Occupational Medicine 2022;**72:**132–141
Advance Access publication on 20 December 2021 https://doi.org/10.1093/occmed/kqab173

Physical work exposure matrix for use in the UK Biobank

E. L. Yanik[1,2], M. J. Stevens[3,4], E. Clare Harris[3,4], K. E. Walker-Bone[3,4], A. M. Dale[5], Y. Ma[5], G. A. Colditz[2] and B. A. Evanoff[5]

[1]Department of Orthopaedic Surgery, Washington University School of Medicine, St. Louis, MO 63110, USA, [2]Department of Surgery, Washington University School of Medicine, St. Louis, MO 63110, USA, [3]MRC Versus Arthritis Centre for Musculoskeletal Health and Work, University of Southampton, Southampton S017 1BJ, UK, [4]MRC Lifecourse Epidemiology Centre, University of Southampton, Southampton, UK, [5]Department of Medicine, Washington University School of Medicine, St. Louis, MO 63110, USA.

Correspondence to: E. L. Yanik, Department of Orthopaedic Surgery, Washington University School of Medicine, 660 S. Euclid Avenue, Campus Box 8233, St. Louis, MO 63110, USA. Tel: (314) 747-3196; fax: 314-273-4949; e-mail: yanike@wustl.edu

Background UK Biobank (UKB) is a large prospective cohort capturing numerous health outcomes, but limited occupational information (job title, self-reported manual work and occupational walking/standing).

Aims To create and evaluate validity of a linkage between UKB and a job exposure matrix for physical work exposures based on the US Occupational Information Network (O*NET) database.

Methods Job titles and UK Standard Occupational Classification (SOC) codes were collected during UKB baseline assessment visits. Using existing crosswalks, UK SOC codes were mapped to US SOC codes allowing linkage to O*NET variables capturing numerous dimensions of physical work. Job titles with the highest O*NET scores were assessed to evaluate face validity. Spearman's correlation coefficients were calculated to compare O*NET scores to self-reported UKB measures.

Results Among 324 114 participants reporting job titles, 323 936 were linked to O*NET. Expected relationships between scores and self-reported measures were observed. For static strength (0–7 scale), the median O*NET score was 1.0 (e.g. audiologists), with a highest score of 4.88 for stone masons and a positive correlation with self-reported heavy manual work (Spearman's coefficient = 0.50). For time spent standing (1–5 scale), the median O*NET score was 2.72 with a highest score of 5 for cooks and a positive correlation with self-reported occupational walking/standing (Spearman's coefficient = 0.56).

Conclusions While most jobs were not physically demanding, a wide range of physical work values were assigned to a diverse set of jobs. This novel linkage of a job exposure matrix to UKB provides a potentially valuable tool for understanding relationships between occupational exposures and disease.

Key words Epidemiologic measurements; job description; occupational exposure; validation study.

Introduction

The UK Biobank (UKB), an open-access prospective cohort of >500 000 people, provides a wealth of opportunities to investigate risk factors for numerous health outcomes [1]. Researchers have made extensive use of UKB, particularly for genetic studies [2], but limited occupational data were collected for all participants (including job title at baseline study visit and self-reported frequency of heavy manual work and walking/standing during work) [3].

Use of job exposure matrices (JEMs) is common in occupational exposure studies, as they provide several advantages over other exposure measurement methods [4,5]. JEMs are less susceptible to information bias than individually self-reported measures [6,7]. Measurement is more efficient and less expensive than other objective methods such as direct observation, making it feasible in large cohort studies [8,9]. And JEMs allow linkage of new information to an epidemiologic study without requiring time or effort from study participants. UKB presents an opportunity to employ a JEM to estimate individual occupational exposures based on job titles. We constructed a novel linkage between jobs reported by UKB participants and US Standard Occupational Classification (SOC) system job codes using several crosswalks. This linkage allowed UKB to be combined with a JEM from the US Occupational Information Network (O*NET). O*NET (https://www.onetcenter.org) is a publicly available data set describing physical and mental demands of >900 occupations defined by the US SOC system.

Key learning points

What is already known about this subject:

- The UK Biobank is a rich resource for epidemiologic investigations, but occupational exposure information is currently limited.

What this study adds:

- We constructed a novel linkage between a job exposure matrix based on the US Occupational Information Network and jobs reported by UK Biobank participants.
- Measurements of occupational demands from this linkage had good face validity and good agreement with two self-reported occupational measures in the UK Biobank.

What impact this may have on practice or policy:

- The linkage of this job exposure matrix provides information on a richer set of occupational measures than currently available in the UK Biobank that can be used in future epidemiologic investigations.

Numerous studies have effectively used O*NET to estimate occupational exposures. One study identified workplace physical exposures associated with carpal tunnel syndrome, yielding similar results to those found using direct observation, a substantially more costly and time-consuming exposure measurement method [4,10]. Other studies have evaluated associations with falls resulting in fracture, joint damage in arthritic patients and work injury in hospital workers [11–13].

Linkage of the O*NET JEM allows UKB to be combined with unique physical work measurements, providing the opportunity to study the influence of physical work requirements on disease development. In the UK, approximately 480 000 workers suffered from work-related musculoskeletal disorders in 2019–20, with the vast majority attributed to physical tasks such as heavy lifting, material manipulation and repetitive tasks [14].

We evaluated the validity of this linkage by examining the face validity of exposures assigned to different occupations and comparing US JEM-based exposures to self-reported occupational and chronic pain measures from UKB. JEM-based exposures were expected to have positive, but imperfect correlations with similar self-reported occupational measures. Higher levels of physical work were expected to be associated with chronic pain.

