White Rose Research Online

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk
Universities of Leeds, Sheffield and York

Deposited via The University of Leeds.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/142957/

Version: Accepted Version

Article:
Lopez Garcia, M, King, MF and Noakes, C (2019) A multi-compartment SIS stochastic model with zonal ventilation for the spread of nosocomial infections: detection, outbreak management and infection control. Risk Analysis, 39 (8). pp. 1825-1842. ISSN: 0272-4332

https://doi.org/10.1111/risa.13300

©2019 Society for Risk Analysis. This is the pre-peer reviewed version of the following article: Lopez Garcia, M , King, MF and Noakes, C (2019) A multi-compartment SIS stochastic model with zonal ventilation for the spread of nosocomial infections: detection, outbreak management and infection control. Risk Analysis. ISSN 0272-4332, which has been published in final form at https://doi.org/10.5518/355. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions.

Reuse
Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

UNIVERSITY OF LEEDS
University of Sheffield
UNIVERSITY of York

A multi-compartment SIS stochastic model with zonal ventilation for the spread of nosocomial infections: detection, outbreak management and infection control

Abstract

In this work we study the environmental and operational factors that influence airborne transmission of nosocomial infections. We link a deterministic zonal ventilation model for the airborne distribution of infectious material in a hospital ward, with a Markovian multi-compartment SIS model for the infection of individuals within this ward, in order to conduct a parametric study on ventilation rates and their effect on the epidemic dynamics. Our stochastic model includes arrival and discharge of patients, as well as the detection of the outbreak by screening events or due to symptoms being shown by infective patients. For each ventilation setting, we measure the infectious potential of a nosocomial outbreak in the hospital ward by means of a summary statistic: the number of infections occurred within the hospital ward until end or declaration of the outbreak. We analytically compute the distribution of this summary statistic, and carry out local and global sensitivity analysis in order to identify the particular characteristics of each ventilation regime with the largest impact on the epidemic spread. Our results show that ward ventilation can have a significant impact on the infection spread, especially under slow detection scenarios or in over-occupied wards, and that decreasing the infection risk for the whole hospital ward might increase the risk in specific areas of the health-care facility. Moreover, the location of the initial infective individual and the protocol in place for outbreak declaration both form an interplay with ventilation of the ward.

Keywords: nosocomial infections, airborne spread, SIS stochastic model, outbreak detection, summary statistic

1 Introduction

The risk of acquiring nosocomial infections is a recognised problem in health-care facilities worldwide.[9] While the transmission routes for some diseases are well documented, the precise mode of transmission is uncertain for many infections, particularly for those pathogens that cause health-care acquired infections (HCAI). Although it is probable that the majority of transmission occurs via contact routes,[25] there is increasing recognition that the air plays an important role in disease spread.[15] Understanding the role that ventilation airflow plays in the dispersion of infectious microorganisms is tantamount to assessing exposure to pathogens and hence infection risk. This study aims to provide an analytical link between airborne hospital infection spread, ventilation design and outbreak management.

Ventilation has been found to have a significant impact on the distribution of infectious material in hospital settings. Examples include *Influenza A* (e.g. H5N1 and H7N9),[24] *M. tuberculosis*,[6] measles[1] and *norovirus*.[28] One of the most infamous examples occurred in 2003 during the severe acute respiratory syndrome (SARS) outbreak in Hong Kong. Analysis of airflow patterns and outbreak data demonstrated that ventilation routes were critical in the short- and long-range spread of aerosolised *coronavirus*.[14] Ventilation is recognised as an important infection control approach in health-care design, with strategies such as mechanical ventilation and pressure zoning set out in international[1] and national guidance.[4]

For Peer Review

Evaluating the influence of ventilation on infection risk typically applies models such as the Wells-Riley equation[23] or a dose-response approach[26] to estimate the influence of ventilation on the number of new cases of an infection. Liao et al.[16] presented a probabilistic transmission dynamic model to assess indoor airborne infection risks and Ko et al.[12,13] developed models for Tuberculosis spread including incorporating a zonal ventilation model. A number of authors have also looked at control strategies including Wein et al.[29] who modelled infection control measures for pandemic influenza, Brienen et al.,[2] where authors analyse the effect of mask use on the spread of influenza, and King et al.[11] who developed a stochastic model to link airborne and contact transmission. It is also worth mentioning the recent work by Carruthers et al.,[3] where a zonal ventilation model similar to the one considered in this paper is linked to a dose-response approach to estimate the risk of infection after an accidental release of bacteria *Francisella tularensis* in a microbiology laboratory.

While these studies enable some understanding of the influence of the environment on transmission, they do not consider relationships between ventilation parameters and the progression and control of an infection outbreak. In an earlier study it was demonstrated that the Wells-Riley model could be coupled to an SI epidemic model to relate ventilation rate and transmission in a fully-mixed environment.[19] In later work a zonal air distribution and a stochastic formulation[20] was considered, and cost-benefits of ventilation from an energy and infection risk perspective were explored.[21]

The model presented in this paper is constructed on a scenario defined in previous work,[20] where the role played by the airflow during a nosocomial outbreak is assessed by linking a deterministic zonal ventilation model with an SI stochastic epidemic model using a computational approach. While the previous approach enabled exploration of the basic interaction between the ventilation and the outbreak, there are a number of limitations:

- The epidemic dynamics are represented through a simple SI epidemic model, not accounting for relevant factors such as the discharge and admission of patients, or the detection and declaration of the outbreak.
- Results reported in[20] have high variability, which is related to the fact that they were obtained by means of stochastic simulations of the epidemic process.
- The large number of parameters associated with each ventilation scenario makes it difficult to identify, from stochastic simulations, the specific factors of the ventilation air distribution that facilitate or mitigate epidemic spread.

We refer the reader to[10] where the limitations of analysing this type of epidemic processes by simulation are discussed in more detail, and where the benefits of following exact analytical approaches instead are highlighted.

Our aim here is to show how this zonal ventilation model can be linked to more complex stochastic epidemic models for the spread of nosocomial pathogens, while accounting for patients admission and discharge, and different outbreak detection and declaration hypotheses. We show how to implement exact analytical procedures for computing summary statistics of the outbreak (statistics measuring outbreak infectiousness), which by means of a perturbation analysis enables identification of specific characteristics of the ventilation setting that are crucial for the spread or control of the infection.

Finally, we carry out a comprehensive numerical study of six ventilation strategies for an hypothetical hospital ward in order to identify particular ventilation characteristics that may promote or inhibit spread of airborne nosocomial infections. Our results explore the interplay between ward ventilation, location of patients, ward over-occupancy and outbreak detection

For Peer Review

management.

2 The model

In Noakes et al.[20] a stochastic model is proposed that links a zonal ventilation model with epidemic dynamics by means of an SI model for the infection spread among individuals in a hospital setting. The combined model can be thought of as an adaptation of the *Wells-Riley* model,[23,26] where each susceptible individual is infected with a *per capita* rate

$$\lambda = \frac{Ipq}{Q},$$

that is proportional to the number I of infective individuals. Here Q is the room ventilation rate ($m^3 s^{-1}$), p is the pulmonary ventilation (breathing) rate ($m^3 s^{-1}$) and q is the unit of infection (*quantum*) as described in Riley et al.[23] In a standard SI stochastic epidemic model, where individuals are either susceptible or infective, individuals do not recover. This can be expressed in terms of a continuous-time Markov chain (CTMC) $\{S(t) : t \geq 0\}$ where the random variable $S(t)$ represents the number of susceptible individuals at time t, and the aim is to compute transient probabilities $p_S(t) = \mathbb{P}(S(t) = S)$, for any $S \in \Omega = \{0, 1, \ldots, N\}$ where N is the number of individuals in the population. These probabilities satisfy the Kolmogorov differential equations

$$\frac{dp_S(t)}{dt} = -\lambda S p_S(t) + \lambda(S+1)p_{S+1}(t), \quad S \in \Omega.$$

Inter-event times are assumed to be exponentially distributed and the stochastic process is simulated using the Gillespie algorithm.[7] In particular, this algorithm generates samples of the inter-event times by using the fact that

$$U \sim U(0,1) \;\Rightarrow\; T = -\frac{ln(U)}{\lambda S} \sim Exp(\lambda S), \tag{1}$$

and then updates the state of the system according to the probabilities of each possible event occurring.

2.1 A zonal ventilation model for linking airflow dynamics and infection rates

In Noakes et al.[20] this Wells-Riley process is adapted to investigate ventilation scenarios in an hypothetical hospital ward split in M *ventilation zones*. The air is assumed to be uniformly mixed within each zone, however, there is incomplete mixing between the zones and unbalanced zone boundaries allow for the effect of directional flow to be examined. In particular the *per capita* infection rate λ_k for susceptible individuals at zone k is defined as $\lambda_k = p_k C_k$, where C_k is the concentration of infectious material at zone k, and p_k is the pulmonary rate of these individuals. We note that this concentration could depend on the amount of infected individuals, i_j, in every zone $1 \leq j \leq M$, due to airflow. In Noakes et al.[20] the spatial distribution of infectious material is represented through the differential equation

$$V_k \frac{dC_k}{dt} = q_k i_k - Q_{o,k} C_k - \sum_j \beta_{kj} C_k + \sum_j \beta_{jk} C_j, \quad 1 \leq k \leq M, \tag{2}$$

For Peer Review

where V_k is the volume of zone k, $q_k i_k$ is the generation rate of infectious quanta, $Q_{o,k}$ represents the extract ventilation rate in zone k, and $\sum_j \beta_{kj} C_k$ and $\sum_j \beta_{jk} C_j$ amount to the volume flow of air between zones k and j. Moreover, each inter-zonal flow rate β_{kj} represents the sum of two contributions

$$\beta_{kj} = \beta_0 + \beta_{Qkj},$$

where β_0 is a global mixing rate and β_{Qkj} is an additional contribution representing net flow across the k/j zonal boundary, from zone k to zone j.

Eq. (2) leads to a *ventilation matrix* that characterises the ventilation air distribution under study,

$$\mathbf{V} = \begin{pmatrix} Q_{o,1} + \sum_k \beta_{1k} & -\beta_{21} & \dots & -\beta_{M-1,1} & -\beta_{M1} \\ -\beta_{12} & Q_{o,2} + \sum_k \beta_{2k} & \dots & -\beta_{M-1,2} & -\beta_{M2} \\ -\beta_{13} & -\beta_{23} & \dots & -\beta_{M-1,3} & -\beta_{M3} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ -\beta_{1M} & -\beta_{2M} & \dots & -\beta_{M-1,M} & Q_{o,M} + \sum_k \beta_{Mk} \end{pmatrix}$$

representing ventilation in a hospital ward divided into M ventilation zones.

By assuming steady-state conditions for airflow, and taking into account Eq. (2), authors then propose in[20] to link infection rates λ_j, $j \in \{1, \dots, M\}$, with the ventilation matrix $\mathbf{V}$ as follows:

$$\begin{pmatrix} \frac{\lambda_1}{p_1} \\ \frac{\lambda_2}{p_2} \\ \vdots \\ \frac{\lambda_M}{p_M} \end{pmatrix} = \mathbf{V}^{-1} \begin{pmatrix} q_1 i_1 \\ q_2 i_2 \\ \vdots \\ q_M i_M \end{pmatrix}. \tag{3}$$

This means that *per capita* infection rates $\{\lambda_1, \dots, \lambda_M\}$ for susceptible individuals at zones $\{1, \dots, M\}$ depend on how many infective individuals $(i_1, \dots, i_M)$ there are at any zone at any given time, computed in an specialised manner (Eq. (3)) that takes into account the ventilation distribution through matrix $\mathbf{V}$. Once this procedure for computing infection rates is proposed, stochastic simulations for the SI epidemic dynamics are carried out in[20] by following steps (i)-(iv),[20] which assume exponentially distributed inter-event times and make use of the property depicted in Eq. (1).

