medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Risk assessment for airborne disease transmission by poly-pathogen aerosols

Freja Nordsiek[1], Eberhard Bodenschatz[1,2,3*], Gholamhossein Bagheri[1],

1 Max Planck Institute for Dynamics and Self-Organization (MPIDS), Göttingen, Niedersachsen, Germany
2 Institute for Dynamics of Complex Systems, University of Göttingen, Göttingen, Niedersachsen, Germany
3 Laboratory of Atomic and Solid State Physics and Sibley School of Mechanical and Aerospace Engineering, Cornell University, Ithaca, New York, USA

* lfpn-office@ds.mpg.de

Abstract

In the case of airborne diseases, pathogen copies are transmitted by droplets of respiratory tract fluid that are exhaled by the infectious and, after partial or full drying, inhaled as aerosols by the susceptible. The risk of infection in indoor environments is typically modelled using the Wells-Riley model or a Wells-Riley-like formulation, usually assuming the pathogen dose follows a Poisson distribution (mono-pathogen assumption). Aerosols that hold more than one pathogen copy, i.e. poly-pathogen aerosols, break this assumption even if the aerosol dose itself follows a Poisson distribution. For the largest aerosols where the number of pathogen in each aerosol can sometimes be several hundred or several thousand, the effect is non-negligible, especially in diseases where the risk of infection per pathogen is high. Here we report on a generalization of the Wells-Riley model and dose-response models for poly-pathogen aerosols by separately modeling each number of pathogen copies per aerosol, while the aerosol dose itself follows a Poisson distribution. This results in a model for computational risk assessment suitable for mono-/poly-pathogen aerosols. We show that the mono-pathogen assumption significantly overestimates the risk of infection for high pathogen concentrations in the respiratory tract fluid. The model also includes the aerosol removal due to filtering by the individuals which becomes significant for poorly ventilated environments with a high density of individuals, and systematically includes the effects of facemasks in the infectious aerosol source and sink terms and dose calculations.

Introduction

It is well known that some diseases such as influenza, the common cold, *Mycobacterium tuberculosis*, measles, Severe Acute Respiratory Syndrome Coronavirus 1 and 2 (SARS-CoV-1 and SARS-CoV-2), are airborne; meaning they are transmitted by particles (also called liquid droplets, aerosols, or droplet nuclei) exhaled by infected individuals. These particles come from the fluid of the lungs, vocal chords, mouth, and nose; which hereafter are all noted as "respiratory tract". Risk of getting infected from such particles for an individual or a population has been subject of numerous studies and analyzes [1, 2, 2–8]. Many of the transmission mitigation strategies rely on results obtained by models that take into account a variety of factors to assess the likelihood of

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

transmission, a good example of which is the World Health Organization's 2009 guidelines *Natural Ventilation for Infection Control in Health-Care Settings* [9]. Two well-known families of models are dose-response and Wells-Riley models, which have been extensively used to model spread of airborne diseases [10].

There are several dose-response models for various diseases in existence which consider the risk of infection for an average dose of pathogen copies, taking full account of the counting statistics [11]. Two common models are the exponential and beta-Poisson models, which are described in great detail by Haas, Rose & Gerba [11]. Many diseases follow the exponential model, which has the added simplicity of having only a single adjustable parameter. Both the exponential and beta-Poisson modelassume that the minimum number of pathogen copies required for infection, the threshold, is one; but other models exist for non-unity thresholds. Both models, along with many others, assume that the number of pathogen copies absorbed follows a Poisson distribution; though modification of the exponential model for doses following a beta or gamma distribution has been conducted [2].

The Wells-Riley model, in its original form, takes the steady state balance of sources and sinks of infectious pathogen copies (in units of quanta) over a period of time in a well-mixed indoor environment such as a room or several rooms connected via ventilation to calculate the average dose received by susceptible individuals over a time period, which is then run through an exponential dose-response model [1]. The original model measures pathogen copies in units of quanta, which is defined as $\mathrm{ID}_{63.21}$ pathogen copies [10]. Sources such as exhalation by infectious individuals in the environment and air exchange with other environments with infectious aerosols and sinks due to fluxes with outside, filtering by the ventilation, filtering by masks, inactivation, settling, and deposition have all been considered as well as full temporal modelling of infectious aerosol concentration rather than assuming steady-state [1, 2, 2–8]. At their heart, it is essentially a conservation of infectious aerosols model, choosing some sources and sinks to explicitly include and considering others to be negligible, to get the pathogen concentration and then the average inhaled dose, before using a dose-response model (usually the exponential model) for the infection risk. Note, in the literature the term "Wells-Riley model" is sometimes used to refer only to when this formulation is used with an exponential model, and the terms "Wells-Riley equation" and "dose-response model" used if other dose-response models are used instead (e.g. [10]). We will use the term "Wells-Riley formulation" to refer to both.

In the past, various generalizations and improvements have been applied to the Wells-Riley formulation for situations beyond its original design and to address its limitations [10]. For example; Nicas, Nazaroff & Hubbard [6] included sink terms for pathogen inactivation, aerosol settling, and deposition as well as less than unitary efficiency of the respiratory tract absorbing infectious aerosols. Wells-Riley formulations have also been combined with SIR (Susceptible-Infectious-Removed) and SEIR (Susceptible-Exposed-Infectious-Removed) models [3, 12]. Noakes & Sleigh [13] made a stochastic model with compartmentalization of the environment into well-mixed subregions that have less mixing with other regions that can work for periods of time longer than the incubation period. Recent Wells-Riley based analyzes during the ongoing SARS-CoV-2 pandemic also include the effects of masks (such as [8]) unless they are investigating scenarios in which individuals are not wearing any mask [7], though including the effect of masks predates the pandemic by decades [2–5].

One of the biggest assumptions of the Wells-Riley formulation is that the indoor environment considered is sufficiently well-mixed [1, 4–8, 10, 13]. Essentially, it assumes that the infectious aerosol concentration is homogeneous enough that the concentration inhaled by susceptible individuals and at all sinks is approximately equal to the volume average concentration [1, 4–8, 10, 13]. The practice of social distancing, using fans to

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

better mix the room, etc. all improve the quality of this assumption, but room conditions in real-world situations can be far away from the well-mixed state.

Besides the well-mixed assumption, there are several other assumptions associated with the Wells-Riley formulations, which are not necessarily always true. As an example, there is an additional loss term that has not been considered yet that is the loss of the infectious aerosols absorbed by the individuals themselves. This is despite the fact that this is exactly the reason that susceptible individuals get infected. In some cases this can be safely neglected, e.g. if the combined breathing volume exchange rate of all individuals in the environment is negligible compared to that of ventilation. But in a poorly ventilated room with many individuals inside, this sink term must be taken into account – not incorporatinhg it leads to false risk predictions.

Another large assumption is that the absorbed doses follow a Poisson distribution, which is implicit in the use of the exponential dose-response model even if not stated explicitly [1, 3, 6–8], though there has been work on doses following beta and gamma distributions [2]. The Poisson distribution assumption requires that the pathogen-carrying aerosols have at most one pathogen inside, i.e. a mono-pathogen assumption. However, this assumption is violated if the pathogen concentration in an infectious individual's respiratory tract is high. For this poly-pathogen situation the Wells-Riley formulation and the dose-response models must be generalized to consider a larger number of pathogen in an individual aerosol explicitly. We will use the term multiplicity to refer to the number of pathogen copies in an aerosol.

Ignoring multiplicity causes the infection risk to be overestimated even though the expected average pathogen dose does not change.Using a modified version of the worked example later in this manuscript, Fig 1 shows this effect on the time required to reach a 50% infection risk for different pathogen concentrations in the respiratory tract fluid with and without considering multiplicity. For low pathogen concentrations and small infection probabilities per pathogen, ignoring multiplicity has only a small effect. But for high pathogen concentrations and/or pathogen copies with a high infection probability per pathogen, ignoring multiplicity has a significant impact. For a respiratory tract pathogen concentration of 10^{11} cm^{-3} where the average number of pathogen copies per aerosol is approximately 6500 for a 50 µm in diameter at production, if the single pathogen infection probability (r) is large enough that multiplicity matters, this means taking into account multiplicities up to approximately 7000.

In this manuscript, we will consider the following generalizations and modifications to the Wells-Riley formulation:

- Fully accounting for the multiplicity of pathogen copies in aerosols and the effect on the dose-response models.
- Additional sink terms due to the filtering of air by people inhaling and then exhaling it back out, including the effects of masks.
- Working exclusively in units of pathogen copies and aerosols instead of quanta (note: quantum is undefinable when accounting for multiplicity).

We will first generalize dose-response models that assume Poisson distributed doses for the distribution that results from poly-pathogen aerosols being present. Then we will develop the general pathogen concentration model that is a generalization of the Wells-Riley formulation. This results in a linear inhomogeneous coupled system of ODEs (Ordinary Differential Equations) for each initial aerosol diameter at production (diameter when exhaled), and one equation for each multiplicity that must be considered. We then derive the general solution, and then simplify the general solution for coefficients that are constant in time. Requirements and heuristics are developed for finding the appropriate cutoff in the multiplicity, M_c. This is important because the

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

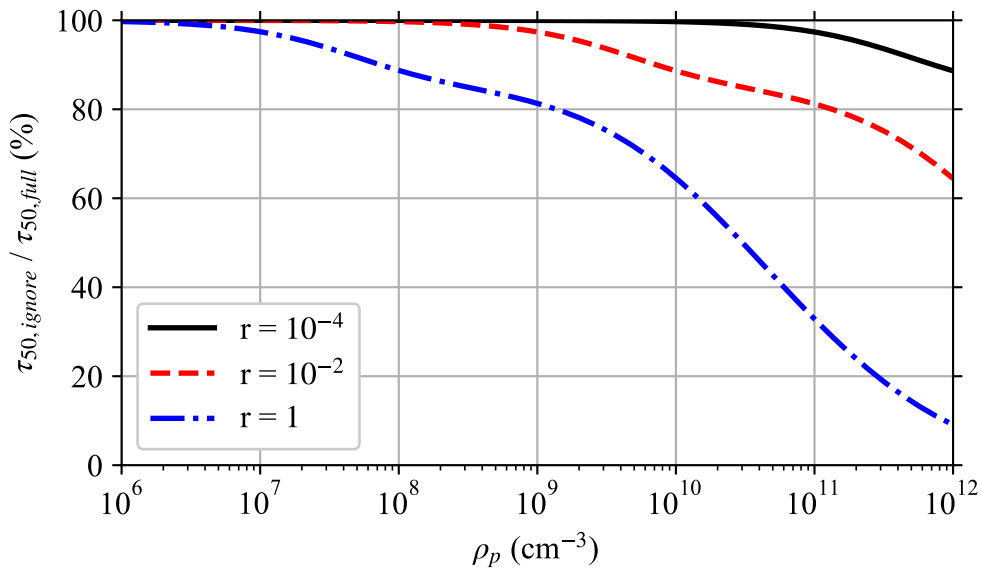


Fig 1. Effect of Ignoring Multiplicity. Ratio of the time required to reach a 50% infection risk when multiplicity is ignored $\tau_{50,ignore}$ to when it is fully accounted for $\tau_{50,full}$ for single pathogen infection probabilities r (an average dose of r^{-1} Poisson distributed pathogen copies gives a mean infection risk of 63.21%) and different pathogen concentrations ρ_p in the respiratory tract fluid of the infectious individual as in the worked example later in the manuscript with a disease following the exponential model, but at steady-state with just the speaking mask-less infectious individual and the risk to a mask-less susceptible individual whose exposure starts after steady state is reached. This is a simplified version of Fig 5.

number of ODEs to solve is equal to M_c; and the computational effort scales as $\mathcal{O}\left(M_c^2\right)$, or worse than for $\mathcal{O}\left(M_c^2\right)$ or $\mathcal{O}\left(M_c^3\right)$ for the different analytical solutions for coefficients constant in time. Some circumstances allow small $M_c = 1$ or close to one. We consider a full hypothetical example situation for SARS-CoV-2 with very high viral loads to apply the generalized Wells-Riley formulation developed in this manuscript. Finally, we discuss the effects of poly-pathogen aerosols, the filtering by the people in the environment, the effects of face-masks, and the model limitations.

Fundamentals

Throughout this manuscript, we will use the Poisson distribution, which describes the probability of counting some number, m, of independent events/objects/etc. as a function of the ensemble mean of the number counted, μ. The Probability Distribution Function (PDF) of the Poisson distribution is

$$P_P\left(\mu, m\right) = e^{-\mu}\left(\frac{\mu^m}{m!}\right) \quad . \tag{1}$$

Most dose-response models assume that the number of pathogen copies absorbed follows a Poisson distribution. For the case of a dose-response model, the average number of pathogen copies absorbed over some period of time would be the μ and then P_P would give the probability that a person absorbed exactly m pathogen copies. For clarity in the rest of this manuscript, we will now define Δ to be the number of

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

pathogen copies absorbed (instead of m) and the average number of pathogen copies absorbed is $\langle\Delta\rangle$, where we have used $\langle\cdot\rangle$ to denote the average. The use of a Poisson distribution for the doses requires that the pathogen copies are independent (i.e. no clumping); and as we will later show, that the number of pathogen copies in aerosols is assumed to be one or zero.

Let $R(\Delta)$ denote the infection probability when exactly Δ pathogen copies are absorbed, and $\mathfrak{R}(\langle\Delta\rangle)$ denote the average infection probability when the average number of pathogen copies absorbed is $\langle\Delta\rangle$. For a disease where the threshold (minimum number of pathogen copies required for infection) is greater than one, the threshold must be included into the its definition of $R(\Delta)$ such that it is zero for Δ less than the threshold, which makes $R(\Delta)$ be a piece-wise function.

There are two ways to construct $\mathfrak{R}(\langle\Delta\rangle)$ from $R(\Delta)$. We use the method of taking the sum over all possible $\Delta \in [1, \infty)$ of the product of the probability of absorbing each particular Δ and the resulting infection risk $R(\Delta)$ [10]. If the number of pathogen copies absorbed follows a Poisson distribution, then

$$\mathfrak{R}\left(\langle\Delta\rangle\right) = \sum_{\Delta=1}^{\infty} P_P\left(\langle\Delta\rangle, \Delta\right) R(\Delta) \quad . \tag{2}$$

The other method instead considers the number of pathogen copies that survive to try to infect, Δ_i, and does a double sum over Δ_i (starting from the threshold) and Δ of the product of the probability of the dose Δ and the probability of exactly Δ_i out of Δ surviving to try to infect [11] (this is **NOT** $R(\Delta)$). The two methods are equivalent, with this extra sum being implicitly included in the definition of $R(\Delta)$. This is why $R(\Delta)$ is a piece-wise function when the threshold is not one. For some models it may be easier to do this other method explicitly rather than try to construct $R(\Delta)$.

The exponential model assumes that all pathogen copies are identical, all people are equally vulnerable to infection, that the pathogen copies are acting independently of each other, and that each pathogen has an equal probability of causing infection r [11]. These assumptions implicitly means that the threshold is one. Each pathogen has a probability $1 - r$ to not infect. Then the exponential model's infection risk for an exact dose Δ is just one minus the probability that all Δ pathogen copies did not infect.

$$R_E(\Delta) = 1 - (1 - r)^{\Delta} \quad . \tag{3}$$

If the dose follows a Poisson distribution, then Eq (2) can be calculated for the exponential model [10], yielding

$$\mathfrak{R}_E\left(\langle\Delta\rangle\right) = 1 - e^{-r\langle\Delta\rangle} \quad . \tag{4}$$

Note that often, the parameter $D \equiv 1/r$ is used instead of r (the symbol k is also used [14]), which is the $\mathrm{ID}_{63.21}$ (Infective Dose required for 63.21% chance of infection). We will be making non-Poissonity corrections to this later.

The beta-Poisson model is essentially the exponential model but considers instead of everyone being equally vulnerable, each person has their own value for r which comes from the beta distribution [10, 11]. The beta distribution PDF [11] is

$$P_B(r) = \frac{\Gamma(\varepsilon + \theta)}{\Gamma(\varepsilon)\Gamma(\theta)} r^{\varepsilon - 1}(1 - r)^{\theta - 1} \quad , \tag{5}$$

where $r \in [0, 1]$ and the symbols ε and θ have been used in place of the conventional alpha and beta parameters respectively to avoid clashing with symbols used later in this manuscript. This means that to get mean infectou risk for a beta-Poisson model $R_{BP}(\Delta)$, we must include an integral over all $r \in [0, 1]$. Specifically,

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

$$R_{BP}(\Delta) = \int_0^1 P_B(r) R_E(\Delta) dr \quad . \tag{6}$$

Since the integral commutes with the sums used to calculate $\mathfrak{R}(\langle\Delta\rangle)$, the integral can be calculated as an outer integral rather than an inner integral yielding [11]

$$\mathfrak{R}_{BP}(\langle\Delta\rangle) = \int_0^1 P_B(r) \mathfrak{R}_E(\langle\Delta\rangle) dr \quad . \tag{7}$$

Wells-Riley formulations, both the original model and many subsequent uses, measure pathogen copies in units of quanta [1, 5–8, 10, 13]. A quanta is defined as $\mathrm{ID}_{63.21}$ pathogen copies [10]. This means that one quantum is equal to $D = 1/r$ pathogen copies. For the case of $r = 1$ such as *Mycobacterium tuberculosis*, one quantum is one pathogen [6, 10]. Using these units, the exponential model from Eq (4) becomes

$$\mathfrak{R}_E\left(\langle\mathbb{Q}\rangle\right) = 1 - e^{-\langle\mathbb{Q}\rangle} \quad , \tag{8}$$

where $\mathbb{Q}$ is the number of absorbed quanta [1, 5–8, 10, 13].

Let N_I be the number of infectious individuals, σ be the average production rate of infectious quanta per infectious individual, λ be the volumetric breathing rate of susceptible individuals, Q be the volumetric rate that clean air is brought into the particular indoor environment, and τ be the time period of exposure of susceptible individuals. Then, in its simplest form, the Wells-Riley Model's infection probability for time periods smaller than the incubation period of the disease [1] is

$$\mathfrak{R}_{WR}\left(\tau\right) = 1 - \exp\left[-\left(\frac{N_I \sigma}{Q}\right)\lambda\tau\right] \quad . \tag{9}$$

For time periods longer than the incubation period of the disease, one must either break the time period into subintervals smaller than the incubation period [1] or model both $\mathfrak{R}$ and the number of infectious and susceptible individuals over time with a SIR or SEIR model [3, 12].

Dose-Response Models for poly-Pathogen Aerosols

General

If the pathogen concentration in an infectious individual's respiratory tract fluid ρ_p is low enough, almost all exhaled pathogen copies will be the only pathogen in their aerosols, i.e. mono-multiplicity aerosols, and poly-multiplicity aerosols can reliably be ignored. We will use the tailing subscript k to denote aerosols with k pathogen copies inside them. An aerosol cannot contain more pathogen copies than will fit in its volume, and there is a limit to how large an aerosol a person can exhale. Let M be the maximum number of pathogen copies that can fit in the largest aerosol that can possibly be exhaled. This is the hard cutoff/limit on k. There also exists a soft cutoff/limit $M_c \leq M$ for which contributions of aerosols with $k > M_c$ is negligible. In a worst case $M_c = M$, but in practice it can be much lower since the pathogen volume fraction of respiratory tract fluid is quite low even at the upper pathogen load for some diseases. For example, SARS-CoV-2 at the very upper end of its concentration range at 10^{11} cm^{-3} [15, 16] would give a volume fraction of approximately 5×10^{-5}, if we treat the virus as a 100 nm sphere (approximate size of the SARS-CoV-2 virus [17]). This is important because an aerosol with a diameter of 1 µm could contain up to approximately 740 spherical pathogen copies with diameter 100 nm, if we assume hard-sphere packing (packing fraction of 74%). An aerosol with a diameter of 10 µm

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

could contain up to approximately 7.4×10^5 of the same pathogen copies for the same packing fraction.

To properly account for higher multiplicities, we must consider the separate doses for each multiplicity. Let Δ_k be the number of **pathogen copies** absorbed from aerosols with multiplicity k, and let m_k be the number of **aerosols** absorbed with multiplicity k. The aerosol and pathogen doses are related by $\Delta_k = km_k$. The total pathogen dose from all aerosols is just the sum of the doses for each multiplicity, which is $\Delta = \sum_{k=1}^{\infty} \Delta_k$. Let $\mu_k = \langle m_k \rangle = \langle \Delta_k \rangle / k$ be the average number of absorbed aerosols with multiplicity k.

As long as the aerosols are randomly distributed in space (well-mixed with no clustering nor avoidance), then the PDF of each m_k follows a Poisson distribution with mean μ_k. Since $\Delta_k = km_k$, the PDF of Δ_k is not a Poisson distribution for $k > 1$. It is instead a scaled-Poisson distribution of the form

$$P_k\left(\mu_k, \Delta_k\right) = \begin{cases} P_P\left(\mu_k, \frac{\Delta_k}{k}\right) & \text{if} \quad \Delta_k \mod k = 0 \quad , \\ 0 & \text{otherwise} \quad . \end{cases}$$

The deviation from the Poisson distribution is most visible in the fact that this distribution has holes. For example with $k = 2$, $P_k = 0$ for all odd Δ_k. Since Δ is the sum of a Poisson distribution $k = 1$ and some number of possibly non-negligible scaled-Poisson distributions, the PDF of Δ will not be a Poisson distribution unless the contributions from $k > 1$ are negligible compared to $k = 1$. So we can't just naively put the expected average dose into dose-response models expecting a Poisson distribution.

Instead, we must change the summation in Eq (2) to get the infection risk $\mathfrak{R}$. Let us consider the p'th moment, $\mathcal{M}_p$, of the infection probabilities as a function of the average aerosol doses μ_k (note, we use p in later sections of this manuscript as a summation index). To determine $\mathcal{M}_p$, we must sum over all possible combinations of exact aerosol doses m_k of each multiplicity for $k \in [1, \infty)$ of the product of the Poisson probabilities of each m_k and the infection risk for the dose raised to the power of p. This is

$$\mathcal{M}_p\left(\mu_1, \ldots, \mu_\infty\right) = \overbrace{\sum_{m_1=0}^{\infty} \cdots \sum_{m_\infty=0}^{\infty}}^{\text{all combinations}} \overbrace{\left[\prod_{k=1}^{\infty} P_P\left(\mu_k, m_k\right)\right]}^{\text{probability of dose}} \overbrace{\left[R\Big(\underbrace{\sum_{k=1}^{\infty} km_k}_{\text{pathogen dose}}\Big)\right]^p}^{\text{infection probability}} \quad , \tag{10}$$

where we have written out the dose Δ inside R. The mean infection risk is the first moment ($p = 1$), which is

$$\mathfrak{R}\left(\mu_1, \ldots, \mu_\infty\right) = \sum_{m_1=0}^{\infty} \cdots \sum_{m_\infty=0}^{\infty} \left[\prod_{k=1}^{\infty} P_P\left(\mu_k, m_k\right)\right] R\left(\sum_{k=1}^{\infty} km_k\right) \quad . \tag{11}$$

Exponential Model Corrections

Then, putting R_E from Eq (3) into Eq (11), the exponential model mean infection risk is

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

$$\begin{aligned}
\mathfrak{R}_E\left(\mu_1,\ldots,\mu_\infty\right) &= \sum_{m_1=0}^{\infty}\cdots\sum_{m_\infty=0}^{\infty}\left[\prod_{k=1}^{\infty}P_P\left(\mu_k,m_k\right)\right]\left[1-(1-r)^{\sum_{k=1}^{\infty}km_k}\right] \\
&= 1-\sum_{m_1=0}^{\infty}\cdots\sum_{m_\infty=0}^{\infty}\prod_{k=1}^{\infty}e^{-\mu_k}e^{(1-r)^k\mu_k}e^{-(1-r)^k\mu_k}\frac{\left[(1-r)^k\mu_k\right]^{m_k}}{m_k!} \\
&= 1-\exp\left[-\sum_{k=1}^{\infty}\left(1-(1-r)^k\right)\mu_k\right] \quad , \qquad (12)
\end{aligned}$$

where the fact that the sum of all probabilities over the Poisson distribution is equal to one has been used extensively. The final sum has a finite number of terms due to the cutoff M as long as the μ_k are finite for $k \leq M$. For small M_c, we can truncate the risk probability and get an easier to calculate approximation. Except for $M_c = 1$, this is different from Eq (4) due to the non-Poissonity in Δ. The expression for the first few values of M_c are

$$\mathfrak{R}_E \approx \begin{cases} 1-e^{-r\mu_1} & \text{if } M_c = 1 \quad , \\ 1-e^{-r\mu_1}e^{-r(2-r)\mu_2} & \text{if } M_c = 2 \quad , \\ 1-e^{-r\mu_1}e^{-r(2-r)\mu_2}e^{-r\left(3-3r+r^2\right)\mu_3} & \text{if } M_c = 3 \quad . \end{cases} \qquad (13)$$

Beta-Poisson Model Corrections

The integral over r commutes with the sums in Eq (10). So as was with the case when multiplicity is not considered in Eq (7), we can get the moments by taking the result for the exponential model and integrating it times the beta distribution PDF over r. This is

$$\mathcal{M}_{BP,p}\left(\mu_1,\ldots,\mu_\infty\right) = \int_0^1 P_B(r)\mathcal{M}_{E,p}\left(\mu_1,\ldots,\mu_\infty\right)dr \quad . \qquad (14)$$

Unfortunately, as is the case for when the dose is Poisson distributed [11], the integral cannot be solved analytically and must be solved numerically or approximated though now it is harder with the extra terms for $M_c > 1$.

General Pathogen Concentration Model

Looking Ahead

Now that we have dose-response models corrected for the multiplicity via Eq (11), we must determine the average **aerosol** doses μ_k for each multiplicity before the infectiou risk can be calculated. We now generalize the Wells-Riley formulation for multi-pathogen aerosols to get this. In the following sections, we will describe the environment, people, aerosols, sources, sinks, etc. to get the model equations. Let $n_k(d_0,t)$ be the concentration density of aerosols with original diameter d_0 and k pathogen copies in them over time, which has units of $[\mathrm{L}]^{-4}$ where [L] is the unit of length since $n_k(d_0,t)dd_0$ is the concentration of infectious aerosols with diameters between d_0 and $d_0 + dd_0$. To get a concentration, $n_k(d_0,t)$ must be integrated with respect to d_0.

In the end, we will get the following system of ODEs (Ordinary Differential Equations) in time t and the original diameter at production d_0 for the n_k, which is

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

$$\frac{dn_k}{dt} = \overbrace{-\alpha(d_0,t)n_k}^{\text{sinks}} + \overbrace{(k+1)\,\gamma(t)n_{k+1} - k\gamma(t)n_k}^{\text{flux from inactivation}} + \overbrace{\beta_k(d_0,t)}^{\text{sources}} \quad , \tag{15}$$

where $\alpha(d_0,t)$ is the sum of all sink term coefficients, $\beta_k(d_0,t)$ is the sum of all sources for each k, $\gamma(t)$ is the pathogen inactivation rate, and we have assumed that the time period considered is shorter than the incubation time of the disease. Then the combined source and sink terms are

$$\beta_k(d_0,t) = \beta_{r,k} + \beta_{I,k} \quad , \tag{16}$$

$$\alpha(d_0,t) = \alpha_o + \alpha_r + \alpha_v + \alpha_g + \alpha_d + \alpha_{I,f} + \alpha_{S,f} + \alpha_{O,f} \quad , \tag{17}$$

which don't depend on $n_k(d_0,t)$ (i.e. no quadratic or higher order terms), though they may depend on t. The different sources and sinks are summarized in Table 1. See their relevant sections for the meanings of their terms, their assumptions, and where they come from.

Table 1. Source And Sink Term Summary Summary of all the source (the β) and sink (the α) terms considered in this manuscript. See their relevant sections for details on where they come from and the meanings of their terms.

