Classification Tree Methods for Development of Decision Rules for Botulism and Cyanide Poisoning

Howell Sasser, PhD^a, Marcy Nussbaum, MS^a, Michael Beuhler, MD^b, Marsha Ford, MD^b

^aDickson Institute for Health Studies, Carolinas Medical Center, Charlotte, NC

^bCarolinas Poison Center, Carolinas Medical Center, Charlotte, NC

ABSTRACT

Introduction: Identification of predictors of potential mass poisonings may increase the speed and accuracy with which patients are recognized, potentially reducing the number ultimately exposed and the degree to which they are affected. This analysis used a decision-tree method to sort such potential predictors.

Methods: Data from the Toxic Exposure Surveillance System were used to select cyanide and botulism cases from 1993 to 2005 for analysis. Cases of other poisonings from a single poison center were used as controls. After duplication was omitted and removal of cases from the control sample was completed, there remained 1,122 cyanide cases, 262 botulism cases, and 70,804 controls available for both analyses. Classification trees for each poisoning type were constructed, using 131 standardized clinical effects. These decision rules were compared with the current case surveillance definitions of one active poison center and the American Association of Poison Control Centers (AAPCC).

Results: The botulism analysis produced a 4-item decision rule with Sensitivity (Se) of 68% and Specificity (Sp) of 90%. Use of the single poison center and AAPCC definitions produced Se of 19.5% and 16.8%, and Sp of 99.5% and 83.2%, respectively. The cyanide analysis produced a 9-item decision rule with Se of 74% and Sp of 77%. The single poison center and AAPCC case definitions produced Se of 10.2% and 8.6%, and Sp of 99.8% and 99.8%, respectively.

Conclusions: These results suggest the possibility of improved poisoning case surveillance sensitivity using classification trees. This method produced substantially higher sensitivities, but not specificities, for both cyanide and botulism. Despite limitations, these results show the potential of a classification-tree approach in the detection of poisoning events.

INTRODUCTION

With recent advances in technology, surveillance for weapons of mass destruction (WMD) agents in real time using reported symptoms has become feasible. An important step in realizing this goal is to develop predictors of human poisoning by WMD agents using reported clinical effects. If a set of strong predictors can be identified, this information may increase the speed and

accuracy with which patients poisoned by WMD are identified, potentially reducing the number of people ultimately exposed and affected. Conversely, the introduction of meaningful numbers of erroneous cases (false positives) can make the tracking process cumbersome [1].

The majority of clinical effects are nonspecific and could plausibly be associated with more than one toxic substance [2,3]. For this reason, it is typically necessary to use a combination of

Keywords: botulism; cyanide; decision trees; models, statistical; poisons, public health aspects; surveillance

Notes: This project was supported by a grant from the National Bioterrorism Hospital Preparedness Cooperative Agreement—North Carolina, administered by the North Carolina Office of Emergency Medical Services.

Acknowledgments: The authors would like to thank Drexel Pratt, Chief, and Holli Hoffman, Assistant Chief, Health Services, of the North Carolina Office of Emergency Medical Services, Division of Facility Services, North Carolina Department of Health and Human Services, for their support of this study.

Corresponding Author: Dr. Howell Sasser, Dickson Institute for Health Studies, Carolinas HealthCare System, P.O. Box 32861, Charlotte, NC 28232-2861. Email: howell.sasser@carolinashealthcare.org

clinical effects (both included and excluded) to define a case. In the cases of cyanide and botulism, the Centers for Disease Control and Prevention (CDC) and the American Association of Poison Control Centers (AAPCC) have proposed sets of criteria for surveillance definitions [4]. However, these are not mandatory, and each of the 61 US poison centers may use a modified definition, a completely different definition, or none at all.

With this situation in mind, we approached the issue from an empirical perspective. Applying a learn-and-test classification tree process to a large sample of poisoning exposure calls, we produced data-based decision rules. The goal of this process was to identify groups of symptoms that were most reliably associated with poisoning by specific agents. These decision rules were then compared with others already in use, to judge their usefulness for poisoning surveillance.

