**Author's response to reviews**

**Title:** Sick-leave track record and other potential predictors of a disability pension  
A population based study of 8,218 men and women followed for 16 years

**Authors:**

Thorne Wallman (thorne.wallman@pubcare.uu.se)  
Hans Wedel (hans.wedel@biostat.se)  
Edward Palmer (edward.palmer@forsakringskassan.se)  
Annika Rosengren (annika.rosengren@gu.se)  
Saga Johansson (saga.johansson@astrazeneca.com)  
Henry Eriksson (henryeriksson@swipnet.se)  
Kurt Svärdsudd (kurt.svardsudd@pubcare.uu.se)

**Version:** 4  
**Date:** 25 December 2008

**Author's response to reviews:** see over
Responses to reviewer’s report

We have now considered the additional points raised by reviewer #2 (Sturla Gjesdal) and we have taken the following actions.

1. Comment: There is a huge literature on predictors of future disability pension in the general population and among persons on sick leave. A large part of this literature used data from population health surveys. An important aspect of these studies is a very long follow-up, which is both a strength and a limitation (causal links may be weakened over time). To mention some of these: Krause et al with data from a health survey in Finland (1), studies from Denmark (2), Sweden (3-8) and some very recent studies from Norway, Hunt study (9, 10). The studies have focused on socioeconomic factors, mental health, self-rated health (SRH) and lifestyle factors. There are also a couple of similar studies which have used data from occupational health surveys and even studies with conscription data (young men). It is surprising that this literature is not used as background data since this study have the same design except for the follow up sickness absence data. The authors reply that they are aware of these studies, and these references (ref 16-25) are now included, however without indicating their importance. It is not very helpful to bundle together 20 references (5-25) In the discussion these references are not mentioned even though the present study seems to contradict previous findings from studies with largely similar design.

Response: As the reviewer quite correctly states the are a large number of studies on this subject. Most of them have been reviewed in the Swedish Council on Technology Assessment in Health Care review of the area, which we refer to. We find it impossible to go into detail with all these studies in the manuscript text, but we give a review of the main findings of the group of studies. The discussion text has been revised as far as possible without making the manuscript even longer (the Discussion section is presently 6 pages!).

2. Comment: The authors mention the problem that socioeconomic data are not updated, and that baseline data are used. For some cases survey data are collected before and for some after baseline, which might be somewhat problematic, put this was not my point. Except for education, and smoking the most important determinants from other similar studies based on health surveys, were not included: Workplace factors, BMI, mental health, and not least self-rated health. This is crucial since the study hypothesis is that there were no signs of impaired health initially in the study (no difference in sickness absence) and that the health problems seemed to develop gradually over a long period of time. The previous studies seemed to show that the health status was different at baseline. Can some kind of selection of “healthy workers” in the surveys explain this discrepancy? The weak effect of education is surprising and should be discussed

Response: The variables collected before and after baseline were mainly variables that are fairly stable over time, such as education. We had no access to data on work place factors. Regarding mental health and SRH see item 5-27. Regarding the effect of education see item 5-16 and 5-27. Regarding healthy worker effect: it is unlikely that there is an important healthy worker effect for several reasons. First, the attrition rate was low in all studies. Secondly, there were numerically small differences in sickness absence history between disability pensioners-to-be and others already from the start, significant due to the high statistical power of the study. Moreover, we had a long follow-up time, up to 16 years, with increasing differences over time. Most other studies had larger differences, possibly due to
shorter follow-up time, and thereby baseline measurements closer to end of follow up, which increases baseline differences.

3. Comment: The relation between “aim of the study” and the “study hypothesis” (that persons who later receive a DP have little sickness absence initially, but have more and more sick-leave during the follow up compared to controls) The aim was (p 5 para 2): to test the efficacy of sick leave track record (3 measures) as a determinant of disability pension. However, thereafter the “study hypothesis” is mentioned, and this is the first main finding (that hypothesis was confirmed) To test this hypothesis is not an aim of the study, however the authors claim that “if the record is a determinant …. (this hypothesis is confirmed) It is not easy to see why this is true. Later the authors refer to the case-referent (pilot) study and figure 1 as evidence. If so, the confirmation of the study hypothesis should be specified as an aim of the pilot study (if that study is sufficiently reliable)

Response: The text has been revised, page 4, 2nd paragraph.

4. Comment: To which extent do the sick-leave measures (track record) depend on the last sick leave? What are the results if this last sick leave period is excluded, or censored after 1 year f.i.? The authors again refer to the pilot study, where follow-up was ended 2 years before the actual granting of the DP (p 9 1 para). But was the same done when calculating the 4 sick leave parameters?