Term	Meaning	Form
$\beta_{r,k}(d_0,t)$	transport from other rooms	$q_r(t)n_{r,k}(d_0,t)$
$\beta_{I,k}(d_0,t)$	production by infectious individuals	$\frac{N_I}{V}\langle\lambda_I(t)n_{I,k}(d_0,t)\left[1-E_{I,m,out}(d_0)\right]\rangle_I$
$\alpha_o(t)$	air exchange with outside	$q_o(t)$
$\alpha_r(t)$	air exchange with other rooms	$q_r(t)$
$\alpha_v(d_0,t)$	filtering by ventilation	$q_v(t)E_v(w(d_0,t)d_0)$
$\alpha_g(d_0,t)$	gravitational settling	$\approx \frac{1}{h}u_g(w(d_0,t)d_0)$
$\alpha_d(d_0,t)$	deposition on surfaces	found elsewhere
$\alpha_{I,f}(d_0,t)$	filtering by infectious individuals inhaling	$\frac{1}{V}\sum_{j=1}^{N_I}\lambda_{I,j}(t)\left[1-S_{I,m,in,j}(d_0,t)S_{I,r,j,k}(d_0)S_{I,m,out,j,k}(d_0)\right]$
$\alpha_{S,f}(d_0,t)$	filtering by susceptible individuals inhaling	$\frac{1}{V}\sum_{j=1}^{N_S}\lambda_{S,j}(t)\left[1-S_{S,m,in,j}(d_0,t)S_{S,r,j,k}(d_0)S_{S,m,out,j,k}(d_0)\right]$
$\alpha_{O,f}(d_0,t)$	filtering by other individuals inhaling	$\frac{1}{V}\sum_{j=1}^{N_O}\lambda_{O,j}(t)\left[1-S_{O,m,in,j}(d_0,t)S_{O,r,j,k}(d_0)S_{O,m,out,j,k}(d_0)\right]$

Environment

Like most Wells-Riley formulations, we consider the infection risk in one sufficiently well-mixed indoor environment such as a room or set of rooms sufficiently coupled together with respect to their air that they have the same infectious aerosol concentration densities. The environment could also be split into coupled well-mixed zones with weaker mixing between them [4, 13], but that shall not be considered here.

Let the volume of the environment be V. Air is exchanged with outside, with other rooms, and circulated internally through the ventilation system. Let Q_o, Q_r, and Q_v be the volumetric rate of air exchange with outdoors, other rooms, and the circulating ventilation of the environment (ventilation system that pulls air out of the environment and puts it back in). These will be normalized by the environment volume; yielding $q_o \equiv Q_o/V$, $q_r \equiv Q_r/V$, and $q_v \equiv Q_v/V$ since target values of these parameters are often the design goals for HVAC systems.

Aerosols

Consider the concentration of infectious aerosols over time. To be completely accurate, we need to consider the concentration density for each multiplicity k as a function of

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

time, current diameter d while in the environment, and the solute content (including inactivated pathogen copies). We have to consider both d and the solute content because an exhaled aerosol's equilibrium diameter is a function of its solute content, the humidity, and the temperature [18]. Higher solute concentrations decrease the vapor pressure of the aerosol, which allows equilibrium to be reached as long as the environment isn't super-saturated close to saturated [18, 19]. For higher humidities, an aerosol will continue to grow by condensation indefinitely, though the growth rate slows towards a crawl for $d > 20$ µm [19, 20]. But such super-saturated conditions can cause clouds/fog, which rarely occur in indoor environments. So we will assume the environment is sub-saturated. If the environment is dry, the aerosols can evaporate at most to the point where they are purely precipitated solid with no water left.

This means that we have four different diameters to consider, which are

d current diameter in the environment (spherical equivalent diameter if it is completely dry or almost dry and the solute causes a non-spherical shape)

$\mathbf{d_e}$ equilibrium diameter in the environment

$\mathbf{d_0}$ wet diameter at production (original diameter), which determines the distribution of initial multiplicities

$\mathbf{d_D}$ spherical equivalent dry diameter when all water is evaporated away and just solute remains (note that the aerosol may no longer be spherical, so the spherical equivalent diameter for the same volume must be used)

For any aerosol; d_0 and d_D are fixed and never change as long as collisional-coalescence and shattering don't occur (can be treated as fixed if these processes are negligible), d_e is dynamic in time if the environment's temperature and/or humidity changes, and d is dynamic in time unless the environment's temperature and humidity exactly match those inside the respiratory tract at the point of production.

Small aerosols respond very quickly to the humidity and temperature, evaporating/condensing to their equilibrium diameter in a very short period of time due to their high surface area to volume ratio [6, 19, 21]. Assuming the environment is well-mixed enough that the time between exhalation from an infectious individual and inhalation by any person is long compared to the evaporation/condensing time scale, we can make the approximation that all aerosols are at their equilibrium diameter when in the environment ($d \approx d_e$). This means that we just need to worry about the equilibrium diameter and its changes, and not the non-immediate response to shifting equilibrium diameters. There is one complication, however. Aerosols will initially stay in the exhaled plume where the humidity is higher, so they won't reach the well-mixed equilibrium diameter till they leave the plume or the plume is diluted and mixed with the environment, which brings us back to the well-mixed environment assumption. We will also make the assumption that the temperatures and humidities in different individuals' respiratory tracts (and the volume under their facemasks if they are wearing any) are similar enough and that the aerosols grow fast enough that the aerosols approximately return to their original diameters at production when inhaled into the respiratory tracts of other people (or the person who generated them if they breath them back in). Combined, our assumption/approximation is

$$d(t) \approx \begin{cases} d_e(t) & \text{if in the environment outside of the respiratory tract} \quad , \\ d_0 & \text{if in anyone's respiratory tract} \quad . \end{cases} \tag{18}$$

Let us define ratios between the remaining diameters: the evaporation ratio w, the dilution ratio δ, and the initial solute ratio ζ as

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

$$w \equiv \frac{d_e}{d_0} \quad , \tag{19}$$

$$\delta \equiv \frac{d_e}{d_D} \quad , \tag{20}$$

$$\zeta \equiv \frac{d_D}{d_0} \quad . \tag{21}$$

Note that w and δ are potentially functions of time, as well as diameter due to the effect of surface curvature (through surface tension) on equilibrium vapor pressure [18, 19]. Also, different solutes have different molar densities, different practical osmotic coefficients, and maximum concentrations before the precipitate; and therefore different functional relationships between the saturation vapor pressure and the concentration [18]. So different solute compositions will cause w and δ to be different even for aerosols with the same ζ.

But, in the folloowing we will make the assumption that the value of ζ and the solute composition (except for the pathogen copies) is approximately constant from each infectious individual to the next and over time with each infectious individual, and we will ignore the contribution of the pathogen copies (both active and inactivated) to the equilibrium vapor pressure and therefore d_e. We will also assume that ζ has no diameter dependence (i.e. attraction and repulsion of solutes from the liquid surface at production has a negligible effect on solute fraction and composition). With these approximations, we have a single constant value of ζ and single functions for w and δ, possibly over time and d_0 (or equivalently d_D), for all infectious aerosols in the environment.

This means we can choose to track one of d_e, d_0, or d_D and always know the other two through the ratios that are the same for all infectious aerosols at the same moment of time with the same value of the chosen diameter parameter. Thus we have two independent variables, t and one diameter parameter.

Processes such as gravitational settling, deposition, filtering or exchange by the ventilation, filtering by facemasks when inhaling are all functions of the current diameter, which is approximately d_e, making d_e convenient. Additionally, any non-drying aerosol instruments can be used in the environment to measure d_e. But, because d_e can change over time for a fixed d_D or d_0, the equations for the aerosol concentration density in terms of t and d_e have a flux term (from evaporation/growth) with a partial derivative with respect to d_e; making the equations PDEs (Partial Differential Equations) which adds complications and can be avoided in the analysis. This can be seen by considering the total time derivative of the aerosol concentration density $\tilde{n}$ expressed in terms of t and d_e, which is

$$\frac{d\tilde{n}(d_e, t)}{dt} = \frac{\partial \tilde{n}}{\partial t} + \frac{\partial \tilde{n}}{\partial d_e}\frac{d_e}{dt} \quad . \tag{22}$$

Since d_D and d_0 are fixed for a given aerosol over time regardless of how the temperature or humidity in the environment might be changing, the equivalent flux term is zero and thus the equivalent functions are ODEs, which are much easier to solve. Thus, we eliminate d_e as a choice for the diameter parameter.

The model in this manuscript can be conducted with either choice of d_0 or d_D, with w appearing in places if d_0 is chosen and both δ and ζ appearing in places, if d_D is chosen. We choose d_0 because then we only need one of the ratios (w only), the diameter limits are easier to express in it, and the literature on the diameter distributions of exhaled aerosols generally work hard to convert their measurements (vary between whether they are d_e or d_D) into expressions in terms of d_0 rather than d_D.

Now, $n_k(d_0, t)$ be the concentration density of aerosols in terms of t and the original diameter d_0. Let $\tilde{n}_k$ be the concentration density in terms of t and d_e, and $\breve{n}_k$ be the

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

concentration density in terms of t and d_D. To make conversions between them; consider the original diameter interval d_0 to $d_0 + dd_0$, and its corresponding intervals d_e to $d_e + dd_e$ and d_D to $d_D + dd_D$. The number of aerosols in each interval must all be equal: $n_k dd_0$, $\tilde{n}_k dd_e$, and $\breve{n}_k dd_D$. Thus, the conversions are

$$\tilde{n}_k = \frac{n_k}{w} \quad , \tag{23}$$

$$\breve{n}_k = \frac{n_k}{\zeta} \quad , \tag{24}$$

$$\tilde{n}_k = \frac{\breve{n}_k}{\delta} \quad . \tag{25}$$

Let $n_{0,k}(d_0)$ be the initial concentration density in the room for a multiplicity k at the initial time $t = t_0$ and $n_{r,k}(d_0, t)$ be the volume averaged concentration density of the air coming in from other rooms. We are assuming that the concentration density outdoors is negligible.

Diameter Limits

For the model, we will limit ourselves for each multiplicity to the range $d_0 \in [d_{m,k}, d_M]$ where $d_{m,k}$ is the minimum aerosol diameter required to hold k pathogen copies, and d_M is a diameter cutoff separating larger aerosols that are more ballistic and gravitationally settle to the ground too quickly to become well mixed and smaller aerosols that more closely follow the flow and mix. Let $K_m(d_0)$ be the largest number of pathogen copies that can fit in an aerosol at production.. We will consider

$$n_k(d_0, t) = 0 \;\; \forall \;\; d \notin [d_{m,k}, d_M], \;\; k > K_m(d_0) \quad . \tag{26}$$

All of these limits have problems, but there is no obvious better choice without adding a lot more complexity to the model.

For a spherical pathogen with diameter d_p, we can use the crude approximation of just considering the total pathogen volume and a packing efficiency $e = 0.74$ (hard pack spheres) with a minimum of 1 and completely neglect the aerosol shape that small number of pathogen copies would force (two pathogen copies, for example, can't be arranged into a configuration that even vaguely resembles a sphere). We can use the same idea to get $K_m(d_0)$. Both of them are

$$d_{m,k} \approx \begin{cases} d_p & \text{if } k = 1 \quad , \\ \left(\frac{k}{e}\right)^{1/3} d_p & \text{if } k > 1 \quad , \end{cases} \tag{27}$$

$$K_m(d_0) \approx \max\left[1, e\left(\frac{d_0}{d_p}\right)^3\right] \quad . \tag{28}$$

At the lower limit near $d_{m,k}$, the pathogen/s take up a disproportionate amount of the space in the aerosol compared to other solutes and the assumption of approximately equal solute concentrations at production is violated and the evaporation ratio has a strong dependence on d_0 and the initial multiplicity, the latter of which we aren't tracking at all. However, as long as the total liquid volume of exhaled aerosols with diameters close to $d_{m,k}$ (say, those whose diameters are small enough that their volume is only a few times larger) is small compared to total liquid volume of the rest of the range in d_0, this problem will have a negligible effect. Additionally, the diameter dependence of many of the sink terms may be much smaller close to $d_{m,k}$ for submicron pathogen copies which means that the effect of assuming the wrong evaporation ratio

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

may be small. The smaller the pathogen, the less issues this will pose. It will be least important for small viruses, and possibly quite important for large bacteria and eukaryotic pathogens..

The upper limit is rather imprecise since there is no single hard separation scale that could be chosen unless the air is completely still in which case one can use a so called "Wells curve" (same Wells as of the Wells-Riley model) for the environment's humidity to determine the largest size that won't settle to the ground before evaporating to their equilibrium diameter, such as the original one [22] or newer ones [21]. But mixing of any sort complicates this. One might think that one could just rely on the fact that the gravitational settling sink term keeps growing with diameter and not bother with the problem. But, the well-mixed assumption breaks down and the lifetime of the aerosols converges towards depending solely on the initial diameter and the height of the infectious individual's mouth and nose from the ground. Additionally, the time to evaporate to the equilibrium diameter increases with increasing size. And from a practical standpoint, it is necessary in order to keep M_c from getting too large since $M_c \sim \mathcal{O}(d_M^3)$ for sufficiently large d_M and pathogen concentration in the infectious individual' respiratory tract fluid ρ_p. If we assume that the aerosols are approximately spherical (reasonably true except potentially when completely dried out) and their density is approximately equal to that of water ρ_w, the aerosols' inertial response times τ_p to fluid motions from Stokes drag (we are assuming they are small enough that contributions beyond Stokes drag are negligible) and gravitational settling terminal velocity u_g are

$$\tau_p = \frac{\rho_w d^2}{18\rho_a \nu_a} \quad , \tag{29}$$

$$u_g = \frac{(\rho_w - \rho_a)\, g d^2}{18\rho_a \nu_a} \approx g\tau_p \quad , \tag{30}$$

where ρ_a is the density of air, ν_a is the kinematic viscosity of air, and g is the acceleration due to gravity.

Both grow quadratically with diameter, which does not lend itself to a well defined cutoff scale. And additionally one must consider that once exhaled, the aerosols will tend to evaporate (relative humidity in the environment is typically lower than in the respiratory tract where it is close to 100%) thereby reducing their inertia and terminal velocities. For 10 µm, 20 µm, and 50 µm diameter aerosols; the terminal velocities at 20 °C and atmospheric pressure are 3.0 $\mathrm{mm\,s^{-1}}$, 1.2 $\mathrm{cm\,s^{-1}}$, and 7.5 $\mathrm{cm\,s^{-1}}$ respectively. However, larger aerosols take longer to evaporate/grow to their equilibrium diameter and therefore will settle at a faster rate initially than their final equilibrium diameter suggests, which makes them even more likely to be lost due to settling than smaller aerosols.

The simulations of Chong *et al.* [23] indicate that 100 µm aerosols are quite ballistic and quickly fall out of the exhaled plume, but 10 µm aerosols are carried along with the plume and stay in the air despite their evaporation being greatly slowed. This suggests that d_M should be chosen somewhere in the 10–100 µm range, which is further supported by the Wells curves found by Xie *et al.* [21]. For lack of a better suggestion; we suggest the use of $d_M = 50$ µm, which will be explored in the Discussion. Before evaporating, the terminal velocity is 7.5 $\mathrm{cm\,s^{-1}}$. If the evaporation ratio is a typical value in the $\frac{1}{2}$–$\frac{1}{5}$ range, the final evaporated diameter would be in the 10–25 µm range and have terminal velocities in the 3–19 $\mathrm{mm\,s^{-1}}$ range which is still in the range that indoor environment air flow can keep suspended (though with a high loss rate).

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

People and Infectious Aerosol Production

We will denote infectious individuals by the subscript I, susceptible individuals by the subscript S, and other individuals by the subscript O. The Other category is all the individuals who are non-infectious non-susceptible. This includes individuals that are immune before they enter the environment (following Jimenez [8]), all of the Removed group SIR and SEIR models except for the individuals who died or leave the environment, and all of the Exposed group in SEIR models. If one wants to make a full SEIR model from the model presented in this manuscript, the two subgroups (Exposed, and the part of Removed that is still within the environment and breathing plus the previously immune individuals) within this group will have to be treated explicitly. Let the number of individuals in category C be N_C. The total number of individuals is $N = N_I + N_S + N_O$. The subscript A will be used to refer to all individuals in all categories. Each count is potentially a function of time as individuals can come in and out of the environment. Let $\langle \cdot \rangle_C$ denote taking the average over all individuals in category C.

Let $\lambda_{C,j}(t)$ be the volumetric breathing rate of the j'th person in category C. Let $E_{C,m,in,j}(d)$ and $E_{C,m,out,j}(d)$ be the filtering efficiency of the mask (if any) of the j'th person in category C for inhalation and exhalation respectively.

The filter efficiencies of most masks vary significantly with aerosol diameter. Note that it is important that the leak rate of the mask be included in its filtering efficiency. These two filtering efficiencies are generally not equal because masks tend to leak more during exhalation than inhalation and aerosols have higher velocities on exhalation than inhalation. We will assume that all infectious aerosols caught by the mask aren't later re-aerosolized.

Let $E_{C,r,j}(d)$ be the filtering/absorption efficiency of the respiratory tract of the j'th person in category C. This term is non-zero, but it is also not equal to one since the respiratory tract does not absorb all infectious aerosols that pass through it [2, 4, 6, 10]. The best example of this is the observation that individuals can inhale smoke (which is composed of many aerosols) and then exhale some of it back out. The diameter will be $d_e = wd$ when passing through the mask on inhalation, and d_0 when in the respiratory tract and when passing through the mask on exhalation since the humidity between the mouth and nose and the mask is high and the distance is short, so there is little time for evaporation. It is often easier to work with the survival efficiencies rather than the filtering efficiencies, defined as

$$S_{C,m,in,j}(d_0,t) = 1 - E_{C,m,in,j}\left(w(d_0,t)d_0\right) \quad , \tag{31}$$

$$S_{C,r,j,k}(d_0) = 1 - E_{C,r,j}(d_0) \quad , \tag{32}$$

$$S_{C,m,out,j,k}(d_0) = 1 - E_{C,m,out,j}(d_0) \quad . \tag{33}$$

We will assume that the number of infectious pathogen copies in each exhaled droplet/aerosol follow a Poisson distribution where the mean count is equal to the aerosol's initial volume times the pathogen load in respiratory tract fluid at the point of production. This excludes diseases where pathogenic agents stick together and clump. Note that this implicitly means we are assuming that the pathogen volume fraction in the respiratory tract fluid is small. Otherwise, the non-Poissonity caused by there being a maximum number of pathogen copies that can fit in a finite sized drop will **NOT** be negligible.

Let $\rho_j\left(d_0,t\right)dd_0$ be the number density in exhaled air of the aerosols with diameters between d_0 and $d_0 + dd_0$ exhaled by the j'th infectious individual at time t. Let $\rho_{p,j}(t)$ be the pathogen concentration in the j'th person's respiratory tract fluid where the aerosols are being produced. The mean/expected multiplicity for infectious aerosols

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

produced by the j'th infectious individual for any d_0 is

$$\langle k \rangle (d_0, t)_j = \frac{\pi}{6} d_0^3 \rho_{p,j}(t) \quad . \tag{34}$$

If the pathogen copies are Poisson distributed in the fluid that makes up the aerosols (no clumping, etc.), then

$$n_{I,j,k}(d_0, t) = \begin{cases} \rho_j \left(d_0, t\right) P_P \left(\langle k \rangle (d_0, t)_j \, , k\right) & \text{if } d_0 \geq d_{m,k} \quad , \\ 0 & \text{if } d_0 < d_{m,k} \quad . \end{cases} \tag{35}$$

Note that no infectious aerosols with multiplicity k can be generated with diameters too small to contain them (i.e. no $d_0 < d_{m,k}$ aerosols).

Sources

We will denote sources by the symbol β with a subscript denoting the individual source. All of them are normalized by the volume of the environment, V.

First, ventilation with other rooms brings infectious aerosols inside at a rate, normalized by the environment volume, of

$$\beta_{r,k}(d_0, t) = q_r(t) n_{r,k}\left(d_0, t\right) \quad . \tag{36}$$

where we have lumped all other rooms that might be exchanging air with the room of interest together rather than summing over them as done by Noakes & Sleigh [13]. A coupled model for multiple rooms would have to split this into a sum and model the whole system. Note that we are assuming, like elsewhere, the aerosols brought in from other rooms reach their equilibrium diameter quickly compared to other processes.

The other source is the infectious individuals exhaling aerosols with pathogen copies in them. The total production from the infectious individuals normalized by the environment volume is the sum of the products of the breathing rate, the exhaled aerosol concentration density, and the survival efficiency of the mask [4, 8]; which is

$$\begin{aligned} \beta_{I,k}(d_0, t) &= \frac{1}{V} \sum_{j=1}^{N_I} \overbrace{\lambda_{I,j}(t) n_{I,j,k}(d_0, t)}^{\text{production rate}} \overbrace{\left[1 - E_{I,m,out,j}(d_0)\right]}^{\text{mask survival}} \\ &= \frac{N_I}{V} \left\langle \lambda_I(t) n_{I,k}(d_0, t) \left[1 - E_{I,m,out}(d_0)\right] \right\rangle_I \quad , \end{aligned} \tag{37}$$

where the j subscript has been dropped in the average. Any terms in the average of a product ($\lambda_{I,j}$, $n_{I,j,k}$, and $1 - E_{I,m,out,j,k}$) that have no correlation with the others can be pulled out to make a product of averages. But any correlated terms cannot be separated, which means it must be kept as an average of a product. As an example, if there are two infectious individuals in a room and one is singing and the other is listening in silence; they will be strongly correlated. The singing person will on average be breathing at a higher rate, could have a higher concentration density of infectious aerosols in their exhaled air, and probably won't be wearing a mask while the listener might be wearing a mask. Now, if all individuals are wearing the same mask, the mask term could be pulled out but the other two terms would remain since they could still be correlated.

Other than not replacing the average of the product with the product of the averages, following aerosols with a particular multiplicity rather than quanta, and consider each diameter separately; this term is identical to the equivalent term by Nazaroff, Nicas & Miller [4] and Jimenez [8] and, if masks are removed, that of the original formulation [1].

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Sinks

Sinks are proportional to the concentration density n_k. We will denote all sinks divided the concentration density by the symbol α with a subscript denoting the individual source. All of them are normalized by the volume of the environment, V. Unlike the sources, none of the sinks (except inactivation, considered separately) depend on the multiplicity and therefore the subscript k is dropped. Note that inactivation is treated separately later since it is a flux term when considering each multiplicity separately, unlike in the traditional formulation where it is a sink.

The volume normalized loss rate coefficients of infectious aerosols due to exchange of clean air with outdoors and other rooms are just the concentration density times the volume normalized flow rate [6, 13] and are

$$\alpha_o(t) = q_o(t) \quad , \tag{38}$$
$$\alpha_r(t) = q_r(t) \quad , \tag{39}$$

respectively.

Let $E_v(d)$ be the filtering efficiency of the circulating ventilation system for aerosols with diameter d. The diameter when an aerosol reaches this filter is $d \approx d_e = w(d_0, t)d_0$. Then the volume normalized loss rate coefficient from the circulating ventilation system [1] is

$$\alpha_v(d_0, t) = q_v(t) E_v\left(w(d_0, t)d_0\right) \quad . \tag{40}$$

Aerosols also gravitationally settle and deposit onto surfaces. We will treat these processes as simple loss rates proportional to their concentration densities just as one does with radioactive decay. The volume normalized loss rates divided by the concentration density, of gravitational settling and deposition are defined to be $\alpha_g(w(d_0, t)d_0)$ and $\alpha_d(w(d_0, t)d_0)$ respectively; which depend on the room geometry, aerosol diameter, and air flow in the room. A possible approximate expression for the settling loss term [6] would be

$$\alpha_g\left(w(d_0, t)d_0\right) \approx \frac{1}{h} u_g\left(w(d_0, t)d_0\right) \quad , \tag{41}$$

where h is the characteristic height of the indoor environment and $u_g(d)$ is the terminal velocity. For small spherical aerosols, Eq (30) provides $u_g(d)$. Larger aerosols need additional diameter corrections [6, 21, 24].

Sinks from Individuals Inhaling Aerosols

Unfortunately, when individuals inhale infectious aerosols, some are absorbed thereby causing a risk of infection. While this phenomena is not desired for susceptible individuals, we must consider the loss rate from this process by the susceptible individuals as well as the infectious individuals and the non-infectious non-susceptible individuals. There are three steps to the filtering process for the j'th person of category C: passing through the mask on inhalation, passing through the respiratory tract, and then passing through the mask on exhalation.

The total survival probability of an aerosol surviving going through all three steps is the product of the individual survival rates. The total filtering efficiency is then one minus the total survival rate. But, there is a time delay between when the aerosols are removed from the environment on inhalation and when the survivors are exhaled back out. As long as this time is short compared to all other time scales such as mixing times in the room, the time scales of all other sinks, the time scale of inactivation, etc.; we can

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

ignore this time delay and consider the re-exhalation to occur at the same time. This assumption implies that we can neglect possible changes in multiplicity by inactivation while the aerosols are in the respiratory tract. In most situations, this is a reasonably good assumption. But, at a swimming pool where people regularly hold their breath for long periods of time, this assumption could be violated for the highest multiplicities since the inactivation rate from k to $k-1$ is proportional to k.

The number of aerosols that are inhaled by a person is equal to $\lambda_{C,j}(t)n_k(d_0,t)$. The volume normalized sink coefficient from this filtering is then

$$\begin{aligned}\alpha_{C,f}(d_0,t) &= \frac{1}{V}\sum_{j=1}^{N_C} \overbrace{\lambda_{C,j}(t)}^{\text{volume rate}} \overbrace{\left[1-\overbrace{S_{C,m,in,j}(d_0,t)}^{\text{mask in}}\overbrace{S_{C,r,j,k}(d_0)}^{\text{resp. tract}}\overbrace{S_{C,m,out,j,k}(d_0)}^{\text{mask out}}\right]}^{\text{total filtering efficiency}} \\ &= \frac{N_C}{V}\left\langle\lambda_C(t)\left\{1-\left[1-E_{C,m,in,j}\left(w(d_0,t)d_0\right)\right]\right.\right. \\ &\quad \left.\left.\bullet\left[1-E_{C,r,j}(d_0)\right]\left[1-E_{C,m,out,j}(d_0)\right]\right\}\right\rangle_C \quad , \end{aligned} \tag{42}$$

where the j subscript has been dropped in the average over category C. As was the case before with the average of a product, only terms that are uncorrelated with the others can be pulled out or be replaced by their average value inside.

Flux: Inactivation

When a pathogen in an aerosol with multiplicity k inactivates, the aerosol's multiplicity changes to $k-1$. We will model inactivation of pathogen copies as exponential decay with inactivation rate $\gamma(t)$, which might depend on time (e.g. dependence on UV light intensity, humidity, etc. that could be fluctuating in time). For aerosols with a multiplicity of k, the volume normalized loss rate to multiplicity $k-1$ is just

$$f_{k,k-1}(t)n_k(d_0,t) = k\gamma(t)n_k(d_0,t) \quad . \tag{43}$$

Two pathogen copies will never inactivate at exactly the same time; so we don't have to consider flux terms beyond the two neighboring multiplicities.

General Concentration Density Equations

All of the sources, sinks, and flux terms can be collected to make the system of total differential equations describing the infectious aerosol concentration density, which is

$$\frac{dn_k}{dt} = -\alpha(d_0,t)n_k + f_{k+1,k}(d_0,t)n_{k+1} - f_{k,k-1}(d_0,t)n_k + \beta_k(d_0,t) \quad . \tag{44}$$

We have assumed that shattering and collisional coalescence of infectious aerosols, whether from turbulent induced collisions or differential gravitational settling, is negligible. Collisional coalescence could begin to be important if there are a significant number of very large aerosols and/or n_k is very large. Particularly, $d > 100$ µm aerosols/droplets, even though they will generally settle to the ground/floor before evaporating to their equilibrium diameter [21, 22], can capture smaller aerosols on their way to the ground/floor [18–20]. This will generally be negligible unless individuals are situated in the environment such that the large aerosols exhaled by one person (who need not be infectious) will fall through the exhaled aerosol plume of an infectious individual , and potentially negligible even then. If the aerosol concentration, including non-infectious aerosols, reach the levels seen in atmospheric clouds, collisional

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

coalescence might also have to be considered along with keeping track of $k = 0$ aerosols; though this is very unlikely in indoor environments except and when there is a lot of smoke or artificial fog machines are in use, like in a discotheque or theater.

