Simulation testing of the disease detection performance of the Bovine Syndromic Surveillance System (BOSSS)

Shephard, R. W.[1,2], Toribio, J. A.[2], Cameron[3], A. R. Thomson, P. C.[2] and Baldock, F.C.[3]
[1]Australian Biosecurity Cooperative Research Centre, Brisbane, Australia; [2]University of Sydney, Sydney, Australia; [3]AusVet Animal Health Services, Brisbane, Australia

The Bovine Syndromic Surveillance System (BOSSS) has been established to collect, collate and analyse observations of cattle disease by a variety of reporters including veterinarians, stock inspectors, farmers and stock workers. The construction and deployment of BOSSS are described elsewhere in these proceedings.

The network of veterinarians observing cattle has declined in rural and remote Australia (Frawley, 2003). BOSSS was developed to assist the 'alert clinician' general disease surveillance system. Data from the general surveillance system has traditionally been examined on a case-by-case basis. Syndromic surveillance systems are being increasingly deployed for human disease surveillance (Lombardo et al., 2003, Tsui et al., 2003, Wagner et al., 2001). These systems collect data on syndrome presentation and population health seeking behaviour. They use pattern detection algorithms to detect change to the distribution of syndromes. The collection, analysis and interpretation of syndrome data from observers of cattle may provide useful support to the existing 'alert clinician' model of general veterinary surveillance.

There have been few studies to examine the usefulness of surveillance systems for early and reliable detection of outbreaks (Buehler et al., 2004). Syndromic surveillance systems are assessed by their ability to detect outbreaks (sensitivity), the false positive alarm rate (ie specificity) and the timeliness of the detection. In order to estimate these parameters with accuracy the system will ideally be evaluated using data with multiple known outbreaks and data from prolonged periods without an outbreak. To date, no major outbreak and very few smaller outbreaks have been detected by functional syndromic surveillance systems. Partly this is because there are very few operational syndromic surveillance systems and an absence of real outbreak data. The only method available to assess potential worth of a syndromic surveillance system component is simulation modelling. Simulation allows detection algorithm performance to be assessed and component sensitivity to be evaluated. The disadvantages of this approach include model development can be complex and it is difficult to assess if baseline data or simulated outbreak data are realistic.

We developed a discrete time stochastic herd-level simulation model to test the BOSSS disease detection system. The programmable language R was used to build the model (R Development Core Team, 2004). A cellular automata herd-level SIR model comprising a rectangular array of farms was used. The distance between each farm was calculated and stored as a variable along with infection status at each time period. A hypothetical infectious disease was defined using arbitrary conditional sign probabilities and within- and between-herd spread parameters. Within-herd spread was modelled using a susceptible-infectious-removed (SIR) approach with arbitrarily assigned alpha and beta coefficients.

Individual farms were randomly assigned at the start of each simulation as reporting or non reporting sources. The number of reports generated per reporting farm per time period was obtained using random sampling from a Poisson distribution that was described during pilot testing. A specific disease was assigned to each report by random sampling from a modified BOSSS disease database adjusted to include the hypothetical disease. The relative individual disease prevalence within each herd provided the selection weight – this was the baseline prevalence for endemic diseases and the within-herd prevalence calculated from the SIR model for the hypothetical disease. The process was repeated until a specific disease was assigned to each report. The positive signs present within a case were assigned using weighted random sampling with conditional probabilities for the signs as sampling weights. Random sampling was again used to determine the number of these case positive signs that were reported. This controlled for observational and reporting biases that can be expected in a voluntary system.

The cellular automata model allows disease to cluster locally yet disseminate through the matrix thereby mimicking reality. Spread of disease occurred via local spread and animal transfer. Local spread risk for a farm was adjusted using the combined effects of within-herd prevalence and distance to other infected farms in the simulation. Animal movements systems allowed disease to spread with risk of transfer dependent in part upon within-herd prevalence.

The model was run with 400 farms for 60 time periods (where a time period was one month). This simulation mimics use of the system in an extensive remote area region of Northern Australia over a period of 5 years. Around half the farms were designated to be reporting farms providing an average of 3 reports and an average of 2.5 signs per report per period. Therefore around 600 reports per month were recorded by the system.

Two algorithms were evaluated using the output from this simulation model. These were the CuSum and WSARE (What's Strange About Recent Events) (O'Brien and Christie, 1997, Wong et al., 2003). The CuSum (i.e. cumulative sum, as it accumulated data over time) can identify an increase in the reporting incidence of a nominated syndrome. A trigger is evoked when the accumulated increase above expectation for a syndrome exceeds a threshold. CuSum parameters can be set to optimise the balance between sensitivity, time to detection and false alarm rate.

