BMJ Open

Analysis of the Dynamics, Outcome, and Prerequisites of the first German SARS-CoV-2 Superspreading Event

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-059809
Article Type:	Original research
Date Submitted by the Author:	07-Dec-2021
Complete List of Authors:	Wessendorf, Lukas; University Hospital Bonn, Virology Richter, Enrico; University Hospital Bonn, Virology Schulte, Bianca; University Hospital Bonn, Institute for Virology Schmithausen, Ricarda Maria; University Hospital Bonn, Hygiene Exner, Martin; University Hospital Bonn Lehmann, Nils; University of Duisburg-Essen, Institute for Medical Informatics, Biometry and Epidemiology Coenen, Martin; Rheinische Friedrich Wilhelms Universitat Bonn, Clinical Study Core Unit, Study Center Bonn (SZB) Fuhrmann, Christine; University Hospital Bonn Kellings, Angelika; University Hospital Bonn Hüsing, Anika; University Hospital Essen Jöckel, Karl-Heinz; Institute of Medical Informatics, Biometry and Epidemiology, University Duisburg- Essen Streeck, Hendrik; German Center for Infection Research
Keywords:	COVID-19, Infection control < INFECTIOUS DISEASES, VIROLOGY

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Analysis of the Dynamics, Outcome, and Prerequisites of the first German SARS-CoV-2 Superspreading Event

Lukas Wessendorf[1], Enrico Richter[1], Bianca Schulte[1], Ricarda Maria Schmithausen[2], Martin Exner[2], Nils Lehmann[3], Martin Coenen[4], Christine Fuhrmann[4], Angelika Kellings[4], Anika Hüsing[3], Karl-Heinz Jöckel[3], and Hendrik Streeck[1,*]

[1] Institute of Virology, University Hospital, University of Bonn, Germany, and German Center for Infection Research (DZIF), partner site Bonn-Cologne

[2] Institute of Hygiene and Public Health, University Hospital, University of Bonn, Germany

[3] Institute of Medical Informatics, Biometry und Epidemiology (IMIBE), University Hospital Essen, Germany

[4] Clinical Study Core Unit, Study Center Bonn (SZB), Institute of Clinical Chemistry and Clinical Pharmacology, University Hospital Bonn, Germany

[*] Corresponding author: Hendrik Streeck, Institute of Virology, University Hospital Bonn, Venusberg-Campus 1, 53127 Bonn, Germany, Hendrik.Streeck@ukbonn.de

Keywords: SARS-CoV-2, COVID-19, superspreading, cohort study, ventilation

Word count: 3903

Abstract

Objectives

Determining the SARS-CoV-2 RT-PCR positivity rate, SARS-CoV-2-specific antibody levels of the participants, and analyzing the conditions and dynamics of superspreading, including ventilation, setting dimensions, distance from infected persons and behavioral patterns.

Design

51 days after the event all participants were asked to give blood samples, pharyngeal swabs and answer a self-administered questionnaire. Metric room coordinates for all tables, seats, and ventilation-points were assessed.

Setting

The superspreading event took place during festivities including 450 people (6-79 years of age) in a building of 27 m x 13.20 m x 4.20 m.

Participants

All persons who took part in the event which led to superspreading of SARS-CoV-2 were invited to participate in this study.

Interventions

No interventions were performed.

Primary and Secondary Outcome Measures

The primary outcome measure was infection status by combining RT-PCR results with ELISA results. Secondary outcomes were symptoms as stated in the questionnaire.

Conclusions

We analyzed infection rates and risk of infection depending on age, alcohol consumption, and ventilation. Overall, 46% of participants had been infected. Spatial distribution of infected participants was associated with proximity to the ventilation system (OR 1.39, 95% KI [0.86; 2.25]). The risk of infection was highly associated with age, thus children (OR: 0.33 [0.267; 0.414]) and young adults had a lower risk than

older participants resulting in an average infection risk increase of 28% per 10 years age difference. Behavioral differences reduced the risk of infection including time spent outside (OR: 0.55 [0.33; 0.91]) or smoking (OR: 0.32 [0.124; 0.81]).

Strengths and Limitations of this Study

- Strength: The setting and the participant group are extremely well-defined.
- Weakness: Some participants left the venue during the event for short times.
- Strength: Participants were invited by only one criteria, namely their presence at the superspreading event; no other preselection/bias took place during enrollment.
- Weakness: The event size was below 1000 people (450), therefore it was not possible to recruit more than 411 study participants.

Article Summary

The scientific literature was searched for the term "superspreading event AND Covid-19 OR Sars Cov 2" and identified published papers from China, South Korea, Europe, and North America. Most researchers analyzed superspreading events within a health care setting e.g. in hospitals or nursing homes, or described the general impact of superspreading events on the global pandemic. Only a few metanalyses of transmission clusters analyzed party occasions (e.g. a nightclub in Berlin, Germany) as superspreading events. These reports describe less than 100 infections and are very limited due to missing data or reporting biases. Therefore, the ability to draw scientific conclusions is also limited. Additionally, to our knowledge, there are no studies, which investigated individual behavior, the location, and role of children during a superspreading event. The research for the study started April 2020 and was concluded in June 2021.

Our report analyzes the first COVID-19 superspreading event in Germany in detail, which was not only a unique setting but also included children and adults in the same room. We demonstrate that nearly half of the participants were infected with SARS-CoV-2 and that the proximity of the seating to the ventilation system was an important risk factor for infection. The data showed that low physical distance including singing and duration of attendance at this event increased the risk of infection, while regular

smoking and spending the break of the event outside lowered the risk of infection. This underlines the benefit of airing to lower the amount of both droplets and aerosols. Furthermore, we found lower infection in children than adults despite being in the same room suggesting differences in infectability in children. Indeed, we observed that an additional 10 years of age is on average associated with 28% increased risk of infection.

Taken together, the results demonstrate the importance of the ventilation system during superspreading events. In particular children and young adults had a lower risk of infection during the event indicating that they have a limited role during this pandemic. Overall, our data demonstrate in detail age-dependent infectability as well as highlights to understand transmission dynamics in order to improve comprehensive public health preparedness measures.

Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is a highly transmissible and pathogenic RNA virus that emerged in late 2019 and has caused a pandemic threatening human health and public safety worldwide.[1] While factors shaping the dynamics of a pandemic are multifactorial, virulence and reproductive number are important properties of a virus.[2] For SARS-CoV-2 there is a substantial over-dispersion of the secondary infection distribution (individual R0) for an individual infected with SARS-CoV-2[2]. An over-dispersed R0 means that most infected people do not transmit (individual R0 = 0) while a minority of infected people are super-spreaders (individual R0 >5). Superspreading has been observed for many infectious pathogens, such as measles or SARS.[3] During the SARS pandemic in 2003 a superspreading event was defined as one infected person infecting eight others.[4] For SARS-CoV-2 it has been estimated that 80% of the infections are caused by 10% of infected individuals highlighting the importance of the cluster factor (k).[2] In Germany an indoor carnival event in the beginning of 2020 is considered as the first major outbreak in a German city and was considered a hotspot during the beginning of the pandemic in Germany.[5] Other SARS-CoV-2 superspreading events worldwide have been linked to indoor gatherings with close proximity of individuals.[6] Nevertheless,

most of the reported superspreading events had less than 100 cases and the reports are limited by missing data or a reporting bias.[6]

Here, we closely examined the prerequisite of a unique super-spreading event in Germany during the SARS-CoV-2 pandemic, where nearly half of the participants became infected including children. We systematically analyzed infection rate, potential individual, and environmental risk factors for infection as well as the role of the ventilation system.

Materials and methods

Study design and sampling

This cross-sectional epidemiological study was conducted 51 days after a carnival celebration in the beginning 2020. Eleven days after the event authorities sent all known participants into quarantine after testing 38 out of 99 individuals PCR-positive. All adults known to have attended the event were invited to participate in the study. About 450 persons attended the event of which 411 participated in the study (**figure 1,** participation rate 91.3%). All study participants provided written informed consent before enrolment.

Self-administered questionnaires included questions about demographic background, symptoms of viral infection as well as detailed information about the behavior during the event. Participants' arrival and exit times were assessed in 1-hour categories. Study participants were asked to provide blood specimens and pharyngeal swabs for further analysis.

Research Ethics Approval

The Ethics Committee of the Medical Faculty of the University of Bonn approved the study (approval number 085/20).

Patient and Public Involvement Statement

The concept and organization of the study on-site was in close cooperation and agreement with the local administration, the county commissioner and the society-leader of the carnival event. Additionally, the community was in need of SARS-CoV-2

testing, because at this time the availability of testing was still limited, therefore they invited the collaboration with our team of scientists and physicians.

Pharyngeal swab and blood preparation

Pharyngeal swabs of participants were performed with FLOQSwabs (Copan) and immediately stored in UTM RT-mini tubes containing UTM Viral Stabilization Media (Copan) at 4 °C. Venous blood was drawn into EDTA tubes (Sarstedt) per volunteer and was transported to the laboratory at the University Hospital Bonn.

Anti-SARS-CoV-2 ELISA

Anti-SARS-CoV-2 IgA and IgG were determined using enzyme-linked immunosorbent assays (ELISA) on the EUROIMMUN Analyzer I platform.[5] According to the manufacturer's instructions a result was considered positive when a ratio (extinction of sample/extinction of calibrator) of 0.8 or higher was reached. The guidelines of the German Medical Association (RiliBÄK) were abided by, including internal and external quality controls.

Reverse transcription polymerase chain reaction (RT-PCR)

Viral RNA was extracted from each 300µl swab sample via the chemagic Viral 300 assay (according to manufacturer's instructions) on the Perkin Elmer chemagic™ Prime™ instrument platform. The presence of two viral target genes (E and RdRP) was assessed in each sample by real time RT-PCR (SuperScript™III One-Step RT-PCR System with Platinum™ TaqDNA Polymerase, Thermo Fisher). The following primers were used, for E gene: E_Sarbeco_F1 and R, and probe E_Sarbeco_P1, for RdRP gene: RdRP_SARSr_F, and R, and probe RdRP_SARSr-P2.[7] In addition, an internal control for RNA extraction, reverse transcription, and amplification was applied to each sample (innuDETECT Internal Control RNA Assay, Analytik Jena #845-ID-0007100). If amplification occurred in both virus-specific reactions samples were considered positive.

SARS-CoV-2 Neutralization Assay

A plaque reduction neutralization test was used to determine SARS-CoV-2 neutralization capacity as previously described.[5] Briefly, plasma samples were heat-

inactivated and supernatant transferred to a new tube and serially two-fold diluted in OptiPROTMSFM (Gibco) performed. 120 mL of each plasma dilution was mixed with 80 plaque-forming units (PFU) of SARS-CoV-2 in 120 mL OptiPRO SFM (GIBCO) cell culture medium and seeded with $1.25x10^5$ Vero E6 cells/well. Subsequently, the inoculum was removed and cells were overlayed with a mixture of carboxymethylcellulose (Sigma) and 2xMEM (Biochrom). Following 3-day incubation, the overlay was removed and the 24 well plates were fixed using a 6% formaldehyde solution and stained with 1% crystal violet in 20% ethanol revealing the formation of plaques. Finally, the neutralizing titers were calculated as the reciprocal of serum dilutions resulting in neutralization of 50% input virus (NT50), read out as reduction in the number of plaques.

Data management and quality control

The Clinical Study Core Unit of the Study Center Bonn (SZB) supported the study by outlining the study protocol and developing the informed consent form as well as participants information sheets with respect to data management and quality control.
The data were gathered on paper-based Case Report Forms (pCRF). Data was entered as double-data-entry into the REDCAp study database programmed and hosted by SZB. Study personnel was trained by experienced members of the SZB. A quality manager was on site to support the study team. Monitoring of trial data and informed consent forms was performed according to the monitoring plan by qualified SZB staff. The ethics committee of the Medical Faculty of the University of Bonn was involved and approved the study (reference no. 085/20)

Spatial information

Metric room coordinates (length and width [m]) for areas, tables, seats and ventilation shafts were assessed via measurements, seating plan and photos from the event. Persons providing multiple positions were considered as spending an equal amount of time on different positions. When exact seating was unclear and information was available on table or greater area localisation (bar, stage), average coordinate values were used.
On the grounds of these coordinates, we calculated pairwise metric distances between all persons and distances to closest inletting and purging airshafts. For all persons their pairwise inverse distances were summarized as mean inverse distance. Inverse metric

distances to persons or airshafts were regarded as representing infectious potential through local proximity, and inverse distances were capped at 2.5 (the inverse of the width of a seat of 0.4 m). Alternatively, we counted all, and all infected persons within adjacent rings of 1.5 m width around each participant as a measure of crowdedness and infectious potential.

Statistical analysis

Associations between positive infection status and exposure variables were analysed via logistic regression models. Exposure variables were included crudely, and adjusted for potential confounding factors age, sex, and duration of attendance as fixed effects. To correct for common household effects a random effects model was used. We present odds ratios with 95% confidence intervals. Because we present data on a single specific event among a limited number of participants, we completely refrain from presenting p-values. All analyses were done with SAS 9.4.

Results

411 out of estimated 450 participants of the event responded to our study invitation, resulting in a response rate of 91.3%. 404 individuals provided plasma samples and 316 pharyngeal swabs (**figure 1**). Genders were represented equally among all 404 participants (48% were male) with a broad range in age ((range 6-79) median age 36 years) and level of education (**table 1**). 297 individuals were residents of the community the event took place in, 103 lived in other parts of the county, and 11 were external visitors.

Overall, 186 out of 404 individuals tested seropositive for IgG- and 161 for IgA-antibodies (**suppl. figure 1**). To confirm seropositivity we performed a plaque reduction neutralization assay (**suppl. figure 2**) demonstrating neutralizing activity against SARS-CoV-2 of their respective antibody responses. Given the low specificity of the IgA assay, IgA seropositivity was not further considered**Error! Bookmark not defined.**. As we tested for seropositivity 51 days after the superspreading event, we additionally performed SARS-CoV-2 RT PCR analysis from pharyngeal swabs to exclude potential recent infections. Indeed, 19 participants tested positive in RT-PCR, and were therefore not considered in the study as there was no likelihood of infection during the superspreading event.

