approved by Yukon's Chief Medical Officer of Health and Whitehorse General Hospital's Manager for Corporate Responsibility. As per Yukon Processes, this manuscript was also sent to Yukon's Health Research Review Committee for their records and no concerns were noted.

Consent for publication

Not applicable.

Availability of data and materials/data permissions

All code developed for this manuscript have been made available at <https://doi.org/10.6084/m9.figshare.12649364.v1>. The data that support the findings are not publically available. The authorities outlined in Yukon's Health Information Privacy and Management Act and a privacy impact assessment approved by Yukon's Chief Medical Officer of Health and Whitehorse General Hospital gave us permission to access the data. These are processes we engage in as part of our routine public health surveillance activities with Yukon Government's Department of Health and Social Services. These steps allowed us to establish an ongoing daily feed of the ED data from the Whitehorse General Hospital to the Department of Health and Social Services. The specific dataset used in this study are therefore available from Yukon Health and Social Services upon reasonable request and with permission of the Yukon government.

Competing interests

The authors declare that they have no competing interests.

Funding

The authors wish to acknowledge funding from the Public Health Agency of Canada to provide a Field Epidemiologist to support the analysis and writing of this project.

Authors' contributions

EB designed and led the study and wrote the original code for the surveillance system. EB and BMH performed validation, analysis and interpretation and drafted and edited the manuscript. BH contributed to study design, provided consultation and oversight, and contributed to proofreading and editing the manuscript. All authors read and approved the final manuscript

Acknowledgements

The authors acknowledge the contributions of Yukon Center for Disease Control and the Whitehorse General Hospital who provided consultation for the project and provided the dataset used for validating the surveillance system. The authors would also like to acknowledge technical contributions from Rohit Kasbekar, and proof-reading services by Alexandra Anderson.

References

1. Gautret P, Steffen R. Communicable diseases as health risks at mass gatherings other than Hajj: What is the evidence? *Int J Infect Dis.* 2016;47:46–52. doi:10.1016/j.ijid.2016.03.007.
2. Kajita E, Luard MZ, Wu H, Hwang B, Mascola L. Harnessing Syndromic Surveillance Emergency Department Data to Monitor Health Impacts During the 2015 Special Olympics World Games. *Public Health Rep.* 2017;132:99S-105S.
3. Fleischauer AT, Gaines J. Enhancing Surveillance for Mass Gatherings: The Role of Syndromic Surveillance. *Public Health Rep.* 2017;132:95S-98S.
4. Lauper U, Chen JH, Lin S. Window of Opportunity for New Disease Surveillance: Developing Keyword Lists for Monitoring Mental Health and Injury Through Syndromic Surveillance. *Disaster Med Public Health Prep.* 2017;11:173–8.
5. Faverjon C, Berezowski J. Choosing the best algorithm for event detection based on the intended application: A conceptual framework for syndromic surveillance. *J Biomed Inform.* 2018;85 August:126–35. doi:10.1016/j.jbi.2018.08.001.
6. May LS, Griffin BA, Bauers NM, Jain A, Mitchum M, Sikka N, et al. Emergency department chief complaint and diagnosis data to detect influenza-like illness with an electronic medical record. *West J Emerg Med.* 2010;11:1–9.
7. May L, Chretien JP, Pavlin JA. Beyond traditional surveillance: Applying syndromic surveillance to developing settings-opportunities and challenges. *BMC Public Health.* 2009;9:1–11.
8. May L, Katz RL, Test E, Baker J. Applications of syndromic surveillance in resource poor settings. *World Med Heal Policy.* 2011;3.
9. Scholer MJ, Ghneim GS, Wu S, Westlake M, Travers DA, Waller AE, et al. Defining and applying a method for improving the sensitivity and specificity of an emergency department early event detection system. *AMIA Annu Symp Proc.* 2007;8:651–5.
10. Chapman WW, Dowling JN, Baer A, Buckeridge DL, Cochrane D, Conway MA, et al. Developing syndrome definitions based on consensus and current use. *J Am Med Informatics Assoc.* 2010;17:595–601.
11. Fleischauer AT, Silk BJ, Schumacher M, Komatsu K, Santana S, Vaz V, et al. The validity of chief complaint and discharge diagnosis in emergency department-based syndromic surveillance. *Acad Emerg Med.* 2004;11:1262–7.
12. Chapman WW, Christensen LM, Wagner MM, Haug PJ, Ivanov O, Dowling JN, et al. Classifying free-text triage chief complaints into syndromic categories with natural language processing. *Artif Intell Med.* 2005;33:31–40.
13. MacRae J, Love T, Baker MG, Dowell A, Carnachan M, Stubbe M, et al. Identifying influenza-like illness presentation from unstructured general practice clinical narrative using a text classifier rule-based expert system versus a clinical expert. *BMC Med Inform Decis Mak.* 2015;15:1–11. doi:10.1186/s12911-015-0201-3.
14. Kotzen M, Sell J, Mathes RW, Dentinger C, Lee L, Schiff C, et al. Using syndromic surveillance to investigate tattoo-related skin infections in New York City. *PLoS One.* 2015;10:1–7.
15. Ivanov O, Wagner MM, Chapman WW, Olszewski RT. Accuracy of three classifiers of acute gastrointestinal syndrome for syndromic surveillance. *AMIA Annu Symp Proc.* 2002;:345–9.
16. Haas SW, Travers D, Waller A, Mahalingam D, Crouch J, Schwartz TA, et al. Emergency Medical Text Classifier: New system improves processing and classification of triage notes. *Online J Public Health Inform.* 2014;6.
17. South BR, Chapman WW, Delisle S, Shen S, Kalp E, Perl T, et al. Optimizing A syndromic surveillance text classifier for influenza-like illness: Does document source matter? *AMIA Annu Symp Proc.* 2008;:692–6.

