A comparison of self-reported data on disability pension status with data from a nationwide administrative register

Ingeborg Hartz[1], Aage Tverdal[2] and Svetlana Skurtveit[2,3]

1) Department of Health and Sports, Hedmark University College, Kirkeveien 47, 2418 Elverum, Norway
2) Division of Epidemiology, Norwegian Institute of Public Health, P.O. Box 4404 Nydalen, 0403 Oslo, Norway
3) Norwegian Centre for Addiction Research, University of Oslo, Norway

Correspondence: Ingeborg Hartz, E-mail: ingeborg.hartz@hihm.no   Telephone: +47 97703853   Telefax: +47 62430300

ABSTRACT

Objectives: Validation studies of self-reported disability pension status have been scarce. The objective of this study was to estimate the sensitivity and specificity, as well as positive and negative predictive values of self-reported disability pension status using an official administrative register as reference standard.
Methods: Data from Cohort of Norway (CONOR) surveys conducted in 2001 in the three Norwegian counties Oslo (HUBRO), Hedmark and Oppland (OPPHED) are included in this study, altogether 17,244 individuals. At the time of investigation, the subjects included in our study-population were aged 30-31, 40-41, 45-47 and 59-61 years. Self-reported data on disability pension status was compared with data from the nationwide population and housing census in Norway (Statistics Norway), performed November 3rd 2001. Data were linked using the unique 11-digit identification number, assigned to all individuals living in Norway.
Results: Sensitivity and specificity for self-reported questions on disability pension were 97.6% (95% CI 91.1-94.1) and 96.8% (96.5-97.1). Positive and negative predictive values were 70.1% (67.9-72.3) and 99.8% (99.7-99.9). Validity measures in a subpopulation, those surveyed in October-November 2001, were almost identical as for the total study-population surveyed in 2001.
Conclusion: The sensitivity of self-reported disability pension status was high. The positive predictive value was lower which may reflect a low prevalence.

INTRODUCTION

The proportion of Norwegians on disability pension has doubled since the 1980s, which pose both economical and societal challenges to the society[1]. The phenomenon has raised the interest of research into the sickness absence and disability pension field.

Questionnaires have the advantage of collecting information on sociodemographic variables, as well as a variety of information on health status, health risk factors, and other variables. Thus, self-reported information is commonly used when sickness absence and disability pension status is evaluated as either exposure or outcome variable in epidemiological studies[2-5]. A generally recognized concern by the use of self-reported data, however, is that the presence of either differential or non-differential misclassification may lead to biased results of the studied associations. In general, key components of measurement error for self-reported data would be related to question comprehension and interpretation, as well as recall abilities by the respondents[6]. Thus, the agreement between retrospectively collected information on self-reported sick-leave days and recorded information on the number of sick-leave days seems to be good, except for longer spells of sickness absence which may be inaccurately recalled[7-9]. Being granted a disability pension is a con-

crete event, presumably contributing to an improved accuracy of self-reported disability pension status, as compared to self-reported sickness absence. However, whereas validation studies on self-reported sickness absence exist, studies on self-reported disability pension status have been scarce[7-9]. The objective of the present study was to estimate measures of validity of self-reported disability pension status among participants in Norwegian health surveys using a national administrative register on disability pension status as the reference standard.

METHODS

Data from Cohort of Norway (CONOR) was included in this study, which is a collection of health data and blood samples from several Norwegian health surveys described elsewhere[10,11]. CONOR data from population based surveys conducted in 2000-01 in the three Norwegian counties Oslo (HUBRO), Hedmark and Oppland (OPPHED) are included in this study, which cover both rural and urban regions. All subjects in selected age cohorts were invited to participate in the respective surveys. At the time of investigation, the subjects were aged 30-31, 40-41, 45-47, 59-61 and 75-76 years. In these age cohorts 58,660 were invited in the HUBRO and 22,327 were invited in the OPPHED

study. 21,361 (36.4%) participated in HUBRO and 12,263 (54.9%) participated in OPPHED and agreed to the storage of blood samples and data for research purposes and agreed to have their data linked to other health registers, altogether 33,624.

