DOI: 10.1111/risa.14295

ORIGINAL ARTICLE

The Wells–Riley model revisited: Randomness, heterogeneity, and transient behaviours

Alexander J. Edwards[1] | **Marco-Felipe King**[2] | **Catherine J. Noakes**[2] | **Daniel Peckham**[3,4] | **Martín López-García**[5]

[1]EPSRC Centre for Doctoral Training in Fluid Dynamics, University of Leeds, Leeds, UK

[2]School of Civil Engineering, University of Leeds, Leeds, UK

[3]Leeds Institute of Medical Research, University of Leeds, Leeds, UK

[4]Leeds Teaching Hospitals NHS Trust, Leeds, UK

[5]School of Mathematics, University of Leeds, Leeds, UK

Correspondence
Alexander J. Edwards, EPSRC Centre for Doctoral Training in Fluid Dynamics, University of Leeds, Woodhouse Lane, Leeds, LS2 9JT, UK.
Email: scaje@leeds.ac.uk

Funding information
Engineering and Physical Sciences Research Council, Grant/Award Number: EP/S022732/1

Abstract

The Wells–Riley model has been widely used to estimate airborne infection risk, typically from a deterministic point of view (i.e., focusing on the average number of infections) or in terms of a per capita probability of infection. Some of its main limitations relate to considering well-mixed air, steady-state concentration of pathogen in the air, a particular amount of time for the indoor interaction, and that all individuals are homogeneous and behave equally. Here, we revisit the Wells–Riley model, providing a mathematical formalism for its stochastic version, where the number of infected individuals follows a Binomial distribution. Then, we extend the Wells–Riley methodology to consider transient behaviours, randomness, and population heterogeneity. In particular, we provide analytical solutions for the number of infections and the per capita probability of infection when: (i) susceptible individuals remain in the room after the infector leaves, (ii) the duration of the indoor interaction is random/unknown, and (iii) infectors have heterogeneous quanta production rates (or the quanta production rate of the infector is random/unknown). We illustrate the applicability of our new formulations through two case studies: infection risk due to an infectious healthcare worker (HCW) visiting a patient, and exposure during lunch for uncertain meal times in different dining settings. Our results highlight that infection risk to a susceptible who remains in the space after the infector leaves can be nonnegligible, and highlight the importance of incorporating uncertainty in the duration of the indoor interaction and the infectivity of the infector when estimating risk.

KEYWORDS
airborne transmission, probability, risk assessment, stochastic, Wells–Riley

1 | INTRODUCTION

Airborne transmission is a common infection route for many respiratory pathogens and can be particularly important due to its facilitation of superspreading events, often requiring tailored mitigations such as face masks, ventilation, or air cleaning. Transmission of airborne infections is complex and the mode of transmission may vary with different pathogens. For example, tuberculosis spreads almost exclusively through the air via inhalation of very small aerosols less than 5 μm diameter (Wang et al., 2021), while influenza is thought to be spread through both small aerosols and larger droplets which can be inhaled or transferred through touch and surface contamination (Milton et al., 2013). For some pathogens, such as influenza and severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), disentangling the preferred form of transmission can be challenging; transmission risk can sometimes be environment-dependent, and possibly include a combination of various routes (Bueno de Mesquita et al., 2020; Wang et al., 2021).

In order to develop a better understanding of transmission mechanisms and to quantify the likely impact of different mit-

igation strategies, one can use a quantitative microbial risk assessment (QMRA) approach. QMRA models assess exposure to particular risks or hazards and use a dose–response approach to characterise the likelihood of an adverse health outcome from exposures to pathogens (Haas, 2020). Previous applications of QMRA include assessing faecal contamination in water systems (Soller et al., 2010), food-borne diseases (Pérez-Rodríguez et al., 2008), and more recently respiratory pathogen transmission, for example, the transmission of influenza and SARS-CoV-2 (Buonanno et al., 2020; Kitajima et al., 2020; Miller et al., 2022; Wilson et al., 2021).

To date, the majority of QMRA models for assessing airborne infection risk within indoor environments use the traditional Wells–Riley approach (Riley et al., 1978; Wells, 1955), which considers that the probability of each single susceptible individual (e.g., sharing an indoor space with a constant number I of infectious individuals) becoming exposed (i.e., infected but not yet infectious) during a time interval $[0, T]$ is

$$P(T) = \frac{E(T)}{S_0} = 1 - \mathrm{e}^{-\frac{Iqp}{Q}T}, \tag{1}$$

where $P(T)$ is the per capita probability of becoming exposed/infected, $E(T)$ is the estimated mean number of exposures (i.e., infections) during $[0, T]$, S_0 is the initial number of susceptible people, p [$\mathrm{m^3\ min^{-1}}$] is the pulmonary breathing rate of each individual, Q [$\mathrm{m^3\ min^{-1}}$] is the extract ventilation rate in the space, and q [quanta $\mathrm{min^{-1}}$] is the quanta production rate by the infectious individuals. A quantum is defined as the number of infectious airborne particles required to infect an individual; in particular, one quantum inhaled by an individual will cause an infection with probability $1 - \mathrm{e}^{-1}$ ($\approx 63.2\%$) (Wells, 1955). The quanta generation rate, q, cannot be experimentally measured, but estimated epidemiologically. If the exposure time, ventilation rate, and number of exposures are known from a particular outbreak, then the quanta generation rate can be calculated using Equation (1) (Sze To & Chao, 2010).

The standard Wells–Riley equation (1) relies on four key assumptions: (i) the air in the indoor space is well-mixed so that the concentration of pathogen in the air is spatially homogeneous, (ii) the concentration of pathogen in the air is at steady-state during $[0, T]$, (iii) individuals behave equally, and (iv) susceptible and infectious individuals remain in the room for the whole time interval $[0, T]$. These are simplifying assumptions which allow for mathematical tractability but can lead to unrealistic predictions, especially in specific scenarios. The well-mixed assumption is commonly referred to as a major limitation of the Wells–Riley model in the literature (Rudnick & Milton, 2003; Sze To & Chao, 2010; Zemouri et al., 2020), since it does not allow one to consider heterogeneities in the concentration of pathogen in the air within the indoor setting (Qian et al., 2009). On the other hand, the steady-state approximation does not allow consideration of transient effects which can be especially relevant under specific environmental conditions (e.g., a moving population (Arino et al., 2016), or under poor levels of ventilation and during small time scales (Edwards et al., 2023)). Despite these limitations, it does offer a quick and simple assessment of the risk in many spaces, and is often viewed as a simpler approach than other methods such as dose–response models or computational fluid dynamics (CFD) methods (Sze To & Chao, 2010; Zhao et al., 2022).

There have been many adaptations made to the traditional Wells–Riley model over the years in an attempt to overcome these limitations arising due to the original assumptions. Considering the influence of spatial effects on the Wells–Riley model using methods such as coupling with CFD (Gao et al., 2008; Qian et al., 2009; Tung & Hu, 2008; Wang et al., 2022), integrated spatial flow fractions (Guo et al., 2021), and general spatial adaptations (Lau et al., 2022; Pantelic & Tham, 2012) are common in attempting to overcome the well-mixed assumption. A recent adaption has been to consider the change of quanta in the space in more detail by using a transport equation for the concentration of pathogen in the space characterising further dynamical features of the outbreak (Boonmeemapasuk & Pochai, 2022; Ding et al., 2022; H. Li et al., 2021; J. Li et al., 2021; Qian et al., 2009; Timpitak & Pochai, 2022). Gammaitoni and Nucci (1997) allowed for varying quanta levels over time with a conservation equation, making the model closer to a dose–response approach, which has then been used extensively (Beggs et al., 2010; King et al., 2021; Knibbs et al., 2011; Rocha-Melogno et al., 2023). Wood et al. (2023) explore a coupling of the quanta conservation equation with an epidemic model, where they analyse important time scales and features arising from the dimensionless dynamical system.

Other adaptations look closer at individual aspects of the Wells–Riley model. For example, levels of excess CO_2 in the air, commonly considered as a proxy for amount of shared air in indoor spaces, was incorporated into the traditional Wells–Riley model by Rudnick and Milton (2003). Zemouri et al. (2020), Cammarata and Cammarata (2021), and Burridge et al. (2022) build on this work with further extensions of the model. Other adaptions include accounting for the deposition of airborne particles (Fisk et al., 2004; Franchimon et al., 2008), or the consideration of mitigation measures such as mask wearing (Dai & Zhao, 2020; Fantozzi et al., 2022; Fennelly & Nardell, 1998; Nazaroff et al., 1998; Nicas, 1996), social distancing (Sun & Zhai, 2020; Shang et al., 2022), and particle filtration or air cleaning (Fisk et al., 2004; Nazaroff et al., 1998). The majority of the models mentioned above typically only consider single-zone environments, but multi-zone versions have also been developed in recent years (Edwards et al., 2023; Ko et al., 2001, 2004; López-García et al., 2019; Noakes & Sleigh, 2009).

Despite the many different generalisations of the Wells–Riley model, many studies (Pavilonis et al., 2021; Rowe et al., 2022; Zhao et al., 2022) still use the classical approach that was originally presented by Wells (1955) and Riley et al. (1978). Moreover, although the Wells–Riley model quantifies individual infection risk in terms of a per capita probability of infection, most of the applications in the literature typically

use deterministic approaches focusing on mean quantities (e.g., mean number of individuals infected in an indoor space during $[0, T]$), typically for single zones and under the steady-state assumption. Stochastic formulations have been applied using Monte Carlo numerical approaches (Noakes & Sleigh, 2009), but those which can deal with more complex scenarios (multi-zone, transient concentration, random length of stay of the infectious individual, or heterogeneities across individuals), and which can allow estimation of the probability distribution of the number of exposed individuals during a specific time interval, are still to be properly formulated from a mathematical point of view. Having these tools would enable a more realistic approach to modelling real scenarios, better representing real-life behaviours and environments, and accounting for stochasticity, making the model more generalisable and applicable.

In this paper, we present stochastic variations to the existing Wells–Riley framework that focus on probability distributions instead of mean quantities. This enables explicit analytical solutions that allow for more general, and accurate applications of the Wells–Riley model when assessing risk. In particular, we analyse the nonnegligible risk of infection to a susceptible person who remains in the space long after the infector leaves. We incorporate randomness in the duration of the indoor interaction, acknowledging that it is not always possible to know exactly how long an infectious individual may be present for, especially when trying to assess risk before the indoor gathering actually takes place. Finally, we address population heterogeneity by considering a probability distribution for the quanta generation rate, accounting for heterogeneous infectivity across individuals. These advancements will allow for a wider use of the traditional Wells–Riley model, increasing its applicability and providing a better representation of the heterogeneity and uncertainty that is present when assessing infection risk in real-life settings.

2 | THE WELLS–RILEY MODEL

The traditional Wells–Riley model estimates airborne infection risk in an indoor setting during a fixed time period, say $[0, T]$ for some $T > 0$. One can classify individuals in the population (N individuals in an indoor setting) according to their disease status, typically Susceptible ($S(t)$) and Infectious ($I(t)$), with $S_0 + I_0 = N$. Here, S_0 is the number of initially susceptible individuals and I_0 is the initial number of infectious individuals.

If the time period, $[0, T]$, is relatively short (i.e., hours rather than days), individuals who become infected are not infectious during the indoor gathering, and the only infectious individuals are the ones initially in the room at time $t = 0$, I_0. It is then more precise to classify individuals as Susceptible ($S(t)$) and Exposed ($E(t)$), where individuals are exposed if they have been infected but are not yet infectious, whereas the number of infectious individuals remains constant, $I(t) = I_0$, and $S(t) + E(t) + I_0 = N$ for all $t \in [0, T]$.

From now on, in this paper, we talk about "infections" and "exposures" interchangeably, and focus on scenarios where the incubation period of the pathogen is longer than the time duration of the indoor interaction, so that individuals become infected but not infectious during the time scales of interest.

A deterministic version of the standard Wells–Riley model, or if one interprets $S(t)$ and $E(t)$ as mean quantities, can be obtained via the ordinary differential equations (ODEs)

$$\begin{aligned} \frac{dS(t)}{dt} &= -pC(t)S(t), \\ \frac{dE(t)}{dt} &= pC(t)S(t), \end{aligned} \tag{2}$$

which represent a Susceptible–Exposed (SE) compartmental epidemic model (Edwards et al., 2023). Infectious quanta is defined in such a way that the rate at which individuals are infected is considered to be proportional to the concentration of pathogen in the air at any given time, $C(t)$ [quanta m^{-3}], where the constant of proportionality is the pulmonary (breathing) rate p [m^3 min^{-1}], leading to the per capita infectivity rate $pC(t)$. Although we consider here an SE compartmental epidemic model to represent a duration of the indoor interaction which is shorter than the incubation period of the pathogen (so that individuals become exposed but not yet infectious), the Wells–Riley methodology has also been linked to other types of compartmental epidemic models, such as the Susceptible-Infected-Susceptible (SIS) (López-García et al., 2019) and Susceptible-Exposed-Infected-Recovered (SEIR) (Wood et al., 2023), to model alternative situations.

In particular, the concentration of pathogen in the air is modelled via the ODE

$$v\frac{dC(t)}{dt} = qI_0 - QC(t), \tag{3}$$

where v [m^3] is the room volume, q [quanta min^{-1}] is the quanta production rate, Q [m^3 min^{-1}] is the extract ventilation rate, and I_0 is the number of infectious individuals in the room during the time interval of interest.

Equation (3) can be solved analytically to obtain a transient solution for $C(t)$. If we let the initial condition be $C(t = 0) = C_0 = 0$ (e.g., time $t = 0$ represents whenever the infector(s) enter(s) the room), then we obtain the following solution:

$$C(t) = \frac{qI_0}{Q}\left(1 - e^{-\frac{Qt}{v}}\right), \quad t \geq 0. \tag{4}$$

More general initial concentrations $C_0 > 0$ may be considered instead to represent specific situations. Moreover, one can solve Equation (3) for steady-state by setting $\frac{dC(t)}{dt} = 0$, which leads to

$$C^* = \frac{qI_0}{Q}. \tag{5}$$

In the following subsections, we show how to exploit these estimates of the concentration of pathogen in the air in order to estimate the number of exposures during the time interval of interest $[0, T]$. Finally, we note that although Q [m^3 min^{-1}] represents an extract ventilation rate and is the parameter considered in most applications of the Wells–Riley model in the literature, there are other mechanisms that can contribute to the removal of pathogenic material from the air, and which can be easily incorporated into these models. In particular the extract ventilation rate Q can be replaced by a more general removal rate R [m^3 min^{-1}], which incorporates other mechanisms such as viral inactivation (biological decay in infectiousness of the pathogen) and viral deposition such as onto surfaces or the ground (Miller et al., 2021), so that

$$R = Q + v\, r_i + v\, r_d,$$

where v [m^3] is the volume of the space, r_i [min^{-1}] is the viral inactivation rate, and r_d [min^{-1}] is the viral deposition rate. We use Q from now on within the next sections to keep notation consistent, but the more general removal rate R can be used instead in all equations in this paper, depending on the scenario under consideration. In fact, general removal rates are considered in the case studies analysed in Section 4.

