# Risk Model of German Corona Warning App – Reloaded

**Jens Braband[1], Hendrik Schäbe [2]***
[1] *TU Braunschweig, Braunschweig, Germany*
[2]*TÜV Rheinland, Köln, Germany*
**schaebe@de.tuv.com*

*Jens Braband*

*Hendrik Schäbe*

**Abstract. Aim.** *In this paper we discuss the risk model of the German Corona Warning App in two versions. Both are based on a general semi-quantitative risk approach that is not state of the art anymore and for some application domains even deprecated. The main problem is that parameter estimates are often only ordinal scale or rank numbers for which such operations as multiplication or division are not clearly specified. Therefore, it may results in underestimation or overestimation of the associated risk.* **Methods.** *The risk models that are used in the apps are analyzed. Comparison of the nomenclature of model parameters, their influence on the result, approaches to the generation of a combined risk assessment is carried out. The effectiveness of the models is analyzed.* **Results.** *It is shown that most of the parameters in the model are used only as binary indicator variable. It has been found that the Corona Warning App uses a much more limited model that does not even assess risk, but relies on one parameter which is weighted exposure time. It has been shown that the application underestimates this parameter and therefore may erroneously reassure users. Thus, it may be concluded that the basic risk model implemented before version 1.7.1., is rather a dosimetric model that depends on the calculated virus concentration and does not depend on exposure and other parameters (excluding some threshold values). It is not even a risk model as defined by many standards. Changes of the risk model in the later version are not fundamental. In particular the later model also assesses not individual risk, but individual exposure according to the results. In addition, the model greatly underestimates the duration of exposure. Although it is reported that about 60% of the app's users have shared positive test results, the absolute number of published results is less than 10% of all positive test results. Therefore, from an individual point of view, the application is effective only in 10% of cases, or even less.* **Conclusions.** *As the Corona Warning App also has other systematic limitations and shortcomings it is advised not to rely on its results but rather on Covid testing or vaccination. In addition, if there are enough virus tests available in the near future, the application will even become outdated. It will be better to develop an application that can assess risks a priori, as a kind of decision support for its users based on their individual risk profile.*

**Keywords:** *Corona Warning App, Risk Model, Model Analysis*

## Introduction

In Germany, an app has been introduced to cope with the Covid pandemic: Corona Warning App (CWA). The CWA computes a risk for persons to have been infected caused by contacts. In a former paper we analyzed, how CWA works and also its risk model [12]. Recently, after some criticism, the model has been changed (from version 1.9 and later). We study in this paper the changes of the risk model and its relevance as an indicator for individual risk. While the particular discussion focusses on the German app only, the so-called Exposure Notification Framework (ENF), on which the CWA relies, has been introduced by Apple and Google as part of a standard interface, and so it can be expected that the results may be generalized [14].

We introduce how the CWA works, then briefly summarize the former risk model, describe the changes. Finally, we compare the risk models and assess the effects of the changes.

## How the app works

First a simplified overview is given in order to understand the risk model. More details are given in [1] and [2]. A much broader view on contract tracing apps has been published recently [14].

After a user has installed the CWA, each day a new anonymous ID is created. About every five minutes the environment is scanned for Bluetooth signals emitted from other mobiles. Data like ID, signal attenuation, duration etc. are collected and aggregated for each day.

If the user receives a positive test result and agrees to publish it, then the associated anonymous IDs for the preceding 14 days are transmitted to the central server, from which it is transmitted to all subscribers and checked against the recorded data in the local app. The actual risk evaluation is performed decentralized by each CWA.

## The Basic Risk Model

The basic model up to version 1.7.1 is defined by four parameters [2], which in a first step are evaluated on a semi-quantitative scale each ranging from 0-8 for each day for each ID that reported a positive test result (see figure 1):

• The Days since Exposure (DE) is the time since exposure to the infected person, a value between 0 and 14, durations longer than 14 days are not taken into account.

