Do disability pensioners have a higher mortality rate than non-pensioners? Adjusting for potential confounding: A commentary on Hult, Stattin, Janlert and Järvholm

Thorne Wallman a,b,*, Kurt Svärdsudd a

a Department of Public Health and Caring Sciences, Family Medicine and Clinical Epidemiology Section, Uppsala University, Box 524, Uppsala, Sweden
b R&D Centre/Centre for Clinical Research, Sörmland County Council, Eskilstuna, Sweden

During the past decades a number of studies have shown a higher mortality among disability pensioners than among non-pensioners (referents) of the same age and sex (Gjesdal, Haug, Ringdal, Maeland, Hagberg, Røraas, et al., 2009; Gjesdal, Svedberg, Hagberg, & Alexanderson, 2009; Gjesdal, Maeland, Svedberg, Hagberg, & Alexanderson, 2008; Karlsson, Carstensen, Gjesdal, & Alexanderson, 2007; Wallman, Wedel, Johansson, Rosengren, Welin, et al., 2006; Ahs, & Westerling, 2006; Medhus, & Kristenson, 2001; Gubéran, & Usel, 1998; Kaprio, Sarna, Fogelholm, & Koskenvuo, 1996; Hasle, Jeune, & Skytte, 1988; Jeune, 1982; Damlund, Gøth, Hasle, Jeune, & Munk, 1982). At face value this is not very surprising, since the overwhelming majority of disability pensioners have an underlying disease or disorder severe enough to cause health problems. Referents also may have similar diseases or disorders, but usually of less severity.

The “over-mortality” among disability pensioners as compared to referents per se is by now a fairly universally accepted finding. The main dispute on this issue has been focussed on whether the “over-mortality” is a natural” one, i.e., caused by the underlying disease, or whether it is caused by other factors, or even by the disability pension itself, or its potential psycho-social consequences. These include loss of the job related part of the social network, and transition from being a self-sufficient person to becoming an ill person dependent on financial governmental support (Ockander & Timpka, 2001). In three recent studies by Gjesdal et al. (2009); Karlsson et al. (2007); Wallman et al. (2006) the over-mortality has been adjusted for the effects of underlying disease through state-of-the-art statistical analysis. All three studies concluded that underlying diseases could only partially explain the over-mortality.

In their article “Timing of retirement and mortality – A cohort study of Swedish construction workers” in this issue of Social Science & Medicine, Hult, Stattin, Janlert, & Järvholm (2010) address the problem of over-mortality, not focussing on potential underlying over-mortality causes, but rather on whether there is an over-mortality or not. The study has a very unusual design from a clinical scientific, epidemiologic, and bio-statistical point of view, and this commentary focuses on four of the design problems.

First, the authors took advantage of an existing dataset based on examinations of some 389,000 construction workers (Hult et al., 2010). From this point of view the study population has a structure similar to the early studies on survival among disability pensioners, i.e., the study population consists of people in one single occupation or related groups of occupations. Moreover, the gender balance is utterly skew. This means that the study population cannot be regarded as a representative sample from the general population. To overcome this problem, most recent studies in this area are based on random samples from the general population with approximately the same number of women and men (Gjesdal et al., 2009; Karlsson et al., 2007; Wallman et al., 2006). The choice of study population in this case was thus not quite optimal, and may produce results at least biased to some extent as
far as the general population, or even the male population, is concerned.

Second, the study population was then reduced, step-by-step, to 32,000 persons or less (Hult et al., 2010). The final study population thus consisted of less than 10% of the original study population. The reason for the downscaling was, according to the authors, to eliminate heterogeneity regarding health status in the study population. The authors call the procedure "adjustments", but this is not adjustment, it is exclusion.

Through the years several ways of handling heterogeneity have been advocated. Fifty years ago exclusion of data not fitting in with expected results was the preferred method to reduce heterogeneity in medical studies. It was euphemistically called "carefully selected patients". The reason for this approach was a feeling that heterogeneity caused more problems than exclusions. Today we know better. The use of exclusions causes potential selection bias, the more extensive the exclusions, the more severe the bias. Selection bias is a threat towards the validity of the results as far as the target population, in this case the general population, is concerned. Handling heterogeneity by exclusions is today definitely not state-of-the-art. It can be handled much more efficiently by analytical tools, whereas the non-validity caused by selection bias is hopeless and can never be measured nor adjusted for.

