Using a Bayesian network to model colonisation with Vancomycin Resistant Enterococcus (VRE)

M Rajmokan[a], A Morton[b,c], K Mengersen[c], L Hall[a] and M Waterhouse[d]

[a] Mohana_rajmokan@health.qld.gov.au

[a] Centre for Healthcare Related Infection Surveillance and Prevention (CHRISP), Level 3, 15 Butterfield Street, Herston, Qld 4006

[b] Infection Management Services, Princess Alexandra Hospital, Brisbane, Australia

[c] School of Mathematical Sciences, Queensland University of Technology, Brisbane, Australia

[d] The Wesley Research Institute, Coronation Drive, Auchenflower, Brisbane, Australia

Keywords:

Bayesian Network; VRE prevalence; VRE transmission

ABSTRACT

Objective: Effective management of multi-resistant organisms is an important issue for hospitals both in Australia and overseas. This study investigates the utility of using Bayesian Network (BN) analysis to examine relationships between risk factors and colonization with Vancomycin Resistant Enterococcus (VRE).

Design: Bayesian Network Analysis was performed using infection control data collected over a period of 36 months (2008-2010).

Setting: Princess Alexandra Hospital (PAH), Brisbane.

Outcome of interest: Number of new VRE Isolates

Methods: A BN is a probabilistic graphical model that represents a set of random variables and their conditional dependencies via a directed acyclic graph (DAG). BN enables multiple interacting agents to be studied simultaneously. The initial BN model was constructed based on the infectious disease physician's expert knowledge and current literature. Continuous variables were dichotomised by using third quartile values of year 2008 data. BN was used to examine the probabilistic relationships between VRE isolates and risk factors; and to establish which factors were associated with an increased probability of a high number of VRE isolates.

Software: Netica (version 4.16).

Results: Preliminary analysis revealed that VRE transmission and VRE prevalence were the most influential factors in predicting a high number of VRE isolates. Interestingly, several factors (hand hygiene and cleaning) known through literature to be associated with VRE prevalence, did not appear to be as influential as expected in this BN model.

Conclusions: This preliminary work has shown that Bayesian Network Analysis is a useful tool in examining clinical infection prevention issues, where there is often a web of factors that influence outcomes. This BN model can be restructured easily enabling various combinations of agents to be studied.

1.1 Introduction:

Vancomycin-resistant Enterococci (VRE) are increasingly important multiple antibiotic resistant organisms (MROs) in Australian hospitals[4]. At the Princess Alexandra Hospital (PAH) in Brisbane, Australia, there was an outbreak of VRE between 1996 and 1999 that was controlled by the application of enhanced environmental cleaning, isolation of colonised patients and reduction of inpatient admissions[1]. Thereafter until 2007 the rate at which new VRE isolates appeared at PAH was low. However, between 2008 and 2010 there was a steady and substantial rise in the prevalence of VRE at PAH. To provide effective control of VRE, a system is required that can predict VRE prevalence and transmission and that can be updated regularly as new evidence accumulates. We describe a Bayesian network (BN) model to analyse the factors that are influential in the occurrence of new VRE isolates.

BNs have become increasingly popular as decision support tools among those researching the use of artificial intelligence, probability and uncertainty. We

have implemented a BN as potentially a most useful approach for inference and prediction of risk factors for multiple antibiotic –resistant organisms (MROs). A key feature of this is the use of expert opinion and existing data to set up the initial model and the continuing enhancement of the model's predictive capabilities as new data accumulate. The system can provide a model for colonisation risk analysis that constantly adapts in response to changes in the patterns of MRO colonisations. The ability to update the initial distribution on receipt of new data makes the BN approach a natural choice for the analysis of MRO colonisation risk factors.

1.2 Analysis of the risk factors influential to VRE isolates:

The list of risk factors that were considered to be potentially influential to VRE colonisation used in the analysis is shown in Table1.