	Not infected N%	Infected N%
Total number	218	186
Female	114 (52%)	89 (48%)
Age		
<18 years	31 (14%)	15 (8%)
18-24 years	30 (14%)	20 (11%)
25-39 years	81 (37%)	43 (23%)
40-64 years	71 (33%)	100 (54%)
65+ years	5 (2%)	8 (4%)
BMI (kg/m^2) (std dev)	24.3 (5.12)	26.2 (5.16)
Participating house hold member (std dev)	2.1 (1.12)	2.4 (1.16)
Highest level of formal education		
None	27 (13%)	13 (7%)
Lower secondary school	27 (13%)	23 (13%)
Secondary school	55 (26%)	71 (39%)
Higher education entrance qualification	54 (25%)	34 (18%)
(Technical) university degree	52 (24%)	43 (23%)
Duration of attendance [h] (std dev)	4.7 (2.06)	5.8 (1.85)
Service team	4 (2%)	22 (12%)
On stage during the event	80 (37%)	62 (34%)
Member of the 'Council of 11'	6 (3%)	18 (10%)
On stage during 'finale'	26 (12%)	48 (26%)
Behavior during break		
Remaining seated	73 (36%)	85 (48%)
Going outside	114 (55%)	72 (39%)
Alcohol consumption [drink] (std dev)	11.3 (7.76%)	12.2 (7.40)
Former smoker	34 (16%)	45 (24%)
Active smoker (>10 cigarettes per day)	54 (25%)	23 (12%)
At least one comorbidity	29 (13%)	28 (15%)
Avg. distance to other participants [m] (std dev)	9.2 (1.68)	9.1 (1.70)
Distance to air inlet [m] (std dev)	6.1 (3.22)	6.0 (3.30)
Distance to air outlet [m] (std dev)	4.8 (2.94)	5.1 (2.87)

Table 1: Distribution of demographic factors and exposure information of interest among study participants who tested positive or negative in serology test of SARS-Cov2-infection.

Overall, we found that 46.0% (95% CI: [41.2%; 51.0%]) tested seropositive who attended the event, which was significantly higher than the overall estimated infection rate in the same community at large at that time. Indeed, officially 3.1% of the community were reported as positive cases at that time and we estimated the infection rate as 15.5% (95% CI:[12.3%; 19.00%])**Error! Bookmark not defined.** for the community. Taken together, an estimated 46% of participants became infected during a single superspreading event.

No association between sex and risk of infection was found ((OR: 1.01 [0.65; 1.58]) for women). On average infected individuals had a higher body mass index (26·2kg/m² compared to 24.3kg/m² for uninfected individuals). Infected participants were more likely to be clustered living in the same household (**table 1**). Having at least one comorbidity, including lung disease (42.3%), cardiovascular disease (53.3%), neurological disease (16.7%), cancer (58.3%) or diabetes (80%), did not increase the risk of infection (OR: 0.64 [0.33; 1.26]). In conclusion, sex and comorbidities did not seem to affect the risk of infection. We next assessed whether age influenced the risk of infection at the event, considering sex, duration of attendance and common household as covariates. Comparison across age-categories showed a lower risk for children (OR: 0.31 [0.14; 0.69]), and also for young adults (18-25 years, OR: 0.53 [0.26; 1.09]) as well as adults between 25 and 40 years (OR: 0.48 [0.28; 0.85]) in comparison to older adults (40 to 65 years) (OR: 1, reference), while seniors had a slightly higher risk (older than 65 years, OR: 1.1 [0.31; 3.97]) (**figure 3**). Our data suggest that an additional 10 years of age are on average associated with 28% increased risk of infection (OR: 1.28 [1.10; 1.48]).

To understand the spreading dynamics of SARS-CoV-2 during the event, we first performed a detailed analysis of potential risk factors and social behavior. The event consisted of speeches, dance, and music performances for a total of five hours, with one large intermission and was hosted at a small community center (320 square meters) with a stage up front and a bar in the back close to the entrance. Alcoholic and nonalcoholic drinks were served in glasses and a food truck was located outside in front of the venue. While most participants were sitting in the hall, a committee of eleven individuals hosting the event were sitting on stage. The eleven people on stage switched after a break. With approximately 450 participants there were about 1.4 individuals per square meter and the tables, each with two benches, were arranged in

two blocks with an alley to the stage (**figure 2**). Infected participants had been seated mostly at tables close to the bar, at the bar, or on stage. One table with 8 out of 11 infected people, was located far away from the bar at the other side of the hall and close to an air inlet. The group sitting on stage showed high numbers of infection (18 infected out of 24, **table 1**). We first analyzed whether the ventilation system influenced the distribution of SARS-CoV-2 infected individuals. It is important to state that the system´s air flow consisted of 75% used and 25% fresh air. The air flow can be described as clockwise. The air system uses vents along one side of the venue and on stage to take in air (**figure 2,** air inlets purple). After 25% of fresh air has been added and the air has been filtered, vents along the other side of the venue return the air into the room (**figure 2,** air outlets blue**)**. All ventilation points received the same amount of air due to throttle valves. For noise protection reasons windows remained closed. The air-system used F7-Filters (ISO ePM ≥ 2,5) and had an air volume flow of 7500 m^3/h.

Most tables located close to the air-inlets and showed no or only few infections (**figure 2,** green) also most surrounding tables showed low numbers of infection (**figure 2,** yellow). Tables close to the air-outlets (**figure 2)** show high (4 or 5 infected per table) and very high (6 or 7 infected per table) numbers of infected individuals. It is important to mention that the overall number of participants per table was not equal for all tables. Greater proximity to air outlets was associated with increased risk of infection with a crude OR=1.39 [0.86; 2.25]. This association remained stable and was hardly attenuated from adjustment for proximity to air inlet, age, gender, duration of attendance, proximity to other infected persons, stage-activity and going outside during the intermission (**figure 4,** multiple adjusted OR=1.26 [0.63; 2.50]). A similar apparent effect for proximity to air inlets (crude OR=1.17 [0.72; 1.89]) disappeared when duration of attendance was added to the model (**figure 4,** multiple adjusted OR=1.01 [0.53; 1.94]). Overall, however, we found the increased risk for individuals located closer to the air outlet remarkably persistent (**figure 4**).

We further studied the sum of the inverse distance to all infected participants as a measure of proximity to either one common virus source or mutual infection. However, there was no evidence for increased risk of infection from greater proximity to other infected persons (**suppl. table 1**). Furthermore, we found no evidence for a single

person being the source of the infection using 401 quantile-plot analysis conducted for each participant as potential source of infection separately (**suppl. figure 3**).

To understand the association of risk with behavior patterns we next investigated the influence of several factors on SARS-CoV-2 infection including time spent outside, smoking, performing on stage and participation during the final act ("Finale") for 30 minutes. Results were all adjusted for age, sex, common household, and duration of attendance. Participation in multiple performances did not increase the risk of infection (OR per performance: 1.08 [0.91; 1.27]) while participation in the last "Finale" indicated a trend towards increased risk of infection (OR: 1.41 [0.65; 3.02]) (**figure 4**). Duration of attendance was persistently and strongly associated with an increased infection risk of 32% with each additional hour spent at the party (OR per hour: 1.32 [1.16; 1.49]). All further analyses were adjusted for this variable as potential confounding factor.

We next determined the level of alcohol consumption as number of drinks (high-proof liquor or beer) and did not observe any influence for the amount of alcohol consumption on the risk of becoming infected (OR per drink: 1.00 [0.96; 1.05]). Furthermore, participants who spent the break outside were less likely to be infected (OR: 0.55 [0.33; 0.91]) compared to individuals who spent the break inside the venue hall (**figure 4**). Interestingly, however, when we determined the impact of being regular smoker (defined as smoking of at least 10 cigarettes a day) on the risk of SARS-CoV-2 infection we observed a reduced risk of infection (OR: 0.32 [0.12; 0.81]) even after adjustment for "time spent outside". In conclusion, duration of attendance at the carnival party increased the risk of infection, the number of alcoholic drinks was not associated with infection risk, while regular smoking and spending the break of the event outside lowers the risk of infection.

We next stratified seropositive individuals by their reported symptoms. Odds-ratios for each symptom were calculated for the timespan of 14 days following the event (**figure 5**). Similar to previous reports[8] loss of smell (OR: 8.78 [4.81; 16.02]) and taste (OR: 10.09 [5.13; 19.88]) were strongest associated with SARS-CoV-2 infection. Other symptoms which were strongly associated with COVID-19 were: sweats and chills (OR: 5.28 [3.08; 9.07]), muscle and joint ache (OR: 5.19 [3.19; 8.44]), fatigue (OR: 4.22 [2.76; 6.45]) and fever (OR: 3.73 [2.10; 6.63]) (**figure 5**). Importantly, 15.1% of

the infected individuals reported no symptoms at all in a period of 14 days after the event. The rate of asymptomatic infections of participants of the event was lower than generally observed in the community the event took place in (36%).[5] Overall, there was a lower proportion of asymptomatic cases among individuals infected after the event compared to members of the community, while loss of smell and taste showed the strongest association with an infection.

Discussion

The high overdispersion characteristics of SARS-CoV2 and its ability to be transmitted via aerosols under certain conditions are one of the main reasons that the beginning of the SARS-CoV2 pandemic was shaped by superspreading events.[9,10] Germany`s first superspreading event was an indoor carnival event in the beginning of 2020 in a rural the community. In this naturally occurring experiment, we demonstrate that nearly half of the participants became infected and demonstrate multiple prerequisites of such an event and risk factors for becoming infected. While our study population is not a representative sample of the general population the event may be regarded as exemplary for similar party occasions and may help reduce the number of infected in the future.

An important factor associated with infection risk was the ventilation system and the individual proximity to the ventilation outlets. Individuals close to the air-outlets that contained air with low amount of fresh air had the highest infection risk compared to those close to the air-inlets. This is in line with previous studies that demonstrated SARS-CoV-2 to be able to become air-borne under certain conditions and that the ventilation system can have an influence on virus spread.[11,12,13] The air filters in the venue were not capable of intercepting virus particles supporting the notion on the importance of proper indoor ventilation systems.[14,15] Indeed, spending the break of the event outside decreased the possibility of infection underscoring the benefit of proper ventilation to lower the amount of aerosols. Due to the nature of the event, the spatial distribution of the participants was not fixed throughout the evening, and not perfectly recapitulated, so this information carries some error. However, allowing for multiple positions per person we used all available information. Assuming further error in the spatial data to be random, this might lead to a dilution of effects, i.e. true associations

may remain undetected. Complementary analyses including e.g. the persons' functions during the event show consistent results, so we see no evidence suggesting bias in our findings.

The consumption of alcoholic drinks did not increase the risk of infection. While it has been assumed that the alcoholic effect of decreased social inhibition may increase likelihood of infection, we did not find any evidence for this association questioning measures of a ban on alcohol to reduce numbers of infected. It is known that current and former smokers disproportionately suffer from severe COVID-19 and their numbers are relatively increased among those patients that need intensive care treatment compared to non-smokers.[16,17] However, it has been previously speculated that the risk of infection is lower for smokers.[18] Furthermore, a meta-analysis of seven studies suggests that smokers have a reduced risk of testing positive for SARS-CoV-2.[19] Interestingly, we also observed a protective effect for an infection with SARS-CoV-2, thus our findings support those statements and show an even greater protective effect. The association might for example be explained by a role of the nicotinic acetylcholine receptor.[20] While we strongly advise that smoking should not be considered as a protective habit to prevent risk of infection, this knowledge may lead to the investigation of a therapeutic or prophylactic treatment on the basis of this molecular target.[21]

Our results indicate a trend that younger people are less likely to be infected compared to older age groups. This trend is strongest for people under 18 but levels out over 40 years of age. The risk of infection for children in superspreading events has not been investigated but the overall risk for infection in children seems to be lower than for adults as a systematic review and its recent update reported, which is further supported by our findings.[22,23] As all individuals were exposed at the same event and time our study is a perfect model for the previously described notion, that children are less likely to become infected. Indeed, a recently published meta-analysis by Viner et al. showed a low susceptibility for children and adolescents (OR of 0.56 (95% CI, 0.37-0.85)) which strongly supports our findings of a lower risk of infection in that age group, which is even lower in our study.[24] Our finding supports the previously shown subordinate influence on the spreading of the virus by children. The finding that each 10 years of age increase the risk of infection during an event indicates that younger people and

their limited role should be considered when measures to contain the pandemic are implemented. Taken together, we could demonstrate important risk factors for infection during a superspreading event, which helps to understand transmission dynamics in order to improve comprehensive public health preparedness measures.

Acknowledgments

We would like to thank the inhabitants of the town in question for their participation. We would also like to thank the local government of the town where the outbreak took place for their support to conduct the study. Furthermore, we would like to thank the following people who helped with the study: Rudi Kirschen, Gero Wilbring, Janett Wieseler, Marek Korencak, Ryan Nattrass, Jernej Pusnik, Maximilian Becker, Ann-Sophie Boucher, Marc Alexander de Boer, Rebekka Dix, Sara Dohmen, Kim Friele, Benedikt Gansen, Jannik Geier, Marie Gronemeyer, Sarah Hundertmark, Nora Jansen, Michael Jost, Louisa Khorsandian, Simon Krzycki, Ekaterina Kuskova, Judith Langen, Silvia Letmathe, Ann-Kathrin Lippe, Jonathan Meinke, Freya Merker, Annika Modemann, Janine Petras, Sophie Marie Porath, Anna Quast, Laurine Reese, Isabel Maria Rehbach, Jonas Richter, Thea Rödig, Eva Schmitz, Tobias Schremmer, Louisa Sommer, Jennifer Speda, Yuhe Tang, Oliver Thanscheidt, Franz Thiele, Johanna Thiele, Julia Tholen, Moritz Transier, Maike van der Hoek, Tillmann Verbeek, Sophia Verspohl, Kira Vordermark, Julian Wirtz, Marina Wirtz, Lisa Zimmer, Philip Koenemann, Adi Yaser, Lisa Anna, Katharina Bartenschlager, Lisa Baum, Roxana Böhmer Romero, Diana de Braganca, Isabelle Engels, Moritz Färber, Carina Fernandez Gonzalez, Lucia Maria Goßner, Victoria Handschuch, Franziska Georgia Liermann, Steffan Meißner, Laura Racenski, Patrick, Denis Raguse, Larissa Reiß, Maximilian Rölle, Franziska Scheele, Chiara Schwippert, Arlene Christin Schwippert, Antonia Seifert, Joshua David Stockhausen, Sofia Waldorf, Leonie Weinhold, Nicolai Trimpop, Julia Reinhardt, Vera Gast, Michelle Yong, Eva Engels, Jonathan Meinke, Susanne Schmidt, Janine Schulte, Saskia Schmitz, Kübra Bayrak, Regina Frizler, Katarzyna Andryka, Sofía Soler, Bastian Putschli, Thomas Zillinger, Marcel Renn, Patrick Müller, Dillon Corvino, Zeinab Abdullah, Katrin Paeschke, Hiroki Kato, Florian Schmidt, Maximilian Baum, Celina Schüter, Daniel Hinze, Martina Schmidt, Arcangelo Ricchiuto, Sonja Gross, Uta Wolber, Tobias Höller, Marion Zerlett, Esther Sib, Benjamin Marx, Souhaib Aldabbagh. We thank the Medizinischer Dienst der

Krankenkasse (MDK) Nordrhein for their generous help in conducting this study. In particular we would like to thank: Tanja Bell, Ina de Hesselle-Taddey, Susanne Goll, Katja Hoffmann-Pruss, Niko Kalamakidis, Verena Kayser, Hildegard Kessler, Norbert Körfer, Daniela Kroll, Florian Messerschmidt, Svenja Peters, Sebastian Schröder, Linda Schroers, Michael Zimmermann, Klaus-Peter Thiele. We thank Stefan Holdenrieder, Alexander Semaan, Bernd Pötzsch, and Georg Nickenig for providing control samples. The idea, the plan, the concept, protocol, the conduct, the data analysis, and the writing of the manuscript of this study were independent of any third parties, including the government of North Rhine-Westphalia, Germany.

