Evaluation of the EsteR Toolkit for COVID-19 Decision Support: Sensitivity Analysis and Usability Study

Rieke Alpers^{1*}, MSc; Lisa Kühne^{2*}, MSc; Hong-Phuc Truong³, MSc; Hajo Zeeb², Dr med; Max Westphal¹, Dr rer nat; Sonja Jäckle⁴, PhD

¹Fraunhofer Institute for Digital Medicine MEVIS, Bremen, Germany

³Fraunhofer Institute for Industrial Mathematics ITWM, Kaiserslautern, Germany

⁴Fraunhofer Institute for Digital Medicine MEVIS, Lübeck, Germany

*these authors contributed equally

Corresponding Author:

Rieke Alpers, MSc Fraunhofer Institute for Digital Medicine MEVIS Max-von-Laue-Str. 2 Bremen, 28209 Germany Phone: 49 421 218 ext 59228 Email: rieke.alpers@mevis.fraunhofer.de

Abstract

Background: During the COVID-19 pandemic, local health authorities were responsible for managing and reporting current cases in Germany. Since March 2020, employees had to contain the spread of COVID-19 by monitoring and contacting infected persons as well as tracing their contacts. In the EsteR project, we implemented existing and newly developed statistical models as decision support tools to assist in the work of the local health authorities.

Objective: The main goal of this study was to validate the EsteR toolkit in two complementary ways: first, investigating the stability of the answers provided by our statistical tools regarding model parameters in the back end and, second, evaluating the usability and applicability of our web application in the front end by test users.

Methods: For model stability assessment, a sensitivity analysis was carried out for all 5 developed statistical models. The default parameters of our models as well as the test ranges of the model parameters were based on a previous literature review on COVID-19 properties. The obtained answers resulting from different parameters were compared using dissimilarity metrics and visualized using contour plots. In addition, the parameter ranges of general model stability were identified. For the usability evaluation of the web application, cognitive walk-throughs and focus group interviews were conducted with 6 containment scouts located at 2 different local health authorities. They were first asked to complete small tasks with the tools and then express their general impressions of the web application.

Results: The simulation results showed that some statistical models were more sensitive to changes in their parameters than others. For each of the single-person use cases, we determined an area where the respective model could be rated as stable. In contrast, the results of the group use cases highly depended on the user inputs, and thus, no area of parameters with general model stability could be identified. We have also provided a detailed simulation report of the sensitivity analysis. In the user evaluation, the cognitive walk-throughs and focus group interviews revealed that the user interface needed to be simplified and more information was necessary as guidance. In general, the testers rated the web application as helpful, especially for new employees.

Conclusions: This evaluation study allowed us to refine the EsteR toolkit. Using the sensitivity analysis, we identified suitable model parameters and analyzed how stable the statistical models were in terms of changes in their parameters. Furthermore, the front end of the web application was improved with the results of the conducted cognitive walk-throughs and focus group interviews regarding its user-friendliness.

(JMIR Form Res 2023;7:e44549) doi: 10.2196/44549



²Department of Prevention and Evaluation, Leibniz Institute for Prevention Research and Epidemiology - BIPS, Bremen, Germany

KEYWORDS

COVID-19; public health; decision support tool; sensitivity analysis; web application; usability study

Introduction

Background

In Germany, the local health authorities played a central role in the management and containment of the COVID-19 pandemic. The employees ordered isolation for infected persons and traced contact persons to quarantine them according to the COVID-19 regulations of the Robert Koch Institute (RKI). The occurring infections were reported by each local health authority to the RKI. During the COVID-19 pandemic, the local health authorities focused only on the management of the pandemic, and during periods of high infection numbers, they were even incapable of conducting contact tracing [1].

The efficiency of processes in local health authorities can be increased using decision support tools. In other working areas, such decision support models have already been developed. For example, in hospitals, decision support tools are used to streamline the diagnosis and prognosis of COVID-19 cases [2]. Furthermore, governments can use models for predicting the number of COVID-19 infections to adapt their current COVID-19 regulations [3].

In the EsteR project, we developed a toolkit for local health authorities to assist with their work. It includes statistical models for 5 specific use cases and a web application to support isolation and quarantine decisions. The 3 use cases for infected persons include the infection period, the time when they were most likely infected; the illness period, the time range when their infected contacts will most likely start to show symptoms; and the infectious period, the time when they are most likely infectious to others. Furthermore, statistical models for 2 group use cases were introduced: a model for estimating the further expected infections from a group event after some attendees develop COVID-19 symptoms and a model for calculating the probability that nobody was infected in a childcare group or school class with one or more index cases and only negative results for tested susceptible group members. For each use case, a statistical model was developed based on the COVID-19 properties reported in the literature. The entire toolkit was already introduced in detail in the article by Jäckle et al [4].

Objectives

In this paper, we presented two complementary evaluations of the EsteR toolkit. First, a sensitivity analysis of the model parameters was carried out. In research, sensitivity analyses have been recently used to analyze the COVID-19 transmission dynamics and which regulations allow for control of the pandemic [5-9]. For our statistical models, the aim of the sensitivity analysis was to determine how a variation in the model parameters influences the answers provided in the tools. In the second step, the usability and applicability of the web application were evaluated. Cognitive walk-throughs analyzed the usability of applications for certain tasks and revealed improvements for easy and efficient use as it has already been

https://formative.jmir.org/2023/1/e44549

RenderX

shown for digital applications targeting health outcomes (eg, disease self-management tools [10,11]). The aim of this study's cognitive walk-throughs was to identify critical usability problems and generate recommendations to improve the efficiency and user acceptance of the tool. In addition, focus group interviews were conducted, which are often applied for implementing and evaluating public health interventions (eg, to elaborate on users' expectations of clinical decision support tools [12,13]). We aimed to reveal containment scouts' perceptions and whether the web application was supportive of their decision tasks.

To our knowledge, this is the first evaluation of a support tool for local health authorities for assessing personal transmission risks. It can support other researchers by serving as an example for evaluations of two important aspects: the statistical models in the back end and the usability in the front end. Our sensitivity analysis demonstrates how the implications of a choice of study can be assessed and how the results can be interpreted from the perspective of real-world implications. We support reproducibility by providing all data and code for the analysis in a public GitHub repository. The usability evaluation showed how the combination of cognitive walk-throughs and focus group interviews helped depict the varying requirements of users with different professional backgrounds. Unlike a clinical decision support tool used only by medical personnel, the staff of local health authorities during the COVID-19 pandemic mostly had only minor previous knowledge in virology and had to deal with a hitherto unknown situation and accompanying tasks.

Methods

Sensitivity Analysis for Statistical Models

Overview

We conducted a simulation by varying the model parameters of the 5 statistical models while keeping the input data fixed. The outcomes obtained using each new parameter set were compared with those obtained using the default values. For the incubation time (used for the infection period and infection spread use case) and the serial interval (used for the illness period use case), the literature review described in the study by Kühne et al [14] identified more than one study reporting suitable parameters. Our selection approach to the default parameters for our web application is described in the first chapter of Multimedia Appendix 1 [15-74] and visualized in Figures S1 and S2 in Multimedia Appendix 1. Table 1 summarizes the main aspects of the simulation. In the following section, we briefly explain each use case, where each subsection corresponds to one column in Table 1. Afterward, we describe the dissimilarity metrics used to compare the outcomes and how we visualized the results. For more details, see also the second chapter of Multimedia Appendix 1. The code and extracted study data can be found on GitHub [75].

Table 1. Overview of the simulation strategy for the sensitivity analysis of the 5 statistical models.^a

	Infection period	Infection spread	Illness period	Infectious period	Group quarantine
COVID-19 prop- erty	• Incubation time	Incubation time	Serial interval	Infectious period	Setting-specific trans- mission
Distribution	• Lognormal	• Lognormal	• Gamma	• Gamma	• Beta binomial
Studies	• [15,19-45]	• [15,19-45]	• [16,19,30,34,36,44-73]	• [17]	• [18]
Model parame- ters	 E(X)^b μ*^c 	 E(X) μ* 	 E(X) SD(X) 	 E(X) SD(X)	 P(K>0)^d E(K K>0)^e
Default values (parameter ranges)	 6.3 (3.9-9.1) 5.4 (3-8.5) 	 6.3 (3.9-9.1) 5.4 (3-8.5) 	 5.54 (2.51-8.72) 3.9 (1.21-5.65) 	 12.89 (8-18) 2.84 (1-5) 	 0.3 (0.24-0.37) 3.3 (1-8)
Outcome	 80% HDR^f distribution 	• Expected infec- tions	80% HDRdistribution	80% HDRdistribution	• Probability of no infec- tion
Metric	 1-IoU^g W1^h 	• $\Delta pred^{i}$	1-IoUW1	1-IoUW1	 Δprob^j

 a For each use case, the COVID-19 property and the distribution type for modeling it, the number of studies reporting suitable parameters, the respective 2 model parameters with their respective default value and the assessed range, and the considered model outcomes and their respective metrics are listed.

^bE(X): mean.

 $^{c}\mu^{*}$: median.

^dP(K>0): transmission probability.

^eE(K|K>0): conditional mean of infected persons in case of transmission.

^fHDR: high-density region.

^gIoU: intersection-over-union.

^hW₁: Wasserstein metric.

ⁱΔpred: difference in predicted infections.

^j∆prob: difference in probability.

