Validity of Workers' Self-Reports. Evaluation of a Question Assessing Lifetime Exposure to Occupational Physical Activity

Anne Møller[1,2,3*], Susanne Reventlow[3], Johan Hviid Andersen[4], Kirsten Avlund[5,6,7] and Ole Steen Mortensen[2,8]

[1]*Department of Occupational and Environmental Medicine, Køge Hospital, Lykkebækvej 1, DK-4600 Køge, Denmark.*
[2]*National Research Centre for the Working Environment, Copenhagen, Denmark.*
[3]*The Research Unit for General Practice and Section of General Practice, Department of Public Health, University of Copenhagen, Copenhagen, Denmark.*
[4]*Danish Ramazzini Centre, Department of Occupational Medicine, Regional Hospital, Herning, Denmark.*
[5]*Section of Social Medicine, Department of Public Health, University of Copenhagen, Copenhagen, Denmark.*
[6]*Centre for Healthy Aging, University of Copenhagen, Copenhagen, Denmark.*
[7]*Danish Aging Research Centre, Universities of Southern Denmark, Aarhus and Copenhagen, Denmark.*
[8]*Department of Occupational and Environmental Medicine, Bispebjerg Hospital, Copenhagen, Denmark.*

Authors' contributions

This work was carried out in collaboration between all authors. AM, JHA, OSM and SR planned the study together with KA. KA was responsible for the CAMB data collection. AM interviewed the participants and made the first draft of the article. JHA, OSM, SR and KA made substantial contributions to the analysis and interpretation of data. They drafted the article critically for important intellectual content and all authors have read and approved the final manuscript.

Research Article

ABSTRACT

Aims: In epidemiological studies exposure assessment based on questionnaires is the most cost-effective method. A question about lifetime exposure to occupational physical

---

*Corresponding author: Email: annemoller1972 @gmail.com;

activity (OPA) was used in a population-based survey (part of the Copenhagen Aging and Midlife Biobank, CAMB). The aim of the study was to validate this question through a three-step process.

Methodology: Firstly, the response process was studied by cognitive interviewing of 7 persons. Secondly, 64 persons participated in semi-structured interviews about their work-life, and expert judgments of exposure to OPA were compared with questionnaire-data. Exposure was 20 years of work in one of four categories of OPA: sedentary, standing and walking, moderate or high OPA. Kappa values were calculated for agreement and interpreted according to Landis and Koch's criteria. Agreement was visualized in Bland-Altman plots. Thirdly, intra- and inter-rater reliability of expert judgments was tested.

Results: Response process: The question had a complicated instruction, and the respondents found it hard to remember, categorize, and summate exposures. Semi-structured interviews: Kappa value for exposure to sedentary work was 'substantial' (0.71) but 'fair' for the other categories of OPA (0.27-0.29). Agreement between questionnaire and interview was higher in sedentary jobs and jobs with high OPA. Intra-rater reliability of expert judgments was 'substantial' or 'moderate' (0.60-0.71). Inter-rater reliability was high in sedentary jobs but lower in the more active jobs.

Conclusion: Self-reports of lifetime exposure to sedentary work are valid in the CAMB cohort, whereas the validity of self-reports of exposure to high levels of occupational physical activity (OPA) are questionable. Thorough pre-testing of questions about lifetime OPA is recommended.

ABBREVIATIONS

OPA: Occupational Physical Activity; CAMB: Copenhagen Aging and Midlife Biobank; QAS: The Question Appraisal System

1. INTRODUCTION

Reliable and valid assessments of occupational physical activity are needed in the study of work and health (Stock et al., 2005; Kwak et al., 2011). In epidemiological studies, which include participants with many different job-titles, exposure assessment based on questionnaires is the most cost-effective method. Many questionnaires and scales assessing occupational physical activity (OPA) have been used, and a recent review found good repeatability in four of 22 questionnaires. However, none of the reviewed questionnaires showed good validity compared to objective measurements (Kwak et al., 2011). This could be partly explained by lack of standardized methods for assessment of OPA and, thereby, lack of a 'gold standard' as reference method. Another explanation is the lack of studies of workers' capability to describe and judge the level of exposure and Stock et al. (2005) suggest that qualitative interdisciplinary methods like 'cognitive interviewing' are used to pre-test questionnaires concerning physical workload.

The questionnaires reviewed by Kwak et al. (2011) and Stock et al. (2005) assessed current OPA by asking questions about usual activity at work, a 'typical workday', or usual activity in the past week or year. Assessment of lifetime exposure to occupational physical activity is an additional challenge and personal interviews have been used to establish a retrospective

job-history, which has been reviewed afterwards by experts, assessing lifetime occupational physical activity (Cassou et al., 1992). However, this is a time-consuming method in large epidemiologic studies and expert judgments have to be validated too.

In the planning of a Danish cohort study (the Copenhagen Aging and Midlife Biobank (CAMB)) (Avlund et al., 2009) we contributed with questions about work-life. CAMB is based on three existing Danish cohorts and aimed at determining the importance of prenatal and perinatal factors, factors in childhood, and factors in early adulthood for early signs of ageing in late midlife. Our study group's main interest is the influence of work on the ageing process and in forthcoming analyses we will study lifetime exposure to occupational physical activity and associations to midlife physical function (Møller et al., 2012).