Methods

UKB recruited nationwide through invitations mailed to people aged 40–69 years registered with the National Health Service [15]. Between 2006 and 2010, 500 187 people were enrolled [1]. Participants gave informed consent and completed detailed questionnaires collecting information on demographics, health behaviours and other characteristics. This included a touchscreen questionnaire that asked about occupational demands (job always/usually/sometimes/rarely 'involves heavy manual or physical work' or 'involves mainly walking or standing') and pain at different anatomic locations [16]. A verbal interview was conducted by trained UKB staff which included recording the participant's current job title and coding it according to the UK SOC System 2000 [17] among participants who were employed or self-employed [16]. The current study was limited to participants reporting a job title. UKB job titles were mapped to the more recent UK SOC 2010 system using an existing crosswalk (9962 job titles mapped; Figure 1) [18]. Some jobs were not listed in this crosswalk and had a UK SOC 2000 code that mapped to multiple UK SOC 2010 codes. In these cases manual review was used to choose the best single match for each job title (470 job titles reviewed by E.L.Y.).

The National Health Service Research Ethics Committee approved UKB. The Washington University Institutional Review Board determined this study to be exempt from oversight.

To characterize physical work requirements associated with each job title we used O*NET (https://www.onetcenter.org). Physical and mental demands are characterized by variables in several domains in O*NET, including Abilities, Work Context and Work Activities. Variables are scored based on job analyst ratings and questionnaires from workers and professionals familiar with each occupation. Variables in the Abilities and Work Activities domains provide scores for occupations on a 0–7 scale with 7 representing the greatest level needed for job performance. Variables in the Work Context domain provide scores on a 1–5 scale with 1 representing no time spent under specified conditions and 5 representing continual/almost continual time spent. Standard US SOC codes use a 6-digit coding system. O*NET uses a more detailed system that adds two decimal places to further distinguish jobs. Our linkage mapped jobs to the standard 6-digit system, so we used the average score for the 96 US SOC codes with >1 corresponding O*NET code.

A linkage between UK SOC 2010 and US SOC 2010 codes was required to use the O*NET JEM to characterize physical occupational exposures in UKB. The International Standard Classification of Occupations

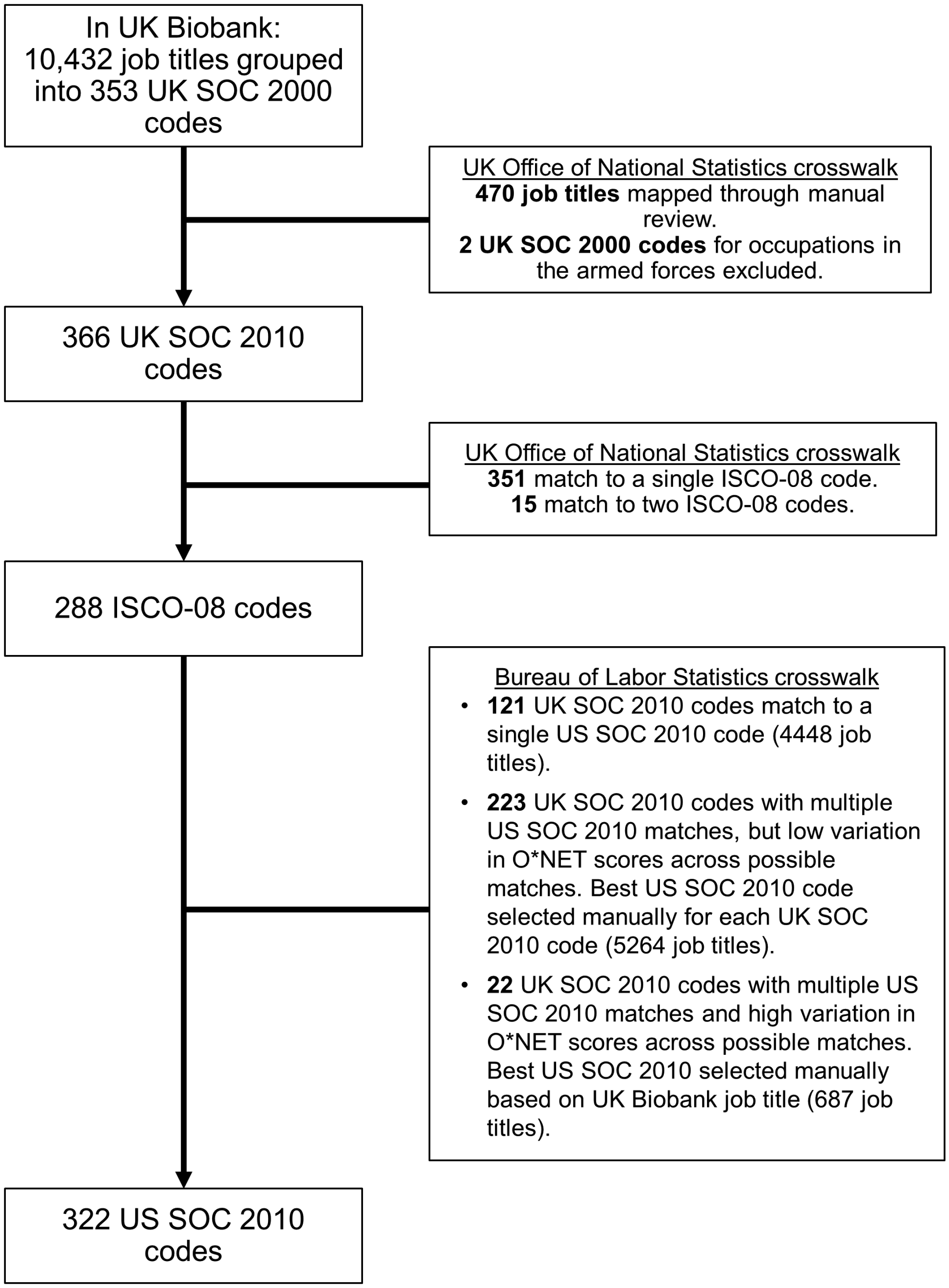

Figure 1. Linkage process between UK Biobank jobs and US SOC 2010 codes.

