Annals of Work Exposures and Health, 2022, Vol. XX, No. XX, 1–12
https://doi.org/10.1093/annweh/wxac032
Original Article

Original Article

Validation of a COVID-19 Job Exposure Matrix (COVID-19-JEM) for Occupational Risk of a SARS-CoV-2 Infection at Work: Using Data of Dutch Workers

Sophie van der Feltz[1], Susan Peters[2], Anjoeka Pronk[1,3], Vivi Schlünssen[4,5], Zara A. Stokholm[6], Henrik A. Kolstad[6], Karin van Veldhoven[7], Ioannis Basinas[8], Martie van Tongeren[8], Alex Burdorf[9] and Karen M. Oude Hengel[1,*]

[1]Netherlands Organisation for Applied Scientific Research TNO, Department of Work Health Technology, Unit Healthy Living, Schipholweg 79-86, 2316 ZL Leiden, The Netherlands; [2]Department Population Health Sciences, Institute for Risk Assessment Sciences, Utrecht University, Yalelaan 2, 3584 CM Utrecht, The Netherlands; [3]Netherlands Organisation for Applied Scientific Research TNO, Department pf Risk Analysis for Products in Development, Princetonlaan 6 3584 CB Utrecht, The Netherlands; [4]Department of Public Health, Environment, Occupation and Health, Danish Ramazzini Centre, Aarhus University, Bartholins Allé 2, 8000 Aarhus C, Denmark; [5]National Research Centre for the Working Environment, Lersø Parkallé 105, 2100 Copenhagen Ø, Denmark; [6]Department of Occupational Medicine, Danish Ramazzini Centre, Aarhus University Hospital, Palle Juul-Jensens Boulevard 99, 8200 Aarhus N, Denmark; [7]London School of Hygiene & Tropical Medicine, Faculty of Epidemiology and Public Health, Department of Non-Communicable Disease Epidemiology, Keppel Street, London WC1E 7HT, UK; [8]Centre for Occupational and Environmental Health, Centre for Epidemiology, Division of Population Health, Health Services Research and Primary Care, School of Health Sciences, Faculty of Biology, Medicine and Health, The University of Manchester, Manchester Academic Health Science Centre, Manchester, UK; [9]Erasmus Medical Center Rotterdam, Department of Public Health, Wytemaweg 80, 3015 CN Rotterdam, The Netherlands

*Author to whom correspondence should be addressed. Tel: +31-6-468-47-269; e-mail: karen.oudehengel@tno.nl

Submitted 22 December 2021; revised 11 April 2022; editorial decision 15 April 2022; revised version accepted 19 April 2022.

Abstract

Objectives: A COVID-19 Job Exposure Matrix (COVID-19-JEM) has been developed, consisting of four dimensions on transmission, two on mitigation measures, and two on precarious work. This study aims to validate the COVID-19-JEM by (i) comparing risk scores assigned by the COVID-19-JEM with self-reported data, and (ii) estimating the associations between the COVID-19-JEM risk scores and self-reported COVID-19.

What's important about this paper

Due to the lack of individual data on exposure to SARS-CoV-2 at the workplace, an international COVID-19 Job Exposure Matrix (COVID-19-JEM) has been developed that can identify job groups with likely exposure at work. This study is the first to validate the COVID-19-JEM by comparing the risk scores assigned by the COVID-19-JEM with risk scores based on self-reported data (i.e. direct validation), and by evaluating the association between COVID-19-JEM assigned risk scores and self-reported COVID-19 (i.e. indirect validation). The direct validation showed moderate to good agreement on the transmission risk dimensions and the social distance dimension, but not on the face covering dimension. The indirect validation showed in general that higher risk scores within each dimension of the COVID-19-JEM are associated with having had COVID-19, especially when the infection with SARS-CoV-2 was thought to have occurred at work.

Methods: Data from measurements 2 (July 2020, n = 7690) and 4 (March 2021, n = 6794) of the Netherlands Working Conditions Survey-COVID-19 (NWCS-COVID-19) cohort study were used. Responses to questions related to the transmission risks and mitigation measures of Measurement 2 were used to calculate self-reported risk scores. These scores were compared with the COVID-19-JEM attributed risk scores, by assessing the percentage agreement and weighted kappa (κ). Based on Measurement 4, logistic regression analyses were conducted to estimate the associations between all COVID-19-JEM risk scores and self-reported COVID-19 (infection in general and infected at work).

Results: The agreement between the COVID-19-JEM and questionnaire-based risk scores was good ($\kappa \geq 0.70$) for most dimensions, except work location ($\kappa = 0.56$), and face covering ($\kappa = 0.41$). Apart from the precarious work dimensions, higher COVID-19-JEM assigned risk scores had higher odds ratios (ORs; ranging between 1.28 and 1.80) on having had COVID-19. Associations were stronger when the infection were thought to have happened at work (ORs between 2.33 and 11.62).

Conclusions: Generally, the COVID-19-JEM showed a good agreement with self-reported infection risks and infection rates at work. The next step is to validate the COVID-19-JEM with objective data in the Netherlands and beyond.

Keywords: COVID-19; Job Exposure Matrix; mitigation measures; precarious work; SARS-CoV-2 infection; transmission risk

Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has been spreading across the world for over 2 years, causing a pandemic of COVID-19. Worldwide, governments have taken measures to control the rapid spread of COVID-19. Beside the development and implementation of vaccinations, primary strategies for this control include social distancing, contact tracing, and wearing face masks (Wiersinga *et al.*, 2020). In many countries, these strategies are incorporated in measures, affecting the entire population. The workplace is one of the key settings in the spread of COVID-19 and plays a role in COVID-19 caused mortality (Baker *et al.*, 2020; Lan *et al.*, 2020; Marinaccio *et al.*, 2020a; Beale *et al.*, 2022; Nafilyan *et al.*, 2021). To minimize social interactions, and thereby reduce infection rates, governments encouraged workers to work from home as much as possible, and restricted opening hours of other sectors several times during the pandemic, although responses varied between countries and stages of the pandemic. However, essential workers who are vital for the core function of society (e.g. nurses, construction workers, and bus drivers), still had to go to their workplaces during the lockdowns. These workers therefore had a higher probability of being exposed to SARS-CoV-2, including in their workplace.