18. Gerbier-Colomban S, Gicquel Q, Millet AL, Riou C, Grando J, Darmoni S, et al. Evaluation of syndromic algorithms for detecting patients with potentially transmissible infectious diseases based on computerised emergency-department data. *BMC Med Inform Decis Mak.* 2013;13:1–10.
19. Gerbier S, Yarovaya O, Gicquel Q, Millet AL, Smaldore V, Pagliaroli V, et al. Evaluation of natural language processing from emergency department computerized medical records for intra-hospital syndromic surveillance. *BMC Med Inform Decis Mak.* 2011;11:50. doi:10.1186/1472-6947-11-50.
20. Conway M, Dowling JN, Chapman WW. Using chief complaints for syndromic surveillance: A review of chief complaint based classifiers in North America. *J Biomed Inform.* 2013;46:734–43.
21. Heffernan R, Mostashari F, Das D, Karpati A, Kulidorff M, Weiss D. *Syndromic Surveillance in Public Health Practice*, New York City. *Emerg Infect Dis.* 2004;10:858–64.
22. Begier EM, Sockwell D, Branch LM, Davies-Cole JO, Jones LVH, Edwards L, et al. The National Capitol Region's Emergency Department syndromic surveillance system: Do chief complaint and discharge diagnosis yield different results? *Emerg Infect Dis.* 2003;9:393–6.
23. Lazarus R, Kleinman KP, Dashevsky I, DeMaria A, Platt R. Using automated medical records for rapid identification of illness syndromes (syndromic surveillance): The example of lower respiratory infection. *BMC Public Health.* 2001;1:1–9.
24. DeLisle S, South B, Anthony JA, Kalp E, Gundlapalli A, Curriero FC, et al. Combining free text and structured electronic medical record entries to detect acute respiratory infections. *PLoS One.* 2010;5.
25. DeYoung K, Chen Y, Beum R, Askenazi M, Zimmerman C, Davidson AJ. Validation of a syndromic case definition for detecting emergency department visits potentially related to marijuana. *Public Health Rep.* 2017;132:471–9.
26. Ising AI, Travers DA, MacFarquhar J, Kipp A, Waller AE. Triage note in emergency department-based syndromic surveillance. *Adv Dis Surveill.* 2006;1:34.
27. Raven MC, Lowe RA, Maselli J, Hsia RY. Comparison of presenting complaint vs discharge diagnosis for identifying “nonemergency” emergency department visits. *JAMA - J Am Med Assoc.* 2013;309:1145–53.
28. Harduar Morano L, Waller AE. Evaluation of the Components of the North Carolina Syndromic Surveillance System Heat Syndrome Case Definition. *Public Health Rep.* 2017;132:40S-47S.
29. Innes G, Murray M, Grafstein E. A consensus-based process to define standard national data elements for a Canadian emergency department information system. *Can J Emerg Med.* 2001;3:277–84.
30. Grafstein E, Bullard MJ, Warren D, Unger B. Revision of the Canadian Emergency Department Information System (CEDIS) Presenting Complaint List Version 1.1. *Can J Emerg Med.* 2008;10:151–61.
31. Lombardo J, Burkom H, Elbert E, Magruder S, Happel Lewis S, Loschen W, et al. A Systems Overview of the Electronic Surveillance System for the Early Notification of Community-Based Epidemics (ESSENCE II). *J Urban Heal.* 2003;80:32–42.
32. The North Carolina Disease Event Tracking and Epidemiologic Collection Tool (NC DETECT). <https://ncdetect.org/>. Accessed 26 Jan 2020.
33. Travers D, Haas SW, Waller AE, Schwartz TA, Mostafa J, Best NC, et al. Implementation of Emergency Medical Text Classifier for syndromic surveillance. *AMIA Annu Symp Proc.* 2013;2013 Cc:1365–74.
34. Public Health Agency of Canada. Overview of influenza monitoring in Canada. 2020. <https://www.canada.ca/en/public-health/services/diseases/flu-influenza/influenza-surveillance/about-fluwatch.html#a2.3>. Accessed 20 Jan 2020.