Simulation Strategy

Infection Period

The underlying distribution in the infection period model was a lognormal distribution for the incubation time [4,15]. For the simulation, the mean parameter μ and the SD parameter σ for a reasonable lognormal distribution were calculated using the lognormal distribution properties for the median μ^* and mean

E(X), namely, $\mu^* = e^{\mu}$ and $E(X) = e^{\mu + \frac{e^2}{2}}$. In the simulations, a scenario of one infected person starting to show symptoms on March 22, 2022, was used. We defined a grid of 100 equidistant values, each from the smallest to the largest reported mean and median identified in the study by Kühne et al [14], and calculated the distribution and the 80% high-density region (HDR) of the infection period of the infected person for each parameter set and the default values.

Infection Spread

RenderX

This model was built using the same incubation time distribution as that of the infection period model. Hence, we used the same parameter ranges in the simulation. A group event scenario with 20 attendees on March 22, 2022, with 3 persons starting to show symptoms until March 26, 2022, was considered. Using the

https://formative.jmir.org/2023/1/e44549

same equidistant parameter sets and default values as for the infection period use case, the total number of expected infections from the group event was calculated.

Illness Period

The illness period of contact persons was modeled using a gamma distribution of the serial interval [4,16]. For this use case, we used the articles from the literature review that reported the mean E(X) and SD(X) of the serial interval. We calculated α and β using the gamma distribution properties $E(X) = \frac{\alpha}{\beta}$ and $SD(X) = \frac{\sqrt{\alpha}}{\beta}$. By choosing the gamma distribution, we implicitly assumed that the serial interval was strictly positive. For the simulation, a scenario with one infected person starting to show symptoms on March 22, 2022, was considered. As with the infection period, a grid of 100 equidistant values for each parameter was defined. All parameter combinations and the default values were used to calculate the distribution and the 80% HDR of the symptom onset period for the first, second, and third contact generation.

Infectious Period

For modeling the infectious period, only one study was available that directly used a gamma distribution [4,17]. For comparability across the use cases, we calculated the mean and SD of that

gamma distribution, as described previously, and defined their own ranges to test them using a similar size as in the illness period use case. As before, a scenario with one infected person starting to show symptoms on March 22, 2022, was considered, and a grid of 100 equidistant values for each parameter was defined. The default parameters and all combinations from that grid were used to calculate the distribution and the 80% HDR for the infectious period.

Risk Assessment for Group Quarantine

In the group quarantine use case, the 2 parameters α and β needed for the beta binomial prior distribution were determined based on the probability P(K>0) that at least one person was infected and the conditional mean number of infected persons in the case of transmission E(K|K>0) [76]. The web application provides a calculation of the probability of no further infections in 2 settings: a childcare group and a school class. Table 1 lists the values for the childcare setting; for the school class, we used default values of P(K>0)=0.12 and E(K|K>0)=1.77 in the EsteR toolkit and conducted the sensitivity analysis for P(K>0) in the range from 0.05-0.21 and for E(K|K>0) in the range from 1-8. In both settings, the ranges of the transmission probability were based on the reported 95% CIs [4,18].

In the simulation, we used 2 typical scenarios for our settings: a childcare group of 14 persons plus 1 index case and a school class of 27 persons plus 2 index cases. For both scenarios, a grid of 100 equidistant values for each parameter and the default values were used to calculate the probability of no further infection based on the changed prior. Furthermore, we conducted a sensitivity analysis for the diagnostic test sensitivity, which influenced the likelihood of the probability calculation. A consistent shift in the range of -0.2 to 0.2 of the day-specific polymerase chain reaction test sensitivity provided by Kucirka et al [74] over all days was simulated, and the probability of no further infections in the childcare scenario was calculated.

Dissimilarity Metrics

For the infection period, illness period, and infectious period use cases, the web application provides an answer sentence containing the 80% HDRs as a range of dates accompanied by a plot of the distributions highlighting the 80% HDRs. Thus, the goal was to test how these outputs change with different model parameters. We decided to compare the 80% HDRs of 1 parameter set (HDR₁) with the 80% HDRs resulting from the default values (HDR₂) using an inverted version of the intersection-over-union (IoU) metric [77]:

$$1 - IoU(HDR_1, HDR_2) = 1 - \frac{|HDR_1 \cap HDR_2|}{|HDR_1 \cup HDR_2|}$$

By inverting the IoU metric, a metric value of 0 indicates a perfect equality of 2 outcomes, and a metric value of 1 corresponds to the situation in which the 80% HDRs do not overlap at all.

These distributions were compared using the Wasserstein metric with p=1 [78], which is defined as $W_1 = \int_{\mathbb{R}} |F(x) - G(x)| dx$ where *F* and *G* denote the cumulative distribution functions of their respective distributions. As the discretized versions of the

https://formative.jmir.org/2023/1/e44549

 $XSI \bullet F($

distributions were used in our calculations, we also estimated the integral as a standardized sum of differences. With W_1 , the changes in the distribution tails were also considered. Although the minimum of 0 of the Wasserstein metric indicates equal distributions, there is no upper bound for the metric as the masses of the 2 distributions can be arbitrarily far away from each other.

For the infection spread, the difference $\Delta pred$ between the number of total symptomatic infections predicted using simulation parameters and predicted using default parameters was determined. In the last use case, the change in the probability of no further infections was calculated as the difference $\Delta prob$ between the probability calculated using simulated and default parameters. Both differences allow for easy and direct interpretation of the changes in the respective outcomes.

Visualization

For all simulations—except the test sensitivity evaluation of the last statistical model—a contour plot was created showing the resulting values; the currently used parameters as red squares; and contour lines with step size 0.2 for the 1-IoU plots, 1 for the W_1 plots, and 0.05 for the $\Delta prob$ plot. In the infection period plot, no contour lines were shown as the $\Delta pred$ values were integers. Furthermore, where available, the published parameter combinations were visualized as different marker shapes according to their period of data collection. For each statistical model, we included the most relevant plot in this paper. All the plots as well as a more detailed description can be found in the third chapter of Multimedia Appendix 1.

Usability Evaluation

Cognitive Walk-Through

For the usability evaluation, we conducted cognitive walk-throughs based on the theory by Lewis and Wharton [79]. The process of a cognitive walk-through includes the definition of a sample task that potential users try to solve. This action sequence is then analyzed to provide feedback to the developers about the users' ability to perform tasks using the web application and which actions likely result in operating errors. The focus was on understanding the learnability of the system for new or occasional users.

For the evaluation, we performed cognitive walk-throughs with potential end users of the web application. The tests included example tasks on contact tracing, one test for each of the 5 use cases (Textbox 1).

The testers were encouraged to speak their thoughts out loud while using the web application. In addition to the example tasks, we asked questions about the usability and applicability of the EsteR toolkit (eg, about the orientation on the starting page and within each tool), as well as the amount and level of detail of the inputs and results. The comprehensiveness of the visualizations was of special interest as these results are the basis for deciding on further handling of the case or contact person. The users were asked how difficult it was to complete the tasks. A 5-point Likert scale was presented indicating the users' impressions of the web application. The describing

attributes included disturbing or supportive, complicated or simple, inefficient or efficient, confusing or clear, and uninteresting or interesting, with the possibility to rate on the scale between these opposites. Moreover, the users were encouraged to provide general feedback. The test was audio recorded, transcribed, and analyzed to provide improvements for the user interface (UI) and suggest a guideline for use.

Textbox 1. Example tasks on contact tracing for each use case.

- 1. A person reports the first symptoms on 17.05. On 14.05, the person was at an event. How do you rate the chance that the person got infected there?
- 2. A mother reports the first symptoms of her children on 18.05. When can she expect to become ill herself?
- 3. A person reports the first symptoms on 13.05. On 10.05, she went on a trip with 3 friends and they sat together in the car for half an hour. How do you rate the chance that the friends got infected?
- 4. On 14.05, a group of 15 adults met. The meeting was held in a closed room and no masks were worn. One person reported the first symptoms on 17.05. How do you rate the chance that people got infected at the meeting?
- 5. Over the weekend, 2 students of a class became ill. The school class of 22 persons met last on 13.05. The infected students were able to contact a few classmates, so 4 tested negative on Sunday. On Monday, the remaining children were given an antigen test, which was negative for all of them. How do you rate the chance that no other children in the class will get sick?

Focus Group Interviews

Focus group interviews are often conducted in public health research as they allow for an exploration of the attitudes, perceptions, and ideas of the target group regarding a topic. The focus group consists of a small group of people with common characteristics, in our case, containment scouts at a local health authority in Germany. The interview is designed to allow for a reflection on the questions asked by the interviewer and a discussion of different answers in the group [80].

We conducted 2 independent focus group interviews with representative containment scouts of the cooperating local health authorities in Berlin-Reinickendorf and Bremerhaven. The containment scouts were asked to share their perceptions in more detail after they conducted the cognitive walk-through and, thus, gained firsthand experience with the web application. The focus was on identifying potential implementation barriers in the local health authorities as well as technical and ergonomic requirements. The containment scouts were asked to rate the effort and benefits of using the web application in their daily work and for additional tasks possibly coming up in the course of the pandemic. The interviews were audio recorded and transcribed. The results were obtained through qualitative content analysis in addition to the impression of the interviewers.

Ethical Considerations

The study participants were recruited from the local health authorities that were already cooperating with the EsteR project during the toolkit development. Before participant recruitment, the local health authorities were given written information about the status of the toolkit and the objectives of its evaluation. Moreover, the study participants were verbally informed about the toolkit and the procedures of the cognitive walk-through and the focus group. All participants gave their consent to take part in the cognitive walk-through and the interviews and agreed to the audio recording. The data were collected, stored, and processed anonymously. The study did not seek ethics approval as the collected data referred to workflows in the specified work setting only and were not considered sensitive to participants' personal interests.