The questionnaire used in CAMB included 100 different questions about health, social and life-style factors; consequently, the space for questions concerning work-life was limited. Based on more than 20 years of experience and several validity studies on assessment of exposures in work-life at The National Research Centre for the Working Environment (Burr et al., 2003), we included a question about OPA in current work in the CAMB questionnaire. The time-frame of the question was changed to cover the entire work-life, to serve as a cumulative exposure assessment in our study of lifetime OPA and ageing. Pilot studies of the CAMB questionnaire resulted in a slightly changed wording of the question. When inclusion into CAMB started, the research assistants reviewing the questionnaires with the participants found that some participants had difficulty in answering this specific question (Question 32, see Appendix 1).

Therefore, we conducted a supplementary small pilot study by introducing the question to a few people. Respondents with sedentary work-life filled out the question about lifetime exposure to OPA satisfactorily, but respondents with exposure to some OPA in work-life had difficulties answering the question. At that time, we were not able to change the question in the CAMB survey. Therefore, we decided to study to which extent we could rely on data from the questionnaire. We planned a three-step process of validation, aiming at answering the following three research questions:

1) How is a question about life-time OPA interpreted and understood by people with a job history of primarily manual work?
2) How is the agreement between exposures to OPA reported in the CAMB questionnaire and information obtained from interviews?
3) How reliable are expert ratings of lifetime occupational physical activity?

The aim of the first step in the process of validation (see Table 1 for overview) was to study the comprehension and interpretation of the question about lifetime OPA because, despite the recommendations made by Stock et al. (2005), we have not seen qualitative methods used in the pre-test of questionnaires about OPA. Furthermore, the aim of the first step was to gain knowledge to be used in the next step of validation. In the second step, the validity of self-reports of lifetime OPA was evaluated, comparing data from questionnaires and from semi-structured interviews. Finally, intra- and inter-rater reliability of expert judgments of OPA used in the semi-structured interviews was evaluated.

Table 1. A three-step validation process. Methods, materials and analyses

Step	Content	Method	Material	Analyses
1	Response process, comprehension and interpretation	Cognitive interviews	7 middle-aged workers primarily working in a hospital	Question Appraisal System, QAS-checklist
2	Validity of self-reports	Semi-structured interviews CAMB questionnaire	64 participants from CAMB[1]	Kappa Bland-Altman plots
3	Intra-rater reliability	Re-test of interview-data from step 2 after 3 months	Primary rater's first judgment in step 2 compared with blinded re-test	Kappa Bland-Altman plots
	Inter-rater reliability	Expert judgments based on interview-data from step 2	3 expert's judgments compared to initial rater's judgments	Bland-Altman plots

[1]*CAMB: Copenhagen Aging and Midlife Biobank*

2. MATERIALS AND METHODS

2.1 Study Design

Participants in CAMB filled in the questionnaire before attending a physical examination. Information about work-life from the questionnaire included a list of the five longest held occupations, current job type and physical, ergonomic, chemical, and psychosocial exposures at work. In the question about lifetime OPA, participants were asked to fill in information about number of years of work in four categories of physical activity: a) sedentary work, b) standing and walking at work, c) moderate OPA and d) high OPA (See Appendix 1).

2.2 Comprehension and Interpretation

Cognitive interviewing has been used since the 80's to improve the quality of survey questions (Willis, 2005; Collins, 2003) and in medical research it has been used in the development of new questionnaires (Watt et al., 2008), revision of existing questionnaires after translation (Andersen et al., 2010), or, before use, in a different cultural setting than the primary one (Napoles-Springer et al., 2006; Cortes et al., 2007). Cognitive interviews study the cognitive aspects of the response process and, thereby, respondents' interpretation and comprehension of questions (Tourangeau et al., 2000).

The respondents received a printed copy of the questions about work-life and were encouraged to 'think aloud' while filling in the questionnaire, as described by Willis (Willis, 2005). However, the 'think aloud' technique is a challenge to some respondents, and we therefore also used 'verbal probing', meaning that the interviewer asks questions (probes) during the interview ('concurrent probing') (Willis, 2005). Probes can either be prepared or spontaneous and are used to explore the comprehension of terms and to catch silent misunderstandings of questions. 'Retrospective probing' was used at the end of the interview

to make a concluding evaluation of the questions concerning work-life (Willis, 2005). Interviews were digitally recorded, and notes and comments were taken during the interview. The interviews were transcribed verbatim.

2.2.1 Population and data collection

From our small pilot study we knew that respondents with sedentary work filled out the question about lifetime exposure to OPA satisfactorily. However, respondents with exposure to some OPA in work-life had difficulties answering the question. Based on this pilot study, a strategic sampling of participants not included in the CAMB study was made. Selection was based on age (minimum 50 years old) and working experience (at least 20 years of non-sedentary work) (Crabtree and Miller, 1999). Participants were primarily recruited among employees at the hospital, and inclusion continued until no further problems in the question of interest were revealed in the interviews, as in 'sampling to redundancy' (Streiner and Norman, 2008). Four men and three women, average age 59 years, were interviewed. Three hospital workers, one secretary with former employment as an assistant nurse, a laboratory assistant and two men with working experience from outside the hospital. Interviews took place in January and February 2010.