Author Contributions

L.W. wrote the manuscript. L.W., RMS, and ER organized and ran the testing center. ER and BS organized and performed sample processing, analysis, and corrected the manuscript. NL and AH performed statistical analysis. MC, CF, and AK monitored the study. ME, KHJ and H.S oversaw the study and corrected the manuscript.

Data statement

All data from the study will be made available upon reasonable request.

Funding

The government of North Rhine-Westphalia (Germany) supported the study with 65,000 Euro.

Conflicts of interest

The authors declare no conflict of interest.

References

1. Chakraborty I, Maity P. COVID-19 outbreak: Migration, effects on society, global environment and prevention. *Sci Total Environ* 2020; **728**: 138882.

2. Lloyd-Smith JO, Schreiber SJ, Kopp PE, Getz WM. Superspreading and the effect of individual variation on disease emergence. *Nature* 2005; **438**(7066): 355-9.

3. De Serres G, Markowski F, Toth E, et al. Largest measles epidemic in North America in a decade--Quebec, Canada, 2011: contribution of susceptibility, serendipity, and superspreading events. *J Infect Dis* 2013; **207**(6): 990-8.

4. Shen Z, Ning F, Zhou W, et al. Superspreading SARS events, Beijing, 2003. *Emerg Infect Dis* 2004; **10**(2): 256-60.

5. Streeck H, Schulte B, Kümmerer BM, et al. Infection fatality rate of SARS-CoV2 in a super-spreading event in Germany. *Nat Commun* 2020; **11**(1): 5829.

6. Leclerc QJ, Fuller NM, Knight LE, Funk S, Knight GM. What settings have been linked to SARS-CoV-2 transmission clusters? *Wellcome Open Res* 2020; **5**: 83.

7. Corman VM, Landt O, Kaiser M, et al. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Euro Surveill* 2020; **25**(3).

8. Schmithausen RM, Döhla M, Schöβler H, et al. Characteristic Temporary Loss of Taste and Olfactory Senses in SARS-CoV-2-positive-Individuals with Mild Symptoms. *Pathog Immun* 2020; **5**(1): 117-20.

9. Kyriakopoulos AM, Papaefthymiou A, Georgilas N, Doulberis M, Kountouras J. The Potential Role of Super Spread Events in SARS-COV-2 Pandemic; a Narrative Review. *Arch Acad Emerg Med* 2020; **8**(1): e74.

10. Patel KP, Vunnam SR, Patel PA, et al. Transmission of SARS-CoV-2: an update of current literature. *Eur J Clin Microbiol Infect Dis* 2020; **39**(11): 2005-11.

11. Lednicky JA, Lauzardo M, Fan ZH, et al. Viable SARS-CoV-2 in the air of a hospital room with COVID-19 patients. *Int J Infect Dis* 2020; **100**: 476-82.

12. Santarpia JL, Rivera DN, Herrera VL, et al. Aerosol and surface contamination of SARS-CoV-2 observed in quarantine and isolation care. *Sci Rep* 2020; **10**(1): 12732.

13. Günther T, Czech-Sioli M, Indenbirken D, et al. SARS-CoV-2 outbreak investigation in a German meat processing plant. *EMBO Mol Med* 2020; **12**(12): e13296.

14. Pokora R, Kutschbach S, Weigl M, et al. Investigation of superspreading COVID-19 outbreak events in meat and poultry processing plants in Germany: A cross-sectional study. *PLoS One* 2021; **16**(6): e0242456.

15. Nazarenko Y. Air filtration and SARS-CoV-2. *Epidemiol Health* 2020; **42**: e2020049.

16. Vardavas CI, Nikitara K. COVID-19 and smoking: A systematic review of the evidence. *Tob Induc Dis* 2020; **18**: 20.

17. Engin AB, Engin ED, Engin A. Two important controversial risk factors in SARS-CoV-2 infection: Obesity and smoking. *Environ Toxicol Pharmacol* 2020; **78**: 103411.

18. Israel A, Feldhamer E, Lahad A, Levin-Zamir D, Lavie G. Smoking and the risk of COVID-19 in a large observational population study. *medRxiv* 2020: 2020.06.01.20118877.

19. Grundy EJ, Suddek T, Filippidis FT, Majeed A, Coronini-Cronberg S. Smoking, SARS-CoV-2 and COVID-19: A review of reviews considering implications for public health policy and practice. *Tob Induc Dis* 2020; **18**: 58.

20. Changeux JP, Amoura Z, Rey FA, Miyara M. A nicotinic hypothesis for Covid-19 with preventive and therapeutic implications. *C R Biol* 2020; **343**(1): 33-9.

21. Paleiron N, Mayet A, Marbac V, et al. Impact of Tobacco Smoking on the Risk of COVID-19: A Large Scale Retrospective Cohort Study. *Nicotine Tob Res* 2021; **23**(8): 1398-404.

22. Li X, Xu W, Dozier M, He Y, Kirolos A, Theodoratou E. The role of children in transmission of SARS-CoV-2: A rapid review. *J Glob Health* 2020; **10**(1): 011101.

23. Li X, Xu W, Dozier M, et al. The role of children in the transmission of SARS-CoV2: updated rapid review. *J Glob Health* 2020; **10**(2): 021101.

24. Viner RM, Mytton OT, Bonell C, et al. Susceptibility to SARS-CoV-2 Infection Among Children and Adolescents Compared With Adults: A Systematic Review and Meta-analysis. *JAMA Pediatr* 2021; **175**(2): 143-56.

Figure Legends

Fig. 1: Study participants. Of the 400 people contacted originally (left) 362 adults and 49 children agreed to enroll in the study. An overview of the number of samples collected is given on the right. Downstream sample processing included centrifugation of blood samples for plasma collection (SARS-CoV-2 ELISAs), and viral RNA extraction from swab samples (SARS-CoV-2 RT-PCR).

Fig. 2: Reconstructed 3D-Model of the venue hall. Self-administered questionnaires included questions about main seating-position of the participants during the evening event as specifying table and seat with the help of a schematic seating plan. Metric room coordinates for all tables, seats, and ventilation-points were assessed and the seating was reconstructed from pictures taken during the event. Therefore, the location of the stage, the bar, the exit as well as the tables and the air-inlets/outlets were reconstructed in a 3D-Model. The original external dimensions of the building were 27m x 13.20m x 4.20m. Tables, where more than 7 infected individuals have stayed are colored in dark red, this includes the stage and bar as well. Air-inlets are colored in violet and the air-outlets in blue. Infected participants had been seated mostly at tables close to the bar, the bar itself and on stage. One table with 8 out of 11 infected people, was located far away from the bar at the other side of the hall and close to an air inlet. The group sitting on stage showed as well high numbers of infection (18 infected out of 24). Greater proximity to air outlets seems to be associated with increased risk of infection with a crude OR=1.39 [0.86; 2.25].

Fig. 3 Odds-Ratio for the likelihood of SARS-CoV-2 infection by age groups. Participants were divided into age groups of 8, 15, or 25 years, participants younger than 18 or older than 65 years. Participants were considered to have been infected during the event if they were SARS-CoV-2 antibody positive (ELISA).

Fig. 4: Odds ratios for the association of SARS-CoV-2 infection with specific activities of the participants and their location in the venue relative to ventilation shafts. The model was additionally adjusted for age, sex, duration of attendance, participation in multiple activities, and cumulative proximity to other infected persons, and common household.

Fig. 5: Odds ratios for symptoms of SARS-CoV-2 antibody-positive participants in the 14 days following the super spreading event. The information on symptoms was derived from the self-administered questionnaire, which was filled out on the day of sample collection. Odds ratio estimates (OR) are shown with confidence intervals

Suppl. Fig. 1: Correlation of SARS-CoV-2 Euroimmun ELISA results for IgA and IgG. The correlation of IgA levels to IgG levels in the same person was significant (r: Pearson coefficient, p<0.0001, 95 % CI, 0.7043 to 0.7902). The dotted lines mark the ratios above which each ELISA result is considered positive.

Suppl. Fig. 2: Correlation of plasma neutralization capacity and IgG ELISA results (Euroimmun) from each donor. The dotted line marks the ratio above which the ELISA result is considered positive. The correlation coefficient (Pearson) was

0.3667 (95 % CI, 0.2275 to 0.4192, p<0.0001). Samples with a negative result in the neutralization assay were set as 0.1 here so as to appear on the logarithmic axis.

Supplemental Figure 3: Quantile plot of observed p-values from analyses of inverse distance [1/m] to single specific study participants as risk factor for corona-virus infection. In case of no association, the ordered log-transformed p-values are expected to lie on, or below the diagonal. Panel A: results from crude analyses, Panel B: analyses were adjusted for age, sex, common household and duration of attendance.

Supplementary table 1: Estimated relative risk of SARS-CoV-2 infection (IGG-positive) from logistic regression on summary measures of spatial proximity between participants in terms of odds ratio estimates (OR) with confidence interval and p-values. a) adjusted for sex, age, common household and duration. b) multivariate analysis, mutually adjusted for distance to ventilation system, participation in (multiple) performances, going out of doors during the intermission, and participating in the grand finale. c) multivariate analysis, mutually adjusted for distance to ventilation system, participation in (multiple) performances, going out of doors during the intermission, and participating in the grand finale and adjusted for sex, age, common household and duration.

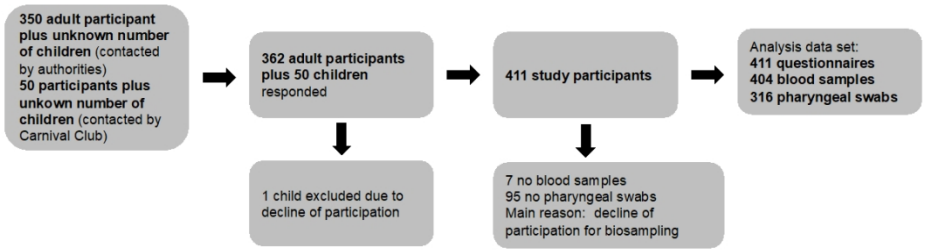


Fig. 1: Study participants. Of the 400 people contacted originally (left) 362 adults and 49 children agreed to enroll in the study. An overview of the number of samples collected is given on the right. Downstream sample processing included centrifugation of blood samples for plasma collection (SARS-CoV-2 ELISAs), and viral RNA extraction from swab samples (SARS-CoV-2 RT-PCR).

342x107mm (96 x 96 DPI)

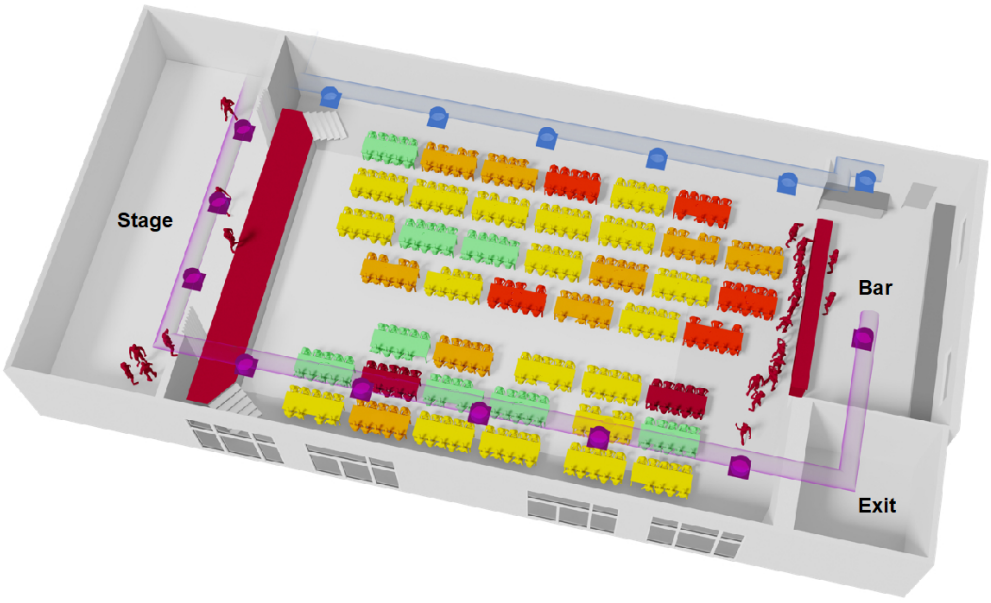


Fig. 2: Reconstructed 3D-Model of the venue hall. Self-administered questionnaires included questions about main seating-position of the participants during the evening event as specifying table and seat with the help of a schematic seating plan. Metric room coordinates for all tables, seats, and ventilation-points were assessed and the seating was reconstructed from pictures taken during the event. Therefore, the location of the stage, the bar, the exit as well as the tables and the air-inlets/outlets were reconstructed in a 3D-Model. The original external dimensions of the building were 27m x 13.20m x 4.20m. Tables, where more than 7 infected individuals have stayed are colored in dark red, this includes the stage and bar as well. Air-inlets are colored in violet and the air-outlets in blue. Infected participants had been seated mostly at tables close to the bar, the bar itself and on stage. One table with 8 out of 11 infected people, was located far away from the bar at the other side of the hall and close to an air inlet. The group sitting on stage showed as well high numbers of infection (18 infected out of 24). Greater proximity to air outlets seems to be associated with increased risk of infection with a crude OR=1.39 [0.86; 2.25].