Our analysis are restricted to data on participants aged 61 years and younger attending surveys conducted in 2001, thus including people below the early age of retirement (contractual pension) only (n=19,125) (Figure 1). Further, a sub-group analysis was performed, including participants reporting DP status close to the date when information on DP status was derived from the administrative register. See description on reference standard for further details. Persons who did not answer the question on disability pension status in the health surveys (n=1,833) were excluded, as well as individuals with missing information in the administrative register on disability pension status (n=46) resulting in a study-population of 17,244 people. Among these, 1093 individuals who attended in October-November 2001 were included in a sub-group analysis.

The participants were asked a wide range of questions covering information on sociodemographics, health status, health risk factors, physical activity and others. The English version of the questionnaire, as well as further details of the surveys, is available at the home page of the Norwegian Institute of Public Health[11]. Questions on receipt of various benefits, disability pension and others, were as follows; "Do you receive any of the following?" with "disability pension?" (answering categories "yes" and "no") as a separate question to be ticked off.

Reference standard

The variable "current activity status" from the nationwide population and housing census in Norway (Statistics Norway), performed November 3rd 2001, was used as a reference standard[12]. This variable is based on information from several administrative registers, and contains information on all citizens in Norway. The variable categorizes Norwegian individuals according to their labour force participation; employed, unemployed, in education, or receipt of national insurance benefit (disability-, retirement- or survivor pension). Data from the health surveys and the administrative register were linked based on the unique 11-digit identification number, assigned to all individuals living in Norway.

Statistical analysis

Validity (accuracy) analysis was performed using an administrative register as the "gold standard".

Two measures of validity, sensitivity and specificity, were calculated for binary measure of disability pension status.

Sensitivity was calculated as the proportion reporting "yes" to the disability pension question, among participants registered as receivers of a disability pension in the official register. Specificity was calculated as the proportion reporting "no" to the disability pension question, among participants registered as non-receivers of a disability pension in the official register.

In addition, prediction values of positive and negative classification were calculated. The positive predictive value (PPV) was calculated as the proportion of those registered as receivers of a disability pension among all participants reporting "yes" to the disability pension question. The PPV is a value representing the extent to which the subjects who answer "yes" actually do receive a disability pension according to the gold standard.

The negative predictive value (NPV) was calculated as the proportion of those registered as non-receivers of a disability pension in the official register, among all participants answering "no" to the question on disability pension. Thus, the NPV is a value representing the extent to which those answering "no" actually are true non-receivers of a disability pension according to the gold standard.

All analyses were done using SPSS 17.0 for Windows. The study protocol was assessed by the Regional Committee for Medical Research Ethics. The record linkage was approved by the Norwegian Data Inspectorate.
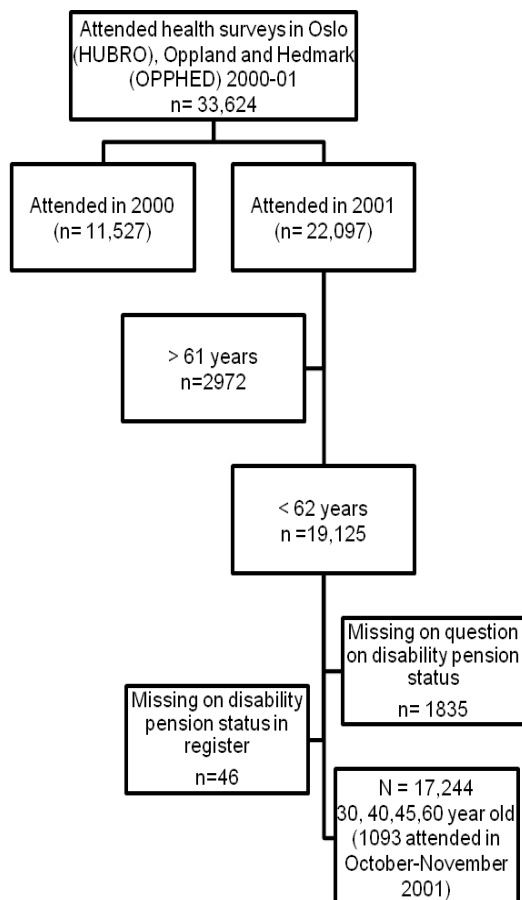

Figure 1. Flow chart for the study population.

