

## RESEARCH REPORTS

### Clinical

M.H. Torres-Urquidy<sup>1,2\*</sup>, G. Wallstrom<sup>1</sup>,  
and T.K.L. Schleyer<sup>2</sup>

<sup>1</sup>Department of Biomedical Informatics, School of Medicine, University of Pittsburgh, 3501 Terrace St., Suite 339, Pittsburgh, PA 15213, USA; and <sup>2</sup>Center for Dental Informatics, School of Dental Medicine, University of Pittsburgh; \*corresponding author, mit7@pitt.edu

*J Dent Res* 88(1):89-94, January, 2009

### ABSTRACT

Oral manifestations of diseases caused by bioterrorist agents could be a potential data source for biosurveillance. This study had the objectives of determining the oral manifestations of diseases caused by bioterrorist agents, measuring the prevalence of these manifestations in emergency department reports, and constructing and evaluating a detection algorithm based on them. We developed a software application to detect oral manifestations in free text and identified positive reports over three years of data. The normal frequency in reports for oral manifestations related to anthrax (including buccal ulcers-sore throat) was 7.46%. The frequency for tularemia was 6.91%. For botulism and smallpox, the frequencies were 0.55% and 0.23%. We simulated outbreaks for these bioterrorism diseases and evaluated the performance of our system. The detection algorithm performed better for smallpox and botulism than for anthrax and tularemia. We found that oral manifestations can be a valuable tool for biosurveillance.

**KEY WORDS:** bioterrorism, dental informatics, dental public health, early detection, oral manifestations.

DOI: 10.1177/0022034508327546

Received September 18, 2007; Last revision September 18, 2008; Accepted September 24, 2008

A supplemental appendix to this article is published electronically only at <http://jdr.sagepub.com/supplemental>.

## Detection of Disease Outbreaks by the Use of Oral Manifestations

### INTRODUCTION

The mailing of letters containing anthrax in 2001 (Jernigan *et al.*, 2001) and the emergence of Severe Acute Respiratory Syndrome in 2003 (Samaranayake and Peiris, 2004) have made the early detection of disease outbreaks a significant concern. Bioterrorist and naturally occurring outbreaks require “extreme timeliness of detection” (Wagner *et al.*, 2001a) to safeguard public health and mitigate deleterious effects.

In this study, we examine a novel data source for electronic biosurveillance (Wagner, 2006): oral manifestations. Some of the diseases resulting from bioterrorist agents cause oral manifestations that are likely to be detected in dental or medical care settings, and some (Flores *et al.*, 2003) have advocated a role for dentists in detecting outbreaks. Oral manifestations could be used for biosurveillance, especially in combination with other data, such as signs and symptoms of respiratory infections (Chapman *et al.*, 2004b) and over-the-counter medication sales (Goldenberg *et al.*, 2002).

Biosurveillance systems use a three-phase approach to detect outbreaks (Centers for Disease Control and Prevention, 2008). First, they gather early symptom data (Wagner *et al.*, 2001b) from systems such as hospital medical records. Second, they aggregate those data in real time, creating an electronic signal. Third, the system issues an alarm (Duchin, 2003) when the signal deviates from the predicted (*i.e.*, sudden spike in clinical cases). The alarm prompts public health officials to investigate and take action.

Biosurveillance systems are evaluated by assessments of sensitivity, specificity, and timeliness (Buehler *et al.*, 2004). This evaluation is made difficult by the lack of true ‘gold standards’ (*i.e.*, data from real outbreaks). Instead, detection algorithms are challenged to detect artificial spikes in the data from simulated outbreaks (Goldenberg *et al.*, 2002; Reis *et al.*, 2003; Wallstrom *et al.*, 2005).

To our knowledge, this is the first study that describes a new detection algorithm based on oral manifestations of bioterrorist agents. For 4 diseases (anthrax, botulism, smallpox, and tularemia), we developed baseline frequencies for oral manifestations based on clinical historical data from the Emergency Department at the University of Pittsburgh Medical Center. We evaluated the performance of the detection algorithm using simulated outbreaks.