Early studies found that occupations with high infection rates were healthcare workers (Marinaccio *et al.*, 2020b; Wu and McGoogan, 2020; Mutambudzi *et al.*, 2021) and other workers in close contact with the general public such as those in hospitality and public transport, but also driving instructors and hairdressers (de Gier *et al.*, 2020). Outbreaks have also been described in industries with a high prevalence of precarious workers, such as agriculture and meat processing (Dyal, 2020; Fassani and Mazza, 2020). In England and

Wales, workers with an elevated risk of infection were healthcare workers, indoor trade workers, process and plant workers, leisure and personal service workers, transport workers and workers working with mobile machines (Beale *et al.*, 2022). This elevated risk was substantially accounted for by the frequency of contacts at the workplace. A Norwegian study found that occupational incidence of COVID-19 differed between waves, depending on the governmental measures (Magnusson *et al.*, 2021). During the first COVID-19 wave the healthcare sector (mainly nurses, physicians, physiotherapists, and dentists) and the public transport sector were most at risk. During the second wave, teachers had an increased risk of COVID-19 compared with other occupations (Magnusson *et al.*, 2021).

As the workplace plays a role in the spread of SARS-COV-2 infections, it is important to assess the occupations at increased risk of exposure to SARS-CoV-2. However, it is challenging to obtain individual exposure data on SARS-CoV-2 infections in large populations, if not impossible considering the time scales involved. Therefore, different tools have been developed during the pandemic to identify levels of COVID-19 risk by occupation to guide policymakers, occupational safety and health (OSH) practitioners, and employers during the pandemic (Baker *et al.*, 2020; Zhang, 2021; Oude Hengel *et al.*, 2022; Williams *et al.*, 2022). Job Exposure Matrices (JEM) have been developed for a wide range of exposures that occur in the workplace and convert occupations into estimates of exposure (Kromhout and Vermeulen, 2001; Peters, 2020). Oude Hengel *et al.* (2022) developed an international COVID-19 Job Exposure Matrix (COVID-19-JEM), consisting of eight dimensions associated with risk of transmission, mitigation measures, and precarious work, to enhance the investigation of the role of the workplace in the spread of SARS-CoV-2 infections and subsequent cases of COVID-19 disease (Oude Hengel *et al.*, 2022). This COVID-19-JEM was developed by occupational exposure experts from three countries—i.e. Denmark, the Netherlands, and the UK.

As with all JEMs and exposure tools, it is important to understand the performance of this COVID-19-JEM. Therefore, the aim of this study is to directly and indirectly validate the expert-based COVID-19-JEM against self-reported data on both SARS-CoV-2 exposure and COVID-19 incidence in the Dutch context.

Methods

COVID-19-JEM

The COVID-19-JEM consists of eight dimensions related to the risk of exposure to SARS-CoV-2 (four determinants of transmission risk, two mitigation measures, and two factors on precarious work; Table 1). All 436 job titles within the International Standard Classification of Occupations from 2008 (ISCO-08) were assigned an exposure risk score ranging from zero to three (none, low, intermediate, high) for each dimension. The ISCO-08 is an international classification of occupations based on tasks and duties (International Labour Office, 2012). The attribution of risk scores to the job titles was performed within each country (Denmark, the Netherlands, and the UK) by three occupational exposure experts. An extensive description of the development and final version of the COVID-19-JEM is described elsewhere (Oude Hengel *et al.*, 2022). The final COVID-19-JEM consisted of specific risk scores per job title for each country. For the current study, we validated the COVID-19-JEM risk scores that were assigned for the Netherlands. The validation was performed with both a direct and an indirect approach, as described in detail below.

Survey data

For the validation of the COVID-19-JEM, the 'Netherlands Working Conditions Survey-COVID-19 (NWCS-COVID-19)' cohort study was used. The NWCS-COVID-19 study is a follow-up study of the annual Netherlands Working Conditions Survey (NWCS) in 2019 (Hooftman *et al.*, 2020) and was conducted with one measurement before the pandemic and four measurements during the pandemic (Oude Hengel *et al.*, 2021). The study population of the NWCS 2019 was selected by Statistics Netherland and consists of a representative group of employees between the age of 15 and 74 years. For the NWCS-COVID-19 cohort study, a group of participants that granted permission in 2019 was approached again to participate in the current study. Participants received follow-up online questionnaires on a variety of topics, including demographics and health and working conditions such as COVID-19-related questions on exposure, measures, and infection. Data on occupation were self-reported at baseline and coded by Statistics Netherlands into 4-digit job codes according to the ISCO-08. The NWCS-COVID-19 study has been approved by TNO's review board (TNO-2020-057 and TNO-2021-101), which is an internal ethics committee who declared that the Medical Research Involving Human Subjects Act does not apply.

Participants of the NWCS-COVID-19 were matched with corresponding COVID-19-JEM risk scores, based on their ISCO-08 code. The COVID-19-JEM experts divided some of the healthcare-related occupations into subsectors and provided each of them with a risk score.

Table 1. Definition of the risk scores provided by the COVID-19-JEM and provided by the NWCS-COVID-19 for all dimensions.[a]