2.1 | Steady-state concentration

From the steady-state concentration solution, Equation (5), to estimate the number of exposed individuals during $[0, T]$ one just needs to set $C(t) = C^*$ during the time interval of interest $[0, T]$, and substitute it back into Equation (2), leading to the solution

$$E(T) = S_0\left(1 - \mathrm{e}^{-\frac{pqI_0}{Q}T}\right).$$

Here, the per capita probability of infection is $E(T)/S_0 = 1 - e^{-\frac{pqI_0}{Q}T}$; that is, the standard Wells–Riley per capita probability of infection given in Equation (1) under steady-state assumptions. From a stochastic point of view, $E(T)$ can be considered as a random variable rather than as a deterministic or mean quantity. Since individuals become infected independently of each other, one can interpret this as a sequence of S_0 Bernoulli experiments, so that the number of exposed individuals in $[0, T]$ follows a Binomial distribution

$$E(T) \sim Bin\left(S_0\ ,\ 1 - \mathrm{e}^{-\frac{pqI_0}{Q}T}\right), \tag{6}$$

which provides a stochastic interpretation of the deterministic estimate. We note that the mean number of exposed individuals during $[0, T]$ is then $S_0\left(1 - \mathrm{e}^{-\frac{pqI_0}{Q}T}\right)$. However, this stochastic interpretation also allows one to estimate the probability of observing a specific number of exposures $E(T) = n$,

$$\mathbb{P}\{E(T) = n\} = \binom{S_0}{n}\left(1 - \mathrm{e}^{-\frac{pqI_0}{Q}T}\right)^n\left(\mathrm{e}^{-\frac{pqI_0}{Q}T}\right)^{S_0-n}, \quad n \in \{0, 1, \dots, S_0\}. \tag{7}$$

This distribution can also be derived from first principles using the *master equation* (i.e., Kolmogorov differential equations) of the stochastic process (see Appendix A).

2.2 | Transient concentration

A particular limitation of the approach in Subsection 2.1 is that the transient concentration of pathogen in the air, $C(t)$ for $t \in [0, T]$, is approximated with the constant steady-state value C^*; this can lead to an overestimation of the predicted quanta concentration levels during early times and be particularly unrealistic in scenarios with low ventilation rates, larger room volumes, or when the infector is only present for a short period of time (Edwards et al., 2023). Alternatively, the transient solution of the concentration of airborne pathogen, Equation (4), can be considered and substituted back into Equation (2) to reach the more precise estimate (Gammaitoni & Nucci, 1997; Sze To & Chao, 2010)

$$E(T) = S_0\left(1 - \mathrm{e}^{\frac{pqI_0}{Q^2}\left[v\left(1-\mathrm{e}^{-\frac{Q}{v}T}\right)-QT\right]}\right).$$

Following the same arguments as before, $E(T)$ can be interpreted as a random variable in a stochastic version of the Wells–Riley model, and follows a Binomial distribution

$$E(T) \sim Bin\left(S_0\ ,\ 1 - \mathrm{e}^{\frac{pqI_0}{Q^2}\left[v\left(1-\mathrm{e}^{-\frac{Q}{v}T}\right)-QT\right]}\right),$$

so that the mean number of exposed individuals during $[0, T]$ is $S_0\left(1 - \mathrm{e}^{\frac{pqI_0}{Q^2}\left[v\left(1-\mathrm{e}^{-\frac{Q}{v}T}\right)-QT\right]}\right)$, and

$$\mathbb{P}\{E(T) = n\} = \binom{S_0}{n}\left(1 - \mathrm{e}^{\frac{pqI_0}{Q^2}\left[v\left(1-\mathrm{e}^{-\frac{Q}{v}T}\right)-QT\right]}\right)^n\left(\mathrm{e}^{\frac{pqI_0}{Q^2}\left[v\left(1-\mathrm{e}^{-\frac{Q}{v}T}\right)-QT\right]}\right)^{S_0-n}$$

for $n \in \{0, 1, \dots, S_0\}$.

3 | BEYOND THE WELLS–RILEY MODEL: RANDOMNESS, HETEROGENEITY, AND TRANSIENT BEHAVIOURS

In this section, we propose extensions of the traditional Wells–Riley model to scenarios that are more representative

of those experienced in real life. In Section 3.1, transient effects are incorporated by analysing the situation where the infectious individual(s) eventually leave(s) the room, whereas the susceptible individuals stay. In Section 3.2, we consider the unknown duration of an interaction with an infectious individual by incorporating a random time for the length of exposure, T. Finally, we explore infectiousness heterogeneity across individuals in Section 3.3 by considering a probability distribution for the quanta production rate q.

3.1 | The infector leaves the room

In Section 2, we estimated the risk for susceptible individuals who share the same room with the initial I_0 infectors during a time period of $[0, T]$ for some time $T > 0$. However, we note that depending on the ventilation settings for the space, the environmental conditions (e.g., temperature, relative humidity), or the biological properties of the airborne pathogen itself, susceptible individuals are also exposed to a (potentially nonnegligible) residual risk even after the infector(s) leave(s) the room, if they remain in situ. For example, a recent study reported evidence to suggest COVID-19 transmission in a shared bathroom between an infectious individual and a susceptible one who entered up to 40 minutes after the infector had left (Jung et al., 2021) and we see further evidence of transmission between neighbouring zones through connected airflow (Eichler et al., 2021; Jung et al., 2021; SAGE, 2021) demonstrating the possibility of airborne transmission despite the absence of an infector.

We focus here on estimating infection risk for susceptible individuals who remain in the room during $[0, T + t]$, whereas the infector(s) leave(s) at time T, for some $T, t > 0$; in Section 4.1, this will apply to a scenario where an infectious healthcare worker (HCW) visits the room of a susceptible hospital patient during a time interval $[0, T]$. We note that our analysis here can be easily adapted to the situation where a susceptible individual enters the room after the infector has already left. The overall infection risk for each susceptible individual during $[0, T + t]$, in terms of the per capita infection probability $P(T + t)$ of the individual being exposed/infected during $[0, T + t]$, can be split according to two distinct periods: the probability of the susceptible individual becoming exposed/infected during the time period during which the infector is present ($[0, T]$), $P_1(T)$, and the probability of the susceptible individual becoming exposed/infected during the time period after the infector leaves the room ($[T, T + t]$), $P_2(T, T + t)$, due to pathogenic material which remains in the air. In particular, the overall probability of infection during $[0, T + t]$ for a susceptible individual in the indoor setting is

$$\begin{aligned} P(T+t) &= \mathbb{P}\{\text{infection in } [0, T+t]\} \\ &= \mathbb{P}\{''\text{infection in } [0,T]'' \text{ or} \\ &\quad ''\text{no infection in } [0,T] \text{ and infection in } [T, T+t]''\} \\ &= P_1(T) + (1 - P_1(T))P_2(T, T+t). \end{aligned} \tag{8}$$

The per capita probability of each susceptible individual to become exposed during $[0, T]$, $P_1(T)$, can be calculated as described in Section 2.2, using the transient concentration of pathogen in the air during $[0, T]$, so that

$$P_1(T) = 1 - \mathrm{e}^{\frac{pqI_0}{Q^2}\left[v\left(1-\mathrm{e}^{-\frac{Q}{v}T}\right)-QT\right]}. \tag{9}$$

To estimate exposure during the second period, when the infector is absent, we consider

$$C(T) = \frac{qI_0}{Q}\left(1 - \mathrm{e}^{-\frac{Q}{v}T}\right), \tag{10}$$

as the initial condition for the concentration of pathogen in the air during the period $[T, T + t]$, while considering $I_0 = 0$ during $[T, T + t]$ (since the infector(s) have left the room). Thus, Equation (3) becomes

$$v\frac{dC(T+t)}{dt} = -QC(T+t),$$

and solving this with the initial condition in Equation (10) gives

$$C(T+t) = \frac{qI_0}{Q}\left(1 - \mathrm{e}^{-\frac{Q}{v}T}\right)\mathrm{e}^{-\frac{Q}{v}t},\ t \geq 0.$$

Hence, to calculate the per capita exposure probability during $[T, T + t]$, $P_2(T, T + t)$, we can solve

$$\frac{dP_2(T, T+t)}{dt} = pC(T+t) = \frac{pqI_0}{Q}\left(1 - \mathrm{e}^{-\frac{Q}{v}T}\right)\mathrm{e}^{-\frac{Q}{v}t},$$

which gives

$$P_2(T, T+t) = \frac{pqvI_0}{Q^2}\left(1 - \mathrm{e}^{-\frac{Q}{v}T}\right)\left(1 - \mathrm{e}^{-\frac{Q}{v}t}\right),\ t \geq 0. \tag{11}$$

As may be common in specific settings, such as hospitals or care homes, susceptible individuals (e.g., a patient) may stay in the room long after the infector(s) (e.g., an HCW) leave(s). Thus, one can use the estimate

$$\lim_{t\to\infty} P_2(T, T+t) = \frac{pqvI_0}{Q^2}\left(1 - \mathrm{e}^{-\frac{Q}{v}T}\right). \tag{12}$$

Finally, combining Equations (8), (9), and (11) gives the overall probability of infection during $[0, T + t]$

$$\begin{aligned} \mathbb{P}\{\text{infection in } [0, T+t]\} = 1 - \mathrm{e}^{\frac{pqI_0}{Q^2}\left[v\left(1-\mathrm{e}^{-\frac{Q}{v}T}\right)-QT\right]} \\ \left(1 - \frac{pqvI_0}{Q^2}\left(1 - \mathrm{e}^{-\frac{Q}{v}T}\right)\left(1 - \mathrm{e}^{-\frac{Q}{v}t}\right)\right). \end{aligned} \tag{13}$$

Setting $t = 0$ we recover the probability of infection during $[0, T]$, $P_1(T)$ in Equation (9). Once again, if $t \to \infty$, that is, if the susceptible individual stays in the indoor space long after the infector has left, the overall per capita probability of infection can be estimated as

$$\lim_{t\to\infty} \mathbb{P}\{\text{infection in } [0, T+t]\} = 1 - \mathrm{e}^{\frac{pqI_0}{Q^2}\left[v\left(1-\mathrm{e}^{-\frac{Q}{v}T}\right)-QT\right]} \left(1 - \frac{pqvI_0}{Q^2}\left(1 - \mathrm{e}^{-\frac{Q}{v}T}\right)\right). \quad (14)$$

3.2 | Uncertainty in the duration of the indoor gathering

Here, we consider estimation of exposure risk during $[0, T]$ when there is randomness or uncertainty in the duration of the event, T. In this case, we consider that T is a random variable rather than a constant, which could represent the typical duration of a particular type of gathering (e.g., lunch across different hospitality venues in our results in Section 4.2) or when analysing a particular outbreak where the duration is not known. We focus on the steady-state modelling framework for simplicity, and consider that $T \sim G(\cdot)$ follows a probability distribution with density function $f_T(t)$. One can compute the per capita probability of infection in $[0, T]$, using Equation (1), as

$$P(T) = \int_0^{+\infty} \left(1 - \mathrm{e}^{-\frac{pqI_0}{Q}t}\right) f_T(t)\, dt\,. \quad (15)$$

While this expression provides quantification for the per capita infection risk (i.e., the probability of infection for an individual attending the indoor gathering, of duration $[0, T]$, where T is random and unknown), the distribution of the number of infections/exposures, $E(T)$, is trickier to find. Given that the time duration is random, infection events are no longer independent Bernoulli trials, and thus $E(T)$ is not a Binomial random variable. This lack of independence can be noticed from the fact that, for example, if a particular individual gets infected during $[0, T]$, with T unknown, this increases the likelihood of the indoor gathering having lasted for a longer period (i.e., T being large), which increases the probability of other individuals having been exposed during the indoor gathering. Still, one can compute the probability distribution of the number of exposures, in terms of probabilities $\mathbb{P}\{E(T) = n\}$ for $n \in \{0, 1, \ldots, S_0\}$, by solving the integral

$$\mathbb{P}\{E(T) = n\} = \int_0^{+\infty} \mathbb{P}\{E(T) = n | T = t\} f_T(t)\, dt\,.$$

Recall, from Equation (6), that $E(T|T = t) \sim Binomial(S_0, 1 - \mathrm{e}^{-\frac{pqI_0}{Q}t})$, and so we get

$$\mathbb{P}\{E(T) = n\} = \int_0^{+\infty} \binom{S_0}{n} \left(1 - \mathrm{e}^{-\frac{pqI_0}{Q}t}\right)^n \left(\mathrm{e}^{-\frac{pqI_0}{Q}t}\right)^{S_0-n} f_T(t)\, dt\,. \quad (16)$$

Although Equations (15) and (16) can easily be solved numerically for any density function $f_T(t)$ of interest, we next explore analytical expressions when specific distributions for T are considered. In particular, we look at the Exponential and Erlang distributions as they are typically used to model waiting times in many applications.

3.2.1 | Exponential distribution

One can model the duration of the indoor gathering in terms of an Exponential distribution with rate λ, $T \sim Exp(\lambda)$, so that $f_T(t) = \lambda \mathrm{e}^{-\lambda t}$ for $t \geq 0$. Then, the per capita probability of infection during $[0, T]$, Equation (15), becomes (see Appendix B for the full derivation)

$$P(T) = 1 - \frac{\lambda}{\frac{pqI_0}{Q} + \lambda}. \quad (17)$$

On the other hand, the probability of observing n infections, $\mathbb{P}\{E(T) = n\}$, can be estimated using Equation (16), which gives

$$\mathbb{P}\{E(T) = n\} = \binom{S_0}{n} \sum_{i=0}^{n} (-1)^i \binom{n}{i} \frac{\lambda}{(S_0-n+i)\frac{pqI_0}{Q}+\lambda}, \quad n \in \{0, 1, \ldots, S_0\}. \quad (18)$$

We refer the reader to Appendix C for a detailed derivation.