### MATERIALS & METHODS

The project consisted of 4 phases: (1) identification of signs and symptoms in the head and neck region caused by 4 bioterrorist agents; (2) development of an algorithm to retrieve emergency department reports; (3) establishment of baseline frequencies of oral manifestations for each disease; and (4) development and evaluation of the detection algorithm.

First, we selected the bioterrorist agents from a list of public health threats (Wagner *et al.*, 2003) identified as significant by organizations such as the Defense Threat

Reduction Agency, Centers for Disease Control and Prevention, and the North Atlantic Treaty Organization. Anthrax, botulism, pneumonic plague, smallpox, and tularemia appeared in all source lists. We then searched MEDLINE (1966 to present), CINAHL (1982 to present), all available years in EMBASE, and the *Science Citation Index* for papers describing clinical manifestations of these diseases. One author (MHTU) reviewed all articles and extracted the terms describing signs/symptoms occurring in the head and neck region (APPENDIX 1). We included only manifestations that were likely to be detected either by a dentist or a physician during a head and neck and/or oral exam. In addition, we recorded the time of onset of oral relative to systemic manifestations. Plague was omitted because of no evidence of oral manifestations.

Second, we developed a set of synonyms and variants for the clinical terms describing oral manifestations (Chapman *et al.*, 2004a), to identify comprehensively the emergency department reports containing evidence for the diseases of interest. We drew this set from the National Library of Medicine's *Unified Medical Language System Metathesaurus*, Release 2004AA. For instance, synonym-variants for the term "oral ulcer" included "mouth ulcer, mouth ulceration, oral ulceration, ulceration of oral mucosa, ulcer of oral mucosa, mouth ulcers, buccal ulceration", and "ulcer buccal". It is important to note that several diseases have similar signs/symptoms. This "overlap" does not necessarily affect the performance of our data source, since the primary purpose is to detect an outbreak early, not to identify the causal agent conclusively. Our compilation included original terms, synonyms/variants, and onset of oral manifestations relative to systemic manifestations (Table 1).

Third, we developed an application in Python (V. 2.4.2, <http://www.python.org/>) to identify reports containing at least one term of interest. Our primary data sources were all 199,691 free-text emergency department reports stored in the University of Pittsburgh Medical Center Presbyterian Hospital's Medical Archiving Record System from 2001 to 2003. The reports were de-identified (Gupta *et al.*, 2004), and the date was substituted with the report's number for the week in the year of record. Our application searched each report for the listed terms (Table 1). When a term was found, the application checked whether the term was negated (*e.g.*, "lacking oral ulcers") (Chapman *et al.*, 2001). Once the application verified a positive report (*i.e.*, that the term was not negated), it recorded a "hit" (case of interest) for the year and week of the report. Additional terms in the same report were ignored. Subsequently, the number of hits was plotted over time in Microsoft Excel (Redmond, WA, USA). The system had a sensitivity (0.98) and specificity (0.93) for term-matching (identification) (Torres-Urquidy, unpublished material).

Finally, our method used historical data (previous weeks) to forecast the number of cases for the upcoming week. The algorithm generated an alert if the observed number differed significantly from the forecast.

We used a four-week moving average to calculate the expected number of cases:

$$\mu_{t+1} = \frac{1}{4}(X_t + X_{t-1} + \dots + X_{t-3})$$

where  $X_t$  is the count of cases for week  $t$ . The algorithm generates an alert if

$$X_{t+1} > \mu_{t+1} + k\hat{\sigma}$$

where  $k$  is a constant that controls the sensitivity and specificity of the algorithm, and  $\hat{\sigma}$  is an estimate of the standard deviation of the forecast error. We computed  $\hat{\sigma}$  empirically by calculating the standard deviation of the forecast errors for the previous 12 weeks (for review, see Wong and Moore, 2006).