(ISCO-08) was used as an intermediary classification system, as pre-existing crosswalks were available to link ISCO-08 to both the UK SOC 2010 and US SOC 2010 systems (Figure 1). UK SOC 2010 codes were linked to ISCO-08 using a crosswalk from the UK Office of National Statistics [19]. While the vast majority were one-to-one matches (351 codes), 15 UK SOC 2010 codes matched to two ISCO-08 codes. In these cases both ISCO-08 codes were retained for linking to all possible US SOC matches.

All matched ISCO-08 codes (288 codes total) were mapped to US SOC 2010 codes through a crosswalk from the Bureau of Labor Statistics [20]. There were 121 UK SOC 2010 codes that mapped through ISCO-08 to a single US SOC 2010 code, while the rest mapped to multiple possible codes (up to 43 different codes in the US SOC system). In these cases, two approaches were taken to identify a best match. O*NET scores for static strength converted to a 0–100 scale ([score/total range] * 100) were assigned to each possible US SOC 2010 code as use of static strength is a good indicator of physical work intensity [21]. For each UK SOC 2010 code the distribution of static strength scores for all possible US SOC matches was assessed. When the distribution had a standard deviation >15 and range >20, the original UKB job titles were examined manually through two independent reviews (one by E.L.Y., another by M.J.S./E.C.H.) and matched to the best possible US SOC 2010 code. Disagreements were resolved by a third independent review (B.A.E. resolved 54 matches). When the distribution had a standard deviation ≤15 or range ≤20 the best possible US SOC match was chosen for each UK SOC code, and all job titles within the UK SOC code were assigned to the same US SOC code (selected by E.L.Y.).

After linkage to O*NET scores via 2010 US SOC codes, we examined distributions of O*NET variables across UKB participants, including means, standard deviations, medians, interquartile ranges and total ranges. UKB job titles with the highest and lowest scores for each O*NET variable were used to assess the face validity of the linkage. We evaluated scores for static strength, dynamic strength and stamina in the Abilities domain (Table 1). In the Work Context domain we evaluated scores for spending time using hands to handle/control/feel objects/tools/controls; spending time walking and running; and spending time standing. In the Work Activities domain we evaluated scores for handling/moving objects; performing general physical activities; interacting with computers; and performing administrative activities.

We pre-specified O*NET variables to compare to each self-reported measure in order to evaluate predicted agreement/disagreement. We predicted that self-reported heavy manual/physical work would be positively correlated with O*NET scores for: static strength; dynamic strength; general physical activities; stamina; handling/moving objects; and spending time using hands to handle/control/feel objects/tools/controls. We anticipated that self-reported occupational walking/standing would be positively correlated with: general physical activities; stamina; spending time standing; and spending time walking/running. In contrast, we predicted that both self-reported heavy manual/physical work and self-reported walking/standing would be negatively correlated with O*NET variables for interacting with computers and performing administrative activities.

Several methods were used to compare JEM-based O*NET scores to self-reported occupational measures. Box plots were constructed to examine the distribution of scores for different O*NET variables across categories of self-reported heavy manual/physical work and self-reported walking/standing (divided into categories of 'Never/rarely', 'Sometimes', 'Usually' or 'Always'). Spearman's rank correlation coefficient was calculated with self-reported responses coded as: 0 = Never/rarely, 1 = Sometimes, 2 = Usually and 3 = Always. ANOVA was used to calculate an R^2 value estimating the proportion of the variance between self-report groups that was explained by each O*NET variable. To determine if prevalent chronic pain was associated with higher physical work levels, associations of self-reported and JEM-based measures with prevalent chronic pain were estimated using logistic regression with adjustments for age, sex and Townsend Deprivation Index.

Results

UKB includes >500 000 people, amongst whom 324 114 people were employed at the baseline assessment visit and declared current job titles for coding. Of these, 178 were excluded because they linked to US SOC 2010 codes that do not have O*NET data (they represent occupations with a wide range of characteristics, e.g. 'Transportation Workers, All Others' and armed services members). All others had a job title that could link to a US SOC 2010 code to allow assignment of job exposure values.

In this selected population of 323 936 people, the median age was 54 years (interquartile range = 48–60), 153 944 were male (48%) and most were of white race (94%). The median Townsend Deprivation Index was −0.21 (interquartile range = −0.36, 0.05), indicating participants lived in geographic areas with less deprivation than the general population. Chronic pain lasting ≥3 months was reported by 14% for the neck/shoulder, 16% for the back, 7% for the hip and 15% for the knee. Chronic pain in ≥1 of these sites was reported by 34% of the population. In total, 285 180 people (88%) self-reported how frequently heavy manual/physical work was required in their current job: 65% reported 'Never/rarely' and 7% reported 'Always' (Table S1, available as Supplementary data at *Occupational Medicine* Online). A similar proportion (285 137 people (88%)) reported how frequently their current job involved mainly walking/standing: 35% reported 'Never/rarely', while 19% reported 'Always'.

As measured by the O*NET variables, most people were not employed in physically demanding occupations, but a range of job characteristics were observed. Within the Abilities domain, the median score for static strength was 1.0 with the highest score of 4.88 found amongst stone masons and the lowest score of 0 identified for numerous jobs including secondary school teachers (Table 1). The median scores for dynamic strength and stamina were 0.62 and 0.75, respectively. For both variables the highest scores were 4.62 (amongst dancers) and the lowest scores were 0 (numerous jobs including secondary school teachers and secretaries). The high scores in this domain reflect the upper limit of scores given to all jobs listed in O*NET, not just the jobs observed in UKB [22].

