The Literature Review of Tools and Methods for Evaluating Physical Activity Program at Workplace

Ni Made Hegard Sukmawati
Faculty of Medicine and Health Sciences, Warmadewa University
Email: sukmawati.hegard@gmail.com

Abstract

Background: workplace is an important setting for health promotion on physical activity programs. Establishing PA intervention in workplace is believed to give more substantial impact than in many other communities. However, the effectiveness of the physical activity intervention in occupational space has been questioned. Physical activity is a convoluted and varying behavior, and the ability to measure the association between physical activity and chronic diseases is strongly dependent on the validity of the tools. Objective: The goal of this study is to review the type of measurements, reliability and validity of instruments/methods of measuring occupational PA. Method: Relevant peer-reviewed journals were collected from two electronic database, Medline and PubMed, using advanced search strategy and eligibility criteria. Results: The searching strategy has generated 413 articles in total and the criteria have narrowed the result to 12 relevant articles. The BRFSS, the IPAQ-L, the MOSPA-Q, the OSPAQ and the OPAQ have been proven to have good reliability. However, these questionnaires showed poor to moderate criterion validity, thus objective measures of occupational PA, such as accelerometer, is still the best option.

Keywords: physical activity, workplace, occupational, validity, reliability

INTRODUCTION

Physical activity (PA) is believed to be an important modifiable risk factor for several chronic disease. Studies have revealed that the lack of physical activity leads to several chronic diseases, such as diabetes, heart disease, stroke, and cancer. It is also believed that routine PA could provide a number of health benefits including reduced risk of morbidity as well as reduced risk of premature mortality. Most of the adult population spend...
their time in workplace, sedentarily without doing significant physical activity (PA). And it is not until recently, workplace is more recognized as an important setting for health promotion to endorse physical activity programs. Establishing PA intervention in workplace is believed to give more substantial impact than in many other communities\(^7\). According to World Health Organization\(^6\), workplace offers several advantages in which considerable number of the working population can be reached and multiple levels of influence on behavior can be targeted. However, the effectiveness of the physical activity intervention in occupational space has been questioned.

There are limited information on the effects of occupational PA on health, as only few studies have adequately examined the outcome of occupational PA with the aim of assessing the health benefits. Available findings have shown discrepancy; where some studies suggest the protective effects of occupational PA against\(^8\)-\(^11\), for example, cardiovascular disease, while others observe no or negative relationship\(^12\)-\(^14\).

Some studies have suggested that the health benefits of PA might vary for different domains of PA\(^11,13,15\). Holtermann, Mortensen\(^16\) has also explained contrasting cardiovascular effects of PA carried out in different domains, such as during work and leisure time.

There are different forms of monitoring and evaluation of a PA program depending on the objectives of the program. And the result of this evaluation should be well communicated to the management/public health workers. Most intervention and epidemiological studies use questionnaires rather than objective measures to monitor and measure PA. The main reason was practicality and feasibility. However, physical activity is a convoluted and varying behavior\(^17\), and the ability to measure the association between physical activity and chronic diseases is strongly dependent on the validity of the tools. Therefore, it is essential to have a reliable measurement instrument as the basis for drawing conclusion regarding to the impact of the program. This study will focus on measurement tools used to monitor and evaluate PA intervention program in workplace setting. The goal of this study is to review the type of measurements, reliability and validity of these instruments/methods; and find optimal methods of measuring occupational PA.

METHODS

Literature Research

Information collected from two electronic database: Medline and PubMed. I searched relevant peer-reviewed journals in PubMed using the following advanced search strategy: (("motor activity"[MeSH Terms] OR ("motor"[All Fields] AND "activity"[All Fields])) OR "motor activity"[All Fields] OR ("physical"[All Fields] AND "activity"[All Fields]) OR "physical activity"[All Fields]) AND (("workplace"[MeSH Terms] OR "workplace"[All Fields]) OR ("workplace"[MeSH Terms] OR "workplace"[All Fields] OR "worksite"[All Fields])) OR occupational[All Fields]) AND ((reliability[All Fields] OR validity [All Fields]) OR accuracy[All Fields]). The similar strategy also applied for Medline using: (physical activity.mp. or Motor Activity/) AND (workplace or worksite or occupational space).mp. AND (Reliability or validity).mp. or accuracy/).