• The Exposure Duration (ED) is the cumulative time of exposure on the day, takes values between 0 and 8.
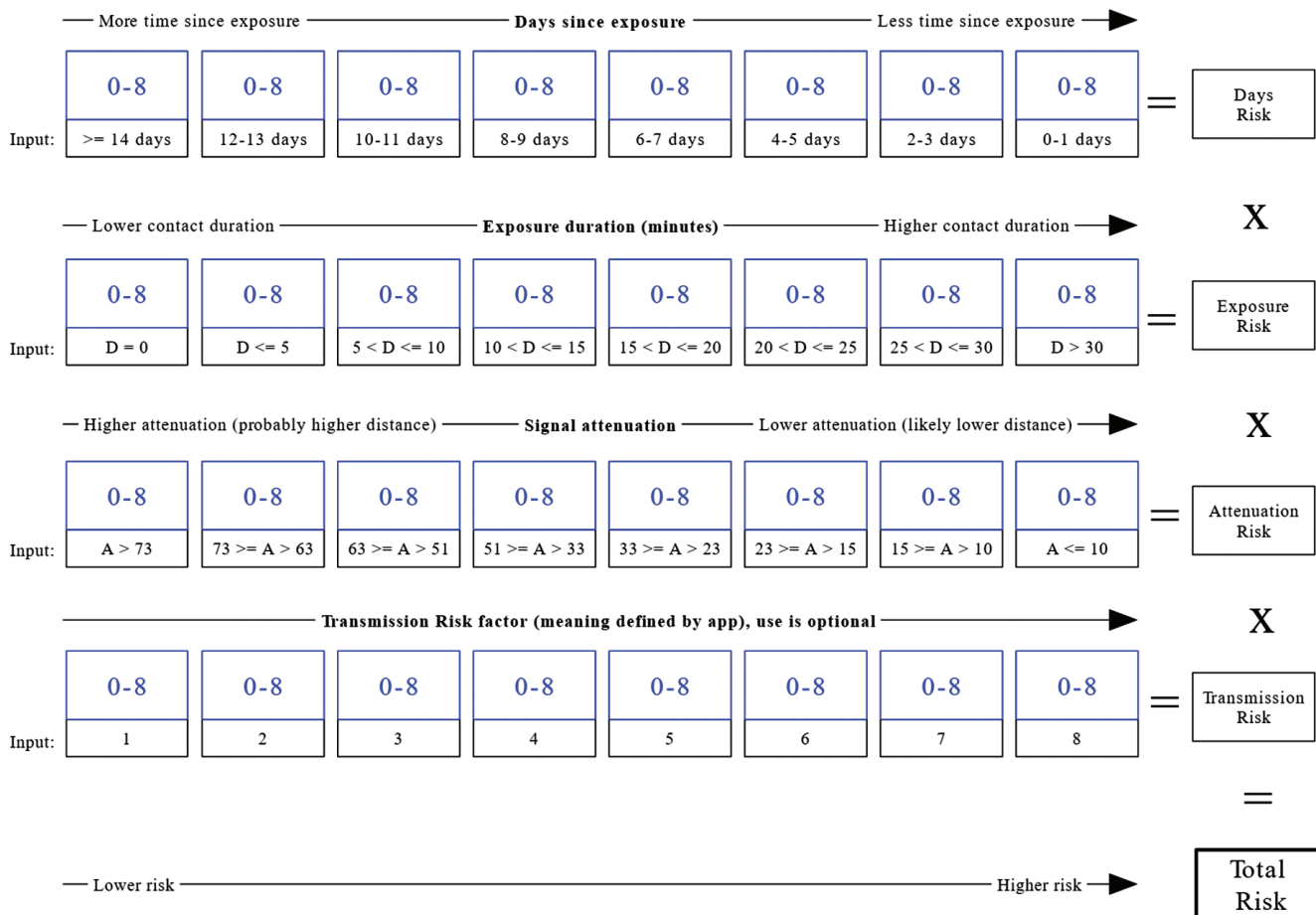


Figure 1: Basic Risk Calculation [2]

• The Bluetooth Signal Attenuation (SA) is used as a measure of the distance to the infected person, takes values between 0 and 8.