306x185mm (96 x 96 DPI)

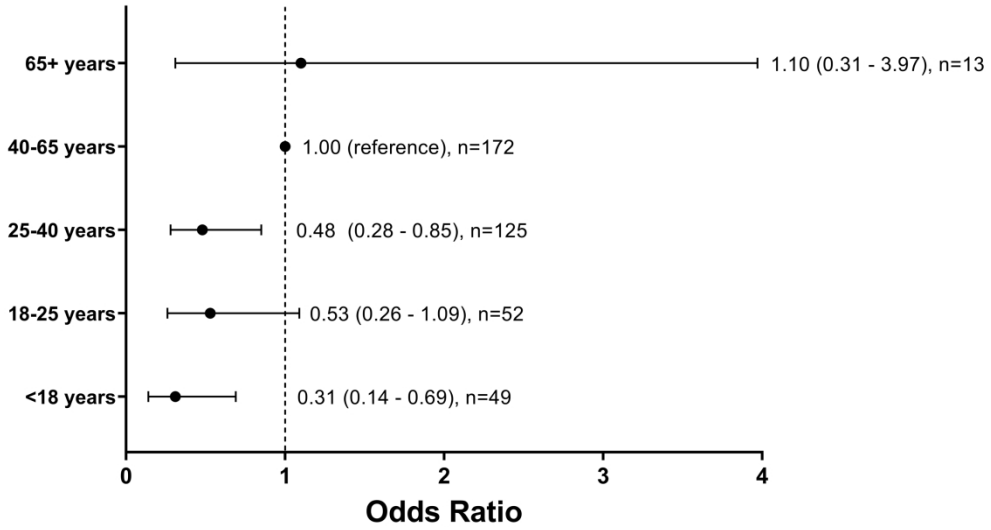


Fig. 3 Odds-Ratio for the likelihood of SARS-CoV-2 infection by age groups. Participants were divided into age groups of 8, 15, or 25 years, participants younger than 18 or older than 65 years. Participants were considered to have been infected during the event if they were SARS-CoV-2 antibody positive (ELISA).

283x160mm (300 x 300 DPI)

Fig. 4: Odds ratios for the association of SARS-CoV-2 infection with specific activities of the participants and their location in the venue relative to ventilation shafts. The model was additionally adjusted for age, sex, duration of attendance, participation in multiple activities, and cumulative proximity to other infected persons, and common household.

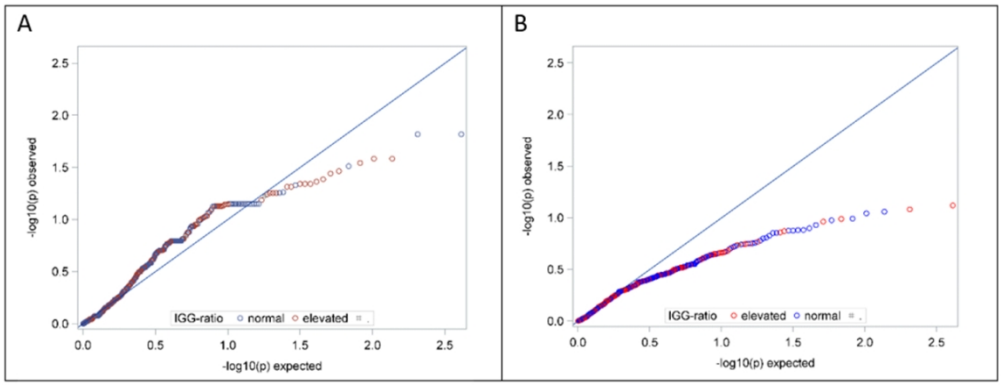
272x124mm (300 x 300 DPI)

Fig. 5: Odds ratios for symptoms of SARS-CoV-2 antibody-positive participants in the 14 days following the super spreading event. The information on symptoms was derived from the self-administered questionnaire, which was filled out on the day of sample collection. Odds ratio estimates (OR) are shown with confidence intervals.

281x217mm (300 x 300 DPI)

277x218mm (300 x 300 DPI)

272x201mm (300 x 300 DPI)



337x129mm (96 x 96 DPI)

	OR	95 % confidence interval		p-value
Proximity of infected persons [sum 1/m]	0.99	0.98	1.01	0.43
Adjusted a)	1	0.98	1.02	0.96
Mutually adjusted b)	0.99	0.97	1.01	0.57
Mutually adjusted c)	0.99	0.97	1.02	0.65
Alternative consideration in distance-bands				
Infected persons within ≤1.5 m [count]	1.01	0.96	1.07	0.68
Infected persons in 1.5 -≤3 m [count]	0.96	0.92	1	0.04
Infected persons in 3 -≤4.5 m [count]	1.03	1	1.06	0.02
Infected persons within ≤1.5 m [count] adjusted a)	1.03	0.97	1.1	0.37
Infected persons in 1.5 -≤3 m [count]	0.96	0.92	1.01	0.11
Infected persons in 3 -≤4.5 m [count]	1.03	1	1.06	0.08
Infected persons within ≤1.5 m [count] mutually adjusted b)	1.01	0.95	1.07	0.73
Infected persons in 1.5 -≤3 m [count]	0.98	0.94	1.02	0.36
Infected persons in 3 -≤4.5 m [count]	1.05	1.02	1.08	0.001
Infected persons within ≤1.5 m [count] mutually adjusted c)	1.02	0.95	1.09	0.64
Infected persons in 1.5 -≤3 m [count]	0.98	0.93	1.03	0.36
Infected persons in 3 -≤4.5 m [count]	1.04	1	1.07	0.04

Supplementary table 1: Estimated relative risk of SARS-CoV-2 infection (IGG-positive) from logistic regression on summary measures of spatial proximity between participants in terms of odds ratio estimates (OR) with confidence interval and p-values. a) adjusted for sex, age, common household and duration. b) multivariate analysis, mutually adjusted for distance to ventilation system, participation in (multiple) performances, going out of doors during the intermission, and participating in the grand finale. c) multivariate analysis, mutually adjusted for distance to ventilation system, participation in (multiple) performances, going out of doors during the intermission, and participating in the grand finale and adjusted for sex, age, common household and duration.

Reporting checklist for cohort study.

Based on the STROBE cohort guidelines.

Instructions to authors

Complete this checklist by entering the page numbers from your manuscript where readers will find each of the items listed below.

Your article may not currently address all the items on the checklist. Please modify your text to include the missing information. If you are certain that an item does not apply, please write "n/a" and provide a short explanation.

Upload your completed checklist as an extra file when you submit to a journal.

In your methods section, say that you used the STROBE cohortreporting guidelines, and cite them as:

von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies.

		Reporting Item	Page Number
Title and abstract			
Title	#1a	Indicate the study's design with a commonly used term in the title or the abstract	1
Abstract	#1b	Provide in the abstract an informative and balanced summary of what was done and what was found	2-3
Introduction			
Background / rationale	#2	Explain the scientific background and rationale for the investigation being reported	4-5
Objectives	#3	State specific objectives, including any prespecified hypotheses	2
Methods			
Study design	#4	Present key elements of study design early in the paper	5
Setting	#5	Describe the setting, locations, and relevant dates, including periods	7

		of recruitment, exposure, follow-up, and data collection	
Eligibility criteria	#6a	Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up.	5
Eligibility criteria	#6b	For matched studies, give matching criteria and number of exposed and unexposed	5
Variables	#7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	5
Data sources / measurement	#8	For each variable of interest give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group. Give information separately for for exposed and unexposed groups if applicable.	7
Bias	#9	Describe any efforts to address potential sources of bias	4
Study size	#10	Explain how the study size was arrived at	5
Quantitative variables	#11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why	5-6
Statistical methods	#12a	Describe all statistical methods, including those used to control for confounding	7-8
Statistical methods	#12b	Describe any methods used to examine subgroups and interactions	7-8
Statistical methods	#12c	Explain how missing data were addressed	8
Statistical methods	#12d	If applicable, explain how loss to follow-up was addressed	8
Statistical methods	#12e	Describe any sensitivity analyses	8
Results			
Participants	#13a	Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible,	8

		included in the study, completing follow-up, and analysed. Give information separately for for exposed and unexposed groups if applicable.	
Participants	#13b	Give reasons for non-participation at each stage	8
Participants	#13c	Consider use of a flow diagram	
8			
Descriptive data	#14a	Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders. Give information separately for exposed and unexposed groups if applicable.	8
Descriptive data	#14b	Indicate number of participants with missing data for each variable of interest	
8			
Descriptive data	#14c	Summarise follow-up time (eg, average and total amount)	
8			
Outcome data	#15	Report numbers of outcome events or summary measures over time. Give information separately for exposed and unexposed groups if applicable.	
8			
Main results	#16a	Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	8-9
Main results	#16b	Report category boundaries when continuous variables were categorized	8-9
Main results	#16c	If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	
8-9			
Other analyses	#17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	10-11

Discussion

Key results	#18	Summarise key results with reference to study objectives	12
Limitations	#19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias.	12
Interpretation	#20	Give a cautious overall interpretation considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence.	11-12
Generalisability	#21	Discuss the generalisability (external validity) of the study results	12
Other Information			
Funding	#22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	15

 This checklist was completed on 01. December 2021 using https://www.goodreports.org/, a tool made by the EQUATOR Network in collaboration with Penelope.ai

BMJ Open

Dynamics, outcomes, and prerequisites of the first SARS-CoV-2 superspreading event in Germany, in February 2020: a cross-sectional epidemiological study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-059809.R1
Article Type:	Original research
Date Submitted by the Author:	23-Feb-2022
Complete List of Authors:	Wessendorf, Lukas; University Hospital Bonn, Virology Richter, Enrico; University Hospital Bonn, Virology Schulte, Bianca; University Hospital Bonn, Institute for Virology Schmithausen, Ricarda Maria; University Hospital Bonn, Hygiene Exner, Martin; University Hospital Bonn Lehmann, Nils; University of Duisburg-Essen, Institute for Medical Informatics, Biometry and Epidemiology Coenen, Martin; Rheinische Friedrich Wilhelms Universitat Bonn, Clinical Study Core Unit, Study Center Bonn (SZB) Fuhrmann, Christine; University Hospital Bonn Kellings, Angelika; University Hospital Bonn Hüsing, Anika; University of Duisburg-Essen, Institute of Medical Informatics, Biometry and Epidemiology Jöckel, Karl-Heinz; University of Duisburg-Essen, Institute of Medical Informatics, Biometry and Epidemiology Streeck, Hendrik; German Centre for Infection Research
Primary Subject Heading:	Infectious diseases
Secondary Subject Heading:	Public health
Keywords:	COVID-19, VIROLOGY, Infection control < INFECTIOUS DISEASES

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Dynamics, outcomes, and prerequisites of the first SARS-CoV-2 superspreading event in Germany, in February 2020: a cross-sectional epidemiological study

Lukas Wessendorf[1], Enrico Richter[1], Bianca Schulte[1], Ricarda Maria Schmithausen[2], Martin Exner[2], Nils Lehmann[3], Martin Coenen[4], Christine Fuhrmann[4], Angelika Kellings[4], Anika Hüsing[3], Karl-Heinz Jöckel[3], and Hendrik Streeck[1,*]

[1] Institute of Virology, University Hospital, University of Bonn, Germany, and German Center for Infection Research (DZIF), partner site Bonn-Cologne

[2] Institute of Hygiene and Public Health, University Hospital, University of Bonn, Germany

[3] Institute of Medical Informatics, Biometry und Epidemiology (IMIBE), University Hospital Essen, Germany

[4] Clinical Study Core Unit, Study Center Bonn (SZB), Institute of Clinical Chemistry and Clinical Pharmacology, University Hospital Bonn, Germany

[*] Corresponding author: Hendrik Streeck, Institute of Virology, University Hospital Bonn, Venusberg-Campus 1, 53127 Bonn, Germany, Hendrik.Streeck@ukbonn.de

Word Count: 4,581

Keywords: SARS-CoV-2, COVID-19, superspreading, cross-sectional study, ventilation

Abstract

Objectives

The first German SARS-CoV-2 outbreak was a superspreading event in Gangelt, North Rhine-Westphalia during indoor carnival festivities called “Kappensitzung” (15th of February 2020). We determined SARS-CoV-2 RT-PCR positivity rate, SARS-CoV-2-specific antibodies, and analyzed the conditions and dynamics of superspreading, including ventilation, setting dimensions, distance from infected persons and behavioral patterns.

Design

In a cross-sectional epidemiological study (51 days post-event), participants were asked to give blood, pharyngeal swabs and complete self-administered questionnaires.

Setting

The SARS-CoV-2 superspreading event took place during festivities in the small community of Gangelt in February 2020. This 5 h event included 450 people (6-79 years of age) in a building of 27m x 13.20m x 4.20m.

Participants

Out of 450 event participants, 411 volunteered to participate in this study.

Primary and Secondary Outcome Measures

Primary outcome: infection status (determined by IgG ELISA). Secondary outcome: symptoms (determined by questionnaire).

Results

Overall, 46% (n=186/404) of participants had been infected, and their spatial distribution was associated with proximity to the ventilation system (OR 1.39, 95% KI [0.86; 2.25]). Risk of infection was highly associated with age: children (OR: 0.33 [0.267; 0.414]) and young adults (age 18-25) had a lower risk of infection than older participants (average risk increase of 28% per 10 year). Behavioral differences were also risk-associated including time spent outside (OR: 0.55 [0.33; 0.91]) or smoking (OR: 0.32 [0.124; 0.81]).

Conclusions

Our findings underline the importance of proper indoor ventilation for future events. Lower susceptibility of children/young adults indicates their limited involvement in superspreading.

Strengths and limitations of this study

- The setting and the participant group are extremely well-defined.
- Participants were invited on the basis of one criterion, namely their presence at the superspreading event; there was no other preselection/bias in the study enrollment and the participation rate was high (91% of those invited).
- The study was conducted 51 days after the event, so it is possible that participants could have become infected unrelated to the event.
- The number of index cases is unknown.

Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is a highly transmissible and pathogenic RNA virus that emerged in late 2019 and has caused a pandemic threatening human health and public safety worldwide.[1] While factors shaping the dynamics of a pandemic are multifactorial, virulence and reproductive number are important properties of a virus.[2] For SARS-CoV-2 there is a substantial over-dispersion of the secondary infection distribution (individual R0) for an individual infected with SARS-CoV-2[2]. An over-dispersed R0 means that most infected people do not transmit (individual R0 = 0) while a minority of infected people are super-spreaders (individual R0 >5). Superspreading has been observed for many infectious pathogens, such as measles or SARS.[3] During the SARS pandemic in 2003 a superspreading event was defined as one infected person infecting eight others.[4] For SARS-CoV-2 it has been estimated that 80% of the infections are caused by 10% of infected individuals highlighting the importance of the cluster factor (k).[2] In Germany an indoor carnival event in the beginning of 2020 is considered as the first major outbreak in a German city and was considered a hotspot during the beginning of the pandemic in Germany.[5] Other SARS-CoV-2 superspreading events worldwide have been linked to indoor gatherings with close proximity of individuals.[6] Nevertheless, most of the reported superspreading events had less than 100 cases and the reports are limited by missing data or a reporting bias.[6]

Here, we closely examined the prerequisite of a unique super-spreading event in Germany during the SARS-CoV-2 pandemic, where nearly half of the participants became infected including children. We systematically analyzed infection rate, potential individual, and environmental risk factors for infection as well as the role of the ventilation system.

Materials and methods

Study design and sampling

This cross-sectional epidemiological study was conducted 51 days after a carnival celebration in the beginning of 2020. Eleven days after the event authorities sent all known participants into quarantine after testing 38 out of 99 individuals PCR-positive.

All adults known to have attended the event were invited to participate in the study. About 450 persons attended the event of which 411 participated in the study (**figure 1,** participation rate 91.3%). All study participants provided written informed consent before enrolment. The Ethics Committee of the Medical Faculty of the University of Bonn approved the study (approval number 085/20).

Self-administered questionnaires included questions about demographic background, symptoms of viral infection as well as detailed information about the behavior during the event. Participants' arrival and exit times were assessed in 1-hour categories. Study participants were asked to provide blood specimens and pharyngeal swabs for further analysis. The local health department supplied data on hospitalizations and fatalities in our cohort (manuscript submitted elsewhere).

Patient and public involvement

This study was designed in close collaboration with both the local health department of Heinsberg and the `Council of 11´ of Gangelt, the organizers of the event described herein. The organizers as well as the city´s head councilman were also involved in recruitment by appealing to the local population to participate in the study. Since the community of Gangelt was the center of the first German outbreak of SARS-CoV-2, there was a great interest from the local public to participate in this study to help understand this new virus and to gain access to early testing. Accordingly, the Ministry of Labor, Health, and Social Affairs of the state government funded this study. In turn, as a service to the public we informed each participant of their PCR and ELISA result via letter and offered a phone hotline for questions about the results.

Spatial information and description of the event

The event took place on February 15th, 2020 and consisted of speeches, dance, and music performances for a total of five hours, with one large intermission. It was a ticketed event, where ticket sale was open to the public. Most of the participants were inhabitants of Gangelt. It was hosted at a small community center (320 square meters) in a single open space with a bar in the front close to the entrance and a stage at the back. The tables, each with two benches, were arranged in two blocks with a center aisle towards the stage. Alcoholic and nonalcoholic drinks were served in glasses and a food truck was located outside in front of the venue. While most participants (about 450 people, 1.4 individuals per square meter) were sitting in the hall, a committee of

eleven individuals hosting the event were sitting on stage. The eleven people on stage switched after a break.

Metric room coordinates (length and width [m]) for areas, tables, benches and ventilation shafts were assessed via measurements, seating plan and photos from the event. Persons providing multiple positions were considered as spending an equal amount of time on different positions. When exact seating was unclear and information was available on table or greater area localisation (bar, stage), average coordinate values were used. On the grounds of these coordinates, we calculated pairwise metric distances between all persons and distances to closest inlet and outlet airshafts. For all persons their pairwise inverse distances were summarized as mean inverse distance. Inverse metric distances to persons or airshafts were regarded as representing infectious potential through local proximity, and inverse distances were capped at 2.5 (the inverse of the width of a seat of 0.4 m). Alternatively, we counted all infected persons within adjacent rings of 1.5 m width around each participant as a measure of crowdedness and infectious potential.

Pharyngeal swab and blood preparation

Pharyngeal swabs of participants were performed with FLOQSwabs (Copan) and immediately stored in UTM RT-mini tubes containing UTM Viral Stabilization Media (Copan) at 4 °C. Venous blood was drawn into EDTA tubes (Sarstedt) per participant and was transported to the laboratory at the University Hospital Bonn.

Anti-SARS-CoV-2 ELISA

Anti-SARS-CoV-2 IgA and IgGs were determined using enzyme-linked immunosorbent assays (ELISA) on the EUROIMMUN Analyzer I platform (EI 2606-9601 A, and EI2606-9601 G, respectively).[5] A result was considered positive when a ratio (extinction of sample/extinction of calibrator) of 0.8 or higher was reached. The guidelines of the German Medical Association (RiliBÄK) were abided by, including internal and external quality controls.

Reverse transcription polymerase chain reaction (RT-PCR)

Viral RNA was extracted from each 300µl swab sample via the chemagic Viral 300 assay (according to manufacturer's instructions) on the Perkin Elmer chemagic™ Prime™ instrument platform. The presence of two viral target genes (E and RdRP) was assessed in each sample by real time RT-PCR (SuperScript™III One-Step RT-

PCR System with Platinum™ TaqDNA Polymerase, Thermo Fisher). The following primers were used, for E gene: E_Sarbeco_F1 and R, and probe E_Sarbeco_P1, for RdRP gene: RdRP_SARSr_F, and R, and probe RdRP_SARSr-P2.[7] In addition, an internal control for RNA extraction, reverse transcription, and amplification was applied to each sample (innuDETECT Internal Control RNA Assay, Analytik Jena #845-ID-0007100). If amplification occurred in both virus-specific reactions samples were considered positive.

SARS-CoV-2 neutralization assay

A plaque reduction neutralization test was used to determine SARS-CoV-2 neutralization capacity as previously described.[5] Briefly, plasma samples were heat-inactivated and supernatant transferred to a new tube and serially two-fold diluted in OptiPRO^{TM}SFM (Gibco). 120 µL of each plasma dilution was mixed with 80 plaque-forming units (PFU) of SARS-CoV-2 in 120 µL OptiPRO^{TM}SFM cell culture medium and used to infect Vero E6 cells ($1.25x10^5$ cells/well seeded into 24-well plates 24 h before). Subsequently, the inoculum was removed and cells were overlayed with a mixture of carboxymethylcellulose (Sigma) and 2xMEM (Biochrom). Following 3-day incubation, the overlay was removed and the 24-well plates were fixed using a 6% (v/v) formaldehyde solution and stained with 1% (w/v) crystal violet in 20% ethanol revealing the formation of plaques. Finally, the neutralizing titers were calculated as the reciprocal of serum dilutions resulting in neutralization of 50% input virus (NT50), read out as reduction in the number of plaques.

Data management and quality control

The Clinical Study Core Unit of the Study Center Bonn (SZB) supported the study by outlining the study protocol and developing the informed consent form as well as participants information sheets with respect to data management and quality control. The data were gathered on paper-based Case Report Forms (pCRF). Data was entered as double-data-entry into the REDCAp study database programmed and hosted by SZB. Study personnel was trained by experienced members of the SZB. A quality manager was on site to support the study team. Monitoring of trial data and informed consent forms was performed according to the monitoring plan by qualified SZB staff. The ethics committee of the Medical Faculty of the University of Bonn was involved and approved the study (reference no. 085/20)

Statistical analysis

Associations between positive infection status (defined as an IgG ratio ≥0.8), and exposure variables were analysed via logistic regression models. Exposure variables were included crudely, and adjusted for the potential confounding factors age, sex, and duration of attendance as fixed effects. To correct for common household effects a random effects model was used. We present odds ratios with 95% confidence intervals. Because we present data on a single specific event among a limited number of participants, we completely refrain from presenting p-values. All analyses were done with SAS 9.4.

Results

411 out of an estimated 450 participants of the event responded to our study invitation, resulting in a response rate of 91.3%. 404 individuals provided plasma samples and 316 pharyngeal swabs (**figure 1**). Genders were represented equally among all 404 participants (n= 201/404, 50% were male) with a broad range in age (6-79 years, median age 36 years) and level of education (**table 1**). 297 individuals were residents of the community the event took place in, 103 lived in other parts of the county, and 11 were external visitors. In total five participants of the event were hospitalized and one participant subsequently died.

Overall, 186 out of 404 individuals tested seropositive for IgG- and 161 for IgA-antibodies (**suppl. figure 1**). To confirm seropositivity we performed a plaque reduction neutralization assay (**suppl. figure 2**) demonstrating neutralizing activity against SARS-CoV-2 of their respective antibody responses. Given the low specificity of the IgA assay, IgA seropositivity was not further considered.[5] 19 participants tested positive in RT-PCR; these were considered infected during the superspreading event only if they were also IgG-positive (this was the case with 16 out of the 19 participants).

Overall, we found that (n= 186/404) 46.0% (95% CI: [41.2%; 51.0%]) tested seropositive who attended the event, which was significantly higher than the overall estimated infection rate in the same community at large at that time. Indeed, officially 3.1% of the community were reported as positive cases at that time, but we estimated the infection rate to be 15.5% (95% CI:[12.3%; 19.0%])[5] for the community. Taken

together, an estimated 46% of participants became infected during a single superspreading event.

No association between the gender of participants and risk of infection was found ((OR: 1.01 [0.65; 1.58]) for women). On average infected individuals had a higher body mass index (26·2kg/m² compared to 24.3kg/m² for uninfected individuals). Infected participants were more likely to be clustered living in the same household (**table 1**). Having at least one comorbidity, including lung disease (n= 11/26, 42.3%), cardiovascular disease (n= 8/15, 53.3%), neurological disease (n= 1/6, 16.7%), cancer (n= 7/12) (58.3%) or diabetes (n= 4/5, 80%), did not increase the risk of infection (OR: 0.64 [0.33; 1.26]). We next assessed whether age influenced the risk of infection at the event, considering gender, duration of attendance and common household as covariates. Comparison across age-categories showed a lower risk for children (OR: 0.31 [0.14; 0.69]), and also for young adults (18-25 years, OR: 0.53 [0.26; 1.09]) as well as adults between 25 and 40 years (OR: 0.48 [0.28; 0.85]) in comparison to older adults (40 to 65 years) (OR: 1, reference), while seniors had a slightly higher risk (older than 65 years, OR: 1.1 [0.31; 3.97]) (**figure 2**). Our data suggests that an additional 10 years of age were on average associated with 28% increased risk of infection (OR: 1.28 [1.10; 1.48]).

To understand the spreading dynamics of SARS-CoV-2 during the event, we first performed a detailed analysis of potential risk factors and social behavior. We first analyzed whether the ventilation system influenced the distribution of SARS-CoV-2 infected individuals. It is important to state that the system´s air flow consisted of 75% used and 25% fresh air. The air flow can be described as clockwise. The air system uses vents along one side of the venue and on stage to take in air (**figure 3,** air inlets purple). After 25% of fresh air has been added and the air has been filtered, vents along the other side of the venue return the air into the room (**figure 3,** air outlets blue**)**. All ventilation points received the same amount of air due to throttle valves. For noise protection reasons windows remained closed. The air-system used F7-Filters (ISO ePM ≥ 2,5) and had an air volume flow of 7500 m^3/h.

Most tables located close to the air-inlets and showed no or only few infections (**figure 3,** green) also most surrounding tables showed low numbers of infection (**figure 3,** yellow). Tables close to the air-outlets show high (4 or 5 infected per table) and very high (6 or 7 infected per table) numbers of infected individuals. Infected participants

had been seated mostly at tables close to the bar, at the bar, or on stage. One table with 8 out of 11 infected people, was located far away from the bar at the other side of the hall and close to an air inlet. The group sitting on stage showed high numbers of infection (18 infected out of 24, **table 1**). Of note is that the overall number of participants per table was not equal for all tables. Greater proximity to air outlets was associated with increased risk of infection with a crude OR=1.39 [0.86; 2.25]. This association remained stable and was hardly attenuated from adjustment for proximity to air inlet, age, gender, duration of attendance, proximity to other infected persons, stage-activity and going outside during the intermission (**figure 4,** multiple adjusted OR=1.26 [0.63; 2.50]). A similar apparent effect for proximity to air inlets (crude OR=1.17 [0.72; 1.89]) disappeared when duration of attendance was added to the model (**figure 4,** multiple adjusted OR=1.01 [0.53; 1.94]). Overall, however, we found the increased risk for individuals located closer to the air outlet remarkably persistent (**figure 4**).

We further studied the sum of the inverse distance to all infected participants as a measure of proximity to either one common virus source or mutual infection. However, there was no evidence for increased risk of infection from greater proximity to other infected persons (**suppl. table 1**). Furthermore, we found no evidence for a single person being the source of the infection from the quantile-plot of p-values from 401 analyses conducted separately for each participant as potential source of infection (**suppl. figure 3**).

To understand the association of risk with behavior patterns we next investigated the influence of several factors on SARS-CoV-2 infection including time spent outside, smoking, performing on stage and participation during the final act (“Finale”) for 30 minutes. Results were all adjusted for age, sex, common household, and duration of attendance. Participation in multiple performances was associated with slightly increased risk of infection (OR per performance: 1.08 [0.91; 1.27]), results for participation in the last “Finale” were stronger (OR: 1.41 [0.65; 3.02]), although neither was significant (**figure 4**). Duration of attendance was persistently and strongly associated with an increased infection risk of 32% with each additional hour spent at the party (OR per hour: 1.32 [1.16; 1.49]). All other analyses were adjusted for this variable as potential confounding factor.