Results

Sensitivity Analysis for Statistical Models

Infection Period

The included studies for the incubation time gathered around a line with $\mu^* \approx E(X) - 1$, recognizable in Figure 1A (corresponding to Figure S3 in Multimedia Appendix 1), and only 7% (2/28) of studies below this line could be considered outliers. All 3 (11%) studies with data from the second half of 2020 and the first half of 2021 were concentrated close to our default model parameters.

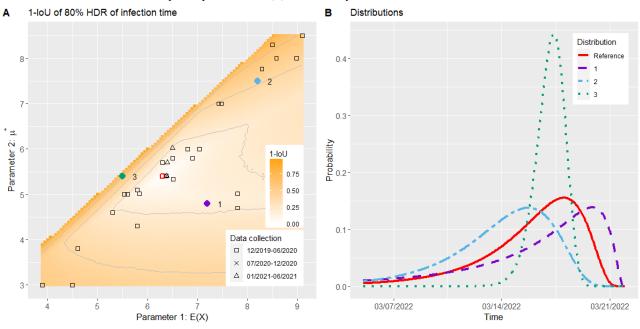
For the infection period, the 1-IoU of the predicted 80% HDRs showed small values along 3 directions starting from the default values: first, toward the lower left corner along the included studies; second, almost horizontally through the 2 outlier studies to the right; and, finally, upward to the right but staying below the upper included studies. The highest 1-IoU values were found along the diagonal, where $\mu^* \approx E(X)$. In the other areas of the plot, the values remained mainly <0.6, and approximately half (15/28, 54%) of the studies lay inside the first contour line of 0.2.

In general, the Wasserstein metric exhibited a similar behavior regarding areas of high and low values (Figure S4 in Multimedia Appendix 1).

The distributions of the 3 parameter sets with different 1-IoU values, together with the distribution from our default parameters, are plotted in Figure 1B. For the lognormal distribution, the density had a narrow, high peak when the mean and median were close together, as can be observed for point 3. In contrast, when the mean and median were farther apart, the SD increased and the peak became smaller. When the mean and median were shifted by the same amount, the distribution width remained intact but was rather shifted. For the parameter set of point 1, the skew to the right of the lognormal distribution increased, whereas for point 2, it decreased. Hence, the Wasserstein metric of point 1 was relatively high compared with that of the 1-IoU, whereas both metrics behaved similarly for point 2.

Alpers et al

Figure 1. Simulation result of the infection period model: 1–intersection-over-union (IoU) of the 80% high-density region (HDR) of the infection time (A) and the distributions derived from 4 specific parameter sets (B). The default parameters are marked in red.

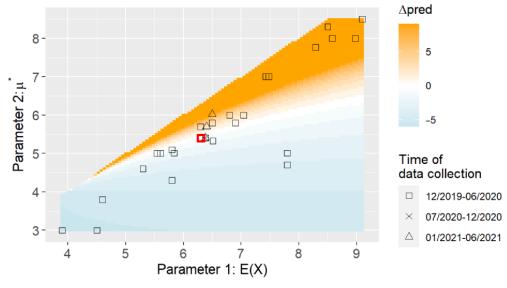


Infection Spread

For the infection spread model, the higher the median or the lower the mean incubation time, the higher the predicted total infections were. However, the change in prediction was far more distinct for the median, which is visible in Figure 2 (Figure S5 in Multimedia Appendix 1). No change in prediction could be observed along the line through the default parameters, with a slope of approximately 0.4. The maximum value of 9 could be found close to the diagonal with $\mu^* \approx E(X)$ fanning out toward the top, whereas the minimum of –6 was located in the lower

left corner. All prediction differences were ≤ 9 because, in the considered setting, the prediction from the default values was 11, and all predictions were limited upward by the total group size of 20. As only mean and median incubation times of at least 3.9 and 3, respectively, were tested and a scenario with 3 observed illnesses after 4 days was considered, the minimal predicted number of total infections was 5, and the minimal difference from the default prediction was -6. The predictions were rounded upward, which resulted in the discrete differences and stepwise changing colors.

Figure 2. Simulation result of the infection spread model: difference in the predicted total infections after a group event. The default parameters are marked in red.



Illness Period

RenderX

As shown in Figure 3A, the included studies for the serial interval spread more than the studies for the incubation time, but they mainly remained within the range of $SD(X) \in [E(X) - E(X)]$

https://formative.jmir.org/2023/1/e44549

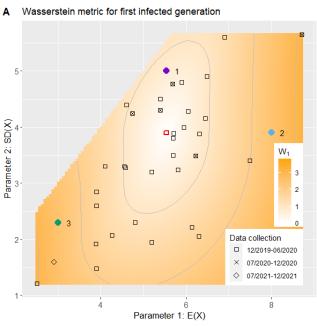
3, E(X) - 0.5]. Only a small number of studies (5/35, 14%) were outside this range and, thus, far away from the default parameters.

In all 3 generations, the 1-IoU of the 80% HDRs of the symptom onset followed a pyramidal course (Figures S6-S8 in Multimedia Appendix 1).

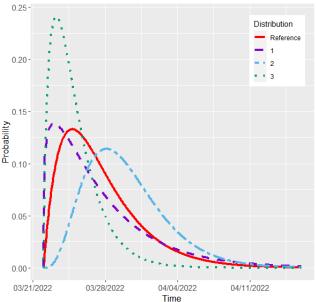
For the Wasserstein metric, Figure 3A (corresponding to Figure S9 in Multimedia Appendix 1) shows that the course of the values has an oval shape stretched vertically. Along the mean, the increase in the Wasserstein metric was not as high as that of the SD. The highest values could be found on the left and right sides of the plot, with a maximum of approximately ≥ 3 . A total of 54% (19/35) of the included studies were within the contour line with a value of 1. For the second and third generations, the shape of the plot remained the same, whereas the horizontal incline increased up to double its original value, as can be seen in Figures S10 and S11 in Multimedia Appendix 1.

The distributions of 1 parameter set with a low Wasserstein metric and 2 with high Wasserstein metrics for the first generation, together with the distribution resulting from our default parameters, are plotted in Figure 3B. Increasing the SD while maintaining the mean shifted the mode and made the distribution more skewed, whereas most of the mass remained in place. When the mean increased, the main mass of the distribution shifted to the right and the skew slightly decreased. For a low mean and SD, the distribution was very narrow and had a high peak around the mean. Hence, both the 1-IoU and the Wasserstein metric were relatively low for point 1 and high for points 2 and 3. In a joint alteration of E(X) and SD(X), the effects may have partly canceled each other out, and the main changes happened in the tails of the distribution. This explains why the 1-IoU followed a pyramidal course, whereas the Wasserstein metric had a vertically stretched oval shape. This effect was strengthened over different generations as, with the convolution we used in our model, the first parameter of the gamma distribution was multiplied by the generation number.

Figure 3. Simulation result of the illness period model: Wasserstein metric for the distribution of the symptom onset of the first contact generation (A) and the distributions of 3 specific parameter sets (B). The default parameters are marked in red. в



Distributions



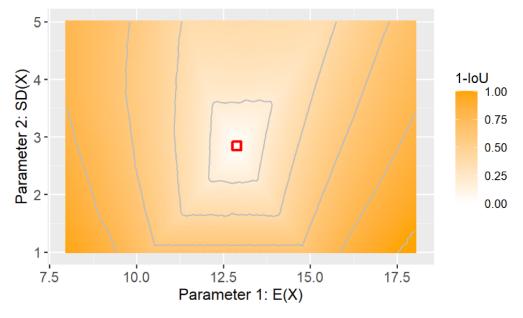
Infectious Period

Both the 1-IoU of the 80% HDRs and the Wasserstein metric plots showed a similar behavior as that of the illness period model as it also used a gamma distribution. The pyramidal shape of the 1-IoU is shown in Figure 4 (Figure S12 in Multimedia

Appendix 1), where the highest values of ≥ 0.8 are in the lower left and right corners. With changes of up to 0.6 in both directions in the mean and SD, the 1-IoU remained <0.2. For the Wasserstein metric plot, see Figure S13 in Multimedia Appendix 1.



Figure 4. Simulation result of the infectious period model: 1-intersection-over-union (IoU) of the 80% high-density region of the infectious period. The default parameters derived from data from the first half of 2020 are marked in red.



Risk Assessment for Group Quarantine

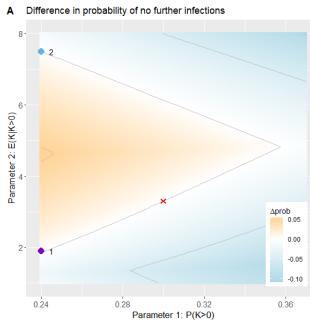
Figure 5A (corresponding to Figure S14 in Multimedia Appendix 1) shows the results of the childcare scenario. In the part of the plot with $E(K|K>0) \in [1.5, 4.7]$, the predicted probability of no further infections increased with a decreasing transmission probability P(K>0) and an increasing conditional mean E(K|K>0). A white line can be detected, which represents the parameter sets with the same probability as the default set. This part was mirrored at the horizontal lines with conditional mean ≈ 1.5 and 4.7, below and above which the predicted probability decreased with increasing E(K|K>0). Similar behavior could be observed in the school scenario; however, the plot was mirrored at a different conditional mean, and the white line with $\Delta prob=0$ had a higher slope (Figure S15 in Multimedia Appendix 1). In Figure 5B, the likelihood and priors with 2 different parameter sets that result in the same predicted probability are plotted. The likelihood had a maximum value of 1 for no further infections and decreased monotonically for higher numbers of newly infected persons. The priors had their maximum value also for no infection but with different behavior. The first prior had higher values for a smaller number of infected persons, whereas the second prior has higher values for >3 infected persons as the higher E(K|K>0), the more mass was shifted from the lower values toward the upper values.