3.2.2 | Erlang distribution

While the Exponential distribution has been widely used in the literature to model waiting times in many applications, partly due to its memoryless property, it can overestimate short times (since the density function decays exponentially from its maximum value at $f_T(0) = \lambda$), which may be unrealistic in specific situations. Alternatively, the Erlang distribution (which is a Gamma distribution with integer shape parameter) can be used, since it allows for a more Log-Gaussian-like shape (unimodal and defined on the positive real numbers); see, for example, Figure 4 where we use the Erlang distribution to model lunch times in different hospitality venues.

If one considers $T \sim Erlang(k,\ \lambda)$, then $f_T(t) = \frac{\lambda^k\, t^{k-1}\, \mathrm{e}^{-\lambda t}}{(k-1)!}$ for $t \geq 0$. Thus, the per capita probability of infection during $[0, T]$ becomes (Appendix B)

$$P(T) = 1 - \left(\frac{\lambda}{\frac{pqI_0}{Q} + \lambda}\right)^k. \quad (19)$$

The number of exposures during $[0, T]$ can be estimated in terms of the probabilities (Appendix C)

$$\mathbb{P}\{E(T) = n\} = \binom{S_0}{n} \sum_{i=0}^{n} (-1)^i \binom{n}{i} \left(\frac{\lambda}{(S_0 - n + i)\frac{pqI_0}{Q} + \lambda} \right)^k, \quad n \in \{0, 1, \dots, S_0\}. \tag{20}$$

We note that, since the $Erlang(k, \lambda)$ can be thought of as a sum of k exponential distributions with rate λ, it is clear that if one sets $k = 1$ in Equations (19) and (20), we recover Equations (17) and (18) for the exponential case, respectively, as expected.

3.3 | Infectiousness heterogeneity

The standard Wells–Riley model relies on the assumption that all infectors, I_0, release pathogenic material [quanta] at a constant and common rate q. However, population heterogeneity in infectiousness is very common for many pathogens (Lloyd-Smith et al., 2005), which is one of the reasons why estimates for the quanta rate q for many pathogens often amount to significantly wide ranges (Mikszewski et al., 2022). The infectivity of a given individual can depend on many different factors including their viral load (Kidd et al., 2021), respiratory activity or behaviour (Buonanno et al., 2020), time since infection (Ma et al., 2021), or symptoms (Wang et al., 2023). Thus, it is of interest to consider the situation where the parameter q follows a probability distribution, $q \sim G(\cdot)$. If, for example, a single infector is in the room, $I_0 = 1$, their infectivity may be unknown and sampled from the corresponding distribution instead, which incorporates such heterogeneity at the population level.

As time T and the quanta production rate q both occur as linear terms in the exponent when calculating the steady-state per capita probability of infection, $P(T)$ in Equation (1), our results in Section 3.2 directly apply here when instead of a random duration T, one considers a random quanta rate $q \sim G(\cdot)$, for some generic probability distribution with density function $f(q)$, whereas T is constant. In particular, one gets

$$P(T) = \int_0^{+\infty} \left(1 - e^{-\frac{pqI_0}{Q}T}\right) f(q)\, dq\,.$$

$$\mathbb{P}\{E(T) = n\} = \int_0^{+\infty} \binom{S_0}{n} \left(1 - e^{-\frac{pqI_0}{Q}T}\right)^n \left(e^{-\frac{pqI_0}{Q}T}\right)^{S_0 - n} f(q)\, dq\,, \quad n \in \{0, 1, \dots, S_0\}.$$

Thus, following the same arguments as in Section 3.2, and if $q \sim Exp(\mu)$ with rate μ, so that $f(q) = \mu e^{-\mu q}$ for $q \geq 0$, we get the per capita probability of infection during $[0, T]$

$$P(T) = 1 - \frac{\mu}{\frac{pI_0}{Q}T + \mu},$$

and the number of exposures during $[0, T]$ can be estimated in terms of the probabilities

$$\mathbb{P}\{E(T) = n\} = \binom{S_0}{n} \sum_{i=0}^{n} (-1)^i \binom{n}{i} \frac{\mu}{(S_0 - n + i)\frac{pI_0}{Q}T + \mu}, \quad n \in \{0, 1, \dots, S_0\}.$$

Alternatively, if one considers instead that $q \sim Erlang(k, \mu)$, so that $f(q) = \frac{\mu^k e^{-\mu q}}{(k-1)!} q^{k-1}$ for $q \geq 0$, the per capita probability of infection during $[0, T]$ becomes

$$P(T) = 1 - \left(\frac{\mu}{\frac{pI_0}{Q}T + \mu} \right)^k,$$

and the number of exposures can be estimated via the probabilities

$$\mathbb{P}\{E(T) = n\} = \binom{S_0}{n} \sum_{i=0}^{n} (-1)^i \binom{n}{i} \left(\frac{\mu}{(S_0 - n + i)\frac{pI_0}{Q}T + \mu} \right)^k, \quad n \in \{0, 1, \dots, S_0\}.$$

4 | RESULTS

In this section, we consider two case studies to show the applicability of the methodology presented in Section 3 in real-life scenarios. Case Study 1 (Section 4.1) refers to exposure risk in a healthcare setting, and investigates risk once the infector has left the room. This uses measured HCW visit times to a single-bed room and explores the infection risk to the susceptible patient who remains in the room long after the infectious HCW leaves, exploiting our results in Section 3.1. Case Study 2 (Section 4.2) explores infection risk when the duration of the outbreak is unknown. In particular, we consider data from the different times spent eating in various restaurant environments, and explore how each hospitality setting can lead to different infection risk depending on the duration of lunch, using methodology from Section 3.2.

4.1 | Case Study 1: Visit from a HCW

In this section, we model a scenario where an infectious HCW enters a patient's single-bed room in order to complete a particular care activity with a given duration, and then leaves. We estimate the infection risk for the susceptible patient who remains in the room for long after the infector has left (i.e., $t \to \infty$). In this section, we leverage our analytical results in Section 3.1, and in particular Equation (14).

The HCW visit times used in this section were taken from a previous study (King et al., 2021) which observed the

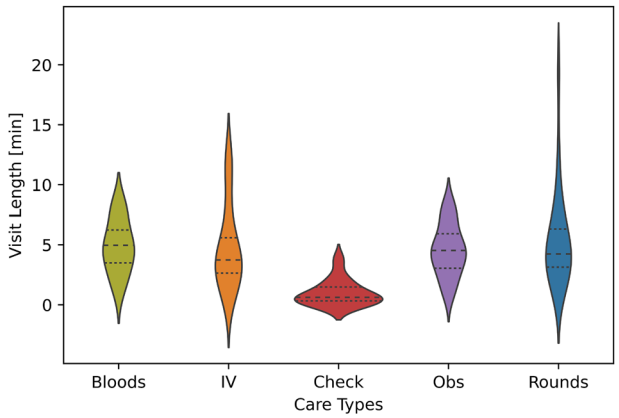

FIGURE 1 Violin plots illustrating the various durations of healthcare worker (HCW) visits when carrying out one of five care activities to a patient as measured by King et al. (2021).

duration of patient care for a variety of activities including taking blood samples ("Bloods", 13 visits recorded), intravenous care ("IV", 20 visits recorded), general checks ("Check", 13 visits recorded), observational care ("Obs", 32 visits recorded), and doctors' rounds ("Rounds", 24 visits recorded), totaling 102 recorded visits. We use the visits as 102 data points for the duration of care, which have mean $\mu = 4.37$ min and standard deviation (SD) $\sigma = 3.15$ min .

We set $v = 28.57$ m^3, which is the volume of a single-bed room on a UK NHS Trust adult respiratory ward used in our previous study (Edwards et al., 2023). For illustrative purposes, and since the quanta production rate for SARS-CoV-2 has been estimated to range from 15 to more than 4000 quanta h^{-1} (Mikszewski et al., 2022), we use $q = 360$ quanta h^{-1}, equating to $q = 6$ quanta min^{-1} in our results, and set the pulmonary breathing rate as $p = 0.01$ m^3 min^{-1} (Noakes & Sleigh, 2009). A range of ventilation rates are explored representing 0.5 Air Changes per Hour (ACH), 1.5 ACH, 3 ACH, and 6 ACH, giving $Q \in \{0.238, 0.714, 1.428, 2.856\}$ [m^3 min^{-1}]. For the viral inactivation and deposition rates, we consider plausible values in ranges proposed in Miller et al. (2021), in particular $r_i \in \{0, 0.0035, 0.007, 0.0105\}$ [min^{-1}] and $r_d \in \{0.005, 0.0115, 0.01815, 0.025\}$ [min^{-1}]. These combined removal mechanisms can result in overall removal rates ranging from $R = 0.381$ to $R = 3.870$ [m^3 min^{-1}], so in this section we explore the values $R \in \{0.381, 1.143, 2.147, 3.870\}$ [m^3 min^{-1}].

In Figure 1, we plot the distribution of visit times (as a violin plot) from HCWs for the five different care types in King et al. (2021). We can observe slight variations across the different activities, with "IV" and "Rounds" containing specific outliers (representing longer times) compared to the other activities. We can also see that "Check" typically has the shortest visit length, and a narrower distribution compared to other care types, with the shape of the violin plot concentrated around its mean. Apart from "Check," the quartiles for the other four types of care activity are similar in value, and we see a similar violin shape in the main body, suggesting a similar distribution of visit duration.

Figure 2 presents violin plots for the resulting infection risk probability $\mathbb{P}\{\text{infection in } [0, \infty)\}$ using Equation (14), for each care type and visit length. This probability estimates the total risk of infection to the susceptible patient, who remains in the room long after the infectious HCW has left. We explore removal rates $R \in \{0.381, 1.143, 2.147, 3.870\}$ [m^3 min^{-1}] in Figure 2A–D. We note that for each removal rate, "Check" typically leads to lower infection risk across care types, which is consistent with it normally having the lowest visit duration (Figure 1). The violin plots in Figure 2 have similar shapes to those in Figure 1 for each care type. As expected, increasing removal rates lead to decreasing infection risk regardless of the care type. Increasing removal rates also leads to decreasing variability (i.e., decreasing SD of the corresponding violin plots), suggesting that increasing ventilation can lead to more homogeneous infection risk across different visits (especially within the same care type). However, the removal rates considered in this section are not able to completely mitigate large infection risk episodes represented by outliers corresponding to "IV" and "Rounds" type of care. This suggests that increasing ventilation might not be enough when HCW visits are significantly long, and that additional mitigation strategies, such as using masks, might be especially important during these cases.

Our results above consider total infection risk, which accounts for the risk during the visit (i.e., during $[0, T]$, where $t = 0$ represents the infectious HCW entering into the patient's room, and T is the time that the infector leaves the room), as well as after the visit (i.e., $[T, \infty)$). It is of interest then to analyse the relative importance of each of these two time periods. In Figure 3A, we plot the total probability of infection during $[0, \infty)$ for the susceptible patient in the single-bed room, given by Equation (14), for all 102 observed visit times (without distinguishing care type) and the different removal rates. We also illustrate the contributing probabilities which arise from each of the two time periods; Figure 3B shows the probability of infection during $[0, T]$ (i.e., when the infector is present, Equation (9)), and Figure 3C shows the probability of infection after the infector has left (Equation (12)). In Figure 3D, we plot the relative contribution of the infection risk during the time period after the infector has left on the total infection risk probability (Figure 3A), that is,

$$\frac{P_{total}(T) - P_1(T)}{P_{total}(T)} \times 100, \tag{21}$$

where $P_{total}(T)$ is computed via Equation (14), and $P_1(T)$ via Equation (9).

When analysing the results in Figure 3A, noting that we have used a log scale for the y-axis, we see that increasing ventilation becomes more effective for longer HCW visits. For relatively short time periods (less than 5 min), we observe small overall probabilities of infection almost regardless of

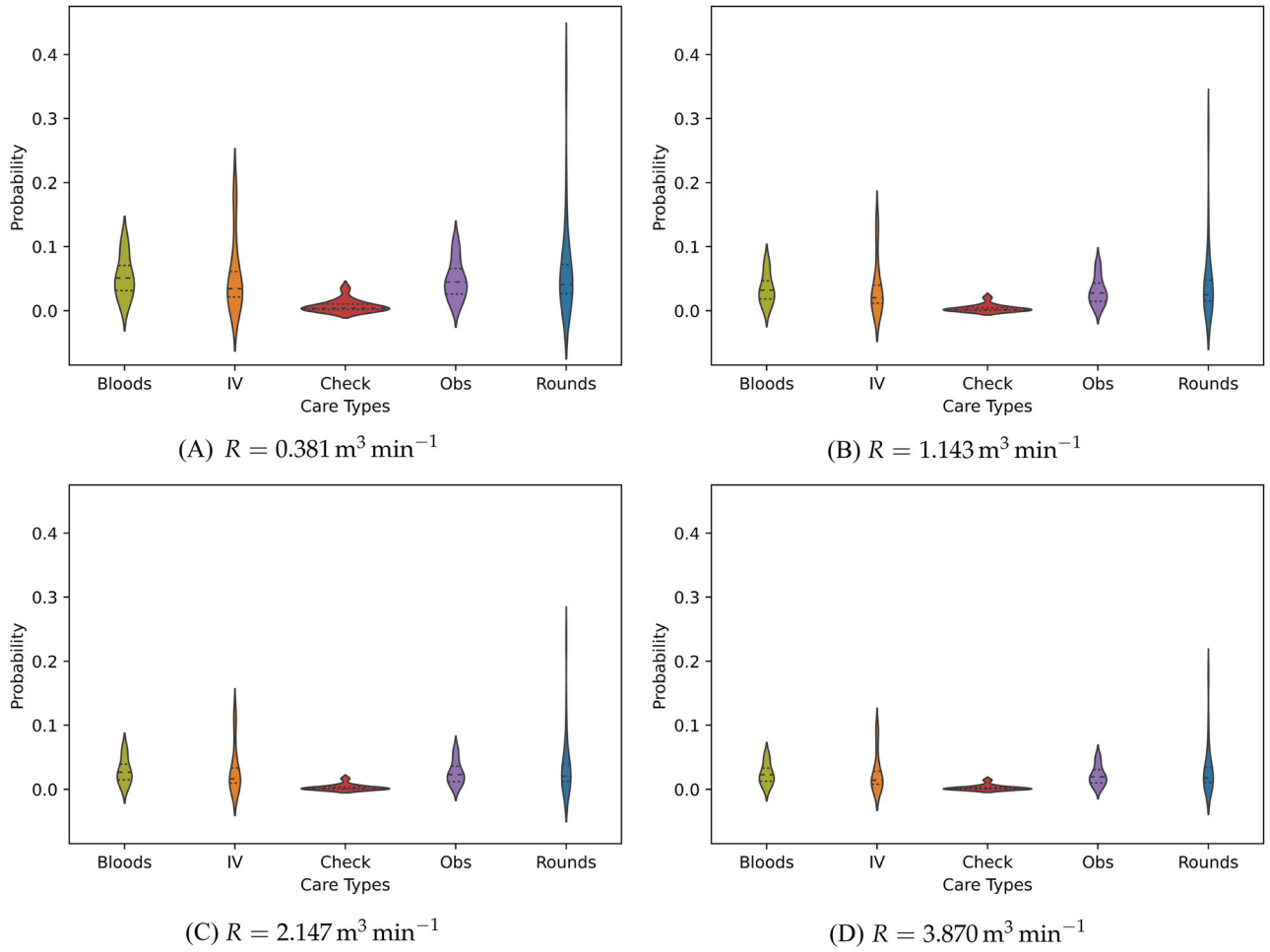

FIGURE 2 Violin plots for the distribution of total risk of infection to the susceptible patient due to the visit of an infectious healthcare worker (HCW) (Equation (14) for $t \to \infty$), across different care types and removal rates: (A) $R = 0.381$ m^3 min^{-1}, (B) $R = 1.143$ m^3 min^{-1}, (C) $R = 2.147$ m^3 min^{-1}, and (D) $R = 3.870$ m^3 min^{-1}.