Response: Any exclusions or truncations must affect the disability pensioner-to-be group and the reference group equally. Exclusion or truncation of single spells affects the groups unequally, since the mean duration is different in the two groups. We therefore choose to exclude, or right truncate as the term is, from one to ten of the last follow-up years, with no effect on the results. A short note has been inserted in the Methods section, page 9, 2nd paragraph.

5. Comment: 4500 words are quite long, and would not be accepted in other epidemiological/public health journals. Since the authors ask for help in order to shorten the paper I have some suggestions in the specific comments.

Response: We are grateful for the reviewer’s efforts. However, the text volume is of limited importance in an internet based journal as opposed to a printed journal where it is very important. When submitting to BMC Public Health we of course have to follow their instructions to authors, not the instructions of other journals. On the other hand, irrelevant or repetitious text should be avoided. We have therefore carefully considered the reviewer’s comments in detail and taken the following actions.

5-1. Comment: Title: Is determinant the best concept here? Predictor might indicate causality to a lesser extent. Pension not grant.

Response: Determinant has been changed to predictor and grant to pension throughout the manuscript.

5-2. Comment: Abstract background: The first sentence is well known, and of little relevance. Sentence 2 also well known. Previous findings concerning predictors for transition SA>DP main background information.
Response: The text has been revised.

5-3. Comment: Abstract Methods: Material is presented, but not methods, dependent and independent variables (are mentioned in results, might be enough?). It is not clear how the track record variables are calculated.

Response: The text has been revised.

5-4. Comment: Abstract Results: Second sentence not results, belongs to data.

Response: The text has been revised.

5-5. Comment: Background: The topic studied is not the Swedish Social Insurance system, but determinants/predictors of DP. The public health importance, previous findings and aims of this study should be summarised. Delete the three first paragraphs (p 3-4) and last para on p 4 (advantages with universal coverage well known). The most important studies (5-25) should be specified and presented. A short presentation of the Swedish system might be placed in methods or as an appendix.

Response: The text has been revised, pages 3 and 4. The public health impact note was requested by reviewer #1. Regarding previous studies see item 1.

5-6. Comment: P5 Last para: Two aims!

Response: The text has been revised and clarified, page 4, 2nd paragraph.

5-7. Comment: Methods P5: Generally: too long, narrative form is not very adequate. Ex first sentence: the following words might be deleted: “performed”, “for this study”. Second sentence: “details are presented elsewhere” is sufficient. However some of these details are presented later of which much is not relevant (postal questionnaires vs on location). Line 4 the word “briefly” might be deleted.

Response: The text has been revised.

5-8. Comment: P6: Classification of educational level: what was used in this study? Second para: type of absence and extent (%) not relevant Should explain better how the sickness absence measures are calculated for each person.

o Duration
o Interval

o Annual (cumulative?) number of days (1.1.-31.12 seem obvious)
Is there one variable for each year, or is this some kind of summary measures (annual means??). Last sentence “in the study population” and “at any time” should be deleted

Response: The 5-point educational level scale was used in the analyses. To keep the volume of Table 1 down only mandatory education was given to illustrate differences between the pension status and sex groups. The type and extent sentence has been shortened. The sickness absence measures are defined in the methods section. The description of the design of the analyses has been clarified in the statistical considerations section, page 9, 3rd paragraph.
5-9. Comment: P7: **Delete the 3 first paragraphs (not relevant in this study)**

Response: we do not share the reviewer’s view that the three paragraphs would be irrelevant. We have to state where from we got the data, we have to give at least a brief information on the causes of disability pension, since we are accounting for these in the analyses, and we have to give information on where we got the mortality data from and how disability diagnoses and causes of death were coded, since we are using the data in the accounting for underlying disease. However, these paragraphs have been shortened somewhat, page 7.

5-10. Comment: P8: **Third line “those who participated” - did not all subjects participate in a survey?**

Response: No, all individuals in the samples did not participate in the surveys. But the attrition rate was moderate.

5-11. Comment: P9: **Second para: Is this not results?? Figure 1 is adjusted for age ? since cases and referents were matched by age (and gender). Third para: line 4, is “still alive” necessary?**

Response: We initially had this piece of information in the Results section, but were persuaded by our co-workers to move it to the Methods section, since the main results are based on the cohort analysis, and using the pilot study information in Results might cause confusion. The only rationale for the pilot study was to gain information on the growth, or lack of growth, of sickness spell durations over time until disability pension. Differences in growth between pensioners-to-be and referents was a prerequisite for the method of analysis we used.