The influence of the test sensitivity on the predicted probability was almost linear, with a slope of approximately 0.2 (Figure S16 in Multimedia Appendix 1).



Alpers et al

Figure 5. Simulation result of the risk assessment for group quarantine model for the childcare scenario (A) and the prior distributions for 2 certain parameter sets and the likelihood (B). The default parameters derived from data from the second half of 2020 are marked in red.



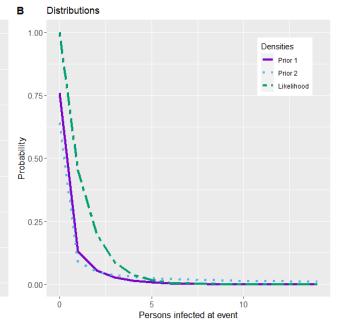
Usability Evaluation

Cognitive Walk-Through

We conducted the cognitive walk-through in the local health authorities of Berlin-Reinickendorf and Bremerhaven with 3 containment scouts each. The users tested the web application version as of May 2022.

Most of the testers (4/6, 67%) were able to complete tasks 1 to 3, apart from testers 2 and 6, who did not manage to apply the task to the tools. Even for the testers who were able to complete the task, the results of task 2 were confusing as the graphic was not intuitive and was described using misleading terms. Tasks 4 and 5 were completed only by testers 4 and 5. For the other testers, the input was not intuitive; therefore, they were not able to complete tasks 4 and 5.

In general, the users rated the web application as supporting and efficient but rather complicated and confusing when using it for the first time (Figure 6). Assuming a Likert scale from 1=disturbing to 5=supporting, the median of 4 refers to most users (4/6, 67%) rating the EsteR toolkit as supporting. Accordingly, the users found the web application to be more *efficient* than *inefficient* (median 4.25) and more interesting than *uninteresting* (median 4.5). Rating the web application as



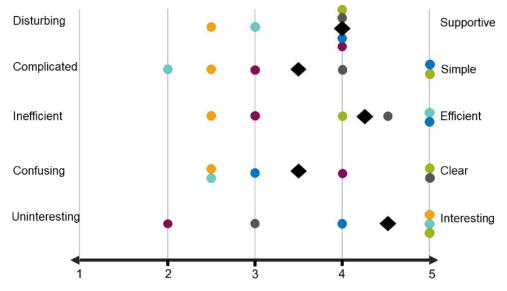
1=simple to 5=complicated revealed heterogeneous results, with a tendency to perceive it as simple (median 3.5). On the Likert scale of 1=clear to 5=confusing, the users tended to find the front end clear (median 3.5). At least 33% (2/6) of the users each agreed in describing the EsteR toolkit as simple, efficient, clear, and interesting. None of the users stated that the web application was fully disturbing, complicated, inefficient, confusing, or uninteresting.

All users (6/6, 100%) stated that the visualization of the results was generally helpful, but some (4/6, 67%) found it hard to interpret. In particular, the interpretation of the 80% and 95% time intervals was unclear to the testers. Some testers (3/6, 50%) expected more detailed information on the event in question, especially for the last 2 tasks. For the testers who tried, it was not possible to generate a PDF document of the results shown in the web application. Nearly all users (5/6, 83%) stated that more guidance and experience would be helpful to use the toolkit efficiently.

The testers indicated that they would adapt the tools for the risk assessment of diffuse events under (partly) unknown circumstances. Moreover, the containment scouts stated that the use cases might help manage groups of >3 people. The detailed results of the 6 cognitive walk-throughs are presented in Multimedia Appendix 2.



Figure 6. Ratings of the web application on a 5-point Likert scale, with 1 indicating full agreement with the negative attribute and 5 indicating full agreement with the positive attribute. The 6 testers are indicated in (1) purple, (2) orange, (3) turquoise, (4) blue, (5) green, and (6) gray. The median is indicated as a black square.



Focus Group Interviews

The interviews were conducted with focus groups in each of the local health authorities of Berlin-Reinickendorf and Bremerhaven. The 2 focus groups consisted of the 3 containment scouts from each local health authority who had previously tested the web application via the cognitive walk-through.

In general, the containment scouts were highly experienced in assessing transmission risks when applying the official guidelines of the federal government agency RKI. The scouts emphasized that they assessed individual cases and, therefore, valued the personal contact with the index case and contact persons via telephone. In this way, they normally gathered detailed information about the event in question. This workflow is time-consuming and could not be maintained during the entire course of the pandemic depending on local transmission clusters, current policies, and guidelines. Therefore, the containment scouts appreciated digital decision tools as they might help in handling the varying workload and human resources. They saw the potential of the web application to offer support, especially in justifying their decision on group quarantine as the toolkit provides a concrete specification of transmission risk. A total of 33% (2/6) of the testers saw the potential of using the web application to train new staff. Unexperienced containment scouts could insert known data and view time frames when transmission may have occurred.

The containment scouts raised concerns that using the web application during common workflows might be too time-consuming. They agreed on preferring short and intuitive features accompanied by less text and explanations. As a trade-off, the functionalities and results needed to be understood. Therefore, more training with the web application would be needed. In one focus group, the idea of a manual with basic explanations of functionalities and their interpretation was discussed.

All the containment scouts (6/6, 100%) were experienced in assessing transmission risk and periods. Most of them (4/6,

67%) expressed that the toolkit's results were in agreement with their expectations based on experience. In total, 33% (2/6) of the containment scouts raised the concern that inexperienced colleagues could tend to rely on the web application's results without questioning their plausibility.

A major concern was the dependence of local health authorities on policies and guidance from the RKI. The local health authorities needed to continuously adapt their daily work to the course of the pandemic. The containment scouts indicated that the web application might not fit their workflows at different stages. Nevertheless, they agreed that contact tracing might be more relevant in the future when the epidemiologic situation and policies change. Further potential was remarked for the first risk assessment for staff in childcare facilities and schools, as well as an adaptation to other infectious diseases.

Discussion

Sensitivity Analysis for Statistical Models

Principal Findings

Infection Period

In the written answers of our web application, the 80% HDRs were rounded to days. The default parameters led to a period of 8 days for the 80% HDR and a 1-IoU of <0.2, which corresponded in almost all cases to at most 2 days that are part of the union but not the intersection of two 80% HDRs. For example, the 80% HDR of point 1 in Figure 1 starts on the same day but ends 1 day after the 80% HDR from the default parameters. Thus, we rated the results as stable regarding the mean and median of the incubation time distribution in the area within the first contour line of 0.2. This area covered half of the studies (14/28, 50%) and especially included the 11% (3/28) of studies that were the most recent in terms of the time of data collection.

Infection Spread

In this use case, the results were more susceptible to parameter deviations than before as, in our underlying group scenario, the prediction was only influenced by the mass of the distribution of the first 4 days. This part of the distribution changed faster than the overall shape and the 80% HDRs when altering the mean or median (Figure 1B). In a different scenario, the deviations in the predicted infections could be even higher in both directions. Hence, it was not possible to determine a range of parameters for which the model is generally stable.

Illness Period

As for the infection period, we reported the 80% HDRs in our web application in days. In this case, the default values led to a period of 9 days, and a 1-IoU of <0.2 again implied for almost all cases a deviation of 2 days at the most. For example, the 80% HDR for point 1 in Figure 3 actually starts and ends on the same day as the one from the default values. As we expected the first generation to be of most interest, we rated the model to be stable regarding the mean and SD of the serial interval in the area within the contour line of 0.2 of the 1-IoU plot of the first contact generation. This area was only slightly smaller than the one from the contour line of W_1 =1 in Figure 3A and covered approximately half (17/35, 49%) of the studies included for this use case.

Infectious Period

In the web application, the default values resulted in a reported 80% HDR with a length of 8 days. For a 1-IoU of <0.2, the difference between two 80% HDRs was again, in most cases, ≤ 2 days. Hence, we again rated the model to be stable regarding the mean and SD in the area within the first contour line of 0.2. Unfortunately, as we were not able to identify other studies that modeled the infectious period, we could not rate the stability of our study selection.

Risk Assessment for Group Quarantine

The nonmonotone course of the prediction when increasing the E(K|K>0) with a fixed P(K>0), as shown in Figure 5A, is due to the complex relationships inside our model: the prediction is derived from the prior distribution for the COVID-19 transmission model for the specific group situations and a likelihood representing the current test situation in the group. In the simulations, the prior was determined by P(K>0) and E(K|K>0), whereas the likelihood remained unchanged as it depended on the number of negative tests per day and the sensitivity of the conducted test types, which were fixed.

To calculate the probability of no further infections, the prior probability for K=0 was divided by the sum of the product of the prior and likelihood. The mass shift in the prior distribution between points 1 and 2 in Figure 5 was caused by the higher E(K|K>0) and resulted in the same probability for the 2 parameter sets in this specific group and test situation. The school scenario was simulated in the same parameter range but with another group and test situation and resulted in a different mirrored behavior than that of the childcare scenario. Thus, it was not possible to determine a range for the model parameters for which the resulting probability is stable as the behavior highly depended on the specific group setting. The variation in the day-dependent test sensitivity shows that it has only a low influence on the predictions, and the model is stable regarding this sensitivity.

Limitations

For each use case, we created a single scenario defined by the input data for which we conducted our simulations. However, many more combinations of different input data could be chosen for the use cases of the infection period for more than one person, the infection spread, and the risk assessment for group quarantine. For the sensitivity analysis, considering all possible combinations would increase the complexity of the simulation and make the interpretation of the results more challenging. Thus, we decided to analyze the infection period only for one person, and we created group situations we typically expect as user input for the infection spread and the risk assessment of group quarantine. Nevertheless, some sensitivity effects might be missed.