the removal rate considered. On the other hand, relatively longer times (more than 10 min) lead to significantly larger infection risk probabilities, and the impact of the removal rate becomes more noticeable. This feature is mirrored in Figure 3B, when focusing on probability of infection during the specific time period when the infectious HCW is still in the room.

Importantly, the impact of increasing removal rates is more substantial when focusing on the probability of infection once the infector leaves (see Figure 3C). Despite the infector only being present for a relatively short period of time (up to 20 min), the infection risk in Figure 3C is not negligible, especially for larger values of T and smaller removal rates R. In particular, increasing removal rates have a considerable impact on the risk of infection after the infector leaves the room, with this difference growing with HCW visit length. When the infector is present for longer than 5 min, we see greater benefit in having a higher removal rate. This is directly related to the fact that the amount of pathogen which can accumulate in the air increases with T, and that in this type of scenario the susceptible individual remains in the room long after the infector has left (i.e., we are considering $t \to \infty$ in Equation (14)).

In Figure 3D, we plot the relative contribution to the total probability of infection from the infection risk during the time period after the infectious HCW has left. This is shown for all visit lengths and each of the four removal rates. We note that the removal rates also have a similar effect here, and with increasing removal rate, the percentage contributed from the second period gets smaller. That is to say, the period of time after the infector leaves becomes less important with larger removal rates (and, thus, with increasing ventilation). This is due to the fact that the removal mechanisms will reduce the concentration of airborne pathogen faster, and thus reduce the risk to the susceptible individual still in the room after the infector leaves. It is clear that for short visit times, the infection risk after the HCW leaves is an important factor to consider, with the majority of the infection risk coming from this period, particularly for smaller removal rates and shorter visit times. Overall, we note the nonnegligible contribution to the overall risk of infection that comes from the period when the infector is not present.

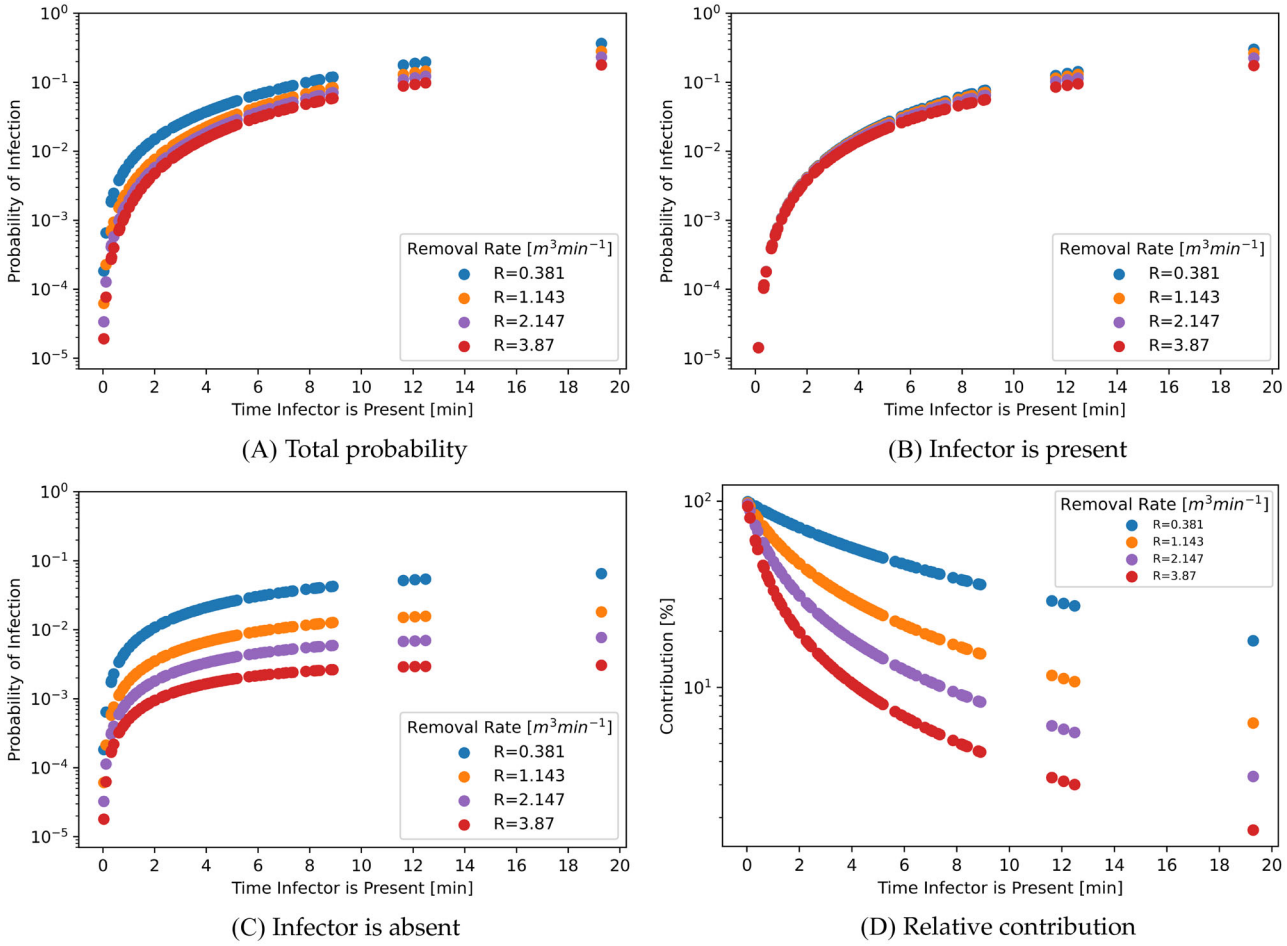

FIGURE 3 Probabilities of infection for various healthcare worker (HCW) visit times, T, and varying removal rates for (A) the total period $[0, \infty)$ (Equation (14)), (B) the period for when the infector is present only, $[0, T]$ (Equation (9)), (C) the period after the infector has left, $[T, \infty)$ (Equation (12)), and (D) Relative contribution towards the total probability of infection from the probability of infection after the infector has left, defined in Equation (21).

A particular feature to highlight across all of the results in Figure 3 is the importance of considering transient effects, especially when modelling infection risk over small time scales. We have illustrated the importance of the period once the infector leaves, when infectious pathogen may remain in the air and contribute to the infection risk for the susceptible patient. This would otherwise be overlooked when using the standard steady-state Wells–Riley model, since in this model the concentration would be considered to be zero as soon as the infector leaves.

4.2 | Case Study 2: Random time spent in different restaurant settings

In this section, we investigate the risk of infection posed by having lunch in different restaurants, and while considering an unknown (random) lunch duration T for different party sizes, leveraging the methodology in Section 3.2.

We consider Erlang distributions to model the duration of lunch, $T \sim Erlang(k\ ,\ \lambda)$, for some parameters (k, λ). In particular, we calibrate these parameters based on data from a study which investigated the impact of social factors on the duration of a meal time, which included different party sizes in a selection of hospitality venues (Bell & Pliner, 2003). This study considered three different venues: a fast-food restaurant, a workplace cafeteria, and a moderately priced restaurant. The party sizes ranged from one person, to 5+ people. Since we focus on analysing infection risk for the group having the meal together, due to an infector being present in this same group, we will consider here party sizes larger than one. For our illustrative results, meal duration for party sizes 5+ in Bell and Pliner (2003) have been used to consider a party size of exactly five individuals. The measured mean meal duration, and its SD, for each party size can be seen for the fast-food restaurant, the workplace cafeteria, and the moderately priced restaurant in Tables 1, 2, and 3, respectively.

As the data for the meal times in Bell and Pliner (2003) are given in the form of the mean and SD, instead of as raw data values, we use these summary statistics to calibrate the Erlang parameters k (shape) and λ (rate). In particular, an

TABLE 1 Observed average meal times, with standard deviation (SD), for various party sizes in a fast-food restaurant (Bell & Pliner, 2003), together with calibrated Erlang distributions to model these.

	Observed		Modelled (Erlang (k , λ))			
Party size	**Mean μ [min]**	**SD σ [min]**	**Mean μ [min]**	**SD σ [min]**	k	λ
2	18.2	6.0	17.8	5.9	9	0.5056
3	18.4	6.8	17.6	6.7	7	0.3979
4	19.7	7.2	18.4	7.0	7	0.3800
5	21.9	5.8	21.5	5.9	14	0.6510

TABLE 2 Observed average meal times, with standard deviation (SD), for various party sizes in a workplace cafeteria (Bell & Pliner, 2003), together with calibrated Erlang distributions to model these.

	Observed		Modelled (Erlang (k , λ))			
Party size	**Mean μ [min]**	**SD σ [min]**	**Mean μ [min]**	**SD σ [min]**	k	λ
2	23.0	7.9	21.7	7.7	8	0.3685
3	33.0	11.3	34.8	11.6	9	0.2584
4	41.1	10.6	41.0	10.6	15	0.3658
5	44.0	14.2	45.8	14.5	10	0.2182

TABLE 3 Observed average meal times, with standard deviation (SD), for various party sizes in a moderately priced restaurant (Bell & Pliner, 2003), together with calibrated Erlang distributions to model these.

	Observed		Modelled (Erlang (k , λ))			
Party size	**Mean μ [min]**	**SD σ [min]**	**Mean μ [min]**	**SD σ [min]**	k	λ
2	44.9	10.8	44.5	10.7	17	0.3849
3	47.2	10.1	47.6	10.1	22	0.4627
4	52.3	8.5	52.5	8.5	38	0.7239
5	58.5	13.1	58.7	13.1	20	0.3409

$Erlang(k, \lambda)$ distribution has mean and SD given as:

$$\mu = \frac{k}{\lambda}, \quad \sigma = \sqrt{\frac{k}{\lambda^2}}. \tag{22}$$

Thus, for each pair of observed (μ, σ) in Tables 1–3, for each party size, the simultaneous equations in Equation (22) are solved to get a corresponding value for k and λ. Since the shape parameter, k, takes integer values, the solution for k was then rounded to the nearest integer. Calibrated parameters are reported in Tables 1–3, and the resulting Erlang distributions are plotted in Figure 4. We note that the rounding method for k leads to a calibrated Erlang which does not precisely match the observed mean (and SD) meal durations, but the relative error is small, so that meal durations are relatively well captured by the resulting Erlang distributions. In particular, the relative error is less than 6.50% for the mean and less than 3.30% for the SD in the fast-food restaurant case, less than 5.65% for the mean and less than 2.85% for the SD in the workplace cafeteria case, and less than 1.64% for the mean and less than 0.83% for the SD in the moderately priced restaurant case, across all party sizes.