Eligibility Criteria

The searching strategy has generated 413 articles in total (122 Medline and 291 in PubMed). All the articles were sorted out based on the relevance to the keywords and then screened for possible inclusion based on title and abstract. The inclusion criteria were 1) the article should focus on physical activity in workplace setting, 2) the study was reliability and/validity study which contained information on measurement properties, 3) the article should be published in English language, and 4) it was published from 1995 to 2019. Studies that focused on specific population, such as pregnant women or elderly were excluded, as well as studies related to specific disease or symptoms. These criteria have narrowed the result to 12 relevant articles (Appendix...
RESULT AND DISCUSSION

Most of the identified literatures were examining the reliability and validity of tools that used to measure behavior during work, such as walking, lifting, sitting, etc. While, intervention such as spatial configuration, perceived physical-social environmental factors, and travel behavior, were represented only by one article respectively. In this review, the studies on behavior PA (sitting, walking, standing, lifting, etc.) are classified as objective measures, subjective measures, and criterion standards. Objective measures such as motion sensor (accelerometer and pedometer) are frequently used in physical activity studies, especially as comparison tools to measure criterion validity. Castillo-Retamal and Hinckson\textsuperscript{18}, in their review reported that accelerometer can be used to capture light intensity PA that rarely detected by self-report measures. But, it also has drawback on detecting upper body movement while sitting or standing. Generally, objective measures offer good reliability and validity (ICC= 0.80-0.96, 14 days test-retest period; \( r = 0.92-0.96, p<0.001 \)). However, these objective measures are overly expensive for population-based studies.

Subjective measures are the most common tools used to measure factors associated with physical activity. In workplace setting, questionnaires are the first preference, followed by motion sensor, and indirect calorimeter\textsuperscript{18}. Subjective techniques are more preferable because they allow assessment in short period of time for large number of samples. In addition, they offer cheaper techniques and have been shown not to significantly disturb work tasks. Kwak, Proper\textsuperscript{19} have conducted a systematic review over 31 articles. This systematic review assessed reliability and validity of 30 questionnaires in regard of work index/activity score, energy expenditure, and duration of activity\textsuperscript{19}. Four prominent questionnaires in repeatability were identified: the BRFSS, the IPAQ-L, the MOSPA-Q, and the OPAQ (ICC 0.76-0.83). Good repeatability in the work index was shown by the BRFSS, and satisfying repeatability in energy expenditure and duration of activity were shown by the IPAQ-L and the MOSPA-Q. While the OPAQ showed good repeatability on measuring duration of activity. All of these result based on strong level of evidence. And all the four questionnaires are deemed to be feasible to be applied in workplace setting as they are short and take less than five minutes to complete. The BRFSS is a single-items measure that is more preferable for rapid assessment of occupational PA level in surveillance study\textsuperscript{20}. A more inclusive assessment of time spent are offered by the MOSPA-Q and the OPAQ\textsuperscript{21,22}. Both are multiple-items measures with various occupational categorize. The same as the MOSPA-Q, the IPAQ-L was also reliable for measuring both energy expenditure and duration of activity, but this questionnaire is more focus on moderate and vigorous intensity PA, thus it is more suitable for assessing health-enhancing PA. However, in term of validity, none of the questionnaires showed good validity against accelerometer. The TCQ was the only questionnaire that showed moderate objective criterion validity on energy expenditure (r=0.5), but there was no appropriate study on its reliability. While, moderate-to- high subjective criterion validity was shown by TOQ in regard of energy expenditure and duration activity (r= 0.57-0.92).\textsuperscript{19}

Jancey, Tye\textsuperscript{25}, particularly studied measurement properties of the OSPAQ and found strong to moderate reliability (ICC 0.66-0.83), and moderate correlation against accelerometer. Comparison made by Chau, Van Der Ploeg\textsuperscript{26} on the OSPAQ and the MOSPA-Q has shown that the OSPAQ was better than the MOSPA-Q in term of reliability and validity, especially in estimating time spent sitting and standing at work. The OSPAQ was also deemed suitable tools to measure multiple health behaviors in epidemiological studies as it needs shorter time for completion and lower burden.

