

RESEARCH ARTICLE

Modelling COVID-19 transmission in supermarkets using an agent-based model

Fabian Ying¹, Neave O'Clery^{2*}

1 Quantitative Research, G-Research, London, United Kingdom, **2** Centre for Advanced Spatial Analysis, UCL, London, United Kingdom

* n.oclery@ucl.ac.uk



OPEN ACCESS

Citation: Ying F, O'Clery N (2021) Modelling COVID-19 transmission in supermarkets using an agent-based model. PLoS ONE 16(4): e0249821. <https://doi.org/10.1371/journal.pone.0249821>

Editor: Martin Chtolongo Simuunza, University of Zambia, ZAMBIA

Received: October 15, 2020

Accepted: March 26, 2021

Published: April 9, 2021

Copyright: © 2021 Ying, O'Clery. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All data can be found under https://github.com/fabianyng/covid19-supermarket-abm/tree/main/covid19-supermarket_abm/example_data.

Funding: G-Research provided support in the form of salaries for the author FY, but did not have any additional role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. The specific roles of these authors are articulated in the 'author contributions' section.

Competing interests: G-Research provided support in the form of salaries for the author FY, but did not have any additional role in the study

Abstract

Since the outbreak of COVID-19 in early March 2020, supermarkets around the world have implemented different policies to reduce the virus transmission in stores to protect both customers and staff, such as restricting the maximum number of customers in a store, changes to the store layout, or enforcing a mandatory face covering policy. To quantitatively assess these mitigation methods, we formulate an agent-based model of customer movement in a supermarket (which we represent by a network) with a simple virus transmission model based on the amount of time a customer spends in close proximity to infectious customers (which we call the exposure time). We apply our model to synthetic store and shopping data to show how one can use our model to estimate exposure time and thereby the number of infections due to human-to-human contact in stores and how to model different store interventions. The source code is openly available under <https://github.com/fabianyng/covid19-supermarket-abm>. We encourage retailers to use the model to find the most effective store policies that reduce virus transmission in stores and thereby protect both customers and staff.

Introduction

As the main provider of food and essential goods, supermarkets remained open in many countries throughout the COVID-19 pandemic in 2020, while the majority of other businesses (such as general retail stores) shut down during periods of government-mandated lockdowns [1, 2]. Supermarkets represent one of the main hubs where a large number of people mix indoors throughout the pandemic and are thus a potential risk area where the virus SARS-CoV-2, which causes COVID-19, may be transmitted. It is therefore vital to find safe ways for customers to shop and minimize virus transmission. Models for customer dynamics and virus transmission are useful towards that goal, as they can be used to estimate the infection risk and assess how different interventions affect the risk. In this article, we propose an agent-based model for customer dynamics which we use to estimate the total amount of *exposure time*, which we define as the total amount of time that customers are in close proximity to infected customers. Using a simple virus transmission model, we estimate the number of infections from exposure time. We thereby ignore fomite and airborne transmission and only consider

design, data collection and analysis, decision to publish, or preparation of the manuscript. The specific roles of these authors are articulated in the 'author contributions' section. This does not alter our adherence to PLOS ONE policies on sharing data and materials.

infections due to respiratory droplet transmission, which the World Health Organisation (WHO) and the US Centre for Disease Control and Prevention (CDC) consider the main mode of transmission of COVID-19 [3, 4]. We apply this model to synthetic data and how to model the following interventions:

- Restricting the maximum number of customers in the store,
- Reducing the rate at which customers enter the store,
- Implementing face mask policy, and
- One-way aisle store layout.

These and other interventions have been used or recommended in supermarkets in the UK and the US [5–7], among other countries.

Related work

Existing work for indoor COVID-19 transmission mainly consider airborne transmission rather than respiratory droplet transmission, with a focus on how ventilation and other environmental factors affect airborne transmission risk. The models are typically either Computational Fluid Dynamics (CFD) models [8–10], airflow models based on partial differential equations (PDEs) [11], or models that make a well-mixed room (WMR) assumption, where virus-carrying aerosols are instantaneously evenly distributed across a room [12–15]. CFD models take a long time to run and are typically intractable for large-scale simulations, particularly in situations where there are many moving objects (such as customers in a store). The WMR room models generally only consider a single room and do not take into account the geometry of a store. PDEs models sit in between CFD and WMR models, allowing for more complicated geometry to be taken into account while being computationally tractable. However, none of these models take close contacts due to customer mobility and the resulting droplet transmission risk into account. Furthermore, existing work has not studied how changes in store policy or layout affect infection risk.

In this work, our focus is on respiratory droplet transmission due to customers coming into close contact with one another. We use an agent-based model as it allows us to take into account the non-linearities of the system due to customer mobility and the layout of the store, while being computationally cheap enough to run a large number of simulations without the need to solve any differential equations, as we ignore airborne transmission. Plata et al. [16] proposed a similar model to ours, but on a higher spatial resolution than what we consider. However, their model is computationally more expensive, and their code is not openly available.

Materials and methods

Store graph

We represent a store as a network (called a *store graph*), in which nodes represent zones and edges connect contiguous zones. We create a store graph from a synthetic store layout following a similar procedure as in [17, 18]. Zones are approximately 2m by 2m and we specify a number of entrance, till, and exit nodes (see Fig 1A). We choose a network representation of a store (instead of modelling on the underlying space directly, see e.g., [16]) for ease of simulation, as it significantly reduces the complexity of the model.