The metrics we used to evaluate detection algorithms were sensitivity, specificity, and timeliness. Because no outbreaks of the 4 diseases occurred in Pittsburgh from 2001 to 2003, we simulated their effect on our baseline data (Wallstrom *et al.*, 2005). For each of the 4 diseases, we

assumed that each case had a probability of 0.4 of visiting an emergency department with oral symptoms. Since our surveillance system uses data from only one emergency department, we used published emergency department utilization data (Pennsylvania Department of Health, 2007) to estimate that each emergency department visit had a 0.0876 probability of being captured by our surveillance system. We estimated the case-detection sensitivity for our system as 0.9 for all 4 diseases. The above values imply that 3.15% of all cases would be expected to appear in the time series generated by our system.

We simulated outbreaks by randomly selecting a week and day of initial exposure. For each disease except smallpox, we assumed a uniform distribution of cases extending throughout the incubation period (anthrax, 1-7 days; botulism, 12-72 hrs; and tularemia, 3-5 days). In the smallpox simulation, we had to account for its high level of contagion. Using a four-component stochastic disease model, we selected an initial number of cases, and assumed that each patient infected 1.5 other individuals (Meltzer *et al.*, 2001), that all patients remained in the region covered by the surveillance system, and that no intervention (*e.g.*, quarantine) was undertaken. For all diseases, we estimated a probability of 0.0315 that a case would visit the emergency department with oral manifestations immediately upon onset and be detected by the system. Simulated cases were aggregated weekly and added to the baseline data.

Because of the uncertainty about whether a case would visit the emergency department with oral manifestations (0.4), and our system would detect this event (0.9), we conducted one-way sensitivity analyses on these two probabilities. Specifically, we measured sensitivity, specificity, and timeliness when the ED visit probabilities are 0.2 and 0.6 (other parameters unchanged). We also evaluated detection performance for case-detection sensitivity values of 0.8 and 1.0, leaving the other parameters unchanged.

For each disease, we simulated 25 outbreaks at random between Week 17 and Week 144. Selecting this interval provided at least 16 weeks of training data for the algorithm and 12 weeks to observe the effect of each outbreak. We simulated outbreaks of anthrax, tularemia, and botulism with 100, 500, and 1000 cases, and for smallpox with 1, 10, and 50 initial cases. For each simulated outbreak, we constructed a time series of emergency department visits, ran our detection algorithm on the series, and determined the number of weeks, if any, from the initial exposure that the algorithm produced an alert. When our algorithm did not detect an outbreak, we assumed that other methods would finally detect it in the week following the last week of the simulated outbreak. We characterized overall sensitivity of detection using receiver operating characteristic curves that display the relationship between false alarm rate (controlled by the parameter  $k$ ) and sensitivity. We also used activity monitor operating characteristic curves (Fawcett and Provost, 1999) to show the relationship between false alarm rate and detection timeliness. This analysis allows public health officials to evaluate a specific surveillance system according to the potential costs of false alarms. The University of Pittsburgh's IRB approved this study as exempt (approval #0406164).

## RESULTS

Readers should note that no outbreaks of anthrax, botulism, smallpox, and tularemia were reported in the Pittsburgh area during the period. Nevertheless, we captured the frequency of related oral manifestations naturally occurring as recorded by physicians from 2001–2003 (Table 2). Out of a total of 199,691 emergency department reports, 30,233 contained at least one term of interest. As shown (Table 2), the highest term frequencies were found for anthrax (including buccal ulcers, sore throat) and tularemia (*e.g.*, tonsillitis), both at approximately 7%, while