MATERIALS AND METHODS

Data

Data for the study were obtained from the Toxic Exposure Surveillance System (TESS), a nationwide database of the majority of calls made to Poison Control Centers. Each record contains basic demographic information, indicators for 131 clinical effects (symptoms), implicated toxin(s), treatments administered, and outcome information as entered electronically by the Specialist in Poison Information (SPI) handling the call. Each clinical effect in the TESS database has an established definition used to standardize coding. The toxic substance details and clinical effects were recorded as reported by the patient, physician, or other reporter, and as interpreted by the SPIs working at the various poison control centers. For this analysis, call records between January 1, 1993 and December 31, 2005 were used. The project was judged exempt from review by the Carolinas Health Systems Institutional Review Board (CHS IRB) because the data used were deidentified and publicly available.

Queries of the TESS system by toxin code (77920, 77983, 127000, and 201060 for botulism; 12000, 12563, 77260, 88000, 201062, and 201101 for cyanide) yielded 45,186 potential cases for cyanide and 26,004 potential cases for botulism before removal of duplication and filtering, as described below. Comparison records for both analyses were drawn from the database of all calls made to a single poison control center (Carolinas Poison Center [CPC]) in 2005. This produced 83,307 potentially usable controls, again, before filtering.

Filtering steps removed duplicate records (an artifact of the extraction process) and excluded cyanide and botulism cases from the control datasets for their respective analyses.

Cases coded as exposures listed in poison control center call records do not necessarily represent a poisoning or overdose. The substance coded reflects a best guess as to the etiological agent(s). Thus, it was necessary to reduce the cyanide and botulism datasets to those observations believed to have been true poisoning cases, termed Poisoned/Cases. We used the database variable

Medical_Outcome for this purpose; the possible codes are listed below:

1. Death
2. Death indirect report
3. Major effect
4. Moderate effect
5. Minor effect
6. Unable to follow—judged as potentially toxic exposure
7. Not followed—judged as nontoxic exposure (clinical effects not expected)
8. Not followed—minimal clinical effects possible (no more than minor effects possible)
9. Unrelated effect—the exposure was probably not responsible for the effect
10. No effect

Choices 1 through 4 were considered true poisonings for the purpose of these analyses. Of note, Minor and No Effect cases, as well as those not followed, were excluded from further analysis. This cut-point balanced internal validity with broad applicability. It was our reasoning that an internally valid model (one with good predictive accuracy) was a necessary first analytical step. If a reliable model could be built to describe the available data, this could then be applied with confidence to new information.

After limitation of cases to the outcomes described above, there remained 1,122 Poisoned/Cases for the cyanide analysis and 262 Poisoned/Cases for the botulism analysis. After omission of duplication and removal of potential cases from the control sample, there were 70,804 Not Poisoned/Controls available for both agent-specific analyses. Random samples from the Not Poisoned/Controls dataset ($n = 11,220$ for cyanide and $n = 3,000$ for botulism) were drawn (without replacement) for comparison with the cyanide and botulism cases.

The data entry mechanism of the TESS system requires that all clinical effects that are reported be coded to reflect the degree to which the caller or SPI believes them to be related to the toxic agent. The available categories are Related, Not related, Unknown if related, or Null (not present or not coded). Thus, a “positive” or “present” symptom could be coded under any of the first three choices. For these analyses, the aim was to assess the predictive value of all reported symptoms, rather than just those judged to be related. For this reason, clinical effects characterized as Related, Not related, or Unknown if related were grouped together and defined as a positive finding, and Null was defined as a negative finding.

Two groups of clinical effects were combined for the cyanide analysis. If either nausea or vomiting was coded as Related, Not related, or Unknown if related, a new aggregated variable (notated as GI in *Figure 3*) was coded as positive. If both were coded as Null, the new variable was coded as negative. Asystole, cardiac arrest, and systemic acidosis were also combined in the same manner (into a variable notated as Cardiac). These decisions were made because it was believed that the SPIs may have keyed in only one of these effects, when in fact one or more others were

also present. For example, if vomiting was present then nausea likely also was present, but may have been omitted as subsumed under the higher-order symptom. Cross-tabulations of the variables to be combined showed meaningful numbers of cases in which agreement between the relevant variables was lacking. This suggests that the combined variables may be more representative, capturing the same phenomenon coded in different ways for different cases.