Table1. The variable name and states of variables used for the analysis:

Variable Name	States of variables	Explanation of Variable name
VRE Isolates	0=Normal; 1=High	VRE new clinical isolates
- VRE Prevalence	0=Normal; 1=High	VRE Prevalence
- Ceph_usage	0=Normal; 1=High	Third generation cephalosporin
- Vancomycin_usage	0=Normal; 1=High	Vancomycin usage
- VRE_carriers_entering_Hospital	0=Normal; 1=High	VRE Carriers enter in to Hospital
- Known_VRE_carriers	0=Normal; 1=High	Known VRE carriers
- Transferred_patients	0=Normal; 1=High	Transferred patients from other facility
- Readmitted patients	0=Normal; 1=High	Readmitted patients
- VRE Transmission	0=Normal; 1=High	VRE Transmission
- Handwashing	0= Satisfactory; 1= Unsatisfactory	Hand washing compliance
- Cleaning Audits	0= Satisfactory; 1= Unsatisfactory	Environmental cleaning audits
- Screening	0= Normal; 1 = High	Screening
- Ward outliers	0= Normal; 1 = High	Ward outliers
- Staffing	0= Satisfactory; 1= Unsatisfactory	Staff level
- Percent_casual	0=Normal; 1=High	Percentage of casual staff
- Staff_1000_OBD	0=Normal; 1=High	Staff per 1000 Occupied bed days
- ISO Ward Overflow	0=Normal; 1=High	Isolation ward overflow
- MRO Prevalence	0=Normal; 1=High	MRO prevalence
- MRO isolates	0=Normal; 1=High	MRO isolates
- Overcrowding	0=Normal; 1=High	Over crowding
- OT_cancellations	0=Normal; 1=High	Operating Theatre cancellations
- ED Access Block	0=Normal; 1=High	Emergency department access block
- Percentage_bed_occupancy	0=Normal; 1=High	Percentage of bed occupancy
- Ward Outliers	0=Normal; 1=High	Ward outliers

1.3 Methods (Bayesian Networks)

A BN is a complex systems model that probabilistically describes the way in which a set of variables interact to influence an outcome. The model is typically represented graphically (as a directed acyclic graph, or DAG) with variables depicted as nodes and the relationships between the variables depicted as directed links. If a link goes from node 1 to node 2, then node 1 is said to be a parent of node 2, and node 2 is a child of node 1. A node can be deterministic, i.e. a function of its parent nodes, or stochastic, with probabilities conditioned on the values of its parent nodes. If a node is discrete, these probabilities can be represented as a conditional probability table (CPT).

A BN is a sensible model for risk assessment of rare events[2]. In this paper, the outcome was defined to be new VRE isolates in a hospital environment. The aims were threefold: to construct a BN to describe potential risk factors associated with this outcome; to quantify the BN using data obtained from PAH; and to evaluate the predictive ability, sensitivity and robustness of the resultant model.

1.4 BN Construction:

The BN model was constructed using 22 variables (Table1) that were identified based on medical literature and the expert knowledge of infectious diseases staff at PAH

.

1.5 BN Quantification:

Infection control data were collected at PAH over a period of 36 months from January 2008 to December 2010. Nineteen variables were dichotomised into 'high' and 'normal' levels based on the third quartile of a subset of the 2008 data, and three variables (cleaning audits, handwashing compliance, staffing) were categorised into 'satisfactory' and 'unsatisfactory'. All variables were based on reported measurements, with the exception of staffing which was based on the following formula[6]

$$staffing = 4\left(\frac{staff\ /1000OBD}{93.0}\right) - \left(\frac{\%casual}{0.461}\right) - 0.25I$$

where 93.0 and 0.461 are the 2008 monthly mean values for staff/1000 OBD and % casual, respectively, and I is an indicator that equals 1 if the data were collected in January or February and 0 otherwise. Conditional probability tables were estimated by cross-tabulation of the relevant data; if a cell in the table was not observed or collected, linear regression models with all parent nodes as predictors were used to obtain estimates.

1.6 BN Evaluation:

The model was analysed using the software packages Netica[5]. Four evaluations of the BN were conducted. First, the probability of a high level of VRE isolates was obtained based on the input data. Second, sensitivity analysis determined the order in which factors influenced outcomes of interest, based on the mutual information between each node and the target node. Mutual information quantifies the extent to which a finding at one node reduces the uncertainty regarding the other node, with higher values indicating a stronger dependency between two nodes. Third, scenario analyses were performed to determine the effect of entering evidence for VRE transmission and VRE prevalence on the response. Finally, the robustness of the model were assessed by quantifying the network using 20 random subsets of data and comparing the results obtained from these partial networks to those from the complete model.