Then, putting the flux terms into Eq (44), we have the following system of ODEs to get the concentration density

$$\frac{dn_k}{dt} = -\alpha(d_0, t)n_k + (k+1)\,\gamma(t)n_{k+1} - k\gamma(t)n_k + \beta_k(d_0, t) \quad . \tag{45}$$

Luckily this is a system of ODEs rather than PDEs with flux terms in diameter (involving derivatives with respect to diameters). This is the advantage of choosing d_0 or d_D instead of d_e. For practical applications, this also means that we can also split the diameter range into bins and solve it for each bin separately since there are no flux terms between bins. (See S3 Appendix for how to bin the model with respect to diameter.)

This is a linear inhomogeneous finite system of coupled ODEs at each d_0. The number of equations in the system is finite since k is non-negative and there is the maximum theoretical multiplicity M. Moreover, we don't even need to care about $k = 0$ since those aerosols are no longer an infection hazard. Additionally, the system that needs to be solved is smaller if $M_c < M$. If $M_c = 1$, then we have only one ODE. This situation occurs if the pathogen load of respiratory tract fluid is low enough that very few aerosols have 2 or more pathogen copies in them.

Note that this model demonstrates superposition with respect to sources since it is linear, as expected intuitively — each aerosol is independent of all others, therefore the response (concentration density and expected dose) from each individual source is independent of all other sources. If $n_{k,1}$ and $n_{k,2}$ are solutions for the same α and γ but different sources $\beta_{k,1}$ and $\beta_{k,2}$ respectively, then the solution for $\beta_k = \beta_{k,1} + \beta_{k,2}$ is $n_k = n_{k,1} + n_{k,2}$.

Infection Risk

Let $\mu_{j,k}$ be the average number of aerosols with multiplicity k absorbed by the j'th susceptible individual from time t_0 to time t. At any particular instant of time, the average number of such aerosols of each original diameter d_0 entering the person's mask if they are wearing a mask or their mouth and nose if they aren't is $\lambda_{S,j}(t)n_k(d_0, t)$. A fraction $S_{S,m,in,j}(d_0, t)$ will survive the mask to enter the respiratory tract [2–5, 8, 10]. A fraction $E_{S,r,j}(d_0)$ of those survivors will be absorbed by the respiratory tract [2, 4, 6, 10], which contributes to the dose. The expected average **aerosol** dose is then the double integral of this over the d_0 and the time between t_0 and t, which is

$$\begin{aligned}
\mu_{j,k}(t) &= \int_{d_{m,k}}^{d_M} d\phi \int_{t_0}^{t} dv \; \overbrace{E_{S,r,j}(\phi)}^{\text{absorption efficiency}} \; \overbrace{S_{S,m,in,j}(\phi, t)}^{\text{survive mask}} \; \overbrace{\lambda_{S,j}(v)n_k(\phi, v)}^{\text{inf. aerosol inhalation rate}} \\
&= \int_{d_{m,k}}^{d_M} d\phi \int_{t_0}^{t} dv \, E_{S,r,j}(\phi)\left[1 - E_{S,m,in,j}\left(w(\phi, v)\phi\right)\right]\lambda_{S,j}(v)n_k(\phi, v),
\end{aligned} \tag{46}$$

where we have ϕ as the integration variable over d_0. We will continue to use ϕ exclusively for this purpose in the rest of the manuscript.

In order to use the $\mu_{j,k}$ in the multiplicity-corrected dose-response model for the particular disease of interest $\mathfrak{R}$, we need to first assume that the **aerosol** dose for each multiplicity follows a Poisson distribution with $\mu_{j,k}$ as the means and that each is independent of each other (no correlations). This requires the well-mixed assumption like many other parts of the model.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

But it also requires that the effect of turbulent inertial clustering is negligible. We will now show that it is negligible except possibly at extremely high aerosol concentrations. It will be negligible if the aerosol Stokes numbers $St = \tau_p/\tau_\eta$ are very small ($St \ll 1$) [25, 26] where τ_p is the aerosol inertial response time scale from Eq (29) and τ_η is the Kolmogorov time scale of the turbulence in the environment, which is $\tau_\eta = \sqrt{\nu_a/\epsilon}$ where ϵ is the turbulent dissipation rate. It will also be small if the typical inter-aerosol distance $\bar{d}_a \sim \mathcal{N}^{-1/3}$, where $\mathcal{N}$ is the total infectious aerosol concentration for all d_0 and k, is much larger than the typical scale of turbulent inertial clustering (i.e. the fraction of aerosols with a neighbor in the clustering range is low). The typical scale of turbulent inertial clustering is about 10η [25, 26] where $\eta = (\nu_a^3/\epsilon)^{1/4}$ is the Kolmogorov length scale of the turbulence. This means that as long as $St \ll 1$ and/or $\mathcal{N}^{-1/3} \gg 10\eta$, the deviations of the aerosol doses from independent Poisson distributions will be negligible. The situation will be worst for the largest $w(d_M, t)d_M$ sized aerosols in high enough humidity that $w(d_M, t) \approx 1$. For a low dissipation rate of $\epsilon = 1$ mW kg^{-1}; $St = 0.06$ for a d_M sized aerosol and the number density limit is $\mathcal{N} \ll 4 \times 10^5$ m^{-3}. The Stokes number is small, so the turbulent inertial clustering's effect will be small even if $\mathcal{N}$ exceeded that limit. For a higher dissipation rate of $\epsilon = 1$ W kg^{-1}; $St = 2.0$ for a d_M sized aerosol and the number density limit is $\mathcal{N} \ll 7 \times 10^7$ m^{-3}. While the Stokes number is large, the number density limit is very high so turbulent inertial clustering's effect will generally be small. For a high for indoors dissipation rate of $\epsilon = 10$ W kg^{-1}; $St = 6.3$ for a d_M sized aerosol and the number density limit is $\mathcal{N} \ll 4 \times 10^8$ m^{-3}. While the Stokes number is large, the number density limit is very high so turbulent inertial clustering's effect will generally be small. Thus, turbulent inertial clustering will have a negligible effect on the Poissonity and independence of the aerosol dose distributions except possibly at extraordinarily high aerosol concentrations.

Model Solution and Simplification

General

There is an analytical solution to Eq (45), though it is not closed form unless the time dependence of α, β, and γ allow it. Eq (45) can be rewritten in matrix-vector form as

$$\frac{d\vec{n}}{dt} = \mathbf{A}(d_0, t)\vec{n}(d_0, t) + \vec{\beta}(d_0, t) \quad , \tag{47}$$

where $\vec{n}(d_0, t)$ and $\vec{\beta}(t)$ are the $n_k(d_0 t)$ and $\beta_k(d_0, t)$ for $k > 0$ in vector form and

$$\mathbf{A} \equiv \begin{bmatrix} -\alpha(d_0,t) - \gamma(t) & 2\gamma(t) & & & \\ & -\alpha(d_0,t) - 2\gamma(t) & 3\gamma(t) & & \\ & & \ddots & \ddots & \\ & & & \ddots & M_c\gamma(t) \\ & & & & -\alpha(d_0,t) - M_c\gamma(t) \end{bmatrix} \quad . \tag{48}$$

is an upper bidiagonal $M_c \times M_c$ square matrix. For any fixed d_0 or bin of d_0, the resulting system of ODEs is particularly amenable to efficient numerical solution even for very large M_c because A is sparse with only one or two elements per row.

The general solution in matrix-vector form, shown in S1 Appendix, is

$$\vec{n}(d_0, t) = \exp\left[\int_{t_0}^{t} A(d_0, x)dx\right]\vec{n}_0(d_0) + \int_{t_0}^{t} \exp\left[\int_{s}^{t} A(d_0, x)dx\right]\vec{\beta}(d_0, s)ds \quad . \tag{49}$$

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Working this out using the structure of the diagonalization of $\mathbf{A}$ in S1 Appendix, the general solution for each k is

$$
\begin{aligned}
n_k(d_0,t) = \exp\left[-\int_{t_0}^{t}\alpha(d_0,x)dx\right]\exp\left[-k\int_{t_0}^{t}\gamma(x)dx\right] \\
\bullet \sum_{p=k}^{M_c}\binom{p}{k} n_{0,p}(d_0)\left[1-\exp\left[-\int_{t_0}^{t}\gamma(x)dx\right]\right]^{p-k} \\
+\sum_{p=k}^{M_c}\binom{p}{k}\int_{t_0}^{t}\beta_p(d_0,s) \\
\bullet \exp\left[-\int_{s}^{t}\alpha(d_0,x)dx\right]\exp\left[-k\int_{s}^{t}\gamma(x)dx\right]\left[1-\exp\left[-\int_{s}^{t}\gamma(x)dx\right]\right]^{p-k} ds \quad , \qquad (50)
\end{aligned}
$$

where $\binom{k}{m} = k!/(m!(k-m)!)$ is the notation for the binomial coefficient k choose m.

Coefficients Constant in Time

We cannot go further in simplifying the general solution from Eq (50) without knowing the time dependence of α, $\vec{\beta}$, and γ. In many situations; α, $\vec{\beta}$, and γ are approximately constant with respect to time. If this is so; the general solution from Eq (50) and its time integral from t_0 to t (needed for the dose) become (see S1 Appendix)

$$
\begin{aligned}
n_k(d_0,t) &= n_{\infty,k} + z^s\left[U_k(d_0,\vec{\beta}(d_0),z) + V_k(\vec{n}_0(d_0),z)\right] \quad , \qquad (51) \\
\int_{t_0}^{t} n_k(d_0,v)dv &= (t-t_0)n_{\infty,k}(d_0) \\
&\quad -U_k(d_0,\vec{n}_0(d_0),1) + z^s U_k(d_0,\vec{n}_0(d_0),z) \\
&\quad -\frac{1}{\gamma}W_k\left(d_0,\vec{\beta},z\right) \quad , \qquad (52)
\end{aligned}
$$

where

$$
\begin{aligned}
z(t) &= e^{-(t-t_0)\gamma} \in (0,1] \quad , \qquad (53) \\
s(d_0) &= \frac{\alpha(d_0)}{\gamma} + k \quad , \qquad (54) \\
V_k(\vec{y},x) &= \sum_{i=k}^{M_c}\binom{i}{k} y_i (1-x)^{i-k} \quad , \qquad (55) \\
U_k(d_0,\vec{y},x) &= -\frac{1}{\gamma}\sum_{i=k}^{M_c}\binom{i}{k} y_i \sum_{p=0}^{i-k}\binom{i-k}{p}\frac{(-1)^p x^p}{s+p} \quad , \qquad (56) \\
W_k(d_0,\vec{y},x) &= \int_1^x dv\, v^{s-1} U_k(d_0,\vec{y},v) \quad , \qquad (57) \\
&= -\frac{1}{\gamma}\sum_{i=k}^{M_c}\binom{i}{k}\beta_i(d_0)\sum_{p=0}^{i-k}\binom{i-k}{p}\frac{(-1)^p\left(z^{s+p}-1\right)}{(s+p)^2} \quad , \qquad (58)
\end{aligned}
$$

and $n_{\infty,k}(d_0)$ is the concentration density as $t \to \infty$ which is

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

$$n_{\infty,k}(d_0) = -U_k\left(d_0, \vec{\beta}, 1\right) = \frac{1}{\gamma}\sum_{i=k}^{M_c}\binom{i}{k}\beta_i(d_0)\sum_{p=0}^{i-k}\binom{i-k}{p}\frac{(-1)^p}{s+p} \quad , \tag{59}$$

Note that s is a function of k and $e^{-(\alpha+k\gamma)(t-t_0)} = z^s$.

It is possible for $\lambda_{S,j}$ to be a function of t but α not be (i.e. there is cancelation). But if $\lambda_{S,j}$ and w are constant, the expected average **aerosol** dose of multiplicity k for the j'th susceptible individual in Eq (46) becomes

$$\mu_{j,k}(t) = \lambda_{S,j}\int_{d_{m,k}}^{d_M} d\phi E_{S,r,j}(d_0)\left(1 - E_{S,m,in,j}\left(w\phi\right)\right)\int_{t_0}^{t} n_k(\phi, v)dv \quad . \tag{60}$$

Calculation of $\vec{n}_k(d_0, t)$, $\vec{n}_{\infty,k}(d_0)$, $\int_{t_0}^{t} n_k(d_0, v)dv$ scales as $\mathcal{O}\left(M_c^3\right)$ due to there being M_c multiplicities and double sums in U_k and W_k that scale as M_c. There is a recursive solution for $\vec{n}_{\infty,k}(d_0)$ which is linear in M_c, and recursive solutions for all the U_k and W_k which are quadratic in M_c. Additionally, the recursive formulas don't require as much numerical precision in the intermediate steps to get a desired final precision as shown in S5 Appendix. From S1 Appendix, the recursive solutions start at $k = M_c$ and proceed downwards to $k = 1$. They are

$$U_k(d_0, \vec{y}, x) = \begin{cases} -\frac{y_{M_c}}{\gamma s} & \text{if } k = M_c \quad , \\ \frac{(k+1)x}{s}U_{k+1}(d_0, \vec{y}, x) - \frac{1}{\gamma s}V_k(\vec{y}, x) & \text{otherwise} \quad , \end{cases} \tag{61}$$

$$W_k(d_0, \vec{y}, x) = \begin{cases} \frac{y_{M_c}}{\gamma s^2}\left(1 - x^s\right) & \text{if } k = M_c \quad , \\ \frac{1}{s}\Big[(k+1)W_{k+1}\left(d_0, \vec{y}, x\right) & \\ \quad + x^s U_k\left(d_0, \vec{y}, x\right) - U_k\left(d_0, \vec{y}, 1\right)\Big] & \text{otherwise} \quad , \end{cases} \tag{62}$$

$$U_k\left(d_0, \vec{y}, 1\right) = \begin{cases} -\frac{y_{M_c}}{\gamma s} & \text{if } k = M_c \quad , \\ \frac{(k+1)x}{s}U_{k+1}(d_0, \vec{y}, 1) - \frac{y_k}{\gamma s} & \text{otherwise} \quad , \end{cases} \tag{63}$$

$$n_{\infty,k} = \begin{cases} \frac{\beta_{M_c}}{\gamma s} & \text{if } k = M_c \quad , \\ \frac{1}{\gamma s}\left[\beta_k + (k+1)\gamma\, n_{\infty,k+1}\right] & \text{otherwise} \quad . \end{cases} \tag{64}$$

$$\tag{65}$$

This recursive analytical solution for $\vec{n}$ is checked against a numerical solution of Eq (47) for a simple case and a very small time step in S2 Appendix. The relative differences for the simple case are very small at less than 10^{-12}. See S5 Appendix for numerical considerations for evaluating the analytical solutions on a computer or solving Eq (47) with a numerical ODE solver. The number of terms for both are discussed, as well as the required precision and maximum magnitude required for floating point numbers used to calculate the analytical solution formulas.

Determining The Cutoff Mc

In order to reduce the number of equations that have to be solved, we need to find a suitable cutoff $M_c < M$ if at all possible, whether for the whole diameter range or for each diameter bin (advantage of doing a separate one for each bin is that M_c tends to be small for the small diameter bins), such that the contribution of all higher multiplicities is less than a threshold $T \in (0, 1]$ fraction of the total contribution from all multiplicities. In many cases, this depends only on the $\rho_{p,j}$ of the infectious individuals and one can skip directly to Eq (74) for the value of M_c to use (shown in Fig 2 for a few

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

$\rho_{p,j}$). However, some cases such as when one starts the model after some number of infectious individuals have left the environment, when there is significant transport from other rooms, etc. require additional heuristics. These heuristics are developed below.

A cutoff is suitable if the total contribution for all $k > M_c$ to the average pathogen dose and therefore infection risk is small compared to the total contribution for $k \leq M_c$. It is almost always true that $M_c < M$, and in many cases it can even be $M_c = 1$. This depends on the distribution of exhaled aerosol sizes and the pathogen concentration ρ_p in the respiratory tract fluid where the aerosols are produced. For very low pathogen loading, one can use $M_c = 1$. Let d_- and d_+ be the bounds in d_0 of the bin (or whole range in which case $d_- = d_{m,1}$ and $d_+ = d_M$) being considered.

The most reliable way to determine M_c is to use the model with the cutoff M and determine M_c afterwards using the result, but that defeats the point of finding M_c since the effort one wants to save has already been expended. So we need heuristics to determine M_c ahead of time. All of them consider the dose contribution from high multiplicity aerosols and consider a simplified $k\mu_{j,k}$ from Eq (46) with a particular concentration density multiplied by the average absorption efficiency of susceptible individuals. For each heuristic, we will define this parameter to be $\mathcal{H}_{h,k}(t)$ where the h denotes the particular heuristic. Then, the heuristic for M_c is that we must find the M_c such that

$$\sum_{k=1}^{M_c} \mathcal{H}_{h,k}(t) \gg \sum_{k=M_c+1}^{\infty} \mathcal{H}_{h,k}(t) \;\; \forall \;\; h, t \geq t_0 \quad . \tag{66}$$

Note that we must take the largest M_c out of the values suggested by the individual heuristics.

An equivalent way to express this heuristic is to look at the ratio of the sum of $\mathcal{H}_{h,k}$ after the cutoff ($k > M_c$) to the total, defined as

$$J_{h,M_c}(t) \equiv \frac{\sum_{k=M_c+1}^{\infty} \mathcal{H}_{h,k}(t)}{\sum_{k=1}^{\infty} \mathcal{H}_{h,k}(t)} \quad . \tag{67}$$

Now, $J_{h,M_c}(t) \in [0,1]$ and is approximately the ratio of the contribution of the higher multiplicities $k > M_c$ aerosols to the total, which we want to be small. An equivalent statement of the heuristics is that one must find the M_c such that $J_{h,M_c} \ll 1 \; \forall \; h, t \geq t_0$. One way to determine M_c is to say pick some threshold $T \in (0,1]$, and then find the smallest M_c such that $J_{h,M_c} \leq T$ for all heuristics. Let $M_{c,h}(T)$ be the smallest value of M_c that satisfies $J_{h,M_{c,h}}(t) \leq T$, which makes it the single heuristic value of M_c. Then, M_c is just the maximum $M_{c,h}$.

First, we define the average absorption efficiency of the susceptible individuals as

$$A_S(d_0, t) \equiv \langle E_{S,r}(d_0) \left[1 - E_{S,m,in}\left(w(d_0,t)d_0\right)\right] \rangle_S \quad . \tag{68}$$

If the α, β, γ, and w are constant in time; it is a lot less effort to calculate $n_{\infty,k}(d_0)$ using Eq (59) than $n_k(d_0)$. Then, each $\mu_{j,k} \sim A_S n_{\infty,k}$. If $q_r(t)$ and $n_{r,k}(d_0,t)$ are non-zero, the doses from them have a similar scaling. If the initial concentration density includes a lot of aerosols with high multiplicities, we will need to set M_c to be large enough to include them even if they won't matter after the initial time. We need to consider this if $n_{0,k} \gg n_{\infty,k}$ for any $k > 1$, and they will have a similar scaling. These heuristics are

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

$$\mathcal{H}_{\infty,k} = k\int_{d_-}^{d_+} A_S(\phi,t)n_{\infty,k}(\phi)d\phi \quad , \tag{69}$$

$$\mathcal{H}_{r,k}(t) = k\int_{d_-}^{d_+} A_S(\phi,t)n_{r,k}(\phi,t)d\phi \quad , \tag{70}$$

$$\mathcal{H}_{0,k}(t) = k\int_{d_-}^{d_+} A_S(\phi,t)n_{0,k}(\phi)d\phi \quad . \tag{71}$$

The last heuristic is similar but considers the infectious individuals inside the environment instead of the concentration density. This has the advantage of not needing to determine $n_{\infty,k}(d_0)$. We essentially take the average over the d_0 interval of $\beta_{I,k}(d_0)$ from Eq (36) times the absorption efficiency of the average susceptible individual. We thus define the infectious individuals heuristic parameter

$$\mathcal{H}_{I,k}(t) \equiv k\int_{d_-}^{d_+} d\phi\, A_S(\phi,t)\sum_{j=1}^{N_I}\lambda_{i,j}(t)n_{I,j,k}(\phi,t)\left[1-E_{I,m,out,j}(\phi)\right] \quad . \tag{72}$$

But there are practical difficulties in using it directly. So instead, we will define the heuristic for each individual infectious individual using the largest diameter in the range d_+, and one would use the maximum M_c indicated by all of these. This has the advantage that there is a simple form for the required M_c, which is derived in S4 Appendix. It is

$$M_{c,I,j}(d_+,T) = 1 + C_P^{-1}\left(\langle k\rangle(d_+,t)_j\,,(1-T)\,C_P\left(\langle k\rangle(d_+,t)_j\,,K_m(d_+)-1\right)\right) \quad , \tag{73}$$

where C_P is the CDF (Cumulative Distribution Function) of the Poisson distribution and $C_P^{-1}(\mu,c)$ is the inverse CDF to find the smallest k for which $C_P(\mu,k)\geq c$. Note that when $K_m(d_+)\gg 1$ and $K_m(d_+)\gg\langle k\rangle(d_+,t)_j\gg 1$, $C_P\left(\langle k\rangle(d_+,t)_j\,,K_m-1\right)\simeq 1$ and

$$M_{c,I,j}(d_+,T) \simeq 1 + C_P^{-1}\left(\langle k\rangle(d_+,t)_j\,,(1-T)\right) \quad . \tag{74}$$

When the assumptions don't apply, this will give an overestimation, so it is usable to get the value of M_c to use. It will just give a bigger value than necessary.

Fig 2 shows $M_{c,I,j}$ as a function of d_0 for several different $\rho_{p,j}$. Increasing $\rho_{p,j}$ approximately just shifts the curves for $M_{c,I,j}$ to the left on a log-scale. Notice the very strong effect of $\rho_{p,j}$ on M_c, with values a little under 7000 being required for the largest diameter bin for $\rho_p = 10^{11}$ cm^{-3} and a value of 2 being required for the same bin for $\rho_p = 10^6$ cm^{-3}. Since M_c increases with d_0, the vast majority of the effort to determine the concentration density and the infection risk will be spent on the largest bins except for small values of ρ_p.

Example for SARS-CoV-2 with High Viral Load

Room, People, and Filter Efficiencies

We consider a hypothetical example based on the ongoing SARS-CoV-2 pandemic — a poorly ventilated seminar room with two infectious individuals with SARS-CoV-2 at the very upper end of viral concentrations (viral load) and one of them continuously coughing. Let the room have volume $V = 200$ m^3 with a height of $h = 4$ m, with

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

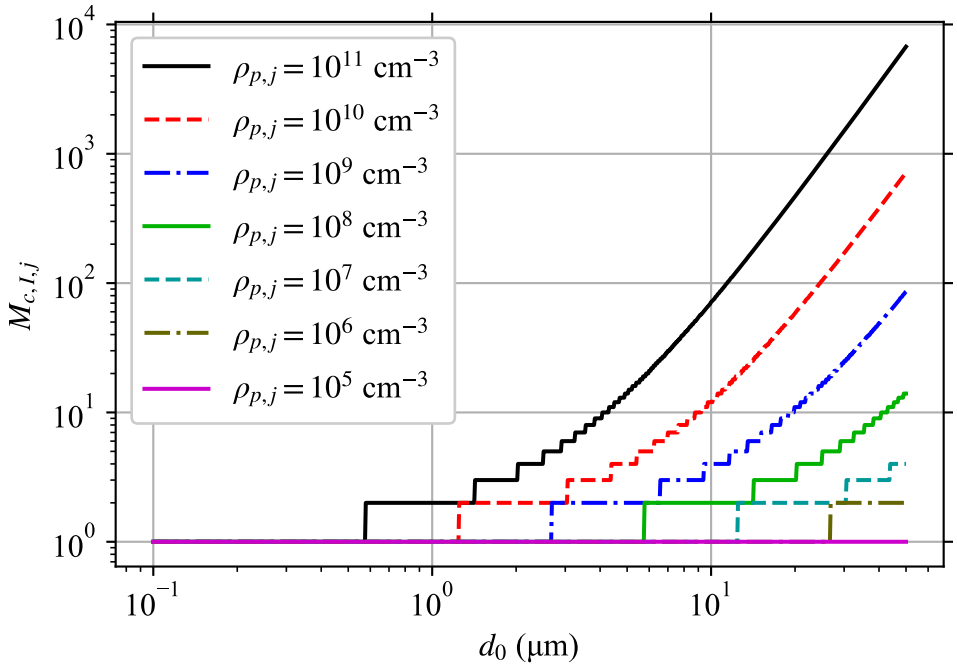


Fig 2. Required $\mathbf{M_c}$ Based on Pathogen Concentration in Infectious Individuals. $M_{c,I,j}$ required to capture 99% of pathogen production for each diameter at aerosol production d_0 from an infectious individual, with each line being a different pathogen concentration in their respiratory tract fluid $\rho_{p,j}$ (see legend).

ventilation $q_r = 0$, $q_v = 0$, and $q_o = 0.5$ hr^{-1}. We will ignore surface tension's effects on w. Let the humidity be such that the evaporation ratio is $w = \frac{1}{3}$, which is a constant with respect to both t and d_0. We ignore deposition ($\alpha_d = 0$). Let there be $N_S = 15$ susceptible individuals in groups of 5 wearing no mask, a simple1 mask, and a simple2 mask (defined later); and no non-infectious non-susceptible individuals ($N_O = 0$). The susceptible individuals will be assumed to be sedentary/passive adults with a breathing rate of $\lambda_{S,j} = 0.3$ m^3 hr^{-1}, which is in the range of mean breathing rates for this activity from the U.S. EPA's *Exposure Factors Handbook* Table 6.2 [27]. The pathogen concentration for SARS-CoV-2 varies widely across individuals, location in the body, and stage of the disease [15, 16, 28, 29], and can sometimes get as high as the 10^{10}–10^{11} cm^{-3} range [15, 16]. We will use this upper range because it makes the model more challenging to solve due to the larger M_c and due to the interest in so called "super-spreading events". The situation is composed of two stages (Stages 1 and 2) that each start when an infectious individual enters the room. Initially, there are no infectious aerosols in the room, meaning $n_{0,k}(d_0) = 0$. Stage 1; at $t = t_0 = 0$, one infectious individual enters the room who is speaking, wearing no mask, breathing at a rate $\lambda_{I,j} = 0.5$ m^3 hr^{-1} (just below an 0.54 m^3 hr^{-1} average value for reading out loud [30]), and has a high respiratory tract fluid pathogen concentration of $\rho_{p,j} = 10^{10}$ cm^{-3}. Stage 2; then at $t = 3$ hr, one more infectious individual enters the room who is continuously coughing while wearing a simple2 mask, breathing at a higher rate of $\lambda_{I,j} = 2.0$ m^3 hr^{-1}, and has a higher respiratory tract fluid pathogen concentration of $\rho_{p,j} = 10^{11}$ cm^{-3} at the very upper range for SARS-CoV-2. We chose this estimated continuous coughing breathing rate by deducing a breathing rate range from Hegland, Troche & Davenport [31] for continuous 3 cough cycles (heavily using their Fig 1), getting a breathing rate range of 1.9–2.3 m^3 hr^{-1} from which we chose

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

$2.0\ \mathrm{m}^3\,\mathrm{hr}^{-1}$.

We use mask filter efficiencies of the functional form

$$E_{C,m,in,j}(d) = E_{C,m,out,j}(d) = E_\infty - (E_\infty - E_0)\, e^{-d/D_{m,c}} \quad , \tag{75}$$

where E_∞ is the aerosol filtering efficiency as $d \to \infty$, E_0 is the aerosol filtering efficiency as $d \to 0$, and $D_{m,c}$ is the scale of the mask efficiency transition. We will use $D_{m,c} = 10$ µm. We consider individuals wearing no masks or one of two types of masks two masks. Their filtering efficiencies are

none (no mask) $E_0 = E_\infty = 0$.

mask simple1 $E_0 = 0.2$ and $E_\infty = 0.8$.

mask simple2 $E_0 = 0.95$ and $E_\infty = 0.99$.