Analysis

Analyses using classification-tree methods were conducted in CART Version 5.0 (Salford Systems, San Diego, CA). The goal of classification-tree analysis is to produce a sequence of decision steps that predicts an outcome based on optimal “splitting” of the available observations into two categories (yes/no, present/absent, etc.). In this analysis, the software began by considering all the potential predictor variables and choosing the one that best split the data into Poisoned/Cases and Not Poisoned/Controls groups or “nodes.” It repeated this process for each “child” node, producing further branches and nodes, until the size of the successive nodes reached a prespecified minimum.

Although a very large tree may accurately reflect the structure of the data used to build it, it is unlikely to be useful in practice. To address this issue, we chose to specify that “terminal” nodes (those with no further splitting) have a number of observations equal to at least 1% of the total analysis population (123 for cyanide and 40 for botulism).

As a way of checking the accuracy of the classification-tree model, some method of testing or validating the tree that has been constructed is often used. If enough data are available, the data can be split into two sets: the learning set and the testing set. It is common to use a learning set of 80% of the data to construct

the tree, then to verify the tree with the remaining 20%. We followed this approach. Importantly, this means that the classification trees shown represent the learning datasets, and the tables comparing the predictive value of the results reflect the validation datasets.

The classification tree diagrams included in this report show the selected splitters in hexagons and terminal nodes in rectangles. For example, for cyanide, the first splitter chosen was dyspnea. Those observations in the learning dataset that responded to dyspnea (Related, Not related, or Unknown if related) were put into the first terminal node to the left of the initial hexagon node. Those observations that did not respond (Null) went into the hexagon node to the right for further splitting. It is not necessarily valid to rank the importance of the splitters from the first to the last in the tree sequence because they are dependent on the structure of the tree (conditional probabilities). However, the tree sequence is a critical part of the model, and inference from the tree is valid only if the splitters are used in order.

The results of the CART analyses were compared with the poisoning surveillance definitions for cyanide and botulism currently in use at CPC and the definitions proposed by AAPCC (*Figures 1 and 2*), automated with the SAS System (SAS Institute, Cary, NC), and applied to the same datasets used for the CART analyses. The sensitivity, specificity, and percent misclassification for each method are reported.

RESULTS

Cyanide

Several possible classification trees were evaluated for sensitivity and specificity. The final tree is shown in *Figure 3*. It included (in order) dyspnea, cardiac, other, GI, cough/choke, tachycardia,

CPC				
At least one of:	AND	At least one of:	AND	None of:
Coma		Acidosis		Tremor
Drowsiness/Lethargy		Cardiac arrest		Tinnitus
Agitation/Irritability				Hematemesis/UGI bleed
Confusion				Oxalate crystals
AAPCC				
At least one of:	AND	At least one of:		
Coma		Acidosis		
Drowsiness/Lethargy				
Agitation/Irritability				
Confusion				

Figure 1: CPC and AAPCC cyanide poisoning definitions.

CPC				
At least one of:	AND	At least one of:	AND	None of:
Blurred vision		Constipation		Burns, superficial
Paralysis		Nausea		Burns, 2nd/3rd degree
Dysphagia		Vomiting		Esophageal injury
Muscle weakness				Esophageal stricture
				Gastric burns
				Oral burns
				Oropharyngeal injury
				Hallucinations/delusions
				Tinnitus
				Deafness

AAPCC		
At least one of:	AND	At least one of:
Blurred vision		Paralysis
Visual defect		Dysphagia
Photophobia		Muscle weakness

Figure 2: CPC and AAPCC botulism poisoning definitions.

erythema/flushing, irritation/pain, and red eye. The sensitivity and specificity by the CART method were 73.9% and 77.2%, respectively. The sensitivity and specificity of the CPC decision algorithm were 10.2% and 99.8%, respectively. The sensitivity and specificity of the AAPCC algorithm were 8.6% and 99.8%, respectively.

Botulism

Models were built including and excluding the variable Other. Other may be coded if a case has clinical effects that cannot be classified by coding one or more of the 130 specific clinical effects. It was unclear what symptoms might commonly be coded under this heading, but it proved to be a moderately strong predictor of botulism. Indeed, the model including Other had greater sensitivity, but poorer specificity, than the model excluding it. Both trees are presented for comparison (*Figures 4 and 5*).