Criterion standard is another option of measurement tool beside of objective and subjective measures. Criterion standards
include indirect calorimeter, direct observation, and doubly labelled water (DLW). They offer excellent reliability and validity on energy expenditure. Indirect calorimeter works by estimating energy expenditure from VO2 consumption and VCO2 production. The benefit of this measure is it can detect small change in the variable studied. However, this measure is not suitable for population-based studies since it can only possible to be used in small sample over short period of time. Other shortcomings of this measures rely on its cost and difficulty to apply into entire work days. However, Pernold, Tornqvist suggested that indirect calorimeter can be employed for epidemiological study as validation tool.

An alternative tool to measure energy consumption was developed by Bernmark, Forsman. Their study was estimating oxygen consumption by measuring heart rate. But, this study showed poor precision, thus, this alternative criterion standard is not reliable for PA measurement.

Non-behavioral variable such as spatial configuration characteristics and perceived physical-environmental factors at workplace are believed to provide insight on occupant movements. For measuring this type of intervention, Duncan, Rashid has developed a self-report instrument named the Office Environmental and Sitting Scale (OFFESS) to measure influence of spatial configuration on sitting behavior. It is reported that OFFESS scales have good internal consistency (α = 0.7 – 0.86) and also reliability (ICC = 0.7 - 0.87, 3 days test-retest period). This questionnaire is suitable for epidemiological studies since the overall length of the OFFESS is relatively short. The work-site Supportive Environment for Active Living Survey (SEALS) were designed by Blunt and Hallam to measure the influence of perceived physical-social environmental factors at workplace to physical activity. This survey has shown good internal consistency and reliability (α = 0.79 – 0.86, Pearson correlation = 0.73 - 0.96, P<0.005). While, article by Petrunoff, Xu (31) measured travel mode and travel time to predict the moderate-to-vigorous physical activity during travel to workplace using an online survey: www.activetravel.net.au/professionals/tools. This survey has moderate criterion validity (ρ=0.75; Kw=0.62, P<0.0001) and good reliability (ρ = 0.83; Kw = 0.82; P<0.0001).

CONCLUSION

Work postures are the most measured variables in occupational PA studies among other predictors of PA level. These variables are commonly measured by subjective measures such as questionnaires. Questionnaires are deemed the most efficient and feasible instruments for estimating various level of physical activity in workplace setting. They offer practicality, and some of them provide good statistical characteristics. The BRFSS, the IPAQ-L, the MOSPA-Q, the OSPAQ and the OPAQ have been proven to have good reliability. However, these questionnaires showed poor to moderate criterion validity, thus objective measures of occupational PA, such as accelerometer, is still the best option.

Limitation

There are three main limitations on this study. One of them is the publication bias, as the search strategy only focus on reliability and validity study which possibly have not identified all studies that measure physical activity at workplace. The number of the source was only 12 articles, 10 of which were single studies with different focus of measurement tools. Because of the small number of journals that met the criteria, the finding of this study may not be generalizable across workplace, occupation, and PA programs. And differences in object measured, type of instruments, and statistical methods that were used, make it difficult to compare the findings of these 12 different studies.

Second limitation of this review is the scoring method that was used to classify the reliability and validity of the instruments. Others might have preferred to use different cut-off point for scoring the data, but in this review, Castillo-Retamal and Hinckson scoring was used to assessed reliability and validity data (good: >0.75, moderate: 0.5-0.75, and poor : <0.5). An-
other potential limitation could be in the judgment on the feasibility of implementing the instruments for occupational PA evaluation, especially for large epidemiological studies. Only 5 out of 12 included studies mentioned the feasibility of the instruments application for large epidemiological studies. Thus, the feasibility criteria mostly based on these 5 articles and my own judgment.