2.2.2 Analysis

The analysis was based on recordings and notes from the interviewer, according to Willis' "The Question Appraisal System" (QAS) (Willis, 2005), using a check-list of seven categories covering the answering process, Table 2. No quantitative measurement of responses was made because the aim of the interviews was primarily to gain an insight into the response process (Watt et al., 2008).

Table 2. The Question Appraisal System

Category	Description
1. Instructions	Look for problems with introductions, instructions, or explanations from the *respondent's* point of view.
2. Clarity	Identify problems related to communicating the *intent* or *meaning* of the question .
3. Assumptions	Determine if there are problems with assumptions made or the underlying logic.
4. Knowledge/ Memory	Check whether respondents are likely to *not* know or have trouble remembering information.
5. Sensitivity/Bias	Assess questions for sensitive nature or wording, and for bias.
6. Response categories	Assess the adequacy of the range of responses to be recorded.
7. Other Problems	Look for problems not identified in steps 1- 6.

2.3 Validity of Self-Reports

The overall aim of the semi-structured interviews was to establish a retrospective job-history, including information about exposures in work-life. The semi-structured interview was based

on an interview-guide, but other questions were allowed to be brought up during the interview (Kvale, 1997). The interview guide was designed for this study based on the knowledge from the cognitive interviews (Step 1).

2.3.1 Population and data collection

75 participants from the CAMB-study were invited to participate in the semi-structured interviews. They were selected strategically, based on their answers about lifetime OPA (Question 32). In order to study possible variations in agreement between exposure groups, 15 participants with at least 20 years of exposure in each of the four categories (a-d) were selected and, in addition, another 15 participants with mixed job-histories. In all other aspects, the selection was random, and the first 15 to fit into the five defined groups of exposure were included. They received a mailed invitation to participate in a telephone interview about their work-life, and the researcher (AM) called them within the next two weeks to set an appointment for the telephone interview. The participants were anonymous in the data material, but coded with a unique registration number from the CAMB-study. At the time of the interview, the interviewer was blinded to the participants' information about exposure status in the questionnaire. The participants were interviewed in May and June 2010, and interviews were digitally recorded.

The interview-guide was based on results from the cognitive interviews and the first question in the retrospective part of the interview was: "Now we are going to talk about your employment since you left school, i.e. all the different jobs you have had during your work-life. When did you finish school, and what did you do afterwards?" The interviewer took notes and was thus able to piece together a story about the entire work-life in cooperation with the respondent. Once the interviewer had an overview of the job-history, she asked more thorough questions about exposures in the work environment. Having finished the interview, the interviewer filled in data about employment and exposures in a database, and went through the recordings of the interviews at least once more. Finally, judgment of level and duration of lifetime OPA was made (answer to question 32), and the judgment was not discussed with the participant.

2.3.2 Analysis

Validity was calculated as kappa coefficients of agreement in exposure using the dichotomized outcome: 20 years of exposure in the specific category or not ("exposed" or "non-exposed"). There is no general consensus about interpretation of kappa values, but we used the slightly adapted guidelines from Landis and Koch's (Altman, 1999) (Strength of agreement: 1.00: Perfect agreement, 0.81-1.00: Very good, 0.61-0.8: Good, 0.41-0.60: Moderate, 0.21-0.40: Fair, <0.2: Poor). However, the kappa coefficient is a dimensionless ratio, and the true agreement or clinical implication of the kappa coefficient is not obvious from the size of the coefficient. Therefore, Bland-Altman plots were used to visualize agreement (Bland and Altman, 1999). For that reason, we calculated an index of OPA taking years of exposure into account (Appendix 2). The OPA-index is based on questionnaire information about years of exposure to OPA in 4 groups, and ranges from 0 and 0.7. An OPA-index of "0" means "no OPA during work-life" and one of "0,7" means "having had OPA throughout the entire work-life". Differences in the OPA-index in the interview and the questionnaire were plotted against their mean, and the lines for the mean-value and the 95% limits of agreement were drawn. If the mean is 0 there is perfect agreement, and the narrower the 95% limits, the better agreement (Bland and Altman, 1999).

2.4 Intra- and Inter-Rater Reliability

2.4.1 Population and data collection

2.4.1.1 Intra-rater reliability

Intra-rater reliability of the expert judgment was evaluated by a test-retest of the OPA-index in all participants. The primary rater, AM, performed a blinded re-judgment of the exposure to OPA three months after the initial judgments, based on the data from the interview about job-history and exposures in work-life.

2.4.1.2 Inter-rater reliability

Three skilled, occupational physicians received information about 34 randomly selected participants from the interview-database, and were asked to judge the level and duration of exposure to OPA (years of exposure in group a-d) in each participant.

2.4.2 Analysis

2.4.2.1 Intra-rater reliability

Kappa values for agreement to exposure in test and re-test were calculated. OPA-index for each participant was calculated, and the difference between the primary OPA-index and the re-tested OPA-index was plotted against the mean of the two indices in a Bland-Altman plot.