The filtering efficiences of both the simple1 and simple2 masks are shown in S6 Fig. The mask parameters were chosen such that they are more efficient at filtering large aerosols/droplets than small ones, with the simple2 mask being better than the simple1 mask. The simple1 and simple2 masks could reasonably correspond to a reasonably well fitted home-made cloth mask and an excellently fitted FFP2 mask, though here we have treated their leak rate to be the same during inhalation as exhalation (not true with most real masks). At the largest sizes, leakage doesn't matter as much since the aerosols are more ballistic. Let us assume that $E_{C,r,j}(d_0) = \frac{1}{2}$ for everyone.

Disease and Infectious Aerosol Production

We assume that an exponential-dose response model is the correct model to use for SARS-CoV-2 since the exponential model works better than the beta-Poisson model for two other human infecting corona viruses (SARS-CoV-1 and HCoV-229E) [14]. In absence of a good value to use for r, we use the same value of r as found for SARS-CoV-1 in mice which is $r = 2.45 \times 10^{-3}$ and the same value of r as found for HCoV-229E in humans which is $r = 5.39 \times 10^{-2}$ [14]. We use $\gamma = 0.64\ \mathrm{hr}^{-1}$ as the inactivation rate for SARS-CoV-2 [32].

We approximate the SARS-CoV-2 pathogen as a sphere with a diameter of 100 nm, which is close to the correct size and the rough shape with the surface proteins removed (actually an ellipsoid) [17]. We use the aerosol size distributions for speaking and coughing from Johnson *et al.* [33], but extrapolate them to smaller diameters (from 800 nm to 100 nm). This is used with Eq (34) and (35) to get the $\beta_{I,k}$ They are shown in the top-right panel of Fig 3. The aerosol size distributions have two peaks at approximately 2 µm and 100 µm. This puts d_M between the trough (between the two peaks) and the second larger diameter peak.

Concentration Densities and Infection Risk

We now find the infectious aerosol concentration densities and doses, and mean infection risks $\mathfrak{R}_E$. First, we split the diameter range between $d_{m,1} = 0.1$ µm and $d_M = 50$ µm into 20 logarithmically spaced bins; and determine the bin average values for the coefficients over each bin by integration following the scheme in S3 Appendix. The infectious individuals source parameters for the i'th bin, $\beta_{I,k}|_i$, are calculated numerically via Simpson's rule for integration with 1000 equal linear width sub-bins in each bin. The particular choice of the mask survival efficiency in Eq (75) and w being constant lets the other binning integrals be calculated analytically.

The model is solved for Stage 1 and then the final values used as initial values for Stage 2 because this makes it so that α and β_k are constant in time when solving the

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

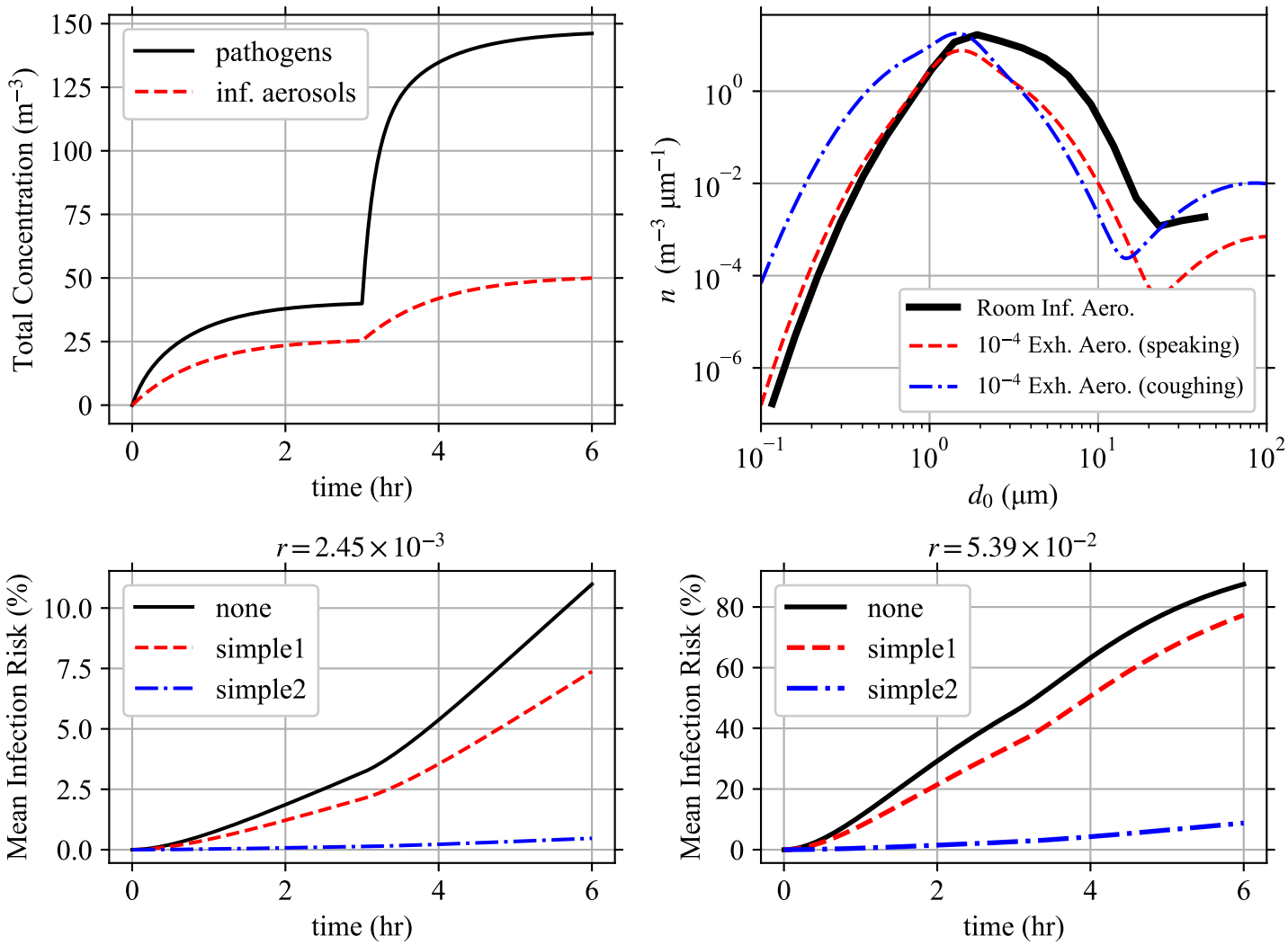


Fig 3. Model Solution for Example Solution to the example case. (Top-Left) The total pathogen and infectious aerosol concentrations over time. (Top-Right) The infectious aerosol concentration densities in the room as a function of d_0 at $t = 6$ hr compared to the aerosol concentration densities being exhaled by speaking and coughing individuals from Johnson *et al.* [33] scaled by 10^{-4} to make them have a comparable magnitude. (Bottom-Left, Bottom-Right) The mean infection risk $\mathfrak{R}_E$ for the susceptible individuals based on the mask they are wearing (none, simple1, or simple2) using the same (Bottom-Left) $r = 2.45 \times 10^{-3}$ (Bottom-Right) $r = 5.39 \times 10^{-2}$.

model (all changes are between stages). For M_c, we used the maximum value of $M_{c,I,j}$ for each infectious individual present at each Stage with $T = 10^{-3}$. Note that M_c stayed the same or increased for each bin going from Stage 1 to Stage 2 with the addition of one more infectious individual.

For the i'th bin, the $n_k|_i(t)$ and $\mu_{j,k}|_i(t)$ are solved analytically if $M_c \leq 500$ using the recursive solution and numerically if $M_c > 500$, both in IEEE-754 `binary64` floating point (also known as double precision and `float64`). This threshold between analytical and numerical solving was chosen to use the analytical solution as much as possible without overflow in V_k (see S5 Appendix). As shown in S5 Appendix, `binary64` numbers provide sufficient precision and allowed maximum magnitude. Note that overflow is easy to spot as infinities, which were not seen so this number format was

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

sufficient to prevent overflow. When doing it numerically, Eq (47) is solved using Runge-Kutta 4 with a time step of 10^{-4} hr, which is required for stability and an accurate solution with the large $\alpha|_i + M_c\gamma$ values in the largest bin. The time integral of n_k to get the dose is done by doing the cumulative sum of the integrand over the time steps and multiplying by the time step. The main results are shown in Fig 3.

The total pathogen concentration is slightly less than double the infectious aerosol concentration in Stage 1, and slightly higher than double in Stage 2. This means that the average multiplicity in both stages is approximately two, and it increases slightly from Stage 1 to Stage 2 which is expected with the higher viral load in the second infectious individual. Also, as expected, increasing r (infection risk of each individual pathogen) increases the infection risk. As expected, susceptible individuals wearing masks decrease their infection risk and increasing exposure increases their infection risk.

Comparing the infectious aerosol concentration density in the room with the aerosol concentration densities exhaled by the infectious individuals as a function of d_0 (see top-right panel of Fig 3); we can see how as d_0 increases, the probability of an aerosol being infectious increases (infectious aerosol concentration density decreases slower after the first peak than the exhaled aerosol concentration densities) but at the largest $d_0 > 15$ µm the increasing α due to stronger gravitational settling causes the infectious aerosol concentration density to grow slower after the trough than the exhaled aerosol concentration densities from the infectious individuals (including the speaking individual who is not wearing a mask). To see the latter, the strengths of the sinks α and total sinks $\alpha + k\gamma$ are shown in Fig 4 and we can see that settling causes α to increase by over a factor of 10 from 100 nm to 50 µm. Fig 4 additionally shows the increase in the total sink strength for the largest multiplicities M_c being considered due to inactivation.

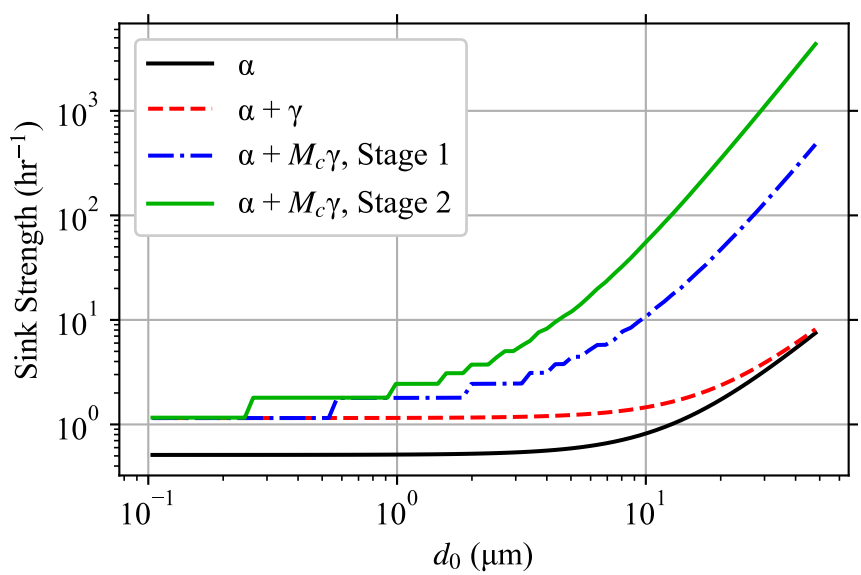


Fig 4. Sink Strength by Bin. The strength of the sink terms for each bin with 80 bins, which is α without inactivation, $\alpha + \gamma$ for $k = 1$, and $\alpha + M_c\gamma$ for $k = M_c$ (different values for Stage 1 and 2).

The pathogen concentrations as a function of d_0 and k right after the beginning and at the end of each Stage are shown in S7 Fig. For large diameters, the concentrations at the beginning of each Stage are initially in a narrow band around the expected multiplicity in each diameter bin but by the end of each Stage the distributions have widened downward as inactivation fills in the lower multiplicities.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

The results of choosing different numbers of bins (5, 20, and 80) is shown in S8 Fig. The difference in the concentration densities between 5 bins and 20 bins is substantial, but the difference between 20 and 80 is small. This means that in our example; for concentration densities, 20 bins is sufficient to capture the variation in $\alpha(d_0)$ and $\beta_k(d_0)$ with respect to diameter, but 5 is too few and 80 is a lot more effort for little gain. But for the $\mathfrak{R}_E$, the difference between the solutions for different number of bins is very small for the smaller $r = 2.45 \times 10^{-3}$, but more noticeable but still small for the larger $r = 5.39 \times 10^{-2}$.

Discussion

Effect of Multiplicity on Dose-Response

We consider a few hypothetical examples to ellucidate the importance of multiplicity in the dose-response using the corrected exponential model in Eq (12). Another dose-response model could be chosen and the resulting values would differ, but the general pattern would be the same.

First, let's reconsider the example case but with all pathogen production forced to be mono-multiplicity. We set the new $\beta_{1,new} = \sum_{k=1}^{M_c} k\beta_k$ and all other $\beta_{k,new} = 0 \; \forall \; k \neq 1$ and then set $M_c = 1$ for all bins. This is equivalent to going to each bin, taking the total aerosol volume production, finding the expected number of pathogen copies in that volume, and redistributing the volume so that each pathogen is alone in an aerosol but not changing d_0 anywhere. Or put equivalently, making Eq (47) track pathogen copies instead of aerosols and ignoring multiplicity. To quantify the difference, we took a simplified version of the example where the second coughing infectious individual was removed, the ρ_p of the first speaking infectious individual was adjusted, and we took the steady state case where $\vec{n}_0 = \vec{n}_\infty$ and calculated the constant $d\mu_{j,k}/dt$ for each susceptible individual. Then using the constant $d\mu_{j,k}/dt$ and an initial dose of zero, we found the time, τ_{50}, required for $\mathfrak{R}_E$ to be 50% (note that the particular choice does not matter, the curve is identical for any chosen risk). This was calculated for the 80 bins in diameter example to keep errors from finite bin width small, and a range of r values up to the maximum value $r = 1$. Ignoring multiplicity causes τ_{50} to be underestimated (overestimation of risk). The underestimate of τ_{50} is shown in Fig 5.

The underestimation increases with increasing ρ_p and r, and decreases when wearing a mask that is more efficient at filtering large aerosols than small aerosols. The largest aerosols have the greatest multiplicities, which means that a mask that filters them out better than small aerosols reduces the effect of ignoring multiplicity. As ρ_p increases, the expected multiplicity range for each d_0 increases which makes ignoring multiplicity underestimate τ_{50} more. For the r values considered here, $\rho_p \leq 10^9$ cm^{-3} underestimates τ_{50} by at most 20% and $\rho_p \leq 10^8$ cm^{-3} underestimates it by at most 12%. While for $\rho_p = 10^{11}$ cm^{-3}, the underestimation up to 67%. To better understand these patterns, we need to consider two more hypothetical situations.

Let the average pathogen dose be $\langle \Delta \rangle = r^{-1}$ and all infectious aerosols have the exact same multiplicity k. Then, the μ for all other multiplicities is zero and $\mu_k = \langle \Delta \rangle / k$. Essentially, we are dividing the same number of pathogen copies among fewer and fewer aerosols as we increase the number of pathogen copies in each one. The mean infection risk for this constant average dose is shown on the left side of Fig 6 as a function of k for four different r. As the multiplicity increases, the mean infection risk decreases even though the average dose is the same. For $k \ll r^{-1}$, the effect of multiplicity on $\mathfrak{R}_E$ is small. It starts to rapidly decrease near $k \sim r^{-1}$ and converges towards zero, because the number of pathogen copies in each aerosol is large enough that each aerosol has a high probability of causing infection by itself but the aerosols

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

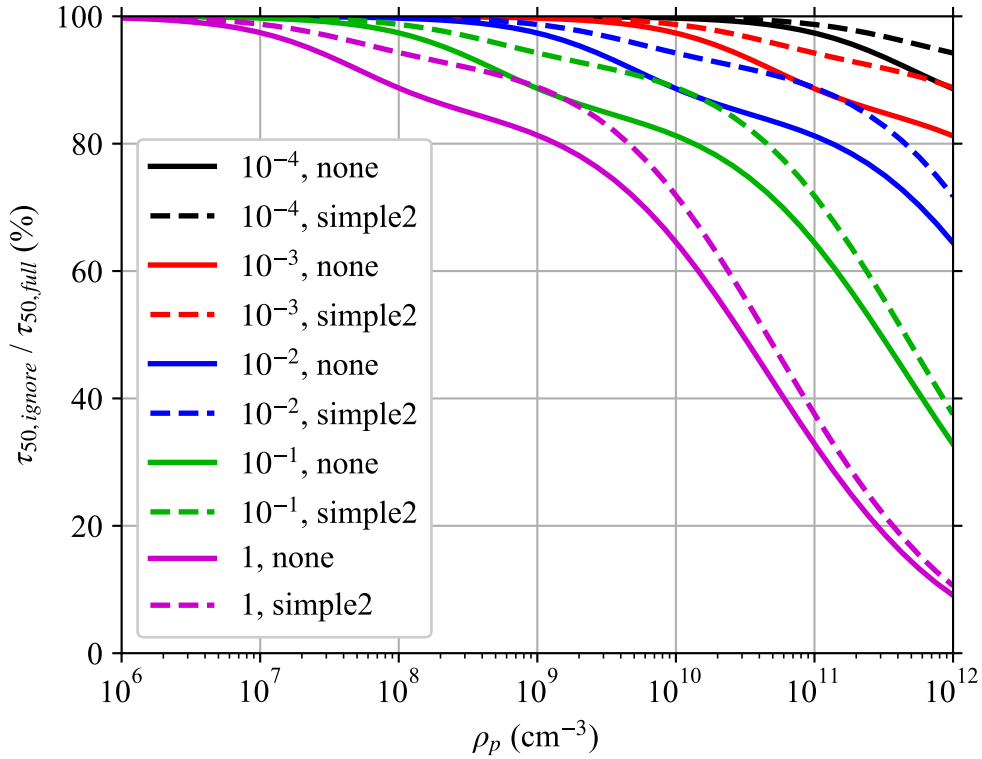


Fig 5. Effect of Ignoring Multiplicity, Full Version. Full version of Fig 1 with more ρ_p and the effect of masks. Plot of the ratio of the time required to reach a 50% infection risk when multiplicity is ignored $\tau_{50,ignore}$ to when it is fully accounted for $\tau_{50,full}$ for different pathogen ρ_p. We are considering the same situation as in the worked example, but at steady-state with just the speaking mask-less infectious individual and the risk to a susceptible individual whose exposure starts after steady state is reached. The ratio is shown for different combinations of mask on the susceptible individual (none and simple2) and for different r. The legend lists the r, mask combinations in the same order as the lines from top to bottom. We assumed a 100 nm diameter spherical pathogen and used 80 diameter bins and chose the M_c (maximum multiplicity considered) heuristic threshold to be $T = 0.01$ (include 99% of pathogen production).

are decreasing in number faster than the risk can increase. The risk per aerosol can't exceed 100% no matter how many pathogen copies are in an aerosol.

Another way to see this is to consider another hypothetical. Let's consider the mean infection risk if all aerosols have multiplicity k as we vary $r\langle\Delta\rangle$ for fixed r. This is shown on the right side of Fig 6 for $r = 10^{-2}$. For low $k \ll r^{-1}$, the infection risk curves are nearly identical. For $k \geq r^{-1}$, the infection risk decreases for increasing k.

Overall, this means that if the typical infectious aerosol multiplicity is on the order of or greater than r^{-1}, there can be a significant decrease in the infection probability for the same average dose. This has implications for large aerosols when the respiratory tract fluid pathogen concentration $\rho_{p,j}$ is large. Large aerosols where $\langle k\rangle \gtrsim r^{-1}$ will contribute less to the infection risk than would otherwise be expected from their resulting average pathogen dose $\langle\Delta_k\rangle$. While we must have $M_c > \langle k\rangle$, M_c is usable as a proxy for which diameters the multiplicity causes a substantial correction to the dose-response. If we were to consider $r = 2.45 \times 10^{-3}$ as was done in the example, Fig 2 shows that this would be important for $d_0 > 15$ µm for a high viral concentration of

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

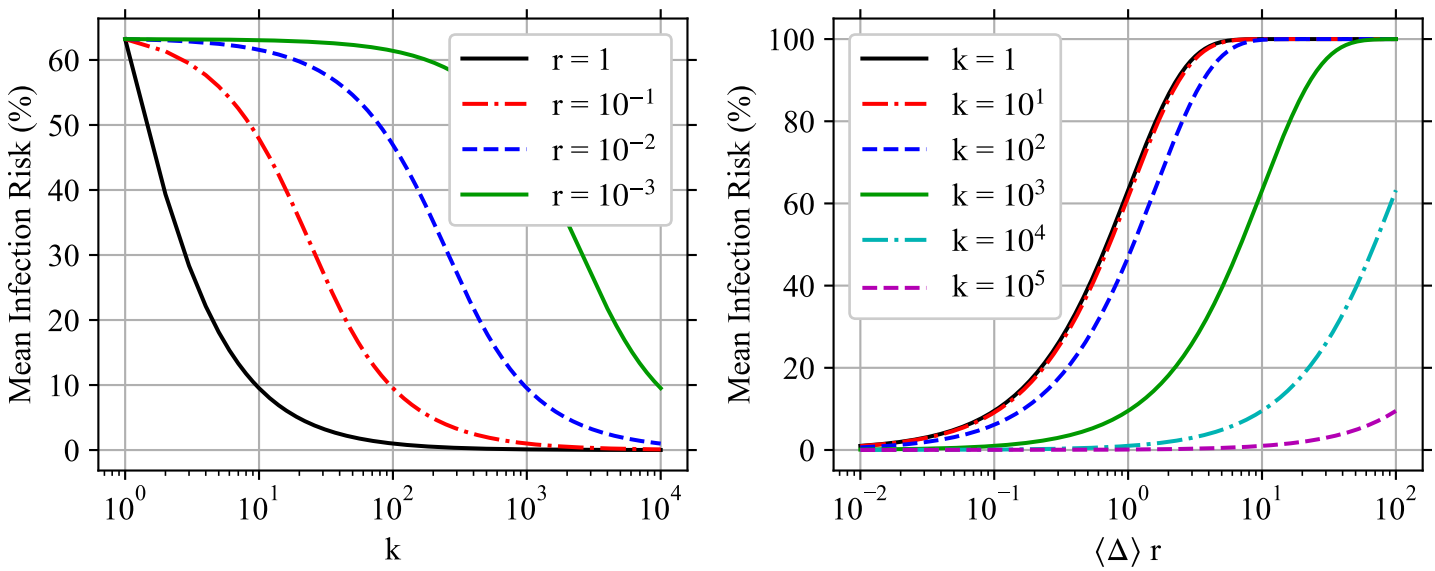


Fig 6. Multiplicity's Impact on Infection Risk. Plots of mean infection risk ($\mathfrak{R}_E$) using the modified exponential dose-response model when all infectious aerosols have the same number of pathogen copies in them k. (Left) The infection risk as a function of k for fixed average dose $\langle\Delta\rangle = r^{-1}$ for different single pathogen infection probabilities r. (Right) The infection risk as a function of the dose scaled by r ($\langle\Delta\rangle\, r$) for different k and the same fixed $r = 10^{-2}$ ($r^{-1} = 100$).

$\rho_{p,j} = 10^{11}$ cm^{-3} and $d_0 > 30$ for the lower but still high viral concentration of $\rho_{p,j} = 10^{10}$ cm^{-3}. If we were to consider $r = 5.39 \times 10^{-2}$ as was also done in the example, Fig 2 shows that this would be important for $d_0 > 5$ µm for a high viral concentration of $\rho_{p,j} = 10^{11}$ cm^{-3} and $d_0 > 10$ for the lower but still high viral concentration of $\rho_{p,j} = 10^{10}$ cm^{-3}.

Going back to the risk overestimation from ignoring multiplicity in Fig 5, decreasing r decreases the underestimation in τ_{50} because the ratio of the average multiplicity in the larger diameter bins to r^{-1} is smaller. A mask that filters large aerosols better than small aerosols reduces the effect of ignoring multiplicity because larger aerosols have higher multiplicities.

Filtering by The People

We introduced the sink terms $\alpha_{C,f}$ for filtering by the individuals in the environment as they inhale aerosols with many being absorbed by their mask or respiratory tract rather than being exhaled back out into the environment. To determine when this sink matters, we need to consider the total volume of air that is filtered and ignore the filtering efficiencies, and compare it to the ventilation. The volumetric rate of air filtration by the individuals normalized by the volume of the environment is

$$q_p(t) = \frac{1}{V}\left[\sum_{j=1}^{N_I}\lambda_{I,j}(t) + \sum_{j=1}^{N_S}\lambda_{S,j}(t) + \sum_{j=1}^{N_O}\lambda_{O,j}(t)\right] = \frac{\sigma_A}{\langle h\rangle}\,\langle\lambda_A\rangle_A \quad , \tag{76}$$

where σ_A is the horizontal area density of all individuals and $\langle h\rangle$ is the average height of the environment.

The mean adult breathing rates from sedentary/passive to high intensity activity ranges between 0.25 m^3 hr^{-1} and 3.2 m^3 hr^{-1} [27]. For sitting, it would be hard to get σ_A to be more than 1 m^{-2} but it would be possible while standing (some public events)

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

though the well-mixed assumption would be breaking down in either case. For a typical room height of $\langle h \rangle = 4$ m, this density limit would yield $\max(q_p) \in [0.063, 0.8]$ hr^{-1}. If the the environment is poorly ventilated (total ventilation rate $q_v + q_o + q_r$ less than 1 hr^{-1}), this high people density would mean the filtering effect of the people would not be negligible compared to the ventilation. But with even moderate ventilation, the contribution of $\alpha_{C,f}$ would be negligible unless all the ventilation is circulating ventilation ($q_o = q_r = 0$) with no filter or a very poor filter. For 1.5 and 2 m social distancing, the maximum σ_A are 0.14 and 0.080 m^{-2} respectively. For a typical room height of $\langle h \rangle = 4$ m, this density limit would yield $\max(q_p) \in [0.005, 0.11]$ hr^{-1} which would be negligible in almost all circumstances. For taller rooms, the contribution would be smaller if the total ventilation rate is held constant.

If the fraction of individuals who are infectious is held constant, then $N_I \sim \sigma_A$. Since $\beta_k \sim N_I$ and $\alpha_{C,f} \sim N_C$ but the non $\alpha_{C,f}$ terms of α stay constant, the source increases faster than the sinks meaning that n_k increases and therefore $\mathfrak{R}$ increases. So, increasing σ_A with everything else held constant increases the risk for the susceptible individuals. Thus, deliberately making $\alpha_{C,f}$ non-negligible is not a viable strategy to decrease risk. If the $\alpha_{C,f}$ dominate over the ventilation, the situation is actually quite hazardous from an infection transmission perspective. It is just that if one ignores the terms, one would overestimate the risk in such a crowded and poorly ventilated space.

Effect of Masks

The filtering effects of masks show up in the source $\beta_{I,k}$, the sinks $\alpha_{C,f}$, and the total dose over time $\mu_{j,k}$. Masks can substantially improve the total filtering efficiency of the people in $\alpha_{C,f}$ since aerosols have to pass through the mask twice, once on inhalation and again on exhalation at a larger diameter (many masks are better at filtering larger diameters than small diameters). But unless the ventilation is poor and there are a lot of people, this increase in $\alpha_{C,f}$ will have only a small effect on the total sink α. Instead, the main contribution is to reducing $\beta_{I,k}$ and $\mu_{j,k}$ which are both linearly proportional to the mask survival efficiency, which can be seen in the example situation.

In the example during Stage 1, there is one infectious individual in the room who is not wearing a mask and the total pathogen concentration reaches about 40 m^{-3} after 3 hr (Fig 3). During Stage 2, an addition infectious individual has entered the room. The second infectious individual's ρ_p is 10 times greater than the first person's and they are breathing at 4 times the rate; which would mean 40 times the pathogen exhalation rate by itself. Additionally, they are coughing rather than speaking, with the resulting larger exhaled aerosol concentration density ρ_j (top-right panel of Fig 3); which increases the number of exhaled pathogen copies further. But, they are wearing a mask which reduces the number of infectious aerosols that survive to reach the environment by a factor of 20–100 depending on the diameter. Due to this, the total pathogen concentration doesn't increase by a factor of over 40 but instead approximately triples to reaches approximately 140 m^{-3}.