	Dimensions (description)	Risk score	Definitions by the COVID-19-JEM	Definitions by the NWCS-COVID-19
Transmission risk	Number of contacts (Number of contacts in close vicinity)	No risk	Homeworkers, not working with others	Homeworkers, or no contacts
		Low risk	Less than 10 contacts per day	Number of contacts <10
		Intermediate risk	Between 10 and 30 contacts per day	Number of contacts ≥10 and ≤30
		High risk	More than 30 contacts per day	Number of contacts >30
	Nature of contacts (Co-workers, general public, or patients with COVID-19)	No risk	Homeworkers, not working with others	Homeworkers or no contacts
		Low risk	Working in work spaces with co-workers only	Only in contact with co-workers and supervisors
		Intermediate risk	Working in work spaces with general public	In contact with general public
		High risk	Working in work spaces with regular contact with suspected or diagnosed COVID-19 patients	In contact with patients who (are suspected to) have COVID-19
	Contaminated workspaces (The risk through contaminated work surfaces and materials)	No risk	Homeworkers, not working with others	Homeworkers or no contaminated workspaces
		Low risk	People frequently sharing materials/surfaces with co-workers (>10 times a day)	Only sharing potentially contaminated workspace with co-workers/supervisors
		Intermediate risk	People sometimes sharing materials/surfaces with patients (<10 times a day)	—[a]
		High risk	People frequently sharing materials/surfaces with patients (>10 times a day)	Sharing potentially contaminated workspace with customers (patients/students/passenger)
	Work location (Indoors or outdoors)	No risk	Homeworkers	Homeworkers
		Low risk	Working mostly outside	Working outside
		Intermediate risk	Working partly inside (1–4 h day^{-1})	—
		High risk	Working mostly inside (>4 h day^{-1})	Working inside
Mitigation factors	Social distance (The possibility to keep at least 1 m of social distance)	No risk	Homeworkers, not working with others	Homeworkers
		Low risk	Always able to maintain social distancing	Always able to maintain distance with contacts
		Intermediate risk	Not always able to maintain social distancing	Sometimes or often able to maintain distance with contacts
		High risk	Never able to maintain social distancing	Never able to maintain distance with either colleagues/supervisors or with customers or both.
	Face covering (The need and usage of face covering)	No risk	Homeworkers, not working with others	Homeworkers
		Low risk	Wearing face covering at worksite	Always wearing personal protective equipment
		Intermediate risk	Wearing face covering during specific activities, but not in between (e.g. talking to co-worker)	Sometimes or often wearing personal protective equipment
		High risk	Involved in activities in proximity of others which cannot be done when wearing face covering (e.g. sports, singing)	Never wearing personal protective equipment

Table 1. Continued

	Dimensions (description)	Risk score	Definitions by the COVID-19-JEM	Definitions by the NWCS-COVID-19
Precarious work	Income insecurity (Proportion of income insecurity due to the pandemic)	No risk	<1%	—[b]
		Low risk	1–10%	—
		Intermediate risk	11–25%	—
		High risk	>25%	—
	Migrants (Proportion of migrants)	No risk	<1%	—
		Low risk	1–10%	—
		Intermediate risk	11–25%	—
		High risk	>25%	—

[a]Data provided by the NWCS-COVID-19 were not sufficient to be able to distinguish between intermediate and high risk for dimensions 3 (contaminated workspaces) and 4 (location).
[b]No data were available on the dimensions 'income insecurity' and 'the proportion of migrants' from the NWCS-COVID-19 study.

Since no data were available on these subsectors, the rounded mean of their COVID-19-JEM risk scores was attributed to their overarching ISCO-08 code.

Direct validation

Study population

Measurement 2 (n = 10 115; 38% response) was used for the direct validation and was conducted in July 2020—after the first wave of the pandemic—in a period with less infections and fewer governmental measurements. We excluded participants with unknown or no paid employment and those without work activities (n = 1405), participants of whom the ISCO-08 job code was unknown (n = 844), and participants of whom it was unclear whether they worked on site or at home (n = 39). Only occupations with at least 10 representatives were included, resulting in an exclusion of another 638 participants. There remained 7690 participants for the direct validation.

Risk scores constructed within Measurement 2 of the NWCS-COVID-19

Self-reported data of Measurement 2 were used to construct variables that could be compared with the first six dimensions of the COVID-19-JEM: number of contacts, nature of contacts, contaminated workspaces, location, social distance, and face covering. The definitions of the risk scores in the COVID-19-JEM and based on the NWCS-COVID-19 data are described in Table 1. The dimensions on precarious work could not be validated since the cohort study provided no accurate information on 'income insecurity' and 'migrant workers'.

Homeworkers and people without contacts at work

The first step was to distinguish between people who worked from home and those working on site, based on questions on the work location and contract hours per week. Answer categories on the work location question were: 'mainly at my own residential address', 'both at my own address and at the employer's address', 'mostly at my employer's address or external locations of my employer', and 'at external location not of employer'. Homeworkers were defined as people who worked mainly at their own address or worked both at home and the employer's address but had an employment contract of <8 h week^{-1}. Site-based workers were participants who worked mostly at their employer's location or external locations or those who worked both at home and at the employer's address but with an employment contract of ≥8 h week^{-1}.

The number of contacts was established based on three questions: (i) the number of patients in health or social care worked with on a typical working day, (ii) the number of people from the general public (or students, passengers, etc.) worked with during a typical working day, and (iii) the number of colleagues worked with on a typical working day.

People working from home or without having contact with others during work received a risk score of zero—no risk—for all six dimensions of the COVID-19-JEM.

Transmission risk

The dimension 'number of contacts' was based on the sum of the three questions on number of contacts to assign four risk categories (i.e. no risk = homeworkers and those with no contacts; low risk = workers with 1–10 contacts per day, intermediate risk = workers with 10–30 contacts per day; high risk = worker with more than 30 contacts per day).

For the dimension 'nature of contacts', we used three dichotomous questions on different types of contacts: (i) contact with or care for patients in health or social care, (ii) contact with the general public, and (iii) direct

physical contact with colleagues. Regarding care for patients, an additional question was asked if these patients (are suspected to) have COVID-19. Participants who answered 'yes' or 'maybe/I do not know' on that latter question were classified to be at high risk. Workers in contact with the general public and non-COVID-19 patients were assigned an intermediate risk, workers only in contact with colleagues were assigned a low risk.

'Contaminated workspaces' refers to sharing surfaces (e.g. desks, door handle) or and materials (e.g. tools, paperwork) that might be infected with SARS-CoV-2. We used four dichotomous questions on touching surfaces or handling items at work that are also handled by the general public or colleagues. Participants who do not share surfaces or items with others were attributed no risk and participants that were only indirectly in contact with co-workers and supervisors were assessed at low risk. Since frequency of contact with potentially contaminated workspaces was not measured, we were not able to distinguish between intermediate and high risk following the rules of the COVID-19-JEM. Therefore, all workers in contact with surfaces or materials that are shared with the general public were assessed at high risk.

The dimension 'location' was investigated with a dichotomous question on working mainly outside or inside. The NWCS-COVID-19 only asked if someone worked mainly outside or inside, while the COVID-19-JEM distinguished three categories [mostly working outside, working partly inside (1–4 h day^{-1}), working mostly inside (>4 h day^{-1})]. Based on this, participants who worked mainly outside were attributed low risk and participants who worked mainly inside were attributed high risk.