**Table 1.** Oral and Head and Neck Manifestations of Bioterrorist Agents

| Disease   | Anthrax  | Botulism   | Smallpox  | Tularemia  |
|---|--|--|---|--|
| Original term(s)                                  | buccal ulcer,<br>hoarseness,<br>oral ulcer,<br>sore throat   | dry mouth  | enanthema   | oral ulcer, tonsillitis  |
| Synonyms and variants                             | macula,<br>macule,<br>mouth ulcer,<br>papulae,<br>pain in the pharynx,<br>pain in throat,<br>papule,<br>pharyngeal pain,<br>pharynx discomfort,<br><br>round ulcer,<br>sore throat,<br>throat discomfort,<br>throat soreness | aptyalia,<br>aptyalism,<br>asialia,<br>asialias,<br>clinical xerostomia,<br>dry mouth,<br>hyposalivation,<br>hyposalivations,<br>hyposecretion of salivary gland,<br><br>mouth became dry,<br>mouth dryness,<br>odynophagia,<br>oral dryness,<br>saliva decreased,<br>salivary hyposecretion,<br>salivary secretion absent,<br>salivary secretion decreased,<br>xerostomia,<br>xerostomias | enanthem,<br>enanthema,<br>maculopapulae,<br>mucous membrane eruption,<br><br>papular rash,<br>papule | buccal ulceration,<br>enlargement of lymph nodes,<br>mouth ulcer,<br>mouth ulceration,<br>papulae, oral mucositis,<br>lymphadenopathies,<br>lymphatic disease, oral ulcer,<br>oral ulceration,<br>purulent tonsillitis<br>with lymphadenitis,<br>ulcer buccal,<br>ulcer of the oral mucosa,<br>ulceration of oral mucosa |
| Time of onset relative to systemic manifestations | simultaneous   | simultaneous, patient may be prompted to visit the ED because of the oral manifestations   | 24 hrs before the systemic rash   | simultaneous   |

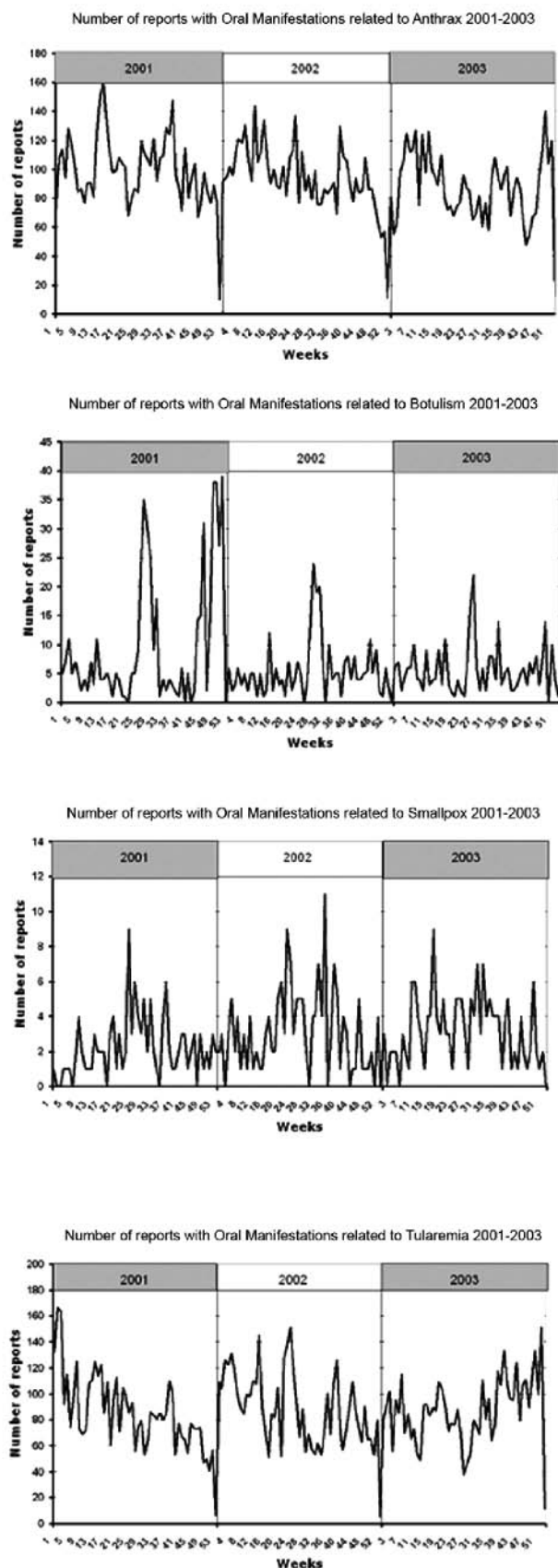
**Table 2.** Frequency of Emergency Department Reports Containing Terms Describing Oral Manifestations of Diseases Caused by Bioterrorist Agents during 2001–2003