REFERENCES
1. Cust AE, Armstrong BK, Friedenreich CM, Slimani N, Bauman A. Physical activity and endometrial cancer risk: a review of the current evidence, biologic mechanisms and the quality of physical activity assessment methods. Cancer Causes & Control. 2007;18(3):243-58.
2. Friedenreich CM. Physical activity and cancer prevention: from observational to intervention research. Cancer Epidemiology and Prevention Biomarkers. 2001;10(4):287-301.
3. Monninkhof EM, Elias SG, Vlems FA, van der Tweel I, Schuit AJ, Voskuil DW, et al. Physical activity and breast cancer: a systematic review. Epidemiology. 2007:137-57.
4. Vainio H, Bianchini F. Weight control and physical activity: iarc; 2002.
5. Warburton DE, Nicol CW, Bredin SS. Health benefits of physical activity: the evidence. Cmaj. 2006;174 (6):801-9.
6. World Health Organization. The workplace as a setting for interventions to improve diet and promote physical activity. Geneva: WHO Press; 2008.
7. Dishman RK, Oldenburg B, O’Neal H, Shephard RJ. Worksite physical activity interventions. Am J Prev Med. 1998;15(4):344-61.
8. Bauman A, Bellew B, Vita P, Brown W, Owen N. Getting Australia active : towards better practice for the promotion of physical activity. Melbourne: National Public Health Partnership; 2002.
9. Barengo NIC, Hu G, Lakka TA, Pekkarinen H, Nissinen A, Tuomilehto J. Low physical activity as a predictor for total and cardiovascular disease mortality in middle-aged men and women in Finland. European Heart Journal. 2004;25(24):2204-11.
10. Greendale GA, Bodin-Dunn L, Ingles S, Haile R, Barrett-Connor E. Leisure, home, and occupational physical activity and cardiovascular risk factors in postmenopausal women: The Postmenopausal Estrogens/Progestins Intervention (PEPI) Study. Archives of internal medicine. 1996;156(4):418-24.
11. Hu G, Jousilahti P, Borodulin K, Barengo NC, Lakka TA, Nissinen A, et al. Occupational, commuting and leisure-time physical activity in relation to coronary heart disease among middle-aged Finnish men and women. Atherosclerosis. 2007;194(2):490-7.
12. Krause N, Brand RJ, Kaplan GA, Kauhanen J, Malla S, Tuomainen T-P, et al. Occupational physical activity, energy expenditure and 11-year progression of carotid atherosclerosis. Scandinavian journal of work, environment & health. 2007:405-24.
13. Kristal-Boneh E, Harari G, Melamed S, Froom P. Association of Physical Activity at Work With Mortality in Israeli Industrial Employees:: The CORDIS Study. Journal of occupational and environmental medicine. 2000;42(2):127-35.
14. Menotti A, Seccareccia F. Physical activity at work and job responsibility as risk factors for fatal coronary heart disease and other causes of death. Journal of Epidemiology & Community Health. 1985;39(4):325-9.
15. Andersen LB, Schnohr P, Schroll M, Hein HO. All-cause mortality associated with physical activity during leisure time, work, sports, and cycling to work. Archives of internal medicine. 2000;160(11):1621-8.
16. Holtermann A, Mortensen OS, Burr H, Søgaard K, Gyntelberg F, Suadicani P. The interplay between physical activity at work and during leisure time-risk of ischemic heart disease and all-cause mortality in middle-aged Caucasian men. Scandinavian journal of work, environment & health. 2009;466-74.

17. Levin S, Jacobs Jr DR, Ainsworth BE, Richardson MT, Leon AS. Intra-individual variation and estimates of usual physical activity. Annals of epidemiology. 1999;9(8):481-8.

18. Castillo-Retamal M, Hinckson EA. Measuring physical activity and sedentary behaviour at work: a review. Work. 2011;40(4):345-57.

19. Kwak L, Proper KI, Hagstromer M, Sjostrom M. The repeatability and validity of questionnaires assessing occupational physical activity—a systematic review. Scand J Work Environ Health. 2011;37(1):6-29.

20. Evenson KR, McGinn AP. Test-retest reliability of adult surveillance measures for physical activity and inactivity. American journal of preventive medicine. 2005;28(5):470-8.

21. Reis JP, DuBose KD, Ainsworth BE, Macera CA, Yore MM. Reliability and validity of the occupational physical activity questionnaire. Medicine and science in sports and exercise. 2005;37(12):2075-83.