Mitigation factors

The dimension 'social distance' was based on two questions on the ability to keep distance from colleagues and/or the general public. These questions were answered on a four-point scale: never, sometimes, often, and always. Participants who were always able to maintain distance were attributed low risk, participants who were sometimes or often able to maintain distance were attributed intermediate risk, and those who were never able to maintain distance were attributed high risk.

We investigated the dimension 'face covering' with one question on wearing personal protective equipment (PPE) at the workplace for protection against COVID-19, with four answer categories ranging from never to always. Participants who were always wearing PPE were attributed low risk. Those who sometimes or often wore PPE were attributed intermediate risk, and those who never wore PPE were attributed high risk. This definition considerably differs from the definition used in COVID-19-JEM's which distinguishes between different types of activities and the likelihood of using face covering (e.g. surgical masks, face shields, respiratory protective equipment) during their performance, while the NWCS-COVID-19 considers the frequency of wearing 'personal protective equipment'. Moreover, PPE is aimed at individual protection, while face covering is used to prevent the spreading of the virus.

Statistical analyses

For the descriptive statistics, continuous variables were summarized with a mean and standard deviation, and categorical variables with frequencies and relative percentages. The statistical analysis was done in two parts: direct validation to compare the risk scores defined by the experts in the JEM and by respondents in the questionnaire, and indirect validation to predict the self-reported occurrence of COVID by the JEM.

The direct validation investigated whether COVID-19-JEM risk scores were in accordance with the risk scores based on the self-reported data in Measurement 2 by the NWCS-COVID-19 questionnaire. The first step was to group the participants of the NWCS-COVID-19 by occupation. Subsequently, per occupation the answers to the questions on number of contacts, nature of contacts, contaminated workspaces, location, social distance, and face covering were summarized into the median risk score per question for each occupation. This procedure allowed a direct comparison of the six risk scores per job, as collated from the NWCS-COVID-19 questionnaire, with the six risk scores of the COVID-19-JEM for each job. Three performance indicators were used to evaluate the reliability and agreement between the risk scores of the COVID-19-JEM and the NWCS-COVID-19, namely (i) agreement score, (ii) weighted kappa, and (iii) variance. The percentage agreement was considered by exact agreement, and agreement with a maximum difference of 1 level of risk between the COVID-19-JEM and the NWCS-COVID-19. An agreement score can range from 0 (0%) to 1 (100%) where the latter means total agreement between both methods. The weighted kappa coefficient also measures agreement but takes into account that agreement may occur by chance. Weighted kappa values were classified according to Cohen, as poor (<0.20), fair (0.21–0.40), moderate (0.41–0.60), good (0.61–0.80), and excellent (0.81–1) agreement (Cohen, 1968). Total variances within the NWCS-COVID-19-based risk scores were calculated, as well as how much of this total variance was explained by variance within occupations (between respondents with the same occupation) and between occupations.

All statistical analyses were conducted with the program 'Rstudio version 4.0.3'.

Indirect validation

Study population

Measurement 4 (n = 8553; 33% response) was used for the indirect validation, and was conducted in March 2021, when governmental measures were further restricted as non-essential shops were only open upon an appointment and with a curfew at 9.00 p.m., but also when vaccination had started among healthcare workers and vulnerable groups. We excluded participants with unknown or no paid employment and those without work activities (n = 1282), participants of whom the ISCO-08 job code was unknown (n = 558), participants where it was unclear whether they worked on site or at home (n = 39). We also excluded 277 participants with missing information about having had COVID-19 or not. This resulted in the inclusion of 6794 participants for the indirect validation.

COVID-19 in general and at work

In Measurement 4, participants were asked 'Do you think you have been infected with COVID-19 in the past 12 months?'.

Answer categories were: 'yes, confirmed by a positive test', 'probably (but not confirmed by a positive test)', 'no, confirmed by a negative test', 'probably not (but not confirmed by a test)', and 'unknown'. The variable was dichotomized into 'have (probably) had COVID-19 (either confirmed by a test or not)' and '(probably) not have had COVID-19'.

Participants who probably have had COVID-19, were asked additionally about the most likely place of infection with three possible answers—i.e. 'unknown', 'probably at work', and 'probably in private setting'—and were dichotomized into 'infected at work' and 'not infected at work' (containing the answer categories unknown or probably in private setting).

Other variables

Age was included as continuous variable and gender as dichotomous variable. Based on the ISCO-08 codes or sector, type of occupation was further categorized into 'blue-collar workers' (manual labour), 'white-collar workers' (non-manual labour), and 'contact profession' (labour requiring physical contact with the general public, students, clients, and patients; Hooftman *et al.*, 2020).

Statistical analyses

For the indirect validation step, logistic regression analyses were used to determine the associations between the COVID-19-JEM risk scores of all eight dimensions, assigned to each respondent based on his or her job, and self-reported COVID-19 in the past 12 months (in general and infected at work). Due to high collinearity, analyses were performed for each dimension separately. In the first step, it was estimated if a higher-risk score on a COVID-19-JEM dimension resulted in a higher odd to be infected with COVID-19. In the second step, the analyses were repeated for infections at the COVID-19 that thought to be occurred at the worksite, compared with those not being infected at the worksite. Odds ratios (ORs) and 95% confidence intervals (95% CIs) were presented. The associations were corrected for age and gender.

Results

In Measurement 2 (July 2020), the 7690 participants reported occupations that were coded to 161 ISCO-08 codes. A small majority of the study population (55%) was female, the average age was 46.7 ± 12.2 and the majority of the participants had a higher education degree (59%) (Table 2). Most of the workers worked (partly) on site (63%) and had a contact-based job (61%). In Measurement 4 (March 2021), the 6794 participants reported jobs that were coded to 330 ISCO-08 codes. These are more than in Measurement 2 since ISCO-08 codes with fewer than 10 representatives were excluded for direct validation. At Measurement 4, a small majority was female (55%) and highly educated (59%) and with an average age of 47.4 ± 11.6. Most participants worked (partly) on site (61%) and had a contact-based job (58%). Of all participants in Measurement 4, 20% have (probably) had COVID-19, of whom almost a third believed they were infected at work (29%).