The reduction in the average dose $\mu_{j,k}$ and therefore infection risk $\mathfrak{R}$ when susceptible individuals wear masks can also be seen in Fig 3. Even the simple1 mask gives some improvement, and the simple2 mask reduces the infection risk by over an order of magnitude.

Let's consider the case where all infectious individuals have the same mask survival efficiency and all susceptible individuals have the same mask survival efficiency. If the effects of masks on α is negligible ($\alpha_{C,f}$ is generally small compared to the other sinks) and $\beta_{r,k}$ is negligible; the combined effect of both infectious and susceptible individuals wearing masks on the dose is quadratic in the survival efficiencies, which has shown up in other Wells-Riley formulations in the past [4, 8]. Due to superposition of sources, $n_k \sim S_{I,m,out}$ since $\beta_{I,k} \sim S_{I,m,out}$. Then, $\mu_{j,k} \sim S_{S,m,in} n_k \sim S_{S,m,in} S_{I,m,out}$, which is

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

a quadratic term. Now $\alpha_{C,f} \sim S_{C,m,in} S_{C,m,out}$ makes the effect stronger (usually only slightly stronger) than quadratic since it only serves to increase α and therefore decrease n_k further. If just the susceptible or just the infectious individuals wear masks, the reduction drops to being stronger than linear (direct contribution of the mask on reducing $\beta_{I,k}$ or reducing $\mu_{j,k}$ plus the effect on $\alpha_{C,f}$). If only non-susceptible non-infectious individuals wear masks, there is still a reduction in the dose but it is small since $\alpha_{O,f}$ is generally small compared to the other sinks, giving a sublinear reduction.

Model Limitations

The biggest limitation to the model presented here, like all Wells-Riley formulations, is the well-mixed environment assumption. In almost all indoor environments, the assumption breaks down to varying degrees — the infectious aerosol concentration densities at the locations of susceptible individuals and all sinks (except possibly inactivation) depend on their locations in the room relative to the sources and the air flow. Social distancing helps with this assumption (reduces direct inhalation of undiluted exhaled puffs of aerosols from infectious individuals), but the assumption is still often dubious. Full fluid dynamics simulations with infectious aerosols simulated as passive scalars or as discrete aerosols such as those done by Löhner *et al.* [24] are the common way to address this limitation, which are considerably more difficult. Further investigation is needed to find simple approximate ways to generalize the Wells-Riley formulation presented in this manuscript for non-well-mixed environments that are easier than full fluid dynamics with suspended aerosols simulations.

Another limitation of the model presented here is that it assumes that all infectious aerosols have the same ζ and solute composition, and therefore the same $w(d_0, t)$. This is more easily circumvented in one case. If the solute concentration and composition is constant over time for each individual source (reasonable assumption over small time spans), the model can be solved for each source individually and then the resulting n_k and $\mu_{k,j}$ summed over the individual solutions. If ζ changes is not constant over time for the sources but the solute composition is constant, then one could generalize the model to additionally track ζ (or equivalently d_D) and initial diameter at production d_0 separately.

Another problem is the choice of diameter limits $d_0 \in [d_{m,k}, d_M]$ for each multiplicity. We have neglected the fact that the solute concentration is much greater for d_0 near the lower limit $d_{m,k}$ as pathogen copies are taking up a very large fraction of the volume and that surface effects may cause additional deviations in the number of pathogen copies in the aerosol from a Poisson distribution. Further work is needed to lift this limitation; though for small pathogens, the total fluid volume and therefore pathogen content in the smallest aerosols where this matters is much less than that of the larger aerosols (see top-right panel of Fig 3) meaning that the effect could be small for small pathogens.

The upper limit d_M is the cutoff where aerosols are so large that they are more ballistic and either settle to the ground before evaporating to equilibrium or still settle too quickly to be mixed even after evaporating to their equilibrium diameter. Based on Xie *et al.* [21] and Chong *et al.* [23], we suggested a value $d_M = 50$ µm. To look at it, we took the example case and re-calculated it for 23 equal log-width bins between 100 nm and 100 µm and considered the concentration densities and mean infection risks if the top 0, 2, and 4 bins were discarded, thereby setting decreasing d_M to 100 µm, 54.8 µm, and 30.1 µm. The time step for the numerical solution had to be reduced to 5×10^{-6} hr due to the increase in M_c at the larger d_M. This is shown in Fig 7. Increasing d_M increases the total pathogen concentration being tracked since a lot of volume is contained in the large diameter aerosols, but the total number concentration does not increase much since these big aerosols are few in number. For the larger $r = 5.39 \times 10^{-2}$, the effect on $\mathfrak{R}_E$ is very small as d_M is increased by a factor of

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

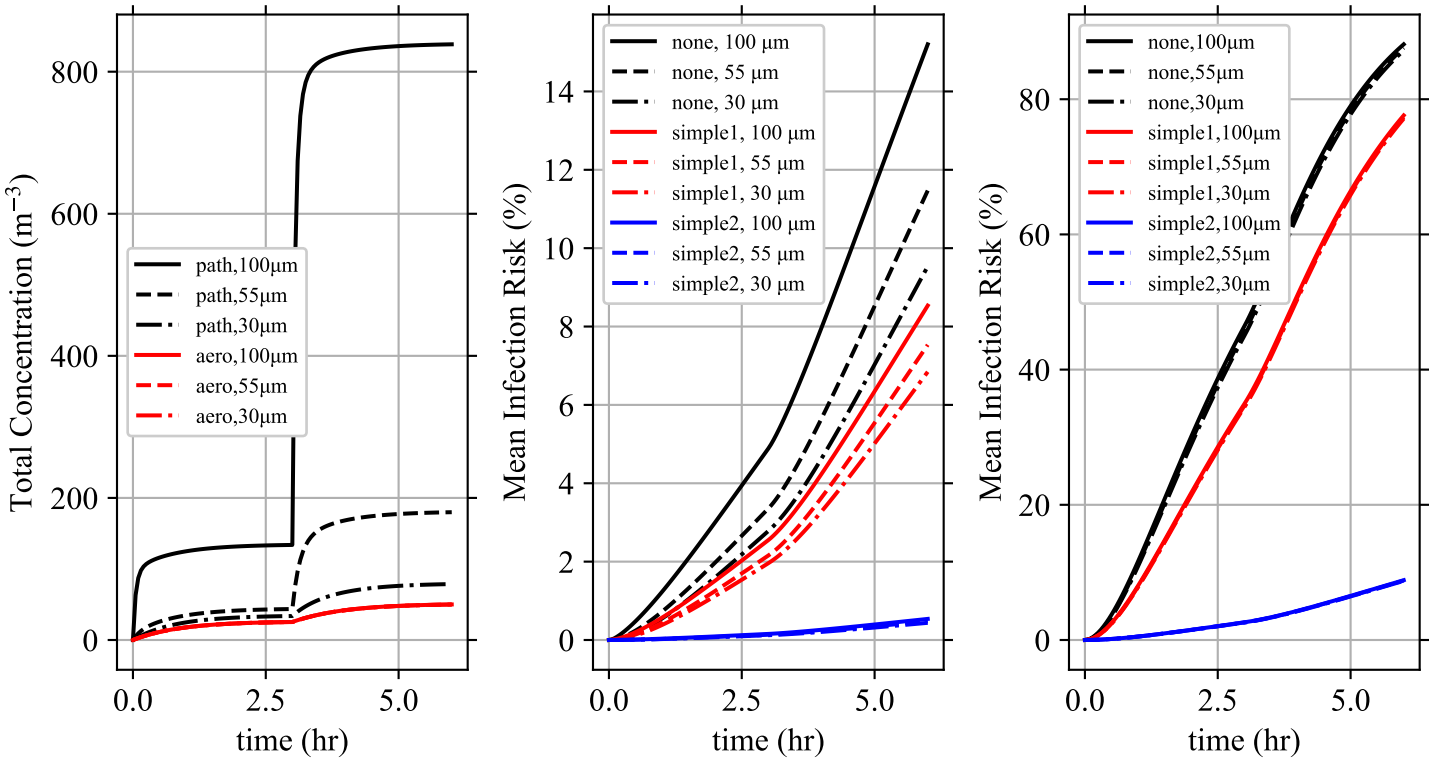


Fig 7. Effect of Upper Diameter Limit $\mathbf{d_M}$. The example situation was calculated for different values of the upper diameter limit d_M (technically, calculated at the largest and then truncated down as needed). (Left) The total pathogen and infectious aerosol concentration densities over time for each d_M. Note that the differences in the total infectious aerosol concentration density are so small that the lines are right on top of each other. The mean infection risk for each combination of masks on a susceptible individual (none, simple1, simple2) for (Middle) $r = 2.45 \times 10^{-3}$ and (Right) $r = 5.39 \times 10^{-2}$.

approximately three. But for the smaller $r = 2.45 \times 10^{-3}$, there is a larger fractional difference in the mean infection risk but the additive difference is no more than 5% for the worst case (no mask). The masks as we have defined them in the example, are better at filtering large particles than small, so they attenuate the effect of increasing d_M on $\mathfrak{R}_E$. More investigation is required on this upper diameter limit. Generalizing the model to track d and d_0 and treating evaporation/growth explicitly over time would help alleviate this problem as the high settling rates and the slower evaporation of the largest aerosols could be treated explicitly.

Conclusions

The number of pathogen copies in infectious aerosols must be taken into account if the number of pathogen copies in poly-multiplicity aerosols is not negligible compared to the number of pathogen copies in mono-multiplicity aerosols. We have generalized the Wells-Riley formulation and two common dose-response models (exponential and beta-Poisson) for poly-multiplicity aerosols and shown how to generalize other dose-response models. The generalized Wells-Riley formulation tracks infectious aerosols for each multiplicity individually rather than quanta as is traditional, which then can be put into the generalized dose-response model of choice. The generalized Wells-Riley formulation results in a linear inhomogeneous coupled system of ODEs, one for each multiplicity, at each initial aerosol diameter at production d_0 (or bin of d_0). The general

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

solution is presented; along with simplified versions for time independent sources, sinks, and humidity and splitting the diameter range into bins. The model is accompanied by an example case for for a poorly ventilated room with SARS-CoV-2, which is presented and solved. The example illustrates how the cutoff multiplicity M_c is determined, the effects of bin size on the solution, and the effects of mask usage on the infection risk. Additional takeaways are

- Ignoring multiplicity causes the infection risk to be over-estimated, which is particularly signficant for high respiratory tract fluid pathogen concentrations and high single-pathogen infection probabilities (see Fig 5).
- The people in the environment filter the air by breathing, which increases the loss rate for infectious aerosols and is included in the model.
- Facemasks on everyone cause a stronger than quadratic reduction in the inhaled dose by susceptible individuals

In summary, we have developed a tractable generalization of the Wells-Riley model for the infection risk from any airborne disease in well mixed indoor environments applicable to both mono- and poly-multiplicity aerosols.

Supporting information

S1 Appendix. Model Solution Derivation. Derivation of the general solution to Eq (47) as well as the constant in time coefficient special solution (both the explicit and recursive forms).

S2 Appendix. Checking Analytical Solution Against Numerical Solution. Checking the recursive analytical solution against solving the system of equations in Eq (47) numerically.

S3 Appendix. Binning Diameter. Shows how the model can be split into discrete diameter bins and each treated separately.

S4 Appendix. $\mathbf{M_c}$ Heuristic for Infectious People Derivation. Derivation of the the individual infectious individual production heuristic for M_c in Eq (73).

S5 Appendix. Numerical Considerations. Considerations for numerically evaluating the analytical model solution and solving the equations numerically; including how the number of terms scales with M_c and the magnitude and precision requirements to avoid numerical overflow and losing accuracy.

S6 Fig. Filtering Efficiencies of simple1 and simple2 Masks from Example. The filtering efficiencies of the simple1 and simple2 masks from the example, whose funcional forms are given by Eq (75), as a function of the diameter.

S7 Fig. Pathogen Concentration by k and Diameter for The Example Situation. The pathogen concentration in the room as a function of d_0 and k, denoted by color, at four different times (listed in the title of each panel) in the example situation. They are (Top-Left) right after the beginning of Stage 1, (Top-Right) at the end of Stage 1, (Bottom-Left) right after the beginning of Stage 2, and (Bottom-Right) at the end of Stage 2. All four panels share the same colorbar, which is in the bottom-right panel.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

S8 Fig. Comparing Different Numbers of Bins in The Model Solution's for The Example Situation. Version of Fig 3, but comparing the model solution for the example situation for 5, 20, and 80 bins. (Top-Left) The total pathogen and infectious aerosol concentrations over time for each number of diameter bins used to solve the model. (Top-Right) The infectious aerosol concentration densities as a function of d_0 at $t = 6$ hr for each number of bins. (Bottom-Left, Bottom-Right) The mean infection risk $\mathfrak{R}_E$ for the susceptible individuals based on the mask they are wearing (none, simple1, or simple2) for each number of bins using the same (Bottom-Left) $r = 2.45 \times 10^{-3}$ and (Bottom-Right) $r = 5.39 \times 10^{-2}$.

Acknowledgments

We would like to thank Oliver Schlenczek for important discussions early in the development of the model, Jan Moláček for comments and discussion during editing, and Hani Kaba and Simone Scheithauer for useful references and discussing those references.

Funding

This work has been partly funded by the BMBF as part of the B-FAST (Bundesweites Netzwerk Angewandte Surveillance und Teststrategie) project (01KX2021) within the NUM (Netzwerk Universitätsmedizin).

References

1. Riley EC, Murphy G, Riley RL. AIRBORNE SPREAD OF MEASLES IN A SUBURBAN ELEMENTARY SCHOOL. American Journal of Epidemiology. 1978;107(5):421–432. doi:10.1093/oxfordjournals.aje.a112560.

2. Nicas M. An Analytical Framework for Relating Dose, Risk, and Incidence: An Application to Occupational Tuberculosis Infection. Risk Analysis. 1996;16(4):527–538. doi:10.1111/j.1539-6924.1996.tb01098.x.

3. Gammaitoni L, Nucci MC. Using a Mathematical Model to Evaluate the Efficacy of TB Control Measures. Emerging Infectious Diseases. 1997;3:335–342. doi:10.3201/eid0303.970310.

4. Nazaroff WW, Nicas M, Miller SL. Framework for Evaluating Measures to Control Nosocomial Tuberculosis Transmission. Indoor Air. 1998;8(4):205–218. doi:10.1111/j.1600-0668.1998.00002.x.

5. Fennelly KP, Nardell EA. The Relative Efficacy of Respirators and Room Ventilation in Preventing Occupational Tuberculosis. Infection Control & Hospital Epidemiology. 1998;19(10):754–759. doi:10.2307/30141420.

6. Nicas M, Nazaroff WW, Hubbard A. Toward Understanding the Risk of Secondary Airborne Infection: Emission of Respirable Pathogens. Journal of Occupational and Environmental Hygiene. 2005;2:143–154. doi:10.1080/15459620590918466.

7. Miller SL, Nazaroff WW, Jimenez JL, Boerstra A, Buonanno G, Dancer SJ, et al. Transmission of SARS-CoV-2 by inhalation of respiratory aerosol in the Skagit Valley Chorale superspreading event. Indoor Air. 2020;00:1–10. doi:10.1111/ina.12751.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

8. Jimenez JL. COVID-19 Aerosol Transmission Estimator; 2020. Online, Version 3.4.19 released 2020 Oct. 27 [Cited 2020 Nov. 26]. Available from: `https://tinyurl.com/covid-estimator`.

9. Atkinson J, Chartier Y, Otaiza F, Pessoa-Silva CL, Ching P, Croome D, et al. Natural Ventilation for Infection Control in Health-Care Settings. World Health Organization; 2009. Available from: `https://www.who.int/water_sanitation_health/publications/natural_ventilation/en/`.

10. Sze To GN, Chao CYH. Review and comparison between the Wells–Riley and dose-response approaches to risk assessment of infectious respiratory diseases. Indoor Air. 2010;20(1):2–16. doi:10.1111/j.1600-0668.2009.00621.x.

11. Haas CN, Rose JB, Gerba CP. Quantitative Microbial Risk Assessment. 2nd ed. John Wiley and Sons, Ltd; 2014.

12. Noakes CJ, Beggs CB, Sleigh PA, Kerr KG. Modelling the transmission of airborne infections in enclosed spaces. Epidemiology and Infection. 2006;134(5):1082–1091. doi:10.1017/S0950268806005875.

13. Noakes CJ, Sleigh PA. Mathematical models for assessing the role of airflow on the risk of airborne infection in hospital wards. Journal of the Royal Society, Interface. 2009;6 Suppl 6:S791–S800. doi:10.1098/rsif.2009.0305.focus.

14. Watanabe T, Bartrand TA, Weir MH, Omura T, Haas CN. Development of a Dose-Response Model for SARS Coronavirus. Risk Analysis. 2010;30(7):1129–1138. doi:10.1111/j.1539-6924.2010.01427.x.

15. Blot M, Jacquier M, Manoha C, Piroth L, Charles PE, study group P. Alveolar SARS-CoV-2 Viral Load Is Tightly Correlated With Severity in COVID-19 ARDS. Clinical Infectious Diseases. 2020;doi:10.1093/cid/ciaa1172.

16. Pan Y, Zhang D, Yang P, Poon LLM, Wang Q. Viral load of SARS-CoV-2 in clinical samples. The Lancet Infectious Diseases. 2020;20(4):411–412. doi:10.1016/S1473-3099(20)30113-4.

17. Yao H, Song Y, Chen Y, Wu N, Xu J, Sun C, et al. Molecular Architecture of the SARS-CoV-2 Virus. Cell. 2020;183(3):730 – 738.e13. doi:10.1016/j.cell.2020.09.018.

18. Pruppacher HR, Klett JD. Microphysics of Clouds and Precipitation. 2nd ed. Dordrecht: Springer Netherlands; 2010.

19. Goody R. Principles of Atmospheric Physics and Chemistry. Oxford University Press; 1995.

20. Shaw RA. Particle-Turbulence Interactions in Atmospheric Clouds. Annual Review of Fluid Mechanics. 2003;35:183–227. doi:10.1146/annurev.fluid.35.101101.161125.

21. Xie X, Li Y, Chwang ATY, Ho PL, Seto WH. How far droplets can move in indoor environments–revisiting the Wells evaporation-falling curve. Indoor Air. 2007;17:211–225. doi:10.1111/j.1600-0668.2007.00469.x.

22. Wells WF. ON AIR-BORNE INFECTION*: STUDY II. DROPLETS AND DROPLET NUCLEI. American Journal of Epidemiology. 1934;20(3):611–618. doi:10.1093/oxfordjournals.aje.a118097.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

23. Chong KL, Ng CS, Hori N, Yang R, Verzicco R, Lohse D. Extended lifetime of respiratory droplets in a turbulent vapour puff and its implications on airborne disease transmission. medRxiv [Preprint]. 2020;medRxiv 2020.08.04.20168468v1. doi:10.1101/2020.08.04.20168468.

24. Löhner R, Antil H, Idelsohn S, Oñate E. Detailed simulation of viral propagation in the built environment. Computational Mechanics. 2020; p. 1–15. doi:10.1007/s00466-020-01881-7.

25. Saw EW, Salazar JPLC, Collins LR, Shaw RA. Spatial clustering of polydisperse inertial particles in turbulence: I. Comparing simulation with theory. New Journal of Physics. 2012;14(10):105030. doi:10.1088/1367-2630/14/10/105030.

26. Saw EW, Shaw RA, Salazar JPLC, Collins LR. Spatial clustering of polydisperse inertial particles in turbulence: II. Comparing simulation with experiment. New Journal of Physics. 2012;14(10):105031. doi:10.1088/1367-2630/14/10/105031.

27. U S EPA. Exposure Factors Handbook. 2011th ed. U.S. Environmental Protection Agency; 2011. Available from: `https://cfpub.epa.gov/ncea/risk/recordisplay.cfm?deid=236252`.

28. Jacot D, Greub G, Jaton K, Opota O. Viral load of SARS-CoV-2 across patients and compared to other respiratory viruses. Microbes and Infection. 2020;doi:10.1016/j.micinf.2020.08.004.

29. To KKW, Tsang OTY, Leung WS, Tam AR, Wu TC, Lung DC, et al. Temporal profiles of viral load in posterior oropharyngeal saliva samples and serum antibody responses during infection by SARS-CoV-2: an observational cohort study. The Lancet Infectious Diseases. 2020;20(5):565 – 574. doi:10.1016/S1473-3099(20)30196-1.

30. Binazzi B, Lanini B, Bianchi R, Romagnoli I, Nerini M, Gigliotti F, et al. Breathing pattern and kinematics in normal subjects during speech, singing and loud whispering. Acta Physiologica. 2006;186(3):233–246. doi:10.1111/j.1748-1716.2006.01529.x.

31. Hegland KW, Troche MS, Davenport PW. Cough expired volume and airflow rates during sequential induced cough. Frontiers in Physiology. 2013;4:167. doi:10.3389/fphys.2013.00167.

32. van Doremalen N, Bushmaker T, Morris DH, Holbrook MG, Gamble A, Williamson BN, et al. Aerosol and Surface Stability of SARS-CoV-2 as Compared with SARS-CoV-1. New England Journal of Medicine. 2020;382(16):1564–1567. doi:10.1056/NEJMc2004973.

33. Johnson GR, Morawska L, Ristovski ZD, Hargreaves M, Mengersen K, Chao CYH, et al. Modality of human expired aerosol size distributions. Journal of Aerosol Science. 2011;42(12):839–851. doi:10.1016/j.jaerosci.2011.07.009.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Risk assessment for airborne disease transmission by poly-pathogen aerosols

Freja Nordsiek[1], Eberhard Bodenschatz[1,2,3*], Gholamhossein Bagheri[1],

1 Max Planck Institute for Dynamics and Self-Organization (MPIDS), Göttingen, Niedersachsen, Germany
2 Institute for Dynamics of Complex Systems, University of Göttingen, Göttingen, Niedersachsen, Germany
3 Laboratory of Atomic and Solid State Physics and Sibley School of Mechanical and Aerospace Engineering, Cornell University, Ithaca, New York, USA

* lfpn-office@ds.mpg.de

S5 Appendix. Numerical Considerations

If α, β_k, γ, and w are all constant with respect to time; the model has both an explicit and recursive solution for the concentration density and dose. Otherwise, it may not be possible to get a closed form analytical solution from the general solution and one would need to solve the model numerically. Of course, even with an analytical solution, one can solve it numerically. There are a number of numerical pitfalls with both the analytical and numerical solutions, which arise as M_c becomes large.

For a numerical solution, the time step of integration must be small compared to the smallest time scale in the model. The smallest time scale could be the time scale on which α, β_k, γ, and/or w change over time; but for large enough M_c it will always be time scale of the total sink for $k = M_c$ aerosols. The total sink for n_k is $-\alpha(d_0, t) - k\gamma(t)$, which scales linearly in M_c for n_{M_c} when M_c is large enough. The time step δt must be $\delta t < (\alpha - M_c\gamma)^{-1}$. Thus the total number of timesteps $N_t \sim M_c$ for large M_c. Since the total number of terms scales linearly in M_c as long as $\partial n_k/\partial t$ is not calculated in vector-matrix form or $\mathbf{A}$ is stored in a sparse format, the total computational effort scales as $\mathcal{O}\left(M_c^2\right)$.

For the analytical solution for coefficients constant in time; there are different difficulties for the explicit solution and the recursive solution. The number of terms scales as $\mathcal{O}\left(M_c^3\right)$ in the explicit solution and $\mathcal{O}\left(M_c^2\right)$ in the recursive solution, since there are M_c equations and they have double and single sums respectively with each sum scaling as M_c.

The analytical solutions have additional difficulties — avoiding numerical overflow and maintaining accuracy. The problem is the binomial coefficients $\binom{i}{k}$ and $\binom{i-k}{p}$ where $i \to M_c$ and $p \in [0, i-k]$. They can be calculated naively by computing each factorial and then doing the multiplication and division ($k!$ will be the largest); or carefully with cancellation handled explicitly which overflows later. And then even if all the binomial coefficients don't overflow, their products and sums can still overflow (the explicit version is worse than the recursive version in this regard due to the products of binomial coefficients).

To see this, we will find the upper bound for $V_k\left(\vec{y}, x\right)$ in the recursive solution. Consider its formula

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

$$V_k\left(\vec{y},x\right)=\sum_{i=k}^{M_c}\binom{i}{k}y_i(1-x)^{i-k} \quad . \tag{1}$$

In the model, all the elements of $\vec{y}$, which is always either $\vec{\beta}$ or $\vec{n}_0$, are non-negative. Since $x\in[0,1]$, $1-x\in[0,1]$ and is therefore also non-negative. This means that all elements in the sum in V_k are non-negative, meaning $V_k\geq 0$ and thus we only need to find the upper bound to determine the risk of overflow (the lower bound, zero, is not a worry). Now, $y_i\leq\max\left(\vec{y}\right)$ for all i where $\max\left(\vec{y}\right)$ shall denote the maximum element of $\vec{y}$. Since $x\in[0,1]$, $0\leq(1-x)^{i-k}\leq 1$. Then we just need to get an upper bound for the binomial coefficients. The binomial coefficient $\binom{j}{m}=k!/m!(j-m)!$ is at its greatest value for fixed j when $m=\lfloor j/2\rfloor$ or $m=\lceil j/2\rceil$ where $\lfloor\cdot\rfloor$ and $\lceil\cdot\rceil$ denote the floor and ceil operators respectively. This value is

$$\binom{j}{\lfloor j/2\rfloor}=\binom{j}{\lceil j/2\rceil}=\frac{j!}{\lfloor j/2\rfloor!\lceil j/2\rceil!} \quad . \tag{2}$$

This will be maximized when $i=j=M_c$. Then, the upper bound for V_k is

$$V_k\left(\vec{y},x\right)\leq\sum_{i=k}^{M_c}\frac{M_c!}{\lfloor M_c/2\rfloor!\lceil M_c/2\rceil!}\max\left(\vec{y}\right)=\frac{(M_c+1-k)M_c!\max\left(\vec{y}\right)}{\lfloor M_c/2\rfloor!\lceil M_c/2\rceil!} \quad . \tag{3}$$

For any M_c, the largest upper bound for any of the V_k will be for when $k=1$, so the upper bound we need to worry about is

$$V_k\left(\vec{y},x\right)\leq\frac{M_cM_c!\max\left(\vec{y}\right)}{\lfloor M_c/2\rfloor!\lceil M_c/2\rceil!} \quad . \tag{4}$$

Now, the factorial function has the bounds [1]

$$\sqrt{2\pi}m^{m+\frac{1}{2}}e^{-m}\exp\left[\frac{1}{12m+1}\right]<m!<\sqrt{2\pi}m^{m+\frac{1}{2}}e^{-m}\exp\left[\frac{1}{12m}\right] \quad . \tag{5}$$

We will still have an upper bound for V_k if we use the factorial lower bounds in place of the factorials in the denominator of Eq (4) and the factorial upper bound for the factorial in the numerator. This leads to

$$V_k\left(\vec{y},x\right)<\frac{\max\left(\vec{y}\right)M_c^{M_c+\frac{3}{2}}\exp\left[\frac{1}{12M_c}+\left\lfloor\frac{M_c}{2}\right\rfloor+\left\lceil\frac{M_c}{2}\right\rceil\right]}{\sqrt{2\pi}\left\lfloor\frac{M_c}{2}\right\rfloor^{\left\lfloor\frac{M_c}{2}\right\rfloor+\frac{1}{2}}\left\lceil\frac{M_c}{2}\right\rceil^{\left\lceil\frac{M_c}{2}\right\rceil+\frac{1}{2}}\exp\left[M_c+\frac{1}{12\lfloor\frac{M_c}{2}\rfloor+1}+\frac{1}{12\lceil\frac{M_c}{2}\rceil+1}\right]} \quad . \tag{6}$$

If we take the $\log_2$, we get the value of the base-2 exponent. Taking the $\log_2$ of both sides,

$$\begin{aligned}\log_2 V_k\left(\vec{y},x\right)<{}&\log_2\left[\max\left(\vec{y}\right)\right]+\left(M_c+\frac{3}{2}\right)\log_2\left(M_c\right)\\&-\left(\left\lfloor\frac{M_c}{2}\right\rfloor+\frac{1}{2}\right)\log_2\left\lfloor\frac{M_c}{2}\right\rfloor-\left(\left\lceil\frac{M_c}{2}\right\rceil+\frac{1}{2}\right)\log_2\left\lceil\frac{M_c}{2}\right\rceil\\&+\left(\log_2 e\right)\left(\frac{1}{12M_c}-\frac{1}{12\lfloor\frac{M_c}{2}\rfloor+1}-\frac{1}{12\lceil\frac{M_c}{2}\rceil+1}\right)\\&-\frac{1}{2}-\frac{1}{2}\log_2\pi \quad ,\end{aligned} \tag{7}$$

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

where we have used the fact that $\lfloor M_c/2 \rfloor + \lceil M_c/2 \rceil = M_c$. To represent all the V_k for a particular M_c without overflowing, a floating point format's maximum supported base-2 exponent, *emax* (using the same notation as the IEEE 754 standard [2]), must be at least this value ($emax \geq \log_2(V_k)$). But, even if the value $\log_2(V_k)$ might not overflow with this minimum value, calculating the binomial coefficient when $\max(\vec{y}) < 1$ could overflow before the multiplication drops the magnitude. So for the minimum *emax*, we must replace the $\max(\vec{y})$ with $\max[1, \max(\vec{y})]$ and we get

$$\begin{aligned} emax(\vec{y}) \geq \Bigg\lceil & \log_2[\max[1, \max(\vec{y})]] + \left(M_c + \frac{3}{2}\right)\log_2(M_c) \\ & - \left(\left\lfloor \frac{M_c}{2} \right\rfloor + \frac{1}{2}\right)\log_2\left\lfloor \frac{M_c}{2} \right\rfloor - \left(\left\lceil \frac{M_c}{2} \right\rceil + \frac{1}{2}\right)\log_2\left\lceil \frac{M_c}{2} \right\rceil \\ & + (\log_2 e)\left(\frac{1}{12M_c} - \frac{1}{12\lfloor \frac{M_c}{2} \rfloor + 1} - \frac{1}{12\lceil \frac{M_c}{2} \rceil + 1}\right) \\ & - \frac{1}{2} - \frac{1}{2}\log_2 \pi \Bigg\rceil \quad . \end{aligned} \tag{8}$$

To see how much of an overestimate this is for *emax*, let's compare it to $\log_2 \max(V_k)$ for a few simple cases. First, lets compare it to $V_k\left(\vec{1}, 0\right)$ since $\vec{y} = \vec{1}$ has all elements equal to the maximum element, $x = 0$ maximizes $(1-x)^{i-k}$, and makes all terms in the sum integers. In Fig 1 (left panel), $\log_2 \max(V_k)$ is compared to *emax* from Eq (8) and the *emax* values for the four smallest IEEE-754-219 binary floating point formats [2]. We calculated $\max(V_k)$ using variable sized integers before being converting to multi-precision floating point numbers for taking the $\log_2$ with the largest supported exponent with the GNU Multiple Precision Floating-point Reliable Library (MPFR, see `https://www.mpfr.org`) and the GNU Multiple Precision Arithmetic Library (GMP, see `https://gmplib.org`) in Python using the gmpy2 package (`https://pypi.org/project/gmpy2`). The minimum *emax* from Eq (8) is only barely larger in a logarithmic sense than $\log_2 \max(V_k)$ except for small M_c, so it isn't an excessive overestimate of the required *emax*.