We note here that Eq. (3) means considering per capita infection rates at each zone $1 \leq j \leq M$ as functions $\lambda_j(i_1, \dots, i_M)$ of the number of infectives $(i_1, \dots, i_M)$ within each ventilation zone at the hospital ward. In subsection 2.2, we now go on to exploit this, to link the zonal ventilation model with a multi-compartment SIS model with detection, to evaluate the infection spread dynamics within the hospital ward until detection of the outbreak. Instead of carrying out stochastic simulations, we present an exact approach for analysing a summary statistic of the outbreak: the total number of infections occurring until the outbreak ends or is detected and declared. This exact approach does not only allow us to compute this quantity of interest, but it also allows one to carry out a sensitivity analysis on the model parameters, so that the impact that different characteristics of the ventilation setting has on this summary statistic can be evaluated.

For Peer Review

2.2 A multi-compartment SIS stochastic model for the infection spread dynamics

At the epidemic level, we assume that patients at each zone i are discharged at rate γ_i, so that γ_i^{-1} amounts to the average length-of-stay (LOS) of patients at zone i. Discharges are immediately replaced by new admitted patients; a reasonable approximation for hospital wards under high demand.[22,30] Moreover, we consider that the nosocomial outbreak will go undetected by health-care workers for some time, and incorporate this fact into our model by considering that each infected individual at zone i can be *discovered/detected* at some rate δ_i. The reciprocal δ_i^{-1} represents the average time until some symptoms arise which alert health-care workers to a patient's infection, or the average time until the infected individual is detected through screening policies put in place at this hospital ward. Figure 1 represents the epidemic dynamics for an individual at zone i.

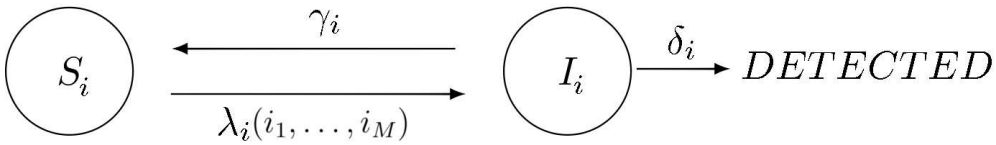

Figure 1: Individual epidemic dynamics for a patient at zone i. Event $I_i \to S_i$ represents the discharge of the infective patient at zone i, immediately replaced by a susceptible one.

This leads to a multi-compartment SIS epidemic model, that can be described as a CTMC $\mathcal{X} = \{(I_1(t), \ldots, I_M(t)) : t \geq 0\}$, where $I_j(t)$ represents the number of infective individuals at zone j at time $t \geq 0$, defined over the space of states $\mathcal{S} = \{(i_1, i_2, \ldots, i_M) : i_j \in \{0, \ldots, N_j\},\ j \in \{1, \ldots, M\}\} \cup \{\Delta\} = \mathcal{C} \cup \{\Delta\}$. N_j is the total number of patients at zone j, leading to $N = \sum_{j=1}^{M} N_j$ patients in the hospital ward. State Δ represents that the nosocomial outbreak has been detected and declared by health-care workers by the first detection of an infected patient in the hospital ward. We note that absorbing state $(0, \ldots, 0)$ represents the end of the outbreak (lack of infective individuals), due to patients discharge (that is, if all the patients infected by the pathogen are discharged before the outbreak is actually detected). We consider Δ also as an absorbing state in this process, since we are only interested in the dynamics of the process until the end or declaration of the outbreak, and the transitions (obtained from diagram in Figure 1) described in Table 1. We note that, according to our comments in subsection 2.1, $\lambda_j(i_1, \ldots, i_M)$ is a function of the state $(i_1, \ldots, i_M)$, representing the per capita infection rate of susceptible individuals at zone j when we have $(i_1, \ldots, i_M)$ infective individuals within the ward, computed from Eq. (3) for each $(i_1, \ldots, i_M) \in \mathcal{C}$.

3 Methodology

Our interest is in analysing the infectious potential of an outbreak in a given hospital ward, for different ventilation configurations and outbreak detection hypotheses. We measure the infectious potential in terms of the following discrete random variable:

$$R \quad = \quad \textit{"number of infections occurred until the end or declaration of the outbreak"},$$

Table 1: Transitions or events occurring in the stochastic process $\mathcal{X}$, and corresponding rates

Event	**Stochastic transition**	**Rate**
Infection of a patient at zone j	$(i_1,\ldots,i_j,\ldots,i_M) \to (i_1,\ldots,i_j+1,\ldots,i_M)$	$\lambda_j(i_1,\ldots,i_M)(N_j - i_j)$
Discharge of an infective patient at zone j	$(i_1,\ldots,i_j,\ldots,i_M) \to (i_1,\ldots,i_j-1,\ldots,i_M)$	$\gamma_j i_j$
Detection of the outbreak	$(i_1,\ldots,i_M) \to \Delta$	$\sum_{j=1}^{M} \delta_j i_j$

which can take values in $\{0,1,2,\ldots\}$. R is used here as a measure of how well ventilation can act as a *preventive infection control strategy*, which is in place even before the actual detection of the outbreak by health-care workers occurs, and *reactive* strategies can be implemented.

3.1 Analysing R

For an initial state $(i_1,\ldots,i_M) \in \mathcal{C}$, our aim is to compute probabilities

$$p_{(i_1,i_2,\ldots,i_M)}(n) \;=\; \mathbb{P}(R = n \mid (I_1(0), I_2(0),\ldots,I_M(0)) = (i_1,i_2,\ldots,i_M)),$$

for $n \geq 0$; that is, the probability distribution of R for some initial state $(i_1,\ldots,i_M)$. We can compute these probabilities from a system of linear equations, which is obtained by a first-step argument. In particular, by proposing notation

$$\begin{aligned} \mathbf{i} &= (i_1,\ldots,i_M), \\ \mathbf{i}^+(s) &= (i_1,\ldots,i_s+1,\ldots,i_M), \\ \mathbf{i}^-(s) &= (i_1,\ldots,i_s-1,\ldots,i_M), \end{aligned}$$

we get

$$p_{\mathbf{i}}(0)\sum_{j=1}^{M}(\lambda_j(\mathbf{i})(N_j - i_j) + (\gamma_j+\delta_j)i_j) \;=\; \sum_{k=1}^{M}\left(\gamma_k i_k p_{\mathbf{i}^-(k)}(0) + \delta_k i_k\right), \tag{4}$$

$$p_{\mathbf{i}}(n)\sum_{j=1}^{M}(\lambda_j(\mathbf{i})(N_j - i_j) + (\gamma_j+\delta_j)i_j) \;=\; \sum_{k=1}^{M}\left(\gamma_k i_k p_{\mathbf{i}^-(k)}(n) + \lambda_k(\mathbf{i})(N_k - i_k)p_{\mathbf{i}^+(k)}(n-1)\right), \tag{5}$$

for $n \geq 1$ and any $(i_1,\ldots,i_M) \in \mathcal{C}$, and with boundary conditions $p_{(0,0,\ldots,0)}(0) = 1$, $p_{(0,0,\ldots,0)}(n) = 0$ for all $n \geq 1$; a detailed explanation on how Eqs. (4)-(5) are obtained is in the Appendix. This means that probabilities for $n = 0$ (Eq. (4)) can be computed by solving a system of

$$\#\mathcal{C} \;=\; \prod_{i=1}^{M}(N_i + 1) \tag{6}$$

linear equations*. Once these are in hand, probabilities for $n \geq 1$ can be computed by solving the system of linear equations given by Eq. (5), which also consists of $\#\mathcal{C}$ equations.

Algorithm 1 in the Appendix computes probabilities $p_{(i_1,\ldots,i_M)}(n)$ for any $(i_1,\ldots,i_M) \in \mathcal{C}$, $n \geq 0$. It works sequentially, computing probabilities $p_{(i_1,\ldots,i_M)}(n)$ for $i_1 + i_2 + \cdots + i_M = I$ for increasing values of $n = 0,1,2,\ldots$ and $I = 1,2,3,\ldots,N$.

*$\#\mathcal{C}$ represents the number of elements (states) in $\mathcal{C}$

For Peer Review

3.2 Local sensitivity analysis

Our analysis allows one to identify the most important characteristics of the ventilation scenario, regarding the infectious potential of the outbreak until detection, by means of computing partial derivatives of the form $\partial E[R]/\partial\theta$ with respect to ventilation parameters $\theta \in \{\beta_0\} \cup \{\beta_{Qij} : i,j \in \{1,\ldots,M\}\} \cup \{Q_{o,i} : i \in \{1,\ldots,M\}\}$. We note that, for an initial state $(i_1,\ldots,i_M) \in \mathcal{C}$, $E[R] = \sum_{n=0}^{+\infty} n p_{(i_1,\ldots,i_M)}(n)$, so that

$$\frac{\partial E[R]}{\partial\theta} = \sum_{n=0}^{+\infty} n \frac{\partial p_{(i_1,\ldots,i_M)}(n)}{\partial\theta}. \tag{7}$$

Partial derivatives $\frac{\partial p_{(i_1,\ldots,i_M)}(n)}{\partial\theta}$ can be computed from direct differentiation of Eqs. (4)-(5). In Eqs. (4)-(5), the only quantities that depend on parameter $\theta \in \{\beta_0\} \cup \{\beta_{Qij} : i,j \in \{1,\ldots,M\}\} \cup \{Q_{o,i} : i \in \{1,\ldots,M\}\}$ are infection rates $\lambda_j(i_1,\ldots,i_M)$, and probabilities $p_{(i_1,\ldots,i_M)}(n)$. Thus we get

$$\sum_{j=1}^{M} \frac{\partial\lambda_j(\mathbf{i})}{\partial\theta}(N_j - i_j)p_{\mathbf{i}}(0) + \sum_{j=1}^{M}\Big(\lambda_j(\mathbf{i})(N_j - i_j) + (\gamma_j + \delta_j)i_j\Big)\frac{\partial p_{\mathbf{i}}(0)}{\partial\theta} = \sum_{k=1}^{M}\gamma_k i_k \frac{\partial p_{\mathbf{i}^-(k)}(0)}{\partial\theta},$$

$$\begin{aligned}\sum_{j=1}^{M} \frac{\partial\lambda_j(\mathbf{i})}{\partial\theta}(N_j - i_j)p_{\mathbf{i}}(n) + \sum_{j=1}^{M}\Big(\lambda_j(\mathbf{i})(N_j - i_j) + (\gamma_j + \delta_j)i_j\Big)\frac{\partial p_{\mathbf{i}}(n)}{\partial\theta} &= \sum_{k=1}^{M}\Big(\gamma_k i_k \frac{\partial p_{\mathbf{i}^-(k)}(n)}{\partial\theta} \\ &+ \frac{\partial\lambda_k(\mathbf{i})}{\partial\theta}(N_k - i_k)p_{\mathbf{i}^+(k)}(n-1) + \lambda_k(\mathbf{i})(N_k - i_k)\frac{\partial p_{\mathbf{i}^+(k)}(n-1)}{\partial\theta}\Big),\end{aligned}$$

for $n \geq 1$, and any $(i_1,\ldots,i_M) \in \mathcal{C}$. Partial derivatives $\frac{\partial p_{(i_1,\ldots,i_M)}(n)}{\partial\theta}$ can then be computed from equations above by following similar arguments than in Algorithm 1 in the Appendix. In order to solve these equations, one needs to have in hand values of $p_{(i_1,\ldots,i_M)}(n)$ (previously computed from Algorithm 1), as well as derivatives $\frac{\partial\lambda_j(i_1,\ldots,i_M)}{\partial\theta}$. These derivatives can be straightforwardly obtained from Eq. (3) as

$$\begin{pmatrix} \frac{1}{p_1}\cdot\frac{\partial\lambda_1(i_1,\ldots,i_M)}{\partial\theta} \\ \frac{1}{p_2}\cdot\frac{\partial\lambda_2(i_1,\ldots,i_M)}{\partial\theta} \\ \vdots \\ \frac{1}{p_M}\cdot\frac{\partial\lambda_M(i_1,\ldots,i_M)}{\partial\theta} \end{pmatrix} = -\mathbf{V}^{-1}\mathbf{V}^{(\theta)}\mathbf{V}^{-1}\begin{pmatrix} q_1 i_1 \\ q_2 i_2 \\ \vdots \\ q_M i_M \end{pmatrix},$$

where $\mathbf{V}^{(\theta)}$ represents the element-by-element partial derivative of matrix $\mathbf{V}$ with respect parameter θ.[8]