Regarding the direct validation, the agreement scores between the COVID-19-JEM assigned risk scores based on expert judgement and the risk scores derived from the NWCS-COVID-19 data are presented in Table 3. The frequencies of the COVID-19-JEM risk scores and NWCS-COVID-19-based risk scores can be found in Supplementary File 1 (available at *Annals of Work Exposures and Health* online). The dimensions 'nature of contacts', 'contaminated workspaces', and 'work location' showed highest exact agreement (>70%), while the dimension 'face covering' showed the lowest exact agreement (42%). Percentage agreement with a maximum difference of 1 risk level between the COVID-19-JEM risk scores and the NWCS-COVID-19-based risk scores was high (>90%) for the dimensions 'number of contacts', 'nature of contacts', and 'social distance', but lower for 'work location' (79%) and 'face covering' (80%). Weighted kappa's were good (ranging between 0.70 and 0.74) for 'number of contacts', 'nature of contacts', 'contaminated workspaces', and 'social distance', but moderate for 'work location' (κ = 0.56) and 'face

Table 2. Characteristics of the study population of the NWCS-COVID-19, Waves 2 and 4.

	Wave 2 July 2020	Wave 4 March 2021
	7690	6794
Gender (*n*, % female)	4204 (54.7%)	3707 (54.6%)
Age (mean ± SD)	46.7 ± 12.2	47.4 ± 11.6
Education[a] (*n*, %)		
Low	601 (7.8%)	482 (7.1%)
Intermediate	2520 (32.9%)	2290 (33.8%)
High	4541 (59.3%)	4001 (59.1%)
Work location[b] (*n*, %)		
At home	2865 (37.3%)	2672 (39.3%)
On site	4825 (62.7%)	4119 (60.7%)
Occupation type[c] (*n*, %)		
Blue-collar profession	648 (8.9%)	690 (10.7%)
White-collar profession	2170 (29.8%)	2034 (31.7%)
Contact profession	4466 (61.3%)	3702 (57.6%)
COVID-19 (*n*, %)		
Infected in past 12 months	—[d]	1336 (19.7%)
Place of infection (*n*, %)		
Probably at work	—	388 (29.1%)

[a]Number of missings in education: *n* = 28 for Wave 2 and *n* = 21 for Wave 4.
[b]Number of missings in work location: *n* = 3 for Wave 4.
[c]Number of missings in occupation type: *n* = 406 for Wave 2 and *n* = 368 for Wave 4.
[d]Not measured during Wave 2.

covering' (κ = 0.41). Risk scores based on self-reported data showed a larger variance between workers within the same occupation than between occupations (Table 3). The proportion of variance explained by occupation ranged from 17% for face covering to 37% for nature of contacts.

Regarding indirect validation, higher COVID-19-JEM risk scores on the four dimensions of transmission risk and two mitigation measures were associated with a higher risk of having had COVID-19 compared with the reference score of 'no risk' (Table 4). Except for the intermediate-risk scores for the dimensions 'work location' and 'face covering', these associations were statistically significant for workers in an occupation with an intermediate or high risk, with ORs ranging from 1.28 (for intermediate risk for dimension 'number of contacts') to 1.80 (for high risk for dimension 'type of contacts'). Occupations with higher proportions of migrant workers showed a decreased risk for having had COVID-19 [intermediate-risk OR: 0.72 (95% CI 0.56–0.92) and high-risk OR: 0.62 (95% CI 0.38–0.96)]. When only considering having had COVID-19 where the infection was reported by the participant to have occurred at the workplace, associations were significant between all transmission and mitigation factors (except low risk at the dimension location), ranging from 2.33 (95% CI 1.55–3.51) to 11.62 (95% CI 7.55–17.92) compared with the reference score 'no risk' (Table 4). Regarding precarious work, no statistically significant associations were observed for intermediate and high risks for the dimension income insecurity and for low risk for the dimension migrant workers. For the migrant worker dimension, the risk of infection at work was reduced in occupations with higher proportion of migrant workers.

Discussion

For the dimensions of transmission risk and mitigation measures, the COVID-19-JEM showed moderate to good agreement scores and weighted kappa's between the COVID-19-JEM risk scores based on expert judgement and the risk scores based on self-reported data from a Dutch working population during the COVID-19 pandemic.

The weighted kappa values between the COVID-19-JEM risk scores and the NWCS-COVID-19-based risk scores were good for the 'number of contacts', 'nature of contacts', 'contaminated workspaces', and 'social distancing', and moderate for 'work location' and 'face covering'. It should be noted that only a small percentage of participants worked outside: 3.3% based on the NWCS-COVID-19 data and 1.3% based on the COVID-19-JEM estimations. When prevalence of an outcome is rare, it becomes harder to obtain a kappa above chance levels (Burn and Weir, 2011). The dimension 'face covering' showed the lowest agreement which may be explained by differences in the definition used for this dimension between the COVID-19-JEM (considering face masks, preventing spread) and the NWCS-COVID-19 (considering PPE, for individual protection). Moreover, the risk scores themselves were differently defined; the COVID-19-JEM determined the risk scores based on the type of situations in which face covering was worn (e.g. only during specific activities), whereas NWCS-COVID-19 only investigated the frequency of wearing PPE against COVID-19. Therefore, the self-reported data were far from ideal to compare with the risk scores assigned by the COVID-19-JEM. An additional explanation is that face masks were not yet obligatory in the Netherlands during Measurement 2 (July 2020). The Dutch JEM experts considered the measures in December 2020 as guideline for the dimension 'face covering', which is when the use of face masks became

Table 3. Percentage agreement and weighted kappa between COVID-19-JEM risk scores and NWCS-COVID-19 risk scores and the variance of the NWCS-COVID-19 risk scores for occupations with ≥10 representatives.