• The Transmission Risk (TR) estimates the level of infectiosity of the person on that day, takes values between 0 and 8.

Then the Total Risk Score (TRS) is evaluated by multiplication of DE, ED, SA and TR, theoretically resulting in scores between 0 and 7168. ENF is a standard interface for CWA introduced by Apple and Google, but the parameters can be chosen by each implementation, and so they differ country by country and app version by app version.

ENF resembles the approach known as Risk Priority Numbers (RPN) and suffers from the same limitations and flaws, which are known for about two decades [5] [6]. For some application sectors the use of RPN is even deprecated [7]. Generally, such an approach is not considered state of the art anymore and should be used with great care.

The major problem is that the scores for the parameters are often only ordinal scale or rank numbers, for which operations like multiplication or division are not well-defined. As a consequence, the results may lead to under- or overestimation of the related risk [8].

However, in the practical implementation of the CWA the model is simplified, and the coinciding ranges are limited [3] by

• Days since Exposure (DE) is set to 5 for values below 14 days, and 0 for above, leading to $\delta_{DE} = 5*I(DE \leq 14)$

• Exposure Duration (ED) is set to 0 for all values up to 10 minutes, and 1 for above, yielding $\delta_{ED} = (ED>10)$

• Signal Attenuation (SA) is set to 0 above 73 dB, and 2 for all values below, i. e. $\delta_{SA} = 2* I(SA \leq 73dB)$

• Transmission Risk (TR) is set to (6, 8, 8, 8, 5, 3, 1, 1, 1, 1, 1, 1, 1) [4], depending on DE, e. g. TR is 6 if DE=1, 8 if DE=2 etc, and 0, if DE=14 or above

Here, I(×) denotes the indicator function, which take value 1 if the expression in brackets holds true, and zero otherwise.

So, most of the parameters are only used as binary indicator variables and in the current configuration the Total Risk Score for a particular day and a particular ID is given by

$$TRS = 10 \cdot \delta_{ED} \cdot \delta_{SA} \cdot \delta_{DE} TR. \qquad (1)$$

So, with the implementation [3] until version 1.7.1 there are only six possible scores: 0, 10, 30, 50, 60, 80. But also a Minimum Risk Score (MRS) of 11 is defined and all risks below are discarded. But the parametrization of the CWA may be changed.

So, we can conclude that the basic risk model as implemented until version 1.7.1 is more a dosimetric model, depending on the estimated virus concentration, rather than on exposure and other parameters (but for some threshold values). It is not even a risk model as per the definition of many standards. Moreover, the risk model has been heavily discretized.

Example:

• Alice receives a positive test result on the 20th of the month, which she reports immediately.

• Bob is often taking the same bus as Alice. A ride takes 10 minutes, and he has met her on the 16th (two rides) and the 9th (one ride). They have set together with a distance of about 1m.