3.3 Spread until the D^{th} individual detection

As outlined above, declaration of the outbreak is identified with the first detection of an infective patient, where each patient is detected at zone j at rate δ_j. If detection of an infective patient occurs because this patient shows symptoms, outbreak declaration might require several ($D > 1$) patients showing some common symptoms, since for some nosocomial pathogens associated symptoms are quite common and pass unnoticed.[5] For example, *norovirus* causes gastrointestinal symptoms such as nausea, vomiting, or diarrhoea that are common to many

For Peer Review

diseases and conditions. The *national guidelines on the management of outbreaks of norovirus infection in health-care settings*,[18] issued by the National Disease Surveillance Centre in Ireland, requires for $D = 2$ patients to show these symptoms in a hospital ward for a potential norovirus outbreak declaration. Once the outbreak has been declared, control strategies such as immediate cleaning and decontamination, frequent handwashing, or cohorting of affected patients are recommended.

Thus, our interest in this subsection is to analyse the summary statistic R when the detection of the outbreak requires for D patients to show symptoms, for some value $D \geq 1$, and results in the subsections above can be seen as the particular case $D = 1$. We define the augmented process $\mathcal{X}^{aug} = \{(I_1(t), \ldots, I_M(t), D(t)) : \ t \geq 0\}$ where the increasing variable $D(t)$ amounts to the number of detected patients up to time $t \geq 0$. We consider that the outbreak is declared once $D(t) = D$, and the space of states of this CTMC is given by

$$\mathcal{S}^{aug} \ = \ \{(i_1, i_2, \ldots, i_M, d) : \ i_j \in \{0, \ldots, N_j\}, \ j \in \{1, \ldots, M\}, \ 0 \leq d \leq D-1\} \cup \{\Delta\}.$$

Thus, $\mathcal{S}^{aug} = \mathcal{C}^{aug} \cup \Delta$, with state Δ representing outbreak declaration (i.e., the detection of the D^{th} infected patient). Events occurring in this process, at different rates, are described in Table 2.

Table 2: Transitions or events occurring in the stochastic process $\mathcal{X}^{aug}$, and corresponding rates

Event	**Stochastic transition**	**Rate**
Infection of a patient at zone j	$(i_1, \ldots, i_j, \ldots, i_M, d) \to (i_1, \ldots, i_j + 1, \ldots, i_M, d)$	$\lambda_j(i_1, \ldots, i_M)(N_j - i_j)$
Discharge of an infective patient at zone j	$(i_1, \ldots, i_j, \ldots, i_M, d) \to (i_1, \ldots, i_j - 1, \ldots, i_M, d)$	$\gamma_j i_j$
Detection of an infective patient	$(i_1, \ldots, i_M, d) \to (i_1, \ldots, i_M, d+1)$, $0 \leq d \leq D-2$	$\sum_{j=1}^{M} \delta_j i_j$
Outbreak declaration	$(i_1, \ldots, i_M, D-1) \to \Delta$	$\sum_{j=1}^{M} \delta_j i_j$

Our arguments in subsections above can be adapted for process $\mathcal{X}^{aug}$. For example, Eqs. (4)-(5) become

$$\begin{aligned} p_{(i_1,\ldots,i_M,d)}(0) \sum_{j=1}^{M} (\lambda_j(i_1, \ldots, i_M)(N_j - i_j) + (\gamma_j + \delta_j) i_j) \ &= \ \sum_{k=1}^{M} \big(\gamma_k i_k p_{(i_1,\ldots,i_k-1,\ldots,i_M,d)}(0) \\ &\quad + \delta_k i_k (1_{d=D-1} + 1_{d<D-1} p_{(i_1,\ldots,i_M,d+1)}(0)) \big), \\ p_{(i_1,\ldots,i_M,d)}(n) \sum_{j=1}^{M} (\lambda_j(i_1, \ldots, i_M)(N_j - i_j) + (\gamma_j + \delta_j) i_j) \ &= \ \sum_{k=1}^{M} \big(\gamma_k i_k p_{(i_1,\ldots,i_k-1,\ldots,i_M,d)}(n) \\ + \lambda_k(i_1, \ldots, i_M)(N_k - i_k) p_{(i_1,\ldots,i_k+1,\ldots,i_M,d)}(n-1) &+ \delta_k i_k 1_{d<D-1} p_{(i_1,\ldots,i_M,d+1)}(n) \big), \quad n \geq 1, \end{aligned}$$

for any $(i_1, \ldots, i_M, d) \in \mathcal{C}^{aug}$, and with boundary conditions $p_{(0,0,\ldots,0,d)}(0) = p_\Delta(0) = 1$, and $p_{(0,0,\ldots,0,d)}(n) = p_\Delta(n) = 0$ for all $n \geq 1$ and $0 \leq d \leq D-1$. 1_A represents a function that takes value 1 if A is satisfied, and 0 otherwise. An adapted version of Algorithm 1, not reported here for the sake of brevity, allows for an efficient solution of this system.

For Peer Review

4 Results

We consider here the hypothetical hospital ward in Figure 2 in,[20] and ventilation settings according to Table 1 in.[20] This hypothetical hospital ward consists of three six-bedded bays connected through a corridor. Each bay is split into two ventilation zones, and each ventilation zone contains three patients. The corridor is split into three ventilation zones and in the first instance has no patients, so that the ward has $N = 18$ patients. Ventilation rate over the whole ward was 27 $m^3 \cdot min^{-1}$ which equated to an air change rate of 3 $AC \cdot h^{-1}$. Diagrams of the ventilation flows and specific rates within and between each zone for each setting are given in Figure 2. Moreover, we set $p_i = 0.01\ m^3 \cdot min^{-1}$ and $q_i = 0.5\ quanta \cdot min^{-1}$ for all patients at all zones and for every ventilation scenario,[20] and assume an average length-of-stay (LOS) for each patient $\gamma_i^{-1} = 7\ days$.

If we order the $M = 9$ ventilation zones as

$$1a \prec 1b \prec c1 \prec 2a \prec 2b \prec c2 \prec 3a \prec 3b \prec c3,$$

the ventilation matrix is given by

$$\mathbf{V} = \begin{pmatrix} Q_{o,1a} + \sum_k \beta_{1a,k} & -\beta_{1b,1a} & \cdots & -\beta_{3b,1a} & -\beta_{c3,1a} \\ -\beta_{1a,1b} & Q_{o,1b} + \sum_k \beta_{1b,k} & \cdots & -\beta_{3b,1b} & -\beta_{c3,1b} \\ -\beta_{1a,c1} & -\beta_{1b,c1} & \cdots & -\beta_{3b,c1} & -\beta_{c3,c1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ -\beta_{1a,c3} & -\beta_{1b,c3} & \cdots & -\beta_{3b,c3} & Q_{o,c3} + \sum_k \beta_{c3,k} \end{pmatrix}$$

and ventilation settings in Table 1 in[20] lead to the ventilation matrices reported in Figure 2.

In subsections 4.1-4.3 and 4.6, we consider that outbreak declaration occurs after one patient shows symptoms, with each patient showing symptoms after an average time δ^{-1} (i.e., $\delta_j = \delta$ for all $1 \leq j \leq M$). Alternative outbreak detection and declaration hypotheses are discussed in subsections 4.4-4.5, while the impact of parameter q in our numerical results is explored in subsection 4.7.

4.1 Impact of ventilation setting on spread dynamics

In Figure 3 we plot the probability mass function of R versus different values of the global mixing rate β_0, the average time δ^{-1} at which each infective patient shows symptoms, and for ventilation scenarios $A - F$. For these results, it is assumed that an infective patient at zone $1a$ starts the outbreak, and we report in Table 3 the mean values $E[R]$ computed for these distributions.

Table 3: Mean number $E[R]$ of infections until end or declaration of the outbreak, for scenarios in Figure 3

β_0	δ^{-1}	A	B	C	D	E	F
9	$12h$	1.75	1.54	1.24	1.75	1.74	1.50
	$48h$	5.01	4.34	3.87	5.20	4.86	4.50
27	$12h$	1.88	1.76	1.61	1.86	1.87	1.77
	$48h$	5.59	5.25	5.03	5.59	5.53	5.37