2.4.2.2 Inter-rater reliability

The difference between the OPA-index judged by the primary rater and each of the three skilled physicians was visualized in one Bland-Altman plot with only one reference-line in y=0, in order to keep the figure simple, and to visualize the agreement which was the primary aim.

3. RESULTS AND DISCUSSION

3.1 Comprehension and Interpretation

3.1.1 Results

Instruction was complicated, aiming at assessing duration (years of exposure in each category), frequency ('mostly') and intensity (level of physical activity in category a) to d)) (Table 3). According to 'Clarity', some respondents were confused about category d) describing 'high speed' and 'heavy and physically demanding work', while they had been working at a 'high speed' but not with heavy work, and 'speed' was not mentioned in the other categories. Questions about employment and exposures back in time caused 'recall problems' in most respondents, and different approaches were used in the search of information, but most participants used first job or graduation as their starting point. 'Computation problems' were obvious in the search for duration of jobs and summation of exposures throughout work-life. Response categories b), c) and d) were overlapping due to vague definitions of levels of physical activity. Category a) was interpreted as office

work/work in front of a computer by everyone, and caused no problems. The distinction between category c) or d) was hard, and some respondents asked for examples of job-titles in the categories. Since the instruction included an option of 'answering in more than one category', some filled in e.g. 40 years of work in both category c) and d) to indicate their difficulties in categorization of exposure. One participant found that her job did not fit into any of the categories and wrote 0 years in all four boxes. Only one of seven respondents understood and answered the question about lifetime occupational physical activity the way it was intended by the researchers.

Table 3. The Question Appraisal System used in the analysis of the question about lifetime occupational physical activity. Some categories and citations are shown in this table

Category	Citations and notes from interviews
Instructions	Most respondents sighed when they read the question and explained that it was hard to understand and impossible to answer correctly
Clarity	The use of "speed" only in category d) was confusing. 'I have always worked fast, but my work has not been hard, but "speed" is not mentioned in category a), b), or c)'. '…..Standing and walking' the respondent "tasted" the word and got confused about the meaning of the expression
Assumptions	In the question constant exposure during a work-day is assumed, but respondents were confused by this assumption: 'I was sitting at the office before lunch, and having heavy work while packaging in the afternoon.'
Knowledge/ Memory	Exposures up to forty years back in time are hard to recall, and the question requires difficult mental calculation.
Response categories	Vague response categories result in wrong answers, since they overlap: 'my job is a mixture… I sit, I walk, I stand, I lift and I laugh…it is hard to choose which category'

3.1.2 Discussion

As we presumed after our pilot study, the cognitive interviews revealed some problems, due to the response process. We found problems in the categorization of physical demands at work and assumptions of constant behavior during a workday and during work-life in the question. Furthermore, it was hard to remember occupational physical activity back in time, and it is known that the higher demands on memory in a question, the less accurate the response will be (Tourangeau et al., 2000). Everyday experiences are liable to imply reconstruction or inference more often than special events. The longer distance in time between an experience in the past and the present, the more difficult it is to remember, not only because of the period of time, but because you may have experienced similar things in the meantime (Tourangeau et al., 2000). However, sedentary jobs were easily categorized as such in the interviews.

The participants were selected strategically and the results from the interviews have low external validity. However, the participants were selected among workers who were assumed to have had some exposures to OPA. In the pilot study of the entire CAMB questionnaire, problems in question 32 were not seen, and though participants in that pilot

study were selected strategically to mirror the CAMB population, there may have been an underrepresentation of manual workers or persons with low educational background.

It may be argued that seven interviews were too few to reach redundancy, but we found that most respondents faced the same problems in the response process. The aim of the cognitive interviews was to explore the response process to be able to design an interview guide for the second step of validation and we gained a useful insight into the problems linked to recall of exposures and reconstruction of lifetime job history.

3.2 Validity of Self-Reports

3.2.1 Results

64 of 75 (85%) participants accepted the invitation, 47% were women, mean age 56,4 years, and mean length of work-life was 39 years (range 22-48). The kappa value for agreement between questionnaire data and interview data for exposure to sedentary work was 'substantial' (0.71) (Table 4). For standing and walking and moderate OPA agreement was 'fair' (kappa 0.23 and 0.37 respectively). Exposure to 20 years of either moderate or high OPA (category c) and d) together) showed 'moderate' agreement (kappa 0.53).

Table 4. Validity of self-reports. Questionnaire versus interview

	Kappa	95% CI
Exposure to a)/ sedentary	0.71	0.50-0.93
Exposure to b)/ standing and walking	0.23	-0.02-0.45
Exposure to c)/ moderate OPA[a]	0.37	0.17-0.57
Exposure to d)/ high OPA[a]	0.27	0.04-0.49
Exposure to c) or d)	0.53	0.35-0.71

[a]*OPA: Occupational physical activity*

Fig. 1 shows the Bland-Altman plot of agreement in OPA-index between interviews and questionnaires. There is satisfactory agreement in low OPA-indices, which means that a sedentary job is categorized equally by the respondent and the rater. The agreement decreases as the OPA-index increases, but for the few high index jobs agreement seems to increase again.