WSARE is a general pattern recognition algorithm that uses Fisher's exact test to identify significant differences to the level of or associations between variables in the current time period compared to a historical time period. Results are adjusted for multiple comparisons using a randomisation test. This algorithm tests all possible variables and two-way associations between variables and therefore is a general detector. The user must define the duration of the testing period and nominate the interval between the current and historical time periods. These choices influence the ability of the algorithm to detect different epidemic types (ie rapid versus slow spreading), the false alarm rate and the timeliness of detection. Both CuSum and WSARE algorithms require training data to set parameters. Parameters were set using the model however recalculation will be essential when real data becomes available because model baseline data often underestimates natural variation (Wagner, 2005).

A CuSum was established for each of the ten individual parent sign categories used in BOSSS. CuSums were modified to monitor proportions instead of counts to control for the effects of varying reporting intensities. CuSum monitoring of individual signs was found to be a very sensitive indicator of disease when the sign had a high conditional probability. Individual sign CuSums had acceptable false positive alarm rates but the combined effect from all ten CuSums was an unacceptably high false alarm rate. This was due in part to multiple comparisons. It is not effective to use multiple CuSums, each monitoring change in a single sign, because of the unacceptably high false alarm rate.

A specific CuSum for tracking multiple sign combinations (ie syndromes) as a detector for individual disease was associated with high sensitivity and low false alarm rate. A CuSum chart monitoring the combined occurrence of three general signs: abnormal behaviour, weight loss and abnormal behaviour as a detector for bovine spongiform encephalopathy (BSE) from simulation modelling is presented in Figure 1. The early and persistent increase in the CuSum at very low prevalence of BSE demonstrates the effectiveness of monitoring sign combinations as detectors for specific disease. This highlights the surveillance value of detailed individual event descriptions provided by BOSSS.

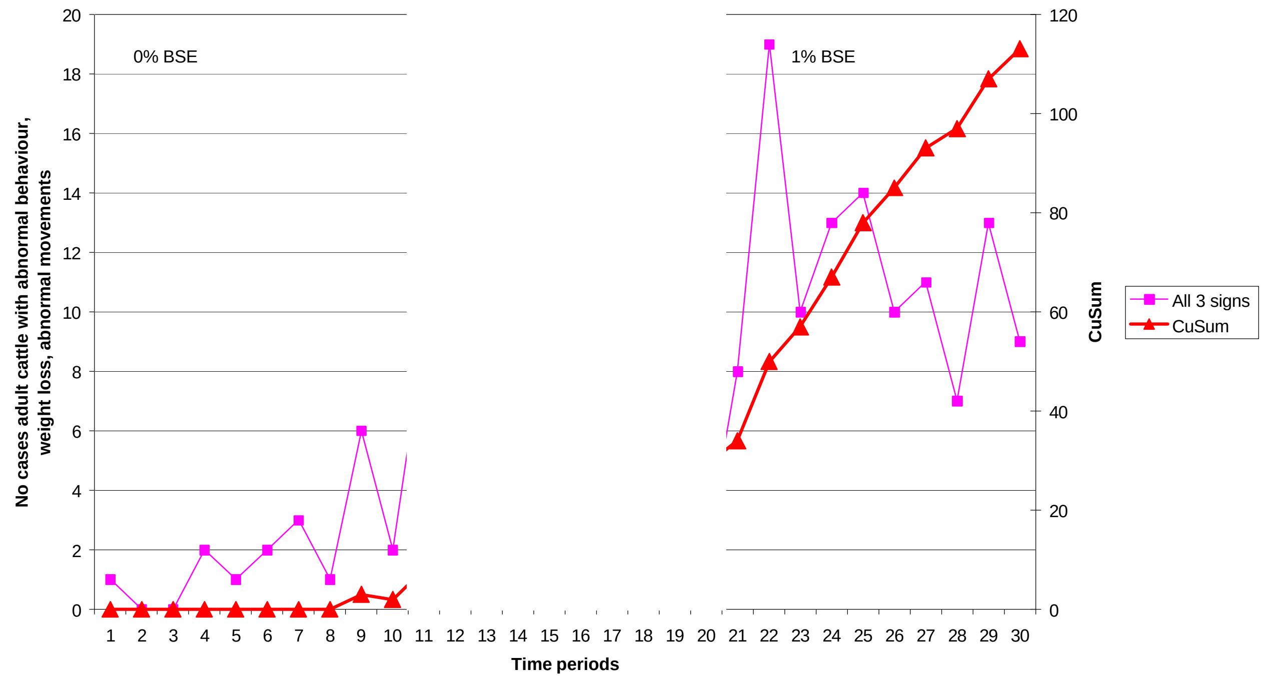

Figure 1 Number of cases with abnormal behaviour, weight loss and abnormal movement and three-sign CuSUM at various BSE prevalences. An alarm cut point for the CuSUM can be applied, This is calculated from the baseline distribution of signs.