	Agreement score				Explained variance of total variance within the NWCS-COVID-19 risk scores (%)	
	Exact	Difference of 1-level in risk score	Weighted kappa (95% CI)	Total variance	Within occupations	Between occupations
COVID-19-JEM dimensions						
Number of contacts	53.4%	94.4%	0.70 (0.68–0.73)	1.25	70.3%	29.7%
Nature of contacts	70.2%	93.8%	0.73 (0.73–0.73)	1.04	63.3%	36.7%
Contaminated workspaces[a]	73.3%	82.6%	0.74 (0.67–0.82)	1.90	69.0%	31.0%
Work location[a]	72.7%	78.9%	0.56 (0.49–0.64)	2.08	76.9%	23.1%
Social distance	55.9%	96.9%	0.72 (0.72–0.72)	1.17	70.7%	29.3%
Face covering	42.2%	80.1%	0.41 (0.32–0.52)	1.72	82.9%	17.1%

[a]The NWCS-COVID-19 could not distinguish between intermediate and high risk in the dimensions 'contaminated workspaces' and 'work location'. Therefore, the COVID-19-JEM was simplified by putting intermediate and high risk together in one category for these two dimensions in the internal validation step.

Table 4. Logistic regression analyses of the association between the COVID-19-JEM attributed risk scores and having been infected with SARS-CoV-2 in general and at work in univariate models corrected for age and gender.

	Dimension	Infection with COVID-19		Infected with COVID-9 at work	
		Risk score	OR[a] (95% CI)	Risk score	OR (95% CI)
Transmission risk	Number of contacts	Low	1.08 (0.91; 1.29)	Low	**2.82 (1.93; 4.13)**
		Intermediate	**1.29 (1.09; 1.53)**	Intermediate	**4.72 (3.35; 6.73)**
		High	**1.42 (1.20; 1.67)**	High	**6.05 (4.36; 8.54)**
	Type of contacts	Low	1.07 (0.89; 1.29)	Low	**2.33 (1.55; 3.51)**
		Intermediate	**1.31 (1.14; 1.51)**	Intermediate	**5.15 (3.76; 7.18)**
		High	**1.80 (1.36; 2.38)**	High	**11.62 (7.55; 17.92)**
	Contaminated workspaces	Low	**1.20 (1.01; 1.42)**	Low	**2.62 (1.85; 3.74)**
		Intermediate	**1.32 (1.13; 1.56)**	Intermediate	**4.50 (3.29; 6.22)**
		High	**1.46 (1.23; 1.74)**	High	**5.15 (3.73; 7.18)**
	Location	Low	1.09 (0.67; 1.70)	Low	0.88 (0.14; 2.88)
		Intermediate	1.17 (0.86; 1.57)	Intermediate	**3.69 (2.11; 6.22)**
		High	**1.28 (1.12; 1.46)**	High	**4.76 (3.51; 6.59)**
Mitigation measures	Social distance	Low	1.08 (0.91; 1.28)	Low	**2.76 (1.91; 4.04)**
		Intermediate	**1.32 (1.14; 1.53)**	Intermediate	**5.03 (3.65; 7.05)**
		High	**1.54 (1.25; 1.89)**	High	**7.54 (5.19; 11.06)**
	Face covering[b]	Low	**1.28 (1.12; 1.47)**	Low	**4.84 (3.56; 6.73)**
		Intermediate	1.20 (0.99; 1.45)	Intermediate	**3.55 (2.41; 5.24)**
Precarious work	Income insecurity	Low	1.05 (0.86; 1.29)	Low	**1.44 (1.04; 1.95)**
		Intermediate	1.06 (0.68; 1.59)	Intermediate	0.81 (0.32; 1.71)
		High	0.91 (0.67; 1.22)	High	0.94 (0.54; 1.53)
	Migrants	Low	0.96 (0.78; 1.20)	Low	0.81 (0.58; 1.15)
		Intermediate	**0.72 (0.56; 0.92)**	**Intermediate**	**0.55 (0.36; 0.83)**
		High	**0.62 (0.38; 0.96)**	**High**	**0.43 (0.16; 0.95)**

Bold number represents the significant association between dimension and infection with COVID-19.

[a]Odds ratio with risk level 'no risk' as reference group.

[b]Since only four participants had a high risk for the dimension 'face covering', these scores were excluded from the analyses.

obligatory in the Netherlands (Rijksoverheid, 2020a). Hence, the NWCS-COVID-19 data may not be adequate to validate the dimension face covering.

The self-reported risk scores showed a larger variance between workers within the same occupation than between occupations. These data suggest that there are substantial individual differences, which cannot be accounted for by a JEM. Possibly, organizational measures or compliance of individuals to the COVID-19 measures differ from one another and could explain the high variance within occupations. However, also different interpretation and recall-bias might explain differences between individuals within the same occupation. Moreover, application of a group-based approach such as a JEM is less likely to result in attenuated risk estimates due to Berkson error, but at the expense of wider CIs.

The results showed indeed that higher-risk scores attributed by the COVID-19-JEM were associated with a higher incidence of self-reported COVID-19, except for the dimensions on precarious work. This observed association is in line with other studies investigating associations between topics similar to the transmission and mitigation factors of the COVID-19-JEM and COVID-19 or other respiratory diseases (Jefferson *et al.*, 2008; Kampf *et al.*, 2020; Vokó and Pitter, 2020; Backer *et al.*, 2021; Beale *et al.*, 2022; Li and Tang, 2021). These findings indicate that the COVID-19-JEM is a useful tool for assessing the risk of having had COVID-19 when data on exposure at individual level are lacking. The only exception is precarious work, and specifically the dimension 'migrant workers' since the risk estimates were in the opposite direction of what was expected. These findings are contrary with literature, because previous studies described that migrant workers are especially vulnerable for catching COVID-19 at the worksite due to poor working and living conditions (Burton-Jeangros *et al.*, 2020; Guadagno, 2020; Liem *et al.*, 2020). One explanation could be that the dimension 'migrant workers' in the COVID-19-JEM was defined as the proportion of first-generation migrants within each occupation, without making a distinction between relatively highly paid and highly educated migrants and those who are low paid and less educated. Based on the recent literature, the COVID-19-JEM definition of migrant workers may be too broad to function as one of the 'precarious work' dimensions (Kreshpaj *et al.*, 2020). Therefore, it is recommended to consider redefining the dimension 'migrants' into 'labour migrants'. Another explanation might be that a selective group of migrant workers participated in the NWCS-COVID-19 data, since questionnaires were in Dutch and potential participants had to live in the Netherlands to be approached by Statistics Netherlands. Finally, labour migrants are possibly tested less frequently, due to their lack of registration, their dependency on their employers and a lack of information in their language (Rijksoverheid, 2020b). Consequently, the actual number of infections in this group might be much higher than the data suggest.