3.2.2 Discussion

Both kappa values and Bland-Altman plots showed that the lower the level of OPA in the job history, the higher the agreement between self-reports and interviews. This is in line with results presented by Torgen et al. (1999) about 6 year recall of workloads, based on questionnaire information and validated by observation. The lower agreement in reports of higher levels of OPA is presumably a result of the problems of the categorization of OPA levels found in the cognitive interviews. Other researchers in this field have experienced problems in self-reported information about exertion and specific working postures (Wiktorin et al., 1993; Mortimer et al., 1999; Viikari-Juntura et al., 1996).

For lack of a 'gold standard' of OPA assessment we have studied the inter-method agreement (Gardner et al., 2010). To validate information from the questionnaire we could have used measurements, logbooks, or observations (Torgen et al., 1999). But as the aim of

the exposure assessment was a lifetime assessment of OPA, this was not possible. Our hypothesis was that the information retrieved by interviews was more valid than self-reports, but this hypothesis has not been tested. However, White et al. (2008) state that interviews are superior to questionnaires if questions are complex and that precise information, e.g. about past exposures, is needed.
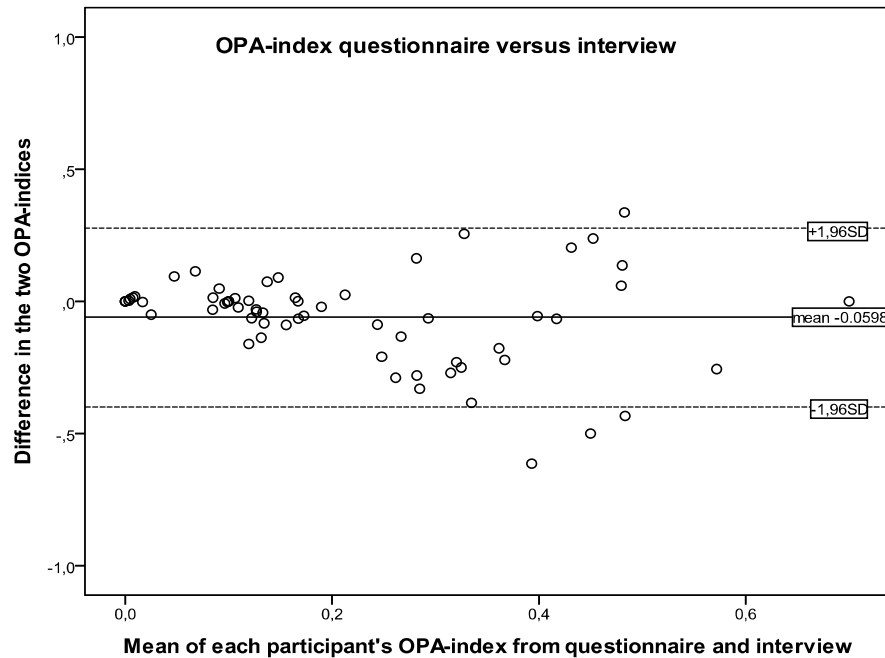

Fig. 1. A question about lifetime exposure to occupational physical activity (OPA) was validated, comparing questionnaire and interview data. An index of OPA was calculated (OPA-index) in each participant based on information from the questionnaire and the interviews. The difference between the two OPA-indices is visualized

In the planning of the study, we chose not to examine the reliability of the question about lifetime OPA because Stock et al. (2005) concluded that the reliability of workers' self-reports about general body postures (e.g. sitting and standing) is 'good to excellent'. We chose to focus on reliability of expert judgments, but, in the light of the results of our study, it would have been interesting also to study the reliability of workers' self-reports.

From the cognitive interviews we knew that categorization of OPA in question 32 was difficult. Highly educated workers may have little or no exposure to OPA (Stock et al., 2005), and thus their jobs are easier to categorize. On the other hand, categorization of jobs with moderate or high levels of OPA may bother respondents with low education. Gender, age, socio-demographics, and musculoskeletal complaints have been hypothesized to influence self-reports of exposure assessment (Sembajwe et al., 2010; Quinn et al., 2007; Viikari-Juntura et al., 1996; Wiktorin et al., 1993; Stock et al., 2005). In forthcoming analyses, it would be interesting to study the effect of these factors in workers' self-reports.

3.3 Intra- and Inter-Rater Reliability

3.3.1 Results

3.3.1.1 Intra-rater reliability

Kappa was 'substantial' for exposure to sedentary work, standing/walking and high OPA (kappa 0.71, 0.62, and 0.64 respectively, Table 5). For exposure to moderate OPA, agreement was 'moderate' (kappa= 0.60). In Fig. 2, intra-rater reliability is shown in a Bland-Altman plot of the agreement in the OPA-index. Intra-rater agreement between initial ratings and blinded ratings three months later was high, but full agreement between the judgments was not obtained.