WSARE was set to use a sliding window of 5 months. Comparison was made with aggregated data from an equivalent 5 month period that finished 6 months before the current time period began. These wide windows were necessary to control for the relative scarcity of data compared to human surveillance systems. Under these constraints, WSARE detected disease with adequate sensitivity and false alarm rate but detection was typically not timely. WSARE detection sensitivity increased as the prevalence of infected farms increased exponentially (Figure 2).

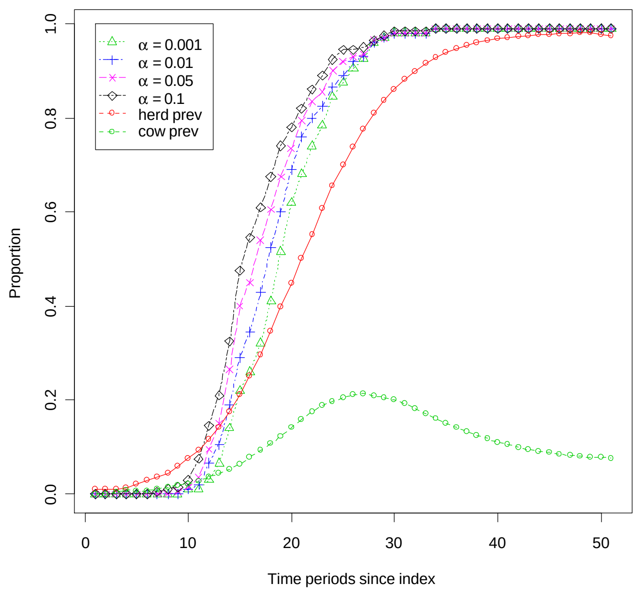

Figure 2 WSARE detection sensitivity at various algorithm detection confidence levels (a), herd prevalence and animal prevalence versus time since index case for simulated disease spread within a population

Sufficient data are necessary to detect disease in a timely manner using WSARE. A narrow time window is essential for timely detection but narrow time windows require a high data collection rate. A voluntary syndrome reporting system may not provide sufficient data to produce timely detection. Greater automation and increased data capture will be essential.

The artificial intelligence system deployed within BOSSS also provides a monitoring system for disease. This case-based analytical system provides the surveillance managers with reasoned interpretation of individual events. The differential list of diseases can be used to guide further surveillance activity as required. This may be as simple as telephone contact with the observer. The final worth of BOSSS as a surveillance tool will be dependent upon the quality and extent of data captured and the systems used to process the data.

References

Frawley, P. T. (2003) AFFA, Canberra, pp. 145.

Lombardo, J., Burkom, H., Elbert, E., Magruder, S., Lewis, S. H., Loschen, W., Sari, J., Sniegoski, C., Wojcik, R. and Pavlin, J. (2003) *Journal of Urban Health,* **80,** i32-42.

O'Brien, S. J. and Christie, P. (1997) *Public Health,* **111,** 255-8.

R Development Core Team (2004) (Ed, Computing, R. F. f. S.) Vienna, Austria.

Tsui, F. C., Espino, J. U., Dato, V. M., Gesteland, P. H., Hutman, J. and Wagner, M. M. (2003) *Journal of the American Medical Informatics Association,* **10,** 399-408.

Wagner, M. M. (Ed.) (2005) *Handbook of Biosurveillance,* Elsevier, Pittsburgh.

Wagner, M. M., Tsui, F. C., Espino, J. U., Dato, V. M., Sittig, D. F., Caruana, R. A., McGinnis, L. F., Deerfield, D. W., Druzdzel, M. J. and Fridsma, D. B. (2001) *Journal of Public Health Management & Practice.,* **7,** 51-9.

Wong, W. K., Moore, A., Cooper, G. and Wagner, M. (2003) *Journal of Urban Health,* **80,** i66-75.