Within the Work Context domain, the median score for spending time using hands to handle/control/feel objects/tools/controls was 2.88 with a highest score amongst hairdressers and beauty therapists and a lowest

Table 1. O*NET characteristics describing physical work in the UK Biobank population

O*NET characteristic[a]	Scale[a]	Range observed	Median (IQR)	UK job titles with the highest scores	UK job titles with the lowest scores
Abilities domain					
Static strength The ability to exert maximum muscle force to lift, push, pull or carry objects.	0–7	0–4.88	1.00 (0–2.38)	Stone mason (*n* = 36), Mason (*n* = 5), Banker mason (*n* = 2), Steel fabricator (*n* = 49)	Secondary school teacher (*n* = 7247), Accountant (*n* = 2188), University lecturer (higher education, university) (*n* = 1967)
Dynamic strength The ability to exert muscle force repeatedly or continuously over time. This involves muscular endurance and resistance to muscle fatigue	0–7	0–4.62	0.62 (0.25–1.28)	Dancer (*n* = 8), Ballet dancer (*n* = 3), Fitness instructor (*n* = 149), Yoga tutor (*n* = 107)	Secondary school teacher (*n* = 7247), Lecturer (higher education, university) (*n* = 1967), University lecturer (*n* = 1597)
Stamina The ability to exert yourself physically over long periods of time without getting winded or out of breath	0–7	0–4.62	0.75 (0–1.88)	Dancer (*n* = 8), Ballet dancer (*n* = 3), Fitness instructor (*n* = 149), Yoga tutor (*n* = 107)	Secretary (*n* = 2968), Accountant (*n* = 2188), University lecturer (higher education, university) (*n* = 1967)
Work Context domain					
Spend time using your hands to handle, control or feel objects, tools or controls How much time in your current job do you spend using your hands to handle, control or feel objects, tools or controls?	1–5	1.3–5	2.88 (2.32–3.52)	Hairdresser (*n* = 855), Beauty therapist (*n* = 147), Barber (*n* = 44), Beautician (*n* = 37)	Consultant psychiatrist (*n* = 112), Psychiatrist (*n* = 57), Psycho-analyst (*n* = 25), Projects manager (*n* = 1263)
Spend time walking and running How much does this job require walking or running?	1–5	1.02–4.73	2.17 (1.89–3.10)	Machinist (textile manufacturing) (*n* = 14), Factory worker (clothing manufacturing) (*n* = 11), Manufacturer (textiles) (*n* = 6)	Travel agent (*n* = 84), Travel consultant (*n* = 74), Agent (travel) (*n* = 34), Travel advisor (*n* = 20)
Spend time standing How much does this job require standing?	1–5	1.49–5	2.72 (2.15–3.55)	Cook (*n* = 468), Cook in charge (*n* = 56), Cook—supervisor (*n* = 48), Cook—general (*n* = 23)	Computer operator (*n* = 136), Secretarial clerk (*n* = 116), Data entry clerk (*n* = 76)
Work Activities domain					
Handling and moving objects Using hands and arms in handling, installing, positioning and moving materials, and manipulating things	0–7	0.34–6.61	2.71 (1.97–3.79)	Stone mason (*n* = 36), Mason (*n* = 5), Banker mason (*n* = 2), Roofer (*n* = 64)	Actuary (*n* = 100), Statistician (*n* = 93), Economist (*n* = 76), Economic consultant (*n* = 40)
Performing general physical activities Performing physical activities that require considerable use of your arms and legs and moving your whole body, such as climbing, lifting, balancing, walking, stooping and handling materials	0–7	0.19–6.52	2.38 (1.48–3.29)	Choreographer (*n* = 6), Stone mason (*n* = 36), Mason (*n* = 5), Banker mason (*n* = 2)	Actuary (*n* = 100), Statistician (*n* = 93), Economist (*n* = 76), Economic consultant (*n* = 40)

IQR, interquartile range.
[a]For the Abilities and Work Activities domains, scores of 7 represent the greatest level of that ability or activity needed for job performance. For the Work Context domain, scores of 1 representing no time spent under specified conditions and scores of 5 represent continual/almost continual time spent under specified conditions. Descriptions taken from questionnaires used for the O*NET Data Collection Program that can be found at https://www.onetcenter.org/questionnaires.html.

score amongst consultant psychiatrists (Table 1). For spending time walking/running the median score was 2.17 with the highest score amongst machinists in textile manufacturing and the lowest score amongst travel agents. For spending time standing the median score was 2.72 with the highest score amongst cooks and the lowest score amongst several jobs including computer operators.

Table 2. O*NET characteristics used as negative controls in the UK Biobank population

O*NET characteristic[a]	Scale[a]	Range observed	Median (IQR)	UK job titles with the highest scores	UK job titles with the lowest scores
Interacting with computers Using computers and computer systems (including hardware and software) to program, write software, set up functions, enter data or process information	0–7	0.06–6.10	3.50 (2.80–3.72)	Director (computing) (*n* = 270), IT director (*n* = 93), Director of IT (*n* = 44), Head of IT (*n* = 28)	Stone mason (*n* = 36), Mason (*n* = 5), Banker mason (*n* = 2), Street cleaner (*n* = 22)
Performing administrative activities Performing day-to-day administrative tasks such as maintaining information files and processing paperwork	0–7	0.38–4.88	3.15 (2.56–3.49)	Human resources manager (*n* = 789), Self development manager (*n* = 234), Personnel manager (*n* = 163)	Dancer (*n* = 8), Ballet dancer (*n* = 3), Laundry assistant (*n* = 54), Carpet cleaner (*n* = 24)

IQR, interquartile range.
[a]Descriptions taken from questionnaires used for the O*NET Data Collection Program that can be found at https://www.onetcenter.org/questionnaires.html.