Table 5. Intra-rater reliability. Test-retest

	Kappa	95% CI
Exposure to a)/ sedentary	0.71	0.48-0.95
Exposure to b)/ standing and walking	0.62	0.42-0.82
Exposure to c)/ moderate OPA[a]	0.60	0.40-0.80
Exposure to d)/ high OPA[a]	0.64	0.31-0.96

a)   *OPA: Occupational physical activity*
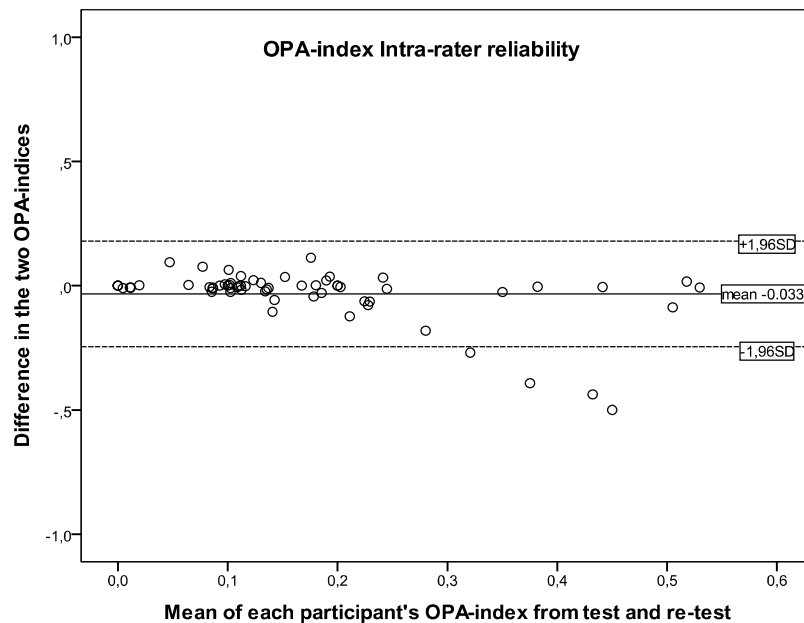
Fig. 2. Intra-rater reliability was evaluated by a blinded re-judgment of exposure to occupational physical activity (OPA) three months after the initial judgment. The difference between the two OPA-indices was visualized against the mean of the indices in a Bland-Altman plot

3.3.1.2 Inter-rater reliability

In Fig. 3, inter-rater reliability is shown, plotting the primary rater against each of the three experts. Inter-rater reliability is high in low OPA-indices but increases with higher OPA-indices. In general, the primary rater tends to score the OPA-index higher than the other experts.
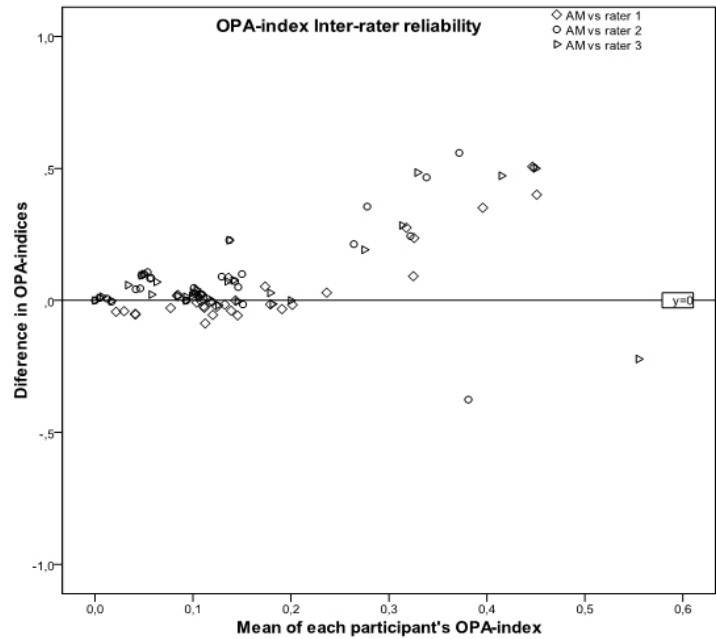

Fig. 3. Three skilled occupational physicians judged the exposure to occupational physical activity (OPA) in 34 participants, based on data from interviews. The difference in the individual OPA-index between the primary rater and each of the three skilled physicians is visualized

3.3.2 Discussion

The reliability of expert judgments of level of OPA in work-life varies according to exposure levels. As seen in the semi-structured interviews, agreement is higher in jobs with lower levels of OPA. Categorization of exposure in group c) or d) was difficult among participants in the cognitive interviews, and, in this third step, it was shown that experts have difficulty in reproducing the categorization of moderate or high level of OPA. The categories are not sufficiently specific for reliable judgment, and we assume that the same results would be found, if reliability of self-reports was tested in the CAMB participants.

According to the reliability of expert judgments, we found good agreement in sedentary jobs but lower agreement in the rating of more physically strenuous jobs. D'Souza et al. found that inter-rater agreement for physical exposure in job-categories was low, except for "sitting position", but their rating procedure was complicated due to heterogeneous exposure-groups (D'Souza et al., 2007). Expert judgments are often seen as a "gold standard" in occupational epidemiology, but risk of misclassification of exposure is still possible. Expert judgments are group-based and individual differences in exposures due to variation in job tasks,

ergonomics and capacity among people with same job-title are not taken into account (Benke et al., 1997).

4. CONCLUSION

In a three-step process, we have studied the validity of workers' self-reports and found that self-reports of lifetime exposure to sedentary work are valid in the CAMB cohort, whereas the validity of self-reports of exposure to moderate and high levels of occupational physical activity is questionable.