35. Fitzner J, Qasmieh S, Mounts AW, Alexander B, Besselaar T, Briand S, et al. Revision of clinical case definitions: Influenza-like illness and severe acute respiratory infection. *Bull World Health Organ*. 2018;96:122–8.
36. Hiller KM, Stoneking L, Min A, Rhodes SM. Syndromic Surveillance for Influenza in the Emergency Department- A Systematic Review. *PLoS One*. 2013;8:6–10.
37. Cadieux G, Buckeridge DL, Jacques A, Libman M, Dendukuri N, Tamblyn R. Accuracy of syndrome definitions based on diagnoses in physician claims. *BMC Public Health*. 2011;11:17.
38. Centers for Disease Control and Prevention. Syndrome Definitions for Diseases Associated with Critical Bioterrorism-associated Agents. 2003.
<https://emergency.cdc.gov/surveillance/syndromedef/pdf/syndromedefinitions.pdf>. Accessed 10 Jan 2020.
39. Ansaldi F, Orsi A, Altomonte F, Bertone G, Parodi V, Carloni R, et al. Emergency department syndromic surveillance system for early detection of 5 syndromes: A pilot project in a reference teaching hospital in Genoa, Italy. *J Prev Med Hyg*. 2008;49:131–5.
40. The SSO Research Group. Syndromic Surveillance Ontology. 2020.
<http://surveillance.mcgill.ca/projects/sso/index.html>. Accessed 14 Jan 2020.
41. Aylward, Bruce (WHO); Liang W (PRC). Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-19). WHO-China Jt Mission Coronavirus Dis 2019. 2020;2019 February:16–24.
<https://www.who.int/docs/default-source/coronaviruse/who-china-joint-mission-on-covid-19-final-report.pdf>.
42. World Health Organization. Coronavirus disease (COVID-2019) situation reports.
<https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports>. Accessed 25 May, 2020.
43. Public Health Agency of Canada. Coronavirus disease (COVID-19): Outbreak Update.
<https://www.canada.ca/en/public-health/services/diseases/2019-novel-coronavirus-infection.html>. Accessed 10 February, 2020.
44. Sloane PD, MacFarquhar JK, Sickbert-Bennett E, Mitchell CM, Akers R, Weber DJ, et al. Syndromic Surveillance for Emerging Infec. *Ann Fam Med*. 2006;4:351–8.

Tables

Table 1:

Case definitions and algorithm logic for syndromes of interest in Yukon's ED-SyS

Syndrome	Definition	Logic/Algorithm
Gastrointestinal (GI)	Match with the following diagnosis of interest OR Diarrhea Concept OR Vomiting/Nausea Concept OR Nausea Only Concept OR Anorexia Concept OR GI Concept GI related EXCLUSIONS TERMS	1. Flag DD logic first if not blank or left without being seen related text 2. If logic step #1 not met and CN/CC logic =1 then flag 3. If DD = "Viral" and DD =0 and CN/CC logic=1 then flag 4. If DD logic met and CN/CC logic met then then flag
Influenza-Like Illness (ILI)	Match with the following diagnosis of interest OR ILI OR Fever Concept AND Cough Concept AND (sore throat Concept OR at least one constitutional sign/symptom: Anorexia Concept OR Dizziness Concept OR Malaise Concept OR Muscle Pain Concept OR Lymph Concept)	1. If DD logic or CN/CC logic met then flag
Respiratory (Resp)	Match with the following diagnosis of interest OR Respiratory Concept OR Cough Concept OR Sore Concept AND (at least one constitutional sign/symptom: Anorexia Concept OR Dizziness Concept OR Fever Concept OR Malaise Concept OR Muscle Pain Concept OR Lymph Concept)	1. Flag DD logic first if not blank or left without being seen related text 2. If logic step #1 not met and CN/CC =1 then flag 3. If DD logic met and CN/CC logic met then then flag
Rash	Match with the following diagnosis of interest OR Rash Concept AND at least one constitutional sign/symptom: (Anorexia term OR Dizziness Concept OR Fever Concept OR Malaise Concept OR Muscle Pain Concept OR Lymph Concept) Rash related EXCLUSION terms	1. If DD logic or CN/CC logic met then flag
Mumps	Match with the following diagnosis of interest OR parotitis OR orchitis OR Swelling Concept AND Face Concept Mumps related EXCLUSIONS	1. Flag DD logic first if not blank or left without being seen related text 2. If logic step #1 not met and CN =1 then flag 3. If DD logic met and CN logic met then then flag
Neurological infections (Neuro)	Match with the following diagnosis of interest OR Fever Concept AND (stiff neck Concept OR Seizures Concept OR Blurred Vision Concept OR Photophobia Concept OR Altered Mental Status Concept OR Headache Concept OR Paralysis Concept)	1. Flag DD logic first if not blank or left without being seen related text 2. If logic step #1 not met and CN/CC logic =1 then flag 3. If DD = "Viral" "Febrile seizure" "Headache nyd" "Malaise nyd" "Sepsis"

and DD logic =0 and CN/CC
logic=1 then flag

4. If DD logic met and CN/CC
logic met then then flag

**Coronavirus
Disease 2019
(COVID-19)**

ED visits prior to April 16 *takes into consideration negating terms

Match with the diagnosis of interest in DD OR

CC (SOB, Cough, Fever) on or after Jan 1, 2020 AND

Match with any of the following:

-countries listed from WHO Sitrep to have "local transmission"-indicates locations where the source of infection is within the reporting location Jan 1, 2020 to March 19, 2020

-Mention of international or travel in CN on or after March 15, 2020

OR

-Match with the diagnosis of interest (COVID) in CN

ED visits on or after April 16 *takes into consideration negating terms

-Match with the diagnosis of interest in DD OR

-Match with the diagnosis of interest in CN OR

-Match with any of the following:

-CC on or after Jan 1, 2020- SOB, Cough, Fever

-CC on or after April 16, 2020- Chestpnonc, Stridor, Wheezing, Sore Throat, Syncope, Genweak, Headache, URTI, Nasal congestion/hayfever, Altered LOC, Confusion, Sensory loss/parathesis, Rep Arrest, Cyanosis

-CN- "Loss of taste" or "Loss of smell"

AND

Match with any of the following:

-countries listed from WHO Sitrep to have local transmission-indicates locations where the source of infection is within the reporting location Jan 1, 2020 to March 19, 2020

Table 2:

Flagged cases for each syndrome from Whitehorse General Hospital ED-Tracker records, Oct. 1, 2018 – Apr. 30, 2019