Yes, in Figure 1 the age and sex and place of residence adjustment was obtained by matching.

Yes, “still alive” is necessary since we adjusted (censored) for non-exposure due to death.

5-12. Comment: P 10: **Second para: Delete 2 first lines after: The analyses .... Last three lines: Cases were followed from baseline to granting of DP, death or end of follow-up. Delete last sentence. Did any cases emigrate?**

Response: We performed several types of analyses, those in the pilot study and those in the main study. The analyses the reviewer is referring to were those in the main study, which must be indicated. Which the censoring events were is essential, the reader should not be forced to guess. The majority of persons leaving the country during the study period did so temporarily, which means that they were still possible to trace in the registers. Less than one percent left the country permanently. The text has been revised, page 9 (see also response to comment 5-8).

5.13. Comment: **Results: Characteristics of the study population – is this results?**

Response: It is customary to provide a description of the study population so that the reader gets a view of what the study population looks like.

5-14. Comment: P11: **Second para, not presented in tables – not very relevant, might be deleted. Third para: Last sentence sufficient?? Data are shown in tables**
Response: Our biostatistician insisted on this paragraph since it is important from an analysis point of view. We choose to present the data as text only not to make the publication even longer. The data on marital status, education, etcetera, are admittedly shown in the table but table contents should be summarised in the text, especially since these data were used as explanatory variables in the analyses.

5-15. Comment: P12: **The strengths of the predictors (!) is reflected by the chi-square values (higher values more important). Is it so simple? Explain - since this is a crucial premise for the study’s conclusions**

Response: Yes, it is that simple! The chi-square value we are talking about is Wald’s chi-square, not to be mixed up with the chi-square used for tests in cross-tables. Wald’s chi-square was developed as a test parameter in logistic and some other regression analysis methods. The p-value in logistic regression is actually based on Wald’s chi-square (always with one degree of freedom), not on the estimate or regression coefficient. It is also the best way of ranking the importance of explanatory variables. The odds ratio is no good ranking variable, since the size of the odds ratio depends not only on effect but also on how the variable is graded. A variable with only two responses, for instance yes/no, gets a higher OR than a variables with many levels, for instance age, all things considered. Say that we have a result where age has OR 1.05 (which it almost always has, irrespective of circumstance, statisticians are even using the OR for age as a test of data accuracy). If age is divided into 10-year age groups (which is usually done by dividing age by 10) the OR becomes 10 times larger, approximately 1.50. For this reason OR cannot be used to rank the importance of explanatory variables, unless all variables have the same grading. Since they seldom have, Wald’s chi-square is our help out of the dark. A short note on this issue has been inserted in the Discussion section, page 17, last paragraph.

5-16. Comment: **Education – how is this variable used here Table one: Mandatory (or basic) education only vs higher. In results this variable decrease the risk of DP, same measure but reversed?? (The low effect of education is perhaps the most surprising finding in this study – even no effect among men in model 2!!)**

Response: To avoid long tabulations only the frequency of the lowest degree of education was presented in the table. In the analyses all five educational levels, as given in the Methods section, were used. Dichotomisation would have meant loss of power due to residual effects within the two classes. The effect of educational level for DP in the model using sickness spell duration as sick leave measure, meant a 18-24% risk decrease from one educational level to the next, and a 13-16% risk decrease per level in the model using cumulative annual sick leave days as measure. All things considered, there is thus a preventive effect of education, but it is a weak factor as compared with sick leave measures, age, and calendar time. A short note on the ranking of impact of education and the other factors in this and other studies has been inserted in the Discussion section, page 17, last paragraph.

5-17. Comment: **P13 Para 1: Figure 3 crude results from “pilot study” – is this “no difference” a robust finding of the study? Delete second sentence here. I wonder whether the case referent study is a valid epidemiological study – pilot usually means “quick and dirty”**

Response: Yes, it is a robust finding. The testing was done using the cohort design, but the
illustration was based on the pilot study since it gave a better possibility to actually see the
time sequence for the various disability pension subgroups. The text has been clarified.

5-18. Comment: **Para 2: This is a crucial point, since one obviously looks at the very long
absences, typically for the Swedish system Duration > 390 days and > 480 days, even 730
and 830 days (2.5 year) Discussion – might be structured (ref BMJ)**

Response: The Discussion text has been clarified, page 18, 3rd paragraph.

5-19. Comment: **P 14: Para 1 Main findings: First two sentences, explain! Finding from
pilot study??**

Response: The text has been revised, page 13, 2nd paragraph.