For Peer Review

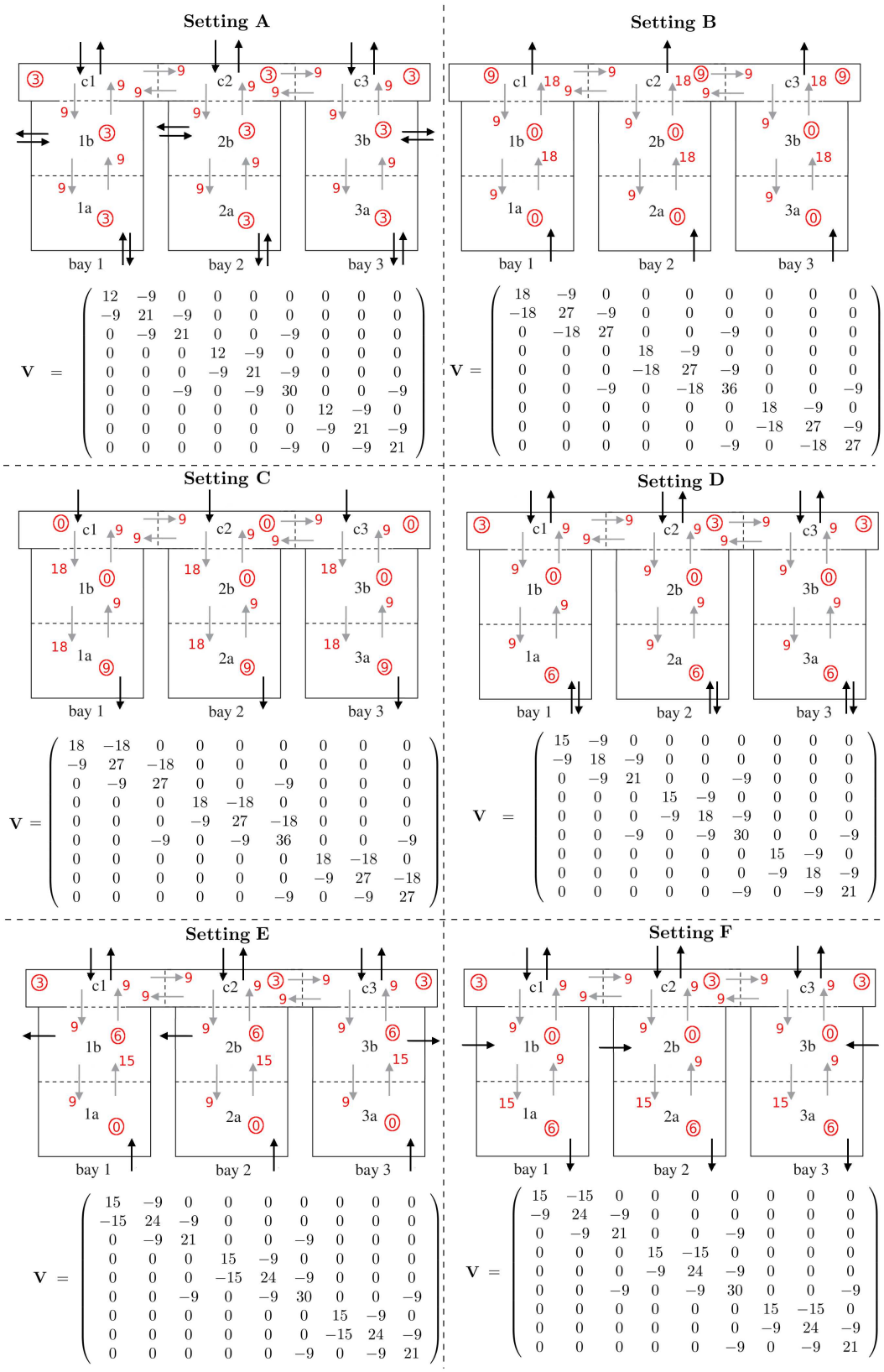

Figure 2: Diagrams of the ventilation settings considered, and corresponding ventilation matrices $\mathbf{V}$ (for $\beta_0 = 9\ m^3 \cdot min^{-1}$). Grey arrows: ventilation between zones (β_{ij} -$m^3 \cdot min^{-1}$-); Black arrows: ventilation supply and extract to the ward; Circled values: extract ventilation rates ($Q_{o,j}$ -$m^3 \cdot min^{-1}$-)

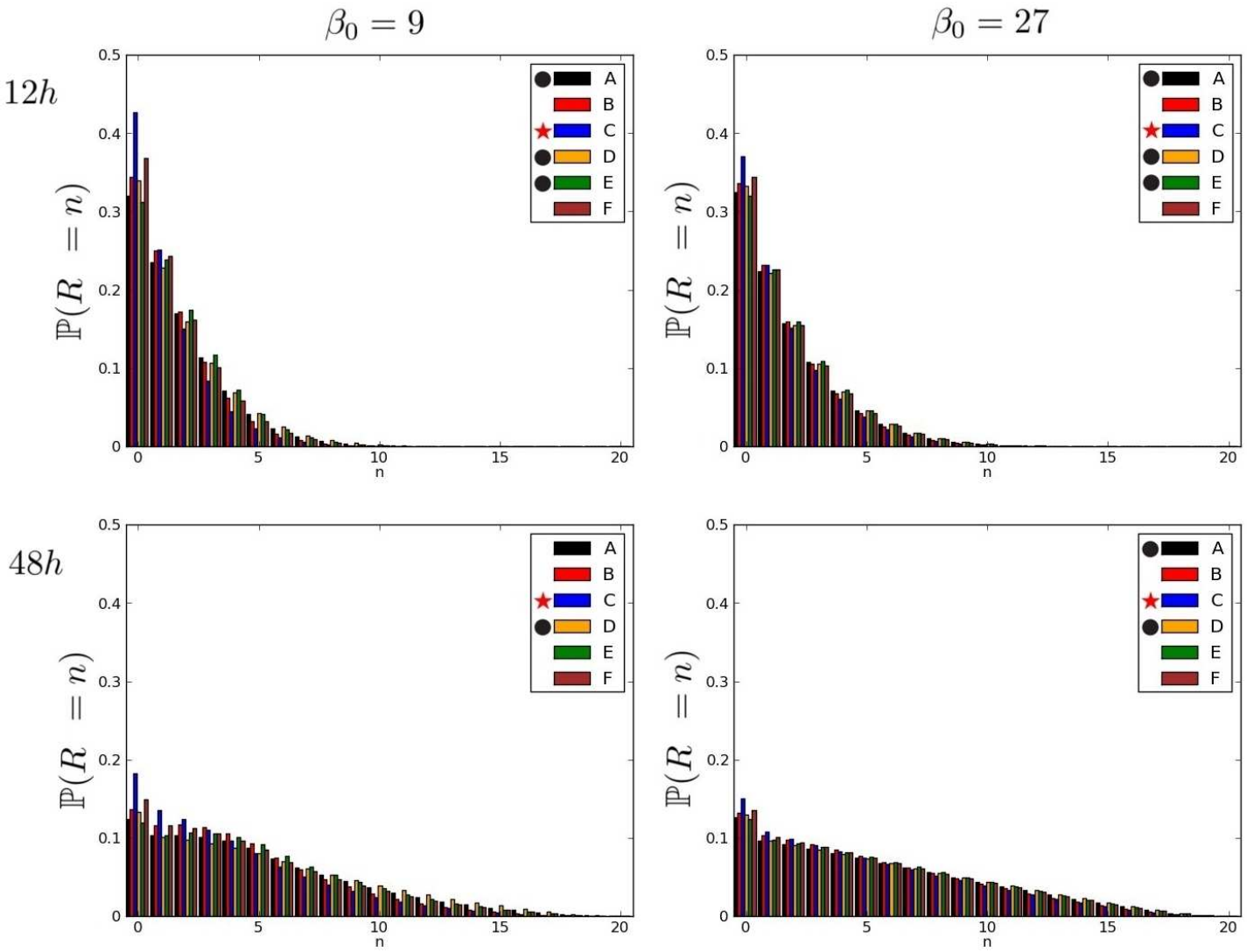

Figure 3: Probability mass function of the number R of infections until the end or detection of the outbreak, for ventilation settings $A - F$. Declaration of the outbreak occurs once one patient shows symptoms, and each infective patient shows symptoms after an average time $\delta^{-1} \in \{12h, 48h\}$. Global mixing rate $\beta_0 \in \{9, 27\}$ $m^3 \cdot min^{-1}$. Initial infective located at zone $1a$. Red stars identify the best ventilation settings in terms of $E[R]$, while black circles identify the worst ones; see Table 3

Ventilation setting C can be identified in Figure 3 and Table 3 as the best one, while settings A, D and E are identified as the worst ones depending on the detection parameter δ and the global mixing rate β_0. We note that ventilation setting C has significant extract ventilation at the initially *infected zone* $1a$, so that the airflow is directed from $1a$ outwards the hospital ward. On the other hand, ventilation setting D represents a well-mixed ward ($\beta_{Qik} = 0$ for all i and k) and no extract ventilation at zone $1b$ ($Q_{o,1b} = 0$), which might favor the spread of pathogens from $1a$ towards other zones within the ward, leading to more infections occurring until outbreak detection.

In general terms, worse scenarios can be identified for values $\delta^{-1} = 48h$ and $\beta_0 = 27m^3 \cdot min^{-1}$, where the long-tailed distribution of R for these scenarios in Figure 3 indicates that large outbreaks occur with significant probability. Larger differences among ventilation settings are also found for value $\delta^{-1} = 48h$. Thus, our results suggest that ventilation of the ward should

be of special concern for pathogens that have longer infectious asymptomatic periods, or in hospital wards with more limited surveillance policies. It is also clear that the average individual detection time δ^{-1} has a higher impact on the infection spread than the specific ventilation setting in the ward, so that outbreak detection seems to *dominate* ventilation regarding infection spread.

4.2 Dependence on location of initial infective

In Figure 4 we plot analogous results to those in Figure 3 when the infective patient starting the outbreak is located at zones $\{1a, 1b, 2a, 2b\}$, for $\delta^{-1} = 12h$ and $\beta_0 = 9\ m^3 \cdot min^{-1}$. Corresponding mean values $E[R]$ are reported in Table 4. We note that zones $3a$ and $3b$ are *equivalent* to zones $1a$ and $1b$, for all ventilation settings in Figure 2, and thus we do not test them. For zones near the corridor (i.e., $1b$ and $2b$), ventilation setting B is identified as the best one, while D is identified as the worst one. We note that ventilation setting B has no extract ventilation at zones $1b$ and $2b$ ($Q_{o,1b} = Q_{o,2b} = 0$), but it directs the airflow instead towards corridor areas. In this setting B, corridor areas have no patients and significant ventilation, with airflow unbalance *from* bays *to* corridor areas acting in practice as an infection control measure. Thus, our results suggest that the spread control ability of a given ventilation setting depends on the location of the patient starting the outbreak as well as the airflow direction. However, from results in Figure 4 and Table 4, ventilation setting D seems to perform poorly regardless of the initial infective location, suggesting that some ventilation settings might be inadvisable regardless of this location (i.e., if this location is unknown).

Table 4: Mean values $E[R]$ for distributions in Figure 4; that is, for different locations of the initial infective

Location	A	B	C	D	E	F
$1a$	1.75	1.54	1.24	1.75	1.75	1.50
$1b$	1.74	1.28	1.57	1.93	1.50	1.73
$2a$	1.79	1.57	1.27	1.79	1.78	1.54
$2b$	1.78	1.32	1.62	1.98	1.54	1.78

When focusing on a particular location for the initial infective, comments above are supported by the sensitivity analysis on the ventilation parameters. For example, in Tables 5-6 we report, for ventilation parameters $\theta \in \{\beta_0, Q_{o,1a}, \ldots, Q_{o,c3}, \beta_{Q1a,1b}, \ldots, \beta_{Qc3,3b}\}$, partial derivatives $\partial E[R]/\partial\theta$ and elasticities $(\partial E[R]/\partial\theta) \cdot (\theta/E[R])$ for ventilation settings B and D, $\delta^{-1} = 12h$, $\beta_0 = 9\ m^3 \cdot min^{-1}$ and an infective patient starting the outbreak at zone $1b$. We note that while dimensionless elasticities are useful for comparison purposes, they equal zero if parameter θ is zero.

Regime B requires airflow towards the corridor in order to expel pathogens from zone $1b$, since $Q_{o,1b} = 0$. Thus, rates $\beta_{Q1b,c1}$, $Q_{o,c1}$, $Q_{o,c2}$ and $Q_{o,c3}$ correspond to significantly large negative elasticities reported in Table 5 (i.e., increasing the values of these rates would lead to decreasing values of $E[R]$). Global mixing rate β_0 has a significant impact (large positive elasticity) favoring disease spread, since increasing the value of β_0 represents increasing the rate at which pathogens flow among all zones, instead of flowing specifically towards the extract ventilation areas (corridors in this setting).