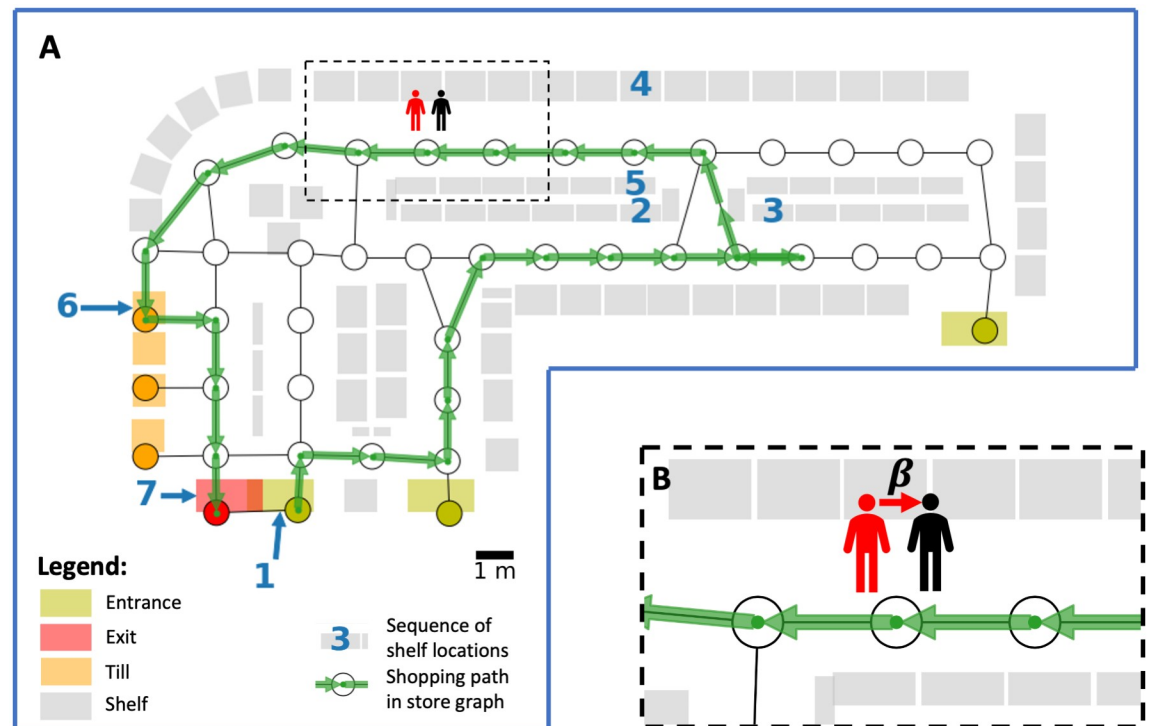


Fig 1. Overview of agent-based model. (A) A network representation of a small supermarket/convenience store with an example shopping path (in green). We generate each shopping path from a sequence of shelf locations (in blue), which correspond to the shelves from a customer picks up their items during a visit and the entrance and the tills. In this example, the customer picks up $K = 4$ items at the shelves marked in blue with 2, 3, 4, and 5. (B) Virus transmission model. A susceptible customer (in black) becomes infected at rate β whenever they are in the same zone as an infectious customer (in red).

<https://doi.org/10.1371/journal.pone.0249821.g001>

Agent-based model

Our agent-based model has two components: a customer mobility model and a virus transmission model. The first component is the customer mobility model for how customers arrive at the store and move. The second component is a model for how the virus transmits in the supermarket.

Customer mobility model. In our agent-based model, customers arrive the store according to a Poisson process with constant rate λ . Each customer starts at a random entrance node (chosen uniformly at random from all entrance nodes) and is assigned a shopping path, chosen uniformly at random from all shopping paths from a synthetic data set or an empirical data set (if available). (In this article, we only consider synthetic data sets, as no empirical ones were available to us.) Each shopping path is a path in the store graph, representing the route that a customer takes in the store. Two consecutive nodes in the shopping path may be identical. This case occurs when a customer picks up one or more items in the zone. A customer traverses the store graph according to its assigned shopping path. At each node, a customer waits a random time T , which is exponentially distributed with mean τ (independent of other waiting times), before traversing to the next node in the shopping path (or staying at the same node, if it is the next node). Once a customer arrives at the final node (which is the exit node) in its shopping path, the customer stays T seconds on the node (with T again exponentially distributed with mean τ) and is then removed from the system. At the beginning of each simulation, the store is empty and customers arrive in the store over a period of H hours

(corresponding to length of the opening hours of the store). After H hours, no new customers arrive and the simulation stops once the last customer leaves the store.

Virus transmission model. Customers are either susceptible or infectious when they enter the store. Each customer that arrives to the store is infectious with independent probability p_I (corresponding to the proportion of infectious customers) and is otherwise susceptible. In our infection mechanism, we assume susceptible customers become infected proportional to the time they spent with infectious customers. We assume that the main mode of transmission is direct transmission via respiratory droplets and neglect airborne transmission and fomite transmission. More formally, we define the *exposure time* E_s for each susceptible customer s as follows. We define the *individual exposure time* $E_s^{(i)}$ to an infectious customer i as the total time that customer s was in the same zone as an infectious customer i during the shopping trip of s . Then $E_s = \sum_i E_s^{(i)}$ is the sum of the individual exposure times. We call a susceptible customer *exposed* if they have positive exposure time. Each exposed customer becomes infected after the shopping trip with probability $\min(\beta E_s, 1)$ for some transmission rate $\beta > 0$. In other words, we model the infection probability of an exposed customer as a linear function of the exposure time with infectious customers. In reality, the infection probability function may take some other form (e.g., a logistic function), but for simplicity and due to lack of validated alternative models, we choose a linear function. We illustrate the virus transmission model in see Fig 1B.

Our model has two main outputs: The total number of infections and the chance of infection for a given simulation period. These two quantities depend linearly on the *total exposure time* (defined as the sum of exposure times) and the mean exposure time per susceptible customer.