Our findings are in line with others concluding that self-administered questionnaires may help to classify groups with heterogeneous occupational tasks but are not suitable for studying quantitative exposure-effect relationships (Stock et al., 2005; Viikari-Juntura et al., 1996).

Introducing a qualitative method like cognitive interviewing in the occupational research field was beneficial to our study. Knowledge about comprehension is essential to the validity and, thus, cognitive interviewing or other methods of pre-testing questions are recommended for use in future planning and pre-testing of questions about work-life. Furthermore, we have shown that it is important to pre-test questionnaires in sub-groups, because many factors may influence the way people answer questions about exposures in their work-life.

ETHICAL APPROVAL

The study was presented to the Ethics committee, but the general approval of the CAMB project covered this project (The CAMB project was approved by the Regional Committee on Biomedical Research Ethics, Capital Region, Registration Number H-A-2008-126). "The Danish Data Protection Agency" refused registration of this project, because the questions in both cognitive and semi-structured interviews were only work-related. All authors hereby declare that all human studies have been examined and approved by the appropriate ethics committee and have therefore been performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki.

ACKNOWLEDGEMENTS

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

Altman, D.G. (1999). Practical statistics for medical research. London, Chapman & Hall.

Andersen, L.G., Groenvold, M., Jorgensen, T., Aadahl, M. (2010). Construct validity of a revised Physical Activity Scale and testing by cognitive interviewing. Scand. J Public Health, 38, 707-714.

Avlund, K., Bruunsgaard, H., Christensen, U., Fiehn, N.-E., Hansen, A.M., Holm-Pedersen, P., Jepsen, E., Molbo, D., Mortensen, E.L., Osler, M., Lund, R. (2009). (CAMB - Copenhagen Aging and Midlife Biobank. Perspectives for future research.) CAMB - Copenhagen Aging and Midlife Biobank. Perspektiver for fremtidig forskning. Miljø og Sundhed supplement 1, 81-88.

Benke, G., Sim, M., Forbes, A., Salzberg, M. (1997). Retrospective assessment of occupational exposure to chemicals in community-based studies: validity and repeatability of industrial hygiene panel ratings. Int J Epidemiol., 26, 635-642.

Bland, J.M., Altman, D.G. (1999). Measuring agreement in method comparison studies. Stat. Methods Med. Res., 8, 135-160.

Burr, H., Bjorner, J.B., Kristensen, T.S., Tuchsen, F., Bach, E. (2003). Trends in the Danish work environment in 1990-2000 and their associations with labor-force changes. Scand. J Work Environ. Health, 29, 270-279.

Cassou, B., Derriennic, F., Iwatsubo, Y., Amphoux, M. (1992). Physical disability after retirement and occupational risk factors during working life: a cross sectional epidemiological study in the Paris area. J Epidemiol. Community Health, 46, 506-511.

Collins, D. (2003). Pretesting survey instruments: an overview of cognitive methods. Qual. Life Res., 12, 229-238.

Cortes, D.E., Gerena, M., Canino, G., Aguilar-Gaxiola, S., Febo, V., Magana, C., Soto, J., Eisen, S.V. (2007). Translation and cultural adaptation of a mental health outcome measure: the BASIS-R(c). Cult. Med. Psychiatry, 31, 25-49.

Crabtree, B.F., Miller, W.L. (1999). Doing Qualitative Research. Sage Publications Inc..

D'Souza, J.C., Keyserling, W.M., Werner, R.A., Gillespie, B., Franzblau, A. (2007). Expert consensus ratings of job categories from the Third National Health and Nutrition Examination Survey (NHANES III). Am. J Ind. Med., 50, 608-616.

Gardner, B.T., Lombardi, D.A., Dale, A.M., Franzblau, A., Evanoff, B.A. (2010). Reliability of job-title based physical work exposures for the upper extremity: comparison to self-reported and observed exposure estimates. Occup. Environ. Med., 16, 1-10.

Kvale, S. (1997). InterView. København, Hans Reitzels.

Kwak, L., Proper, K.I., Hagstromer, M., Sjostrom, M. (2011). The repeatability and validity of questionnaires assessing occupational physical activity - a systematic review. Scand. J Work Environ. Health, 37, 6-29.

Mortimer, M., Hjelm, E.W., Wiktorin, C., Pernold, G., Kilbom, A., Vingard, E. (1999). Validity of self-reported duration of work postures obtained by interview. MUSIC-Norrtalje Study Group. Appl. Ergon., 30, 477-486.

Møller, A, Mortensen, O.S., Reventlow, S., Skov, P.G., Andersen, J.H., Rubak, T.S., Hansen, A.M., Andersen, L.L., Lund, R., Osler, M., Christensen, U., Avlund, K. (2012). Lifetime Occupational Physical Activity and Musculoskeletal Aging in Middle-Aged Men and Women in Denmark: Retrospective Cohort Study Protocol and Methods. JMIR Res Protoc, 1(2):e7.

Napoles-Springer, A.M., Santoyo-Olsson, J., O'Brien, H., Stewart, A.L. (2006). Using cognitive interviews to develop surveys in diverse populations. Med. Care, 44, 21-30.