According to results in Table 6, ventilation setting D could be significantly improved by increasing extract ventilation (especially at zones 1a, $1b$ and $c1$), as well as increasing airflow from $1b$ to $1a$ and to $c1$. This is directly related to the fact that, since there is no extract

For Peer Review

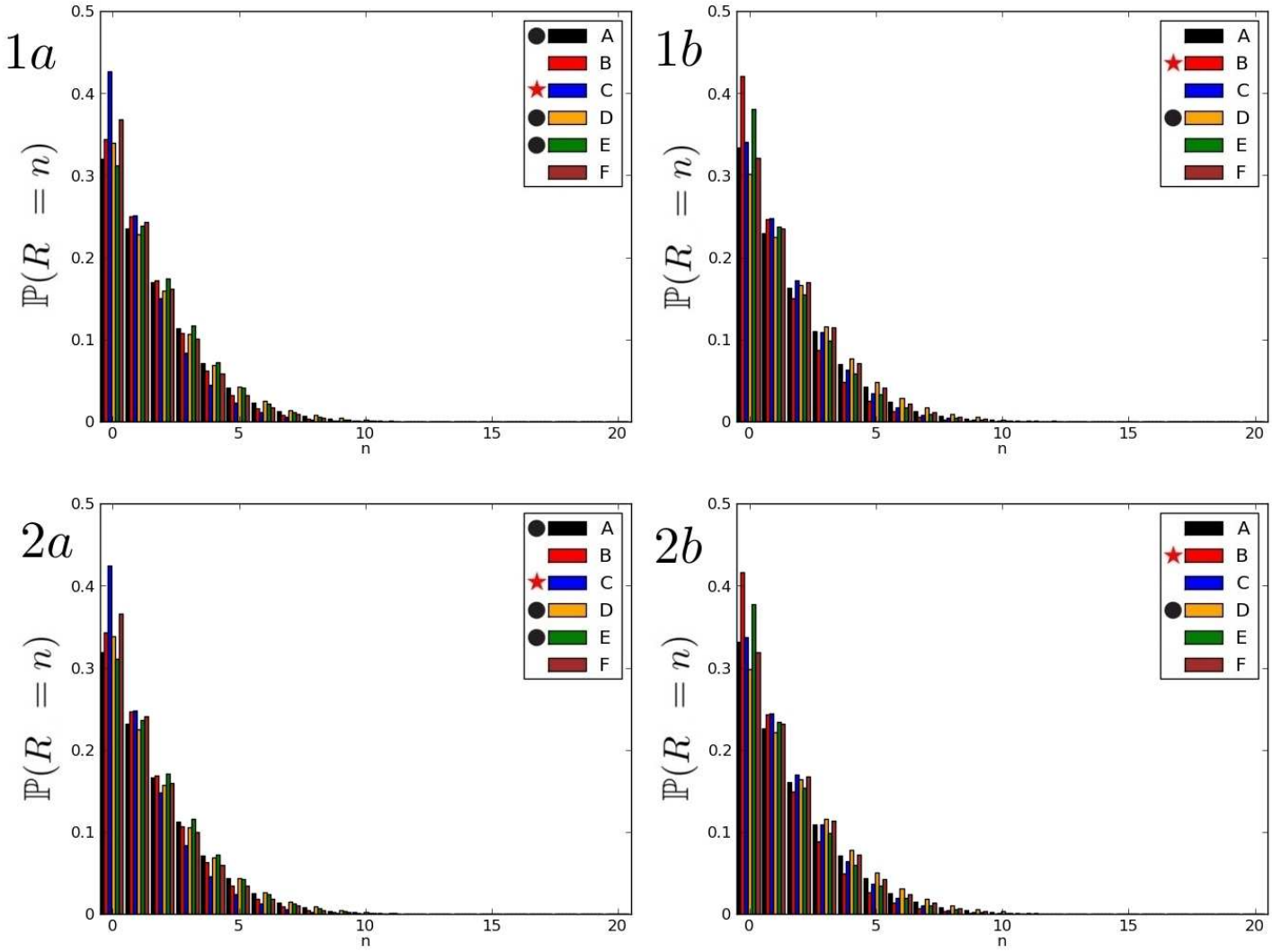

Figure 4: Probability mass function of the number R of infections until the end or detection of the outbreak, for ventilation settings $A - F$. Declaration of the outbreak occurs once one patient shows symptoms, and each infective patient shows symptoms after an average time $\delta^{-1} = 12h$. Initial infective located at zones $\{1a, 1b, 2a, 2b\}$. Global mixing rate $\beta_0 = 9\ m^3 \cdot min^{-1}$

ventilation at zone $1b$, infectious material in this zone can only be expelled by directing it towards adjacent zones $1a$ and $c1$.

4.3 Decreasing hospital ward infection spread risk might increase risk at specific bays

It is clear that the number, R, of infections occurring until the end or detection of the outbreak can be split according to where these infections actually occur as

$$R \quad = \quad R(1) + R(2) + R(3);$$

where $R(j)$ is the number of infections occurring at bay j. Although probabilities $\mathbb{P}(R(j) = n)$ can be analytically computed by adapting arguments in Section 3, details are omitted here for

Table 5: Partial derivatives and elasticities of $E[R]$ with respect to ventilation parameters, for setting B. Average individual detection time $\delta^{-1} = 12h$, global mixing rate $\beta_0 = 9\ m^3 \cdot min^{-1}$, and an initially infective patient starting the outbreak at zone $1b$

Partial Derivatives						Elasticities					
β_0	0.0230	$\beta_{Q2b,c2}$	−0.0141	$Q_{o,1b}$	−0.0890	β_0	0.2899	$\beta_{Q2b,c2}$	−0.1042	$Q_{o,1b}$	0
$\beta_{Q1a,1b}$	−0.0201	$\beta_{Qc2,2b}$	0.0243	$Q_{o,c1}$	−0.0398	$\beta_{Q1a,1b}$	−0.1483	$\beta_{Qc2,2b}$	0	$Q_{o,c1}$	−0.2876
$\beta_{Q1b,1a}$	0.0207	$\beta_{Qc2,c3}$	0.0019	$Q_{o,2a}$	−0.0209	$\beta_{Q1b,1a}$	0	$\beta_{Qc2,c3}$	0	$Q_{o,2a}$	0
$\beta_{Q1b,c1}$	−0.0382	$\beta_{Qc3,c2}$	−0.0008	$Q_{o,2b}$	−0.0274	$\beta_{Q1b,c1}$	−0.2735	$\beta_{Qc3,c2}$	0	$Q_{o,2b}$	0
$\beta_{Qc1,1b}$	0.0301	$\beta_{Q3a,3b}$	−0.0032	$Q_{o,c2}$	−0.0221	$\beta_{Qc1,1b}$	0	$\beta_{Q3a,3b}$	−0.0240	$Q_{o,c2}$	−0.1615
$\beta_{Qc1,c2}$	0.0097	$\beta_{Q3b,3a}$	0.0060	$Q_{o,3a}$	−0.0120	$\beta_{Qc1,c2}$	0	$\beta_{Q3b,3a}$	0	$Q_{o,3a}$	0
$\beta_{Qc2,c1}$	−0.0041	$\beta_{Q3b,c3}$	−0.0081	$Q_{o,3b}$	−0.0159	$\beta_{Qc2,c1}$	0	$\beta_{Q3b,c3}$	−0.0601	$Q_{o,3b}$	0
$\beta_{Q2a,2b}$	−0.0056	$\beta_{Qc3,3b}$	0.0141	$Q_{o,c3}$	−0.0132	$\beta_{Q2a,2b}$	−0.0420	$\beta_{Qc3,3b}$	0	$Q_{o,c3}$	−0.0971
$\beta_{Q2b,2a}$	0.0106	$Q_{o,1a}$	−0.1089			$\beta_{Q2b,2a}$	0	$Q_{o,1a}$	0		

Table 6: Partial derivatives and elasticities of $E[R]$ with respect to ventilation parameters, for setting D. Average individual detection time $\delta^{-1} = 12h$, global mixing rate $\beta_0 = 9\ m^3 \cdot min^{-1}$, and an initially infective patient starting the outbreak at zone $1b$

Partial Derivatives						Elasticities					
β_0	−0.0032	$\beta_{Q2b,c2}$	−0.0198	$Q_{o,1b}$	−0.1746	β_0	−0.0151	$\beta_{Q2b,c2}$	0	$Q_{o,1b}$	0
$\beta_{Q1a,1b}$	0.0184	$\beta_{Qc2,2b}$	0.0241	$Q_{o,c1}$	−0.0872	$\beta_{Q1a,1b}$	0	$\beta_{Qc2,2b}$	0	$Q_{o,c1}$	−0.1356
$\beta_{Q1b,1a}$	−0.0248	$\beta_{Qc2,c3}$	0.0046	$Q_{o,2a}$	−0.0411	$\beta_{Q1b,1a}$	0	$\beta_{Qc2,c3}$	0	$Q_{o,2a}$	−0.1278
$\beta_{Q1b,c1}$	−0.0212	$\beta_{Qc3,c2}$	−0.0026	$Q_{o,2b}$	−0.0718	$\beta_{Q1b,c1}$	0	$\beta_{Qc3,c2}$	0	$Q_{o,2b}$	0
$\beta_{Qc1,1b}$	0.0124	$\beta_{Q3a,3b}$	0.0053	$Q_{o,c2}$	−0.0590	$\beta_{Qc1,1b}$	0	$\beta_{Q3a,3b}$	0	$Q_{o,c2}$	−0.0918
$\beta_{Qc1,c2}$	0.0167	$\beta_{Q3b,3a}$	−0.0074	$Q_{o,3a}$	−0.0293	$\beta_{Qc1,c2}$	0	$\beta_{Q3b,3a}$	0	$Q_{o,3a}$	−0.0912
$\beta_{Qc2,c1}$	−0.0090	$\beta_{Q3b,c3}$	−0.0140	$Q_{o,3b}$	−0.0516	$\beta_{Qc2,c1}$	0	$\beta_{Q3b,c3}$	0	$Q_{o,3b}$	0
$\beta_{Q2a,2b}$	0.0074	$\beta_{Qc3,3b}$	0.0171	$Q_{o,c3}$	0.0434	$\beta_{Q2a,2b}$	0	$\beta_{Qc3,3b}$	0	$Q_{o,c3}$	−0.0675
$\beta_{Q2b,2a}$	−0.0102	$Q_{o,1a}$	−0.1017			$\beta_{Q2b,2a}$	0	$Q_{o,1a}$	−0.3165		

the sake of brevity, and results in Table 7 are obtained from 10^6 stochastic simulations of the process.

In Table 7, we report values of $E[R] = E[R(1)] + E[R(2)] + E[R(3)]$ for $\beta_0 = 9\ m^3 \cdot min^{-1}$, $\delta^{-1} = 48h$, ventilation settings A, D and E, and an infective patient starting the outbreak at zone $1a$. Results suggest that epidemic spread can be limited by *switching* ward ventilation from setting D to A, and further containment is obtained by switching to ventilation setting E. However, infection risk in bay 1 (in terms of $E[R(1)]$) behaves contrarily; although the global hospital ward infection risk (in terms of $E[R]$) is lower for setting E, this is at the expense of expelling pathogens from the infected zone $1a$ towards zones $1b$ and $c1$, and thus posing a greater risk to patients in bay 1.

Table 7: Mean values of $E[R]$, $E[R(1)]$, $E[R(2)]$ and $E[R(3)]$ for $\beta_0 = 9\ m^3 \cdot min^{-1}$, $\delta^{-1} = 48h$ and ventilation settings A, D and E. Initially infective individual at zone $1a$

Regime	$E[R]$	$E[R(1)]$	$E[R(2)]$	$E[R(3)]$
A	5.01	2.65	1.34	1.02
D	5.20	2.56	1.47	1.17
E	4.86	2.69	1.25	0.92

For Peer Review

4.4 Interplay with detection management

As explained in subsection 3.3, detection and declaration of an outbreak in the hospital ward may require several patients showing symptoms, and not only one. In Figure 5 we plot analogous results to those in Figure 3 when declaration of the outbreak occurs after $D = 2$ patients show symptoms (each after average time δ^{-1}). The corresponding mean values of $E[R]$ are reported in Table 8.

