strong evidence that mild steel welding fumes, in addition to stainless steel, induce chronic inflammation and are immunosuppressive, and this was confirmed in molecular epidemiology studies of workers. We continued studies using metabolomic approaches in a repeated measures design and found welding fume exposure-related changes in blood in pathways related to disturbances in unsaturated fat metabolism, as in the signaling lipids Sphingosine 1-phosphate (S1P) and sphingasine 1-phosphate (SA1P). Global metabolomic profiling also revealed several metabolic changes after welding fume exposure, mainly involved in the lipid pathway [glucocorticoid class (cortisol, corticosterone, and cortisone), acylcarnitine class, and DiHOME species (9,10-DiHOME and 12,13-DiHOME)], amino acid utilization (isoleucine, proline and phenylalanine), and S-(3-hydroxypropyl) mercapturic acid (3-HPMA): compounds are all associated with inflammation.

Conclusion There is strong mechanistic evidence in humans for inflammatory and metabolic changes that promote carcinogenicity of welding fumes in humans.

S-495 INTRODUCTION TO THE KEY CHARACTERISTICS APPROACH TO INTERPRETING MECHANISTIC DATA

[1]Martyn Smith. [1]*University of California, Berkeley School of Public Health, United States*

10.1136/OEM-2021-EPI.449

The key characteristics (KCs) of human carcinogens were recently introduced as the basis of a uniform approach for searching, organizing, and evaluating mechanistic evidence to support cancer hazard identification (PMID:30521319). The KCs comprise the properties of known human carcinogens, including their ability to, be genotoxic; be immunosuppressive; or modulate receptor-mediated effects (PMID: 26600562). Established human carcinogens commonly exhibit one or more of these characteristics, and therefore, data on these characteristics can provide independent evidence of carcinogenicity when human data are lacking (PMID: 29562322). Such data can also help in interpreting the relevance and importance of findings of cancer in animals and in humans. In its 2017 report on 'Using 21st Century Science to Improve Risk-Related Evaluations', the NRC opined that the KCs approach 'avoids a narrow focus on specific pathways and hypotheses and provides for a broad, holistic consideration of the mechanistic evidence.' They further suggested that key characteristics be developed for other endpoints, such as endocrine disruption and reproductive toxicity. These have recently been published (PMID: 31719706; 31322437; 31199676) and KCs for hepato-, immuno-, neuro- and cardiovascular toxicants are in the final stages of development. We have also recently published the findings of two expert committees who described approaches to studying carcinogenicity of chemical mixtures using the KCs (PMID: 33784186) and identified biomarkers that can be used to measure the KCs of carcinogens in humans, animals and cell culture (PMID: 32152214). A uniform approach to applying these biomarkers in occupational studies of different epidemiologic design needs to be developed so that the most relevant biomarkers of each KC are measured in exposed human populations, thereby improving hazard identification and risk assessment.

S-496 RESPIRATORY DISEASE IN THE AGRICOH CONSORTIUM FROM SIX CONTINENTS

[1]I Annesi-Maesano. [1]*INSERM, France*

10.1136/OEM-2021-EPI.450

Objectives Farming is filled with respiratory hazards: pesticide vapors, dusty fields, dangerous hydrogen sulfide accumulations in manure pits and pump sumps, nitrogen dioxide in conventional silos, and many others. Despite the recognition of respiratory hazards, this problem has not been fully investigated at the world level after harmonization of data. We used data from the AGRICOH consortium, a collective of prospective cohorts of agricultural workers, to assess respiratory disease prevalence among adults in 18 cohorts representing over 200,000 (118,520 men, 92,712 women) farmers, farmworkers, and their spouses from six continents.

Methods Cohorts collected data between 1992–2016 and ranged in size from 200 to >128,000 individuals; 44% of participants were female. Farming practices varied from subsistence farming to large scale industrial agriculture. All cohorts provided respiratory outcome information for their cohort based on their study definitions. The majority of outcomes were based on self-report using standard respiratory questionnaires; the greatest variability in assessment methods was associated with chronic obstructive pulmonary disease (COPD).

Results The median prevalence was 18.6% for cough, 13.3% for phlegm, and 15.0% for wheeze and was higher in men than in women, with the greatest difference for phlegm (17% vs. 10%). For asthma, the median prevalence was 7.2% and was higher in women (7.8% vs. 6.5%). The relative proportion of allergic asthma varied among cohorts. In two of eight cohorts for women and two of seven cohorts for men, allergic asthma was more common than non-allergic asthma. The median prevalence of COPD was 4.5%. Men suffered more than women from COPD (5.5% vs. 4.0%).

Conclusion These findings indicate that respiratory outcomes are common among farmers around the world despite the differences in agricultural production. Both allergic and non-allergic asthma exist among farmers.

S-497 COMPARISON OF REPORTED RELATIVE RISKS FOR HEALTH CARE, TRANSPORT, AND FOOD PROCESSING WORKERS

[1]S Rhodes, M Gittins, K Stocking, M Van Tongeren, N Pearce. [1]*The University of Manchester, United Kingdom*

10.1136/OEM-2021-EPI.451

Background Risk of SARS-CV-2 infection has been strongly linked to occupation, with specific occupational sectors such as health care, food production, and transport, particularly affected. To better understand the potential risks by occupational sector we investigated the reported risks of COVID-19 infection and mortality for employees in the three sectors.