| Year          | 2001    | 2002   | 2003   | Total   |
|---------------|---------|--------|--------|---------|
| ED Reports    | 61,808  | 68,046 | 69,837 | 199,691 |
| Anthrax       |         |        |        |         |
| Frequency     | # 5256  | 5009   | 4627   | 14,892  |
|               | % 8.50  | 7.36   | 6.63   | 7.46    |
| Weekly mean # | 99.1    | 94.5   | 87.3   | 93.6    |
| Botulism      |         |        |        |         |
| Frequency     | # 506   | 293    | 296    | 1095    |
|               | % 0.82  | 0.43   | 0.42   | 0.55    |
| Weekly mean # | 9.7     | 5.6    | 5.5    | 7.25    |
| Smallpox      |         |        |        |         |
| Frequency     | # 118   | 163    | 171    | 452     |
|               | % 1.91  | 0.24   | 0.24   | 0.23    |
| Weekly mean # | 2.22    | 3.13   | 3.28   | 2.87    |
| Tularemia     |         |        |        |         |
| Frequency     | # 4,520 | 4,666  | 4,608  | 13,794  |
|               | % 7.31  | 6.86   | 6.60   | 6.91    |
| Weekly mean # | 85.28   | 88.03  | 86.94  | 86.75   |

those for botulism (dry mouth) and smallpox (enanthema) were relatively low (0.55% and 0.23%). This finding makes intuitive sense, because both anthrax and tularemia have symptoms that occur quite commonly in other diseases, such as the common cold. Inspection of the graphs displaying weekly frequencies did not reveal any obvious patterns, as shown in the detected reports with oral manifestations (Fig. 1). (End-of-year valleys are artifacts resulting from the de-identification process.)

Next, we report on the performance of the detection algorithm (Fig. 2). The receiver operating characteristic curves show

the algorithm sensitivity as a function of false alarms *per* year, while the activity monitor operating characteristic curves do so for the week of detection from the beginning of the outbreak. The algorithm performed best for smallpox and botulism in detecting actual outbreaks, especially with a large number of cases. For instance, the algorithm can detect a smallpox epidemic of 50 initial cases with a sensitivity of 80%, with slightly over two false alarms *per* year. For botulism (1000 cases), the false alarm rate at the same level of sensitivity is about 4/yr. The sensitivity characteristics for anthrax and tularemia, in contrast,



**Figure 1.** Detected reports with oral manifestations.

resemble the line of no discrimination, indicating poor performance of our algorithm. The activity monitor operating characteristic curves present a similar impression with regard to timeliness of detection. For instance, a botulism outbreak (500 cases) is detected within one week at a false alarm rate of 4/yr. Smallpox outbreaks take longer to detect (between 6 and 7 wks) at the same false alarm rate. For anthrax and tularemia, detection timeliness ranges between 2 and 3, and 1 and 2 wks, respectively, given 4 false alarms/yr. However, the poor sensitivity of the algorithm for the two diseases reduces its application.