Moreover, it is important to realize that the performance of the COVID-19-JEM can vary over time together with the changing governmental measures between and within governments. The COVID-19-JEM was based on the situation in which working from home was encouraged when possible, wearing facemasks was obligatory, but without closure of any of the sectors. This is virtually the situation when Measurement 2 was conducted, a period with fewer COVID-19 cases and a relaxation of government measures in the Netherlands (Rijksoverheid, 2020c). This meant, for example, that some workers (like workers in restaurants and cafés) worked (inside) again, public transport was used more frequently, and healthcare professions were less excessively exposed to COVID-19 patients. However, when measures become stricter, the risk scores of different occupations might change and become higher or lower depending on the sector. Moreover, we need to be aware that only one third of the infected study population thought they were infected at the worksite. Even though this is a substantial proportion which is in line with what others have reported (Lan *et al.*, 2020; Beale *et al.*, 2022), it also means that it is not the only key setting of being infected with SARS-CoV-2.

A strength of the study was that information about the transmission risks and mitigation factors was gathered during the pandemic by the NWCS-COVID-19, which allowed us to conduct a direct validation. However, it should be noticed that the questions in the NWCS-COVID-19 did not directly match the COVID-19-JEM dimensions, which resulted in the need to merge some risk groups and therewith might have resulted in a less precise direct validation. Another strength of the study was that we were able to distinguish between COVID-19 for which the infection was believed to have occurred at the worksite and those that were believed to have occurred elsewhere. The results showed that associations were higher for all dimensions of transmission risk and mitigation measures when participants believed they were infected at work. This is not surprising since the JEM was developed to estimate exposure risk at work. However, the results may suffer from reporting bias, as those working in occupations which have widely been reported to be associated to risk of infection (e.g. healthcare workers, public transport workers) may overestimate the likelihood that the infection occurred in the

workplace rather than elsewhere. On the other hand, workplace infections may be underreported in other occupations (e.g. construction workers), which could lead to underreporting of infection with SARS-CoV-2 at work. Moreover, classifying people that did not know their place of infection as '(probably) not infected at the worksite' could have led to classification bias.

A limitation of the current study is that we were not able to validate the COVID-19-JEM across different time periods with different governmental measures. It could be hypothesized that the validation might differ across time periods, because fewer or more workers were at the worksites. Another limitation of the current study is that data on the dimensions on precarious work were not included in the NWCS-COVID-19, and it is recommended to validate this also in future studies. Even though a large representation of different job titles was included in the NWCS-COVID-19, no data were available on self-employed workers and both lower educated people and migrants were probably underrepresented. Furthermore, data on COVID-19 were measured subjectively rather than objectively. While these limitations might have led to some biased results, it is unlikely the main conclusions would have changed.

Based on our findings, several recommendations are formulated for future research. First, validation of the COVID-19-JEM with objective test data is needed, for example using a test negative design to address any differences between occupations in the likelihood of getting tested (Vandenbroucke and Pearce, 2019). Moreover, future research should evaluate the performance of the COVID-19-JEM during different time periods, as governmental measures differ over time. After validation of the COVID-19-JEM, this JEM may be an important tool to assess the role of the worksite for the probability of exposure to SARS-CoV-2. Thereby, the COVID-19-JEM can also be of importance for OSH practitioners, employers, and employees to assess the relevant exposure risks aiming to quickly identify and implement appropriate job-task-specific controls (Descatha *et al.*, 2022).

Conclusion

The direct validation showed moderate to good agreement between the COVID-19-JEM and self-reported exposure risk scores, except for the dimension on face covering. Apart from the precarious work dimensions, the dimensions of the COVID-19-JEM were associated with having had COVID-19, especially when the infection with SARS-CoV-2 occurred at work. These dimensions will be useful factors in predicting which groups in the Netherlands are likely to be exposed to SARS-CoV-2 at work. The next step will be to validate the COVID-19-JEM in the Netherlands and abroad by using objective data.

Supplementary Data

Supplementary data are available at *Annals of Work Exposures and Health* online.

Authors' contribution

S.v.d.F., S.P., and K.M.O.H. contributed to the analysis of the study. S.v.d.F. drafted the initial manuscript. All authors were involved in the conception of this study, made critical revisions, and approved the final manuscript.

Conflict of interest

The authors declare no conflict of interest.

Data availability

Data are stored at TNO, Unit Healthy Living in the Netherlands. Data are available upon reasonable request by the last author.

References

Backer JA, Mollema L, Vos ER *et al.* (2021) The impact of physical distancing measures against COVID-19 transmission on contacts and mixing patterns in the Netherlands: repeated cross-sectional surveys in 2016/2017, April 2020 and June 2020. *Euro Surveill*; **26**: 2000994.

Baker MG, Peckham TK, Seixas NS. (2020) Estimating the burden of United States workers exposed to infection or disease: a key factor in containing risk of COVID-19 infection. *PLoS One*; **15**: 1–8.

Beale S, Patel P, Rodger A *et al.* (2022) Occupation, work-related contact, and SARS-CoV-2 anti-nucleocapsid serological status: findings from the virus watch prospective cohort study, *Occup Environ Med*; **0**: 1–7.

Burn CC, Weir AA. (2011) Using prevalence indices to aid interpretation and comparison of agreement ratings between two or more observers. *Vet J*; **188**: 166–70.

Burton-Jeangros C, Duvoisin A, Lachat S *et al.* (2020) The impact of the Covid-19 pandemic and the lockdown on the health and living conditions of undocumented migrants and migrants undergoing legal status regularization. *Front Public Health*; **8**: 1–12.

Cohen J. (1968) Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychol Bull*; **70**: 213–20.

de Gier B, Lima PDOB, van Gaalen RD *et al.* (2020) Occupation- and age-associated risk of SARS-CoV-2 test positivity, the Netherlands, June to October 2020. *Euro Surveill*; **25**: 1–8.

Descatha A, Fadel M, Sembajwe G *et al*. (2022) Using the COVID-19 job exposure matrix for essential workplace preparedness. *J Occup Environ Med*; **64**: e39–40.

Dyal JW. (2020) COVID-19 among workers in meat and poultry processing facilities—19 states, April 2020. *MMWR Morb Mortal Wkly Rep*; **69**: 557–61.