Table 3. Comparisons between O*NET variables and occupational characteristics surveyed in the UK Biobank

	ANOVA R^2	Spearman's correlation coefficient (95% CI)*
Comparisons with UK Biobank 'Job involves heavy manual or physical work' variable		
Static strength	0.27	0.50 (0.50, 0.50)
Dynamic strength	0.29	0.50 (0.49, 0.50)
Stamina	0.26	0.50 (0.50, 0.51)
Spend time using your hands to handle, control or feel objects, tools or controls	0.14	0.32 (0.32, 0.33)
Handling & moving objects	0.28	0.51 (0.50, 0.51)
Performing general physical activities	0.26	0.49 (0.49, 0.49)
Interacting with computers	0.18	−0.37 (−0.38, −0.37)
Performing administrative activities	0.15	−0.35 (−0.35, −0.34)
Comparisons with UK Biobank 'Job involves mainly walking or standing' variable		
Stamina	0.24	0.49 (0.48, 0.49)
Spend time walking and running	0.29	0.52 (0.52, 0.52)
Spend time standing	0.34	0.56 (0.56, 0.56)
Performing general physical activities	0.22	0.46 (0.46, 0.46)
Interacting with computers	0.17	−0.39 (−0.40, −0.39)
Performing administrative activities	0.15	−0.34 (−0.34, −0.34)

Spearman's rank correlation coefficient was calculated with self-reported responses coded as: 0 = Never/rarely, 1 = Sometimes, 2 = Usually and 3 = Always. ANOVA was used to calculate an R^2 value estimating the proportion of the variance between self-report groups (without an imposed numerical ordering) that was explained by each O*NET variable.
*All *P*-values for Spearman's correlation tests were <0.001.

Within the Work Activities domain, the median score for handling/moving objects was 2.71 with a highest score for stone masons. The median score for performing general physical activities was 2.38 with a highest score for choreographers. For both these variables the lowest scores were observed amongst several jobs including actuaries and statisticians. The median score for interacting with computers was 3.5 with a highest score for computing and information technology directors (Table 2). The median score for performing administrative activities was 3.15 with a highest score of 4.88 for numerous jobs including human resources managers.

When comparing self-reported heavy manual/physical work to O*NET measures we observed the expected positive correlations with static strength; dynamic strength; stamina; spending time using hands to handle/control/feel objects/tools/controls; handling/moving objects; and performing general physical activities (Table 3). This was observed qualitatively in box plots, and demonstrated quantitatively by Spearman's rank correlation coefficients (Figure 2; Table 3). The strongest correlation for self-reported manual/physical work was with handling/moving objects (Spearman's coefficient: 0.51; 95% confidence interval (CI): 0.50–0.51), while the weakest correlation was with spending time using hands to handle/control/feel objects/tools/controls (Spearman's coefficient: 0.32; 95% CI: 0.32–0.33). Comparing O*NET scores across the different groups