In Figure 4, the resulting Erlang distributions can be seen for each party size, for all of the dining locations. Figure 4A shows relatively small variability between the meal duration for different party sizes when dining in a fast-food restaurant, suggesting that, regardless of the number of people, typical mean length remains the same. In the other two cases, there is evidence of more variability between the time spent across different party sizes. These results are rather intuitive to what we would expect to see in everyday life. Typically, fast-food settings are more informal and encourage a quicker dining experience regardless of the number of individuals present, whereas in more formal settings such as the workplace cafe or restaurant, once seated you may be more likely to stay for longer, especially for larger party sizes. Now, we aim to explore the impact of these random meal durations, across different party sizes and hospitality venues, on the infection risk for the group (in terms of number of infections), if an infector is present within it. That is, for a party size N, we

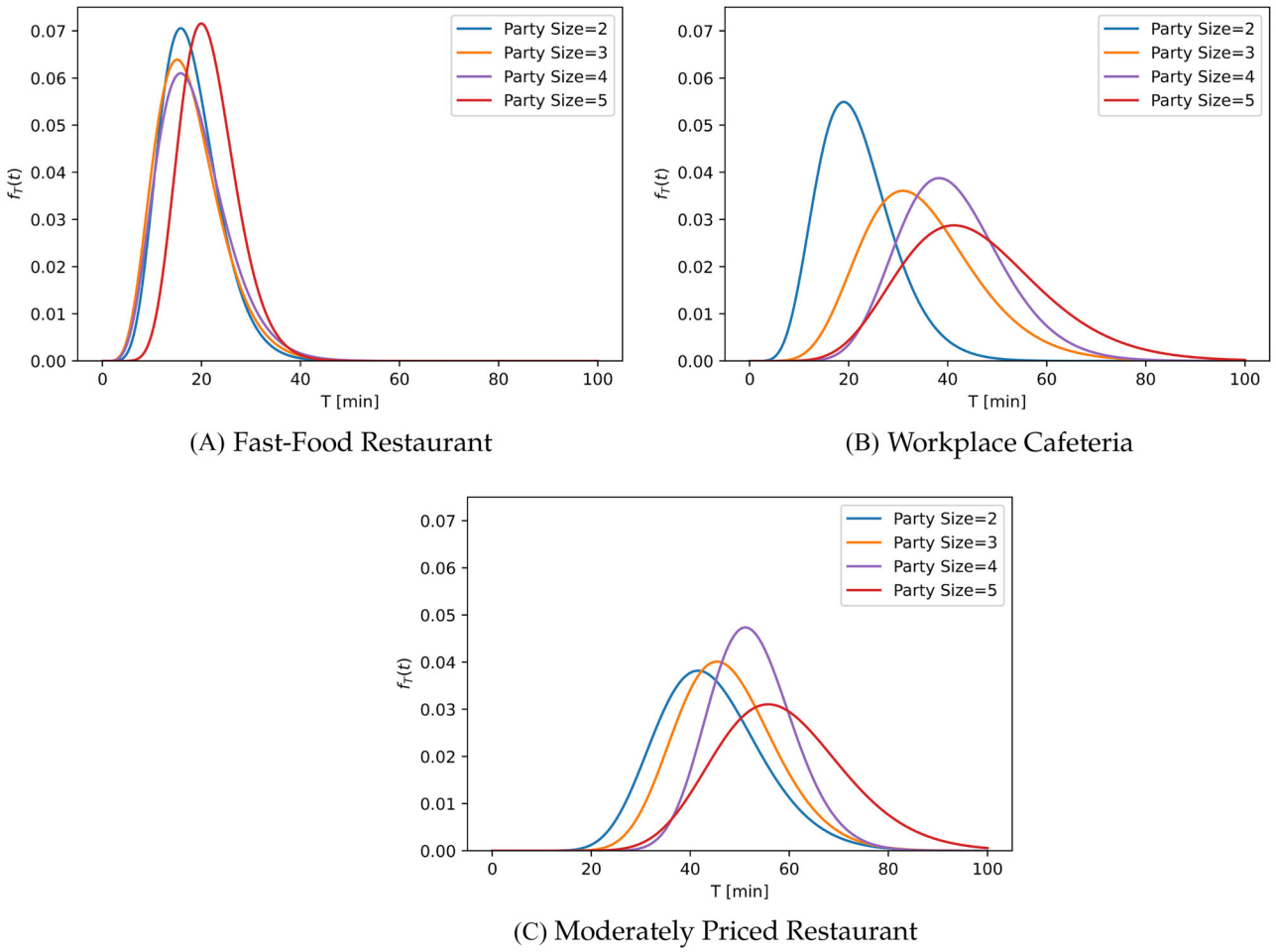

FIGURE 4 Calibrated Erlang density functions for each party size in the (A) fast-food restaurant, (B) workplace cafeteria, and (C) moderately priced restaurant.

set the initial number of infectors to be $I_0 = 1$, and the initial number of susceptible individuals $S_0 = N - I_0$, and do not consider any interaction with other individuals in the venue. We set the pulmonary breathing rate to be $p = 0.01$ m^3 min^{-1} (Noakes & Sleigh, 2009), and the quanta production rate $q =$ 6 quanta min^{-1} as in the previous case study. The volumes of each space are considered to be equal, so that our results focus on analysing the impact of meal duration (which depends on hospitality venue type and party size) as the key factor of interest, while all other parameters/factors remain the same. For illustrative purposes, we consider space dimensions of $3 \times 10 \times 10$ (height [m] $\times$ width [m] $\times$ depth [m]) giving $v =$ 300 m^3. We explore ventilation rates of 3 and 1 ACH, giving extract ventilation $Q = 15$ m^3 min^{-1} and $Q = 5$ m^3 min^{-1}, respectively. For the viral inactivation and removal rates, we chose the midpoint of the ranges proposed in Miller et al. (2021), leading to $r_i = 0.00525$ min^{-1} and $r_d = 0.015$ min^{-1}. Combining these with the extract ventilation rates lead us to consider removal rates $R = 21.075$ m^3 min^{-1} (corresponding to 3 ACH) and $R = 11.075$ m^3 min^{-1} (corresponding to 1 ACH).

Figure 5 shows the probability distribution of the number of infections for each party size and hospitality venue, when considering 3 ACH ventilation rate. The mean number of infections is plotted as a vertical line for each histogram, and these mean values are reported in the legend. Results suggest a relatively low risk for all scenarios, with the most likely outcome being zero infections (highest probability at zero). This is consistent across all three dining settings. In all of the cases, the mean number of exposures is less than 1, with the highest mean being 0.61 in the restaurant case with a party size of 5, as one would expect. In general, larger party sizes lead to higher risk of some infection(s) occurring: the probability of getting exactly one infection during the meal is $20 - 37\%$ for a party size of 5, but less than 12% for a party size of 2. We also note that, in reality, the probability of having an infector present would significantly increase for larger party sizes, which is not explored in these results (where we assume that a single infector is present, regardless of the scenario). Similarly, in reality both the risk of an infector being present and the total number of new infections would also be influenced by the number of other people present in the restaurant. Overall, the predicted number of infections in the moderately

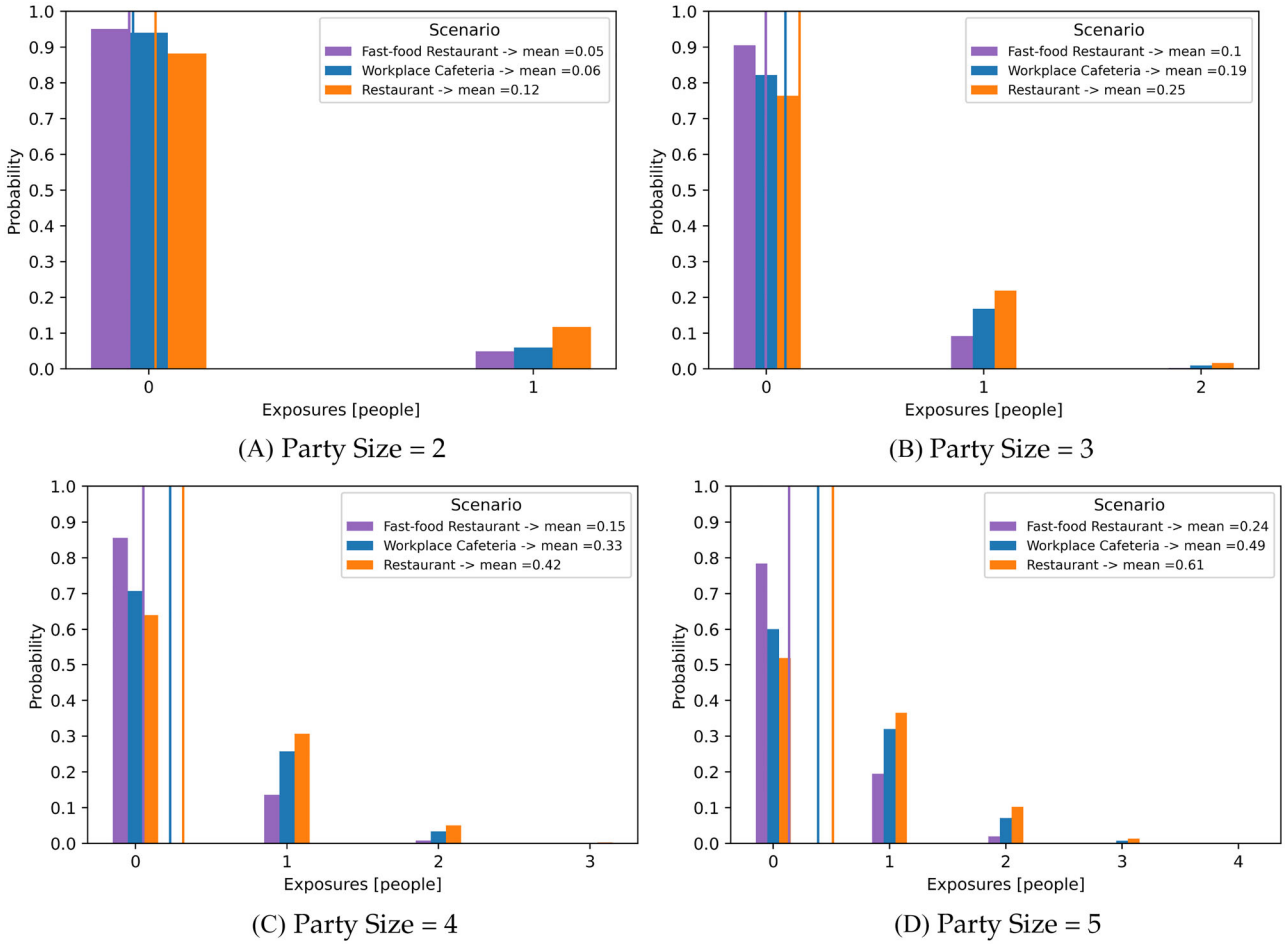

FIGURE 5 The probability of having a given number of exposures for different hospitality venues and party sizes. Volume of $v = 300$ m^3, and ventilation rate of 3 Air Changes per Hour (ACH). Mean numbers of infections are reported in the legend and plotted as vertical lines for each histogram.

priced restaurant is larger than in the workplace cafeteria, this one being larger than in the fast-food venue, which is directly related to the meal durations observed in Figure 4.

Another possibility for the low risk could be due to the good levels of ventilation. It is possible that the ventilation is muting some of the nuances due to the small variations present in the distributions. For additional comparison, the probability of a given number of exposures for the three different restaurant cases can be seen for four different party sizes and a reduced ventilation rate of 1 ACH in Figure 6.

When we reduce the ventilation in these spaces, the differences in risk becomes more noticeable, and we observe a larger probability across all non-zero predicted exposures, with a smaller probability of zero predicted exposures. In particular, the probability of zero exposures in a party size of 5 in a moderately priced restaurant has reduced from over 50% in the 3 ACH case, to less than 30% in the 1 ACH case. This increase in risk is also seen in the predicted mean number of exposures, where the highest mean across all scenarios is now 1.08, compared to 0.61 in the 3 ACH case.

In Figures 5 and 6, the focus is in the population level risk, in terms of the number of predicted exposures during the meal. One can instead look at the risk posed to an individual who is attending the event (per capita probability), and study how this per capita risk is affected by the meal duration. The traditional Wells–Riley model (Equation (1)) offers a per capita probability of an individual becoming infected in a given scenario, for a given time duration T. In Section 3.2.2, we introduced the per capita probability for an individual when the exact time is not known, but is instead represented through an Erlang distribution (Equation (19)). Predictions given by these Equations (1) and (19) can be now compared when considering the meal duration times for the different hospitality venues. In this analysis, we consider the room volume $v = 300$ m^3, pulmonary breathing rate $p = 0.01$ m^3 min^{-1}, quanta production rate $q = 6$ quanta min^{-1}, and removal rate $R = 11.075$ m^3 min^{-1}.

To better understand the difference between per capita probabilities for known and unknown exposure times, we take a sample of 100 random time durations from each Erlang distribution across each dining setting and party size in Figure 4. For each sampled meal duration, Equation (1) provides the per capita infection risk probability for that specific duration of the meal. Thus, Equation (1) applied to 100 sampled times (from each Erlang distribution in Figure 4, corresponding to each hospitality venue and party size) gives 100 values

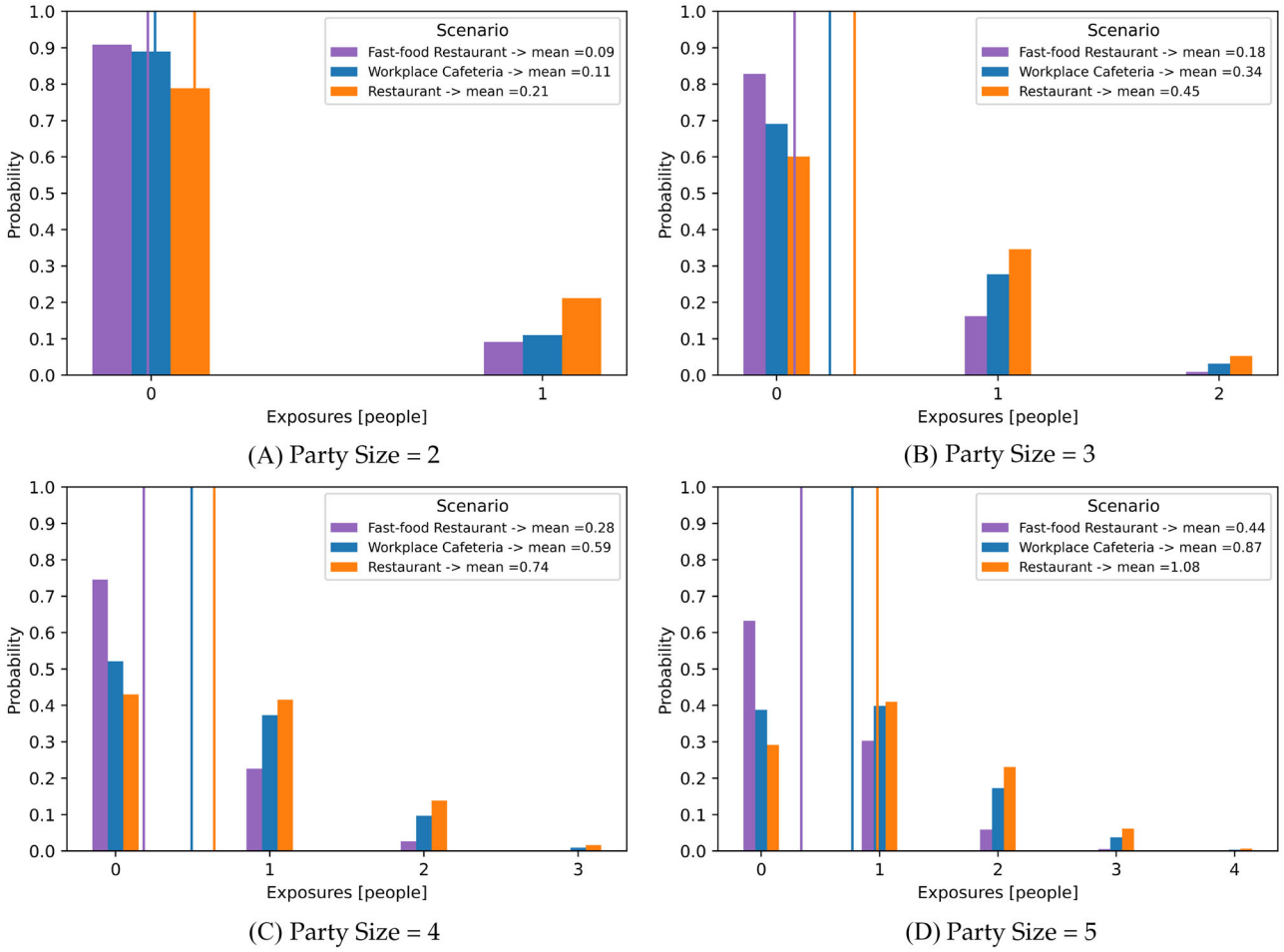

FIGURE 6 The probability of having a given number of exposures for different hospitality venues and party sizes. Volume of $v = 300$ m^3, and ventilation rate of 1 Air Changes per Hour (ACH). Mean numbers of infections are reported in the legend and plotted as vertical lines for each histogram.

of $P(T)$, the per capita probability of infection estimated by the traditional Wells–Riley model. These 100 values of $P(T)$ are plotted as violin plots in Figure 7. On the other hand, each Erlang distribution in Figure 4 leads to a single overall prediction $P(T)$ from Equation (19), which estimates the per capita infection risk probability accounting for the uncertainty in meal duration encoded by the corresponding Erlang distribution. This single prediction of $P(T)$ is superimposed on each of the corresponding violin plots, as a black dot.