22. Roeykens J, Rogers R, Meeusen R, Magnus L, Borns J, DE MEIRLEIR K. Validity and reliability in a Flemish population of the WHO-MONICA Optional Study of Physical Activity Questionnaire. Medicine & Science in Sports & Exercise. 1998;30(7):1071-5.

23. Campbell L, Pannett B, Egger P, Cooper C, Coggon D. Validity of a questionnaire for assessing occupational activities. Am J Ind Med. 1997;31(4):422-6.

24. Kurtze N, Rangul V, Hustvedt BE, Flanders WD. Reliability and validity of self-reported physical activity in the Nord-Trondelag Health Study (HUNT 2). Eur J Epidemiol. 2007;22(6):379-87.

25. Jancey J, Tye M, McGann S, Blackford K, Lee AH. Application of the Occupational Sitting and Physical Activity Questionnaire (OSPAQ) to office based workers. BMC Public Health. 2014;14:762.

26. Chau JY, Van Der Ploeg HP, Dunn S, Kurko J, Bauman AE. Validity of the occupational sitting and physical activity questionnaire. Med Sci Sports Exerc. 2012;44(1):118-25.

27. Pernold G, Tornqvist EW, Wiktorin C, Mortimer M, Karlsson E, Kilbom A, et al. Validity of occupational energy expenditure assessed by interview. AIHA J (Fairfax, Va). 2002;63(1):29-33.

28. Bernmark E, Forsman M, Pernold G, Wiktorin C. Validity of heart-rate based measurements of oxygen consumption during work with light and moderate physical activity. Work. 2012;41 Suppl 1:5475-6.

29. Duncan MJ, Rashid M, Vandelanotte C, Cutumisu N, Plotnikoff RC. Development and reliability testing of a self-report instrument to measure the office layout as a correlate of occupational sitting. The international journal of behavioral nutrition and physical activity. 2013;10:16.

30. Blunt GH, Hallam JS. The worksite supportive environments for active living survey: development and psychometric properties. Am J Health Promot. 2010;25(1):48-57.

31. Petrunoff NA, Xu H, Rissel C, Wen LM, van der Ploeg HP. Measuring workplace travel behaviour: validity and reliability of survey questions. J Environ Public Health. 2013;2013:423035.
Appendix 1. Search Result

**PUBMED**

| Query                                                                 | Items found |
|-----------------------------------------------------------------------|-------------|
| 1  Search physical activity                                           | 391480      |
| 2  Search ((workplace) OR worksite) OR occupational                   | 291051      |
| 3  Search ((reliability) OR validity) OR accuracy                     | 375801      |
| 4  Search ((physical activity) AND ((workplace) OR worksite) OR occupational) AND ((reliability) OR validity) OR accuracy | 291         |
| sort by: [relevance]                                                  |             |

**MEDLINE**

| No | Query                                                                 | Items found |
|----|-----------------------------------------------------------------------|-------------|
| 1  | physical activity.mp. or exp Motor Activity/                          | 291527      |
|    | (occupational or workplace or worksite).mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier] | 251404      |
| 3  | (Reliability or validity).mp. or accuracy/ [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier] | 176583      |
| 4  | 1 and 2 and 3                                                        | 153         |
| 5  | limit 5 to (english language and yr="1995 - 2019")                   | 122         |
## Appendix 2: Data Extraction Table