Second, we will compare it to typical $\vec{y}$ used in the model. We will base $\vec{y}$ on $\vec{\beta}_{I,k}$ for single infectious person but remove all of the environment parameters and person specific parameters except for the expected average multiplicity $\langle k \rangle_j = \frac{\pi}{6} d_0^3 \rho_{p,j}$. We shall use the vector

$$\psi_k = \frac{V}{\rho_j \lambda_{I,j} S_{I,m,out,j}} \beta_{I,k} = P_P\left(\langle k \rangle_j, k\right) \quad . \tag{9}$$

For a range of $\langle k \rangle_j$, M_c was determined with the single infectious person production heuristic $M_{c,I,j}$ for thresholds T of 10^{-1} and 10^{-9}. Then the $V_k\left(\vec{\psi}, 0\right)$ were calculated in binary floating point with a 256 bit mantissa, $emax = 32768$, and minimum exponent $emin = -32767$ with gmpy2, MPFR, and GMP as before. The binomial coefficients calculated exactly before conversion to floating point. To improve the accuracy, the sum in Eq (1) was done as a sorting sum where all the terms were sorted in a list, the two smallest terms removed, those terms added together and inserted back into list; and this process repeated till only a single term (the total sum) remained. The right panel of Fig 1 compares $\log_2 \max(V_k)$ against the *emax* calculated from Eq (8) and the maximum exponents of some of the smallest IEEE 754 binary floating point formats. We can see that Eq (8) overestimates the required *emax* by quite a bit in a logarithmic

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

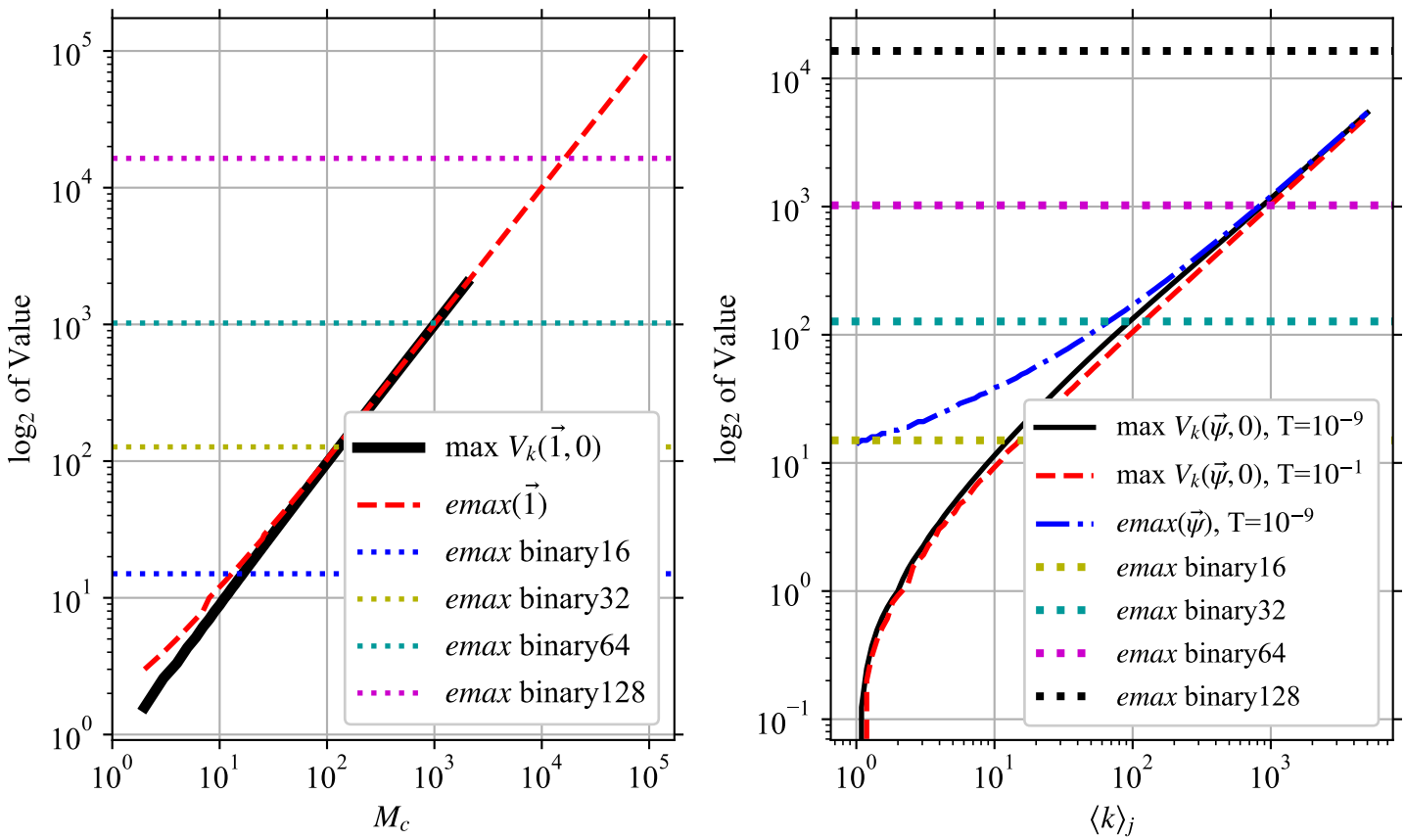


Fig 1. Comparison of Maximum $\mathbf{V_k}$ Values to Required emax Overestimate And Floating Point Format Limits. $\log_2 \max(V_k)$ is compared to the overestimate of the required $emax(\vec{y})$ from Eq (8) as well as the maximum exponents supported by the four smallest IEEE-754-2019 binary floating point formats [2]. Note that the x87 FPU 80-bit floating point format has the same maximum as `binary128` [3]. (Left) For $\vec{y} = \vec{1}$ and $x = 0$ as a function of M_c. (Right) For $\vec{y} = \vec{\psi}$ from Eq (9) and $x = 0$ as a function of the average multiplicity $\langle k \rangle_j = \frac{\pi}{6} d_0^3 \rho_{p,j}$.

sense for $\langle k \rangle_j < 10$, but not by much in a logarithmic sense for larger $\langle k \rangle_j$. So or small $\langle k \rangle_j$, it suggests an overkill *emax* but `binary16` (half precision) and `binary32` (single precision) are typically the smallest floating point formats with hardware support on most computers and the bulk of the computational effort is spent on diameter bins with the largest M_c, so there isn't much reason to use a smaller format for small $\langle k \rangle_j$. `binary64` (double precision) is sufficient for $\vec{\psi}$ this case up to $\langle k \rangle_j = 1000$. The exact values of other terms in $\vec{\beta}_I$ would determine if `binary64` would be safe for that expected multiplicity with $\vec{\beta}_I$. The extra *emax* requirement would additively increase by $\log_2 \left[\max\left(1, \frac{1}{V} \rho_j \lambda_{I,j} S_{I,m,out,j}\right)\right]$. `binary128` and x87 FPU 80-bit floating point numbers which have the same *emax* [3] would have some headroom for the magnitude of the other coefficients in $\vec{\beta}_I$ even for $\langle k \rangle_j = 10^4$.

To investigate the required floating point precision to calculate n_k, we will look at the required precision in bits required to calculate all $U_k(d_0, \vec{y}, x)$ for a given M_c to within a specified relative tolerance δ of the exact values. We chose

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

$$\begin{aligned} \alpha &= 1 \quad , \\ \gamma &= 3 \quad , \\ \vec{y} &= \vec{1} \quad , \\ x &\in \left\{0, \frac{5}{13}, 1\right\} \quad , \end{aligned}$$

which makes $U_k\left(d_0, \vec{y}, x\right)$ a rational number for any M_c. For several M_c, the $U_k\left(d_0, \vec{y}, x\right)$ were calculated exactly using variable sized rational arithmetic. Then, the smallest floating point precision was found for which the maximum relative error in the $U_k\left(d_0, \vec{y}, x\right)$ calculated in floating point in that precision (note, the binomial coefficents were calculated first as variable sized integers and then converted to floating point) is less than δ using the explicit and recursive formulas for $U_k\left(d_0, \vec{y}, x\right)$. The calculations were done using gmpy2, MPFR, and GMP as before. The required precisions for δ of 10%, 10 PPM (Parts per Million), and 1 PPB are shown in Fig 2 and compared to the precisions of various standard floating point formats.

Using the explicit formula, even quadruple precision (`binary128`) is insufficient by $Mc = 80$ for this $\vec{y}$ even for a tolerance of 10%. But using the recursive formula, double precision (`binary64`) is good enough for a tolerance of 1 PPB even for the largest $M_c = 5000$ that was checked. From this and the recursive solution's number of terms scaling as $\mathcal{O}\left(M_c^2\right)$ instead of $\mathcal{O}\left(M_c^3\right)$, the recursive solution is much more amenable to calculations than the explicit solution.

Considering both Fig 1 and 2, double precision (`binary64`) seems to be suitable, depending on the largest values in the $\vec{y}$, for moderate M_c of a few hundred. To go to M_c of a few thousand, quadruple precision (`binary128`) or x87 FPU 80-bit floating point numbers are required. Above this, either multi-precision floating point with higher exponents must be used or the model must be solved numerically.

The number of terms to compute scales quadratically in M_c for numerical solution and recursive analytical solution, and cubically for the explicit analytical solution; and the analytical solutions have the additional problems of avoiding numerical overflow as well as the computational effort to calculate some terms increasing with M_c. For any particular fixed size number format, there is an M_c above which the analytical result will overflow and one must solve the model numerically instead.

References

1. Robbins H. A Remark on Stirling's Formula. The American Mathematical Monthly. 1955;62(1):26–29. doi:10.2307/2308012.

2. IEEE Floating-Point Working Group. IEEE Standard for Floating-Point Arithmetic. IEEE Std 754-2019 (Revision of IEEE 754-2008). 2019; p. 1–84. doi:10.1109/IEEESTD.2019.8766229.

3. Intel. Intel 64 and IA-32 Architectures Software Developer's Manual, Volume 1: Basic Architecture; 2020. Available from: `https://software.intel.com/content/www/us/en/develop/articles/intel-sdm.html`.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

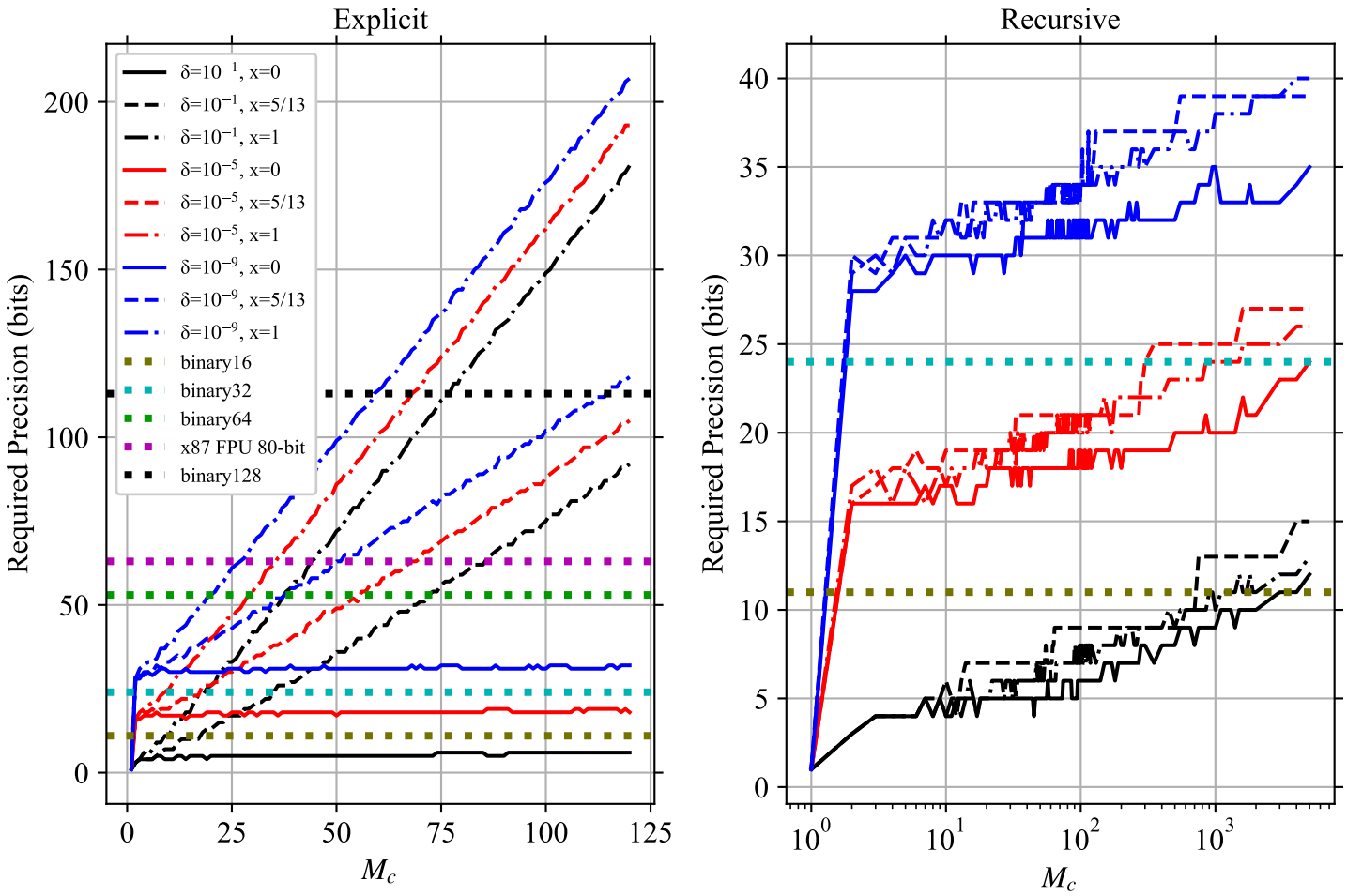


Fig 2. Required Floating Point Precision to Calculate $\mathbf{U_k}$. The required floating point precision in bits to get the largest relative error in the $U_k(d_0, \vec{y}, x)$ to be less than the tolerance δ for three values of x using the (Left) explicit formula and (Right) recursive formula. Horizontal lines show the precisions provided by the four smallest IEEE-754-2019 binary floating point formats [2] and the x87 FPU 80-bit floating point format [3]. Both panels share the same legend, which is in the left panel.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Risk assessment for airborne disease transmission by poly-pathogen aerosols

Freja Nordsiek[1], Eberhard Bodenschatz[1,2,3*], Gholamhossein Bagheri[1],

1 Max Planck Institute for Dynamics and Self-Organization (MPIDS), Göttingen, Niedersachsen, Germany
2 Institute for Dynamics of Complex Systems, University of Göttingen, Göttingen, Niedersachsen, Germany
3 Laboratory of Atomic and Solid State Physics and Sibley School of Mechanical and Aerospace Engineering, Cornell University, Ithaca, New York, USA

* lfpn-office@ds.mpg.de

S4 Appendix. $\mathbf{M_c}$ Heuristic for Infectious People Derivation

The M_c heuristic for production by all infectious people is

$$\mathcal{H}_{I,k}(t) \equiv k \int_{d_-}^{d_+} d\phi\, A_S(\phi,t) \sum_{j=1}^{N_I} \lambda_{i,j}(t) n_{I,j,k}(\phi,t) \left[1 - E_{I,m,out,j}(\phi)\right] \quad . \tag{1}$$

For this heuristic, we can simplify $J_{I,M_c}(t)$ since $n_{I,j,k}(d_0,t)$ is the product of a scalar and a Poisson probability until $d_{m,k} > d_0$ after which it is zero. Now $n_{I,j,k}(d_0,t) = 0$ for $k > K_m(d_0)$. The sums and the integrals all commute, so we can change their order. We then get inner sums of the following form in both the numerator and the denominator

$$\begin{aligned}
Z(\mu, m, K_m) &= \sum_{k=m}^{K_m} k P_P(\mu, k) \\
&= \sum_{k=m}^{K_m} k e^{-\mu} \frac{\mu^k}{k!} \\
&= \mu \sum_{k=m}^{K_m} e^{-\mu} \frac{\mu^{k-1}}{(k-1)!} \\
&= \begin{cases} 0 & \text{if } m > K_m \\ \mu\left[C_P(\mu, K_m - 1) - C_P(\mu, m-2)\right] & \text{otherwise} \end{cases} \\
&= \mu\left[C_P(\mu, K_m - 1) - C_P(\mu, \min(m-2, K_m - 1))\right] \quad ,
\end{aligned} \tag{2}$$

where m is the starting index and $C_P(\mu,k)$ is the CDF (Cumulative Distribution Function) of the Poisson distribution with mean μ for count k, with $C_P(\mu,k) = 0$ for $k < 0$. Note that

$$Z(\mu, 1, K_m) = \mu C_P(\mu, K_m - 1) \quad . \tag{3}$$

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Then

$$J_{I,M_c}(t) = \left\{ \int_{d_-}^{d_+} d\phi\, A_S(\phi,t) \bullet \sum_{j=1}^{N_I} \lambda_{I,j}(t)\left[1 - E_{I,m,out,j}(\phi)\right] \rho_j(\phi,t) Z\left(\langle k\rangle(\phi,t)_j, M_c+1, K_m\right) \right\} \Bigg/ \left\{ \int_{d_-}^{d_+} d\phi\, A_S(\phi,t) \sum_{j=1}^{N_I} \lambda_{I,j}(t)\left[1 - E_{I,m,out,j}(\phi)\right] \rho_j(\phi,t) Z\left(\langle k\rangle(\phi,t)_j, 1, K_m\right) \right\} \quad . \tag{4}$$

Our heuristic is strictly for all of the J_{h,M_c}. But, we could widen the set of heuristics by doing one for each infectious person. If we also replace the integral by an evaluation at d_+, the largest d_0 in the interval, and ignore the filtering efficiency of the infectious person's mask on the way out; then there exists some $j \in I$ for which $M_{c,I,j} \geq M_{c,I}$, meaning that the maximum $M_{c,I,j}$ is an overestimate for $M_{c,I}$. If we use this heuristic, we could end up having to use a larger M_c than we really need; but this method has the advantage of easy calculation. This new heuristic for the j'th infected person is

$$\begin{aligned} \mathcal{J}_{I,M_{c,I,j},j}(t) &= \frac{Z\left(\langle k\rangle(d_+,t)_j, M_{c,I,j}+1, K_m(d_+)\right)}{Z\left(\langle k\rangle(d_+,t)_j, 1, K_m(d_+)\right)} \\ &= 1 - \frac{C_P\left(\langle k\rangle(d_+,t)_j, \min(M_{c,I,j}, K_m(d_+)) - 1\right)}{C_P\left(\langle k\rangle(d_+,t)_j, K_m(d_+) - 1\right)} \quad , \end{aligned} \tag{5}$$

where $K_m(d_0)$ is evaluated for $d_0 = d_+$.

With the threshold $T \in (0,1]$, we get

$$C_P\left(\langle k\rangle(d_0,t)_j, \min(M_{c,I,j}, K_m(d_+)) - 1\right) \geq (1-T)\, C_P\left(\langle k\rangle(d_0,t)_j, K_m(d_+) - 1\right) \quad , \tag{6}$$

and then must find the smallest $M_{c,I,j} \leq K_m(d_+)$ for which this is true. This value exists because $M_{c,I,j} = K_m(d_+)$ trivially makes the statement true. Then,

$$M_{c,I,j}(d_+,T) = 1 + C_P^{-1}\left(\langle k\rangle(d_+,t)_j, (1-T)\, C_P\left(\langle k\rangle(d_+,t)_j, K_m(d_+) - 1\right)\right) \quad , \tag{7}$$

where $C_P^{-1}(\mu,c)$ is the inverse CDF to find the smallest k for which $C_P(\mu,k) \geq c$.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Risk assessment for airborne disease transmission by poly-pathogen aerosols

Freja Nordsiek[1], Eberhard Bodenschatz[1,2,3*], Gholamhossein Bagheri[1],

1 Max Planck Institute for Dynamics and Self-Organization (MPIDS), Göttingen, Niedersachsen, Germany
2 Institute for Dynamics of Complex Systems, University of Göttingen, Göttingen, Niedersachsen, Germany
3 Laboratory of Atomic and Solid State Physics and Sibley School of Mechanical and Aerospace Engineering, Cornell University, Ithaca, New York, USA

* lfpn-office@ds.mpg.de

S3 Appendix. Binning Diameter

Since each d_0 can be treated in isolation of all other d_0 in the model's system of equations, we can safely split the $d_0 \in [d_{m,1}, d_M]$ range into finite sized diameter bins and solve each separately and independently of the other bins. We can then apply the general solution to each bin using coefficients averaged over the diameter range in the bin with suitable weights.

Let the i'th strictly increasing bin boundary be $d_{b,i}$ for $i \in [0, N_b]$ where N_b is the number of bins, $d_{b,0} = d_{m,1}$, and $d_{b,N_b} = d_M$. The i'th bin will refer to the bin $d_0 \in [d_{b,i-1}, d_{b,i})$ for $i \in [1, N_b]$. In the limit that the width of the widest bin goes to zero, we get the exact answer. For finite width bins, the more that $\alpha(d_0, t)$ and $\beta_k(d_0, t)$ vary with respect to d_0 in the bin, the less accurate the result from applying binning will be. In practice, this means that we could choose to do a single bin to make it easier to compute the answer but suffer accuracy problems or we could choose a smaller bin width that is small compared to the scale that $\alpha(d_0, t)$ and $\beta_k(d_0, t)$ change over and get good accuracy at the expense of effort. An added benefit of binning is that we can choose different M_c for each bin. In many cases, M_c will be one except for the largest diameter bins.

For any choice of bins, we must determine suitable average values of $n_{0,k}(d_0)$, $n_{r,k}(d_0, t)$, $\alpha(d_0, t)$, and $\beta_k(d_0, t)$. One possible scheme would be

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

$$
\begin{aligned}
n_{0,k}|_i &= \langle n_{0,k}(d_0)\rangle_i \quad , && (1)\\
n_{r,k}|_i(t) &= \langle n_{r,k}(d_0,t)\rangle_i \quad , && (2)\\
\beta_{r,k}|_i(t) &= q_r(t)\,\langle n_{r,k}(d_0,t)\rangle_i \quad , && (3)\\
\beta_{I,k}|_i(t) &= \frac{N_I}{V}\left\langle \lambda_I(t)\,\langle n_{I,k}(d_0,t)\,[1-E_{I,m,out}(d_0)]\rangle_i\right\rangle_I \quad , && (4)\\
\beta_k|_i(t) &= \beta_{r,k}|_i(t)+\beta_{I,k}|_i(t) \quad , && (5)\\
\alpha_v|_i(t) &= q_v(t)\,\langle E_v\,(w(d_0,t)d_0)\rangle_i \quad , && (6)\\
\alpha_g|_i(t) &= \langle \alpha_g\,(w(d_0,t)d_0)\rangle_i \approx \frac{g\,(\rho_w-\rho_a)\,w(d_0,t)^2\left(d_{b,i}^3-d_{b,i-1}^3\right)}{54h\rho_a\nu_a\,(d_{b,i}-d_{b,i-1})} \quad , && (7)\\
\alpha_d|_i(t) &= \langle \alpha_d\,(w(d_0,t)d_0)\rangle_i \quad , && (8)\\
\alpha_{C,f}|_i(t) &= \frac{N_C}{V}\Bigg\langle \lambda_C(t)\Bigg\langle 1-[1-E_{C,m,in,j}\,(w(d_0,t)d_0)] \\
&\qquad \bullet\,[1-E_{C,r,j}(d_0)]\,[1-E_{C,m,out,j}(d_0)]\Bigg\rangle_i\Bigg\rangle_C \quad , && (9)\\
\alpha|_i(t) &= \alpha_o(t)+\alpha_r(t)+\alpha_v|_i(t)+\alpha_g|_i(t)+\alpha_d|_i(t) \\
&\quad +\alpha_{I,f}|_i(t)+\alpha_{S,f}|_i(t)+\alpha_{N,f}|_i(t) \quad , && (10)
\end{aligned}
$$

where $\alpha_o(t)$ and $\alpha_r(t)$ do not require taking bin averages at all since they don't depend on d_0, we have assumed that w is approximately constant with respect to diameter across the bin in evaluating $\alpha_g|_i(t)$ which is essentially ignoring the effect of surface tension (the quality of the approximation gets better as d_0 increases), the notation $F|_i(t)$ denotes the bin average value of $F(d_0,t)$ to use for the i'th bin, and the bin average of some quantity $\langle F(d_0,t)\rangle_i$ is taken by doing

$$\langle F(d_0,t)\rangle_i \equiv \frac{1}{d_{b,i}-d_{b,i-1}}\int_{d_{b,i-1}}^{d_{b,i}} F(\phi,t)d\phi \quad . \tag{11}$$

If a bin is small enough compared to the variation in $F(d_0,t)$ with respect to d_0, one could just approximate the average as the value of F for an arbitrary $d_0 \in [d_{b,i-1}, d_{b,i}]$.