Quinn, M.M., Sembajwe, G., Stoddard, A.M., Kriebel, D., Krieger, N., Sorensen, G., Hartman, C., Naishadham, D., Barbeau, E.M. (2007). Social disparities in the burden of occupational exposures: results of a cross-sectional study. Am. J Ind. Med., 50, 861-875.

Sembajwe, G., Quinn, M., Kriebel, D., Stoddard, A., Krieger, N., Barbeau, E. (2010). The influence of sociodemographic characteristics on agreement between self-reports and expert exposure assessments. Am. J Ind. Med., 53, 1019-31.

Stock, S.R., Fernandes, R., Delisle, A., Vezina, N. (2005). Reproducibility and validity of workers' self-reports of physical work demands. Scand. J Work Environ. Health, 31, 409-437.

Streiner, D.L., Norman, G.R. (2008). Health measurement scales, a practical guide to their development and use. Oxford, Oxford University Press.

Torgen, M., Winkel, J., Alfredsson, L., Kilbom, A. (1999). Evaluation of questionnaire-based information on previous physical work loads. Stockholm MUSIC 1 Study Group. Musculoskeletal Intervention Center. Scand. J Work Environ. Health, 25, 246-254.

Tourangeau, R., Rips, L.J., Rasinski, K. (2000). The psychology of survey response. New York, Cambridge University Press.

Viikari-Juntura, E., Rauas, S., Martikainen, R., Kuosma, E., Riihimaki, H., Takala, E.P., Saarenmaa, K. (1996). Validity of self-reported physical work load in epidemiologic studies on musculoskeletal disorders. Scand. J Work Environ. Health, 22, 251-259.

Watt, T., Rasmussen, A.K., Groenvold, M., Bjorner, J.B., Watt, S.H., Bonnema, S.J., Hegedus, L., Feldt-Rasmussen, U. (2008). Improving a newly developed patient-reported outcome for thyroid patients, using cognitive interviewing. Qual. Life Res., 17, 1009-1017.

White, E., Armstron, B.K., Saracci, R. (2008). Principles of exposure measurement in epidemiology: collecting, evaluating and improving measures of disease risk factors. Oxford University Press.

Wiktorin, C., Karlqvist, L., Winkel, J. (1993). Validity of self-reported exposures to work postures and manual materials handling. Stockholm MUSIC I Study Group. Scand. J Work Environ. Health, 19, 208-214.

Willis, G.B. (2005). Cognitive interviewing. A tool for improving questionnaire design. Sage Publications Inc..

APPENDIX

Appendix 1

Question 32 in Danish from the questionnaire and an English translation. Groups of OPA below.

32. Når du ser tilbage på hele dit arbejdsliv:
   (Du må gerne svare i mere end én kategori)

a) Hvor mange år af dit arbejdsliv har du haft <u>mest</u> stillesiddende arbejde, som ikke kræver fysisk anstrengelse?
b) Hvor mange år af dit arbejdsliv har du haft <u>mest</u> stående eller gående arbejde, som ikke kræver fysisk anstrengelse?
c) Hvor mange år af dit arbejdsliv har du haft <u>mest</u> stående eller gående arbejde med en del løfte- eller bærearbejde?
d) Hvor mange år af dit arbejdsliv har du haft <u>mest</u> tungt eller hurtigt arbejde, som er fysisk anstrengende?

32. Looking back on your entire working life:
   (You may answer in more than one category)

a) For how many years of your working life have you had <u>mostly</u> sedentary work without physical strain?
b) For how many years of your working life have you had <u>mostly</u> standing or walking work without major physical activity?
c) For how many years of your working life have you worked <u>mostly</u> standing or walking with some lifting and carrying?
d) For how many years of your working life have you had to work <u>mostly</u> at a high speed, with heavy and physically demanding work?

Categories of occupational physical activity (OPA) used in the study according to question 32:

a) : Sedentary work
b) : Standing and walking
c) : Moderate OPA
d) : High OPA

Appendix 2

The OPA- index

We created an arbitrary index of occupational physical activity (OPA), based on answers about lifetime OPA in the questionnaire. It was necessary to construct an index eliminating overlap, while many respondents had written 40 years of employment in more than one category. Filling out both category c) and d) was interpreted as the job having included elements of both moderate and high physical activity.

In the index, it is assumed that category d) has OPA 70% of the time and c) has 20% OPA. Having a job that primarily includes standing and walking/b) has OPA 10% of the time, and a job mostly sedentary has 0% of OPA. The index is a summation of OPA in years divided with the total duration of employment.

An example:

a) A bricklayer working for 40 years in the same job categorized as d):
   OPA-index: (0.7*40)/40= 0.7.
b) A nurse working in a clinical department for 20 years categorized as c), and in an administrative job for 10 years categorized as a).
   OPA-index: (0.2*20+0*10)/(10+20)=0.13.
c) A confectioner working for 30 years wrote "30 years" in b), c) and d) in the questionnaire. Therefore, total years of exposure are 90 years.
   OPA-index: (0.7*30+0.2*30+0.1*30)/(30+30+30)=0.33.