• For the 16th DE=4 and so TR=8. Both SA and ED are above the threshold and set to 2 and 1, respectively, so TRS=80.
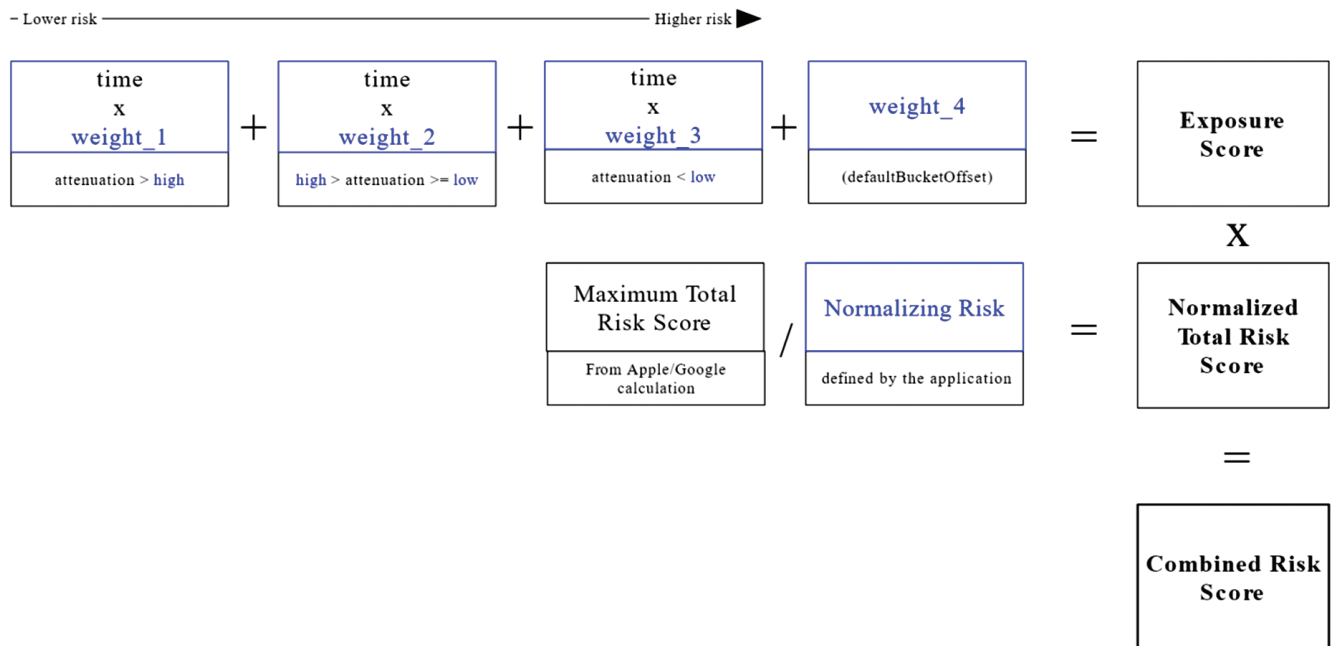


Figure 2: Combined Risk Calculation [2]

• For the 9th DE=11 and so TR=1. However, ED is below the threshold and set to 0. So TRS=0. Otherwise, the TRS would have been 5, which is below the MRS and would have been discarded anyhow.

## Combined Risk Model

In a second step each CWA combines the scores for different encounters calculated by the basic risk model. Let R1, R2, …Rn denote the individual TRS for different days and different IDs that are above the MRS.

In a first step the maximum value Rmax of the different TRS is determined. Then the ED of all the n encounters is summed up in three different classes: close, medium and far. Let these durations be t1, t2 and t3. For each of the classes a weight is defined and additionally a weight offset, which are denoted by w1, w2, w3 and w4, respectively. Note that in practice the weights for the close and medium classes outweigh the others. Also, an Average Risk Score (ARS), currently 50 [3], is defined. Then the Total Combined Risk (TCR) is calculated as (see figure 2)

$$TCR = \left(t_1 w_1 + t_2 w_2 + t_3 w_3 + w_4\right)\frac{R_{max}}{ARS}. \tag{2}$$

Surprisingly the TCR is in fact not a risk, but an exposure as the result is in minutes. The first term (in brackets) is a weighted exposure time which is adjusted by a relative factor (dimensionless) which depends on the maximum virus concentration compared with some average.

So overall, without full mathematical exactness, we can characterize the approach used until version 1.7.1 [3] by the German Corona warning app as

$$TCR = \left(t_1 w_1 + t_2 w_2 + t_3 w_3 + w_4\right)\cdot$$
$$\cdot\frac{max\left\{10\cdot\delta_{ED}\cdot\delta_{DE}\delta_{SA}\cdot TR\right\}}{ARS} \approx \left(t_1 + t_2/2\right)\frac{10\,max\left\{TR\right\}}{ARS}. \tag{3}$$

Basically, this is not a risk in the narrow sense, it is a weighted exposure duration, the units of TCR are not expected damage per time or similar, but just minutes of exposure. And the weight applied is just a measure of relative infectiosity, which expresses the size of TR relative to a normalizing factor of 5, which is assumed as average infectiosity. And the impact of the factor is limited, the highest possible value being 1.6.