We next determined the level of alcohol consumption as number of drinks (high-proof liquor or beer) and did not observe any influence for the amount of alcohol consumption on the risk of becoming infected (OR per drink: 1.00 [0.96; 1.05]). Furthermore, participants who spent the break outside were less likely to be infected (OR: 0.55 [0.33; 0.91]) compared to individuals who spent the break inside the venue hall (**figure 4**). Interestingly, however, when we determined the impact of being regular smoker (defined as smoking of at least 10 cigarettes a day) on the risk of SARS-CoV-2 infection we observed a reduced risk of infection (OR: 0.32 [0.12; 0.81]) even after adjustment for "time spent outside". Taken together, our results demonstrated that the duration of attendance at the carnival party correlated with an increased risk of infection, but the number of alcoholic drinks was not associated with infection risk, while regular smoking and spending the break of the event outside showed a negative correlation with the risk of infection.

We next stratified seropositive individuals by their reported symptoms. Odds-ratios for each symptom were calculated for the timespan of 14 days following the event (**figure 5**). We identified that loss of smell (OR: 8.78 [4.81; 16.02]) and taste (OR: 10.09 [5.13; 19.88]) exhibited the strongest association with SARS-CoV-2 infection. Other symptoms which were strongly associated with COVID-19 were: sweats and chills (OR: 5.28 [3.08; 9.07]), muscle and joint ache (OR: 5.19 [3.19; 8.44]), fatigue (OR: 4.22 [2.76; 6.45]) and fever (OR: 3.73 [2.10; 6.63]) (**figure 5**). Importantly, 15.1% (28/186) of the infected individuals reported no symptoms at all in a period of 14 days after the event. The rate of asymptomatic infections of participants of the event was lower than generally observed in the community where the event took place (36%).[5] Overall, there was a lower proportion of asymptomatic cases among individuals infected after the event compared to members of the community, while loss of smell and taste showed the strongest association with an infection.

Discussion

The high overdispersion characteristics of SARS-CoV2 and its ability to be transmitted via aerosols under certain conditions are one of the main reasons that the beginning of the SARS-CoV2 pandemic was shaped by superspreading events.[8,9] Germany`s

first superspreading event was an indoor carnival event in the beginning of 2020 in a rural community. In this naturally occurring experiment, we found that nearly half of the participants became infected and determined multiple prerequisites for superspreading and risk factors for becoming infected. While our study population is not a representative sample of the general population the event may be regarded as exemplary for similar party occasions and may help reduce the number of those infected in the future. At the time of the event described herein SARS-CoV-2 had not diversified yet, but ever since many variants of the virus have arisen and have taken turns dominating the global pandemic. Therefore, the results shown here need to be viewed as qualified in describing a superspreading event under the circumstances in the beginning of the pandemic. However, they help us to understand infection dynamics and requisites for infection with this virus family, ultimately giving a frame of reference for similar studies conducted throughout the alpha, delta, and omicron waves of the COVID-19 pandemic.

An important factor associated with infection risk was the ventilation system and the individual proximity to the ventilation outlets. Individuals close to the air-outlets that contained air with low amount of fresh air had the highest infection risk compared to those close to the air-inlets. This was particularly interesting, because we did not see any increased risk of infection from greater proximity to other infected persons, which indicates that ventilation was perhaps more important than physical proximity. Our findings are in line with previous studies that demonstrated SARS-CoV-2 to be able to become air-borne under certain conditions and that the ventilation system can have an influence on virus spread.[10,11,12] The air filters in the venue were not capable of intercepting virus particles supporting the notion on the importance of proper indoor ventilation systems.[13,14] Indeed, spending the break of the event outside decreased the possibility of infection underscoring the benefit of proper ventilation or fresh air to lower the amount of aerosols. Due to the nature of the event, the spatial distribution of the participants was not fixed throughout the evening, and not perfectly recapitulated, so this information carries some error. However, allowing for multiple positions per person we used all available information. Assuming further error in the spatial data to be random, this might lead to a dilution of effects, i.e. true associations may remain undetected. Complementary analyses including e.g. the persons' functions during the event show consistent results, so we see no evidence suggesting bias in our findings.

Nevertheless, the infection-rate might be overestimated as the study was conducted 51 days after the event as participants could have become infected not related to the event. However, this weakness is limited by the official shut down of the community shortly after the event: A detailed timeline of the containment measures put in place after the superspreading event is included in Streeck et al.[5]. Briefly, a strict home quarantine for all attendees of the carnival event was imposed after 38 out of 99 participants tested positive for SARS-CoV-2. In addition, 13 days after the event the town went into full lockdown, including the closing of schools, childcare and outpatient care facilities, and restrictions of public access to the town. These concerted containment measures proved so effective that the peak of new infections in the community was already reached 27 days after the event.

The consumption of alcoholic drinks did not increase the risk of infection. While it has been assumed that the alcoholic effect of decreased social inhibition may increase likelihood of infection, we did not find any evidence for this association questioning measures of a ban on alcohol to reduce numbers of infected. It is known that current and former smokers disproportionately suffer from severe COVID-19 and their numbers are relatively increased among those patients that need intensive care treatment compared to non-smokers[15,16]. However, it has been previously speculated that the risk of infection is lower for smokers.[17] Furthermore, a meta-analysis of seven studies suggests that smokers have a reduced risk of testing positive for SARS-CoV-2.[18] Interestingly, we also observed that regular smoking lowered the risk of infection. The association might for example be explained by a role of the nicotinic acetylcholine receptor (nAChR).[19] Because other viruses, such as rabies virus, have been known to bind nAChRs, it was hypothesized recently, that SARS-CoV-2 spike protein might bind nAChRs as a coreceptor for infection.[20,21] Indeed, *in silico* molecular docking simulations predicted binding of spike to nAChRs.[22] If this interaction proves to be of advantage to the virus, then nicotine or its derivatives which bind nAChRs could compete with SARS-CoV-2 for binding and thereby reduce interactions of the virus with its target cells. Currently, at least one prospective observational study is being undertaken on the effects of smoking on COVID-19 infection rates, including a smoking cessation control group on nicotine substitutes.[23] While we strongly advise that smoking should not be considered as a protective habit to prevent risk of infection, this

knowledge may lead to the investigation of a therapeutic or prophylactic treatment on the basis of this molecular target.[24]

Our results indicate a trend that younger people are less likely to be infected compared to older age groups. This trend is strongest for people under 18 but levels out over 40 years of age. The risk of infection for children in superspreading events has not been investigated but the overall risk for infection in children seems to be lower than for adults as a systematic review and its recent update reported, which is further supported by our findings.[25,26] Considering the risk of infection with SARS-CoV-2 in general however, in a meta-analysis Madewell et al. conclude that the secondary attack rate in households is lower to children contacts than to adult contacts[27]. Many primary articles and meta-analyses point out the confounding effect of SARS-CoV-2 infections being mostly asymptomatic in young children has on the identification of children as index persons. To some extent, this problem could be avoided in our study since all participants of the event were invited to take part, regardless of age. As all individuals were exposed at the same event and time our study is a very suitable model for the previously described notion, that children are less likely to become infected. Indeed, a recently published meta-analysis by Viner et al. showed a low susceptibility for children and adolescents (OR of 0.56 (95% CI, 0.37-0.85)) which strongly supports our findings of a lower risk of infection in that age group, which is even lower in our study[28]. Our finding supports the previously shown minor influence on the spreading of the virus by children. The finding that for every 10 additional years of age the risk of infection increases during an event indicates that younger people and their limited role should be considered when measures to contain the pandemic are implemented. It should be mentioned that although children had similar exposure compared to adults and probably spent even less time outside the venue hall, the behaviors of children may be different compared to adults. Therefore, we cannot exclude that our findings of lower seroprevalence in children might be biased by factors very specific to this particular event. Taken together, we demonstrate important risk factors for infection during a superspreading event, which helps to understand transmission dynamics in order to improve comprehensive public health preparedness measures, including mandatory ventilation during indoor events and age-adjusted measures according to different risk of infection.

As to the strengths and limitations of this study, the participant group is extremely well-defined and there was no bias or preselection during enrollment as there was only one

criteria for invitation, namely presence at the event. Because of the time between the event and the study it is possible that participants were infected unrelated to the event, but the official shut down of the community limits this risk. The number of index cases during the event is not known and it is possible that a high number of individuals were already infectious. In addition, the identification of a past SARS-CoV-2 infection via serological test is not perfect and according to the manufacturer their IgG detection is 94.4 % sensitive (on samples collected >10 days after beginning of symptoms or direct detection of virus) and 99.0 % specific (for a ratio ≥0.8). For our infection rate analysis this predicts 2 false positives and 10 false negative IgG results. However, when field-tested by the UK National Health Service (NHS) the same assay showed 74.7 % sensitivity (62 false-negatives in our data set) and the same specificity of 99.0 %.

Acknowledgments

We would like to thank the inhabitants of the town of Gangelt for their participation. We would also like to thank the local government of the town where the outbreak took place for their support to conduct the study. Furthermore, we would like to thank the following people who helped with the study: Rudi Kirschen, Gero Wilbring, Janett Wieseler, Marek Korencak, Ryan Nattrass, Jernej Pusnik, Maximilian Becker, Ann-Sophie Boucher, Marc Alexander de Boer, Rebekka Dix, Sara Dohmen, Kim Friele, Benedikt Gansen, Jannik Geier, Marie Gronemeyer, Sarah Hundertmark, Nora Jansen, Michael Jost, Louisa Khorsandian, Simon Krzycki, Ekaterina Kuskova, Judith Langen, Silvia Letmathe, Ann-Kathrin Lippe, Jonathan Meinke, Freya Merker, Annika Modemann, Janine Petras, Sophie Marie Porath, Anna Quast, Laurine Reese, Isabel Maria Rehbach, Jonas Richter, Thea Rödig, Eva Schmitz, Tobias Schremmer, Louisa Sommer, Jennifer Speda, Yuhe Tang, Oliver Thanscheidt, Franz Thiele, Johanna Thiele, Julia Tholen, Moritz Transier, Maike van der Hoek, Tillmann Verbeek, Sophia Verspohl, Kira Vordermark, Julian Wirtz, Marina Wirtz, Lisa Zimmer, Philip Koenemann, Adi Yaser, Lisa Anna, Katharina Bartenschlager, Lisa Baum, Roxana Böhmer Romero, Diana de Braganca, Isabelle Engels, Moritz Färber, Carina Fernandez Gonzalez, Lucia Maria Goßner, Victoria Handschuch, Franziska Georgia Liermann, Steffan Meißner, Laura Racenski, Patrick, Denis Raguse, Larissa Reiß, Maximilian Rölle, Franziska Scheele, Chiara Schwippert, Arlene Christin Schwippert, Antonia Seifert, Joshua David Stockhausen, Sofia Waldorf, Leonie Weinhold, Nicolai Trimpop, Julia Reinhardt, Vera Gast, Michelle Yong, Eva Engels, Jonathan Meinke,

Susanne Schmidt, Janine Schulte, Saskia Schmitz, Kübra Bayrak, Regina Frizler, Katarzyna Andryka, Sofía Soler, Bastian Putschli, Thomas Zillinger, Marcel Renn, Patrick Müller, Dillon Corvino, Zeinab Abdullah, Katrin Paeschke, Hiroki Kato, Florian Schmidt, Maximilian Baum, Celina Schüter, Daniel Hinze, Martina Schmidt, Arcangelo Ricchiuto, Sonja Gross, Uta Wolber, Tobias Höller, Marion Zerlett, Esther Sib, Benjamin Marx, Souhaib Aldabbagh. We thank the Medizinischer Dienst der Krankenkasse (MDK) Nordrhein for their generous help in conducting this study. In particular we would like to thank: Tanja Bell, Ina de Hesselle-Taddey, Susanne Goll, Katja Hoffmann-Pruss, Niko Kalamakidis, Verena Kayser, Hildegard Kessler, Norbert Körfer, Daniela Kroll, Florian Messerschmidt, Svenja Peters, Sebastian Schröder, Linda Schroers, Michael Zimmermann, Klaus-Peter Thiele. We thank Stefan Holdenrieder, Alexander Semaan, Bernd Pötzsch, and Georg Nickenig for providing control samples. The idea, the plan, the concept, protocol, the conduct, the data analysis, and the writing of the manuscript of this study were independent of any third parties, including the government of North Rhine-Westphalia, Germany.

Contributors

L.W., E.R., B.S. and H.S. wrote the manuscript. L.W., R.M.S., and E.R. organized and ran the testing center. M.E. inspected the event venue and examined ventilation and air filtration systems. ER and BS organized and performed sample processing, experiments, analyses, and corrected the manuscript. N.L. and A.H. performed statistical analysis. M.C., C.F., and A.K. monitored the study. M.E., K.H.J. and H.S. oversaw the study and corrected the manuscript.

Data availability statement

All data from the study will be made available upon reasonable request.

Funding

The government of North Rhine-Westphalia (Germany) supported the study with 65,000 Euro.

Competing interests

The authors declare no conflict of interest.

References

1. Chakraborty I, Maity P. COVID-19 outbreak: Migration, effects on society, global environment and prevention. *Sci Total Environ* 2020; **728**: 138882.

2. Lloyd-Smith JO, Schreiber SJ, Kopp PE, Getz WM. Superspreading and the effect of individual variation on disease emergence. *Nature* 2005; **438**(7066): 355-9.

3. De Serres G, Markowski F, Toth E, et al. Largest measles epidemic in North America in a decade--Quebec, Canada, 2011: contribution of susceptibility, serendipity, and superspreading events. *J Infect Dis* 2013; **207**(6): 990-8.

4. Shen Z, Ning F, Zhou W, et al. Superspreading SARS events, Beijing, 2003. *Emerg Infect Dis* 2004; **10**(2): 256-60.

5. Streeck H, Schulte B, Kümmerer BM, et al. Infection fatality rate of SARS-CoV2 in a super-spreading event in Germany. *Nat Commun* 2020; **11**(1): 5829.

6. Leclerc QJ, Fuller NM, Knight LE, Funk S, Knight GM. What settings have been linked to SARS-CoV-2 transmission clusters? *Wellcome Open Res* 2020; **5**: 83.

7. Corman VM, Landt O, Kaiser M, et al. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Euro Surveill* 2020; **25**(3).

8. Kyriakopoulos AM, Papaefthymiou A, Georgilas N, Doulberis M, Kountouras J. The Potential Role of Super Spread Events in SARS-COV-2 Pandemic; a Narrative Review. *Arch Acad Emerg Med* 2020; **8**(1): e74.