1.7 Results

The constructed BN is depicted in Figure 1. Based on this model and the input data, the baseline probability of a high number of VRE isolates was 0.0224. According to this model, VRE transmission and VRE prevalence were the two factors that were directly linked to new VRE isolates. Based on the input data, the probability of high level of VRE transmission was 0.174 and the probability of high level of VRE prevalence was 0.121.

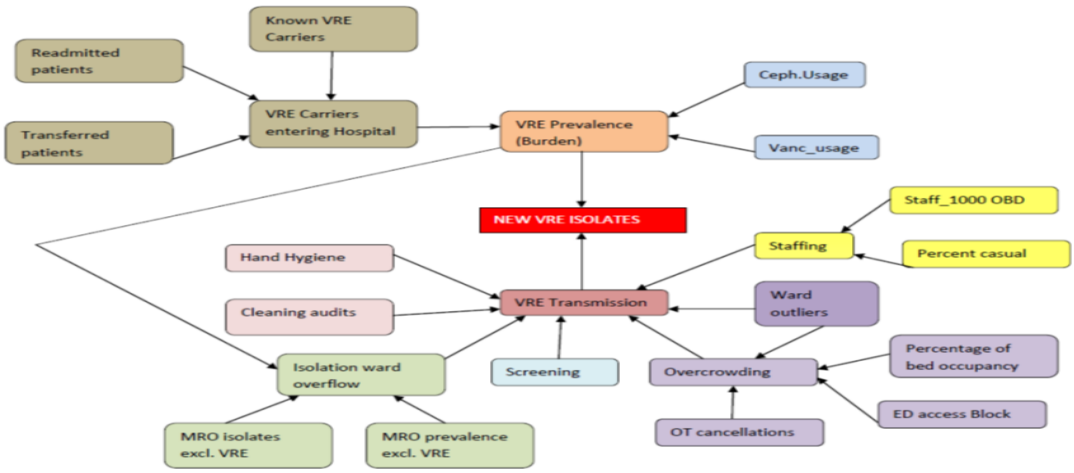

Figure1. Structure of the BN used to model colonization with VRE

Table 2 orders the factors from most to least influential with respect to the number of new VRE isolates. VRE transmission was the most influential factor with respect to VRE new isolates. In terms of mutual information, VRE prevalence was the next most influential factor, but was only about a third (34.6%) as important as VRE transmission. The next four influential factors were vancomycin usage, screening, hand washing and cleaning audits respectively; these three factors were less than 5% as important as VRE transmission. The factors that considerably influenced VRE transmission were, in order, number of screenings, handwashing compliance, cleaning audits, isolation ward overflow and ward outliers.

Table2. The probability, p, that the factor is at the specified level, with nodes listed from most to least influential with respect to VRE isolates based on the mutual information.

Factor	Level	Mutual information	Importance relative to VRE transmission (%)	*p*
New VRE isolates	-	0.15497		0.022
VRE transmission	High	0.02487		0.174
VRE Prevalence	High	0.0086	34.6	0.121
Vancomycin usage	High	0.00064	2.6	0.447
Screening	High	0.00042	1.7	0.579
Hand washing	Unsatisfactory	0.00035	1.4	0.526
Cleaning audits	Unsatisfactory	0.00032	1.3	0.500
Ceph. usage	High	0.00022	0.9	0.289
VRE Carriers Entering Hospital	High	0.0002	0.8	0.281
Ward outliers	High	0.00015	0.6	0.316
Over crowding	High	0.00007	0.3	0.169
Staffing	Unsatisfactory	0.00004	0.2	0.222
Isolation ward overflow	High	0.00003	0.1	0.422
Readmitted patients	High	0.00001	0.04	0.684
Known VRE Carriers	High	0.00001	0.04	0.553
Staff per 1000 OBD	High	0.00001	0.04	0.684
MRO Isolates	High	0	0	0.368
Transferred patients	High	0	0	0.263
MRO Prevalence	High	0	0	0.105
Operating Theatre Cancellations	High	0	0	0.132
Percentage bed occupied	High	0	0	0.132
Emergency Department Access block	High	0	0	0.368
% casual	High	0	0	0.368

Table 3 shows the results of the scenario analysis for VRE transmission and VRE prevalence. Interest focused on the change in the probability of a high level of VRE isolates from 0.0224, obtained under baseline standard conditions. The results confirm that VRE transmission was the most influential risk factor.