Flagged visits by separate fields					Flagged visits by combined fields				
	Original Case Definition		Final Case Definition			Original Case Definition		Final Case Definition	
Syndrome	# Hits	TP (PPV%)	# Hits	TP (PPV%)	Syndrome	# Hits	TP (PPV%)	# Hits	TP (PPV%)
Gastrointestinal					Gastrointestinal				
CN	1,579	360 (22.8)	746	327 (43.8)	CN	1,579	360 (22.8)	746	327 (43.8)
CC	866	303 (35.0)	295	118 (63.7)	CN/CC	1644	367 (22.3)	759	338 (44.5)
DD	485	249 (51.3)	236	205 (86.9)	CN/CC/DD	601	293 (48.8)	316	249 (78.8)
Respiratory					Respiratory				
CN	639	477 (74.6)	625	467 (74.7)	CN	639	477 (74.6)	625	467 (74.7)
CC	19	4 (21.1)	19	4 (21.1)	CN/CC	662	487 (73.6)	647	477 (73.7)
DD	76	70 (92.1)	76	70 (92.1)	CN/CC/DD	548	460 (83.9)	542	457 (84.3)
ILI					ILI				
CN	361	311 (86.1)	356	310 (87.1)	CN	361	311 (86.1)	356	310 (87.1)
CC	0	0 (0)	0	0 (0)	CN/CC	380	324 (85.2)	375	323 (86.1)
DD	141	135 (95.7)	141	135 (95.7)	CN/CC/DD	482	420 (87.1)	477	419 (87.8)
Rash					Rash				
CN	9	5 (55.6)	9	5 (55.6)	CN	9	5 (55.6)	5	9 (55.6)
CC	0	0 (0)	0	0 (0)	CN/CC	56	32 (57.1)	56	32 (57.1)
DD	10	9 (90.0)	10	9 (90.0)	CN/CC/DD	64	40 (62.5)	64	40 (62.5)
Neurological					Neurological				
CN	201	70 (34.8)	193	68 (35.2)	CN	201	70 (34.8)	193	68 (35.2)
CC	0	0 (0)	0	0 (0)	CN/CC	167	91 (54.5)	160	89 (55.6)

DD	2	2 (100.0)	2	2 (100.0)	CN/CC/DD	6	5 (83.3)	19	17 (89.5)
Mumps					Mumps				
CN	43	16 (37.2)	40	16 (40.0)	CN	43	16 (37.2)	40	16 (40.0)
DD	17	16 (94.1)	17	16 (94.1)	CN/DD	57	29 (50.9)	17	16 (94.1)
<p>Abbreviations: CN, clinical notes; CC, chief complaint; DD discharge diagnosis; TP True-positives; PPV positive predictive value</p> <p>Data Source: Meditech ED Tracker</p> <p>(a) True-positive to case definition were those for which the ED visit appeared to be related to the disease/condition of interest</p> <p>(b) Integrated can also include OR statements and/or AND statements using information in the CN, CC, DD fields within each term used</p>									

Table 3:

Positive Predictive Value (PPV) of selected CEDIS codes used to inform local inclusion or exclusion criteria for each syndrome using Whitehorse General Hospital ED-Tracker records from Oct. 1, 2018 – Apr. 30, 2019 (n=19,023 records)

Syndrome and CEDIS code	Description	Hits (PPV%)	Inclusion / Exclusion
Gastrointestinal §			
CEDIS-254	<i>"diarrhea"</i>	49 (22.8)	Inclusion
CEDIS-257	<i>"vomiting and/or nausea"</i>	170 (35.0)	Inclusion
CEDIS-852	<i>"fever"</i>	53 (51.3)	Inclusion
CEDIS-407	<i>"head injury"</i>	67 (0.0%)	Exclusion
CEDIS-307	<i>"UTI"</i>	30 (0.0%)	Exclusion
CEDIS-753	<i>"withdrawl"</i>	11 (0.0%)	Exclusion
CEDIS-251	<i>"abdominal pain"</i>	572 (20.1)	Exclusion
Respiratory §			
CEDIS-653	<i>"cough"</i>	278 (85.6)	Inclusion
CEDIS-103	<i>"sorethroat"</i>	105 (84.8)	Inclusion
CEDIS-852	<i>"fever"</i>	103 (84.5)	Inclusion
ILI			
CEDIS-653	<i>"cough"</i>	284 (77.5)	Inclusion
CEDIS-103	<i>"sorethroat"</i>	97 (57.7)	Inclusion
CEDIS-852	<i>"fever"</i>	68 (92.6)	Inclusion
Rash			
CEDIS-708	<i>"rash"</i>	71 (73.2)	Inclusion
Neurological			
CEDIS-852	<i>"fever"</i>	33 (72.7)	Inclusion
CEDIS-404	<i>"headache"</i>	25 (68.0)	Inclusion
Mumps			
CEDIS-709	<i>"localized swelling/redness-abscess"</i>	17 (77.8)	Inclusion
CEDIS-104	<i>"neck swelling/pain"</i>	9 (73.2)	Inclusion
CEDIS-101	<i>"dental gum"</i>	3 (33.3)	Exclusion
CEDIS-102	<i>"facial trauma"</i>	3 (0.0)	Exclusion
Abbreviations: CEDIS Canadian Emergency Department Information System; PPV positive predictive value			