| Literature | Description/Setting | Sample size (N) | Object of measurement | Type of Instrument | Measurement method |
|------------|---------------------|----------------|-----------------------|--------------------|--------------------|
| Benmark E, et al, 2012 | Workplace (unspecified) | 23 participants | Light to moderate PA (91 type of work task) | Heart rate measurement (Validation tool: Oxygen consumption) | Estimated oxygen consumption based on heart rate |
| Blunt, G. H., & Hallen, J. S., 2010 | University worksite | 683 participants | Perceived physical & social environmental factors influence PA (e.g. distance to facilities) | Self-report: The Worksite Supportive Environments for Active Living Survey | Cronbach’s α & Test-retest |
| Campbell, L, et al, 1997 | Review: Measuring PA and Sedentary behavior at Work | 11 articles | Lifting activities and Non-lifting activities (walking, squatting, kneeling etc) | Self-report: Questionnaire (Validation tools: direct observation and pedometer) | % Agreement |
| Castillo-Retamal, M, et al, 2011 | Office-based occupation | 307 participants | Relationship of spatial configuration on office work with duration and frequency of occupational sitting | Questionnaires: OSPAQ and MOSPA-Q (Validation tool: accelerometer) | Test-retest (intraclass correlation coefficient) & Validity (Spearman’s ρ) |
| Chau, J. Y., 2012 | 2 medium-sized workplace | 99 participants | Occupational sitting, standing, and walking | Questionnaires: OSPAQ and MOSPA-Q (Validation tool: accelerometer) | Test-retest (intraclass correlation coefficient) & Validity (Spearman’s ρ) |
| Duncan, M. J., et al, 2013 | Office-based occupation | 307 participants | Relationship of spatial configuration on office work with duration and frequency of occupational sitting | Self-report: Online survey | Cronbach’s α to evaluate internal consistency and test re-test reliability (retest ICC) |
| Jancey, J., et al, 2014 | University Offices | 99 participants | The effects for setting, standing and walking | Self-report: The Occupational Sitting and Physical Activity Questionnaire (Validation tools: OSPAQ & Accelerometer) | Test-retest Study: Reliability was assessed by paired t-test and intraclass correlations (ICC) via a two-way mixed model based on absolute agreement. Difference and agreement were measured by comparison of mean self-reported data with accelerometer data using the Pearson’s correlation coefficient and Bland-Altman plots |
| Kwak, L, et al, 2011 | Systematic Review | 31 papers: 30 types of questionnaire | Occupational PA | 30 types of questionnaire (Validation tools: accelerometer, maximal oxygen uptake, body composition measures) | Reproducibility: Test-retest & Intraclass correlation coefficients and weighted Cohen's kappa. Validity: Correlation coefficient, plus sensitivity specificity, Pearson’s product moment or Spearman’s rank |
| Pernold, G., et al, 2002 | Various occupations | 27 participants | 1)Posture: Sitting, standing, and walking; 2)Intensity of PA | Test-oriented interview (Validation tools: oxygen consumption and observation) | The product’s moment correlation coefficient |
| Pernovsk, N. A., et al, 2013 | Hospital setting | 65 participants | Agreement between tools and moderate-to-vigorous PA | Self-reported online survey and four-day travel diary & Accelerometer | Reliability & Validity: kappa statistic, Spearman’s correlation coefficient, t-test |
| Reis, J. P., et al, 2005 | Broad range of occupations | 41 participants | Sitting, standing, walking, and heavy labor activities | Self-report: The Occupational Physical Activity Questionnaire (OPAQ) (Validation tools: the Behavioral Risk Factor Surveillance System (BRFSS)) | Test-retest reliability: intraclass correlation coefficient. Validity: Spearman correlation, kappa correlation |
| Kurota, N, et al, 2007 | Occupational setting (not specified) | 108 participants | PA & Energy expenditure | Nordic-Trialling Health Study (HUNT 2) questionnaire (Validation tools: VO2max, International Physical Activity Questionnaire (IPAQ) & AcclReg) | Reliability: Spearman’s coefficient |

DOI: [https://doi.org/10.22225/wmj.5.2.2374.85-93](https://doi.org/10.22225/wmj.5.2.2374.85-93)
Appendix 3. Questionnaire abbreviation

| Questionnaire Abbreviation | Description                                      |
|----------------------------|--------------------------------------------------|
| BRFSS                     | Behavioral Risk Factor Surveillance System       |
| IPAQ-L                    | International Physical Activity Questionnaire-Long |
| MOSPA-Q                   | Monica Optional Study of Physical Activity Questionnaire |
| OPAQ                      | Occupational Physical Activity Questionnaire     |
| OSPAQ                     | Occupational Sitting and Physical Activity Questionnaire |
| TCQ                       | Tecumseh Community Questionnaire                 |
| TOQ                       | Tecumseh Occupational Activity Questionnaire      |