5-20. Comment: **Para 2 start of “strengths of study”, some of this is said before**

Response: The text has been shortened somewhat.

5-21. Comment: **Para 3 back to main findings? Appropriate methods were used (main
finding?)**

Response: As indicated in the text it is necessary to check the prerequisites for the methods
used. Logistic regression was found to be an appropriate method in the main study since the
pilot study indicated that sick leave measures grew constantly over time, and not only during
the last few years before disability pension. For this reason the pilot study was essential in
determining these prerequisites. It was in no way “quick and dirty”, it took more than a year
to perform.

5-22. Comment: **P 15 Para 2: More strengths (advantages) also some repetitions**

Response: Yes, a few words have appeared previously in the text, but we decided to keep
them to make the advantage section more comprehensive with these words than if they were
deleted.

5-23. Comment: **Para 3: Limitation? Might use one sentence elsewhere: 75% of
population was covered**

Response: A note on advantages and limitations is usually required in the Discussion section.
It would be difficult to understand what we mean by only mention the 75% elsewhere. The
main message of this paragraph was that northern Sweden was not covered, “which might be
different from the rest of the country”. The difference, which we did not want to spell out, is
that sick leave, disability pension and other forms of community support are much more
accepted as a mean of subsistence by the general public in the north than in the rest of the
country.

5-24. Comment: **Para 4: More on the Swedish system (also strength of study) – universal
coverage is the main point here.**

Response: The description of the system in previous sections has been shortened or deleted.
We decided to keep this paragraph for two reasons. First, foreign readers, especially
Americans, assume that our social insurance system is similar to that in their country, and secondly, since it contains essential information on coverage and, most importantly, the lack of right truncation (limitation of time on sick leave), which most other studies are affected by. Truncation almost always creates bias in the analyses.

5-25. Comment: **P16 Para 1**: The very long sick-leave spells in Sweden are now also mentioned as an advantage “from a scientific point of view”

Response: See item 5-24.

5-26. Comment: **Para 3: possible bias 1> small and Para 4 possible bias 2> small and P17 Para 1: possible bias 3, not important and Para 2 possible bias 4 not important. Those 4 possible biases are refuted, and this might be done in much less space (nearly 2 pages)

Response: We have discussed these comments thoroughly in the author group. Our statistician very decisively claims that the discussion on possible sources of bias is one of the most important parts of the discussion section. The professional reader must be allowed to assess the credibility of the results. In addition, reviewer #1 had comments in the opposite direction, not less but more discussion on potential bias.

5-27. Comment: **Are there no real limitations? Should be a separate subsection In my view:** Lack of baseline health variables (SRH, mental health a.s.f) and Peculiarities of the Swedish system I would like a discussion of the weak effect of SES – educational level

Response: There is a paragraph on study limitations in page 14, 3rd paragraph. Regarding lack of baseline variables: we had access to self-rated health. When checking the results by inclusion of a seven-grade self-rated health variable in the final model the results were unchanged. However, self-rated health is a poor measure of presence of illness or disease but an excellent measure of wellbeing. We had access to much better measures of disease in the form of all hospital diagnoses from 1971 and onwards and these were used in the analyses of impact of underlying disease. Mental health was part of these data. Regarding text revision see item 5-15.

5-28. Comment: **Para 3: Start of comparison with previous studies. Ref 29 deals with duration of a single SA period, and DP 3 years later. Refs 4,6-7, 10,12-13,30,33 are summarised, if none are mentioned might be even shorter**

Response: See item 1.

5-29. Comment: **P18 Para 2: Again “numerous studies”, none is specified! Which are the most important – Månsson et al?? and NB very important variables are not mentioned: SRH! Mental health variables! Mykletun et al. The last two sentences are mentioned before (and later)**

Response: See item 1. Regarding SRH and mental health see item 5-27.

5-30 Comments: **Para 3: DP diagnoses no effect, main finding? Robust? Ref 6 deals with sickness absence diagnoses, not DP diagnoses**
Response: Regarding DP diagnoses see item 5-17. The comment on ref 6 is of course correct. Thank you! A note that sickness absence diagnoses were used has been added.

5-31. Comment: Para 4 Implications: Might hold only for Sweden, since the probabilities are estimated according to very long spells, since most DP cases are granted after several years on sick-leave. In other systems/countries a similar measurement might be “time after latest work day when sickness absent” but “sickness absence track record” might be useful only (or mostly) in Sweden

Response: The text has been revised, page 18, 3rd paragraph.

5-32. Comment: P19 Conclusions: Exactly the same words have been said before

Response: Yes, but this is our summary and conclusion.