The sensitivity analyses showed that the sensitivity and timeliness were impervious to changes in case-detection sensitivity and moderately sensitive to large changes in the probability of reporting oral symptoms (APPENDICES 2-5). For example, the time to detect a botulism outbreak with 500 cases and 4/yr false alarms varied from 1.03 to 0.95 and 1.11 wks when the case-detection sensitivity was changed from 0.9 to 0.8 and 1.0, respectively. However, when the probability of reporting oral symptoms varied from 0.4 to 0.2 and 0.6, timeliness changed to 0.74 and 1.48 wks.

## DISCUSSION

This study characterized the frequency of oral manifestations in a population and developed a model for the detection of bioterrorist attacks by monitoring these manifestations in emergency department reports. In this model, the performance differential among different diseases may have occurred for several reasons. One major factor is that oral manifestations for smallpox and botulism occurred much less frequently (0.23% and 0.55%, respectively) in emergency department reports at baseline than those for anthrax and tularemia (7.46% and 6.91%). Any increase in corresponding oral manifestations thus had a proportionally larger effect for smallpox and botulism than for anthrax and tularemia. Second, the incubation periods for anthrax and tularemia are much longer than for the other two diseases. The uniform distribution of simulated cases over the incubation period thus “diluted” the signal compared with the diseases with a shorter incubation period. In the case of smallpox, the highly contagious nature of this disease had the effect of amplifying the initial signal rapidly, and thus may be primarily responsible for the superior performance of the algorithm.

The main objective of this study was to identify certain oral manifestations and explore the feasibility of using these as a data source for biosurveillance, not to construct a definitive algorithm. As such, our model is contingent on assumptions for which, at present, little support exists. First, our assumptions about oral manifestations of bioterrorist diseases are based on a literature review, not on systematic studies conducted by oral health researchers. Future studies should focus on determining oral manifestations in detail, and how health professionals report them. Second, we do not know how likely it is that patients would seek care for oral manifestations during an outbreak. With the exception of those for smallpox, none is prodromal, making systemic manifestations equally useful for monitoring. Third, the ‘moving average’ analysis we used is only one approach for detection. More sophisticated univariate detection methodologies exist, and other factors, such as day-of-week effects and seasonality, should be utilized in refining