We note that values in Table 8 are significantly larger than those in Table 3, since outbreak declaration takes longer to occur allowing for more infections to take place. This increase is significantly larger than the differences that can be observed, in Table 3, between different ventilation settings, suggesting again that detection policy is likely to dominate ventilation as an infection control strategy. Under slow detection scenarios ($\delta^{-1} = 48h$), we observe in Figure 5 a clear bi-modality for the distribution of R. Thus, our model predicts that under slow detection, a two-output situation can be expected: either the initially infective patient is discharged before infecting any other patient (so that $R = 0$), or this patient infects a second patient, leading to a large outbreak (represented by the second mode in Figure 5).

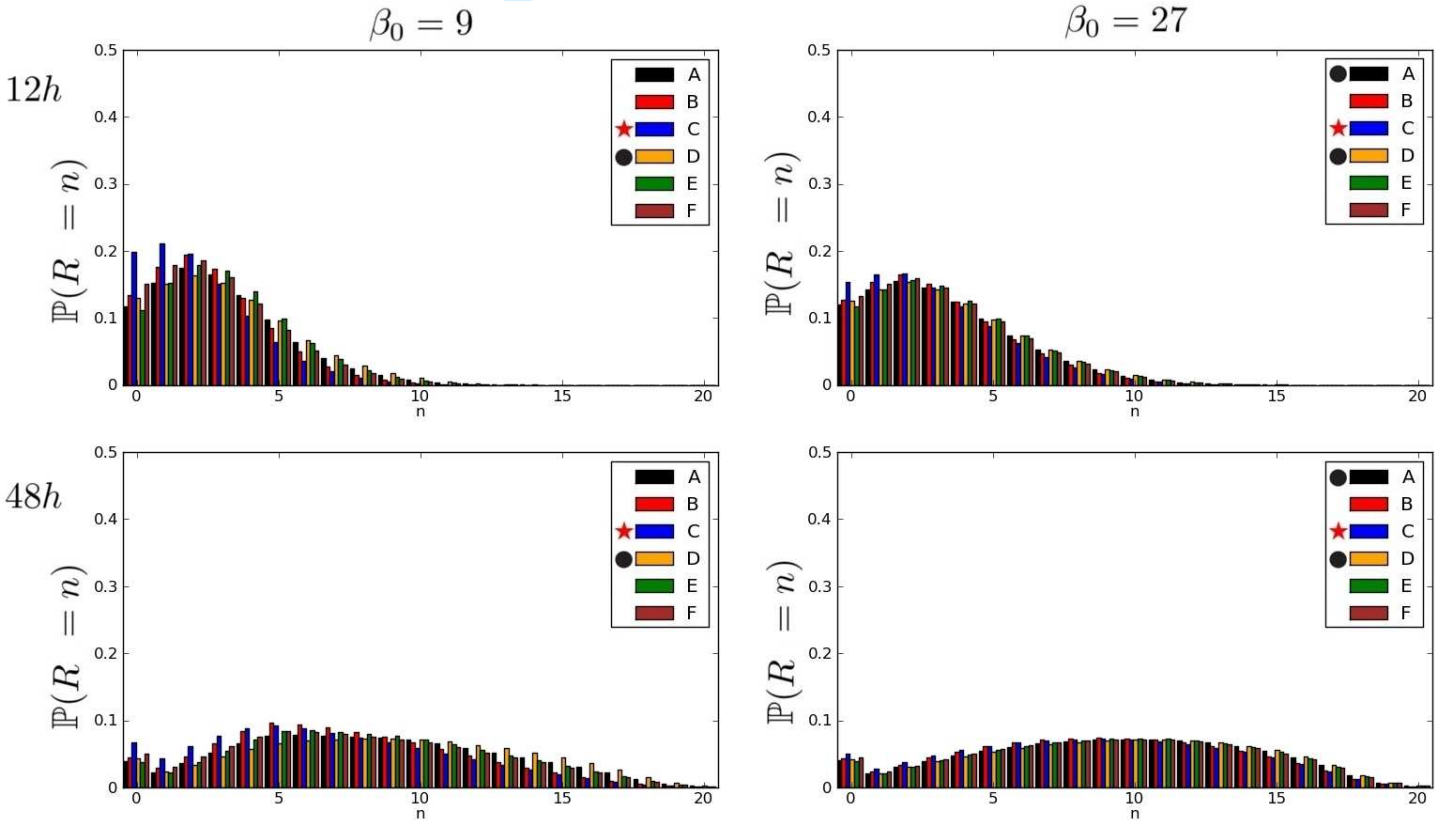

Figure 5: Probability mass function of the number R of infections until the end or detection of the outbreak, for ventilation settings $A - F$. Declaration of the outbreak occurs once two patients show symptoms, and each infective patient shows symptoms after an average time $\delta^{-1} \in \{12h, 48h\}$. Global mixing rate $\beta_0 \in \{9, 27\}$ $m^3 \cdot min^{-1}$. Initial infective located at zone $1a$

Table 8: Mean values of $E[R]$ for scenarios in Figure 5.

β_0	δ^{-1}	A	B	C	D	E	F
9	$12h$	3.20	2.78	2.33	3.26	3.14	2.79
	$48h$	8.42	7.31	6.78	8.79	8.10	7.72
27	$12h$	3.52	3.28	3.07	3.50	3.49	3.34
	$48h$	9.37	8.87	8.63	9.39	9.26	9.10

4.5 Screening at admission

In subsections 4.1-4.4, we analyse infection spread under the assumption that each individual at zone j is detected (by showing symptoms) at rate δ_j, with $\delta_j = \delta$ for all j, and where the outbreak is detected and declared after one (or several) infective patients are detected. This leads to the contribution $\sum_{j=1}^{M} \delta_j i_j$ in Eqs. (4)-(5). However, if the detection of the outbreak is due instead to the screening of the newly admitted patient who starts the outbreak, and results of this screening arrive after an average time δ^{-1}, then the outbreak is detected at rate δ, and one needs to replace $\sum_{j=1}^{M} \delta_j i_j$ by δ in Eqs. (4)-(5).

Under this hypothesis, we plot in Figure 6 the probability mass function of the number R of infections until the end or detection of the outbreak, when the results of this screening (and thus, the declaration of the outbreak) arrive after an average time $\delta^{-1} \in \{4h, 8h, 12h, 24h\}$. Corresponding mean values of $E[R]$ are reported in Table 9. If results arrive after $\delta^{-1} = 4$ hours, ventilation has a less significant impact on the nosocomial spread, and low values of $E[R]$ are reported in Table 9. The number of infections until the end or detection of the outbreak proportionally increases with the delay δ^{-1} in obtaining the screening results. In particular, for $\delta^{-1} = 24$ hours significant differences in $E[R]$ can be noticed among the different ventilation settings, and a marked bi-modality can be observed in Figure 6.

Table 9: Mean values $E[R]$ for scenarios in Figure 6

δ^{-1}	A	B	C	D	E	F
$4h$	1.12	0.92	0.66	1.11	1.11	0.89
$8h$	2.66	2.13	1.63	2.72	2.59	2.16
$12h$	4.24	3.43	2.74	4.36	4.11	3.54
$24h$	8.15	6.91	5.94	8.36	7.91	7.17

4.6 Ventilation and over-occupancy

In this subsection our aim is to shed some light on the interplay between ventilation, nosocomial spread and over-occupancy of the hospital ward. We represent hospital ward over-occupancy by locating three additional patients at the corridor areas; in particular, we set $N = 21$ and locate one additional patient at each of the $\{c1, c2, c3\}$ zones. This practice is common in UK hospitals during times of high demand. We assume that the outbreak is detected and declared after the first patient shows symptoms, each patient showing symptoms after an average time $\delta^{-1} = 12h$. For interzonal mixing $\beta_0 = 9\ m^3 \cdot min^{-1}$, we report in Table 10 the mean number R of infections until the end or detection of the outbreak, for ventilation settings $A - F$, and for the initial infective patient being located at zones $\{1a, c1, c2, c3\}$.

We first note that results in Table 10 are exactly the same for an initially infective individual being located at zones $c1$ and $c3$. This is explained by noting that bays 1 and 3 are completely

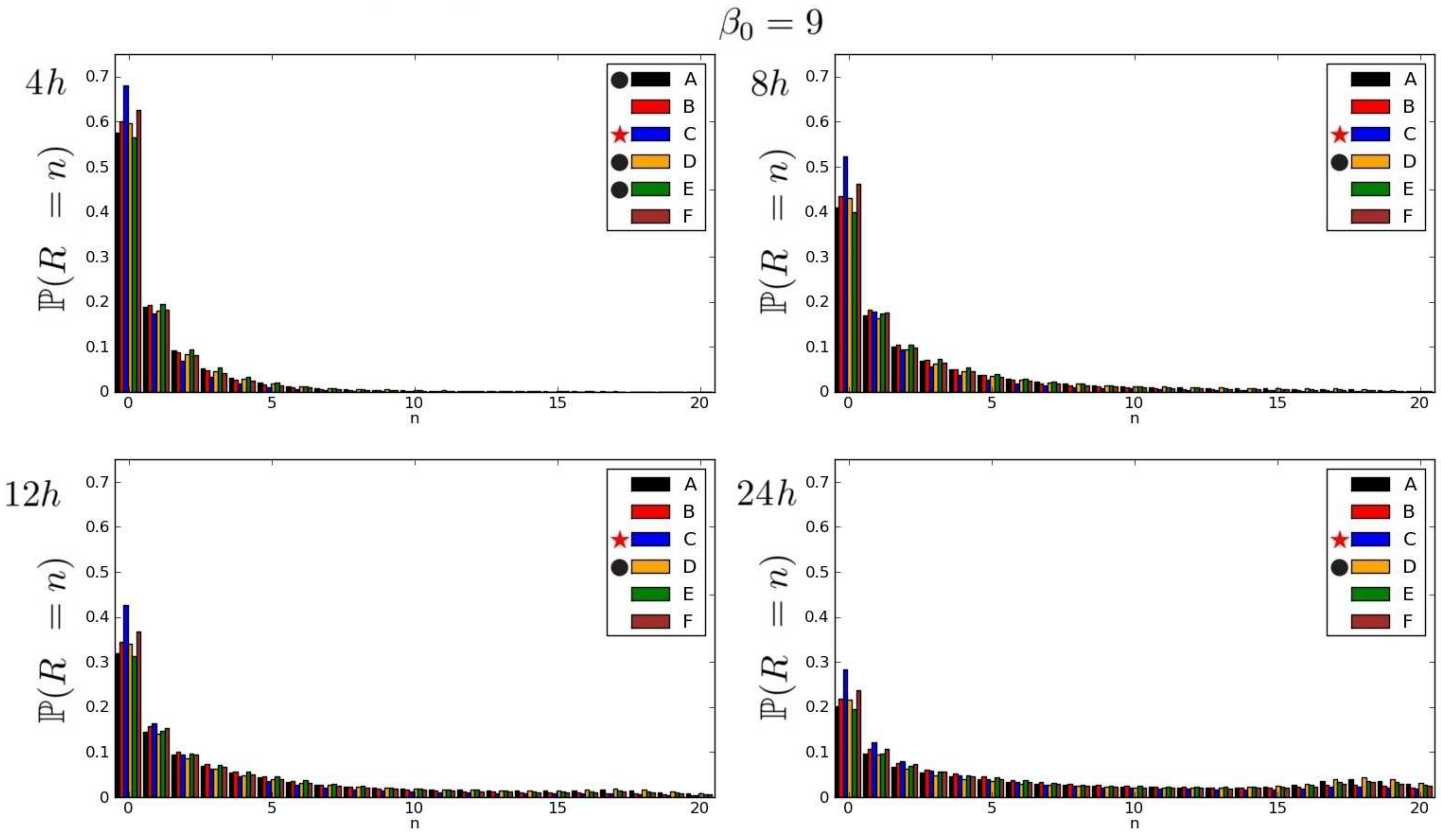

Figure 6: Probability mass function of the number R of infections until the end or detection of the outbreak, for ventilation settings $A - F$. Declaration of the outbreak occurs after an average time $\delta^{-1} \in \{4h, 8h, 12h, 24h\}$, which is independent of the number of infectives and represents a screening event. Global mixing rate $\beta_0 = 9\ m^3 \cdot min^{-1}$. Initial infective located at zone $1a$