From Figure 7, we see a similar pattern to that observed across the probability distributions for the number of exposures in Figures 5 and 6. In particular, the scenarios with the typically shorter meal time durations experience the lowest per capita infection risk probability, that is, the fast-food restaurant has the lowest per capita risk. This is consistent also across the party sizes with larger groups typically experiencing a high individual per capita risk of infection, as expected. We can also note how the per capita probability obtained via Equation (19), which estimates infection risk while accounting for the uncertainty in the meal duration time encoded by the Erlang distribution, is able to summarise all of the distribution from the traditional Wells–Riley per capita (violin plot) into a single point (black dot), across Figure 7A–C. Thus, by using our methodology in Section 3, one can estimate infection risk in terms of $P(T)$ while accounting for the uncertainty in the time duration of the indoor interaction.

An interesting feature to highlight on these plots is the change in the per capita infection risk probability with party size. It is expected that the per capita infection risk probability may vary across the different dining settings (i.e., that the individual infection risk depends on the venue/indoor setting). However, one would expect that the individual infection risk does not depend on party size, if the number of infectors present is fixed to $I_0 = 1$; that is, Equations (1) and (19) are both independent of the total population size, N. Through Figure 7, we see that as the party size increases, so does the per capita probability of infection. This is due to the fact that the Erlang distributions across the party sizes and settings account for the fact that individuals are likely to spend longer in a particular venue if there are more people dining (Figure 4). This highlights the importance of social and behavioural factors in determining infection risk, which is often overlooked when using traditional QMRA techniques such as the Wells–Riley model, but which can be better

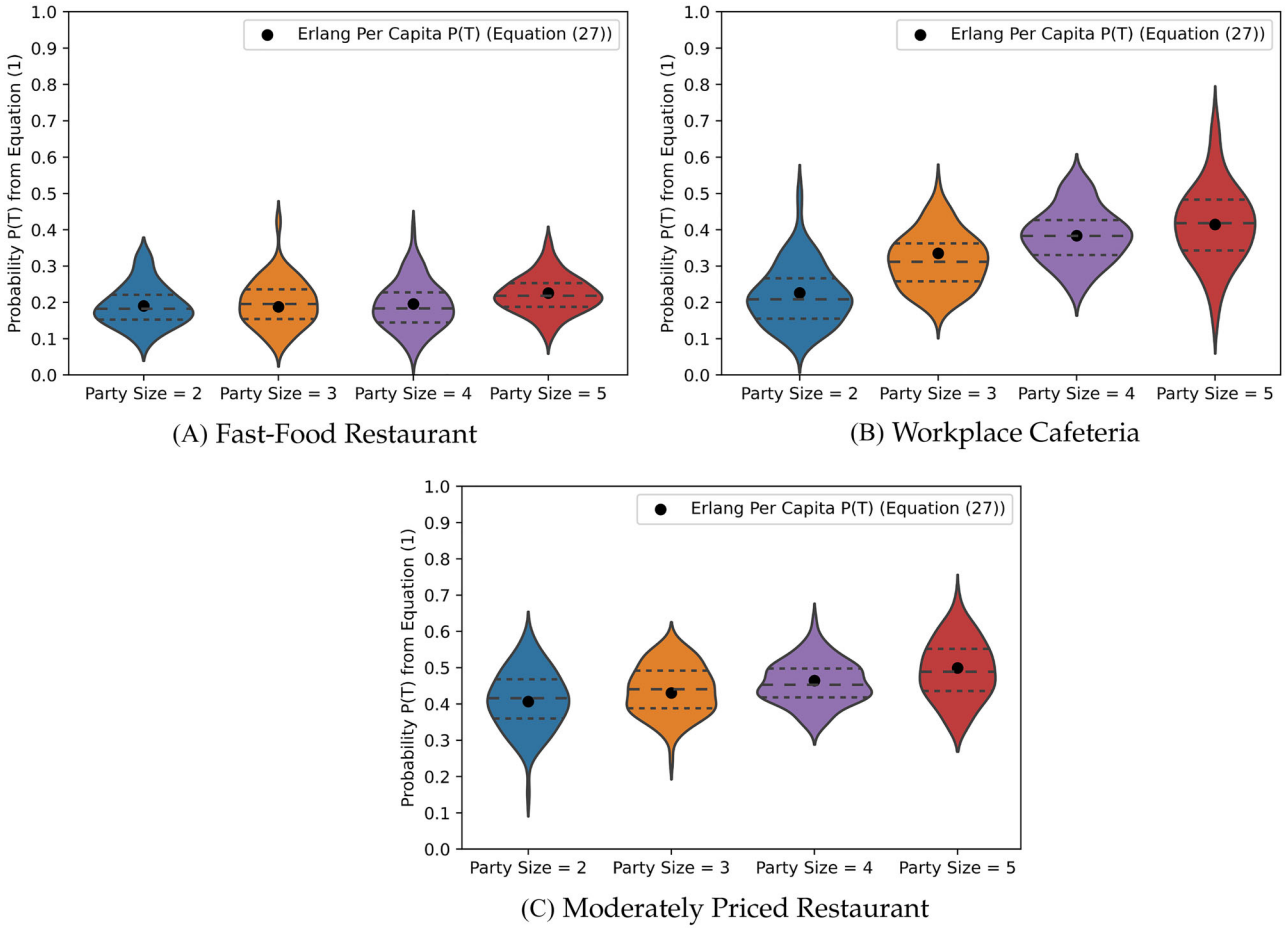

FIGURE 7 Violin plots for the distribution of per capita $P(T)$ for the traditional Wells–Riley model (Equation (1)) with duration T sampled from the Erlang distributions (Figure 4), along with the per capita $P(T)$ for unknown T using an Erlang distribution (Equation (19)) for each party size in the (A) fast-food restaurant, (B) workplace cafeteria, and (C) moderately priced restaurant.

explored via our newly proposed methodology by integrating uncertainty in some model parameters.

5 | DISCUSSION

In this study, we have extended the traditional Wells–Riley framework, to assess risk in three situations: where the infector leaves the room but the susceptible remains; the duration of the indoor interaction is random/unknown; or there exists heterogeneity in infectivity across infectious individuals. We have been able to compute the per capita probability of infection for each susceptible individual in these scenarios, and to estimate the probability distribution for the number of exposures/infections.

When using the traditional Wells–Riley model under the steady-state assumption, the risk of infection is only non-zero when the infector is present. However, the suspension of infectious aerosols means that the risk remains possibly long after the infector leaves. By using a transient concentration of pathogen solution, we are able to model the gradual growth and decay of the airborne pathogen concentration, and as a result, the accompanying infection risk. This allows one to compute the probability of infection whilst the infector is present, and once they leave, giving a holistic risk assessment to a susceptible individual, acknowledging that the risk of infection stretches beyond the presence of an infector. This aids our understanding of what governs infection risk, and how this may differ across the two time periods. We provide the analytical solution for this probability for a single zone where the susceptible individual is present for a particular time period (Equation (13)) or for long after the infector leaves (Equation (14)).

The applicability of these results has been shown via a first case study (Section 4.1), where we leveraged real-life HCW visit times data, and estimate risk experienced by the susceptible patient before and after the infector (HCW) leaves the space. We show that post-departure risk is nonnegligible, and becomes even more important for shorter visit times, or under smaller removal rates. In the results, we assume that the susceptible individual is present for $t \to \infty$ after the infector leaves. This is more relevant for hospitals wards, such as an adult respiratory ward where patients are often admitted as inpatients for longer periods. It is possible that the percentage contribution from the period after the infector leaves is much less if the susceptible individual also leaves shortly after.

It is possible that we see such a large contribution from the second period due to small time scales used for the presence of the infector. Even though these are realistic times (taken from measured data), if we considered alternative scenarios where the infector was present for longer periods, we may see a smaller overall contribution from the second period. This would be scenarios where the first period displays steady-state-like behaviour. We can see hints of this in Figure 3D, for example, in the case where we have the highest removal rate ($R = 3.97\ \mathrm{m}^3\ \mathrm{min}^{-1}$) and the longest visit time (19.28 min). This will most likely be the closest to steady-state behaviour and thus, the contribution from the second period is very small (1.71%), and could be seen as negligible. Simultaneously, this further illustrates the importance of including the period after the infector leaves for shorter visits, and poor removal rates as its relevance can vary depending on the scenario parameters.

An example of the stochasticity that exists in real life, which is overlooked by the traditional Wells–Riley model, is not knowing the duration of an outbreak, or the time that the infector is present in the space. In these scenarios, it may be more suitable to consider a random duration, in the form of a distribution. This way, we can estimate infection risk while accounting for realistic time distributions that are representative of typical behaviour experienced in the corresponding setting. We explored this methodology in Section 3.2, providing explicit solutions for steady-state concentration in a single room for two well-known distributions commonly used to model waiting times; Exponential and Erlang. A limitation of this methodology is that we use the steady-state solution for the concentration of airborne pathogen in the model, so future work should be devoted to generalise these results for a transient concentration of pathogen in the air.

In Case Study 2 (Section 4.2), we explore the possibility of not knowing the duration of the indoor interaction. In particular, we consider infection risk during lunch in different dining settings, and for different party sizes, by fitting an Erlang distribution to measured meal duration data. The resulting Erlang distributions varied across the scenarios with long meal durations experienced more so in the workplace cafeteria or in the moderately priced restaurant, where as the fast-food restaurant had much smaller meal durations and less variation across the party sizes. This pattern was consistent with the infection risk, where settings of typically longer meal durations lead to a higher risk of infection, as does the increase in the party size. We observed how decreasing the removal rate increased the risk of infection across all scenarios, but the distribution of the risk of possible predicted exposures was still dominated by the distributions of meal durations.

When analysing the per capita probability of infection (Figure 7), we saw the importance of considering social factors when assessing infection risk. Intuitively, the per capita probability of infection should be unaffected by change in population size under the assumption that a single infector is present, but we saw how the typical time spent in each scenario is highly influenced by social setting and party size, leading to variations in the per capita probability that are often overlooked when the exact exposure time is known. The per capita probability of infection for an Erlang-distributed time duration, represented as a single point in Figure 7, is closely aligned to the mean value of the standard Wells–Riley per capita probability of infection across 100 time durations sampled from the corresponding Erlang distribution. Thus, our single estimate is able to quantify infection risk while accounting for the uncertainty in the duration of the indoor interaction. This demonstrates the importance of incorporating uncertainty in the duration of the indoor interaction when estimating infection risk, rather than using a single estimate of this duration.

6 | CONCLUSION

In this work, we have presented new mathematical formulations to assess infection risk that build on the existing Wells–Riley framework. Through the inclusion of stochasticity, we are able to take the existing methodology presented in the Wells–Riley model, and use this to derive the probability distribution of the number of exposures/infections. Moreover, we extended these results in order to assess risk of infection in the period after the infector leaves, and when considering an unknown random outbreak duration, or quanta generation rate, in terms of probability distributions. We illustrated this methodology through two case studies: one case study that shows that the period after the infector leaves can contribute nonnegligible risk, and another case study which illustrates the risk for various restaurant scenarios where the exact duration of each outbreak is unknown. We have successfully developed new, explicit formulas for the infection probability in these scenarios, increasing the applicability of the Wells–Riley framework, with a better representation of the characteristics we face when trying to model real-life outbreaks.

AUTHOR CONTRIBUTIONS

Alexander J. Edwards: Writing—review and editing; writing—original draft; visualization; validation; software; resources; project administration; methodology; investigation; funding acquisition; formal analysis; data curation; conceptualization. **Marco-Felipe King**: Writing—review and editing; conceptualization; funding acquisition; supervision. **Catherine J. Noakes**: Writing—review and editing; conceptualization; funding acquisition; supervision. **Daniel Peckham**: Writing—review and editing; conceptualization; funding acquisition; supervision. **Martín López-García**: Project Administration; writing—review and editing; conceptualization; funding acquisition; supervision. Compliance with ethical standards.

ACKNOWLEDGMENTS

AJE is funded by the Engineering and Physical Sciences Research Council (EPSRC) Centre for Doctoral Training (CDT) in Fluid Dynamics [Grant Number EP/S022732/1].

CONFLICT OF INTEREST STATEMENT

During the COVID-19 pandemic C. J. Noakes was a participant in the UK Scientific Advisory Group for Emergencies (SAGE) and co-chaired the SAGE Environment and Modelling Sub-Group. The other authors declare no conflicts of interest.

ETHICAL APPROVAL

This study does not contain any studies with human or animal subjects performed by any of the authors.

DATA AVAILABILITY STATEMENT

The code and data are available at https://github.com/scaje/stochastic_WR_paper_AJE.git.

ORCID

Alexander J. Edwards https://orcid.org/0000-0001-5313-8490

REFERENCES

Allen, L. J. (2010). *An introduction to stochastic processes with applications to biology*. CRC press.