Fassani F, Mazza J. (2020) *A vulnerable workforce: migrant workers in the COVID-19 pandemic*. Luxembourg: Publications Office of the European Union. ISBN 978-92-76-18957-2, doi:10.2760/914810, JRC120730

Guadagno L. (2020) Migrants and the COVID-19 pandemic: an initial analysis. Migration Research Series No. 60. International Organization for Migration. Available at https://publications.iom.int/system/files/pdf/mrs-60.pdf

Hooftman WE, Mars GMJ, van Dam LMC *et al*. (2020) *Nationale Enquete arbeidsomstandigheden 2019 [Netherlands Working Conditions Survey 2019]*. Leiden, The Netherlands: TNO.

International Labour Office (ILO). (2012) International Standard Classification of Occupations—Volume 1: structure, group definitions and correspondence tables. Available at https://www.ilo.org/wcmsp5/groups/public/@dgreports/@dcomm/@publ/documents/publication/wcms_172572.pdf

Jefferson T, Foxlee R, Del Mar C *et al*. (2008) Physical interventions to interrupt or reduce the spread of respiratory viruses: systematic review. *BMJ*; **336**: 77–80.

Kampf G, Todt D, Pfaender S *et al*. (2020) Persistence of coronaviruses on inanimate surfaces and their inactivation with biocidal agents. *J Hosp Infect*; **104**: 246–51.

Kreshpaj B, Orellana C, Burström B *et al*. (2020) What is precarious employment? A systematic review of definitions and operationalizations from quantitative and qualitative studies. *Scand J Work Environ Health*; **46**: 235–47.

Kromhout H, Vermeulen R. (2001) Application of job-exposure matrices in studies of the general population: some clues to their performance. *Eur Respir Rev*; **11**: 80–90.

Lan FY, Wei CF, Hsu YT *et al*. (2020) Work-related COVID-19 transmission in six Asian countries/areas: a follow-up study. *PLoS One*; **15**: e0233588.

Li C, Tang H. (2021) Study on ventilation rates and assessment of infection risks of COVID-19 in an outpatient building. *J Build Eng*; **42**: 1–12.

Liem A, Wang C, Wariyanti Y *et al*. (2020) The neglected health of international migrant workers in the COVID-19 epidemic. *Lancet Psychiatry*; **7**: e20.

Magnusson K, Nygard KM, Methi F *et al*. (2021) Occupational risk of COVID-19 in the 1st vs 2nd wave of infection. *Euro Surveill*; **26**: 1–8.

Marinaccio A, Boccuni F, Rondinone BM *et al*. (2020a) Occupational factors in the COVID-19 pandemic in Italy: compensation claims applications support establishing an occupational surveillance system. *Occup Environ Med*; **77**: 818–21.

Marinaccio A, Guerra R, Iavicoli S. (2020b) Work a key determinant in COVID-19 risk. *Lancet Glob Health*; **8**: e1368.

Mutambudzi M, Niedwiedz C, Macdonald EB *et al*. (2021) Occupation and risk of severe COVID-19: prospective cohort study of 120 075 UK Biobank participants. *Occup Environ Med*; **78**: 307–14.

Nafilyan V, Pawelek P, Ayoubkhani D *et al*. (2021) Occupation and COVID-19 mortality in England: a national linked data study of 14.3 million adults. *Occup Environ Med*; 0: 1–9.

Oude Hengel KM, Bouwens L, Zoomer T *et al*. (2021) *Reading guide: the Netherlands Working Conditions Survey—COVID-19: four waves (2019–2020–2021)*. Leiden, The Netherlands: TNO.

Oude Hengel KM, Burdorf A, Pronk A *et al*. (2022) Exposure to a SARS-CoV-2 infection at work: development of an international job exposure matrix (COVID-19-JEM). *Scand J Work Environ Health*; **8**: 61–70.

Peters S. (2020) Although a valuable method in occupational epidemiology, job-exposure-matrices are no magic fix. *Scand J Work Environ Health*; **46**: 231–4.

Rijksoverheid. (2020a) Mondkapje verplicht vanaf 1 december [Facemask mandatory as of the 1st of December]. Dutch. Available at https://www.rijksoverheid.nl/actueel/nieuws/2020/11/19/mondkapje-verplicht-vanaf-1-december#:~:text=Om%20verspreiding%20van%20het%20coronavirus,dan%20een%20mondkapje%20gaan%20dragen

Rijksoverheid. (2020b) Geen tweederangsburgers. Aanbevelingen om misstanden bij arbeidsmigranten in Nederland tegen te gaan [No second-class citizens. Recommendations to counteract abuses among migrant workers in the Netherlands]. Available at https://www.rijksoverheid.nl/documenten/rapporten/2020/10/30/tweede-advies-aanjaagteam-bescherming-arbeidsmigranten

Rijksoverheid. (2020c) Juni 2020: Versoepeling coronamaatregelen en testen voor iedereen. [English title: June 2020: Relaxation of corona measures and tests for everyone]. Dutch. Available at https://www.rijksoverheid.nl/onderwerpen/coronavirus-tijdlijn/juni-2020-versoepeling-coronamaatregelen-en-testen-voor-iedereen

Vandenbroucke JP, Pearce N. (2019) Test-negative designs: differences and commonalities with other case-control studies with "other patient" controls. *Epidemiology*; **30**: 838–44.

Vokó Z, Pitter JG. (2020) The effect of social distance measures on COVID-19 epidemics in Europe: an interrupted time series analysis. *Geroscience*; **42**: 1075–82.

Wiersinga WJ, Rhodes A, Cheng AC *et al*. (2020) Pathophysiology, transmission, diagnosis, and treatment of coronavirus disease 2019 (COVID-19): a review. *JAMA*; **324**: 782–93.

Williams K, Cherrie JW, Dobbie J *et al*. (2022) The development of a COVID-19 control measures risk matrix for occupational hygiene protective measures. *Ann Work Expo Health*; **66**: 269–75.

Wu Z, McGoogan JM. (2020) Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 72 314 cases from the Chinese Center for Disease Control and Prevention. *JAMA*; **323**: 1239–42.

Zhang M. (2021) Estimation of differential occupational risk of COVID-19 by comparing risk factors with case data by occupational group. *Am J Ind Med*; **64**: 39–47.