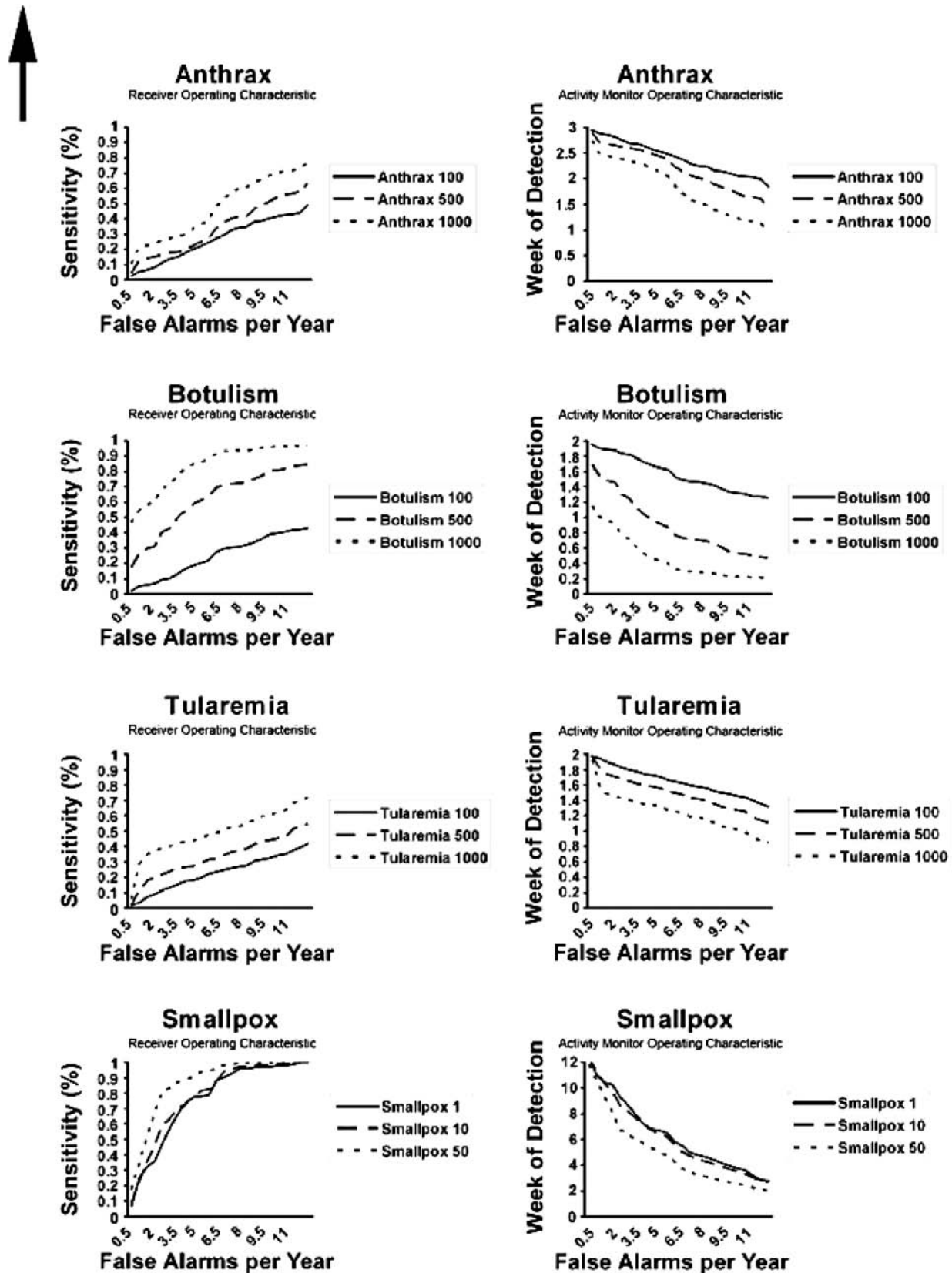


Figure 2. Performance of the detection algorithm.

the algorithm. Last, the lack of data from real outbreaks limited our ability to evaluate the proposed detection system. A reasonable approach in leveraging our method would be to combine the developed signal with others to improve detection performance.

In conclusion, our study determined the prevalence of specific oral manifestations in a population and showed that this is a viable novel method for their monitoring. In this way, dentists and physicians could contribute to effective and efficient biosurveillance.

## ACKNOWLEDGMENTS

The authors thank Wendy Chapman, Rebecca Crowley, Melissa Saul, and Robert Weyant for their assistance. M.H.T.U. received support from CONACyT, Mexico (#167967) and NIDCR/NIH (1R21DE018548-01). G.W. was supported by a grant from the CDC (R01PH000025). This work is solely the responsibility of its authors and does not necessarily represent the views of the funding sources. Preliminary results of this project were presented as a poster at the 84<sup>th</sup> General Session & Exhibition of the IADR (2006) in Brisbane, Australia.