To get the actual number of aerosols $\mathcal{N}_{a,i}$ in each bin, we need to multiply by the bin width, which is

$$\mathcal{N}_{a,i}(t) = (d_{b,i}-d_{b,i-1})\,n_k|_i(t) \quad . \tag{12}$$

We also need to get the average **aerosol** dose in each bin by suitably averaging $\mu_{j,k}(t)$ over the bin. One possible scheme is

$$\mu_{j,k}|_i(t) = \int_{t_0}^{t} dv\,\langle E_{S,r,j}(d_0)\,[1-E_{S,m,in,j}\,(w(d_0,v)d_0)]\rangle_i\,\lambda_{S,j}(v)\mathcal{N}_{a,i}(v) \quad , \tag{13}$$

where $\mathcal{N}_{a,i}(t)$ has taken the place of $n_k|_i(t)$

Regardless of the scheme, the average dose for each multiplicity is calculated by summing over the bins as

$$\mu_{j,k}(t) = \sum_{i=1}^{N_b}\mu_{j,k}|_i(t) \quad . \tag{14}$$

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Risk assessment for airborne disease transmission by poly-pathogen aerosols

Freja Nordsiek[1], Eberhard Bodenschatz[1,2,3*], Gholamhossein Bagheri[1],

1 Max Planck Institute for Dynamics and Self-Organization (MPIDS), Göttingen, Niedersachsen, Germany
2 Institute for Dynamics of Complex Systems, University of Göttingen, Göttingen, Niedersachsen, Germany
3 Laboratory of Atomic and Solid State Physics and Sibley School of Mechanical and Aerospace Engineering, Cornell University, Ithaca, New York, USA

* lfpn-office@ds.mpg.de

S2 Appendix. Checking Analytical Solution Against Numerical Solution

The recursive analytical solution for $\vec{n}$ will be compared against a numerical solution. Lets consider, for some fixed d_0, the following case.

$$
\begin{aligned}
M_c &= 10 \\
\alpha &= \frac{3}{2} \quad [\mathrm{T}]^{-1} \\
\beta &= 50 + 20k \quad [\mathrm{L}]^{-4}\,[\mathrm{T}]^{-1} \\
\gamma &= \frac{1}{10} \quad [\mathrm{T}]^{-1} \\
n_{0,k} &= \left| \frac{M_c}{2} - |4-k| \right|^4 \quad [\mathrm{L}]^{-4}
\end{aligned}
$$

where [T] is the unit of time and [L] is the unit of length. This example starts with both some $n_{0,k} < n_{\infty,k}$ and some $n_{0,k} > n_{\infty,k}$ and a source whose strength increases with k. The system was solved from $t = t_0 = 0$ [T] to $t = 100$ [T] in steps of 10^{-2} [T] for the analytical solution and 10^{-3} [T] for the numerical solution. The system of ODEs was solved numerically using the standard Runge-Kutta 4 in IEEE 754 `binary64` floating point (commonly known as `float64` or double precision). The small time step was chosen in order to check that the differences between the two solutions are small. The analytical solution to the concentration densities $n_k(d_0, t)$ over time is shown in the left panel in Fig 1, along with the normalized residual between the analytical and numerical solutions (absolute value of the difference divided by the analytical solution) in the right panel. The concentration densities decay or grow from $\vec{n}_0$ towards $\vec{n}_\infty$ as we expect. The differences between the analytical and numerical calculations are small (less than 10^{-12}); sometimes reaching the smallest relative differences that can be represented in IEEE 754 `binary64` numbers with their 53 bit mantissas [1], which are 2.2×10^{-16} (numerical bigger than analytical by a fraction of 2^{-52}) and 1.1×10^{-16} (numerical smaller than analytical by a fraction 2^{-53}).

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

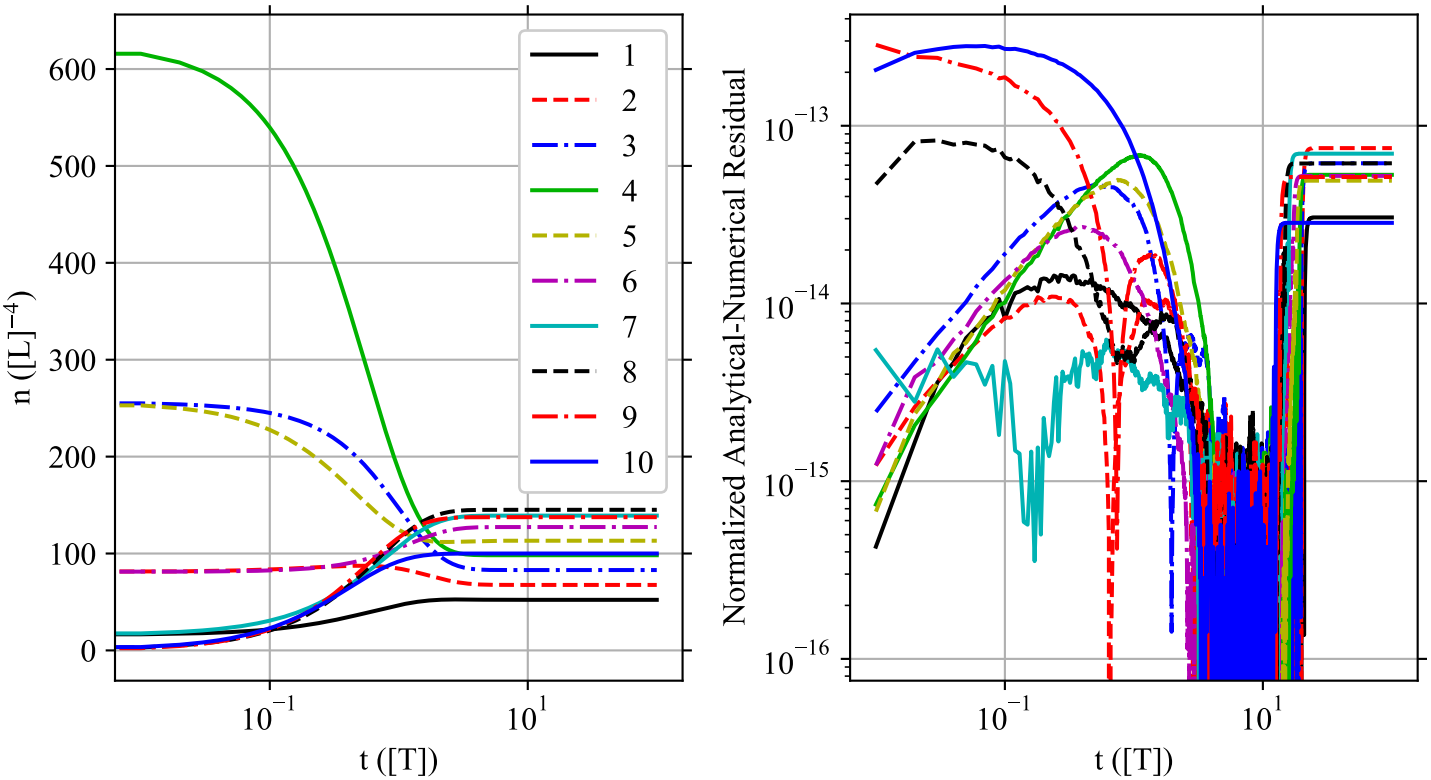


Fig 1. Numerical Validation of Analytical Solution. Comparison of the analytical and numerical solutions of $\vec{n}$ for one particular case and d_0. (Left) Analytical solution (recursive form) to the infectious aerosol concentration density $n_k(t)$ over time, and (Right) the normalized residual between the analytical and numerical solutions $(|n_{k,analytical} - n_{k,numerical}|\,/n_{k,analytical}$ over all time steps except t_0 where some $n_{0,k}$ are zero. Each k is drawn as a separate line, labeled by the value of k. Both panels share the same legend, which is in the Left panel. The numerical solution was done by Runge-Kutta 4 with a time step of 10^{-3} [T] using IEEE 754 `binary64` arithmetic.

References

1. IEEE Floating-Point Working Group. IEEE Standard for Floating-Point Arithmetic. IEEE Std 754-2019 (Revision of IEEE 754-2008). 2019; p. 1–84. doi:10.1109/IEEESTD.2019.8766229.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Risk assessment for airborne disease transmission by poly-pathogen aerosols

Freja Nordsiek[1], Eberhard Bodenschatz[1,2,3*], Gholamhossein Bagheri[1],

1 Max Planck Institute for Dynamics and Self-Organization (MPIDS), Göttingen, Niedersachsen, Germany
2 Institute for Dynamics of Complex Systems, University of Göttingen, Göttingen, Niedersachsen, Germany
3 Laboratory of Atomic and Solid State Physics and Sibley School of Mechanical and Aerospace Engineering, Cornell University, Ithaca, New York, USA

* lfpn-office@ds.mpg.de

S1 Appendix. Model Solution Derivation

General Solution

The model is a finite coupled system of linear inhomogeneous ODEs for any fixed d_0. In matrix-vector form, it is

$$\frac{d\vec{n}}{dt} = \mathbf{A}(d_0,t)\vec{n}(d_0,t) + \vec{\beta}(d_0,t) \quad , \tag{1}$$

where $\vec{n}(d_0,t)$ and $\vec{\beta}(t)$ are the $n_k(d_0 t)$ and $\beta_k(d_0,t)$ for $k > 0$ in vector form and

$$\mathbf{A} \equiv \begin{bmatrix} -\alpha(d_0,t) - \gamma(t) & 2\gamma(t) & & & \\ & -\alpha(d_0,t) - 2\gamma(t) & 3\gamma(t) & & \\ & & \ddots & \ddots & \\ & & & \ddots & M_c\gamma(t) \\ & & & & -\alpha(d_0,t) - M_c\gamma(t) \end{bmatrix} \quad . \tag{2}$$

is an upper bidiagonal $M_c \times M_c$ square matrix.

It is well-known that the complementary solution (the one for when $\vec{\beta} = \vec{0}$) is $\vec{n}_C(d_0,t) = \mathbf{X}(d_0,d_0,t)\mathbf{X}^{-1}(d_0,t_0)\vec{n}_0$ where $\mathbf{X}(d_0,t) = \exp\left[\int^t \mathbf{A}(d_0,x)dx\right]$ if $\mathbf{A}(d_0,t)$ commutes with itself at all combinations of time at fixed d_0 ($\mathbf{A}(d_0,x)\mathbf{A}(d_0,y) = \mathbf{A}(d_0,y)\mathbf{A}(d_0,x) \;\; \forall \;\; x,y \in \mathbb{R}$) [1]. As we will later show from its diagonalization, $\mathbf{A}$ commutes with itself for any two combinations of time with fixed d_0. Let its diagonalization be $\mathbf{A} = \mathbf{B}\mathbf{\Lambda}\mathbf{B}^{-1}$, which is gotten via eigenvalue decomposition. We will use the notation $\mathbf{c}_{m,k}$ to denote the m'th row and k'th column of a matrix $\mathbf{C}$.

It can be shown that the eigen values of a bidiagonal matrix are its diagonal elements. The eigenvalue matrix $\mathbf{\Lambda}$ is a diagonal matrix where the k'th diagonal element is the k'th eigen value. Putting them in the same order as the diagonal elements of $\mathbf{A}$, they are

$$\ell_k(d_0,t) = -\alpha(d_0,t) - k\gamma(t) \quad . \tag{3}$$

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Except in the degenerate but trivial case $\gamma(t) = 0$, there are M_c distinct eigenvalues (not degenerate), and in the degenerate case $\mathbf{A} = -\alpha\mathbf{I}$ and is trivially diagonalized ($\mathbf{\Lambda} = \mathbf{A}$ and $\mathbf{B} = \mathbf{B}^{-1} = \mathbf{I}$). So $\mathbf{A}$ has M_c linearly independent eigen vectors and is therefore diagonalizable. To complete the diagonalization, we need to find the eigen vector $\vec{b}_k(d_0, t)$ for each eigen value by solving

$$\mathbf{A}\vec{b}_k = \ell_k \vec{b}_k \quad . \tag{4}$$

We can begin construction of the eigenvector by setting all elements $m > k$ to zero and element $m = k$ to one. For elements $m \geq k$, the eigen value-vector relationship above trivially holds. For element $m < k$, this becomes the following recursive relationship for the m'th row in terms of row $m + 1$

$$\begin{aligned} \ell_m b_{m,k} + (m+1)\gamma b_{m+1,k} &= \ell_k b_{m,k} \\ b_{m,k} &= -\left(\frac{m+1}{k-m}\right) b_{m+1,k} \quad , \end{aligned} \tag{5}$$

which starts with $b_{k,k} = 1$. From this, element $m < k$ is

$$b_{m,k} = (-1)^{k-m} \frac{(m+1)(m+2)\cdots k}{(k-m)(k-m-1)\cdots 1} = (-1)^{k-m} \binom{k}{m} \quad , \tag{6}$$

where $\binom{k}{m} = k!/(m!(k-m)!)$ is the notation for the binomial coefficient k choose m. Conveniently, this means that all the elements of the eigen vectors are integers and there is no dependence on t nor d_0. The full eigen vector matrix is

$$b_{m,k} = \begin{cases} (-1)^{k-m}\binom{k}{m} & \text{if } m < k \quad , \\ 1 & \text{if } m = k \quad , \\ 0 & \text{if } m > k \quad , \end{cases} \tag{7}$$

where the first case is technically also true along the diagonal ($m = k$) since it is equal to one but it is more clear to write the diagonal elements explicitly as one. The inverse of $\mathbf{B}$ is equal to the matrix $\mathbf{B}$ but with all elements replaced by their absolute values. The elements of $\mathbf{B}^{-1}$ are thus

$$\left(b^{-1}\right)_{m,k} = \begin{cases} \binom{k}{m} & \text{if } m < k \quad , \\ 1 & \text{if } m = k \quad , \\ 0 & \text{if } m > k \quad . \end{cases} \tag{8}$$

This was originally guessed after looking at the numerically computed inverse, but we will prove it here. Consider the m'th row and the k'th column of $\mathbf{B}\mathbf{B}^{-1}$ which is just $\left(\mathbf{B}\mathbf{B}^{-1}\right)_{m,k} = \sum_{p=1}^{M_c} b_{m,p}\left(b^{-1}\right)_{p,k}$. The elements of the sum are zero unless $m \leq p \leq k$, which means that $\left(\mathbf{B}\mathbf{B}^{-1}\right)_{m,k} = 0$ for $m > k$. For $m \leq k$, the zero elements at the beginning and end of the sum drop and we get a sum between m and k that is

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

$$
\begin{aligned}
\left(\mathbf{B}\mathbf{B}^{-1}\right)_{m,k} &= \sum_{p=m}^{k}(-1)^{p-m}\binom{p}{m}\binom{k}{p} \\
&= \sum_{q=0}^{s}(-1)^{p}\binom{q+m}{m}\binom{k}{q+m} \\
&= \sum_{q=0}^{s}(-1)^{q}\frac{k!}{m!q!(k-q-m)!} \\
&= \frac{k!}{m!s!}\sum_{q=0}^{s}(-1)^{q}\binom{s}{q} \\
&= \begin{cases} 1 & \text{if } m=k \quad , \\ 0 & \text{if } m<k \quad , \end{cases}
\end{aligned} \tag{9}
$$

where we have defined $q \equiv p - m$ and $s \equiv k - m$, the last sum is trivially equal to 1 if $m = k$ since then $s = 0$ and only the first term is present, and the sum is zero for $s > 0$ $(k > m)$ as proven by Aupetit (Eq 7 in Appendix) [2]. All non-diagonal elements are zero and all diagonal elements are one, meaning that we have the identity matrix $\mathbf{I}$. Thus Eq 8 is indeed the matrix inverse of Eq 7.

Thus we have diagonalized $\mathbf{A}$. We can see that while the eigen values depend on t and d_0 if α or γ do; $\mathbf{B}$ (and therefore $\mathbf{B}^{-1}$) do not and are therefore constant. Note that $\mathbf{A}(d_0, t)$, $\mathbf{B}$, $\mathbf{B}^{-1}$, and $\mathbf{\Lambda}(d_0, t)$ are all upper-triangular matrices. Now we can show that $\mathbf{A}$ commutes with itself for any two times but fixed d_0. For two arbitrary times $x, y \in \mathbb{R}$, the product is

$$
\begin{aligned}
\mathbf{A}(x)\mathbf{A}(y) &= \mathbf{B}\mathbf{\Lambda}(x)\mathbf{B}^{-1}\mathbf{B}\mathbf{\Lambda}(y)\mathbf{B}^{-1} \\
&= \mathbf{B}\mathbf{\Lambda}(y)\mathbf{\Lambda}(x)\mathbf{B}^{-1} \\
&= \mathbf{B}\mathbf{\Lambda}(y)\mathbf{B}^{-1}\mathbf{B}\mathbf{\Lambda}(x)\mathbf{B}^{-1} \\
&= \mathbf{A}(y)\mathbf{A}(x) \quad ,
\end{aligned}
$$

since all diagonal matrices, such as $\mathbf{\Lambda}(t)$, commute with each other. Thus $\mathbf{A}$ commutes with itself for any two times (but fixed d_0). Since $\mathbf{A}$ commutes with itself for any two times, $\int^t \mathbf{A}(d_0, x)dx$ also commutes itself for any two times (integral can be turned into an infinite Riemann sum over $\mathbf{A}$ which shows that it commutes with itself). Then the multiplied matrix exponentials $\exp\left[\int^t \mathbf{A}(d_0, x)dx\right] \exp\left[-\int^t \mathbf{A}(d_0, x)dx\right]$ become the matrix exponential of their sums. Choosing the constant of integration such that$\vec{n}_C(d_0, 0) = \vec{n}_0(d_0)$, we get for the complete complementary solution

$$
\vec{n}_C(d_0, t) = \exp\left[\int_{t_0}^{t} \mathbf{A}(d_0, x)dx\right]\vec{n}_0 \quad . \tag{10}
$$

Then the general solution to Eq (1) is

$$
\vec{n}(d_0, t) = \exp\left[\int_{t_0}^{t} A(d_0, x)dx\right]\vec{n}_0(d_0) + \int_{t_0}^{t} \exp\left[\int_{s}^{t} A(d_0, x)dx\right]\vec{\beta}(d_0, s)ds \quad , \tag{11}
$$

which we got by guessing based on the equivalent solution for a single equation rather than a system and then checking it (checking it is straightforward).

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Now we need the matrix exponential $\exp\left[\int_s^t \mathbf{A}dx\right]$ for some scalar s. All t and d_0 dependence in the diagonalization of $\mathbf{A}$ is confined to the eigen value matrix $\mathbf{\Lambda}(d_0, t)$, which greatly simplifies finding this matrix exponential. First, we can rewrite the eigenvalue matrix as

$$\mathbf{\Lambda}(d_0, t) = -\alpha(d_0, t)\mathbf{I} - \gamma(t)\mathbf{G} \quad , \tag{12}$$

where $\mathbf{G}$ is the diagonal matrix

$$\mathbf{G} = \begin{bmatrix} 1 & & \\ & \ddots & \\ & & M_c \end{bmatrix} \quad . \tag{13}$$

Since $\mathbf{A}$ is diagonalized, all diagonal matrices commute with each other with respect to multiplication, $\mathbf{B}$ is not a function of time, and the matrix exponential of the sum of two matrices is the product of their exponentials if they commute; the matrix exponential of $\int_s^t \mathrm{A}(d_0, s)dx$ is

$$\begin{aligned} \exp\left[\int_s^t \mathrm{A}(d_0, x)dx\right] &= \mathbf{B}\,\exp\left[\int_s^t \mathbf{\Lambda}(d_0, x)dx\right]\mathbf{B}^{-1} \\ &= \mathbf{B}\,\exp\left[-\mathbf{I}\int_s^t \alpha(x)dx - \mathbf{G}\int_s^t \gamma(x)dx\right]\mathbf{B}^{-1} \\ &= \mathbf{I}\exp\left[-\int_s^t \alpha(d_0, x)dx\right]\mathbf{B}\exp\left[-\mathbf{G}\int_s^t \gamma(x)dx\right]\mathbf{B}^{-1} \\ &= u(d_0, t, s)\mathbf{H}(t, s) \quad , \end{aligned} \tag{14}$$

where

$$u(d_0, t, s) = \exp\left[-\int_s^t \alpha(d_0, x)dx\right] \quad , \tag{15}$$

$$H(t, s) = \mathbf{B}\exp\left[-\mathbf{G}\int_s^t \gamma(x)dx\right]\mathbf{B}^{-1} \quad . \tag{16}$$

We can put this into the general solution for the system of ODEs from Eq (11) to get

$$\vec{n}(d_0, t) = u(d_0, t, t_0)\mathbf{H}(t, t_0)\vec{n}_0(d_0) + \int_{t_0}^t u(d_0, t, s)\mathbf{H}(t, s)\vec{\beta}(d_0, s)ds \quad . \tag{17}$$

Now we must determine $\mathbf{H}$. Since $\mathbf{G}$ is a diagonal matrix

$$\exp\left[-\mathbf{G}\int_s^t \gamma(x)dx\right] = \begin{bmatrix} e^{-\int_s^t \gamma(x)dx} & & \\ & \ddots & \\ & & e^{-M_c\int_s^t \gamma(x)dx} \end{bmatrix} \quad . \tag{18}$$

When this multiplies with $\mathbf{B}$, it essentially scales the columns by the diagonal elements. The resulting matrix is an upper triangular matrix. For the m'th row and k'th column, $h_{m,k}(t, s) = 0$ if $m > k$; and for $m \leq k$ it is

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

$$
\begin{aligned}
h_{m,k}(t,s) &= \sum_{i=m}^{k} b_{m,i}(b_{i,k})^{-1} \exp\left[-i\int_s^t \gamma(x)dx\right] \\
&= \sum_{i=m}^{k} \frac{(-1)^{i-m}k!}{m!(i-m)!(k-i)!} \exp\left[-i\int_s^t \gamma(x)dx\right] \\
&= \sum_{p=0}^{k-m} \frac{(-1)^{p}k!}{m!p!(k-m-p)!} \exp\left[-(p+m)\int_s^t \gamma(x)dx\right] \\
&= \frac{k!}{m!(k-m)!} \sum_{p=0}^{k-m} (-1)^p \binom{k-m}{p} \exp\left[-(p+m)\int_s^t \gamma(x)dx\right] \\
&= \binom{k}{m} \sum_{p=0}^{k-m} (-1)^p \binom{k-m}{p} \exp\left[-(p+m)\int_s^t \gamma(x)dx\right] \quad , \qquad (19) \\
&= \binom{k}{m} \exp\left[-m\int_s^t \gamma(x)dx\right] \sum_{p=0}^{k-m} \binom{k-m}{p} \left[-\exp\left[-\int_s^t \gamma(x)dx\right]\right]^p \\
&= \binom{k}{m} \exp\left[-m\int_s^t \gamma(x)dx\right] \left[1-\exp\left[-\int_s^t \gamma(x)dx\right]\right]^{k-m} \quad , \qquad (20)
\end{aligned}
$$

where the last step uses the definition of binomial coefficients. One could use either Eq (19) or (20) to calculate the elements or do further derivations.

Dropping out of vector form, we can now write Eq (17) for each multiplicity as

$$
\begin{aligned}
n_k(d_0,t) = u(d_0,t,t_0) &\exp\left[-k\int_{t_0}^t \gamma(x)dx\right] \\
&\bullet \sum_{p=k}^{M_c} \binom{p}{k} n_{0,p}(d_0) \left[1-\exp\left[-\int_{t_0}^t \gamma(x)dx\right]\right]^{p-k} \\
&+ \sum_{p=k}^{M_c} \binom{p}{k} \int_{t_0}^t u(d_0,t,s)\beta_p(d_0,s) \\
&\bullet \exp\left[-k\int_s^t \gamma(x)dx\right] \left[1-\exp\left[-\int_s^t \gamma(x)dx\right]\right]^{p-k} ds \quad , \qquad (21)
\end{aligned}
$$

or with $u(d_0,t,t_0)$ substituted out as

$$
\begin{aligned}
n_k(d_0,t) = \exp\left[-\int_{t_0}^t \alpha(d_0,x)dx\right] &\exp\left[-k\int_{t_0}^t \gamma(x)dx\right] \\
&\bullet \sum_{p=k}^{M_c} \binom{p}{k} n_{0,p}(d_0) \left[1-\exp\left[-\int_{t_0}^t \gamma(x)dx\right]\right]^{p-k} \\
&+ \sum_{p=k}^{M_c} \binom{p}{k} \int_{t_0}^t \beta_p(d_0,s) \\
\bullet \exp\left[-\int_s^t \alpha(d_0,x)dx\right] &\exp\left[-k\int_s^t \gamma(x)dx\right] \left[1-\exp\left[-\int_s^t \gamma(x)dx\right]\right]^{p-k} ds \quad . \qquad (22)
\end{aligned}
$$

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Solution for Coefficients Constant in Time

Eq (22) can be simplified if α, $\vec{\beta}$, and/or γ are constant with respect to time (or approximately so). If $\alpha(d_0)$ is constant with respect to time, then

$$u(d_0, t, s) = e^{-(t-s)\alpha(d_0)} \quad . \tag{23}$$

If γ is constant (it has no d_0 dependence, so being constant with respect to time makes it a constant outright), then Eq (19) and (20) for expressing the upper triangle of $\mathbf{H}$ become

$$h_{m,k}(t,s) = \binom{k}{m} \sum_{p=0}^{k-m} (-1)^p \binom{k-m}{p} e^{-(p+m)(t-s)\gamma} \quad , \tag{24}$$

$$h_{m,k}(t,s) = \binom{k}{m} e^{-m(t-s)\gamma} \left[1 - e^{-(t-s)\gamma}\right]^{k-m} \quad . \tag{25}$$

If α, $\vec{\beta}$, and γ are all constant with respect to time; then for multiplicity k,

$$\begin{aligned}
&\left[\int_{t_0}^{t} u(d_0,t,s)\, \mathbf{H}(t,s)\vec{\beta}(d_0,s)ds\right]_k \\
&\quad = \sum_{i=k}^{M_c} \binom{i}{k} \beta_i(d_0) \int_{t_0}^{t} e^{-(\alpha(d_0)+k\gamma)(t-s)} \left[1 - e^{-(t-s)\gamma}\right]^{i-k} ds \\
&\quad = \sum_{i=k}^{M_c} \binom{i}{k} \beta_i(d_0) \sum_{p=0}^{i-k} \binom{i-k}{p} \int_{t_0}^{t} e^{-(\alpha(d_0)+k\gamma)(t-s)} \left[-e^{-(t-s)\gamma}\right]^{p} ds \\
&\quad = \sum_{i=k}^{M_c} \binom{i}{k} \beta_i(d_0) \sum_{p=0}^{i-k} \binom{i-k}{p} (-1)^p \int_{t_0}^{t} e^{-[\alpha(d_0)+(k+p)\gamma](t-s)} ds \\
&\quad = \sum_{i=k}^{M_c} \binom{i}{k} \beta_i(d_0) \sum_{p=0}^{i-k} \binom{i-k}{p} \frac{(-1)^p}{\alpha(d_0)+(k+p)\gamma} \\
&\qquad \bullet \left[1 - e^{-[\alpha(d_0)+(k+p)\gamma](t-t_0)}\right] \quad .
\end{aligned} \tag{26}$$

Then if α, $\vec{\beta}$, and γ are all constant; the general solution from Eq (21) becomes

$$\begin{aligned}
n_k(d_0,t) &= e^{-(\alpha(d_0)+k\gamma)(t-t_0)} \sum_{i=k}^{M_c} \binom{i}{k} n_{0,i}(d_0) \left[1 - e^{-(t-t_0)\gamma}\right]^{i-k} \\
&+ \sum_{i=k}^{M_c} \binom{i}{k} \beta_i(d_0) \sum_{p=0}^{i-k} \binom{i-k}{p} \frac{(-1)^p}{\alpha(d_0)+(k+p)\gamma} \left[1 - e^{-[\alpha(d_0)+(k+p)\gamma](t-t_0)}\right] \quad .
\end{aligned} \tag{27}$$