And, if we take additionally into account the uncertainty and spread of all the input parameters, e. g. noise in the signal attenuation, uncertainty in the exposure duration and infectiosity [4], or arbitrariness of the weights chosen, then the model boils down to a quite simple formula and decision procedure:

$$TCR \approx \frac{maxTR}{5}\sum ED, \tag{4}$$

• Estimate the minutes that the person was exposed closely to infected persons (∑ED)

• Weight the exposure ED by the infectiosity of the most infected person (max TR/5)
• Take action if the result is HIGH
Example (continued)
• Additionally, Charlie has received a positive test result on the 20th, but he reported it only on the 21st. But he has installed the app only a week ago.
• He has been on the same bus on the same days, but with some larger distance to Bob (2m)
• For the 16th, DE=5 (evaluated on the 21st) and so TR=5. Both SA and ED are above the threshold and set to 2 and 1, so TRS=50.
• For the 9th, Charlie had not installed the app yet, so there are no data.
• As Alice sat close to Bob, and Charlie in medium distance, t1=t2=20 minutes. The weights are currently set [3] to w1=1, w2=0.5, w3=w4=0, and so the weighted sum results in 30 minutes
• So the TCR amounts to 30 x 80 / 50, which gives 48 minutes.
• The warnings of the app are issued based on the TCR, which are configured [3] as LOW for values up to 15, and HIGH from 15 onwards. So finally Bob would get a HIGH risk warning.

Note that the TCR is almost independent of the distance measured by Bluetooth Signal Attenuation. There is only a loose threshold defined and the exposure duration is weighted into two distance classes. However, it is known, see e. g. the FAQ by RKI [9], that transmission is mainly by aerosols, and the risk increases when the exposure distance decreases. Close contact to infected persons is also known to be a factor in superspreading events. The importance of the distance is also supported by the fact that it is a major factor in the German Corona rules AHA, where the first letter stands for Abstand=distance.

## Updated Risk Model

From version 1.9 on the risk model has been changed several times [10], with the main intent of improving the granularity of the risk assessment and also the accumulation of risk i. e. several low-risk encounters in the former model may constitute a high risk in the new model. As a caveat it should be noted that some parameters seem to have been adapted several times and that even on the GitHub repository [11] the parameters in the documentation are not consistent. The information presented here is from end of March 2021 and in doubt the authors have used the parameters directly from the configuration files of the repository rather than from the documentation.

As a main means to achieve this 30-minute evaluation windows are introduced, and a contact is only counted if within this window
1. The signal attenuation (SA) is below the threshold of 73 dB for at least 5 minutes (ED), and

Figure 3: Direct mapping of TRL to scaling factor [11]

2. Transmission Risk (TR) level must be at least 3 (compare figure 1)

If these criteria are not met, then the contact is discarded. Compared to formula (1) this seems not a big change but for the second condition, which was no present before.

Close encounters with SA below 55dB will be counted full, while all other encounters will be discounted by 50%. Note that this does not match with the interface defined by figure 1, but the weight used are the same as before, compare formula (2).

For the TR level a new weighting is introduced, ranging from 0.6 to 1.6, see figure 3. This seems to be a shortcut compared to the calculations in formula (2). Most of the factors are the same as before, e. g. for TR=8, 6, 5, 3, but some have been discarded or slightly changed.

Finally, the measured encounter times (more that 5 minutes within a 30-minute window) are weighted by their proximity and multiplied by the TR level weight (from figure 3) and summed up for a particular day. If the sum is below 15 minutes, but above 5 minutes (some sources suggest 13 minutes [11]) the overall risk is LOW or GREEN, otherwise it is HIGH or RED. This is evaluated per day and the app displays the number of days with low or high risk instead of the numbers of encounters with low or high risk (as up to version 1.7.1).