Arino, J., Sun, C., & Yang, W. (2016). Revisiting a two-patch SIS model with infection during transport. *Mathematical Medicine and Biology*, *33*(1), 29–55.

Beggs, C. B., Shepherd, S. J., & Kerr, K. G. (2010). Potential for airborne transmission of infection in the waiting areas of healthcare premises: Stochastic analysis using a Monte Carlo model. *BMC Infectious Diseases*, *10*(1), 1–8.

Bell, R., & Pliner, P. L. (2003). Time to eat: The relationship between the number of people eating and meal duration in three lunch settings. *Appetite*, *41*(2), 215–218.

Boonmeemapasuk, W., & Pochai, N. (2022). A risk model of airborne transmission and vaccine efficacy in an outpatient room with a ventilation system. *Engineering Letters*, *30*(2), 644–651.

Bueno de Mesquita, P. J., Noakes, C. J., & Milton, D. K. (2020). Quantitative aerobiologic analysis of an influenza human challenge-transmission trial. *Indoor Air*, *30*(6), 1189–1198.

Buonanno, G., Morawska, L., & Stabile, L. (2020). Quantitative assessment of the risk of airborne transmission of SARS-CoV-2 infection: Prospective and retrospective applications. *Environment International*, *145*, 106112.

Buonanno, G., Stabile, L., & Morawska, L. (2020). Estimation of airborne viral emission: Quanta emission rate of SARS-CoV-2 for infection risk assessment. *Environment International*, *141*, 105794.

Burridge, H. C., Fan, S., Jones, R. L., Noakes, C. J., & Linden, P. (2022). Predictive and retrospective modelling of airborne infection risk using monitored carbon dioxide. *Indoor and Built Environment*, *31*(5), 1363–1380.

Cammarata, A., & Cammarata, G. (2021). Dynamic assessment of the risk of airborne viral infection. *Indoor Air*, *31*(6), 1759–1775.

Dai, H., & Zhao, B. (2020). Association of the infection probability of COVID-19 with ventilation rates in confined spaces. In *Building simulation* (Vol. *13*, pp. 1321–1327). Springer.

Ding, S., Lee, J. S., Mohamed, M. A., & Ng, B. F. (2022). Infection risk of SARS-CoV-2 in a dining setting: Deposited droplets and aerosols. *Building and Environment*, *213*, 108888.

Edwards, A. J., Benson, L., Guo, Z., López-García, M., Noakes, C. J., Peckham, D., & King, M.-F. (2023). A mathematical model for assessing transient airborne infection risks in a multi-zone hospital ward. *Building and Environment*, *238*, 110344.

Eichler, N., Thornley, C., Swadi, T., Devine, T., McElnay, C., Sherwood, J., Brunton, C., Williamson, F., Freeman, J., Berger, S., Ren, X., Storey, M., de Ligt, J., & Geoghegan, J. L. (2021). Transmission of severe acute respiratory syndrome coronavirus 2 during border quarantine and air travel, New Zealand (Aotearoa). *Emerging Infectious Diseases*, *27*(5), 1274–1278.

Fantozzi, F., Lamberti, G., Leccese, F., & Salvadori, G. (2022). Monitoring CO_2 concentration to control the infection probability due to airborne transmission in naturally ventilated university classrooms. *Architectural Science Review*, *65*(2), 306–318.

Fennelly, K. P., & Nardell, E. A. (1998). The relative efficacy of respirators and room ventilation in preventing occupational tuberculosis. *Infection Control & Hospital Epidemiology*, *19*(10), 754–759.

Fisk, W. J., Sepppanen, O., Faulkner, D., & Huang, J. (2004). Economic benefits of an economizer system: Energy savings and reduced sick leave. Technical Report, LBNL-54475, Lawrence Berkeley National Laboratory.

Franchimon, F., Pernot, C. E., Khoury, E., & Bronswijk, J. (2008). The feasibility of indoor humidity control against avian influenza. *Proceedings of the 11th International Conference on Indoor Air Quality and Climate, Indoor Air*.

Gammaitoni, L., & Nucci, M. C. (1997). Using a mathematical model to evaluate the efficacy of TB control measures. *Emerging Infectious Diseases*, *3*(3), 335–342.

Gao, N., Niu, J., Perino, M., & Heiselberg, P. (2008). The airborne transmission of infection between flats in high-rise residential buildings: Tracer gas simulation. *Building and Environment*, *43*(11), 1805–1817.

Guo, Y., Qian, H., Sun, Z., Cao, J., Liu, F., Luo, X., Ling, R., Weschler, L. B., Mo, J., & Zhang, Y. (2021). Assessing and controlling infection risk with Wells-Riley model and spatial flow impact factor (SFIF). *Sustainable Cities and Society*, *67*, 102719.

Haas, C. N. (2020). Quantitative microbial risk assessment and molecular biology: Paths to integration. *Environmental Science & Technology*, *54*(14), 8539–8546.

Jung, J., Lee, J., Jo, S., Bae, S., Kim, J. Y., Cha, H. H., Lim, J. Y., Kwak, S. H., Hong, M. J., Kin, E. O., Bae, J. Y., Kang, C., Sung, M., Park, M. S., & Kim, S. H. (2021). Nosocomial outbreak of COVID-19 in a hematologic ward. *Infection & Cheomotherapy*, *53*(2), 332–341.

Kidd, M., Richter, A., Best, A., Cumley, N., Mirza, J., Percival, B., Mayhew, M., Megram, O., Ashford, F., White, T., Moles-Garcia, E., Crawford, L., Bosworth, A., Atabani, S. F., Plant, T., & McNally, A. (2021). S-variant SARS-CoV-2 lineage B1.1.7 is associated with significantly higher viral load in samples tested by TaqPath polymerase chain reaction. *Journal of Infectious Diseases*, *223*(10), 1666–1670.

King, K. G., Delclos, G. L., Brown, E. L., Emery, S. T., Yamal, J. M., & Emery, R. J. (2021). An assessment of outpatient clinic room ventilation systems and possible relationship to disease transmission. *American Journal of Infection Control*, *49*(6), 808–812.

King, M., Wilson, A., López-García, M., Proctor, J., Peckham, D., Clifton, I., Dancer, S., & Noakes, C. (2021). Why is mock care not a good proxy for predicting hand contamination during patient care? *Journal of Hospital Infection*, *109*, 44–51.

Kitajima, M., Ahmed, W., Bibby, K., Carducci, A., Gerba, C. P., Hamilton, K. A., Haramoto, E., & Rose, J. B. (2020). SARS-CoV-2 in wastewater: State of the knowledge and research needs. *Science of the Total Environment*, *739*, 139076.

Knibbs, L. D., Morawska, L., Bell, S. C., & Grzybowski, P. (2011). Room ventilation and the risk of airborne infection transmission in 3 health care settings within a large teaching hospital. *American Journal of Infection Control*, *39*(10), 866–872.

Ko, G., Burge, H. A., Nardell, E. A., & Thompson, K. M. (2001). Estimation of tuberculosis risk and incidence under upper room ultraviolet germicidal irradiation in a waiting room in a hypothetical scenario. *Risk Analysis*, *21*(4), 657–674.

Ko, G., Thompson, K. M., & Nardell, E. A. (2004). Estimation of tuberculosis risk on a commercial airliner. *Risk Analysis: An International Journal*, *24*(2), 379–388.

Lau, Z., Griffiths, I. M., English, A., & Kaouri, K. (2022). Predicting the spatio-temporal infection risk in indoor spaces using an efficient airborne transmission model. *Proceedings of the Royal Society A*, *478*(2259), 20210383.

Li, H., Shankar, S. N., Witanachchi, C. T., Lednicky, J. A., Loeb, J. C., Alam, M. M., Fan, Z. H., Mohamed, K., Eiguren-Fernandez, A., & Wu, C.-

Y. (2021). Environmental surveillance and transmission risk assessments for SARS-CoV-2 in a fitness center. *Aerosol and Air Quality Research*, *21*(11), 210106.

Li, J., Cheng, Z., Zhang, Y., Mao, N., Guo, S., Wang, Q., Zhao, L., & Long, E. (2021). Evaluation of infection risk for SARS-CoV-2 transmission on university campuses. *Science and Technology for the Built Environment*, *27*(9), 1165–1180.

Lloyd-Smith, J. O., Schreiber, S. J., Kopp, P. E., & Getz, W. M. (2005). Superspreading and the effect of individual variation on disease emergence. *Nature*, *438*(7066), 355–359.

López-García, M., King, M. F., & Noakes, C. J. (2019). A multicompartment SIS stochastic model with zonal ventilation for the spread of nosocomial infections: Detection, outbreak management, and infection control. *Risk Analysis*, *39*(8), 1825–1842.

Ma, J., Qi, X., Chen, H., Li, X., Zhang, Z., Wang, H., Sun, L., Zhang, L., Guo, J., Morawska, L., Grinshpun, S. A., Biswas, P., Flagan, R. C., & Yao, M. (2021). Coronavirus disease 2019 patients in earlier stages exhaled millions of severe acute respiratory syndrome coronavirus 2 per hour. *Clinical Infectious Diseases*, *72*(10), e652–e654.

Mikszewski, A., Stabile, L., Buonanno, G., & Morawska, L. (2022). The airborne contagiousness of respiratory viruses: a comparative analysis and implications for mitigation. *Geoscience Frontiers*, *13*(6), 101285.

Miller, D., King, M.-F., Nally, J., Drodge, J. R., Reeves, G. I., Bate, A. M., Cooper, H., Dalrymple, U., Hall, I., López-García, M., Parker, S. T., & Noakes, C. J. (2022). Modeling the factors that influence exposure to SARS-CoV-2 on a subway train carriage. *Indoor Air*, *32*(2), e12976.

Miller, S. L., Nazaroff, W. W., Jimenez, J. L., Boerstra, A., Buonanno, G., Dancer, S. J., Kurnitski, J., Marr, L. C., Morawska, L., & Noakes, C. (2021). Transmission of SARS-CoV-2 by inhalation of respiratory aerosol in the Skagit Valley Chorale superspreading event. *Indoor Air*, *31*(2), 314–323.

Milton, D. K., Fabian, M. P., Cowling, B. J., Grantham, M. L., & McDevitt, J. J. (2013). Influenza virus aerosols in human exhaled breath: Particle size, culturability, and effect of surgical masks. *PLoS Pathogens*, *9*(3), e1003205.

Nazaroff, W. W., Nicas, M., & Miller, S. L. (1998). Framework for evaluating measures to control nosocomial tuberculosis transmission. *Indoor Air*, *8*(4), 205–218.

Nicas, M. (1996). An analytical framework for relating dose, risk, and incidence: An application to occupational tuberculosis infection. *Risk Analysis*, *16*(4), 527–538.

Noakes, C. J., & Sleigh, P. A. (2009). Mathematical models for assessing the role of airflow on the risk of airborne infection in hospital wards. *Journal of the Royal Society Interface*, *6*(Suppl 6), S791–800.

Pantelic, J., & Tham, K. W. (2012). Assessment of the mixing air delivery system ability to protect occupants from the airborne infectious disease transmission using Wells–Riley approach. *HVAC&R Research*, *18*(4), 562–574.

Pavilonis, B., Ierardi, A. M., Levine, L., Mirer, F., & Kelvin, E. A. (2021). Estimating aerosol transmission risk of SARS-CoV-2 in New York city public schools during reopening. *Environmental Research*, *195*, 110805.

Pérez-Rodríguez, F., Valero, A., Carrasco, E., García, R. M., & Zurera, G. (2008). Understanding and modelling bacterial transfer to foods: A review. *Trends in Food Science & Technology*, *19*(3), 131–144.

Qian, H., Li, Y., Nielsen, P. V., & Huang, X. (2009). Spatial distribution of infection risk of SARS transmission in a hospital ward. *Building and Environment*, *44*(8), 1651–1658.

Riley, E., Murphy, G., & Riley, R. (1978). Airborne spread of measles in a suburban elementary school. *American Journal of Epidemiology*, *107*(5), 421–432.

Rocha-Melogno, L., Crank, K., Bergin, M. H., Gray, G. C., Bibby, K., & Deshusses, M. A. (2023). Quantitative risk assessment of COVID-19 aerosol transmission indoors: A mechanistic stochastic web application. *Environmental Technology*, *44*(9), 1201–1212.

Rowe, B. R., Canosa, A., Meslem, A., & Rowe, F. (2022). Increased airborne transmission of COVID-19 with new variants, implications for health policies. *Building and Environment*, *219*, 109132.

Rudnick, S., & Milton, D. (2003). Risk of indoor airborne infection transmission estimated from carbon dioxide concentration. *Indoor Air*, *13*(3), 237–245.

SAGE. (2021). *EMG/Transmission Group/SPI-B: COVID-19 transmission in hotels and managed quarantine facilities (MQFs)*. https://www.gov.uk/government/publications/emg-covid-19-transmission-in-hotels-and-managed-quarantine-facilities-mqfs-9-september-2021

Shang, Y., Dong, J., Tian, L., He, F., & Tu, J. (2022). An improved numerical model for epidemic transmission and infection risks assessment in indoor environment. *Journal of Aerosol Science*, *162*, 105943.

Soller, J. A., Schoen, M. E., Bartrand, T., Ravenscroft, J. E., & Ashbolt, N. J. (2010). Estimated human health risks from exposure to recreational waters impacted by human and non-human sources of faecal contamination. *Water Research*, *44*(16), 4674–4691.

Sun, C., & Zhai, Z. (2020). The efficacy of social distance and ventilation effectiveness in preventing COVID-19 transmission. *Sustainable Cities and Society*, *62*, 102390.

Sze To, G. N., & Chao, C. Y. H. (2010). Review and comparison between the Wells–Riley and dose-response approaches to risk assessment of infectious respiratory diseases. *Indoor Air*, *20*(1), 2–16.

Timpitak, W., & Pochai, N. (2022). A mathematical model of risk assessment on airborne infection in a room with an outlet ventilation system. *Engineering Letters*, *30*(2), 898–903.

Tung, Y.-C., & Hu, S.-C. (2008). Infection risk of indoor airborne transmission of diseases in multiple spaces. *Architectural Science Review*, *51*(1), 14–20.