Table 10: Mean values $E[R]$ for a hospital ward with over-occupancy. Patients show symptoms after an average time $\delta^{-1} = 12h$, and we set $\beta_0 = 9\ m^3 \cdot min^{-1}$

Location of initial infective	A	B	C	D	E	F
$1a$	2.01	1.82	1.37	2.04	2.02	1.72
$c1$	1.95	1.14	1.95	2.19	1.71	1.92
$c2$	1.99	1.18	2.00	2.23	1.75	1.96
$c3$	1.95	1.14	1.95	2.19	1.71	1.92

symmetric for all ventilation settings as noted in subsection 4.2; see diagrams and matrices in Figure 2. On the other hand, when the initial infective patient is located at zone $1a$, results in Table 10 can be compared to those in Table 3 for $(\delta^{-1}, \beta_0) = (12h, 9\ m^3 \cdot min^{-1})$. For zone $1a$, values of $E[R]$ are larger in Table 10 than in Table 3; if the nosocomial outbreak is initiated by an infective patient at zone $1a$, more infections during this outbreak should be expected under ward over-occupancy. This might not be only related to having more patients in the ward under over-occupancy (21 instead of 18), but also to the potential of patients in the corridor to act as *infection links* between bays. For example under over-occupancy, an infective individual at zone $1a$ might infect individuals in bay 2 by, as a first step, infecting individuals in the corridor areas. These people then might more easily infect individuals at bay 2, before being discharged, due to being in closer proximity and depending on the particular ventilation

setting in place in the ward.

Infection dynamics related to the scenario above highly depend on the particular ventilation setting under study, which can be noticed by inspecting rows corresponding to zones $\{c1, c2, c3\}$ in Table 10. While an individual at zone $1a$ has a larger infectious potential (in terms of $E[R]$) than individuals located at the corridor when ventilation settings B and E are in place, this is not the case for ventilation settings C, D and F, and these infectious potentials are comparable under ventilation setting A which represents a well-mixed scenario. Our results then indicate that over-occupancy leads in general to higher airborne spread risks, and that this increase can be especially significant depending on the specific ventilation in place.

4.7 The unit of infection

We note that parameter q is highly pathogen-dependent, ranging from $q \sim 0.01\ quanta \cdot min^{-1}$ for rhinovirus to $q \sim 10\ quanta \cdot min^{-1}$ for measles.[20] We perform a parametric analysis by varying q to assess the sensitivity of our conclusions, and report expected infections $E[R]$ until the end of the outbreak in Table 11. In particular, we are interested in the mean number $E[R]$ of infections if the outbreak is detected upon the first patient showing symptoms, each patient showing symptoms after an average time $\delta^{-1} = 12h$, and where we consider $\beta_0 = 9\ m^3 \cdot min^{-1}$. We note that, as expected, increasing values of q lead to increasing mean number $E[R]$ of infections. However, this does not seem to affect the relative infectiousness of ventilation setting C, which is identified as the best scenario regardless of the value of q. On the other hand, less advantageous ventilation schemes are dependent on value of q. For example, ventilation setting A and E can be identified as the worst for $q = 0.1\ quanta \cdot min^{-1}$, while setting D can be seen as the worst one for $q = 50.0\ quanta \cdot min^{-1}$.

Table 11: Mean number $E[R]$ of infections until end or declaration of the outbreak, for $\beta_0 = 9\ m^3 \cdot min^{-1}$ and when outbreak declaration occurs after the first patient shows symptoms. Each patient shows symptoms after an average time $\delta^{-1} = 12h$. Different values of q ($quanta \cdot min^{-1}$) considered

q	A	B	C	D	E	F
0.1	0.41	0.36	0.27	0.38	0.42	0.34
0.5	1.75	1.54	1.24	1.75	1.74	1.50
1.0	3.06	2.66	2.26	3.14	3.00	2.69
10.0	11.43	10.44	10.16	11.75	11.13	10.92
50.0	15.54	15.12	15.05	15.66	15.41	15.35

5 Discussion and Conclusions

In this work, we link a zonal ventilation model for the generation and airborne spread of infectious material within a hospital ward, with a multi-compartment SIS Markovian model for the infection of patients within this ward. Our model incorporates the possibility of considering a wide range of ventilation settings, the discharge and arrival of patients within the ward, as well as different hypotheses regarding how outbreak detection and declaration occurs. Moreover, it allows us to explore the interplay between ventilation, outbreak management, ward over-occupancy and the location of the infective patient starting the outbreak.

Our results suggest that detection time dominates ventilation when the variable of interest is the number of infections occurring before the declaration or end of the outbreak, with longer

For Peer Review

detection times leading to significantly more infections happening. Longer detection times can arise when analysing pathogens with long infectious asymptomatic periods, when declaration of an outbreak requires for several patients to show symptoms, or when this declaration depends on screening events for which results take longer to arrive. The interplay between ventilation of the hospital ward and location of the initially infective patient starting the outbreak implies that recommendations on where to locate potentially infected (e.g. newly admitted) patients in a given hospital ward could be issued depending on the ventilation in place in the ward. Our model also predicts that decreasing the infection spread risk in the hospital ward can sometimes go at the expense of increasing the risk in particular areas of the ward.

Similar models have already been considered in the literature for linking zonal ventilation scenarios with epidemic spread models,[3,13,20] where epidemic dynamics are usually analysed by means of stochastic simulations. To the best of our knowledge, this is the first time that this link is carried out by defining in detail the continuous-time Markov chain for the infection spread, where infection rates at each ventilation zone are in fact functions $\lambda_j(i_1,\ldots,i_M)$ of the number of infectives at each zone at any given time, and where $(i_1,\ldots,i_M)$ represents in fact a state of the CTMC under study. This detailed mathematical construction allows for the analytical computation of summary statistics (such as R in this work), and for carrying out a local sensitivity analysis that allows one to identify the particular factors of each ventilation setting having the most significant impact on the infection spread.

It should be noted that the primary objective in this study is to demonstrate this detailed mathematical analysis and how it can be applied to evaluate the relative influence of different parameters. The model is applied to a hypothetical hospital ward, which, while it is representative of multi-bed ward environments in many hospitals, is a very simplified model of reality. The results demonstrate that the ventilation flow settings may influence the dispersion of airborne pathogens and hence the risk of transmission, however these should be interpreted with caution. We assume a steady state ventilation scenario with the flow pattern replicated exactly between neighboring bays 1, 2 and 3. In reality, the flows will not be exactly identical for every bay, and other factors such as heat sources and movement of people will alter the mixing with and between zones. In particular, corridor ventilation often has a directional flow due to wider spacing of ventilation supply/extract grilles which may hinder or improve the control of infection. However, the analysis we have carried out gives some clear insight into why particular directional flows influence risk, and the relative importance of detection strategies, ventilation control and occupancy.

It is clear that some of our conclusions could be highly dependent on the hospital ward structure, and therefore the flexibility of our methodology comes into play. It can be applied to any hospital ward of interest by appropriately adapting the corresponding ventilation matrix $\mathbf{V}$. Although carrying out a detailed mathematical analysis of a number of potentially different hospital ward structures is out of the scope of this paper, we include a short numerical study of an alternative hospital ward in the Supplementary Material. The aim of this is two-fold: (i) to show how our methodology can be easily implemented for a different hospital ward to that in Figure 2 by just adapting the ventilation matrix $\mathbf{V}$, illustrating how this matrix varies with hospital ward structure; and (ii) to show that while some of our conclusions might be hospital ward structure dependent, others seem to be valid for a wide range of hospital ward structures (e.g. detection dominates ventilation as well in this alternative hospital ward).

In this paper, we go beyond the SI epidemic model in,[20] proposing an SIS-type model which allows us to incorporate patients' arrival/discharge and outbreak detection and declaration. We note that this epidemic model structure would be especially relevant for pathogens with no or

For Peer Review

short (i.e. negligible compared to the average patient's length of stay) non-infectious or latent periods, and where the infectious period is long enough so that recovery of patients does not occur before discharge (or detection). Depending on the hospital ward under analysis and the average length of stay of patients in this ward, this could be the case for influenza or norovirus. Pathogens with non-negligible incubation periods (e.g. 7-21 days for measles) might require more complex stochastic epidemic model structures such as the SEIR (*susceptible-exposed-infective-recovered*). On the other hand, when analysing hospital wards with longer average patient's length of stay, so that individuals may become infected and recover during their stay, SIRS-type epidemic models would be required, to represent the recovery of patients ($I \to R$) before discharge ($R \to S$) occurs. We note here that in principle, the methodology outlined in Section 3 can be extended to any of these compartmental-based epidemic models for the disease spread dynamics, where the link between the deterministic zonal ventilation model for the airflow dynamics and the stochastic epidemic model for disease spread dynamics would still be as in Section 2. In a similar way, more complex epidemic model structures could allow one to study the infection spread dynamics after outbreak detection and declaration occurs. In this paper, we have focused instead on the impact of ventilation on disease spread until the end or declaration of the outbreak. Considering these alternative compartmental-based epidemic model structures could be the aim of future work.

Finally, we note that when carrying out our analysis, the main computational effort lies in solving systems of linear equations, where the number of equations is determined by the number of states of the corresponding CTMC, given by Eq. (6) in our model. Limitations of our approach are then of computational nature, since highly complex epidemic models (here, a multi-compartment SIS stochastic model with detection) linked to large hospital wards split in many different ventilation zones (here, $M = 9$ zones with three *empty* zones and six zones containing three patients each) would lead to an intractable number of equations, and stochastic simulation approaches would prevail.

Data, Software and Reproducibility

Computer codes (in *Python*) in order to reproduce our numerical results are available at an on-line repository.[31]

Acknowledgment

This work is supported by... (to be included after acceptance, to fulfil conditions of blind review process)

References

[1] Atkinson J, Chartier Y, Pessoa-Silva CL, Jensen P, Li Y (2009) *Natural Ventilation for Infection Control in Health-Care Settings*. WHO Publication/Guidelines. ISBN: 978-92-4-154785-7.

[2] Brienen NCJ, Timen A, Wallinga J, Van Steenbergen JE, Teunis PFM (2010) *The effect of mask use on the spread of influenza during a pandemic*. Risk Analysis, 30, 1210-8.