Data

We use a synthetically-created store layout and shopping path data with 10^6 paths. The (synthetic) store is a small store with around 80 shelves, four tills, three entrances, and one exit (see Fig 1). We synthetically generate 10^6 shopping paths as follows. For each path, we first sample the number K of items that the customer picks up during their shop. We choose a log-normal distribution for K where the underlying normal distribution has mean μ and standard deviation σ [19]. We then choose a sequence of locations (s_1, \dots, s_{K+3}) , where s_1 is a random entrance, s_2, \dots, s_{K+1} are random shelves (chosen uniformly at random from all shelves with replacement), s_{K+2} is a random till, and s_{K+3} is the exit. The locations s_2, \dots, s_{K+1} represent the K shelves where a customer buys their items in their shopping trip. We map each location s_i to the corresponding node v_i in the store graph that contains the shelf, entrance, till, or exit location. We then generate the shopping path from the node sequence (v_1, \dots, v_{K+3}) by choosing a shortest path in the store graph that visits each of these nodes in sequence. We show an example sequence of locations and its corresponding shopping path in Fig 1A.

Parameter values

In our simulations, we set the default arrival rate to be 2.55 customers per minute. This is based on the mean number of baskets per store over a 91-day period across 17 UK supermarkets as reported in [17] and the typical UK store opening period of 14 hours [20]. (In [17], Ying *et al.* used a data set of 13672 mean baskets per store over 91 days, representing a sample of 7% of the total number of baskets. Therefore, the customer arrival rate is $13672 / (0.07 \times 91 \times 14 \times 60) = 2.55$ customers per minute over this period.) We assume that each basket corresponds to a single customer (rather than groups of customers). At the time of writing, many UK supermarkets advise or restrict customers to shop alone [21].

We choose $\mu = 0.07$ and $\sigma = 0.76$ for the parameters of the log-normal distribution for sampling the item basket size K . These values correspond to the ones reported in [19] for small supermarkets and convenience stores. We infer the mean wait time τ at each node from the mean shopping time of 5.95 minutes for small stores as reported in [19] and the mean shopping path length of 29.66 nodes based on the synthetic data described above. This yields a mean wait time of τ equals to $5.95/29.66 = 0.2$ minutes per node. Note that data on the mean shopping length and the arrival rate is from before the COVID-19 pandemic; these values may differ during a pandemic due to a change in customer shopping behaviour.

We use $p_I = 1.87\%$ as the proportion of infectious customers based on data reported from UK's Office for National Statistics (ONS) Infection Survey from January 2021 [22]. In this survey, a random sample of the population is tested for COVID-19 to estimate the overall proportion of people who have COVID-19 at that particular point in time.

The transmission rate β is much harder to estimate than the previous parameters because very little data exists. In [23], the mean probability of transmission per contact is estimated to be 2.11×10^{-8} . However, they do not specify the duration of a contact. We use this parameter and assume that the mean contact duration is 15 minutes to obtain a rate of transmission to be $\beta = 2.11 \times 10^{-8}/15 = 1.41 \times 10^{-9}$ per minute. (We assume that the probability of transmission is proportional to the contact duration).

We summarize the parameter values that we use in Table 1.

Analysis environment

We implemented our agent-based model in Python 3.6 using SimPy 4 [24]. We ran our simulations on an Amazon Web Service cluster (ml.c5.4xlarge with 32 GB RAM and 16 vCPU) [25], although all simulations can be performed on a standard desktop computer.

Results and discussion

We demonstrate in this section how to use our model, what metrics we can record in it, and compare different interventions on our synthetic store.

Using the default parameters listed in Table 1, we perform 1000 simulations, each simulating a day in our synthetic store. Customers stay on average 5.97 minutes in the store, with on average 14.96 customers present in the store at any given time. With $p_I = 1.87\%$ infected customers, the total exposure time (i.e., the exposure time summed over all susceptible customers) is on average 94.98 min per day. Multiplying this with $\beta = 1.41 \times 10^{-9}$ infections per minute of

Table 1. Parameter values in our agent-based model. We list uncertainty bounds, when available.

Parameter	Default value	Reference/Assumption
Parameters (μ, σ) of log-normal distribution for basket size	(0.07, 0.76)	[19]
Arrival rate (λ)	2.55 customer/min	[17]
Mean wait time at each node (τ)	0.2 min	inferred from [19]
Percentage of infectious customers (p_I)	1.87% 95% conf. interval: (1.79%, 1.95%)	[22]
Transmission rate (β)	1.41×10^{-9} per min 95% conf. interval: (1.15, 1.66×10^{-9})	inferred from [23] (assuming mean contact duration of 15 minutes)
Length of opening hours (H)	14 hours	[20]
Store layout	fixed	Layout of synthetic store