The variables in the final model including Other were (in order): muscle weakness, nausea, "other," and abdominal pain. The variables and order in the final model excluding Other were the same except for that omission. The sensitivity and specificity of the model including Other were 68.2% and 90.1%, respectively. Those for the model excluding Other were 63.6% and 92.5%, respectively. The sensitivity and specificity of the CPC decision algorithm for botulism were 19.5% and 99.5%, respectively. The sensitivity and specificity of the AAPCC algorithm were 16.8% and 83.2%, respectively.

DISCUSSION

To the best of our knowledge, this is the first published report of the accuracy of various surveillance definition models for botulism and cyanide poisoning. These results suggest the possibility of improved surveillance sensitivity using classification trees. The decision-tree method was able to produce substantially higher sensitivities for both cyanide and botulism. However, this came at the expense of specificity. The CART models' specificities were lower than those of the other decision rules.

The ideal decision rule would incorporate both capturing as many cases as possible (high sensitivity) and excluding as many non-cases as possible (high specificity). The degree to which lower performance in one aspect or the other can be tolerated likely will depend on the goal of surveillance. The potential cost of missed cases must be weighed against the cost of unnecessary activations of medical and/or public safety alerts and mechanisms in response to erroneous positive findings.

A number of factors are important in placing these findings in perspective. First, analysis took place after the fact. Each case record contained all the information collected during the course of the case, irrespective of the time at which each clinical effect developed or was reported. This meant that the decision rules could contain clinical effects that manifested at different time points in the course of the event. Prospective assessment of a case might not permit concurrent evaluation of all the necessary clinical effects. This might cause underreporting of cases in real-life conditions, especially with more complex decision paths.

Figure 3: Classification tree for cyanide poisoning prediction.

Second, the analysis relies on the clinical effects reported. It is important to note that failure to report a symptom (or the absence of coding for the symptom in the form of a clinical effect in the database) does not mean that the symptom was not present. This may be exacerbated by the likelihood of underreporting of symptoms in the context of a stressful event.

Third, the clinical effects listed for each case are as reported to, and recorded by, the SPI who took the call. If the caller was a

parent, relative, or other nonmedical associate of the case, clinical effects may not have been reported completely or accurately. This likely was mitigated to some extent by the severity of the cases included in this analysis. Few if any would have been managed at home, suggesting that a medical entity was involved at some point in every case. The SPI may also have missed symptoms, or made invalid inferences in mapping reported symptoms to the clinical effects available for coding. Variations in the training of

Figure 4: Classification tree for botulism poisoning prediction, excluding "Other."

Figure 5: Classification tree for botulism poisoning prediction, including "Other."

SPIs, particularly the emphasis given to coding accuracy, may contribute to this. A comparison of poison center and hospital records suggests that these issues are of some importance [5].

Despite these limitations, these results show the potential of a systematic approach to the classification of poisoning events. By developing empirically derived predictive definitions from the very data stream it is intended to analyze, this analytical method has arguable advantages over others that rely on case reports or textbooks. The cyanide and botulism definitions presented here appear to be useful for those engaged in real-time surveillance using poison center data, or clinical data that can be broken into individual clinical symptoms. This work also highlights the importance of determining the goal of surveillance with respect to sensitivity and specificity, and demonstrates the apparent variability in those dimensions inherent in the rules employed by several poison centers that currently are conducting surveillance.

This project was supported by a grant from the National Bioterrorism Hospital Preparedness Cooperative Agreement—North Carolina,

administered by the North Carolina Office of Emergency Medical Services.

REFERENCES

1. Beuhler JW, Hopkins RS, Overhage JM, Sosin DM, Tong V. Framework for evaluating public health surveillance systems for early detection of outbreaks. *MMWR Morb Mortal Wkly Rep (Recommendations and Reports)* **2004**;53(RR05):1–11.
2. Shapiro RL, Hatheway C, Swerdlow DL. Botulism in the United States: A clinical and epidemiologic review. *Ann Intern Med* **1998**;129:221–228.
3. Yen D, Tsai J, Wang LM, et al. The clinical experience of acute cyanide poisoning. *Am J Emerg Med* **1995**;13:524–528.
4. Wolkin AF, Patel M, Watson W, et al. Early detection of illness associated with poisonings of public health significance. *Ann Emerg Med* **2006**;47:170–176.
5. Hoyt BT, Rasmussen R, Griffin S, Smilkstein MJ. Poison center data accuracy: A comparison of rural hospital chart data with the TESS database. *Acad Emerg Med* **1999**;6:851–855.