Prediction of local COVID-19 spread in Heidelberg [version 1; peer review: 1 approved, 2 approved with reservations]

Lisa Koeppel1, Claudius Gottschalk1, Andreas Welker2, Britta Knorr2, Claudia M. Denkinger1

1Division of Tropical Medicine, Heidelberg University, Heidelberg, Germany
2Gesundheitsamt des Rhein-Neckar-Kreises, Heidelberg, Germany

Abstract
Since the first identified case of COVID-19 in Wuhan, China, the disease has developed into a pandemic, imposing a major challenge for health authorities and hospitals worldwide. Mathematical transmission models can help hospitals to anticipate and prepare for an upcoming wave of patients by forecasting the time and severity of infections. Taking the city of Heidelberg as an example, we predict the ongoing spread of the disease for the next months including hospital and ventilator capacity and consider the possible impact of currently imposed countermeasures.

Keywords
COVID-19, compartmental model, mortality, hospitalization

This article is included in the Disease Outbreaks gateway.
Introduction
Since the first appearance of the novel coronavirus disease (COVID-19) in Wuhan, China, in December 2019, the disease has spread worldwide to be classified as a pandemic by the WHO on 11 March 2020.

In Germany the first confirmed COVID-19 case was recorded on 28 January 2020. Contact tracing and isolating of affected individuals interrupted the chains of transmission. However, with tourists returning from holidays in high-risk regions of northern Italy combined with a carnival celebration near Cologne, the virus spread to 13 out of 16 federal states within only a month. The exponential increase in new confirmed cases in Germany has reached 16,290 positively tested cases by 20 March 2020, with many more cases presumed undetected due to a possible transmission from minimally symptomatic or asymptomatic persons carrying the virus.

The spread of COVID-19 in other countries, such as China and Italy, indicate the major challenges German hospitals, especially intensive care facilities, will have to face in the weeks to come. On 17 March 2020, the Robert Koch Institute declared the population risk as “high”. For the elderly and individuals with pre-existing health conditions, there is an increased probability of severe disease requiring hospitalization and critical care treatment, as well as higher mortality rates.

The spread of the disease varies locally and depends on several factors, such as the temporal progression of the infection in the specified region, the size and the age distribution of the population, the available capacities and the implemented countermeasures. Thus, to prepare hospitals and assist health authorities in the most efficient way, the resulting challenge and burden on the health care system should be modelled locally. As an exemplary case, in this manuscript we are concerned with the spread of the COVID-19 epidemic for the city of Heidelberg in Germany.

The first COVID-19 positive case in Heidelberg was confirmed on 28 February 2020, with eight more cases sporadically appearing in the following two weeks. The vast majority were home-comers from vacations in a high-risk region in Italy and their immediate contacts. All positive tested individuals were isolated and several hundred contacts were traced and quarantined to prevent further spread. Only thereafter, the total numbers of positively tested COVID-19 cases rose to 20 on March 15, and 34 on March 16.

The capabilities to cope with the increasing burden on the health system on a local scale depend on the preparedness of available local resources in the hospitals and the implemented countermeasures (isolation, quarantine, social distancing) to slow down the disease spread. In such an extraordinary situation, it is crucial for hospitals to rely on future predictions about the upcoming need of beds as well as ventilators. Predictions from mathematical modelling can assist authorities in their decision making on suitable countermeasures to slow down the viral transmission and prepare the health systems.