<https://doi.org/10.1371/journal.pone.0249821.t001>

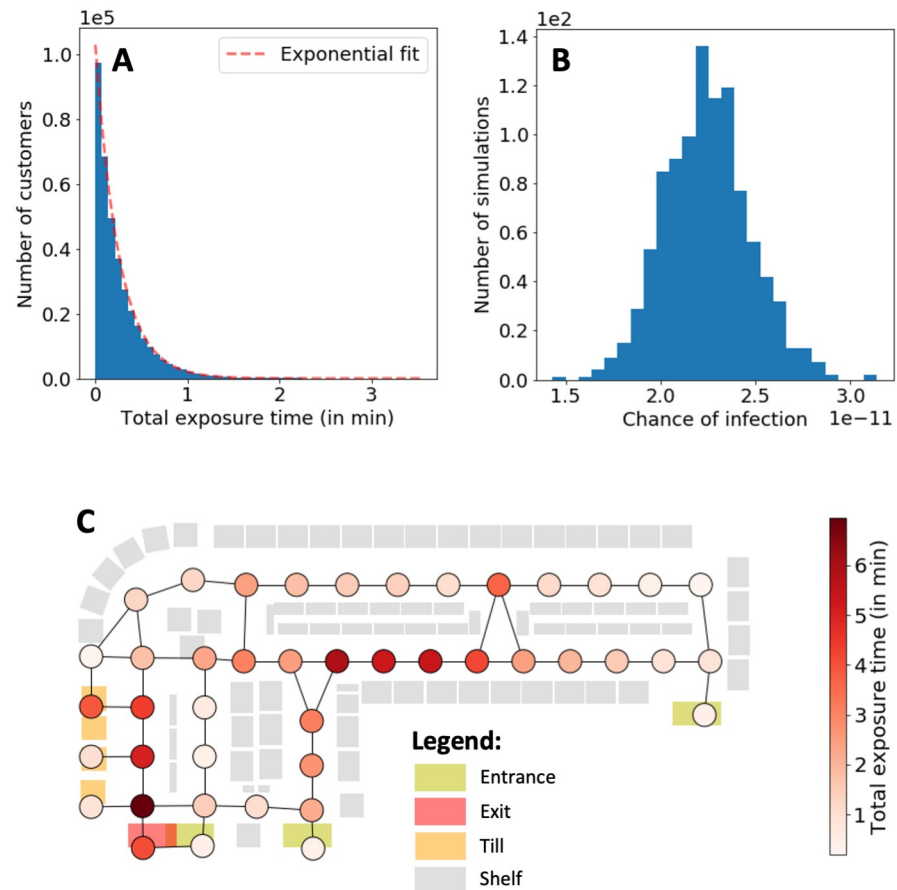


Fig 2. Exposure times and chance of infection from 1000 simulations. (A) Histogram of total customer exposure time for all exposed customers (i.e., susceptible customers with positive exposure time) across 1000 simulations. The distribution can be approximated by an exponential distribution. (B) Histogram of chance of infection per susceptible customer across 1000 simulations. (C) Total exposure time per node. Nodes in the centre and near the tills of the store show significantly higher amount of exposure time than others.

<https://doi.org/10.1371/journal.pone.0249821.g002>

exposure time, we estimate an average of 1.34×10^{-7} infections per day. Each susceptible customer is exposed 0.0452 minutes (or about 2.71 seconds) on average per visit to an infected customer and the estimated a chance of infection is 6.37×10^{-11} . The mean individual exposure time is dominated by a majority of susceptible customers that were not exposed (i.e., who had no exposure with infected customers); only 17.69% of susceptible customers had any exposure. Among the exposed customers, the exposure time appears to be exponentially distributed with a mean exposure time of 0.26 minutes. We note that the majority of exposed customers have a short exposure time of less than 1 minute (see Fig 2A). The longest exposure time of a customer in our simulations is 3.5 minutes. Therefore, if we used an alternative virus transmission model where an infection can only occur after 15 minutes of cumulative exposure time, for example, we would not record any infections in our simulations. The CDC and many other national health agencies at the time of writing [26–28] define a close contact (within 2 metres) to occur when there is 15 minutes of cumulative exposure. However, infections may still occur when a customer has additional exposure in a different setting (e.g., in another shop) during a 24-hour window [27, 29]. Therefore, reducing exposure time is still desirable, even if the maximum exposure time is short in a supermarket.

Table 2. Simulation results. We show the mean and standard deviation of each metric across 1000 simulations.

Metric	Mean	Standard deviation
Number of customers	2143	46.76
Number of susceptible customers	2103	46.52
Number of infected customers	39.28	5.44
Mean number of customers in store	14.96	0.45
Mean shopping time (in min)	5.97	0.06
Total exposure time (in min)	94.98	15.38
Exposure time per sus. customer (in min)	0.0452	0.0073
Exposure time per exposed customer (in min)	0.26	0.27
Percentage of sus. customers with exposure	17.69%	2.40%
Number of infections	1.34×10^{-7}	2.17×10^{-8}
Chance of infection per sus. customer	6.37×10^{-11}	1.03×10^{-11}

<https://doi.org/10.1371/journal.pone.0249821.t002>

The chance of infection per susceptible customer does not show large deviations across simulations (see Fig 2B). We list all metrics that we record in our simulations in Table 2.

Our models also allow us to record the total exposure time for each node v (defined as the sum of the individual exposure times that occurred on v). We see in Fig 2C that near the tills and centre of the store shows the highest amount of exposure time, thereby revealing mobility flow bottlenecks in the store that may be mitigated by changing to a different layout.

Restricting the maximum number of customers in store

In line with UK government guidelines [30], many stores restrict the maximum number C_{\max} of customers in a store. We can add this restriction to our model by simulating a queue outside of the store, where customers queue up if we have C_{\max} or more customers in the store. Customers from the queue only enter the store when the number of customers in the store is below C_{\max} . In our model, the estimated chance of infection and number of infections also decreases significantly when decreasing the maximum number of customers in the store (see Fig 3A+3C). We also note that the mean number of infections plateaus as we increase the C_{\max} beyond 20, as the number of customers typically does not exceed 20 in our simulations.

Reducing customer arrival rate

Another way of reducing the number of customers in the store is to restrict the rate at which customers enter the store. We can incorporate this in our model by varying the arrival rate λ . We see that the chance of infection increases linearly with λ while the number of infections increases quadratically (see Fig 3B+3D). The linear and quadratic scaling are not unsurprising: The number of customers (both infectious customers and susceptible customers) in the store increases linearly with the arrival rate. Therefore, we expect the exposure time (and hence the chance of infection) for each susceptible customer to increase linearly with the arrival rate λ . For the quadratic scaling of the number of infections, note that the total number of infections is the product of the number of susceptible customers and the chance of infection. Both quantities scale linearly with the arrival rate, which gives a quadratic scaling for the number of infections.