## REFERENCES

- Buehler JW, Hopkins RS, Overhage JM, Sosin DM, Tong V, CDC Working Group (2004). Framework for evaluating public health surveillance systems for early detection of outbreaks: recommendations from the CDC Working Group. *MMWR Recomm Rep* 53:1-11.
- Centers for Disease Control and Prevention (2008). Syndromic surveillance: an applied approach to outbreak detection. <http://www.cdc.gov/ncphi/diss/nndss/syndromic.htm> (accessed Sept 24, 2008)
- Chapman WW, Bridewell W, Hanbury P, Cooper GF, Buchanan BG (2001). A simple algorithm for identifying negated findings and diseases in discharge summaries. *J Biomed Inform* 34:301-310.
- Chapman WW, Dowling JN, Wagner MM (2004a). Fever detection from free-text clinical records for biosurveillance. *J Biomed Inform* 37:120-127.
- Chapman WW, Fiszman M, Dowling JN, Chapman BE, Rindfleisch TC (2004b). Identifying respiratory findings in emergency department reports for biosurveillance using MetaMap. *Stud Health Technol Inform* 107(Pt 1):487-491.
- Duchin JS (2003). Epidemiological response to syndromic surveillance signals. *J Urban Health* 80(2 Suppl 1):i115-i116.
- Fawcett T, Provost F (1999). Activity monitoring: noticing interesting changes in behavior. Proceedings of the Fifth International Conference on Knowledge Discovery and Data Mining. <http://doi.acm.org/10.1145/312129.312195> (accessed Sept 24, 2008)
- Flores S, Mills SE, Shackelford L (2003). Dentistry and bioterrorism. *Dent Clin North Am* 47:733-744.
- Goldenberg A, Shmueli G, Caruana RA, Fienberg SE (2002). Early statistical detection of anthrax outbreaks by tracking over-the-counter medication sales. *Proc Natl Acad Sci USA* 99:5237-5240.
- Gupta D, Saul M, Gilbertson J (2004). Evaluation of a deidentification (De-Id) software engine to share pathology reports and clinical documents for research. *Am J Clin Pathol* 121:176-186.
- Jernigan JA, Stephens DS, Ashford DA, Omenaca C, Topiel MS, Galbraith M, *et al.* (2001). Bioterrorism-related inhalational anthrax: the first 10 cases reported in the United States. *Emerg Infect Dis* 7:933-944.
- Meltzer MI, Damon I, LeDuc JW, Millar JD (2001). Modeling potential responses to smallpox as a bioterrorist weapon. *Emerg Infect Dis* 7:959-969.
- Pennsylvania Department of Health (2007). Data from the Annual Hospital Questionnaire, Reporting Period July 1, 2001—June 30, 2002. Pennsylvania Department of Health, Bureau of Health Statistics and Research. <http://www.dsf.health.state.pa.us/health/lib/health/facilities/hosamb/2001-2002/H0204.PDF> (accessed Sept 24, 2008)
- Reis BY, Pagano M, Mandl KD (2003). Using temporal context to improve biosurveillance. *Proc Natl Acad Sci USA* 100:1961-1965.
- Samaranayake LP, Peiris M (2004). Severe acute respiratory syndrome and dentistry: a retrospective view. *J Am Dent Assoc* 135:1292-1302.
- Wagner MM (2006). Introduction. In: Handbook of biosurveillance. Wagner MM, Moore AW, Aryel RM, editors. Burlington: Elsevier Academic Press, pp. 3-12.
- Wagner MM, Tsui FC, Espino JU, Dato VM, Sittig DF, Caruana RA, *et al.* (2001a). The emerging science of very early detection of disease outbreaks. *J Public Health Manag Pract* 7:51-59.
- Wagner M, Aryel R, Dato VM, Krenzelok E, Fapohunda A, Sharma R (2001b). Availability and comparative value of data elements required for an effective detection system. Contract No. 290-00-0009. Rockville, MD: Agency for Healthcare Research and Quality.
- Wagner MM, Dato V, Dowling JN, Allswede M (2003). Representative threats for research in public health surveillance. *J Biomed Inform* 36:177-188.
- Wallstrom GL, Wagner M, Hogan W (2005). High-fidelity injection detectability experiments: a tool for evaluating syndromic surveillance systems. *MMWR Morb Mortal Wkly Rep* 54(Suppl):85-91.
- Wong WK, Moore AW (2006). Classical time-series methods for biosurveillance. In: Handbook of biosurveillance. Wagner MM, Moore AW, Aryel R, editors. Burlington: Academic Press, pp. 217-234.