Taking this all into account, it seems as if the general semi-quantitative framework ENF as defined in figure 1 is not really applied anymore. Instead for each encounter i of a particular day a Transmission Risk Score similar to (1) is evaluated (with some changes to the indicator functions):

$$TRS_i = \delta_{ED} \cdot \delta_{SA} \cdot ED \cdot TR, \qquad (5)$$

where ED is already understood as the weighted encounter times similar to (2)

$$ED = t_1 w_1 + t_2 w_2 + t_3 w_3 + w_4. \qquad (6)$$

Finally, the risk is combined per day by simple summation

$$TCR = \sum_{per\ day} TRS_i. \qquad (7)$$

Example (reloaded):
• For the encounter with Alice on the 16th DE=4 and so originally TR=8, which by figure 3 is then recoded to TR=1.6. Both SA and ED are above the threshold and the weight is 1 and both times are counted full, respectively, so both encounters score TRS=16 minutes, individually. Summing them up a RED warning of HIGH risk is displayed to Bob for this day.
• For the 9th DE=11 and so TR=1, which is recoded to TR=0. So for this risk TRS=0 even as the other thresholds are fulfilled. Otherwise with the lowest TR=0.6 the TRS would have been 6, which would have been at least a GREEN warning with LOW risk for this day.
• Note that the RED warning is independent from the encounter with Charlie on the 16th. For this day Charlies' parameters are the DE=5 with TR=5 recoded to TR=1. As the distance was larger the ED is halved to 5, so both encounters would score TRS=5, summing up to 10 for the day, which standalone would lead to a GREEN warning and combined with Alice's scores would lead to combined score of 26.

If we take into account that some scaling factors have been left out in comparison with the previous approach, e. g. for DE and SA, then the results in the example are quite comparable. The most obvious differences are that
• the term related to max{TR} in formulae (3) and (4) has been simplified and has been substituted by a fixed factor based on TR as in figure 3, and
• that the risks are accumulated and reported per day and not combined for all encounters as in formula (4)

## Discussion

The changes in the risk model are not fundamental, so the general criticism from [12] remains valid. In particular the CWA does not estimate individual risk, but individual exposure, as already the units of the resulting TCR shows, which is [min] and not expected harm or expected severity or similar expressions of risk.

Also, the CWA grossly underestimates the exposure duration for the following reasons:
• There have been about 25 million downloads of CWA so far [13], but it is unknown how many installations are active. If we assume optimistically that 30% of the population would have an active app, then less than 10% of the encounters could be registered, as both parties need to have the app activated. Studies indicate [14] that from 60% coverage of the population by CWA, it could become an effective tool.
• Although it is reported that about 60% of the app users have shared positive test results [13], the absolute number of about 250,000 shared results is again less than 10% of all positive test results (about 2,820,000)

• A large number of short contacts (below 5 minutes), that could be close and infectious, are not detected and registered by the app due to energy saving reasons the app scans only about every 5 minutes its environment

This holds also from an individual user's perspective: a user with active CWA meets only other users with active CWA in less that 30% of the encounters, many short encounters, e. g. in the supermarket or at work, will be discarded and only in 60% of the possible cases other infected users will report their positive results. So, also from the individual perspective the CWA is effective only in about 10% of the cases, or even less.

## The Good, the Bad and the Ugly, or: Summary

The risk model of the German CWA has several interesting, somewhat puzzling and also disturbing properties:

1) It has increased the number of users and is also compatible with apps from some neighboring countries, but its coverage is still rather limited.

2) The main advantage is that the CWA records contacts automatically that would otherwise perhaps not have been noticed by the user. But due to data privacy the user can't learn from the warning as it is not reported when and how long the exposure has been. Therefore the CWA can also not support contact tracing.

3) Also the number of positive tests reported is only a small percentage of all positive tests

4) In the narrow definition of many standards, it is not a complete risk model, as it estimates only single parameters of risk, but not a comprehensive risk, more a kind of exposure measure. A partial explanation can be based on the decentralized architecture of the app and the incomplete and inaccurate information it uses

5) Even worse, it does not cover many situations with short exposure that might still go along with a high dose of virus transmission. This may be due to the pressure by the mobile industry to save energy.

6) Also, it is not an a priori risk assessment in order to decide whether a particular action is acceptable, it assesses risk only a posteriori, when it can't be reduce anymore.