Note that restricting C_{\max} is equivalent to reducing the arrival rate λ , as both methods reduce the rate at which customers enter the store.

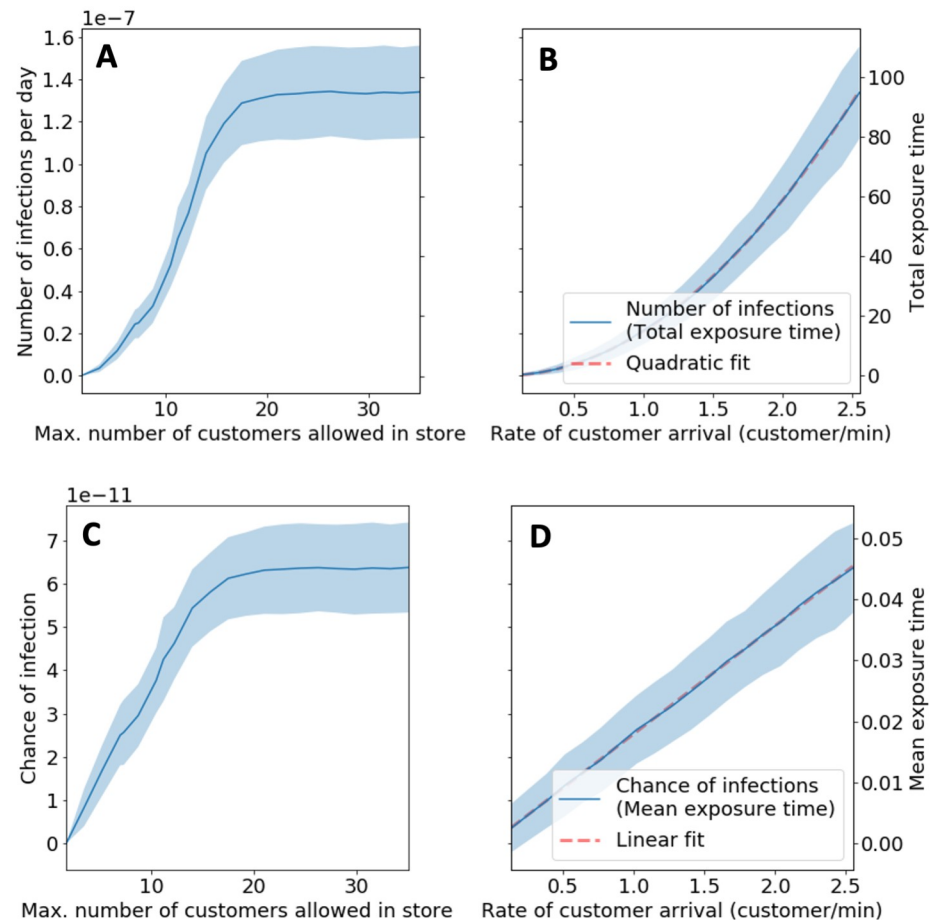


Fig 3. Change in number of infections and chance of infection by reducing maximum number of customers in store or arrival rate. (A + B) Mean number of infections (with the shaded area showing the standard deviation) as a function of maximum number C_{\max} of customers and customer arrival rate λ (respectively). As the number of infections is a linear function of the total exposure time, we also show the total exposure time on the right vertical axis. (C + D) Mean chance of infection for each susceptible customer (with the shaded area showing the standard deviation) as a function of maximum number C_{\max} of customers and customer arrival rate λ (respectively). As the chance of infection is a linear function of the mean exposure time (per susceptible customer), we also show the mean exposure time on the right vertical axis. In subfigures (A) and (C), the mean number of infections and mean chance of infection plateaus, as the number of customers typically does not exceed 20 in our simulations.

<https://doi.org/10.1371/journal.pone.0249821.g003>

Face masks

Similar to [31], we model the implementation of a face mask policy via a reduction in the transmission rate. For example [32], estimated the relative (transmission) risk reduction to be $RRR = 0.17$. We incorporate this by multiplying β with RRR , reducing the number of infections and chance of infection by the same factor: the number of infections decreases from 1.34×10^{-7} to 2.28×10^{-8} ; the chance of infection decreases from 6.37×10^{-11} to 1.08×10^{-11} .

One-way aisle layout

A number of stores have implemented one-way systems to assist with social distancing and potentially redistributing the flow of customers. We can also assess this policy in our framework by changing the store graph to a directed graph, where some edges are uni-directional. For the synthetic store, we show an example one-way aisle layout in Fig 4A which we call the