Methods
The transmission model and the parameters
We model the transmission dynamics of the COVID-19 outbreak by an individual level susceptible (S) – exposed (E) – infected (I) – removed (R; SEIR) compartmental model. The transitions between the compartments are described through the following set of differential equations

\[
\frac{dS}{dt} = -R_0 \gamma SI
\]

\[
\frac{dE}{dt} = R_0 \gamma SI - \kappa E
\]

\[
\frac{dI}{dt} = \kappa E - \gamma I
\]

\[
\frac{dR}{dt} = \gamma I
\]

Where \( \gamma \) is the inverse of the infectious period and \( \kappa \) is the inverse of the incubation period. Extended data, Table 1 lists all parameter choices for our model. For the main part of this manuscript we assume an incubation period of 5 days and an infectious period each of 2.9 days. The infectiousness of the disease is given by \( R_0 \) and is defined as the mean number of people that get infected by a single infected individual over the whole infectious period. Due to the limited data about the COVID-19 pandemic, there is much uncertainty in the estimation of \( R_0 \). As baseline we follow the value of 2.2 obtained from the COVID-19 outbreak on the Diamond Princess ship. For a sensitivity analysis, we also consider a lower \( R_0 \) of 1.5 with respect to the estimation of Du et al. (2020) and a more severe outbreak progression with 3 in accordance to the estimation of \( R_0 \) ranging between 2.2 and 3.6 for the SARS epidemic in 2003.

According to the office for statistics and urban development Heidelberg, for our model we assume a closed population of 150,000, without individual entering or leaving the system. We consider this as a realistic scenario because we do not expect major demographic changes to occur within the six months of the projections.

The true number of initial infected individuals remains unknown due to the asymptomatic progression in some cases. Furthermore, due to limited testing capacities, only high-risk individuals (i.e. people who had prior close contact to infected individuals or people returning from a high risk region according to the WHO definition) were tested, resulting in missing positive cases in the population that resulted from possible community transmission. To capture this uncertainty, we assume that the outbreak starts on 12 March 2020 with 10 initial infectious individuals who have not been isolated and are able to transmit the disease.

Results
For now, we assume that no countermeasures have been set in place and predict the future dynamics of the outbreak for the upcoming period of 6 months. Figure 1 displays the number of infected individuals per day for varying values of \( R_0 \). In the baseline
Figure 1. Number of infected individuals per day for values of $R_0$ between 1.5 and 3 for sensitivity analysis. In all simulations, the number of initial infections was 10 and the mean incubation and infectious period is 5 days.

In the mild outbreak scenario, there will be 77952 (52%) total infections by 12 September 2020, whereas in the severe scenario we will have a rapid growth and a shorter doubling of infections time with 141,084 (94%) infections in total.

For sensitivity analysis, we also investigated, with respect to expert opinion, shortening and prolonging the incubation period [3.5–5.5 days] as well as the infectious period [1.5–4.5 days] (see Extended data, Figure A1). The model showed more robustness to this kind of variation on the overall outbreak curve with a shift in max 42 days for the peak of the outbreak curve and difference of 8856 infections during peak time (compared to 90 days and 12,635 when varying $R_0$).

When varying the number of first infected individuals, the dynamics of the outbreak stay the same. The more infected individuals we start the simulation with, the further we assume the outbreak has progressed already (see Extended data, Figure A2).

Countermeasures for reducing transmission

From these scenarios, it is obvious that lowering $R_0$ leads to a lowering of the final size of the epidemic, prolonging the peaking and minimizing the peak. The latter is the most important, as it avoids an outbreak trajectory, in which a large number of people get sick at the same time with overwhelming effects for the health system.

In the hope to lower the amount of secondary cases resulting from one primary case, on 22 March 2020, the German government announced the implementation of extraordinary measures to reduce social contact and thus transmission for at least two weeks. These include regulations on social distancing such as the closure of schools, kindergartens, shops, restaurants, imposing a visiting ban for hospitals and elderly homes, as well as preventing transmission amplification events and social interactions in public with more than two persons. Such countermeasures can be considered in the model by diminishing the transmission intensity $R_0$.