For Peer Review

[3] Carruthers J, Lopez-García M, Gillard JJ, Laws TR, Lythe G, Molina-París C (2018) *A novel stochastic multi-scale model of Francisella tularensis infection to predict risk of infection in a laboratory.* Frontiers in Microbiology, 9, 1165.

[4] Department of Health (2007) *Health Technical Memorandum HTM 03-01: Specialised ventilation for healthcare premises, Part A: Design and Validation.* The Stationary Office, London.

[5] Ekkert J (2015) *Gastrointestinal (GI) infection outbreak guidelines for health care facilities.* Interior Health.

[6] Escombe AR, Huaroto L, Ticona E, Burgos M, Sanchez I, Carrasco L, Farfan E, Flores F, Moore DAJ (2010) *Tuberculosis transmission risk and infection control in a hospital emergency department in Lima, Peru.* The International Journal of Tuberculosis and Lung Disease, 14, 1120-6.

[7] Gillespie DT (1976) *A general method for numerically simulating the stochastic time evolution of coupled chemical reactions.* Journal of Computational Physics, 22, 403-34.

[8] Gómez-Corral A, López-García M (2018) *Perturbation analysis in finite LD-QBD processes and applications to epidemic models.* Numerical Linear Algebra with Applications, 25, e2160.

[9] Harbarth S, Sax H, Gastmeier P (2003) *The preventable proportion of nosocomial infections: an overview of published reports.* Journal of Hospital Infection, 54, 258-66.

[10] Keeling MJ, Ross JV (2008) *On methods for studying stochastic disease dynamics.* Journal of the Royal Society Interface 5, 171-81.

[11] King MF, Noakes CJ, Sleigh PA (2015) *Modelling environmental contamination in hospital single and four-bed rooms.* Indoor Air 25(6): 694-707

[12] Ko G, Burge HA, Nardell EA, Thompson KM (2001) *Estimation of Tuberculosis risk and incidence under upper room ultraviolet germicidal irradiation in a waiting room in a hypothetical scenario.* Risk Analysis, 21, 657-74.

[13] Ko G, Thompson KM, Nardell EA (2004) *Estimation of Tuberculosis risk on a commercial airliner.* Risk Analysis, 24, 379-88.

[14] Li Y, Huang X, Yu ITS, Wong TW, Qian H (2005) *Role of air distribution in SARS transmission during the largest nosocomial outbreak in Hong Kong.* Indoor Air, 15, 83-95.

[15] Li Y, Leung GM, Tang JW, Yang X, Chao CYH, Lin JZ, Lu JW, Nielsen PV, Niu J, Qian H, Sleigh AC, Su HJJ, Sundell J, Wong TW, Yuen PL (2007) *Role of ventilation in airborne transmission of infectious agents in the built environment - A multidisciplinary systematic review.* Indoor Air, 17, 2-18.

[16] Liao C-M, Chang C-F, Liang H-M (2005) *A probabilistic transmission dynamic model to assess indoor airborne infection risks.* Risk Analysis, 25, 1097-107.

[17] López-García M (2016) *Stochastic descriptors in an SIR epidemic model for heterogeneous individuals in small networks.* Mathematical Biosciences, 271, 42-61.

[18] National Disease Surveillance Centre (2003) *National guidelines on the management of outbreaks of norovirus infection in health-care settings.* ISBN: 0-9540177-4-9.

[19] Noakes CJ, Beggs CB, Sleigh PA, Kerr KG (2006) *Modelling the Transmission of Airborne Infections in Enclosed Spaces.* Epidemiology and Infection, 134(5), 1082-1091

[20] Noakes CJ, Sleigh PA (2009) *Mathematical models for assessing the role of airflow on the risk of airborne infection in hospitals.* Journal of the Royal Society Interface, rsif20090305.

For Peer Review

[21] Noakes CJ, Sleigh PA, Khan A (2012) *Appraising Healthcare ventilation from combined infection control and energy perspectives.* HVAC&R Research, 18(4), 658-670.

[22] Pelupessy I, Bonten MJ, Diekmann O (2002) *How to assess the relative importance of different colonization routes of pathogens within hospital settings.* Proceedings of the National Academy of Sciences, 99, 5601-5.

[23] Riley EC, Murphy G, Riley RL (1978) *Airborne spread of measles in a suburban elementary school.* American Journal of Epidemiology, 107: 421-432.

[24] Reed D, Kemmerly S (2009) *Infection control and prevention: a review of hospital-acquired infections and the economic implications.* The Ochsner Journal, 9, 27-31.

[25] Sax H, Allegranzi B, Chraïti M-N, Boyce J, Larson E, Pittet D (2009) *The World Health Organization hand hygiene observation method.* American Journal of Infection Control, 37, 827-34.

[26] Sze TGN, Chao CYH (2010) *Review and comparison between the Wells-Riley and dose-response approaches to risk assessment of infectious respiratory diseases.* Indoor Air, 20, 2-16.

[27] Tang JW, Li Y, Eames I, Chan PKS, Ridgway GL (2006) *Factors involved in the aerosol transmission of infection and control of ventilation in health-care premises.* Journal of Hospital Infection, 64, 100-14.

[28] Teunis PFM, Moe CL, Liu P, Miller SE, Lindesmith L, Baric RS, Pendu JL, Calderon RL (2008) *Norwalk virus: how infectious is it?* Journal of Medical Virology, 80, 1468-76.

[29] Wein LM, Atkinson MP (2009) *Assessing infection control measures for pandemic influenza.* Risk Analysis, 29, 949-62.

[30] Wolkewitz M, Dettenkofer M, Bertz H, Schumacher M, Huebner J (2008) *Environmental contamination as an important route for the transmission of the hospital pathogens VRE: modeling and prediction of classical interventions.* Infectious Diseases: Research and Treatment, 1.

[31] https://github.com/matml/A-multi-compartment-SIS-stochastic-model-with-zonal-ventilation-for-nosocomial-outbreaks

Appendix

First-step arguments

We explain here in detail how Eq. (4) is obtained, in order to compute probabilities

$$p_{(i_1,i_2,\ldots,i_M)}(n) \quad = \quad \mathbb{P}(R = n \mid (I_1(0), I_2(0), \ldots, I_M(0)) = (i_1, i_2, \ldots, i_M)),$$

for any initial state $(i_1, \ldots, i_M) \in \mathcal{C}$. These equations are obtained by following a first-step argument, and applying the total probability law regarding the following event occurring in the stochastic process. In particular, let us denote by $(i_1, i_2, \ldots, i_j, \ldots, i_M) \to (i_1, i_2, \ldots, i_j + 1, \ldots, i_M)$ the event representing that, if the process is initially at state $(i_1, i_2, \ldots, i_j, \ldots, i_M)$, it moves to state $(i_1, i_2, \ldots, i_j{+}1, \ldots, i_M)$ in the next *jump* (that is, that we have $(i_1, i_2, \ldots, i_j, \ldots, i_M)$ infectives within the hospital ward at current time, and the next event that occurs is an infection at zone j). From the theory of CTMCs, it is well-know that the probability of this event

For Peer Review

occurring is

$$\mathbb{P}((i_1, i_2, \ldots, i_j, \ldots, i_M) \to (i_1, i_2, \ldots, i_j + 1, \ldots, i_M)) = \frac{\lambda_j(i_1, \ldots, i_M)(N_j - i_j)}{\sum_k (\lambda_k(i_1, \ldots, i_M)(N_k - i_k) + (\gamma_k + \delta_k)i_k)};$$

that is, it is equal to the ratio between the rate corresponding to this event and the sum of all the rates corresponding to all the possible events that can actually occur (infection at any zone, discharge at any zone or detection of an infective individual at any zone). Same argument applies to any other possible event, so that $\mathbb{P}((i_1, i_2, \ldots, i_j, \ldots, i_M) \to (i_1, i_2, \ldots, i_j - 1, \ldots, i_M))$ is given by

$$\frac{\gamma_j i_j}{\sum_k (\lambda_k(i_1, \ldots, i_M)(N_k - i_k) + (\gamma_k + \delta_k)i_k)}$$

and $\mathbb{P}((i_1, i_2, \ldots, i_j, \ldots, i_M) \to \Delta)$ is given by

$$\frac{\sum_j \delta_j i_j}{\sum_k (\lambda_k(i_1, \ldots, i_M)(N_k - i_k) + (\gamma_k + \delta_k)i_k)}.$$

Thus, we can apply the total probability law so that $p_{(i_1,\ldots,i_M)}(0) = \mathbb{P}(R = 0)$ is equal to

$$\begin{aligned}
&\sum_{k=1}^{M} \Big(\mathbb{P}(R = 0 \mid (i_1, \ldots, i_M) \to (i_1, \ldots, i_k + 1, \ldots, i_M)) \mathbb{P}((i_1, \ldots, i_M) \to (i_1, \ldots, i_k + 1, \ldots, i_M)) \\
&+ \mathbb{P}(R = 0 \mid (i_1, \ldots, i_M) \to (i_1, \ldots, i_k - 1, \ldots, i_M)) \mathbb{P}((i_1, \ldots, i_M) \to (i_1, \ldots, i_k - 1, \ldots, i_M)) \Big) \\
&+ \mathbb{P}(R = 0 \mid (i_1, \ldots, i_M) \to \Delta) \mathbb{P}((i_1, \ldots, i_M) \to \Delta) \\
&= \frac{1}{\sum_{j=1}^{M} (\lambda_j(i_1, \ldots, i_M)(N_j - i_j) + (\gamma_j + \delta_j)i_j)} \sum_{k=1}^{M} \left(\gamma_k i_k p_{(i_1,\ldots,i_k-1,\ldots,i_M)}(0) + \delta_k i_k \right),
\end{aligned}$$

corresponding to Eq. (4).

Algorithms

Algorithm 1 (*For computing the probability distribution of* R)

For all $(i_1, \ldots, i_M) \in \mathcal{C}$, compute and store $\lambda_j(i_1, \ldots, i_M)$, for $1 \le j \le M$, from Eq. (3);

$n = 0$;

$p_{(0,0,\ldots,0)}(0) = 1$;

For $I = 1, 2, \ldots, N$*:*

For all $\mathbf{i} = (i_1, i_2, \ldots, i_M)$ *s.t.* $\sum_{j=1}^{M} i_j = I$*:*

$$p_{\mathbf{i}}(0) = \frac{1}{\sum_j (\lambda_j(i_1,\ldots,i_M)(N_j - i_j) + (\gamma_j + \delta_j)i_j)} \sum_k \left(\gamma_k i_k p_{(i_1,\ldots,i_k-1,\ldots,i_M)}(0) + \delta_k i_k \right)$$

For Peer Review

For $n = 1, 2, 3, \ldots$:

$p_{(0,0,\ldots,0)}(n) = 0$;

For $I = 1, 2, \ldots, N$:

For all $\mathbf{i} = (i_1, i_2, \ldots, i_M)$ *s.t.* $\sum_{j=1}^{M} i_j = I$:

$$p_{\mathbf{i}}(n) = \frac{1}{\sum_j (\lambda_j(i_1,\ldots,i_M)(N_j - i_j) + (\gamma_j + \delta_j) i_j)} \sum_k \Big(\gamma_k i_k p_{(i_1,\ldots,i_k-1,\ldots,i_M)}(n) + \lambda_k(i_1, \ldots, i_M) \times (N_k - i_k) p_{(i_1,\ldots,i_k+1,\ldots,i_M)}(n-1) \Big);$$

For Peer Review