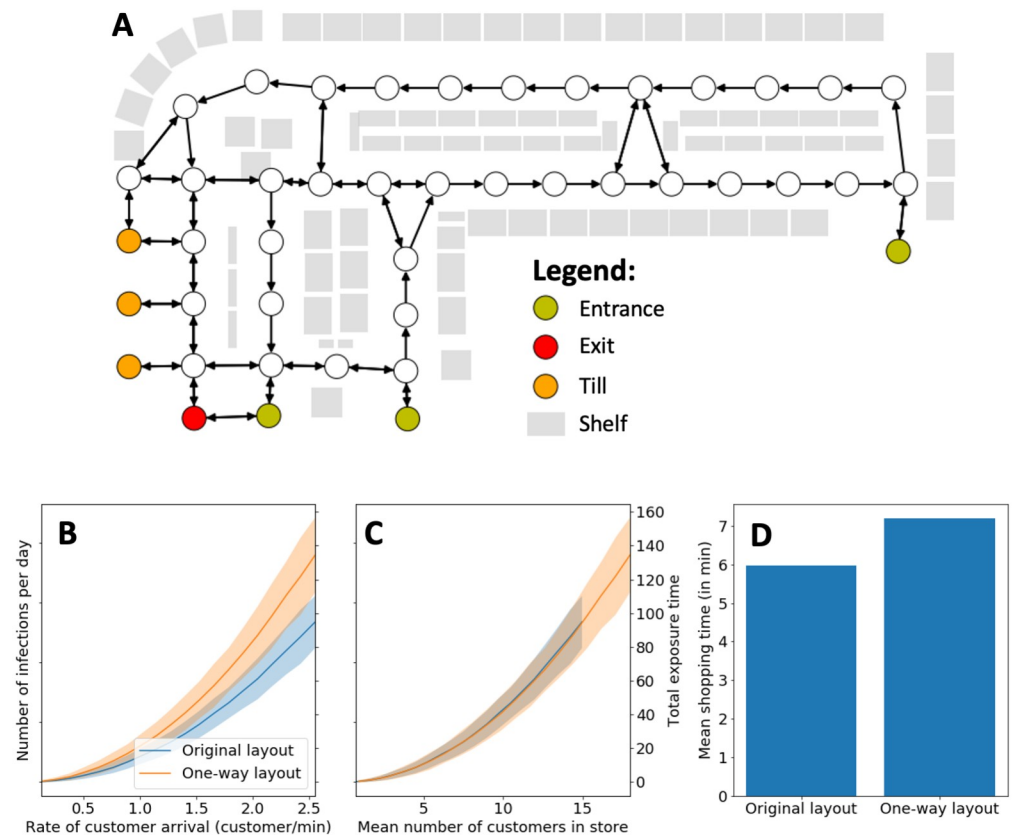


Fig 4. Effect of one-way aisle layout on infections. (A) Store layout with one-way aisles. (B + C) Number of infections in a store as a function of the customer arrival time and mean number of customers (respectively). We show on the right vertical axis the total exposure time, as the number of infections is proportional to the total exposure time in our model. The one-way layout increases the number of infections with the same arrival rate (see subfigure D). It appears that the number of infections mainly depends on how many customers are in the store on average (see subfigure C). (D) Mean customer shopping time. The one-way layout increases the time that customers spend in the store, so more customers are in the store and thereby increase the number of infections.

<https://doi.org/10.1371/journal.pone.0249821.g004>

one-way store layout. We need to change the shopping paths in our data, as they may violate the uni-directionality of the one-way store layout. For each path, we first consider again the node sequence (v_1, \dots, v_{K+3}) from which the path was generated. (As a reminder, v_1 is an entrance node, v_{K+2} is a till node, v_{K+3} is an exit node, and the intermediate nodes v_2, \dots, v_{K+1} are locations where customers bought one or more items.) The corresponding path for the one-way store layout is then a shortest path that visits each of the nodes v_1, \dots, v_{K+3} in sequence in the one-way store layout. In other words, we assume that the one-way layout does not change the order in which customers buy their items; it merely changes the route that they take between items (or between an item and an entrance/till).

According to our model, the one-way layout increases the number of infections (see Fig 4B). To explain this phenomenon, we can look at the mean shopping time between the two layouts. In the one-way store layout, the shopping path necessarily is at least as long as in the original store layout. Therefore, on average, the mean shopping time increases (see Fig 4D), so more customers are in the store at the same time and thus more infections occur. When we fix the mean number of customers in the store, the number of infections is largely the same (see Fig 4C).

Combination of interventions

We can group the interventions that we listed above as follows:

Group 1: Control the in-flow of customers (by restricting C_{\max} or λ)

Group 2: Reduction of virus transmissibility (by implementing a face mask policy)

Group 3: Change of store layout (by using a one-way aisle layout)

Interventions between different groups are independent of each other and we can combine interventions from different groups for increased effectiveness. In [Table 3](#), we list a number of possible interventions and some combinations of them. For example, based on our results above, reducing customer arrival rate by 50% leads to 75% fewer infections and 50% smaller chance of infection. However, combining this intervention with a face mask policy, we achieve a 96% in number of infections and 92% reduction in the chance of infections. Practitioners should seek to find the most effective intervention for each group in their stores, and use the combination of interventions that achieve the largest individual reduction in infections or chance in infection for each group.

Limitations

Our model comes with a number of limitations. Firstly, a customer's path in our model does not depend on other customers, whereas we expect a customer's path to change depending on, for example, the crowdedness of other zones. Unlike in [\[16\]](#), our spatial topology is more coarsely grained, where customers are either in the same zone or not, and we therefore do not take the precise distance between customers into account. We also assume that the chance of infection is proportional to the exposure time, whereas in reality it may be non-linear (e.g., represented by a logistic function to model infectious dosage). In our simulations, we used a constant arrival rate and random shopping paths that do not change with time, while we expect time-varying arrival rates and shopping path distributions in reality. However, it is relatively straightforward to incorporate modifications to the transmission function or to the arrival rate into our agent-based model. Lastly, all results that we present in this article are from simulations on a synthetic data set using a transmission rate β which was estimated in a different setting. Therefore, we anticipate our point estimate for β may not be very accurate, and the concrete results that we presented may have limited generalisation power. The number of infections of store that we estimated should therefore not be taken at face value. Nonetheless, even without an accurate measure of β , we anticipate that the exposure time is a relevant

Table 3. Comparison of interventions and combinations of interventions. We compare possible interventions and combinations of interventions by their effect on the number of infections and chance of infection based on our agent-based model. We can achieve the largest amount of reduction by combining the face mask policy with reducing arrival rate or the maximum number of customers.