Figure 2 displays the progression of the outbreak in Heidelberg with respect to different reductions in social contact and thus disease transmission from Monday, March 23 onwards. It illustrates, that restricting social contact will have a major impact on the number of infected individuals per day, on the overall outbreak size and flatten the curve. In fact, a permanent transmission reduction by 30% will yield a total of 89,022 infected individuals by mid-September and 3,877 during peak time (Figure 2a). Decreasing social contact by 50% will massively flatten the curve, with a total of 6,927 and 223 infected individuals at peak time. This will not stop the outbreak, but it will slow it down to a major extent. Considering the transmission on a scale of one year in time, with the permanent reduction of 50% of transmission, we will have 15% of the population of Heidelberg infected.

However, such a permanent lock-down of the population will not be feasible due to economical and psychological health reasons. Thus, in Figure 2b, we assume that after three weeks social life will be partially resumed and transmission rates will increase again by 50%. In this case, a transmission reduction by 70% at first with a following increase in contact rates (transmission...
reduction lowered to 35%) will lead to a total of 48,468 people infected by September. Compared to the worst-case scenario (permanent $R_0 = 2.2$) this is a decrease of 62% (78,106) in the total number of people infected.

If after three weeks contact rates will revert to their original state, the outbreak curve will only be shifted by weeks with no effect on the height nor on the outbreak size (see Extended data, Figure A3).

Deaths, hospitalizations and ventilator capacity

As the severity of the disease and thus the number of expected deaths from the COVID-19 outbreak depends on the age of the individual, we obtained the specific age distribution for the city of Heidelberg: <55 (72.4%), 55–65 (11.06%), 65–75 (7.77%), and >75 years of age (8.77%). Adapting from Dowd et al. (2020)\cite{dowd2020}, the mortality probabilities for each age category are 0.1% (<55), 1.5% (55–65), 7% (65–75), and 20% (>75 years of age). By means of these percentages we calculated the cumulative number of deaths in Heidelberg as proportions of the people in the “Removed (R)” compartment over time (Figure 3) for different scenarios of transmission reduction. We model the delay between end of infectious period and time of death by 6.2 days and follow Sanche et al. (2020)\cite{sanche2020}, who assume that the duration from hospital admittance to death is 11.2 days.

As preparation for the health system to plan the bed and ventilator capacity that needs to be enabled, it is important to predict the number of patients who require hospitalization and ventilation.

It is difficult to give estimates for the proportion of hospitalization from all infected individuals, because there is an inherent bias in the existing data\cite{chastre2020}. In the beginning of an epidemic, people are hospitalized for isolation purposes but not because of their severe symptoms. Besides, due to asymptomatic or minimally symptomatic infections that do not present for testing, the true number of infections remains unknown and thus estimates of hospitalization might be overly pessimistic. For this simulation study, we nevertheless assume, that 5% of all infected individuals need hospitalization. Because the duration of hospitalization varies in the literature, we take the smaller and more conservative estimate of 11.5 days\cite{chen2020,dowd2020,sanche2020} (Figure 4a). As the curves are proportions of the number of infected individuals, changing the proportion of hospitalizations will only change the scale of the plot. A sensitivity analysis reducing the duration of hospitalization to seven days is shown in Figure A3 in the appendix.

For calculating the number of ventilators needed, we assume that 90% of all people who died from the disease were ventilated before. The remaining 10% possess a patient do not resuscitate/ do not intubate (DNR/DNI) status. We further assume that 60% of ventilated die. This leads to 1.5 (0.9/0.6) people ventilated per death count. According to Chen et al. (2020)\cite{chen2020}, the average duration of ventilation is 3-20 days. Supposing a conservative estimate of 7 days, we can model the capacity in the Intensive Care Unit given the calculated death numbers.

According to the Science Media Center Germany, routine capacity for ventilation of the University Hospital Heidelberg is 112 spaces\cite{science_media_center}. Even if we unrealistically assume that all spaces can be used for COVID-19 patients, the capacity will not be enough within the following six months, unless we have a permanent reduction in the disease transmission by 50% (Figure 4b).