Furthermore, the literature review was from March 2022 and resulted in papers that mainly analyzed data from early 2020. Thus, work published since March 2022 is not included in the review or in our analysis. For the infectious period and the risk assessment for group quarantine, we even identified no further studies and, thus, had to define plausible model parameter ranges for the sensitivity analysis on our own.

Comparison With Prior Work

Sensitivity analysis is a widely used tool, and in the literature, it has been used for various COVID-19–related research questions (eg, it has been used to analyze epidemiological models [5,6,8] or understand the transmission of COVID-19 [7]). Weng and Yi [9] conducted a sensitivity analysis to estimate the COVID-19 incubation time based on a systematic literature review. The results of the sensitivity analysis showed that the incubation time varied when excluding or including certain studies. Our statistical model for the infection period and spread was also based on the incubation period. Our data obtained from a literature review [14] also showed a high heterogeneity.

Usability Evaluation

Principal Findings

Cognitive Walk-Through

The cognitive walk-throughs revealed critical issues of the web application and users' understanding of features and results. The concerns and misunderstandings gathered led to an update of features of the web application.

In this study, the cognitive walk-throughs showed that the tools were not easy to understand as not all testers were able to complete the example tasks that included basic information on containment. In total, the participants in our study provided similar feedback that varied according to their personal attitudes toward digital tools in general and their experience with containment. In particular, use cases 4 and 5 on group meetings were not intuitive. Accordingly, the input UI was restructured in such a way that the inputs are aligned with the tools. Furthermore, the testers' feedback showed that the results

XSL•FO RenderX

statements and graphics could be misunderstood or not understood at all. The visualization of the results was generally stated to be helpful, and some testers (2/6, 33%) highlighted that it should be as easy to understand as possible, either self-explanatory or accompanied by a brief explanation. Thus, we added more detailed descriptions to the graphics and explanatory texts. On request, the results section of each tool in the web application provides an example of how to interpret the calculated probabilities.

The cognitive walk-throughs disclosed functionalities that were not working properly and minor bugs. This led to adaptations of the back end. The function to create a PDF document of the results shown in the web application was not working. Thus, this function was removed from the UI.

Most of the testers (5/6, 83%) expected to be able to use the tools properly after trying out some more inputs. However, more guidance would be appreciated. The wording used to guide users in inserting the inputs was adapted. The UI of use cases 4 and 5 was adapted with a clear workflow to enter the inputs and more detailed descriptions of the inputs and results. All users were highly experienced in COVID-19 risk assessment and used their expertise to interpret the results of the tools. Usually, the containment scouts process detailed information on the event in question for a more precise risk assessment. We discussed the number of possible inputs to the models and decided to provide general tools that are applicable to various events. In general, the containment scouts stated that the tools might be helpful in justifying their decision on whether to order quarantine for contact persons.

Focus Group Interviews

The feedback from the focus group interviews showed the containment scouts' perception of the values and concerns regarding the web application and revealed suggestions to improve usability. The suggestions were implemented when applicable within the tools and relevant for the broader purpose of the toolkit.

In the focus group interviews, the containment scouts were open and interested in integrating the digital decision support tool into their workflows when suitable. The scouts highlighted the importance of efficient contact tracing to contain COVID-19 outbreaks. Most of them (5/6, 83%) stated that they would use the EsteR toolkit or recommend it to colleagues in case the policies changed and set contact tracing back in the focus of pandemic control. Tools of the web application should have a high comprehensibility to allow for use in daily work by local health authorities. This indicates that the UI should be as simple and clear as possible. In this way, the web application can also serve as an efficient tool instead of being unnecessarily time-consuming, which some containment scouts were concerned about. The apprehension that the tools might be too time-consuming was already expected beforehand as we consulted an executive of a local health authority regularly while developing the web application.

Another important concern is the rapidly changing policies of the RKI. The statistical models behind the web application are based on the COVID-19 transmission properties obtained from our latest literature research [14]. The tool for calculating the infectiousness period of persons without symptoms was based on the RKI policies until May 2, 2022, which are displayed in the results of the tool. At the time of the interviews, contact tracing was not a priority for pandemic containment. The local health authorities were not in charge of ordering quarantine for suspected infected persons. Therefore, the containment scouts highlighted that the web application should be as flexible as possible to adhere to changing RKI regulations. The tools are generally based on the evidence found in international peer-reviewed literature. In this way, single tools can be adapted and applied to specific cases.

The target group was rather heterogeneous in terms of experience with knowledge of COVID-19 transmission risks and with digital tools. In the interviews, the containment scouts discussed the idea of using the web application to familiarize new employees with the work tasks and support the decision-making process. To address this suggestion, we simplified the tools in the web application and added more detailed descriptions and helpful explanations on request. In this way, the web application can be used intuitively for the first time without reading a manual or receiving further training.

Limitations

We expected more differential feedback for the use of the web application as each local health authority in Germany works slightly differently. Only 6 containment scouts from 2 local health authorities participated in the cognitive walk-through. Unfortunately, it was not possible to reach out to another local health authority for the usability evaluation.

Furthermore, user feedback was limited to their first impressions and whether they could solve the tasks on the first or second try. Related to the feedback that they would prefer to try out some of the web application functionalities intuitively and get an impression of how to use it, we will provide the web application to the users before conducting the cognitive walk-throughs next time. We would expect more specific feedback and further thoughts on implementation after users have already tried it out. However, users should not already be familiar with the web application to gather their perceptions for intuitive use.

Furthermore, the usability evaluation was limited to only 6 participants from 2 local health authorities.

Our results might be biased as the containment scouts we approached were highly experienced with contact tracing. Their perceptions of efficient decision support coincided with feedback on guidance and explanations. Nevertheless, the participants had different backgrounds using digital tools. Therefore, we assessed different levels of expertise in working with the UI. However, a major limitation is that the containment scouts were not mainly working on contact tracing at the time of the interview owing to the currently valid regulations. By that time, the RKI policy did not focus on contact tracing, and persons suspected of being infected were quarantined. Therefore, the containment scouts had to think back to the time when they had to justify their decision on whether to order quarantine. Nevertheless, they expected the regulations to be adapted to the

XSL•FO

course of the pandemic, so contact tracing might be important again during times of higher infection numbers or higher hospitalization rates.

Comparison With Prior Work

Although the cognitive walk-through method has been applied for UI evaluation of digital applications in general [79], it has recently been applied to digital applications targeting health outcomes. For example, user-centered cognitive walk-throughs have been applied to evaluate a diabetes self-management tool. This method has been shown to be sensitive in identifying critical usability problems and generating recommendations to improve the effectiveness, efficiency, and user acceptance of the tool [10]. Moreover, a decision support tool in the clinical context of self-medication has recently been evaluated through cognitive walk-throughs. Dargan [11] found that their tool was generally understandable and helpful for users. The participants provided recommendations to improve content and layout that the developers could adapt to improve the usability and satisfaction of the decision support tool.

Focus group interviews have been applied for usability evaluation (eg, of clinical decision support tools [12,13]). Kastner et al [12] assessed physicians' attitudes toward an osteoporosis disease management tool and consequently designed a functional prototype of the tool. The interviews revealed the importance of the use of the tool being as time-saving and handy as possible, as well as specific recommendations for features to be implemented. Ahearn and Kerr [13] conducted focus group interviews to explore the intended users' attitudes and expectations for integrating such tools into their workflows. The participants stated that the output should be as narrow as possible to not overwhelm the user with less important information.

Accordingly, in our study, we identified functionalities that were not working properly and features that led to misuse and misunderstanding, obtained recommendations for layout improvements, and obtained feedback that the web application would help familiarize the user with the context and support the decision-making process. The containment scouts highlighted the importance of efficient contact tracing to contain COVID-19 outbreaks. We assessed similar concerns as in the focus group interviews conducted by Kastner et al [12] and Ahearn and Kerr [13], namely, that the features need to be easy and fast to use to implement the tool in the workflow.

Conclusions

In this study, we validated every statistical model of the EsteR toolkit through a sensitivity analysis of the model parameters derived from a literature review [14]. The results of this simulation study revealed how stable the statistical models used are and, thus, how carefully their parameters must be chosen. For the single infected person use cases (infection period, illness period, and infectious period), we were able to identify ranges for all model parameters for which the results were stable. However, the infection spread for groups was very sensitive to parameter changes, and this use case, as well as the risk assessment of group quarantine, highly depended on the specific group setting. These 2 use cases would benefit the most from an update of the model parameters to the exact situation at hand. Furthermore, the data sets provided by the literature review showed a high heterogeneity.

In the second step, we conducted cognitive walk-throughs and focus group interviews with containment scouts to evaluate the usability of the web application. On the basis of the results, the UI of the web application was adapted to make it more intuitive and easier to learn and reduce the time of use.

For now, the EsteR toolkit will not be improved further as the COVID-19 pandemic does not play a major role for local health authorities anymore. However, in the case of a new pandemic arising, the toolkit can be easily adapted and extended as soon as studies providing pathogen properties such as incubation time are published. This evaluation study quantifies the general usefulness of the EsteR toolkit and demonstrates the extensive evaluation of a decision support tool in the public health sector, providing all data and code for reproducing the sensitivity analysis.

Acknowledgments

The authors thank Stefanie Grimm, Peter Schmitz, and Amelie Lucker from the Fraunhofer Institute for Industrial Mathematics (ITWM) and Sophia Brüssermann from the Leibniz Institute for Preventive Research and Epidemiology for their fruitful collaboration in the EsteR project. Furthermore, the authors thank the local health authorities in Bremerhaven and Berlin-Reinickendorf for evaluating their web application in the cognitive walk-throughs and focus group interviews. The authors also thank 2 anonymous reviewers for critically reading their manuscript and providing valuable comments for improvement. This work was funded by the German Federal Ministry of Education and Research (*Bundesministerium für Bildung und Forschung* [BMBF]; Project EsteR; funding code: 13GW0542).