Intervention	Relative change in number of infections	Relative change in chance of infection
Restrict maximum number of customers C_{\max} in store to 75% of mean number of customers	-50%	-33%
Reduce customer arrival rate by 50%	-75%	-50%
Face mask policy	-83%	-83%
Face mask policy + reduce arrival rate by 50%	-96%	-92%
Face mask policy + Restrict C_{\max} to 75% of mean number of customers	-92%	-89%
One-way layout	+42%	+42%

<https://doi.org/10.1371/journal.pone.0249821.t003>

metric to measure the relative risk of transmission. Validation of our results is likely to be difficult, as no data on the number of infections in supermarkets exists. A possible way to estimate this is to survey people who have gotten infected and ask where they likely got infected. While this will likely add some biases, such data could be helpful. The closest such data to the best of our knowledge is the survey data by Public Health England on the percentage of infected people who have visited a supermarket prior to getting a positive COVID-19 test [33]. However, one cannot infer from this data directly how many of those people were infected in the supermarket.

Conclusion

We presented a model for modelling virus transmission (in particular, SARS-CoV-2 transmissions, which causes COVID-19, but it is more generally applicable) in supermarkets based on an agent-based model of customers traversing from zone to zone and being exposed to potential virus infection when in the same zone as an infected customer. We measured the risk of virus transmission by the total time that susceptible customers spent in the same zone as infected customers (and called this time the exposure time). We demonstrated the capabilities of the model by applying it to synthetic data with model parameters specific to SARS-CoV-2. We showed how one can use the model to identify hotspots and bottlenecks, estimate the reduction in virus transmissions in different scenarios such as restricting the maximum number of customers in a store or implementing a one-way aisle system. In particular, in our synthetic store, a directed store layout does not help in reducing the exposure time, as it increases the customer shopping time, so more customers are in the store at any given time. The best policy among those that we tested is to restrict the arrival rate of customers or the maximum number of customers together with a mandatory face mask policy; doing so can significantly reduce the number of infections and the chance of getting infected in a supermarket. We invite retailers to use our model to identify bottlenecks that lead to crowded zones as well as to inform them on the best store policy. In future work, it would be interesting to understand which store network features or size of stores lead to lower number of infections and lower chance of infection.

Supporting information

S1 Fig. Sensitivity analysis. We plot the mean number of infections and chance of infection (with the shaded area showing the standard deviation) as a function of (A) + (D) traversal time τ , (B) + (E) proportion p_I of infectious customers, and (C) + (F) transmission parameter β . As the number of infections is a linear function of the total exposure time, we also show the total exposure time on the right vertical axis in subfigures (A)–(C). Similarly, we show the mean exposure time on the right vertical axis in subfigures (D)–(F). (TIF)

S1 Appendix.
(PDF)

Acknowledgments

This work was undertaken as a contribution to the Rapid Assistance in Modelling the Pandemic (RAMP) initiative, coordinated by the Royal Society. We thank David Romano-Critchley and Sumanas Sarma from Sainsbury's for providing the synthetic store data. We also thank Mason A. Porter, Sam D. Howison, Mariano Beguerisse-Díaz, John Fitzgerald, Mattie Landman, Mike Batty, and Philip Wilkinson for helpful discussions.

Author Contributions

Conceptualization: Fabian Ying, Neave O’Clery.

Data curation: Fabian Ying.

Formal analysis: Fabian Ying.

Investigation: Fabian Ying, Neave O’Clery.

Methodology: Fabian Ying, Neave O’Clery.

Project administration: Fabian Ying, Neave O’Clery.

Resources: Fabian Ying.

Software: Fabian Ying.

Supervision: Neave O’Clery.

Validation: Fabian Ying.

Visualization: Fabian Ying.

Writing – original draft: Fabian Ying.

Writing – review & editing: Fabian Ying, Neave O’Clery.

References

1. Europe’s coronavirus lockdown measures compared;. <https://www.politico.eu/article/europes-coronavirus-lockdown-measures-compared/>.
2. Hale T, Angrist N, Cameron-Blake E, Hallas L, Kira B, Majumdar S, et al. Oxford COVID-19 Government Response Tracker. Blavatnik School of Government. 2020; 25.
3. Coronavirus disease (COVID-19): How is it transmitted?;. <https://www.who.int/news-room/q-a-detail/coronavirus-disease-covid-19-how-is-it-transmitted>.
4. Scientific Brief: SARS-CoV-2 and Potential Airborne Transmission;. <https://www.cdc.gov/coronavirus/2019-ncov/more/scientific-brief-sars-cov-2.html>.
5. Coronavirus in-store safety: which supermarkets are doing it best?;. <https://www.thegrocer.co.uk/supermarkets/coronavirus-in-store-safety-which-supermarkets-are-doing-it-best/645177.article>.
6. What Grocery and Food Retail Workers Need to Know about COVID-19;. <https://www.cdc.gov/coronavirus/2019-ncov/community/organizations/grocery-food-retail-workers.html>.
7. Guidance for food businesses on coronavirus (COVID-19);. <https://www.gov.uk/government/publications/covid-19-guidance-for-food-businesses/guidance-for-food-businesses-on-coronavirus-covid-19>.
8. Birnir B, Angheluta L. The build-up of droplet/aerosols carrying the sars-cov-2 coronavirus, in confined spaces. medRxiv. 2020.
9. Shao S, Zhou D, He R, Li J, Zou S, Mallery K, et al. Risk assessment of airborne transmission of COVID-19 by asymptomatic individuals under different practical settings. *Journal of aerosol science*. 2021; 151:105661. <https://doi.org/10.1016/j.jaerosci.2020.105661> PMID: 32968325
10. Vuorinen V, Aarnio M, Alava M, Alopaeus V, Atanasova N, Auvinen M, et al. Modelling aerosol transport and virus exposure with numerical simulations in relation to SARS-CoV-2 transmission by inhalation indoors. *Safety Science*. 2020; 130:104866. <https://doi.org/10.1016/j.ssci.2020.104866> PMID: 32834511
11. Lau Z, Kaouri K, Griffiths IM. Modelling Airborne Transmission of COVID-19 in Indoor Spaces Using an Advection-Diffusion-Reaction Equation. arXiv preprint arXiv:201212267. 2020.
12. Jimenez JL. 2020 COVID-19 Aerosol Transmission Estimator;. <https://docs.google.com/spreadsheets/d/16K1OQkLD4BjgBdO8ePj6ytf-RpPMLJ6aXFg3PrIQBbQ/edit#gid=519189277>.
13. Dai H, Zhao B. Association of infected probability of COVID-19 with ventilation rates in confined spaces: a Wells-Riley equation based investigation. *MedRxiv*. 2020.