RESULTS

Sensitivity and specificity for self-reported questions on disability pension were 97.6% (95% CI 91.1-94.1) and 96.8% (96.5-97.1). Positive and negative predictive values were 70.1% (67.9-72.3) and 99.8% (99.7-99.9), respectively. See Table 1 for further details. Whereas sensitivities, specificities and negative predictive values were almost identical across gender and age-groups, positive predictive values increased with narrower confidence interval in older age-groups, in particular among men. Hence, the PPV of self-reported disability pension status increased from 56% (37.9-73.2) among 31-32 year old men to 76% (71.7-79.8) in 60-61 year old men. Parallel, the PPV increased from 65.9% (50.0-79.1) among the 31-32 year old to 74.8% (71.2-78.2) among the 60-61 year old women.

Validity measures in a subpopulation, those surveyed in October-November 2001, were almost identical as for the total study-population surveyed in 2001 (See Table 2 for details). Sensitivity and specificity for self-reported questions on disability pension were 99.9% (95% CI 93.9-99.9) and 95.6% (94.0-96.7). Positive and negative predictive values were 69.9% (61.6-77.0) and 99.9% (99.3-99.9), respectively.

DISCUSSION

We examined the validity of self-reported disability pension status with data from a nationwide administrative register as the reference standard. To our knowledge, no studies are published on validation of self-reported disability pension status. An estimated sensitivity and specificity above 90% indicate that the validity of self-reported disability status is high. The use of a questionnaire seems to be a valid method for correctly identifying the true positives and true negatives; those that according to the official register either do or do not receive a disability pension. A main strength of the study is a complete national register considered to be a true "gold standard", and the use of the unique 11-digit identification number, which ensured valid data linkage between the questionnaire and the reference register.

The participants disability pension status may, however, be misclassified in two ways: (1) if they report reception of a disability pension and have no disability pension granted according to register information (false positives) or (2) if they self-report no disability pension but have had a disability pension granted according to information from register (false negatives). Whereas the estimated NPV of self-reported disability pension status was close to 100%, and thereby the presence of false negatives was low, the overall PPV was 70%. The PPV depend not only on the quality of a test, but also on the prevalence. A low PPV may be due to a low prevalence or it may be due to poor performance of the test[13].

It is noted that the PPV was distinctly higher in the older than in the younger age group and this supports the idea that a low prevalence of disability pension does have an impact on the PPV. Another possible explanation for the somewhat lower PPV could be that individuals reporting to be on a disability pension early in 2001, actually have become reintegrated into the job-market later that year, as registered per November 3rd 2001. However, sub-group analysis of individuals who attended (reported DP status) in October-November 2001, revealed validity measures which were almost identical as in the total study-population surveyed throughout 2001. And, in general, once being granted a disability pension, very few disability pensioners in Norway do become reintegrated into job-market. In this context a reintegration rate at 30% among those reporting to be disability pensioners in the selected health surveys seems a less likely explanation for the somewhat lower PPV. A third explanation may be due to question comprehension and interpretation. In the questionnaire the respondents were ask to tick off (yes/no) whether they received each of the following benefits: sickness benefit, old-age pension, rehabilitation benefit, disability pension, unemployment benefit, social welfare benefits, and social benefit for single parent. It may for example be so that individuals ticking off yes on a disability pension, actually are on rehabilitation benefits which in most cases should precede grant of a disability pension. Unfortunately, the available register-information in this study does not include information on rehabilitation benefits.

Table 1. Distribution of self-reported disability pension (DP) status among participants in Norwegian Health surveys surveyed in 2001, according to disability pension status in a nationwide administrative register (November 3rd 2001).