Wang, C. C., Prather, K. A., Sznitman, J., Jimenez, J. L., Lakdawala, S. S., Tufekci, Z., & Marr, L. C. (2021). Airborne transmission of respiratory viruses. *Science*, *373*(6558), eabd9149.

Wang, X., Ma, S., Zhao, B., Deng, G., She, H., Xu, K., Hao, L., Deng, Y., Li, Q., Yu, Z., & Zhu, X. (2023). Correlations between the viral loads and symptoms in the SARS-CoV-2-infected patients. *MedComm*, *4*(4), e324.

Wang, Z., Galea, E. R., Grandison, A., Ewer, J., & Jia, F. (2022). A coupled computational fluid dynamics and Wells-Riley model to predict COVID-19 infection probability for passengers on long-distance trains. *Safety Science*, *147*, 105572.

Wells, W. F. (1955). *Airborne contagion and air hygiene: An ecological study of droplet infections*. Harvard University Press, Cambirdge, MA.

Wilson, A. M., Weir, M. H., Bloomfield, S. F., Scott, E. A., & Reynolds, K. A. (2021). Modeling COVID-19 infection risks for a single hand-to-fomite scenario and potential risk reductions offered by surface disinfection. *American Journal of Infection Control*, *49*(6), 846–848.

Wood, S. G., Craske, J., & Burridge, H. C. (2023). Relating quanta conservation and compartmental epidemiological models of airborne disease outbreaks in buildings. *Scientific Reports*, *13*(1), 17335.

Zemouri, C., Awad, S., Volgenant, C., Crielaard, W., Laheij, A., & De Soet, J. (2020). Modeling of the transmission of coronaviruses, measles virus, influenza virus, *Mycobacterium tuberculosis*, and *Legionella pneumophila* in dental clinics. *Journal of Dental Research*, *99*(10), 1192–1198.

Zhao, X., Liu, S., Yin, Y., Zhang, T., & Chen, Q. (2022). Airborne transmission of COVID-19 virus in enclosed spaces: An overview of research methods. *Indoor Air*, *32*(6), e13056.

How to cite this article: Edwards, A. J., King, M.-F., Noakes, C. J., Peckham, D., & López-García, M. (2024). The Wells–Riley model revisited: Randomness, heterogeneity, and transient behaviours. *Risk Analysis*, 1–22. https://doi.org/10.1111/risa.14295

APPENDIX A

Probability distribution of the number of exposures in $[0, T]$ in the Wells–Riley model under steady-state conditions.

We consider here the stochastic version of the standard Wells–Riley model under steady-state concentration of pathogen in the air. In particular, we consider $N = S_0 + I_0$ individuals in an indoor setting during $[0, T]$. The number of infectors over time remains constant, I_0, and the interest is in analysing the number of exposures that occur, $E(T)$. To do this, we consider the random variable $S(t)$, representing the number of susceptible individuals at time t. According to Equation (2), one considers that each susceptible individual becomes exposed, independently of each other, at rate pC^*, with $C^* = \frac{qI_0}{Q}$. Thus, the variable $S(t)$ follows a pure death process (see Allen, 2010, Section 6.4.2) taking decreasing values in $\{S_0, S_0 - 1, \dots, 2, 1, 0\}$, with initial condition $S(0) = S_0$.

The number of exposures in $[0, T]$ is a random variable which can be analysed in terms of the number of susceptibles, since $E(T) = S_0 - S(T)$. The number of susceptible individuals at any given time, $S(t)$, can be studied in terms of the probabilities $p_n(t) = \mathbb{P}\{S(t) = n\}$, $n \in \{0, 1, \dots, S_0\}$, which obey the Forward Kolmogorov equations (Allen, 2010, Section 5.6), typically referred to as the *master equation* of the stochastic process,

$$\frac{dp_n(t)}{dt} = pC^*(n+1)p_{n+1}(t) - pC^*np_n(t), \quad n \in \{0, 1, \dots, S_0 - 1\},$$
$$\frac{dp_{S_0}(t)}{dt} = -pC^*S_0p_{S_0}(t).$$

These differential equations can be solved in $[0, T]$ using probability generating function techniques (see Allen, 2010, Section 6.4.2), leading to

$$p_n(T) = \mathbb{P}\{S(T) = n\} = \binom{S_0}{n} e^{-npC^*T}\left(1 - e^{-pC^*T}\right)^{S_0-n}, \quad n \in \{0, 1, \dots, S_0\}.$$

Finally, and since $E(T) = S_0 - S(T)$ so that $\mathbb{P}\{S(T) = n\} = \mathbb{P}\{E(T) = S_0 - n\}$, and $C^* = \frac{qI_0}{Q}$, one gets

$$\mathbb{P}\{E(T) = n\} = \binom{S_0}{n}\left(1 - e^{-\frac{pqI_0}{Q}T}\right)^n\left(e^{-\frac{pqI_0}{Q}T}\right)^{S_0-n}, \quad n \in \{0, 1, \dots, S_0\},$$

which proves Equation (7). Here, we have also used the fact that $\binom{S_0}{n} = \binom{S_0}{S_0-n}$. Thus, the distribution of the number of exposed individuals during $[0, T]$, under steady-state conditions, is indeed a Binomial distribution

$$E(T) \sim Bin\left(S_0, 1 - e^{-\frac{pqI_0}{Q}T}\right).$$

APPENDIX B

Per capita probability of infection during $[0, T]$ under Exponentially distributed or Erlang distributed random times.

Let us consider that T is exponentially distributed with rate λ, $T \sim Exp(\lambda)$, so that $f_T(t) = \lambda e^{-\lambda t}$ for $t \geq 0$. Then, the per capita probability of infection during $[0, T]$ is

$$\begin{aligned} P(T) &= \int_0^{+\infty}\left(1 - e^{-(\frac{pqI_0}{Q})t}\right)\lambda e^{-\lambda t} \\ &= \int_0^{+\infty}\lambda e^{-\lambda t}\,dt - \int_0^{+\infty}\lambda e^{-(\frac{pqI_0}{Q}+\lambda)t}\,dt \\ &= 1 + \frac{\lambda}{\frac{pqI_0}{Q}+\lambda}\left[e^{-(\frac{pqI_0}{Q}+\lambda)t}\right]_{t=0}^{+\infty} \end{aligned}$$

so that

$$P(T) \;=\; 1 - \frac{\lambda}{\frac{pqI_0}{Q} + \lambda}\,,$$

which corresponds to Equation (17).

If one considers instead $T \sim Erlang(k,\ \lambda)$, so that $f_T(t) = \frac{\lambda^k\ t^{k-1}\ \mathrm{e}^{-\lambda t}}{(k-1)!}$ for $t \geq 0$, the per capita probability of infection during $[0, T]$ is given by

$$\begin{aligned} P(T) &= \int_0^{+\infty} \left(1 - \mathrm{e}^{-(\frac{pqI_0}{Q})t}\right) \frac{\lambda^k\ t^{k-1}\ \mathrm{e}^{-\lambda t}}{(k-1)!}\ dt \\ &= \int_0^{+\infty} \frac{\lambda^k\ t^{k-1}\ \mathrm{e}^{-\lambda t}}{(k-1)!}\ dt \;-\; \int_0^{+\infty} \frac{\lambda^k\ \mathrm{e}^{-(\frac{pqI_0}{Q}+\lambda)t}\ t^{k-1}}{(k-1)!}\ dt \\ &= \underbrace{\frac{\lambda^k}{(k-1)!}\int_0^{+\infty} \mathrm{e}^{-\lambda t}\ t^{k-1}\ dt}_{(\dagger)} \;-\; \underbrace{\frac{\lambda^k}{(k-1)!}\int_0^{+\infty} \mathrm{e}^{-(\frac{pqI_0}{Q}+\lambda)t}\ t^{k-1}\ dt}_{(\dagger\dagger)}\,. \end{aligned}$$

One can compute (†)

$$\int_0^{+\infty} \frac{\lambda^k}{(k-1)!}\mathrm{e}^{-\lambda t}\ t^{k-1}\ dt \;= 1$$

since it corresponds to integrating the density function of an $Erlang\,(k, \lambda)$. One can also compute (††)

$$\begin{aligned} \int_0^{+\infty} \frac{\lambda^k t^{k-1}}{(k-1)!}\mathrm{e}^{-(\frac{pqI_0}{Q}+\lambda)t} dt &= \left(\frac{\lambda}{\frac{pqI_0}{Q}+\lambda}\right)^k \underbrace{\int_0^{+\infty} \frac{t^{k-1}}{(k-1)!}\left(\frac{pqI_0}{Q}+\lambda\right)^k \mathrm{e}^{-(\frac{pqI_0}{Q}+\lambda)t}\ dt}_{\text{Equals 1, since it is the density of an } Erlang\left(k, \frac{pqI_0}{Q}+\lambda\right)} \\ &= \left(\frac{\lambda}{\frac{pqI_0}{Q}+\lambda}\right)^k. \end{aligned}$$

Thus,

$$P(T) = 1 - \left(\frac{\lambda}{\frac{pqI_0}{Q}+\lambda}\right)^k,$$

which corresponds to Equation (19).

APPENDIX C

Probability distribution of the number of exposures during $[0, T]$ for Exponentially distributed and Erlang distributed random times.

Let us consider that T is exponentially distributed with rate λ, $T \sim Exp(\lambda)$, so that $f_T(t) = \lambda e^{-\lambda t}$ for $t \geq 0$. Then, the probability of observing exactly n exposures during $[0, T]$, for $n \in \{0, 1, \ldots, S_0\}$, is given by

$$\begin{aligned}
\mathbb{P}\{E(T) = n\} &= \int_0^{+\infty} \binom{S_0}{n} \left(1 - e^{-\frac{pqI_0}{Q}t}\right)^n \left(e^{-\frac{pqI_0}{Q}t}\right)^{S_0-n} \lambda e^{-\lambda t}\, dt \\
&= \binom{S_0}{n} \lambda \int_0^{+\infty} \underbrace{\left(1 - e^{-\frac{pqI_0}{Q}t}\right)^n}_{\text{Use Binomial Expansion}} \left(e^{-\frac{pqI_0}{Q}t}\right)^{S_0-n} e^{-\lambda t}\, dt \\
&= \binom{S_0}{n} \lambda \int_0^{+\infty} \left(e^{-\frac{pqI_0}{Q}t}\right)^{S_0-n} e^{-\lambda t} \sum_{i=0}^{n} (-1)^i \binom{n}{i} e^{-i\frac{pqI_0}{Q}t}\, dt \\
&= \binom{S_0}{n} \lambda \sum_{i=0}^{n} (-1)^i \binom{n}{i} \int_0^{+\infty} e^{-\left[(S_0-n+i)\frac{pqI_0}{Q}+\lambda\right]t}\, dt \\
&= -\binom{S_0}{n} \lambda \sum_{i=0}^{n} (-1)^i \binom{n}{i} \left[\frac{1}{(S_0-n+i)\frac{pqI_0}{Q}+\lambda} e^{-\left[(S_0-n+i)\frac{pqI_0}{Q}+\lambda\right]t}\right]_{t=0}^{+\infty},
\end{aligned}$$

which leads to

$$\mathbb{P}\{E(T) = n\} = \binom{S_0}{n} \sum_{i=0}^{n} (-1)^i \binom{n}{i} \left(\frac{\lambda}{(S_0-n+i)\frac{pqI_0}{Q}+\lambda}\right)^k, \quad n \in \{0, 1, \ldots, S_0\},$$

corresponding to Equation (18).

Alternatively, if we consider $T \sim Erlang(k,\ \lambda)$, so that $f_T(t) = \frac{\lambda^k\, t^{k-1}\, e^{-\lambda t}}{(k-1)!}$ for $t \geq 0$, the probability of observing exactly n exposures during $[0, T]$, for $n \in \{0, 1, \ldots, S_0\}$, is given by

$$\begin{aligned}
\mathbb{P}\{E(T) = n\} &= \int_0^{+\infty} \binom{S_0}{n} \left(1 - e^{-\frac{pqI_0}{Q}t}\right)^n \left(e^{-\frac{pqI_0}{Q}t}\right)^{S_0-n} \frac{\lambda^k\, t^{k-1}\, e^{-\lambda t}}{(k-1)!}\, dt \\
&= \binom{S_0}{n} \frac{\lambda^k}{(k-1)!} \int_0^{+\infty} \underbrace{\left(1 - e^{-\frac{pqI_0}{Q}t}\right)^n}_{\text{Use Binomial Expansion}} \left(e^{-\frac{pqI_0}{Q}t}\right)^{S_0-n} e^{-\lambda t}\, t^{k-1}\, dt \\
&= \binom{S_0}{n} \frac{\lambda^k}{(k-1)!} \int_0^{+\infty} \left(e^{-\frac{pqI_0}{Q}t}\right)^{S_0-n} e^{-\lambda t}\, t^{k-1} \sum_{i=0}^{n} (-1)^i \binom{n}{i} e^{-i\frac{pqI_0}{Q}t}\, dt \\
&= \binom{S_0}{n} \frac{\lambda^k}{(k-1)!} \sum_{i=0}^{n} (-1)^i \binom{n}{i} \int_0^{+\infty} \underbrace{e^{-\left[(S_0-n+i)\frac{pqI_0}{Q}+\lambda\right]t}}_{\text{Let } \gamma_i = (S_0-n+i)\frac{pqI_0}{Q}+\lambda}\, t^{k-1}\, dt \\
&= \binom{S_0}{n} \sum_{i=0}^{n} (-1)^i \binom{n}{i} \left(\frac{\lambda}{\gamma_i}\right)^k \underbrace{\int_0^{+\infty} e^{-\gamma_i t} \frac{(\gamma_i)^k t^{k-1}}{(k-1)!}\, dt}_{\text{Equals 1, since it is the density of an } Erlang(k,\gamma_i)} \\
&= \binom{S_0}{n} \sum_{i=0}^{n} (-1)^i \binom{n}{i} \left(\frac{\lambda}{\gamma_i}\right)^k.
\end{aligned}$$

Finally, recalling that $\gamma_i = (S_0 - n + i)\frac{pqI_0}{Q} + \lambda$, we get

$$\mathbb{P}\{E(T) = n\} = \binom{S_0}{n} \sum_{i=0}^{n} (-1)^i \binom{n}{i} \left(\frac{\lambda}{(S_0 - n + i)\frac{pqI_0}{Q} + \lambda} \right)^k, \quad n \in \{0, 1, \ldots, S_0\},$$

which corresponds to Equation (20).