7) And finally, most important, it underestimates the true exposure time or virus dose by a large factor, possibly 5 to 10, so it reassures its users to feel safe if it states no or only a low risk, but in reality, the exposure maybe much higher.

In summary, the authors would recommend not to rely on the results of the CWA due to the many shortcomings of its approach and its risk model, some of which are systematic and can not be improved by further updates. A test or even better, vaccination, should always be preferred. Also, it seems that in the near future, when sufficient Corona tests are available, the CWA will even become obsolete. It is rather recommended to develop an app that may estimate risks a priori as a kind of decision support for its users based on their individual risk profile.

## References

1. RKI: So funktioniert die Corona-Warn-App im Detail. Available at: https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/WarnApp/Funktion_Detail.pdf (accessed at 06.10.2020).

2. Команда CWA: Corona-Warn-App Solution Architecture. Available at: https://github.com/corona-warn-app/cwa-documentation/blob/master/solution_architecture.md (accessed at 06.10.2020).

3. CWA Team: Wie ermittelt die Corona-Warn-App ein erhöhtes Risiko? Available at: https://github.com/corona-warn-app/cwa-documentation/blob/master/translations/cwa-risk-assessment.de.md (accessed at 06.10.2020)

4. CWA Team: Эпидемиологическая мотивация уровня риска передачи, 2020-06-15. Available at: https://github.com/corona-warn-app/cwa-documentation/blob/master/transmission_risk.pdf (accessed at 06.10.2020).

5. Bowles J.: An Assessment of RPN Prioritization in a Failure Modes Effects and Criticality Analysis // Proc. RAMS2003, Tampa, January 2003.

6. Braband J. Improving the Risk Priority Number Concept // Journal of System Safety. 2003. No. 3. P. 21–23.

7. Durivage M. Is It Time To Say Goodbye To FMEA Risk Priority Number (RPN) Scores? // Pharmaceutical Online. Available at: https://www.pharmaceuticalonline.com/doc/is-it-time-to-say-goodbye-to-fmea-risk-priority-number-rpn-scores-00012020-04-27 (accessed at 06.10.2020).

8. Braband J. Beschränktes Risiko // Qualität und Zuverlässigkeit. 2008. No. 53(2). P. 28–33.

9. Robert-Koch-Institut: SARS-Cov-2 Steckbrief. Available at: https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Steckbrief.html (accessed at 06.10.2020).

10. SAP Deutschland: Risikoberechnung der Corona-Warn-App nach detaillierten Tests weiter angepasst, 23.02.2021. Available at: https://www.coronawarn.app/de/blog/2021-02-23-corona-warn-app-risk-calculation-optimization/ (accessed at 06.10.2020).

11. Available at: https://github.com/corona-warn-app (accessed at 06.10.2020).

12. Braband J., Schäbe H. Analysis of the Risk Model of German Corona Warning App, Reliability: Theory & Applications. 2021. No. 1 (61). Vol. 16,. – P. 62-70

13. Robert-Koch-Institut: Kennzahlen der Corona-Warn-App, 17.02.2021. Available at: https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/WarnApp/Archiv_Kennzahlen/Kennzahlen_18022021.pdf?__blob=publicationFile (accessed at 30.03.2020).

14. Landau S. People Count – Contact Tracing Apps and Public Health. The MIT Press, 2021.

## About the authors

**Jens Braband,** Dr. rer. nat., Principal Key Expert for RAMSS at Siemens Mobility GmbH, and Honorary Professor, TU Braunschweig, Germany. Email: j.braband@ tu-braunschweig.de

**Hendrik Schäbe,** Dr. rer. nat. habil., Chief Expert on Reliability, Operational Availability, Maintainability and Safety, TÜV Rheinland InterTraffic, Cologne, Germany, e-mail: schaebe@de.tuv.com

## The authors' contribution

The contribution of the authors consists of an analysis of the risk models of the Corona applications The analysis has been carried out by Jens Braband, Hendrik Schäbe has supported with the analysis of the model and some considerations on the application.

## Conflict of interests

The authors declare the absence of a conflict of interests.