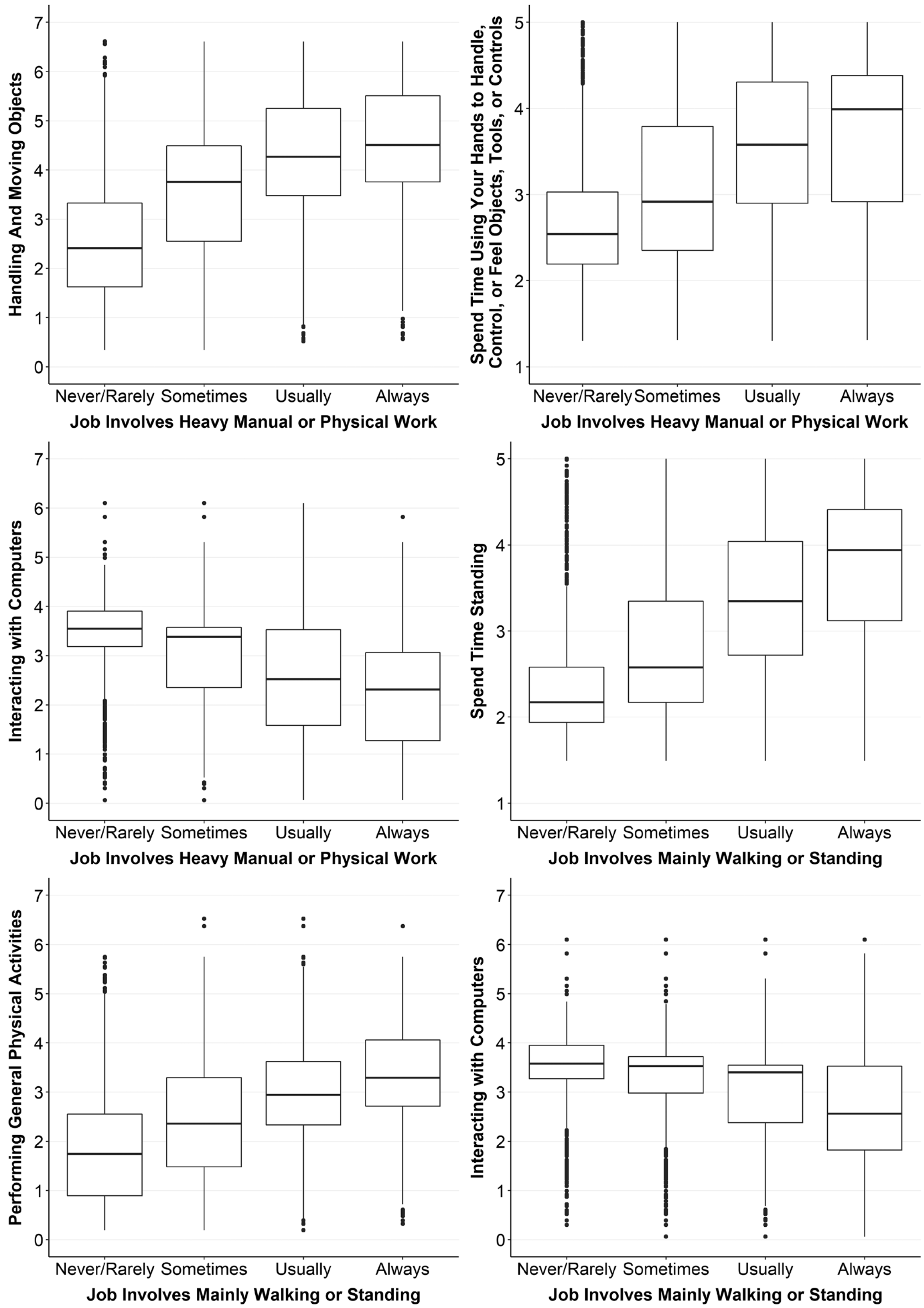

Figure 2. Distributions of select O*NET scores by self-reported heavy manual/physical work and self-reported occupational walking and standing. Boxes represent the interquartile range (IQR) with the bottom of the box at the first quartile, the top of the box at the third quartile and the middle line at the median value. The lower and upper whiskers extend 1.5 * (IQR) beyond the first and third quartiles, respectively, with any observations outside these ranges shown as dots.

of self-reported manual/physical work using ANOVA, the highest R^2 value was 0.29 for dynamic strength and the lowest R^2 value was for 0.14 for spending time using hands to handle/control/feel objects/tools/controls. Expected negative correlations were also observed between self-reported manual/physical work and both

Table 4. Comparisons between O*NET variables and self-reported occupational exposures by self-reported chronic pain at interview

	Chronic neck/shoulder pain (n = 40 301)	No chronic neck/shoulder pain (n = 243 780)	Odds ratio* (95% CI)	P*
Self-reported exposure				
Job involves heavy manual or physical work, n (%)				
Never/rarely	24 256 (60)	160 647 (66)	Ref.	Ref.
Sometimes	9376 (23)	51 642 (21)	1.22 (1.19, 1.25)	<0.001
Usually	3157 (8)	16 147 (7)	1.36 (1.30, 1.41)	<0.001
Always	3512 (9)	15 344 (6)	1.59 (1.53, 1.65)	<0.001
O*NET variables				
Static strength, median (IQR)	1.25 (0, 2.50)	1.00 (0, 2.25)	1.08 (1.07, 1.09)	<0.001
Dynamic strength, median (IQR)	0.85 (0.25, 1.38)	0.62 (0.25, 1.28)	1.13 (1.11, 1.14)	<0.001
Stamina, median (IQR)	1.00 (0, 2.00)	0.75 (0, 1.88)	1.08 (1.07, 1.10)	<0.001
Spend time using your hands to handle, control or feel objects, tools, median (IQR)	2.92 (2.32, 3.59)	2.88 (2.31, 3.51)	1.10 (1.09, 1.12)	<0.001
Handling & moving objects, median (IQR)	2.91 (2.03, 3.86)	2.71 (1.86, 3.76)	1.07 (1.06, 1.08)	<0.001
Performing general physical activities, median (IQR)	2.5 (1.49, 3.29)	2.38 (1.48, 3.29)	1.08 (1.07, 1.09)	<0.001
	Chronic hip/knee/back pain (n = 78 430)	No chronic hip/knee/back pain (n = 206 286)	Odds ratio* (95% CI)	P*
Self-reported exposure				
Job involves mainly walking or standing, n (%)				
Never/rarely	24 333 (31)	76 085 (37)	Ref.	Ref.
Sometimes	23 920 (31)	63 299 (31)	1.17 (1.15, 1.19)	<0.001
Usually	12 131 (15)	29 916 (15)	1.25 (1.22, 1.28)	<0.001
Always	18 046 (23)	36 986 (18)	1.49 (1.45, 1.52)	<0.001
O*NET variables				
Stamina, median (IQR)	1.00 (0, 2.00)	0.62 (0, 1.88)	1.14 (1.13, 1.15)	<0.001
Spend time walking/running, median (IQR)	2.23 (1.89, 3.10)	2.15 (1.88, 3.07)	1.17 (1.16, 1.18)	<0.001
Spend time standing, median (IQR)	2.79 (2.17, 3.63)	2.60 (2.15, 3.43)	1.12 (1.11, 1.13)	<0.001
Performing general physical activities, median (IQR)	2.55 (1.49, 3.37)	2.38 (1.35, 3.29)	1.11 (1.10, 1.11)	<0.001

IQR, interquartile range.
*Estimates from logistic regression adjusting for age, sex and Townsend Deprivation Index.

interacting with computers and performing administrative activities (Spearman's coefficients of −0.37 and −0.35, respectively).

We observed expected positive correlations when comparing self-reported occupational walking/standing to O*NET measures characterizing stamina, spending time walking/running, spending time standing and performing general physical activities (Figure 2; Table 3). The strongest correlations for self-reported occupational walking/standing were with the O*NET measures capturing spending time standing (Spearman's correlation coefficient = 0.56, R^2 = 0.34) and spending time walking/running (Spearman's coefficient = 0.52, R^2 = 0.29), while the weakest positive correlation was with performing general physical activities (Spearman's coefficient = 0.46, R^2 = 0.22). Expected negative correlations were observed with O*NET measures of interacting with computers and performing administrative activities (Spearman's coefficients of −0.39 and −0.34, respectively).