As an example of downscaling Hult et al. (2010) exclude all persons born before 1920 and after 1932. The reason for these exclusions is obscure. It appears that they want to have an approximately equal follow-up time for the whole study population and they do not want to do mathematical adjustments. However, they are using a Cox's proportional hazards regression model, which is highly "mathematical" and is an excellent tool to take differences in follow-up into account. There is thus incongruence in reasoning and choice of analysis methods.

Third, in their downscaling of the study population, the authors excluded data asymmetrically (Hult et al., 2010), i.e., they excluded groups with supposedly bad health from the disability pension group but not from the referent group. They seem to reason that only the disability pensioners, but not those who continue to work, suffer from potentially fatal diseases. Furthermore, in their argument they appear to regard most disability pension diagnoses as consequences of the work environment.

This is an illusion. All disability diagnoses prevail among disability recipients as well as among referents, even though the prevalence is lower among the latter, and the fraction receiving disability pension due to damage from work is small (http://statistik.forsakringskassan.se/rfv/html/LAF_2008.html). Excluding groups with presumed bad health only in the disability pension group causes an imbalance, making the disability pension group "more healthy" and the referents relatively "more unhealthy". In this way the differences in survival between the groups attenuate with successive exclusions. The results of the proportional hazards regression analyses in this way become more and more biased.

Fourth, all those who died before age 65 were excluded (Hult et al., 2010). A large proportion of outcomes were thus discarded. It is well known that the cumulative mortality for groups with different hazard functions diverge in the beginning and then the gap is closing at higher ages, to be nil when all are dead. By excluding all deaths before age 65 (so called left-truncation of data) the differences in the mortality rate are biased towards a null result, even in the case of no health status bias. Moreover, the statistical power, i.e., the possibility of finding existing relationships, which is highly dependent on the number of persons reaching an outcome, is reduced by this procedure.

In conclusion, the procedures employed by Hult et al. (2010) are not according to the textbook and, instead, seem to be an unorthodox and insufficiently rigorous analytical approach. The optimum solution would have been to set up an analysis model in which all adjustments needed would have been introduced. The analyses should have been performed on the complete original study population, with its high statistical power, even though it is not perfect for the study purpose. Heterogeneity could then fairly easily have been handled by covariates, stratification, or other means of variance reduction and confounding elimination or reduction. Given the considerable size of the study population that type of strategy would have provided highly reliable and valuable results. The analysis of the over-mortality issue is not easily achieved. A possible way forward might be to let other research groups analyse the problem using this dataset.

References


Gjesdal, S., Haug, K., Ringdal, P., Maeland, J.G., Hagberg, J., Røraas, T., et al. (2009). Disability pension due to damage from work is small (http://statistik.forsakringskassan.se/rfv/html/LAF_2008.html). Excluding groups with presumed bad health only in the disability pension group causes an imbalance, making the disability pension group “more healthy” and the referents relatively “more unhealthy”. In this way the differences in survival between the groups attenuate with successive exclusions. The results of the proportional hazards regression analyses in this way become more and more biased.


Ockander, M., & Timpka, T. (2001). A female lay perspective on the establishment existing relationships, which is highly dependent on the number of persons reaching an outcome, is reduced by this procedure.

In conclusion, the procedures employed by Hult et al. (2010) are not according to the textbook and, instead, seem to be an unorthodox and insufficiently rigorous analytical approach. The optimum solution would have been to set up an analysis model in which all adjustments needed would have been introduced. The analyses should have been performed on the complete original study population, with its high statistical power, even though it is not perfect for the study purpose. Heterogeneity could then fairly easily have been handled by covariates, stratification, or other means of variance reduction and confounding elimination or reduction. Given the considerable size of the study population that type of strategy would have provided highly reliable and valuable results. The analysis of the over-mortality issue is not easily achieved. A possible way forward might be to let other research groups analyse the problem using this dataset.

References


Gjesdal, S., Haug, K., Ringdal, P., Maeland, J.G., Hagberg, J., Røraas, T., et al. (2009). Disability pension due to damage from work is small (http://statistik.forsakringskassan.se/rfv/html/LAF_2008.html). Excluding groups with presumed bad health only in the disability pension group causes an imbalance, making the disability pension group “more healthy” and the referents relatively “more unhealthy”. In this way the differences in survival between the groups attenuate with successive exclusions. The results of the proportional hazards regression analyses in this way become more and more biased.