9. Patel KP, Vunnam SR, Patel PA, et al. Transmission of SARS-CoV-2: an update of current literature. *Eur J Clin Microbiol Infect Dis* 2020; **39**(11): 2005-11.

10. Lednicky JA, Lauzardo M, Fan ZH, et al. Viable SARS-CoV-2 in the air of a hospital room with COVID-19 patients. *Int J Infect Dis* 2020; **100**: 476-82.

11. Santarpia JL, Rivera DN, Herrera VL, et al. Aerosol and surface contamination of SARS-CoV-2 observed in quarantine and isolation care. *Sci Rep* 2020; **10**(1): 12732.

12. Günther T, Czech-Sioli M, Indenbirken D, et al. SARS-CoV-2 outbreak investigation in a German meat processing plant. *EMBO Mol Med* 2020; **12**(12): e13296.

13. Pokora R, Kutschbach S, Weigl M, et al. Investigation of superspreading COVID-19 outbreak events in meat and poultry processing plants in Germany: A cross-sectional study. *PLoS One* 2021; **16**(6): e0242456.

14. Nazarenko Y. Air filtration and SARS-CoV-2. *Epidemiol Health* 2020; **42**: e2020049.

15. Vardavas CI, Nikitara K. COVID-19 and smoking: A systematic review of the evidence. *Tob Induc Dis* 2020; **18**: 20.

16. Engin AB, Engin ED, Engin A. Two important controversial risk factors in SARS-CoV-2 infection: Obesity and smoking. *Environ Toxicol Pharmacol* 2020; **78**: 103411.

17. Israel A, Feldhamer E, Lahad A, Levin-Zamir D, Lavie G. Smoking and the risk of COVID-19 in a large observational population study. *medRxiv* 2020: 2020.06.01.20118877.

18. Grundy EJ, Suddek T, Filippidis FT, Majeed A, Coronini-Cronberg S. Smoking, SARS-CoV-2 and COVID-19: A review of reviews considering implications for public health policy and practice. *Tob Induc Dis* 2020; **18**: 58.

19. Changeux JP, Amoura Z, Rey FA, Miyara M. A nicotinic hypothesis for Covid-19 with preventive and therapeutic implications. *C R Biol* 2020; **343**(1): 33-9.

20. Stefano ML, Kream RM, Stefano GB. A Novel Vaccine Employing Non-Replicating Rabies Virus Expressing Chimeric SARS-CoV-2 Spike Protein Domains: Functional Inhibition of Viral/Nicotinic Acetylcholine Receptor Complexes. *Med Sci Monit* 2020; **26**: e926016.

21. Dormoy V, Perotin JM, Gosset P, Maskos U, Polette M, Deslée G. Nicotinic receptors as SARS-CoV-2 spike co-receptors? *Med Hypotheses* 2022; **158**: 110741.

22. F. Oliveira AS, Ibarra AA, Bermudez I, et al. Simulations support the interaction of the SARS-CoV-2 spike protein with nicotinic acetylcholine receptors. *bioRxiv* 2020: 2020.07.16.206680.

23. clinicaltrials.gov. Impact of Smoking and Nicotine on the Risk of Being Infected With COVID-19 (MAGIC). 2021. https://clinicaltrials.gov/ct2/show/NCT04429815.

24. Paleiron N, Mayet A, Marbac V, et al. Impact of Tobacco Smoking on the Risk of COVID-19: A Large Scale Retrospective Cohort Study. *Nicotine Tob Res* 2021; **23**(8): 1398-404.

25. Li X, Xu W, Dozier M, He Y, Kirolos A, Theodoratou E. The role of children in transmission of SARS-CoV-2: A rapid review. *J Glob Health* 2020; **10**(1): 011101.

26. Li X, Xu W, Dozier M, et al. The role of children in the transmission of SARS-CoV2: updated rapid review. *J Glob Health* 2020; **10**(2): 021101.

27. Madewell, Z. J., Yang, Y., Longini, I. M., Jr., Halloran, M. E., Dean, N. E. Household Transmission of SARS-CoV-2: A Systematic Review and Meta-analysis. JAMA Netw Open 2020; 3(12): e2031756

28. Viner RM, Mytton OT, Bonell C, et al. Susceptibility to SARS-CoV-2 Infection Among Children and Adolescents Compared With Adults: A Systematic Review and Meta-analysis. *JAMA Pediatr* 2021; **175**(2): 143-56.

Tables

		Not infected N%	Infected N%
Total number		218	186
Female		114 (52%)	89 (48%)
Age:	<18 years	31 (14%)	15 (8%)
	18- 24 years	30 (14%)	20 (11%)
	25-39 years	81 (37%)	43 (23%)
	40-64 years	71 (33%)	100 (54%)
	65+ years	5 (2%)	8 (4%)
BMI [kg/m²] (Std Dev)		24.3 (5.12)	26.2 (5.16)
Participating household member (Std Dev)		2.1 (1.12)	2.4 (1.16)
Highest level of formal education:			
	None	27 (13%)	13 (7%)
	lower secondary school	27 (13%)	23 (13%)
	secondary school	55 (26%)	71 (39%)
	higher education entrance qualification	54 (25%)	34 (18%)
	(technical) university degree	52 (24%)	43 (23%)
Duration of attendance [h] (Std Dev)		4.7 (2.06)	5.8 (1.85)
Service team		4 (2%)	22 (12%)
On stage during event		80 (37%)	62 (34%)
Member of „Council of 11“		6 (3%)	18 (10%)
On stage during “Finale”		26 (12%)	48 (26%)
Behavior during break:			
	Remaining seated	73 (36%)	85 (48%)

Going outside	114 (55%)	72 (39%)
Alcohol consumption [drink] (Std Dev)	11.3 (7.76)	12.2 (7.40)
Former smoker	34 (16%)	45 (24%)
Active smoker (≥10 cigarettes/day)	54 (25%)	23 (12%)
At least one comorbidity	29 (13%)	28 (15%)
Avg. distance to other participants [m] (Std Dev)	9.2 (1.68)	9.1 (1.70)
Distance to air inlet [m] (Std Dev)	6.1 (3.22)	6.0 (3.30)
Distance to air outlet [m] (Std Dev)	4.8 (2.94)	5.1 (2.87)

Table 1: Distribution of demographic factors and exposure information of interest among study participants who tested positive or negative in serology test of SARS-Cov2-infection
`Council of 11´ stands for the hosts of the events located on stage (personnel switched during the break). `Finale´ describes the final presentation of the event with all performers on stage.

Figure Legends

Figure 1: Study participants
Enrollment and flow of participants through the study. Downstream sample processing included centrifugation of blood samples for plasma collection (SARS-CoV-2 ELISAs), and viral RNA extraction from swab samples (SARS-CoV-2 RT-PCR).

Figure 2: Odds ratios for the likelihood of SARS-CoV-2 infection by age groups

Participants were divided into age groups of 8, 15, or 25 years, participants younger than 18 or older than 65 years. Participants were considered to have been infected during the event if they were SARS-CoV-2 antibody positive (ELISA).

Figure 3: Reconstructed 3D-Model of the venue hall
The venue was a single open space with a stage on one end and a bar as well as the exit on the opposite end. Distribution of tables and seating was as indicated by table and chairs symbols. Please note that the people pictured are illustrative and do not represent individual participants. Self-administered questionnaires included questions about main seating-position of the participants during the evening event as specifying table and seat with the help of a schematic seating plan. Metric room coordinates for all tables, seats, and ventilation-points were assessed and the seating was reconstructed from pictures taken during the event. Therefore, the location of the stage, the bar, the exit as well as the tables and the air-inlets/outlets were reconstructed in a 3D-Model. The original external dimensions of the building were 27m x 13.20m x 4.20m. Tables, where more than 7 infected individuals have stayed are colored in dark red, this includes the stage and bar as well. Air-inlets are colored in violet and the air-outlets in blue. Infected participants had been seated mostly at tables close to the bar, the bar itself and on stage. One table with 8 out of 11 infected people, was located far away from the bar at the other side of the hall and close to an air inlet. The group sitting on stage showed as well high numbers of infection (18 infected out of 24).

Figure 4: Odds ratios for the association of SARS-CoV-2 infection with specific activities of the participants and their location in the venue relative to ventilation shafts

The model was additionally adjusted for age, sex, duration of attendance, participation in multiple activities, and cumulative proximity to other infected persons, and common household.

Figure 5: Odds ratios for symptoms of SARS-CoV-2 antibody-positive participants in the 14 days following the super spreading event

The information on symptoms was derived from the self-administered questionnaire, which was filled out on the day of sample collection. Odds ratio estimates (OR) are shown with confidence intervals.

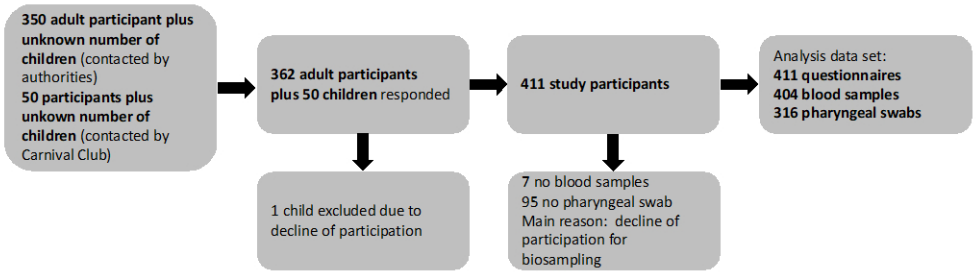


Fig. 1: Study participants. Enrollment and flow of participants through the study. Downstream sample processing included centrifugation of blood samples for plasma collection (SARS-CoV-2 ELISAs), and viral RNA extraction from swab samples (SARS-CoV-2 RT-PCR).

338x123mm (78 x 78 DPI)

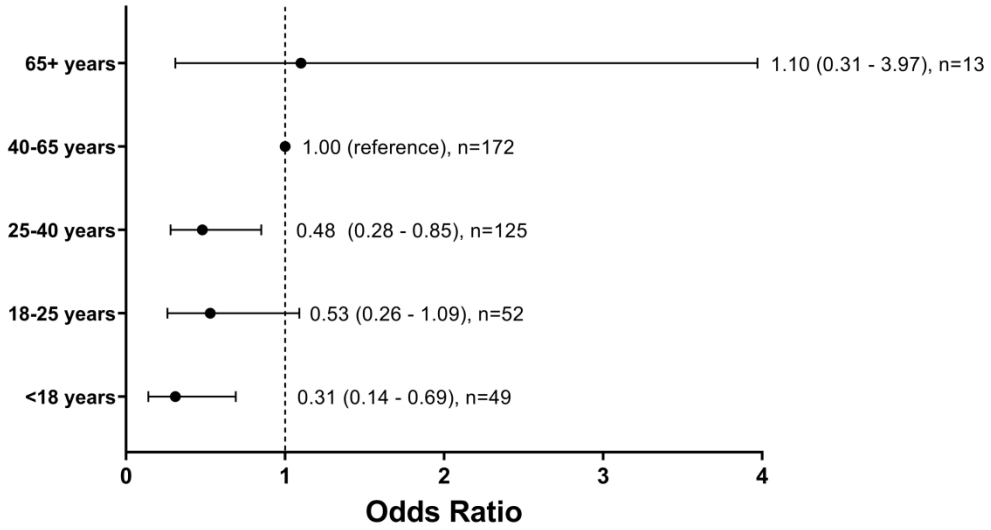


Fig. 2 Odds-Ratio for the likelihood of SARS-CoV-2 infection by age groups. Participants were divided into age groups of 8, 15, or 25 years, participants younger than 18 or older than 65 years. Participants were considered to have been infected during the event if they were SARS-CoV-2 antibody positive (ELISA).

283x160mm (300 x 300 DPI)

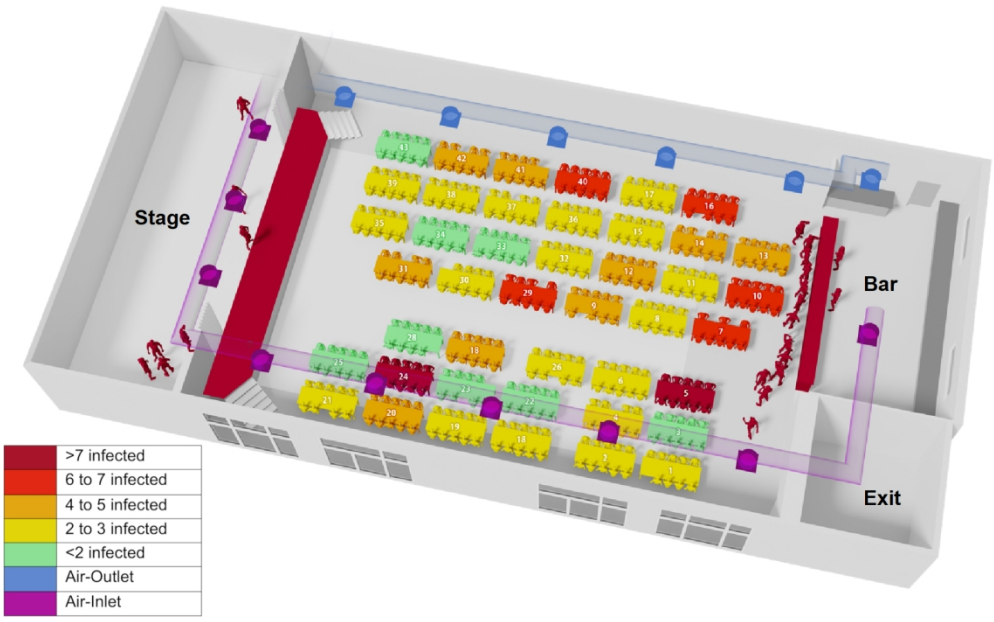


Fig. 3: Reconstructed 3D-Model of the venue hall. The venue was a single open space with a stage on one end and a bar as well as the exit on the opposite end. Distribution of tables and seating was as indicated by table and chairs symbols. Please note that the people pictured are illustrative and do not represent individual participants. Self-administered questionnaires included questions about main seating-position of the participants during the evening event as specifying table and seat with the help of a schematic seating plan. Metric room coordinates for all tables, seats, and ventilation-points were assessed and the seating was reconstructed from pictures taken during the event. Therefore, the location of the stage, the bar, the exit as well as the tables and the air-inlets/outlets were reconstructed in a 3D-Model. The original external dimensions of the building were 27m x 13.20m x 4.20m. Tables, where more than 7 infected individuals have stayed are colored in dark red, this includes the stage and bar as well. Air-inlets are colored in violet and the air-outlets in blue. Infected participants had been seated mostly at tables close to the bar, the bar itself and on stage. One table with 8 out of 11 infected people, was located far away from the bar at the other side of the hall and close to an air inlet. The group sitting on stage showed as well high numbers of infection (18 infected out of 24).