Higher physical workplace exposure levels were associated with chronic musculoskeletal pain across all exposures, assessed by both self-reported and JEM-based measures (Table 4).

Discussion

We successfully created a linkage between job titles reported in UKB and a JEM based on the US O*NET database. We were able to assign a range of values measuring numerous dimensions of physical work to UKB participants. We observed higher O*NET scores in UK occupations known to be more physically demanding, and confirmed expected relationships between O*NET measures and self-reported exposures. Both findings provide evidence of a valid linkage, and the potential utility of JEM-based exposure estimates.

The ranges of O*NET scores for jobs among UKB participants were similar to those reported for occupations in the USA [22]. For instance, maximum values for static strength (4.88) and dynamic strength (4.62) were identical to the maximum values assigned to US job titles. Importantly, the O*NET JEM assigned distinctly higher values to jobs that would be expected to have higher exposure.

Most O*NET measures for physical work showed moderate positive correlations with self-reported measures of occupational manual/physical work and time spent walking/standing. These correlations were similar to those found in a study comparing physical exposures measured by O*NET and self-report in a US population (most correlation statistics in 0.30–0.60 range) [21]. Correlations were stronger for comparisons with O*NET variables that were expected to more closely capture the self-reported measures. For instance, O*NET static strength scores correlated more strongly with self-reported manual/physical work than scores for time spent using your hands. The weaker correlations are likely due to capturing different underlying exposures (many tasks involving the hands are not considered heavy manual work). Greater physical work exposure was associated with common chronic pain conditions, for both self-reported and JEM-estimated exposures. Similar findings have been reported in both cross-sectional and longitudinal studies using other physical work JEMs [23–25]. While an association with self-reported data could partly be due to differential reporting of exposure between symptomatic and asymptomatic workers, JEM-estimated exposures are not subject to such reporting bias. This highlights a benefit of the addition of JEM-based exposure measures for drawing inferences from future UKB studies.

A JEM has several advantages for occupational exposure measurement in UKB. There is less potential for differential information bias than exposure questionnaires completed by study participants [6]. Additionally, a JEM provides an efficient, inexpensive way to estimate current and past occupational exposures [4]. For UKB, the O*NET JEM provides an opportunity to characterize jobs in a number of novel dimensions without requiring additional participant contact. Beyond assessing physical work measures, other job characteristics relevant to health outcomes could be examined in the future. The O*NET JEM includes information on workplace social interactions, including scores for 'Interacting with Others' and 'Physical Proximity', and work environment, including scores for exposure to 'Radiation' and 'Very Hot or Cold Temperatures'. One other JEM has been developed for use in the UK Biobank, but this JEM only captures chemical airborne exposures [26]. Consequently, our linkage provides a powerful tool for future occupational epidemiology studies in UKB.

There are important limitations to use of a JEM, which cannot capture exposure variability between workers employed in the same occupation or differences in exposure over time due to changes in work duties. There are also limitations specific to our study. First, while O*NET captures numerous work requirements, it was not built to assess risks for particular diseases, and may lack the desired specificity for some research questions. Second, despite our best efforts to match UK jobs to US jobs, some exposure misclassification likely occurred due to the differences in these classification systems, and due to differences in work duties in similarly named jobs. There is prior evidence that JEMs can be useful in a cross-national setting, though this depends upon the exposures examined and the countries across which the JEM is being transported [27–30]. O*NET has previously been linked to a French cohort where it corresponded well with other measures of physical work [27,28]. Finally, we expect UKB to have less representation of physically demanding jobs than the general population, as participants generally live in more affluent areas.

In conclusion, we linked a JEM based on the US O*NET database to the UKB population to characterize physical work exposures. The linked occupational variables captured exposure variation across jobs and produced expected correlations with self-reported occupational measurements. This tool provides novel information on a richer set of occupational measures than currently available in UKB that can be used in future epidemiologic investigations. This JEM opens the door to unique studies of occupational exposures, such as explorations of the interactions between genetics and the work environment to influence disease risk.

Funding

This work was supported by the United States National Institutes of Health (K01AR073318 to E.L.Y.) and the United States National Institute for Occupational Safety and Health (R01 OH011076 to B.A.E.). The funders did not influence the design of the study, analysis or interpretation of the data, the writing of this report or the decision to publish.

Competing interests

None declared.

Acknowledgements

This research has been conducted using the UK Biobank Resource (http://www.ukbiobank.ac.uk/) under Application Number 27034.

References

1. Collins R. What makes UK Biobank special? *Lancet* 2012;**379**:1173–1174.
2. Conroy M, Sellors J, Effingham M *et al*. The advantages of UK Biobank's open-access strategy for health research. *J Intern Med* 2019;**286**:389–397.
3. De Matteis S, Jarvis D, Young H *et al*. Occupational self-coding and automatic recording (OSCAR): a novel