14. Buonanno G, Stabile L, Morawska L. Estimation of airborne viral emission: Quanta emission rate of SARS-CoV-2 for infection risk assessment. *Environment International*. 2020; 141:105794. <https://doi.org/10.1016/j.envint.2020.105794> PMID: 32416374
15. Sun C, Zhai Z. The efficacy of social distance and ventilation effectiveness in preventing COVID-19 transmission. *Sustainable cities and society*. 2020; 62:102390. <https://doi.org/10.1016/j.scs.2020.102390> PMID: 32834937
16. Plata S, Sarma S, Lancelot M, Bagrova K, Romano-Critchley D. Simulating human interactions in supermarkets to measure the risk of COVID-19 contagion at scale. *arXiv preprint arXiv:200615213*. 2020.
17. Ying F, Wallis AOG, Beguerisse-Díaz M, Porter MA, Howison SD. Customer mobility and congestion in supermarkets. *Physical Review E*. 2019; 100(6):062304. <https://doi.org/10.1103/PhysRevE.100.062304> PMID: 31962461
18. Ying F. Customer Mobility and Congestion in Supermarkets [D.Phil. thesis]. University of Oxford; 2019.
19. Sorensen H, Bogomolova S, Anderson K, Trinh G, Sharp A, Kennedy R, et al. Fundamental patterns of in-store shopper behavior. *Journal of Retailing and Consumer Services*. 2017; 37:182–194. <https://doi.org/10.1016/j.jretconser.2017.02.003>
20. Supermarket coronavirus opening times: What time do Tesco, Asda and Aldi open and close?;. <https://www.thesun.co.uk/money/11206101/supermarket-coronavirus-open-close/>.
21. Shopping rules: latest restrictions at Asda, Sainsbury's, Tesco & more supermarkets explained;. <https://www.lovemoney.com/news/94072/supermarket-policies-aldi-asda-lidl-tesco-sainsburys-morrisons-waitrose-coronavirus-pandemic>.
22. Office for National Statistics: Coronavirus (COVID-19) Infection Survey;. <https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/datasets/coronaviruscovid19infectionsurveydata>.
23. Tang B, Bragazzi NL, Li Q, Tang S, Xiao Y, Wu J. An updated estimation of the risk of transmission of the novel coronavirus (2019-nCov). *Infectious Disease Modelling*. 2020; 5:248–255. <https://doi.org/10.1016/j.idm.2020.02.001>
24. SimPy 4.0.1;. <https://simpy.readthedocs.io/en/latest/index.html>.
25. Amazon Web Services;. <https://aws.amazon.com/>.
26. Public Health Guidance for Community-Related Exposure;. <https://www.cdc.gov/coronavirus/2019-ncov/php/public-health-recommendations.html>.
27. Operational Considerations for Adapting a Contact Tracing Program to Respond to the COVID-19 Pandemic;. <https://www.cdc.gov/coronavirus/2019-ncov/global-covid-19/operational-considerations-contact-tracing.html>.
28. Guidance for contacts of people with confirmed coronavirus (COVID-19) infection who do not live with the person;. <https://www.gov.uk/government/publications/guidance-for-contacts-of-people-with-possible-or-confirmed-coronavirus-covid-19-infection-who-do-not-live-with-the-person/guidance-for-contacts-of-people-with-possible-or-confirmed-coronavirus-covid-19-infection-who-do-not-live-with-the-person>.
29. Pringle JC, Leikauskas J, Ransom-Kelley S, Webster B, Santos S, Fox H, et al. COVID-19 in a Correctional Facility Employee Following Multiple Brief Exposures to Persons with COVID-19—Vermont, July–August 2020. *Morbidity and Mortality Weekly Report*. 2020; 69(43):1569. <https://doi.org/10.15585/mmwr.mm6943e1> PMID: 33119564
30. Keeping workers and customers safe during COVID-19 in shops and branches; 2020. <https://assets.publishing.service.gov.uk/media/5eb9703de90e07082fa57ce0/working-safely-during-covid-19-shops-041120.pdf>.
31. Li T, Liu Y, Li M, Qian X, Dai SY. Mask or no mask for COVID-19: A public health and market study. *PloS one*. 2020; 15(8):e0237691. <https://doi.org/10.1371/journal.pone.0237691> PMID: 32797067
32. Chu DK, Aki EA, Duda S, Solo K, Yaacoub S, Schünemann HJ, et al. Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: a systematic review and meta-analysis. *The Lancet*. 2020;. [https://doi.org/10.1016/S0140-6736\(20\)31142-9](https://doi.org/10.1016/S0140-6736(20)31142-9) PMID: 32497510
33. Covid: Supermarkets revealed as place visitors and workers most likely to be exposed to coronavirus in PHE data;. <https://www.itv.com/news/2020-11-19/covid-supermarkets-revealed-as-place-visitors-and-workers-are-most-likely-to-be-exposed-to-coronavirus>.