Data Availability

The analyzed data extracted from the included studies on incubation time and serial interval and the code used for the sensitivity analysis are available in the GitHub repository [75].

Authors' Contributions

RA implemented the simulation for the sensitivity analysis; evaluated the results; and wrote the Methods, Results, and Discussion sections for the statistical model evaluation, as well as the simulation report in Multimedia Appendix 1. LK planned, conducted, and evaluated the cognitive walk-throughs and focus group interviews and wrote the Methods, Results, and Discussion sections

for the usability evaluation. HPT implemented the metrics for the sensitivity analysis and wrote the Methods, Results, and Discussion sections for the statistical model evaluation. HZ supported the development of the methodological concept and the interpretation of the results of the usability evaluation and reviewed the first manuscript draft. MW supported the development of the methodological concept and the interpretation of the results of the statistical model evaluation and reviewed the first manuscript draft. SJ created the concept of the paperwork, wrote the Introduction section, reviewed the first manuscript draft, and conducted a literature search for related work on sensitivity analysis.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Simulation report. [PDF File (Adobe PDF File), 859 KB-Multimedia Appendix 1]

Multimedia Appendix 2

Results of the cognitive walk-throughs. [PDF File (Adobe PDF File), 159 KB-Multimedia Appendix 2]

References

- Linden M, Dehning J, Mohr SB, Mohring J, Meyer-Hermann M, Pigeot I, et al. Case numbers beyond contact tracing capacity are endangering the containment of COVID-19. Dtsch Arztebl Int 2020 Nov 13;117(46):790-791 [FREE Full text] [doi: 10.3238/arztebl.2020.0790] [Medline: 33533714]
- 2. Jiang X, Coffee M, Bari A, Wang J, Jiang X, Huang J, et al. Towards an artificial intelligence framework for data-driven prediction of coronavirus clinical severity. Comput Materials Continua 2020;63(1):537-551 [doi: 10.32604/cmc.2020.010691]
- 3. Mandal M, Jana S, Nandi SK, Khatua A, Adak S, Kar T. A model based study on the dynamics of COVID-19: prediction and control. Chaos Solitons Fractals 2020 Jul;136:109889 [doi: 10.1016/j.chaos.2020.109889] [Medline: 32406395]
- 4. Jäckle S, Alpers R, Kühne L, Schumacher J, Geisler B, Westphal M. EsteR a digital toolkit for COVID-19 decision support in local health authorities. Stud Health Technol Inform 2022 Aug 17;296:17-24 [doi: <u>10.3233/SHTI220799</u>] [Medline: <u>36073484</u>]
- Ma C, Li X, Zhao Z, Liu F, Zhang K, Wu A, et al. Understanding dynamics of pandemic models to support predictions of COVID-19 transmission: parameter sensitivity analysis of SIR-type models. IEEE J Biomed Health Inform 2022 Jun;26(6):2458-2468 [doi: 10.1109/jbhi.2022.3168825]
- Hussain T, Ozair M, Ali F, Rehman SU, Assiri TA, Mahmoud EE. Sensitivity analysis and optimal control of COVID-19 dynamics based on model. Results Phys 2021 Mar;22:103956 [FREE Full text] [doi: 10.1016/j.rinp.2021.103956] [Medline: 33623733]
- Biglarbeigi P, Ng KY, Finlay D, Bond R, Jing M, McLaughlin J. Sensitivity analysis of the infection transmissibility in the UK during the COVID-19 pandemic. PeerJ 2021;9:e10992 [FREE Full text] [doi: 10.7717/peerj.10992] [Medline: 33665041]
- Wang Z, Aydin O. Sensitivity analysis for COVID-19 epidemiological models within a geographic framework. In: Proceedings of the 1st ACM SIGSPATIAL International Workshop on Modeling and Understanding the Spread of COVID-19. 2020 Presented at: SIGSPATIAL '20: 28th International Conference on Advances in Geographic Information Systems; Nov 3, 2020; Seattle WA USA [doi: 10.1145/3423459.3430755]
- 9. Weng Y, Yi GY. Estimation of the COVID-19 mean incubation time: systematic review, meta-analysis, and sensitivity analysis. J Med Virol 2022 Sep 16;94(9):4156-4169 [FREE Full text] [doi: 10.1002/jmv.27841] [Medline: 35524338]
- Georgsson M, Staggers N, Årsand E, Kushniruk A. Employing a user-centered cognitive walkthrough to evaluate a mHealth diabetes self-management application: a case study and beginning method validation. J Biomed Inform 2019 Mar;91:103110 [FREE Full text] [doi: 10.1016/j.jbi.2019.103110] [Medline: 30721757]
- 11. Dargan C. Patient-facing antihemophilic factor clinical decision support tool. SHARE @ Children's Mercy. 2022 May 3. URL: <u>https://scholarlyexchange.childrensmercy.org/researchdays/GME_Research_Days_2022/ResearchDay2/10</u> [accessed 2022-06-22]
- Kastner M, Li J, Lottridge D, Marquez C, Newton D, Straus SE. Development of a prototype clinical decision support tool for osteoporosis disease management: a qualitative study of focus groups. BMC Med Inform Decis Mak 2010 Jul 22;10(1):40 [FREE Full text] [doi: 10.1186/1472-6947-10-40] [Medline: 20650007]
- 13. Ahearn MD, Kerr SJ. General practitioners' perceptions of the pharmaceutical decision-support tools in their prescribing software. Med J Aust 2003 Jul 07;179(1):34-37 [doi: 10.5694/j.1326-5377.2003.tb05415.x] [Medline: 12831382]

- Kühne L, Brüssermann S, Santis K, Zeeb H. EsteR Decision support for German health departments by risk modelling in order to contain the COVID-19 pandemic. OSF Preprint posted online on January 31, 2022 [FREE Full text] [doi: 10.31219/osf.io/vq64r]
- Xin H, Wong JY, Murphy C, Yeung A, Taslim Ali S, Wu P, et al. The incubation period distribution of coronavirus disease 2019: a systematic review and meta-analysis. Clin Infect Dis 2021 Dec 16;73(12):2344-2352 [doi: <u>10.1093/cid/ciab501</u>] [Medline: <u>34117868</u>]
- Son H, Lee H, Lee M, Eun Y, Park K, Kim S, et al. Epidemiological characteristics of and containment measures for COVID-19 in Busan, Korea. Epidemiol Health 2020 Jul;42(7):e2020035-e2020802 [FREE Full text] [doi: 10.4178/epih.e2020035] [Medline: 32512664]
- 17. He X, Lau EH, Wu P, Deng X, Wang J, Hao X, et al. Temporal dynamics in viral shedding and transmissibility of COVID-19. Nat Med 2020 May 15;26(5):672-675 [doi: 10.1038/s41591-020-0869-5] [Medline: 32296168]
- Schoeps A, Hoffmann D, Tamm C, Vollmer B, Haag S, Kaffenberger T, et al. Surveillance of SARS-CoV-2 transmission in educational institutions, August to December 2020, Germany. Epidemiol Infect 2021 Sep 22;149:e213 [doi: 10.1017/s0950268821002077]
- 19. Böhm S, Woudenberg T, Chen D, Marosevic DV, Böhmer MM, Hansen L, et al. Epidemiology and transmission characteristics of early COVID-19 cases, 20 January–19 March 2020, in Bavaria, Germany. Epidemiol Infect 2021 Mar 02;149:e65 [doi: 10.1017/s0950268821000510]
- Cai Y, Liu J, Yang H, Wang M, Guo Q, Huang D, et al. Association between incubation period and clinical characteristics of patients with COVID-19. J Int Med Res 2020 Sep 23;48(9):300060520956834 [FREE Full text] [doi: 10.1177/0300060520956834] [Medline: 32967488]
- 21. Dai J, Yang L, Zhao J. Probable longer incubation period for elderly COVID-19 cases: analysis of 180 contact tracing data in Hubei province, China. Risk Manage Healthcare Policy 2020 Aug; Volume 13:1111-1117 [doi: 10.2147/rmhp.s257907]
- 22. Deng Y, You C, Liu Y, Qin J, Zhou XH. Estimation of incubation period and generation time based on observed length-biased epidemic cohort with censoring for COVID-19 outbreak in China. Biometrics 2021 Sep 28;77(3):929-941 [FREE Full text] [doi: 10.1111/biom.13325] [Medline: 32627172]
- 23. Elias C, Sekri A, Leblanc P, Cucherat M, Vanhems P. The incubation period of COVID-19: a meta-analysis. Int J Infect Dis 2021 Mar;104:708-710 [FREE Full text] [doi: 10.1016/j.ijid.2021.01.069] [Medline: 33548553]
- 24. Hu S, Wang W, Wang Y, Litvinova M, Luo K, Ren L, et al. Infectivity, susceptibility, and risk factors associated with SARS-CoV-2 transmission under intensive contact tracing in Hunan, China. Nat Commun 2021 Mar 09;12(1):1533 [FREE Full text] [doi: 10.1038/s41467-021-21710-6] [Medline: 33750783]
- Huang S, Li J, Dai C, Tie Z, Xu J, Xiong X, et al. Incubation period of coronavirus disease 2019: new implications for intervention and control. Int J Environ Health Res 2022 Aug 04;32(8):1707-1715 [doi: <u>10.1080/09603123.2021.1905781</u>] [Medline: <u>33818217</u>]
- 26. Ki M, Task Force for 2019-nCoV. Epidemiologic characteristics of early cases with 2019 novel coronavirus (2019-nCoV) disease in Korea. Epidemiol Health 2020 Feb 09;42:e2020007 [FREE Full text] [doi: 10.4178/epih.e2020007] [Medline: 32035431]
- 27. Kong T. Longer incubation period of coronavirus disease 2019 (COVID-19) in older adults. Aging Med (Milton) 2020 Jun 22;3(2):102-109 [FREE Full text] [doi: 10.1002/agm2.12114] [Medline: 32661509]
- 28. Li L, Han ZG, Qin PZ, Liu WH, Yang Z, Chen ZQ, et al. Transmission and containment of the SARS-CoV-2 Delta variant of concern in Guangzhou, China: a population-based study. PLoS Negl Trop Dis 2022 Jan 5;16(1):e0010048 [FREE Full text] [doi: 10.1371/journal.pntd.0010048] [Medline: 34986169]
- 29. Linton NM, Kobayashi T, Yang Y, Hayashi K, Akhmetzhanov AR, Jung SM, et al. Incubation period and other epidemiological characteristics of 2019 novel coronavirus infections with right truncation: a statistical analysis of publicly available case data. J Clin Med 2020 Feb 17;9(2):538 [FREE Full text] [doi: 10.3390/jcm9020538] [Medline: 32079150]
- 30. Ma S, Zhang J, Zeng M, Yun Q, Guo W, Zheng Y, et al. Epidemiological parameters of COVID-19: case series study. J Med Internet Res 2020 Oct 12;22(10):e19994 [FREE Full text] [doi: 10.2196/19994] [Medline: 33001833]
- McAloon C, Collins Á, Hunt K, Barber A, Byrne AW, Butler F, et al. Incubation period of COVID-19: a rapid systematic review and meta-analysis of observational research. BMJ Open 2020 Aug 16;10(8):e039652 [FREE Full text] [doi: 10.1136/bmjopen-2020-039652] [Medline: 32801208]
- 32. Men K, Li Y, Wang X, Zhang G, Hu J, Gao Y, et al. Estimate the incubation period of coronavirus 2019 (COVID-19). Comput Biol Med 2023 May;158:106794 [FREE Full text] [doi: 10.1016/j.compbiomed.2023.106794] [Medline: 37044045]
- Qin J, You C, Lin Q, Hu T, Yu S, Zhou XH. Estimation of incubation period distribution of COVID-19 using disease onset forward time: a novel cross-sectional and forward follow-up study. Sci Adv 2020 Aug 14;6(33):eabc1202 [FREE Full text] [doi: 10.1126/sciadv.abc1202] [Medline: 32851189]
- 34. Ren X, Li Y, Yang X, Li Z, Cui J, Zhu A, et al. Evidence for pre-symptomatic transmission of coronavirus disease 2019 (COVID-19) in China. Influenza Other Respir Viruses 2021 Jan 07;15(1):19-26 [FREE Full text] [doi: 10.1111/irv.12787] [Medline: 32767657]
- 35. Tan WY, Wong LY, Leo YS, Toh MP. Does incubation period of COVID-19 vary with age? A study of epidemiologically linked cases in Singapore. Epidemiol Infect 2020 Sep 02;148:e197 [doi: 10.1017/s0950268820001995]