The concentration density as $t \to \infty$ is

$$n_{\infty,k}(d_0) = \sum_{i=k}^{M_c} \binom{i}{k} \beta_i(d_0) \sum_{p=0}^{i-k} \binom{i-k}{p} \frac{(-1)^p}{\alpha(d_0)+(k+p)\gamma} \quad . \tag{28}$$

This term is present in Eq (27), so we can re-express it as

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

$$n_k(d_0,t) = n_{\infty,k}(d_0) + e^{-(\alpha(d_0)+k\gamma)(t-t_0)} \sum_{i=k}^{M_c} \binom{i}{k} \Bigg\{ n_{0,i}(d_0) \left[1 - e^{-(t-t_0)\gamma}\right]^{i-k} - \beta_i(d_0) \sum_{p=0}^{i-k} \binom{i-k}{p} \frac{(-1)^p e^{-p(t-t_0)\gamma}}{\alpha(d_0)+(k+p)\gamma} \Bigg\} \quad , \tag{29}$$

Calculation of the average aerosol dose requires the time integral of n_k. Completing the integral on all but the last term (will do an alternative evaluation of it later), it is

$$\begin{aligned} \int_{t_0}^{t} n_k(d_0,v)dv &= (t-t_0)n_{\infty,k}(d_0) \\ &+ \sum_{i=k}^{M_c} \binom{i}{k} n_{0,i}(d_0) \sum_{p=0}^{i-k} \binom{i-k}{p} \frac{(-1)^p}{\alpha(d_0)+(k+p)\gamma} \left[1 - e^{-[\alpha(d_0)+(k+p)\gamma](t-t_0)}\right] \\ &- \int_{t_0}^{t} dv\, e^{-(\alpha(d_0)+k\gamma)(v-t_0)} \sum_{i=k}^{M_c} \binom{i}{k} \beta_i(d_0) \sum_{p=0}^{i-k} \binom{i-k}{p} \frac{(-1)^p e^{-p(v-t_0)\gamma}}{\alpha(d_0)+(k+p)\gamma} \quad . \end{aligned} \tag{30}$$

And if we completely evaluate the last integral,

$$\begin{aligned} \int_{t_0}^{t} n_k(d_0,v)dv &= (t-t_0)n_{\infty,k}(d_0) \\ &+ \sum_{i=k}^{M_c} \binom{i}{k} \sum_{p=0}^{i-k} \binom{i-k}{p} \frac{(-1)^p}{\alpha(d_0)+(k+p)\gamma} \Bigg\{ n_{0,i}(d_0) \left[1 - e^{-[\alpha(d_0)+(k+p)\gamma](t-t_0)}\right] \\ &+ \beta_i(d_0) \left(\frac{e^{-[\alpha(d_0)+(k+p)\gamma](t-t_0)} - 1}{\alpha(d_0)+(k+p)\gamma} \right) \Bigg\} \quad . \end{aligned} \tag{31}$$

Recursive Solution for Coefficients Constant in Time

For $\mathbf{n_{\infty,k}}$

We can rewrite $n_{\infty,k}(d_0)$ as a recursive expression, which greatly reduces the effort in calculating $\vec{n}_\infty(d_0)$. At steady state where $\vec{n}_k = \vec{n}_\infty$ (such as when $t \to \infty$), $dn_k/dt \to 0$ and Eq 1 becomes

$$n_{\infty,k}(d_0) = \frac{1}{\alpha(d_0)+k\gamma} \left[\beta_k(d_0) + (k+1)\gamma\, n_{\infty,k+1}(d_0)\right] \quad , \tag{32}$$

which can be calculated recursively starting from $k = M_c$ and then descending to $k = 1$. At $k = M_c$, the value is

$$n_{\infty,M_c}(d_0) = \frac{\beta_{M_c}(d_0)}{\alpha(d_0)+M_c\gamma} \quad , \tag{33}$$

since $n_{\infty,k+1} = 0$ for $k \geq M_c$.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

For $\mathbf{n_k}$

We have so far been unable to make a recursive form for $\vec{n}$ in Eq (29) where the total number of terms scales as $\mathcal{O}(M_c)$, but there is a recursive form that reduces the number of terms that must be evaluated to scale as $\mathcal{O}\left(M_c^2\right)$ instead of $\mathcal{O}\left(M_c^3\right)$. We can re-express the innermost sum in Eq 29 in terms of a Gauss hypergeometric function. The Gauss hypergeometric function [3] is

$$ {}_2\mathrm{F}_1(a,b;c;z) = \sum_{p=0}^{\infty} \frac{(a)_p (b)_p \, z^p}{(c)_p \, p!} \quad , \tag{34} $$

and the Pochhammer symbol [3] for integer p is

$$ (x)_p = \begin{cases} x(x+1)\ldots(x+p-1) & \text{if } p > 0 \quad , \\ 1 & \text{if } p = 0 \quad . \end{cases} \tag{35} $$

Note that ${}_2\mathrm{F}_1(a,b;c;z) = {}_2\mathrm{F}_1(b,a;c;z)$. For the special case that $a = -m$ is a negative integer, like in our case, it is [3]

$$ {}_2\mathrm{F}_1(-m,b;c;z) = \sum_{p=0}^{m} \binom{m}{p} \frac{(-1)^p (b)_p z^p}{(c)_p} \quad . \tag{36} $$

This lets us re-express Eq 29 as

$$ n_k(d_0,t) = n_{\infty,k}(d_0) + z^s \left[U_k(d_0, \vec{\beta}(d_0), z) + V_k(\vec{n}_0(d_0), z) \right] \quad , \tag{37} $$

where

$$ V_k(\vec{y}, x) = \sum_{i=k}^{M_c} \binom{i}{k} y_i (1-x)^{i-k} \quad , \tag{38} $$

and

$$ \begin{aligned} U_k(d_0, \vec{y}, x) &= -\sum_{i=k}^{M_c} \binom{i}{k} y_i \sum_{p=0}^{i-k} \binom{i-k}{p} \frac{(-1)^p x^p}{\alpha(d_0) + (k+p)\gamma} \\ &= -\frac{1}{\gamma s} \sum_{i=k}^{M_c} \binom{i}{k} y_i \, {}_2\mathrm{F}_1(-(i-k), s; s+1; x) \quad , \end{aligned} \tag{39} $$

and we have defined

$$ s(d_0) = \frac{\alpha(d_0)}{\gamma} + k \quad , \tag{40} $$

$$ z(t) = e^{-(t-t_0)\gamma} \quad . \tag{41} $$

The last definitions mean that $e^{-(\alpha+k\gamma)(t-t_0)} = z^s$. Note that since $t \geq t_0$, $z \in (0,1]$ which means we are in the domain of the Gauss hypergeometric functions of the form in Eq (36)

We need to separate the sum in $U_k(d_0, \vec{y}, x)$ into the $i = k$ case and the rest, which is a sum over the same range as for n_{k+1} but with different terms inside. Additionally, $\binom{i}{k} = \frac{k+1}{i-k}\binom{i}{k+1}$ and ${}_2\mathrm{F}_1(0, s; s+1; x) = 1$ Then,

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

$$U_k(d_0, \vec{y}, x) = -\frac{1}{\gamma s}\left[y_k + (k+1)\sum_{i=k+1}^{M_c}\binom{i}{k+1}\frac{y_i}{i-k}{}_2\mathrm{F}_1(-(i-k), s; s+1; x)\right] \quad . \quad (42)$$

To build a recursive relationship from k to $k+1$, we need to relate ${}_2\mathrm{F}_1(-(i-k), s; s+1; x)$ to ${}_2\mathrm{F}_1(-(i-k)+1, s+1; s+2; x)$ since $k \to k+1$ changes $s \to s+1$ and $-(i-k) \to -(i-k)+1$. To do this, we will use the contiguous relations/functions which can used to relate any three Gauss hypergeometric functions of the form ${}_2\mathrm{F}_1(a+m_1, b+m_2; c+m_3; z)$ for different combinations of integers m_1, m_2, and m_3 [3]. We use two from Rakha, Rathie & Chopra (their Eq 1.17 and 1.20) [3] to relate these two hypergeometric functions to each other and ones which can be evaluated exactly. Combining the two contiguous relations,

$$\begin{aligned}{}_2\mathrm{F}_1(-(i-k), s; s+1; x) = {}_2\mathrm{F}_1(-(i-k)-1, s+1; s+1; x) \\ + \frac{x}{s+1}\Big[\\ (i-k){}_2\mathrm{F}_1(-(i-k)+1, s+1; s+2; x) \\ + (s+1){}_2\mathrm{F}_1(-(i-k), s+2; s+2; x)\Big] \end{aligned} \quad (43)$$

Now, we have two Gauss hypergeometric functions, with $m \geq 0$, of the form

$$\begin{aligned}{}_2\mathrm{F}_1(-m, c; c; z) &= \sum_{p=0}^{m}\binom{m}{p}\frac{(-1)^p (c)_p z^p}{(c)_p} \\ &= \sum_{p=0}^{m}\binom{m}{p}(-z)^p \\ &= (1-z)^m \end{aligned} \quad (44)$$

Putting this in, we get

$${}_2\mathrm{F}_1(-(i-k), s; s+1; x) = (1-x)^{i-k} + \frac{(i-k)x}{s+1}{}_2\mathrm{F}_1(-(i-k)+1, s+1; s+2; x) \quad . \quad (45)$$

Inserting this into Eq (42) and stepping the binomial coefficient in the sum with the first term back down to $\binom{i}{k}$,

$$\begin{aligned}U_k(d_0, \vec{y}, x) = -\frac{1}{\gamma s}\Bigg[y_k + \sum_{i=k+1}^{M_c}\binom{i}{k}y_i(1-x)^{i-k} \\ + \frac{(k+1)x}{s+1}\sum_{i=k+1}^{M_c}\binom{i}{k+1}y_i\,{}_2\mathrm{F}_1(-(i-k)+1, s+1; s+2; x)\Bigg] \quad . \end{aligned} \quad (46)$$

We can re-arrange this to be a recursive relationship and combine the first two terms in the brackets into a single sum (the first term is the $i = k$ case of the first sum).

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Noting that ${}_2\mathrm{F}_1(0, s; s+1; x) = 1$ for evaluating U_{M_c} with Eq (39), the recursive relationship is

$$U_k(d_0, \vec{y}, x) = \begin{cases} -\frac{y_{M_c}}{\gamma s} & \text{if } k = M_c \quad , \\ \frac{(k+1)x}{s} U_{k+1}(d_0, \vec{y}, x) - \frac{1}{\gamma s} V_k(\vec{y}, x) & \text{otherwise} \quad . \end{cases} \tag{47}$$

The number of terms in evaluating all the $U_k(d_0, \vec{y}, x)$ at a specific d_0 and t scales as $\mathcal{O}\left(M_c^2\right)$ with this recursive definition, rather than $\mathcal{O}\left(M_c^3\right)$ if we use the original definition in Eq (39) since the Gauss hypergeometric function itself makes for an inner sum from Eq 36. And therefore we have also made the number of terms required to evaluate $\vec{n}$ at a specific d_0 and t scale as $\mathcal{O}\left(M_c^2\right)$, since now there are only single sums in each n_k. Additionally, no binomial coefficients are multiplied together anymore which makes it easier to avoid numerical overflow, and the $(-1)^p$ term is gone which causes accuracy problems when doing the inner sum at any given fixed precision.

Looking back at the recursive form of $n_{\infty,k}(d_0)$ in Eq (32) and (33), we can see that

$$n_{\infty,k}(d_0) = -U_k(d_0, \vec{\beta}(d_0), 1) \quad . \tag{48}$$

For Time Integral of $\mathbf{n_k}$

Now, we can build a recursive relationship for $\int_{t_0}^{t} n_k dt$ from Eq (30). The double sums can all be re-expressed in terms of U_k, getting

$$\begin{aligned} \int_{t_0}^{t} n_k(d_0, v) dv = (t - t_0) n_{\infty,k}(d_0) - U_k(d_0, \vec{n}_0(d_0), 1) \\ + z^s U_k(d_0, \vec{n}_0(d_0), z) + \int_{t_0}^{t} dv\, z(v)^s U_k(d_0, \vec{\beta}(d_0), z(v)) \quad . \end{aligned} \tag{49}$$

Changing integration variables in the last integral from t to $z(v)$, we get

$$\begin{aligned} \int_{t_0}^{t} n_k(d_0, v) dv = (t - t_0) n_{\infty,k}(d_0) - U_k(d_0, \vec{n}_0(d_0), 1) \\ + z^s U_k(d_0, \vec{n}_0(d_0), z) - \frac{1}{\gamma} W_k\left(d_0, \vec{\beta}, z\right) \quad , \end{aligned} \tag{50}$$

where

$$W_k\left(d_0, \vec{y}, x\right) = \int_1^x dv\, v^{s-1} U_k(d_0, \vec{y}, v) \quad , \tag{51}$$

with a re-use/reassignment of the integration variable.

Then, writing $U_k(d_0, \vec{y}, x)$ out, using Eq (45) to step the Gauss hypergeometric function to $k+1$, and identifying $W_{k+1}\left(d_0, \vec{y}, x\right)$ and $U_k\left(d_0, \vec{y}, x\right)$;

medRxiv preprint doi: https://doi.org/10.1101/2020.11.30.20241083; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

$$
\begin{aligned}
W_k\left(d_0,\vec{y},x\right) &= -\frac{1}{\gamma s}\int_1^x dv\, v^{s-1}\sum_{i=k}^{M_c}\binom{i}{k}y_i\,{}_2\mathrm{F}_1(-(i-k),s;s+1;v) \\
&= -\frac{1}{\gamma s}\sum_{i=k}^{M_c}\binom{i}{k}y_i\left[\int_1^x v^{s-1}(1-v)^{i-k}dv\right. \\
&\quad \left.+\int_1^x \frac{(i-k)v^s}{s+1}{}_2\mathrm{F}_1(-(i-k)+1,s+1;s+2;v)dv\right] \\
&= -\frac{1}{\gamma s}\sum_{i=k}^{M_c}\binom{i}{k}y_i\sum_{p=0}^{i-k}\binom{i-k}{p}\int_1^x(-1)^p v^{p+s-1}dv \\
&\quad -\frac{k+1}{s}\left[\frac{1}{(s+1)\gamma}\sum_{i=k+1}^{M_c}\binom{i}{k+1}y_i\right. \\
&\quad \left.\bullet\int_1^x v^s{}_2\mathrm{F}_1(-(i-k)+1,s+1;s+2;v)dv\right] \\
&= \frac{k+1}{s}W_{k+1}\left(d_0,\vec{y},x\right)-\frac{x^s}{\gamma s}\sum_{i=k}^{M_c}\binom{i}{k}y_i\sum_{p=0}^{i-k}\binom{i-k}{p}\frac{(-1)^p x^p}{s+p} \\
&\quad +\frac{1}{\gamma s}\sum_{i=k}^{M_c}\binom{i}{k}y_i\sum_{p=0}^{i-k}\binom{i-k}{p}\frac{(-1)^p}{s+p} \\
&= \frac{1}{s}\left[(k+1)W_{k+1}\left(d_0,\vec{y},x\right)+x^sU_k\left(d_0,\vec{y},x\right)-U_k\left(d_0,\vec{y},1\right)\right] \quad .(52)
\end{aligned}
$$

This is recursive since it depends on itself from $k+1$ and $U_k\left(d_0,\vec{y},x\right)$ and $U_k\left(d_0,\vec{y},1\right)$ which are recursive. All that remains is to evaluate it for $k=M_c$, and then Eq (50) is complete. This is

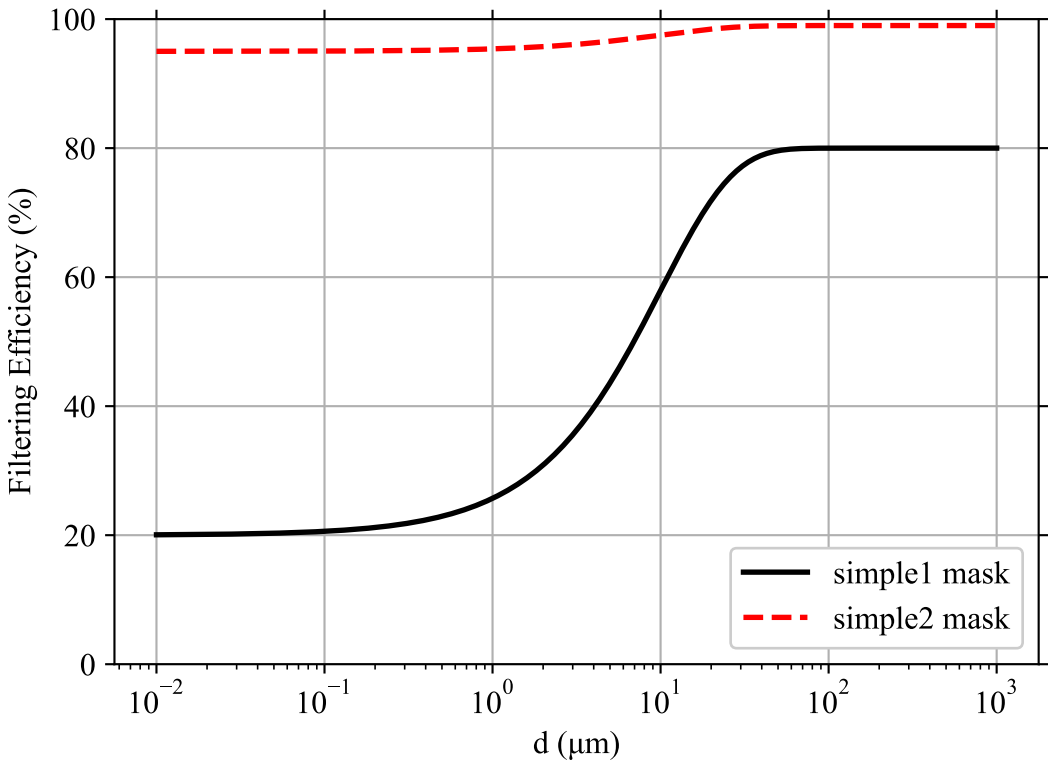
$$
\begin{aligned}
W_{M_c}\left(d_0,\vec{y},x\right) &= -\frac{1}{\gamma s}\int_1^x dv\, v^{s-1}\sum_{i=M_c}^{M_c}\binom{i}{M_c}y_i\,{}_2\mathrm{F}_1(-(i-M_c),s;s+1;v) \\
&= -\frac{y_{M_c}}{\gamma s}\int_1^x v^{s-1}\,{}_2\mathrm{F}_1(0,s;s+1;v)dv \\
&= -\frac{y_{M_c}}{\gamma s}\int_1^x v^{s-1}\,dv \\
&= \frac{y_{M_c}}{\gamma s^2}\left(1-x^s\right) \qquad (53)
\end{aligned}
$$

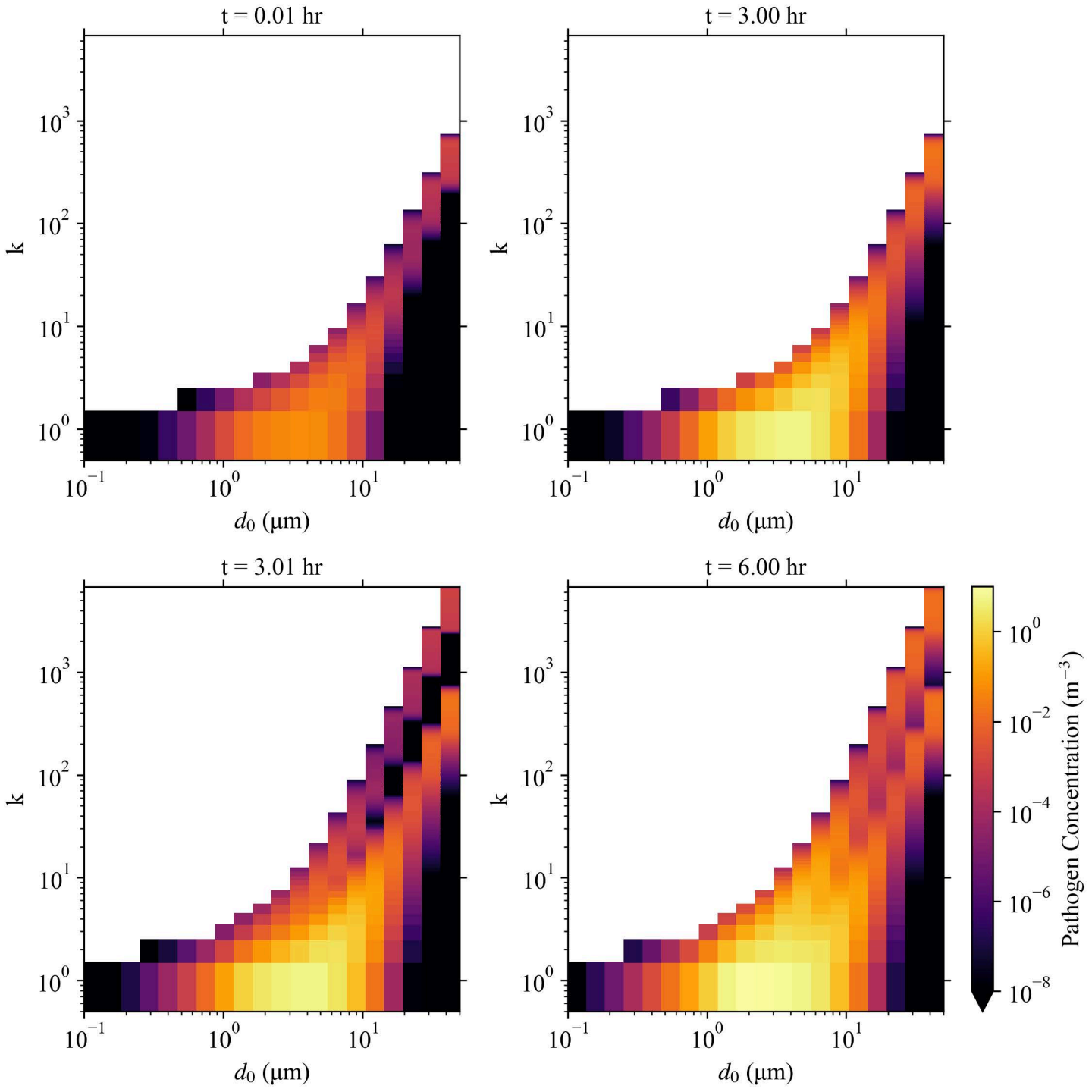
References

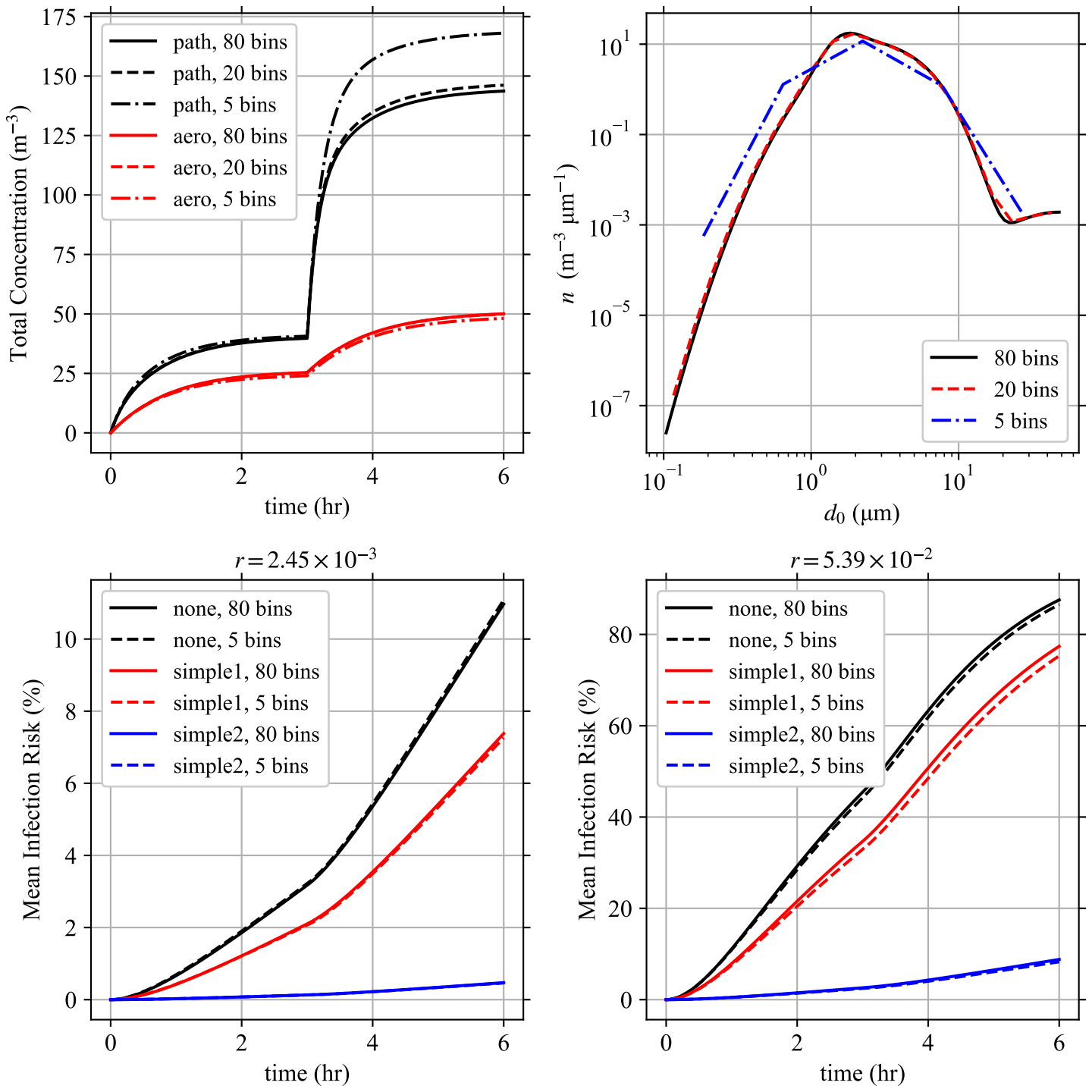
1. Zhu J, Morales CH. On linear ordinary differential equations with functionally commutative coefficient matrices. Linear Algebra and its Applications. 1992;170:81 – 105. doi:10.1016/0024-3795(92)90412-4.

2. Aupetit M. Nearly homogeneous multi-partitioning with a deterministic generator. Neurocomputing. 2009;72(7):1379 – 1389. doi:10.1016/j.neucom.2008.12.024.

3. Rakha MA, Rathie AK, Chopra P. On some new contiguous relations for the Gauss hypergeometric function with applications. Computers & Mathematics with Applications. 2011;61(3):620 – 629. doi:10.1016/j.camwa.2010.12.008.


100
80
60
40
20
0
Filtering Efficiency (%)
10^{-2}
10^{-1}
10^{0}
10^{1}
10^{2}
10^{3}
d (μm)
simple1 mask
simple2 mask


t = 0.01 hr
t = 3.00 hr
t = 3.01 hr
t = 6.00 hr
k
d_0 (μm)
10^{-1}
10^{0}
10^{1}
10^{2}
10^{3}
Pathogen Concentration (m^{-3})
10^{-8}
10^{-6}
10^{-4}
10^{-2}
10^{0}


Total Concentration (m^{-3})
time (hr)
path, 80 bins
path, 20 bins
path, 5 bins
aero, 80 bins
aero, 20 bins
aero, 5 bins
n (m^{-3} μm^{-1})
d_0 (μm)
80 bins
20 bins
5 bins
$r = 2.45 \times 10^{-3}$
$r = 5.39 \times 10^{-2}$
Mean Infection Risk (%)
none, 80 bins
none, 5 bins
simple1, 80 bins
simple1, 5 bins
simple2, 80 bins
simple2, 5 bins