web-based tool to collect and code lifetime job histories in large population-based studies. *Scand J Work Environ Health* 2017;**43**:181–186.
4. Evanoff B, Zeringue A, Franzblau A, Dale AM. Using job-title-based physical exposures from O*NET in an epidemiological study of carpal tunnel syndrome. *Hum Factors* 2014;**56**:166–177.
5. Descatha A, Evanoff BA, Andersen JH *et al*. JEMINI (Job Exposure Matrix InterNatIonal) initiative: a Utopian possibility for helping occupational exposure assessment all around the world? *J Occup Environ Med* 2019;**61**:e320–e321.
6. Viikari-Juntura E, Rauas S, Martikainen R *et al*. Validity of self-reported physical work load in epidemiologic studies on musculoskeletal disorders. *Scand J Work Environ Health* 1996;**22**:251–259.
7. Dale AM, Strickland J, Gardner B, Symanzik J, Evanoff BA. Assessing agreement of self-reported and observed physical exposures of the upper extremity. *Int J Occup Environ Health* 2010;**16**:1–10.
8. Trask C, Mathiassen SE, Wahlström J, Forsman M. Cost-efficient assessment of biomechanical exposure in occupational groups, exemplified by posture observation and inclinometry. *Scand J Work Environ Health* 2014;**40**:252–265.
9. Trask C, Teschke K, Village J *et al*. Measuring low back injury risk factors in challenging work environments: an evaluation of cost and feasibility. *Am J Ind Med* 2007;**50**:687–696.
10. Dale AM, Ekenga CC, Buckner-Petty S *et al*. Incident CTS in a large pooled cohort study: associations obtained by a job exposure matrix versus associations obtained from observed exposures. *Occup Environ Med* 2018;**75**:501–506.
11. Verma SK, Sorock GS, Pransky GS, Courtney TK, Smith GS. Occupational physical demands and same-level falls resulting in fracture in female workers: an analysis of workers' compensation claims. *Inj Prev* 2007;**13**:32–36.
12. Zhou W, Chandran V, Cook R, Gladman DD, Eder L. The association between occupational-related mechanical stress and radiographic damage in psoriatic arthritis. *Semin Arthritis Rheum* 2019;**48**:638–643.
13. d'Errico A, Punnett L, Cifuentes M *et al*.; Promoting Healthy and Safe Employment In Healthcare Research Team. Hospital injury rates in relation to socioeconomic status and working conditions. *Occup Environ Med* 2007;**64**:325–333.
14. Health and Safety Executive. Work-Related Musculoskeletal Disorder Statistics (WRMSDs) in Great Britain, 2020. 2020. www.hse.gov.uk/statistics/ (9 July 2021, date last accessed).
15. Manolio TA, Weis BK, Cowie CC *et al*. New models for large prospective studies: is there a better way? *Am J Epidemiol* 2012;**175**:859–866.
16. UK Biobank. Data Showcase. UK Biobank, 2021. https://biobank.ndph.ox.ac.uk/showcase/ (20 July 2021, date last accessed).
17. United Kingdom Office for National Statistics. SOC 2000: Previous Version of the Standard Occupational Classification. United Kingdom Office for National Statistics, 2016. https://www.ons.gov.uk/methodology/classificationsandstandards/standardoccupationalclassificationsoc/socarchive (9 July 2021, date last accessed).
18. United Kingdom Office for National Statistics. SOC 2010 Volume 2: The Structure and Coding Index. https://www.ons.gov.uk/methodology/classificationsandstandards/standardoccupationalclassificationsoc/soc2010/soc2010volume2thestructureandcodingindex (9 July 2021, date last accessed).
19. United Kingdom Office for National Statistics. SOC 2010: Previous Version of the Standard Occupational Classification. United Kingdom Office for National Statistics, 2016 https://www.ons.gov.uk/methodology/classificationsandstandards/standardoccupationalclassificationsoc/soc2010 (9 July 2021, date last accessed).
20. United States Bureau of Labor Statistics. Crosswalks Between the 2010 SOC and Systems Used by Other Federal and International Statistical Agencies United Kingdom Office for National Statistics. United States Bureau of Labor Statistics, 2012. https://www.bls.gov/soc/soccrosswalks.htm (9 July 2021, date last accessed).
21. Gardner BT, Lombardi DA, Dale AM, Franzblau A, Evanoff BA. Reliability of job-title based physical work exposures for the upper extremity: comparison to self-reported and observed exposure estimates. *Occup Environ Med* 2010;**67**:538–547.
22. O*NET 24.2 Database [Internet]. National Center for O*NET Development, 2020. https://www.onetcenter.org/ (13 May 2020, date last accessed).
23. Badarin K, Hemmingsson T, Hillert L, Kjellberg K. Physical workload and increased frequency of musculoskeletal pain: a cohort study of employed men and women with baseline occasional pain. *Occup Environ Med* 2021;**78**:558–566.
24. Solovieva S, Pehkonen I, Kausto J *et al*. Development and validation of a job exposure matrix for physical risk factors in low back pain. *PLoS One* 2012;**7**:e48680.
25. Madsen IEH, Gupta N, Budtz-Jørgensen E *et al*. Physical work demands and psychosocial working conditions as predictors of musculoskeletal pain: a cohort study comparing self-reported and job exposure matrix measurements. *Occup Environ Med* 2018;**75**:752–758.
26. Dimakakou E, Johnston HJ, Streftaris G, Cherrie JW. Evaluation of the suitability of an existing job-exposure matrix for the assessment of exposure of UK Biobank participants to dust, fumes, and diesel exhaust particulates. *Int J Environ Res Public Health* 2020;**17**:4919.
27. Yung M, Evanoff BA, Buckner-Petty S, Roquelaure Y, Descatha A, Dale AM. Applying two general population job exposure matrices to predict incident carpal tunnel syndrome: a cross-national approach to improve estimation of workplace physical exposures. *Scand J Work Environ Health* 2020;**46**:248–258.
28. Evanoff B, Yung M, Buckner-Petty S *et al*. Cross-national comparison of two general population job exposure matrices for physical work exposures. *Occup Environ Med* 2019;**76**:567–572.
29. Lavoué J, Pintos J, Van Tongeren M *et al*. Comparison of exposure estimates in the Finnish job-exposure matrix FINJEM with a JEM derived from expert assessments performed in Montreal. *Occup Environ Med* 2012;**69**:465–471.
30. Offermans NS, Vermeulen R, Burdorf A *et al*. Comparison of expert and job-exposure matrix-based retrospective exposure assessment of occupational carcinogens in the Netherlands Cohort Study. *Occup Environ Med* 2012;**69**:745–751.