Discussion

This manuscript modelled the future of COVID-19 spread within the city of Heidelberg and showed the various possible scenarios over the coming weeks and months. Considering the current exponential growth of cases in Germany\cite{bundesministerium}, we have to assume a current $R_0$ between 2.2 and 3.

Our model has several limitations. Our model does not account for the percentage of people with previous conditions who are...
at increased risk. We also did not include any movement data in our model and did not account for tourists visiting Heidelberg, which we assume to be a realistic scenario given that governmental restrictions will keep movement to a minimum in the following months. We also did not account for deaths due to selective treatment of COVID-19 patients over others with different health conditions when bed capacity is exceeded.

Although the number of infected individuals and therefore the number of deaths can be reduced by prolonging the outbreak, severe impacts on the economy as well as effects on psychological and social life can be expected if restrictions on social behaviour and interactions will be imposed for a longer time. In an optimistic scenario with a permanent reduction in transmission, in twelve months 50% of the population in Heidelberg will be infected. This clearly states the urgency for developing and distributing vaccines among the population.

The predictions of our model are estimations and include inherent uncertainty because the model parameters are derived from only limited clinical and epidemiological data of the COVID-19 outbreak dynamics and estimates thereof. Moreover, the future behaviour of the population, their adherence to movement and social contact restriction and the timely development of a suitable vaccine determine the ongoing of the spread. To better understand, where we are on the epidemic curve, requires more testing. The currently implemented restrictive testing strategy focuses on people returning from high-risk regions and their immediate contacts and does not capture ongoing community transmission. Community transmission is likely going to be more relevant for transmission at this stage of the epidemic as it is more difficult capture and more dispersed. Currently available testing capacity would not allow testing of all persons with respiratory symptoms in the city of Heidelberg, but at a minimum sentinel testing of a representative subset of the population should be considered. Ultimately a seroprevalence survey will be necessary to completely understand the true number of cases in the population.

We anticipate the modelling outcomes can be useful for health authorities in their decision-making process to plan resources and
bed capacities ahead of time and implement suitable preparations to meet the upcoming medical need.

Data availability
Underlying data
All data underlying the results are available as part of the article and no additional source data are required.

Extended data

This project contains the following extended data:
- Figure A1: 500 Simulations with fixed R0=2.2 and changing values for κ-U [0.1818,0.2857] and γ-U [0.2222,0.6666].
- Figure A2: Outbreak simulations with varying number of first infected individuals on 12 March 2020.
- Figure A3: Outbreak simulations with different reductions in transmission from Monday, 23 March 2020.
- Table 1: Choices of model parameters.

Extended are available under the terms of the Creative Commons Attribution 4.0 International license (CC-BY 4.0).

References

5. Personal communication with Andreas Welker.
Open Peer Review

Current Peer Review Status: 🟢 🟢 🟢

Version 1

Reviewer Report 24 August 2020

https://doi.org/10.5256/f1000research.25429.r69089

© 2020 Hatami F. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Farhad Hatami
Centre for Health Informatics, Computation and Statistics, Lancaster Medical School, Lancaster University, Lancaster, UK

In Koppel et al., the authors propose a SEIR (Susceptible, Exposed, Infected, Removed or Recovered) differential equations model estimating the spread of COVID-19 in the city of Heidelberg.

The proposed modelling approach is performed on a closed population with no new entry/leave. As declared by the authors, a more realistic approach could extend the current model with taking into account the scenario where people with different underlying health conditions/immune systems would have an unexpectedly higher rate of susceptibility/infection, or a second scenario where a "recovered" case does not necessarily mean "not susceptible" anymore. However, for the sake of this paper, the deployed model benefits from high robustness which makes it a reliable prediction algorithm.

A sensitivity analysis of the model has been performed via applying on a range of assumptions for countermeasures and infection rates. As a result, changing the number of initially infected individuals at the beginning of the pandemic seems to only shift the distribution linearly with no effect on the peak. Considering the limitations in the intensive care unit, a set of policies have been introduced to help the authorities in Germany.