- Wang X, Zhou Q, He Y, Liu L, Ma X, Wei X, et al. Nosocomial outbreak of COVID-19 pneumonia in Wuhan, China. Eur Respir J 2020 Jun 04;55(6):2000544 [FREE Full text] [doi: 10.1183/13993003.00544-2020] [Medline: 32366488]
- Wang X, Pan Y, Zhang D, Chen L, Jia L, Li X, et al. Basic epidemiological parameter values from data of real-world in mega-cities: the characteristics of COVID-19 in Beijing, China. BMC Infect Dis 2020 Jul 20;20(1):526 [FREE Full text] [doi: 10.1186/s12879-020-05251-9] [Medline: 32689956]
- 38. Wang P, Lu J, Jin Y, Zhu M, Wang L, Chen S. Statistical and network analysis of 1212 COVID-19 patients in Henan, China. Int J Infect Dis 2020 Jun;95:391-398 [FREE Full text] [doi: 10.1016/j.ijid.2020.04.051] [Medline: 32339715]
- Wei Y, Wei L, Liu Y, Huang L, Shen S, Zhang R, et al. Comprehensive estimation for the length and dispersion of COVID-19 incubation period: a systematic review and meta-analysis. Infection 2022 Aug 18;50(4):803-813 [FREE Full text] [doi: 10.1007/s15010-021-01682-x] [Medline: 34409563]
- 40. Wu J, Huang Y, Tu C, Bi C, Chen Z, Luo L, et al. Household transmission of SARS-CoV-2, Zhuhai, China, 2020. Clin Infect Dis 2020 Nov 19;71(16):2099-2108 [FREE Full text] [doi: 10.1093/cid/ciaa557] [Medline: 32392331]
- Xiao Z, Xie X, Guo W, Luo Z, Liao J, Wen F, et al. Examining the incubation period distributions of COVID-19 on Chinese patients with different travel histories. J Infect Dev Ctries 2020 Apr 30;14(4):323-327 [FREE Full text] [doi: 10.3855/jidc.12718] [Medline: 32379707]
- 42. Xiao Z, Guo W, Luo Z, Liao J, Wen F, Lin Y. Examining geographical disparities in the incubation period of the COVID-19 infected cases in Shenzhen and Hefei, China. Environ Health Prev Med 2021 Jan 18;26(1):10 [FREE Full text] [doi: 10.1186/s12199-021-00935-3] [Medline: 33461491]
- 43. Zaki N, Mohamed EA. The estimations of the COVID-19 incubation period: a scoping reviews of the literature. J Infect Public Health 2021 May;14(5):638-646 [FREE Full text] [doi: 10.1016/j.jiph.2021.01.019] [Medline: 33848893]
- 44. Zhao H, Zhang Z, Lun W, Chen Z, Lu X, Li J, et al. Transmission dynamics and successful control measures of SARS-CoV-2 in the mega-size city of Guangzhou, China. Medicine (Baltimore) 2021 Dec 03;100(48):e27846 [FREE Full text] [doi: 10.1097/MD.000000000027846] [Medline: 35049185]
- 45. Zhao S, Tang B, Musa SS, Ma S, Zhang J, Zeng M, et al. Estimating the generation interval and inferring the latent period of COVID-19 from the contact tracing data. Epidemics 2021 Sep;36:100482 [FREE Full text] [doi: 10.1016/j.epidem.2021.100482] [Medline: 34175549]
- 46. Aghaali M, Kolifarhood G, Nikbakht R, Saadati HM, Hashemi Nazari SS. Estimation of the serial interval and basic reproduction number of COVID-19 in Qom, Iran, and three other countries: a data-driven analysis in the early phase of the outbreak. Transbound Emerg Dis 2020 Nov 16;67(6):2860-2868 [FREE Full text] [doi: 10.1111/tbed.13656] [Medline: 32473049]
- 47. Bae S, Kim H, Jung TY, Lim JA, Jo DH, Kang GS, et al. Epidemiological characteristics of COVID-19 outbreak at fitness centers in Cheonan, Korea. J Korean Med Sci 2020 Aug 10;35(31):e288 [FREE Full text] [doi: 10.3346/jkms.2020.35.e288] [Medline: 32776726]
- 48. Baskaradoss JK, Alsumait A, Malik S, Ariga J, Geevarghese A, Francis R, et al. Epidemiological characteristics of COVID-19 cases among Indians residing in Kuwait. East Mediterr Health J 2021 Aug 26;27(8):745-754 [FREE Full text] [doi: 10.26719/emhj.21.006] [Medline: 34486710]
- 49. Bi Q, Wu Y, Mei S, Ye C, Zou X, Zhang Z, et al. Epidemiology and transmission of COVID-19 in 391 cases and 1286 of their close contacts in Shenzhen, China: a retrospective cohort study. Lancet Infect Dis 2020 Aug;20(8):911-919 [FREE Full text] [doi: 10.1016/S1473-3099(20)30287-5] [Medline: 32353347]
- 50. Chaw L, Koh WC, Jamaludin SA, Naing L, Alikhan MF, Wong J. Analysis of SARS-CoV-2 transmission in different settings, Brunei. Emerg Infect Dis 2020 Nov;26(11):2598-2606 [FREE Full text] [doi: 10.3201/eid2611.202263] [Medline: 33035448]
- Gouvea-Reis FA, Oliveira PD, Silva DC, Borja LS, Percio J, Souza FS, et al. COVID-19 outbreak in a large penitentiary complex, April-June 2020, Brazil. Emerg Infect Dis 2021 Mar;27(3):924-927 [FREE Full text] [doi: 10.3201/eid2703.204079] [Medline: 33434475]
- 52. Gupta M, Parameswaran GG, Sra MS, Mohanta R, Patel D, Gupta A, et al. Contact tracing of COVID-19 in Karnataka, India: superspreading and determinants of infectiousness and symptomatic infection. PLoS One 2022 Jul 11;17(7):e0270789 [FREE Full text] [doi: 10.1371/journal.pone.0270789] [Medline: 35816497]
- Kwok KO, Wong VW, Wei WI, Wong SY, Tang JW. Epidemiological characteristics of the first 53 laboratory-confirmed cases of COVID-19 epidemic in Hong Kong, 13 February 2020. Euro Surveill 2020 Apr;25(16):2000155 [FREE Full text] [doi: 10.2807/1560-7917.ES.2020.25.16.2000155] [Medline: 32347198]
- 54. Kwok KO, Wei WI, Huang Y, Kam KM, Chan EY, Riley S, et al. Evolving epidemiological characteristics of COVID-19 in Hong Kong from January to August 2020: retrospective study. J Med Internet Res 2021 Apr 16;23(4):e26645 [FREE Full text] [doi: 10.2196/26645] [Medline: 33750740]
- 55. Lau YC, Tsang TK, Kennedy-Shaffer L, Kahn R, Lau EH, Chen D, et al. Joint estimation of generation time and incubation period for coronavirus disease 2019. J Infect Dis 2021 Nov 22;224(10):1664-1671 [FREE Full text] [doi: 10.1093/infdis/jiab424] [Medline: 34423821]

- Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. N Engl J Med 2020 Mar 26;382(13):1199-1207 [FREE Full text] [doi: 10.1056/NEJMoa2001316] [Medline: 31995857]
- 57. Li M, Liu K, Song Y, Wang M, Wu J. Serial interval and generation interval for imported and local infectors, respectively, estimated using reported contact-tracing data of COVID-19 in China. Front Public Health 2020;8:577431 [FREE Full text] [doi: 10.3389/fpubh.2020.577431] [Medline: 33490015]
- 58. Liu T, Qi L, Yao M, Tian K, Lin M, Jiang H, et al. Serial interval and reproductive number of COVID-19 among 116 infector-infectee pairs Jingzhou City, Hubei Province, China, 2020. China CDC Wkly 2020 Jul 03;2(27):491-495 [FREE Full text] [doi: 10.46234/ccdcw2020.118] [Medline: 34594686]
- 59. Najafi F, Izadi N, Hashemi-Nazari SS, Khosravi-Shadmani F, Nikbakht R, Shakiba E. Serial interval and time-varying reproduction number estimation for COVID-19 in western Iran. New Microbes New Infect 2020 Jul;36:100715 [FREE Full text] [doi: 10.1016/j.nmni.2020.100715] [Medline: 32566233]
- 60. Nishiura H, Linton NM, Akhmetzhanov AR. Serial interval of novel coronavirus (COVID-19) infections. Int J Infect Dis 2020 Apr;93:284-286 [FREE Full text] [doi: 10.1016/j.ijid.2020.02.060] [Medline: 32145466]
- 61. Ping K, Lei M, Gou Y, Tao Z, Yao G, Hu C, et al. Epidemiologic characteristics of COVID-19 in Guizhou Province, China. J Infect Dev Ctries 2021 Mar 31;15(3):389-397 [FREE Full text] [doi: 10.3855/jidc.12818] [Medline: 33839714]
- 62. Quach HL, Nguyen KC, Hoang NA, Pham TQ, Tran DN, Le MT, et al. Association of public health interventions and COVID-19 incidence in Vietnam, January to December 2020. Int J Infect Dis 2021 Oct;110 Suppl 1:S28-S43 [FREE Full text] [doi: 10.1016/j.ijid.2021.07.044] [Medline: 34332082]
- 63. Rajendrakumar AL, Nair AT, Nangia C, Chourasia PK, Chourasia MK, Syed MG, et al. Epidemic landscape and forecasting of SARS-CoV-2 in India. J Epidemiol Glob Health 2021 Mar;11(1):55-59 [FREE Full text] [doi: 10.2991/jegh.k.200823.001] [Medline: 32959618]
- 64. Reed IG, Walker ES, Landguth EL. SARS-CoV-2 serial interval variation, Montana, USA, March 1-July 31, 2020. Emerg Infect Dis 2021 May;27(5):1486-1491 [FREE Full text] [doi: 10.3201/eid2705.204663] [Medline: 33900189]
- 65. Saurabh S, Verma MK, Gautam V, Kumar N, Goel AD, Gupta MK, et al. Transmission dynamics of the COVID-19 epidemic at the district level in India: prospective observational study. JMIR Public Health Surveill 2020 Oct 15;6(4):e22678 [FREE Full text] [doi: 10.2196/22678] [Medline: 33001839]
- 66. Shi P, Gao Y, Shen Y, Chen E, Chen H, Liu J, et al. Characteristics and evaluation of the effectiveness of monitoring and control measures for the first 69 Patients with COVID-19 from 18 January 2020 to 2 March in Wuxi, China. Sustain Cities Soc 2021 Jan;64:102559 [FREE Full text] [doi: 10.1016/j.scs.2020.102559] [Medline: 33101882]
- Song JS, Lee J, Kim M, Jeong HS, Kim MS, Kim SG, et al. Serial intervals and household transmission of SARS-CoV-2 omicron variant, South Korea, 2021. Emerg Infect Dis 2022 Mar;28(3):756-759 [FREE Full text] [doi: 10.3201/eid2803.212607] [Medline: 35107418]
- 68. Wang K, Zhao S, Liao Y, Zhao T, Wang X, Zhang X, et al. Estimating the serial interval of the novel coronavirus disease (COVID-19) based on the public surveillance data in Shenzhen, China, from 19 January to 22 February 2020. Transbound Emerg Dis 2020 Nov 19;67(6):2818-2822 [FREE Full text] [doi: 10.1111/tbed.13647] [Medline: 32452648]
- 69. Won YS, Kim JH, Ahn CY, Lee H. Subcritical transmission in the early stage of COVID-19 in Korea. Int J Environ Res Public Health 2021 Jan 31;18(3):1265 [FREE Full text] [doi: 10.3390/ijerph18031265] [Medline: 33572542]
- 70. Wong J, Chaw L, Koh WC, Alikhan MF, Jamaludin SA, Poh WW, et al. Epidemiological investigation of the first 135 COVID-19 Cases in Brunei: implications for surveillance, control, and travel restrictions. Am J Trop Med Hyg 2020 Oct;103(4):1608-1613 [FREE Full text] [doi: 10.4269/ajtmh.20-0771] [Medline: 32815514]
- 71. Xia W, Liao J, Li C, Li Y, Qian X, Sun X, et al. Transmission of corona virus disease 2019 during the incubation period may lead to a quarantine loophole. MedRxiv 2020 Mar [doi: 10.1101/2020.03.06.20031955]
- 72. Yang L, Dai J, Zhao J, Wang Y, Deng P, Wang J. Estimation of incubation period and serial interval of COVID-19: analysis of 178 cases and 131 transmission chains in Hubei province, China. Epidemiol Infect 2020 Jun 19;148:e117 [FREE Full text] [doi: 10.1017/S0950268820001338] [Medline: 32594928]
- 73. Zhao S, Gao D, Zhuang Z, Chong MK, Cai Y, Ran J, et al. Estimating the serial interval of the novel coronavirus disease (COVID-19): a statistical analysis using the public data in Hong Kong from January 16 to February 15, 2020. Front Phys 2020 Sep 17;8 [FREE Full text] [doi: 10.3389/fphy.2020.00347]
- 74. Kucirka LM, Lauer SA, Laeyendecker O, Boon D, Lessler J. Variation in false-negative rate of reverse transcriptase polymerase chain reaction-based SARS-CoV-2 tests by time since exposure. Ann Intern Med 2020 Aug 18;173(4):262-267 [FREE Full text] [doi: 10.7326/M20-1495] [Medline: 32422057]
- 75. alpersr / EsteR-sensitivity-analysis. GitHub. URL: https://github.com/alpersr/EsteR-sensitivity-analysis [accessed 2023-05-26]
- 76. Jäckle S, Röger E, Dicken V, Geisler B, Schumacher J, Westphal M. A statistical model to assess risk for supporting COVID-19 quarantine decisions. Int J Environ Res Public Health 2021 Aug 31;18(17):9166 [FREE Full text] [doi: 10.3390/ijerph18179166] [Medline: 34501757]
- 77. Padilla R, Netto SL, da Silva EA. A survey on performance metrics for object-detection algorithms. In: Proceedings of the 2020 International Conference on Systems, Signals and Image Processing (IWSSIP). 2020 Presented at: 2020 International

Conference on Systems, Signals and Image Processing (IWSSIP); Jul 1-3, 2020; Niteroi, Brazil [doi: 10.1109/iwssip48289.2020.9145130]

- 78. Kantorovich LV. Mathematical methods of organizing and planning production. Manage Sci 1960 Jul;6(4):366-422 [doi: 10.1287/mnsc.6.4.366]
- 79. Lewis C, Wharton C. Chapter 30 Cognitive walkthroughs. In: Handbook of Human-Computer Interaction. Amsterdam: North-Holland; 1997.
- 80. Krueger R, Casey MA. Focus Groups A Practical Guide for Applied Research, Third edition. Thousand Oaks, California, United states: SAGE Publications; 2000.

Abbreviations

HDR: high-density region IoU: intersection-over-union RKI: Robert Koch Institute UI: user interface

Edited by A Mavragani; submitted 05.12.22; peer-reviewed by L Yang, I Iyamu; comments to author 24.04.23; revised version received 14.05.23; accepted 15.05.23; published 27.06.23 <u>Please cite as:</u> Alpers R, Kühne L, Truong HP, Zeeb H, Westphal M, Jäckle S Evaluation of the EsteR Toolkit for COVID-19 Decision Support: Sensitivity Analysis and Usability Study JMIR Form Res 2023;7:e44549 URL: https://formative.jmir.org/2023/1/e44549 doi: 10.2196/44549 PMID:

©Rieke Alpers, Lisa Kühne, Hong-Phuc Truong, Hajo Zeeb, Max Westphal, Sonja Jäckle. Originally published in JMIR Formative Research (https://formative.jmir.org), 27.06.2023. 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 work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on https://formative.jmir.org, as well as this copyright and license information must be included.