Methods We performed a rapid review of observational studies reporting COVID-19 risk for employees in health care, food manufacturing, and transport sectors. All studies published in the peer-review and pre-print literature between March 2020 and June 2021 were considered. The primary outcome measure was COVID-19 infection, with COVID-19 related mortality and hospitalisation considered as secondary measures. We

extracted odds ratios/relative risks (and standard errors) comparing workers in each sector to within study reference category. To avoid complications due to study heterogeneity, we simply produced descriptive forest plots where possible utilising the odds ratios that have been adjusted for demographics.

Results Eight studies were identified from Norway (one), and California (one), and the UK (five). Three studies at time of writing were peer reviewed, and five of the eight were cohort studies. Food production was the least well investigated with limited evidence of increased infection, severe infection or mortality compared to the 'other' groups. Healthcare and transport show wide variation in the odds ratios reported. Public facing roles did indicate greater infection risk, specifically in the first wave.

Conclusions Considerable study heterogeneity is present, particularly with respect to the chosen reference groups, meaning objective comparisons are limited here. However, as would be expected public facing roles, especially in the first wave, did appear to experience increased infection. Further prospective work with subject level data is needed to better understand the occupational risks.

S-498 METHODOLOGICAL ISSUES IN STUDYING OCCUPATIONAL DIFFERENCES IN RISK OF COVID-19

[1]Neil Pearce. [1]*London School of Hygiene and Tropical Medicine, United Kingdom*

10.1136/OEM-2021-EPI.452

It is well-established that certain occupations that have frequent contact with patients (e.g. health care workers) or the public (e.g. bus drivers), may have increased risks of COVID-19. However, estimating these risks involves a number of important methodological problems. Not everyone who has a SARS-CoV-2 infection gets symptoms; not everyone with symptoms gets tested; not all tests yield valid results. Therefore, even just estimating the incidence or prevalence of COVID-19 in particular occupational groups is difficult. Moreover, these selection pressures may be different for different occupational groups (e.g. health care workers may be more likely to be tested than some other occupations). In addition, unlike most other occupational exposures, you can get infected at work, or get the same infection at home, potentially yielding quite severe confounding by lifestyle and living conditions. These methodological issues are not insurmountable, but require careful study design and data analysis. In particular: (i) when not everyone is being tested, then comparisons between and within occupations can be biased, but this bias can be minimised using the test-negative design; and (ii) comparisons between occupations require careful and rigorous adjustment for work-related 'living conditions', i.e. the fact that workers in insecure low-paid jobs may also have a high risk of contracting SARS-CoV-2 infections outside of the workplace.

S-499 THE APPLICATION OF ARTIFICIAL INTELLIGENCE IN THE CODING OF OCCUPATIONAL INFORMATION

[1]Anil Adisesh, Christopher JO Baker. [1]*University of Calgary, Canada*

10.1136/OEM-2021-EPI.453

Many research studies seek to identify the social determinants of health and occupation is an important predictor, both at the level of the individual as well as for populations. Whereas job titles are usually solicited during interviews or by questionnaire, before being able to use this information the responses need to be categorized using a coding system, such as the Canadian National Occupational Classification (NOC).

Manual coding is the usual method, which is a time-consuming and error-prone activity with variable or inconsistent outcomes from teams of coders. In recent work the ACA-NOC algorithm[1] was developed to perform automated coding based on matching job title text with the NOC's job titles and textual descriptions. This algorithm was benchmarked on a small sample manually coded data set with subject matter experts subsequent review of coding discrepancies to facilitate functional improvements to the algorithm. Performance levels achieved illustrated the viability of the approach albeit larger benchmarking data sets were required.

CanPATH[2]has collected data from approximately 330,000 volunteer Canadians, including information about health, lifestyle, occupation, environment and behavior. We report on the further benchmarking and further development of this algorithm in CanPATH funded project using over 60,000 manually coded job titles from the constituent Alberta Tomorrow Project. The algorithm was also applied to over 100,000 un-coded job titles from Atlantic PATH, including the Core questionnaire and occupational history data.

The core outcome of the project identified that auto-coding results are comparable to manual coding in accuracy and superior in speed e.g. 2 years of manual coding (64,000 records) can be auto coded in 72 hours. The algorithm was considered ready for deployment in operational settings: point of care, decision support for manual coders.

Additional insights gained during the project revealed that (i) NOC and ATP data sets have a distribution bias where some NOC categories were over or under-represented and numerous non-standard lexical features were found in job titles and NOC job descriptions, (ii) benchmarking datasets from ATP included coding errors that were corrected by expert coders leading to the creation of gold standard test sets for further algorithm improvement studies, (iii) a study on 17 categories of occupations initially difficult to code, identified some job categories with near 90% coding accuracy.

Automated coding of job titles to the NOC has been shown to be both practicable to good levels of accuracy and shown to significantly accelerate manual coding efforts from years to autocoding in a matter of hours without decrease in accuracy. Autocoding can replace costly, error prone manual labor with accurate point-of-care auto-coding such that patient occupation information during healthcare encounters could now supplement existing administrative data sets in electronic health record systems. This data can be used better to understand the socioeconomic consequences of health conditions, advise patients about returning to work with a health condition, recognizing occupations at risk of disease e.g. as in the COVID-19 pandemic.

REFERENCES

1. Bao H, Baker CJO, Adisesh A. Occupation coding of job titles: iterative development of an automated coding algorithm for the canadian national occupation classification (ACA-NOC). *JMIR Form Res* 2020 Aug 5;4(8):e16422. doi:10.2196/16422
2. CanPath the Canadian Partnership for Tomorrow Project. https://canpath.ca/ accessed: 01.09.2021