155x96mm (300 x 300 DPI)

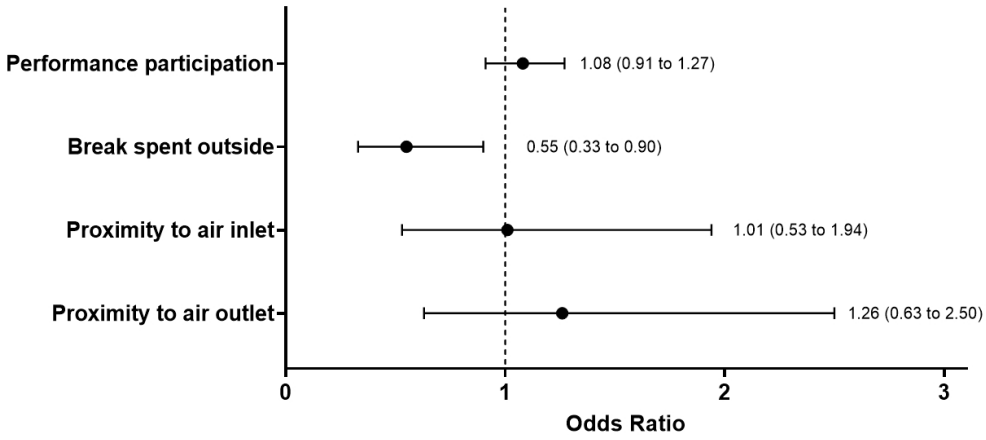


Fig. 4: Odds ratios for the association of SARS-CoV-2 infection with specific activities of the participants and their location in the venue relative to ventilation shafts. The model was additionally adjusted for age, sex, duration of attendance, participation in multiple activities, and cumulative proximity to other infected persons, and common household.

259x115mm (120 x 120 DPI)

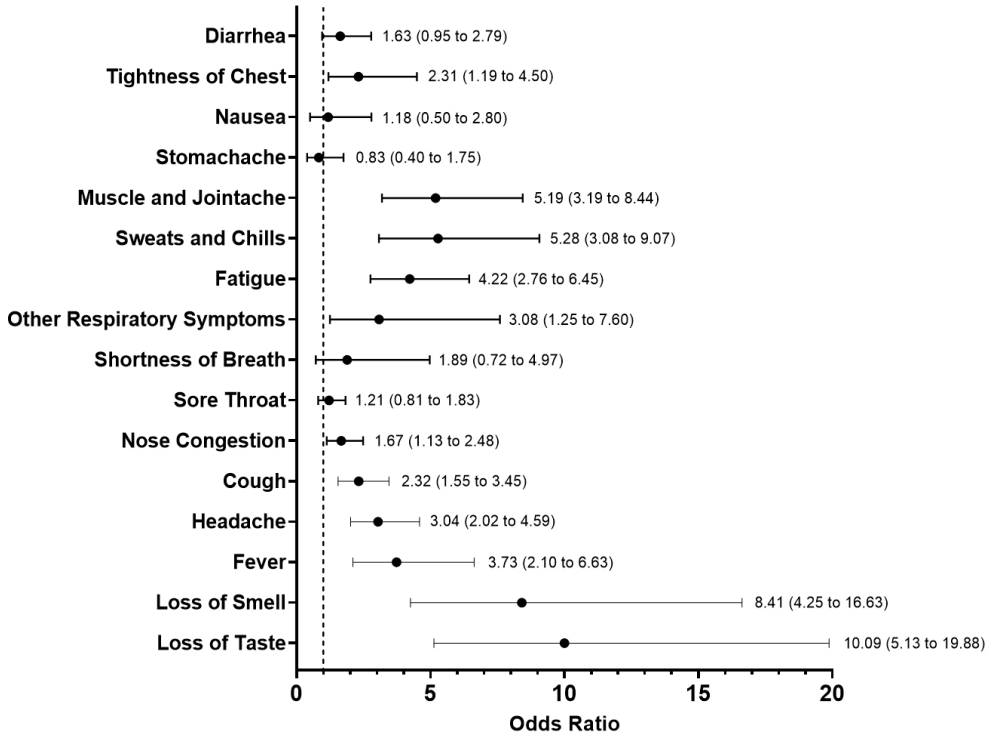
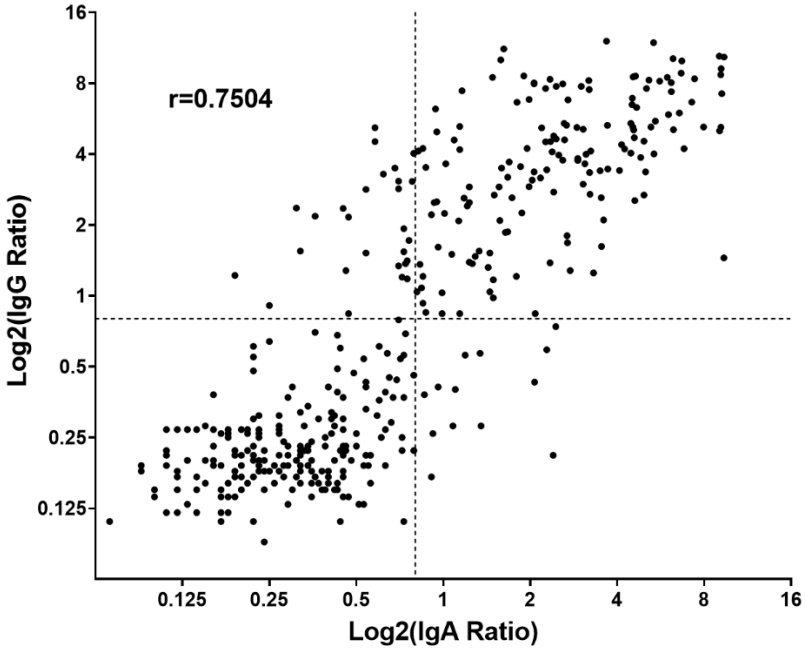


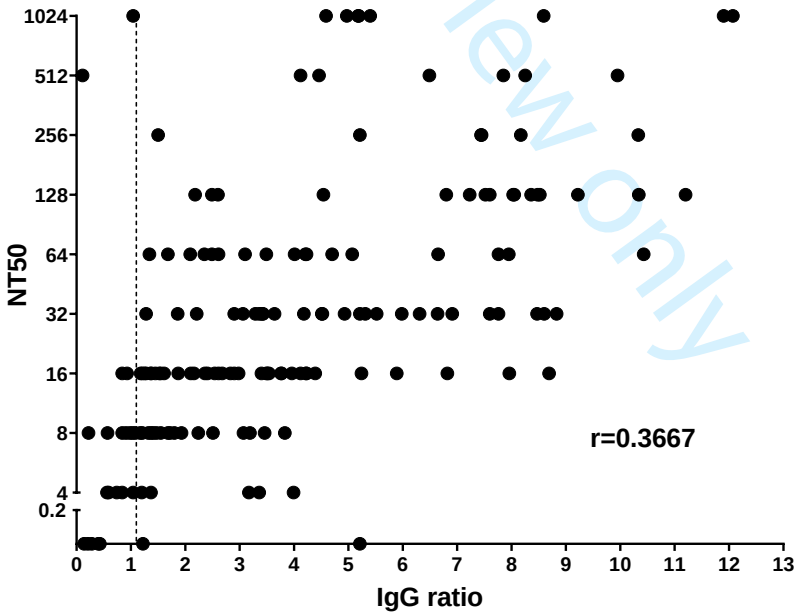
Fig. 5: Odds ratios for symptoms of SARS-CoV-2 antibody-positive participants in the 14 days following the super spreading event. The information on symptoms was derived from the self-administered questionnaire, which was filled out on the day of sample collection. Odds ratio estimates (OR) are shown with confidence intervals.

280x208mm (120 x 120 DPI)

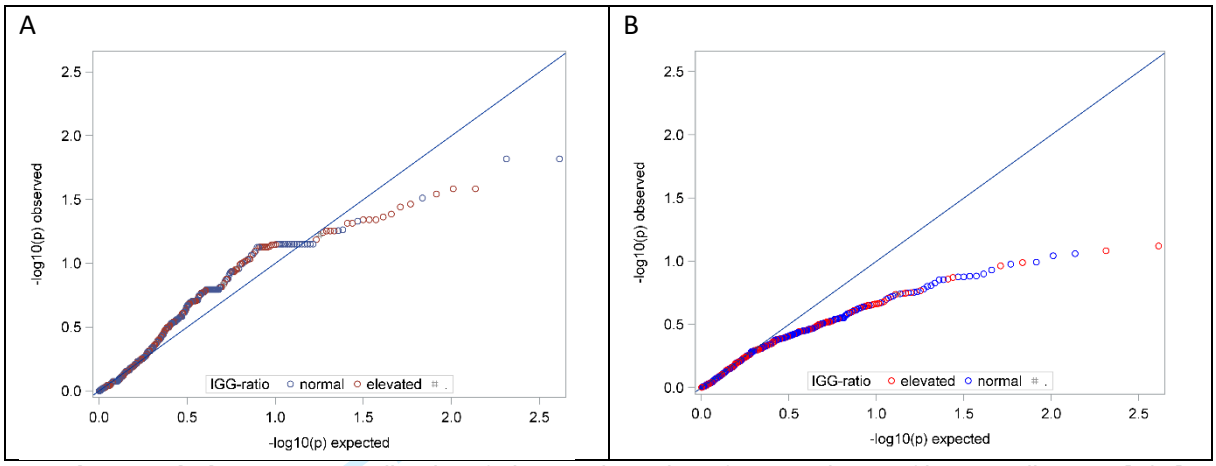
Supplementary Material



Suppl. Fig. 1: Correlation of SARS-CoV-2 Euroimmun ELISA results for IgA and IgG. The correlation of IgA levels to IgG levels in the same person was significant (r: Pearson coefficient, p<0.0001, 95 % CI, 0.7043 to 0.7902). The dotted lines mark the ratios above which each ELISA result is considered positive.



Suppl. Fig. 2: Correlation of plasma neutralization capacity and IgG ELISA results (Euroimmun) from each donor. The dotted line marks the ratio above which the ELISA result is considered positive. The correlation coefficient (Pearson) was 0.3667 (95 % CI, 0.2275 to 0.4192, p<0.0001). Samples with a negative result in the neutralization assay were set as 0.1 here so as to appear on the logarithmic axis.



Supplemental Figure 3: Quantile plot of observed p-values from analyses of inverse distance [1/m] to single specific study participants as risk factor for corona-virus infection. In case of no association, the ordered log-transformed p-values are expected to lie on, or below the diagonal. Panel A: results from crude analyses, Panel B: analyses were adjusted for age, sex, common household and duration of attendance.

		OR	95% confidence interval		p-value
proximity to infected persons [sum 1/m]		**0,99**	0,98	1,01	0,430
	adjusted a)	**1,00**	0,98	1,02	0,957
	mutually adjusted b)	**0,99**	0,97	1,01	0,571
	mutually adjusted c)	**0,99**	0,97	1,02	0,646
Alternative consideration in distance-bands					
Infected persons within ≤1.5m [count]		**1,01**	0,96	1,07	0,681
Infected persons in 1.5 - ≤3m [count]		**0,96**	0,92	1,00	0,043
Infected persons in 3 - ≤4.5m [count]		**1,03**	1,00	1,06	0,023
Infected within ≤1.5m [count]	adjusted a)	**1,03**	0,97	1,10	0,366
Infected in 1.5 - ≤3m [count]		**0,96**	0,92	1,01	0,113
Infected in 3 - ≤4.5m [count]		**1,03**	1,00	1,06	0,083
Infected within ≤1.5m [count]	mutually adjusted b)	**1,01**	0,95	1,07	0,734
Infected in 1.5 - ≤3m [count]		**0,98**	0,94	1,02	0,359
Infected in 3 - ≤4.5m [count]		**1,05**	1,02	1,08	0,001
Infected within ≤1.5m [count]	mutually adjusted c)	**1,02**	0,95	1,09	0,638
Infected in 1.5 - ≤3m [count]		**0,98**	0,93	1,03	0,363
Infected in 3 - ≤4.5m [count]		**1,04**	1,00	1,07	0,041

Supplementary table 1: Estimated relative risk of SARS-CoV-2 infection (IGG-positive) from logistic regression on summary measures of spatial proximity between participants in terms of odds ratio estimates (OR) with confidence interval and p-values. a) adjusted for sex, age, common household and duration. b) multivariate analysis, mutually adjusted for distance to ventilation system, participation in (multiple) performances, going out of doors during the intermission, and participating in the grand finale. c) multivariate analysis, mutually adjusted for distance to ventilation system, participation in (multiple) performances, going out of doors during the intermission, and participating in the grand finale and adjusted for sex, age, common household and duration.

STROBE Statement—Checklist of items that should be included in reports of ***cross-sectional studies***

	Item No	Recommendation	Page No
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	3
Objectives	3	State specific objectives, including any prespecified hypotheses	3
Methods			
Study design	4	Present key elements of study design early in the paper	3,4
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	3-7
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants	3,4
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	2
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	*4*
Bias	9	Describe any efforts to address potential sources of bias	3
Study size	10	Explain how the study size was arrived at	3
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	4
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	4
		(*b*) Describe any methods used to examine subgroups and interactions	5
		(*c*) Explain how missing data were addressed	4
		(*d*) If applicable, describe analytical methods taking account of sampling strategy	
		(*e*) Describe any sensitivity analyses	14
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	19
		(b) Give reasons for non-participation at each stage	19
		(c) Consider use of a flow diagram	19
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	7
		(b) Indicate number of participants with missing data for each variable of interest	7
Outcome data	15*	Report numbers of outcome events or summary measures	7,8
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	9,10

		(*b*) Report category boundaries when continuous variables were categorized	8,10
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	9-11
Discussion			
Key results	18	Summarise key results with reference to study objectives	11-13
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	14
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	13,14
Generalisability	21	Discuss the generalisability (external validity) of the study results	11
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	16

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.