	Register-information on DP status		
Self-reported DP	Yes	No	Total
Yes	1198 (TP)	511 (FP)	1709
No	29 (FN)	15506 (TN)	15535
Total	1227	16017	17244

TP: True positives, FP: false positives, FN: false negatives,
TN: true negatives

Table 2. Distribution of self-reported disability pension (DP) status among participants in Norwegian Health surveys surveyed in October-November 2001, according to disability pension status in a nationwide administrative register (November 3rd 2001).

	Register-information on DP status		
Self-reported DP	Yes	No	Total
Yes	102 (TP)	44 (FP)	146
No	1 (FN)	946 (TN)	947
Total	103	990	1093

TP: True positives, FP: false positives, FN: false negatives,
TN: true negatives

In conclusion, using self-reported information, as obtained through questionnaires, as a source when disability pension status is evaluated as either exposure or outcome variable in epidemiological studies should take into account the potential overestimation of disability pensioner prevalence.

In general, a nationwide register will provide the most valid information on disability pension status over a longer period. Such a register will on the other hand lack information on major confounding factors for specific outcomes, which have to be collected in questionnaires or other high quality registers if they exist and its allowance is given for linking of databases. In the end, information from registers and surveys are both necessary to carry out a valid epidemiologic study of disability pension status as either exposure or outcome variable.

REFERENCES

1. Norwegian Ministry of Health and Social Care. Official Norwegian Report (In Norwegian: Norges offentlige utredninger (NOU)) 2000:27. Sickness absence and disability pension. Available at: http://www.regjeringen. no/Rpub/NOU/20002000/027/PDFA/NOU200020000027000DDDPDFA.pdf
2. Arber S, Bote M, Meadows R. Gender and socio-economic patterning of self-reported sleep problems in Britain. *Soc Sci Med* 2009; **68**: 281-9.
3. Hartz I, Lundesgaard E, Tverdal A, Skurtveit S. Disability pension is associated with the use of benzodiazepines 20 years later: A prospective study. *Scand J Public Health* 2009; **37**: 320-6.
4. Johnell K, Månsson N, Sundquist J, Melander A, Blennow G, Merlo J. Neighborhood social participation, use of anxiolytic-hypnotic drugs, and women's propensity for disability pension: a multilevel analysis. *Scand J Public Health* 2006; **34**: 41-8.
5. Allebeck P, Mastekaasa A. Swedish Council on Technology Assessment in Health Care (SBU). Chapter 5. Risk factors for sick leave – general studies. *Scand J Public Health* 2004; **63**: 49-108.
6. West S, Strom B, Poole C. Validity of pharmacoepidemiologic drug and diagnosis data,. In: Strom B (ed). Pharmacoepidemiology, 4th edn. Chichester: Wiley, 2005.
7. Ferrie JE, Kivimaki M, Head J, Shipley MJ, Vahtera J, Marmot MG. A comparison of self-reported sickness absence with absences recorded in employers' registers: evidence from the Whitehall II study. *Occup Environ Med* 2005; **62**: 74-9.
8. Voss M, Stark S, Alfredsson L, Vingard E, Josephson M. Comparisons of self-reported and register data on sickness absence among public employees in Sweden. *Occup Environ Med* 2008; **65**: 61-7.
9. Hensing G. Methodological aspects in sickness-absence research. *Scand J Public Health* 2004; **32**: 44-8.
10. Naess O, Sogaard AJ, Arnesen E, Beckstrom AC, Bjertness E, Engeland A, et al. Cohort profile: cohort of Norway (CONOR). *Int J Epidemiol* 2008; **37**: 481-5.
11. CONOR (The Cohort of Norway), Norwegian Institute of Public Health, Oslo, Norway. Available at: http://www.fhi.no/eway/default.aspx?pid=238&trg=MainLeft_5853&MainArea_5811=5853:0:15,2818:1:0:0: ::0:0&MainLeft_5853=5825:56612::1:5857:2:::0:0.
12. Utne H. The Population and Housing Census Handbook 2001. Statistics Norway, Department of Social Statistics, 2005/2. Available at: http://www.ssb.no/english/subjects/02/01/doc_200502_en/doc_200502_en.pdf.
13. Pepe MS. The statistical evaluation of medical tests for classification and prediction. Oxford: Oxford University Press, 2003.