Table3. The probability, p, that VRE isolates is high for various levels of VRE transmission and VRE prevalence. We use "-" to indicate that no evidence has been entered for factor.(For example, the first row of considers the scenario where VRE transmission is normal with probability 1, while VRE prevalence is unchanged from the baseline level)

VRE transmission	VRE Prevalence	Probability (p)
Normal	-	0.0074↓ from 0.024
High	-	0.094 ↑ from 0.024
-	Normal	0.0152 ↓from 0.024
-	High	0.0752 ↑ from 0.024
Normal	Normal	0 ↓ from 0.024
Normal	High	0.0601↑from 0.024
High	Normal	0.087 ↑ from 0.024
High	High	0.148 ↑ from 0.024

Results from the robustness assessment are shown in Table 4. It shows the proportion of times a node had the same rank in the sensitivity to finding analysis when compared to the analysis using the model quantified using the complete dataset. The nodes VRE transmission, VRE prevalence that ranked highly based on the complete dataset, tended to have the same importance when using the reduced datasets. The nodes that were identified as having a very low rank based on the complete dataset (little or no association with VRE isolates), tended to have the same indication when using the reduced datasets. On thirteen occasions vancomycin usage was the third most influential factor and on seven occasions screening was the third most influential factor based on the reduced datasets.

Table4. Proportion of times each node had the same rank in the sensitivity to findings analysis for VRE isolates based on 20 reduced datasets, when compared to the model quantified using the complete dataset.

Factor	Proportion agreement
VRE transmission	1.00
VRE Prevalence	1.00
Vancomycin usage	0.65
Screening	0.50
Hand washing	0.40
Cleaning audits	0.65
Ceph_usage	0.35
VRE Carriers Entering in Hospital	0.60
Ward Outliers	0.65
Over crowding	0.55
Staffing	0.30
Isolation ward overflow	0.20
Readmitted patients	0.50
Known VRE Carriers	0.35
Staff per 1000 OBD	0.30
Transferred patients	0.15
MRO Isolates	0.10
MRO Prevalence	0.25
Percentage bed occupancy	0.50
Operating Theatre Cancellations	0.40
Emergency Department access block	0.30
% casual	0.60

1.8 Conclusion:

The study indicated that Transmission appears to be more important than Prevalence even when the latter is increased by substantial numbers of carriers referred from other institutions. It also suggests that Hand Hygiene and Cleaning have a relatively minor effect with respect to new VRE isolates. It is being recognised that currently used cleaning methods and their surveillance require change[3] and it appears that hand hygiene may need to be at a higher level to control VRE than MRSA[6]

Since several nodes and their connections seem redundant, it will be useful to consider some pruning of the BN, e.g. the nodes Readmitted and Transferred may now be redundant because carriers in these categories are now being identified as Known carriers as shown in Figure1.

The mutual interdependence of prevalence and transmission could not be assessed using this DAG model in Netica. This requires further evaluation.

References:

1. Bartley P, Schooneveldt J, Looke D, et.al.(2001), The relationship of a clonal outbreak of *Enterococcus faecium*vanA to methicillin-resistant Staphylococcus aureus incidence in an Australian hospital, *Journal of Hospital Infection* , **48(1),** 43-54.

2. Brooker P. (2011), Experts, Bayesian Belief Networks, rare events and aviation risk estimates. *Safety Science,* **49(8-9),** 1142-1155.

3. Cooper R, Griffith C, Malik R, Obee P and Looker N,(2007), Monitoring the effectiveness of cleaning in four British hospitals, *American Journal of Infection Control* 2007, **35(3),** 338-341.

4. Johnson P, Ballard S, Grabsch E, et.al.,(2010), A Sustained Hospital Outbreak of Vancomycin-Resistant *Enterococcus faecium* Bacteraemia due to Emergence of vanB *E. faecium* Sequence Type 203, *The Journal of Infectious Diseases,***202(8),**1278-1286.

5. Netica. Norsys Systems Corp. Vancouver, Canada.

6. Waterhouse M, Morton A, Mengersen K, et.al.(2011), Investigating the role of overcrowding in MRSA transmission using a Bayesian Network, Results for a single public hospital, *Journal of Hospital Infection* , **78(2),** 92-96.