While using a well-known robust algorithm, and bearing in mind that this is not supposed to be a methodological research, the paper still lacks some validation tests. A validation test on the existing benchmarking experiment can be carried out in several ways. At the time of this review, there might be access to the real data of COVID-19 spread in Heidelberg in order to compare with the estimated results. Alternatively, the same approach (of course with a different rate of daily infection) could be applied on a data of SARS to test the performance of the model.

Is the work clearly and accurately presented and does it cite the current literature?
Yes
Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Partly

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Biostatistics, Machine Learning

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 10 August 2020

https://doi.org/10.5256/f1000research.25429.r68119

© 2020 Basu A. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Arindam Basu
School of Health Sciences, University of Canterbury, Canterbury, UK

The authors used a compartmental model that they referred to as a SEIR model where they divided the population into four compartments: Susceptible (S), Exposed (E), Infected (I), and Removed (R; Removed == "Recovered" or "Died", hence cannot be reinfected, although this assumption could be modified but for the sake of argument, we will accept that those who recovered could not be reinfected and hence removed).

Now the model could be specified differently as their scheme of the equations was not clear to me. The susceptibles pool depleted at a constant rate of exposure (typically mixing rates and rate of contact with another infected person), but they seem to have substituted and mixed up with an SIR model (susceptible, infected, and Recovered/Removed model).

S -> E -> I -> R should be the correct order of their model specification, so one way to formulate would be
\[
\begin{align*}
\frac{dS}{dt} &= - \text{Rate of Exposed} \\
\frac{dE}{dt} &= \text{Rate of Exposed} - \text{Rate of Infected} \\
\frac{dI}{dt} &= \text{Rate of Infected} - \text{Rate of removal} \\
\frac{dR}{dt} &= \text{Rate of accumulation of removed persons from the pool}
\end{align*}
\]

But from reading their formulae it does not appear that they went with these equations, nor did they explain the basis or provide a reproducible code.

Hence, although the epidemic curves look plausible, it is difficult to interpret how they were specified or their true significance. The authors may want to clarify this aspect of the methods section.

**Is the work clearly and accurately presented and does it cite the current literature?**
Partly

**Is the study design appropriate and is the work technically sound?**
Partly

**Are sufficient details of methods and analysis provided to allow replication by others?**
No

**If applicable, is the statistical analysis and its interpretation appropriate?**
Partly

**Are all the source data underlying the results available to ensure full reproducibility?**
No

**Are the conclusions drawn adequately supported by the results?**
No

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Epidemiology

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

https://doi.org/10.5256/f1000research.25429.r68118

© 2020 Mahmoud M. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
Mohammad Reza Mahmoud  
Department of Statistics, Faculty of Science, Fasa University, Fasa, Iran

In this work, the authors predict the ongoing spread of the disease including hospital and ventilator capacity for the next months in the city of Heidelberg and consider the possible impact of currently imposed countermeasures.

This is an interesting contribution to the existing literature, but the paper suffers from several shortcomings listed in the following comments:

- The paper should be checked by a native.
- A Conclusion section should be added.
- The novelty and contribution should be clearly bolded.
- The introduction should be updated by recent researches. The authors should consider some works about statistical methods (including time series models or regression models) that can be applied to model and predict medical datasets.
- Comparative Analysis should be given in more details.
- It’s better to suggest some subjects for future works.

Is the work clearly and accurately presented and does it cite the current literature?  
Yes

Is the study design appropriate and is the work technically sound?  
Yes

Are sufficient details of methods and analysis provided to allow replication by others?  
Yes

If applicable, is the statistical analysis and its interpretation appropriate?  
Yes

Are all the source data underlying the results available to ensure full reproducibility?  
Yes

Are the conclusions drawn adequately supported by the results?  
Partly

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Data Analysis, Statistics, Data Modelling

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.
The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com