§ These syndromes were applied to a subset of the total data containing 8246 records from November 1, 2018 to January 31, 2019

Figures

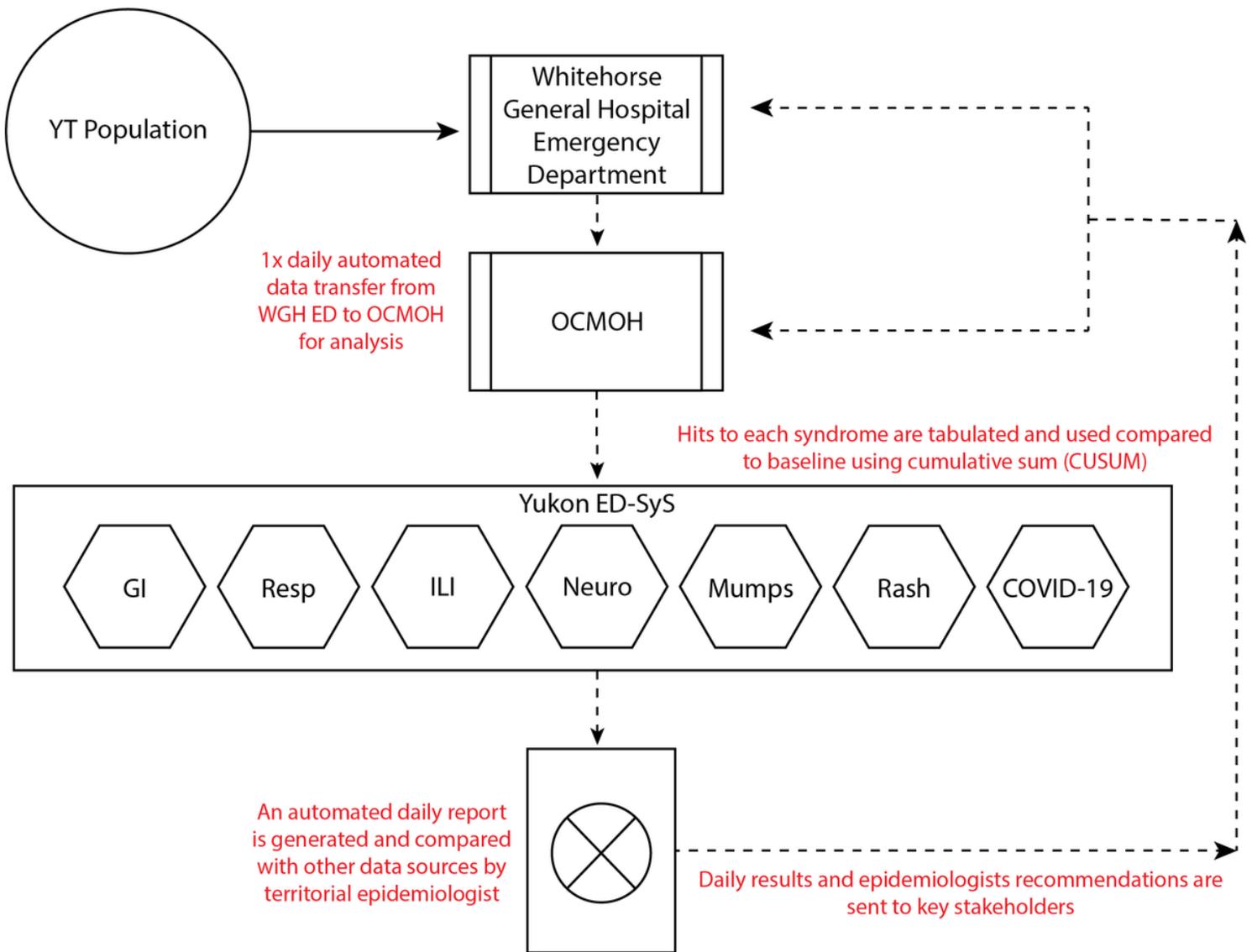


Figure 1

Overview of Yukon's automated ED-SyS. Daily data transfers are exchanged between Whitehorse General Hospital's ED-Tracker system and the Yukon Office of the Chief Medical Officer of Health (OCMOH).

