The epidemiology of common mental disorders from age 20 to 50: results from the Zurich cohort study

Commentary on: Angst *et al.* (2015). The epidemiology of common mental disorders from age 20 to 50: results from the prospective Zurich cohort Study. *Epidemiology and Psychiatric Sciences* (2016), **25**, 24–32.

In the epidemiology of the common mental disorders, to have studied some 4500 adults in the community over 30 years is an exceptional undertaking. Angst and his team deserve our unqualified appreciation for this achievement. For those who have followed large-scale prospective studies of the general population over the past half-century, it is now possible to identify the principal conclusions to be drawn from this epidemiological genre. But in working towards such a synthesis, there are invariably limitations to what can be quantitatively aggregated. The samples differ in their provenance, they are drawn from different societies, refusal rates vary, there is different attrition over time and ascertainment methods are rarely the same. Indeed, the very words used in interviews may carry different meanings in different locations. The investigators themselves often change, as do some of their aims. Indeed, the last two waves of the Zurich study, in 1999 and 2008, were conducted by Dr Wulf Rössler and his team following Angst's retirement. These reservations apart, the study has a number of notable assets. Here we look at these, to accord the study its rightful place in constructing a global picture of this class of morbidity in the human population.

We note how the sample started: because military service has been obligatory for all healthy males in Switzerland, the investigators seized this opportunity to construct a cohort of them. To achieve an acceptance rate of 99.7% is astonishing and surely unique in a democracy. The sample of women recruited from the same community was probably as satisfactory as could possibly be achieved, though the acceptance rate was a much more modest 75%, still much better than usually can be obtained. Subsequent attrition, which is inevitable and impossible to overcome, was apparently considerable as author's figure 1 shows. It would have been non-random, rather than having an equal probability of respondents dropping out, irrespective of their health. The final sample examined in 2008 seems to have been less than a tenth of the original cohort. So we do not know with reasonable certainty about the extent of bias in the later estimates. Admittedly, analysis of the consequences of attrition is reported to have found nothing of concern.

In the first examination, a two-phase design was adopted. To this writer, such a strategy has been surprisingly underused in mental health epidemiology, despite its great efficiency. This was shown by Duncan-Jones & Henderson (1978) in a much more modest cross-sectional study in Canberra. In a two-phase design, a brief screening instrument with appropriately broad psychometric properties is first administered to the entire sample. In terms of latent trait theory, the screening instrument should cover a wide range of severity, so it should have a wide threshold and shallow slope to detect a range of severity (Grayson, 1988; Andrich & Van Schoubreck, 1989). A more detailed examination is then made of the high scorers, but importantly, also of a proportion of the middle and low scorers. It is then easy to weight back the findings in each segment to make a statistically acceptable estimate of morbidity in the entire sample, although only a part of it has had the full examination. This saves both money and the uninformative examination of many people with few or no symptoms. The authors used this with good effect.

To be strongly endorsed is the period of recall they used in estimating incidence. Angst *et al.* cite some of the most informative publications on this. By making the reference period only the last 12 months instead of the longer interval since the previous interview, the investigators minimised the very serious underestimate that we now know takes place and which they

Address correspondence to: S. Henderson, National Institute for Mental Health Research, The Australian National University, Canberra, A.C.T. 0200, Australia.

(Email: ashenderson@netspace.net.au)

discuss. This underestimate is at its most apparent in so-called life-time estimates used in many large-scale national surveys. The stem usually starts with the words, 'Have you ever…..?' At a WPA Epidemiology meeting where life-time estimates were under discussion, Ernest Gruenberg once scoffed, 'Oh yes, the McCarthy item!' He was referring to the CIA's enquiries about a person's possible membership of the Communist Party in earlier life. Hopefully, this practice in surveys will now be abandoned. The Zurich study produces the highest rates ever reported for life-time prevalence, 82% in both sexes, or 74% if tobacco dependence is excluded. These are remarkably high estimates, but they are reached when the most rigorous methodology is used to minimise underreporting. What is the significance of this finding? Angst *et al*. justifiably draw attention to what they describe as the 'nearly universal' disturbance that takes place in mid-adulthood. Furthermore, they confirm the consistently higher rates for affective disorders in women, while it is men who lead in alcohol and substance dependence.

What has the Zurich Study, and indeed its equivalents elsewhere, achieved scientifically and for mental health services, bearing in mind that the ultimate aim of all epidemiology is prevention. One could ask how such studies start. What questions did the investigators have in mind at the beginning? Can the findings be applied to improve prevention? What does it take for the team to cohere over many years, to retain their morale in the face of uncertainty and disappointments, and what technical problems emerge, apart from drop-out among the researchers themselves? These are the matters that would make a most appealing narrative.

Angst *et al*. have persuasively shown both the general public and health policy administrators that the common mental disorders are ubiquitous and spare few of the population. Their distribution is higher in some groups than in others, and they are mostly recurrent or very long lasting. Rationally justified entry points for prevention are proposed. Within the research community, the team have generated instruments and methods of considerable utility, not least for possible application in primary care. In examining a sample of the same individuals prospectively over a long period of time, the Zurich Study informs us about the long-term course of disorders, just as studies of the same population, but not the same respondents, over a decade or more can identify changes in the pattern of morbidity (Kessler *et al*. 2005; Slade *et al*. 2009).

If Angst *et al* were starting again, would they do anything differently? Would they still do the study, with hindsight? If it were indeed to be undertaken, the present writer can identify three areas to consider, all of them additions. First, measures to improve sample retention could be introduced, using all ethically acceptable means to promote respondents interest and commitment to the study. In this, the approach to prospective respondents is critical at each wave, with personal contact being very desirable where this is practicable. The cost, though, can be prohibitive. Next, biological measures such as cytokines or a cheek-swab for DNA now have some appeal, especially when these are linked to indicators of psychological vulnerability. This is an emerging field worth keeping in mind, albeit with caution (Niculescu *et al*. 2015). Lastly, thought might be given to the feasibility of introducing an intervention. An example would be that a random proportion of the cohort be offered exposure to some activity known to promote mental health. Much interest is currently given to internet methods for this, particularly since these are inexpensive, readily accessed and can be repeated over time. But all of these are only embellishments on what is indisputably one of the signal achievements of psychiatric epidemiology.

Scott Henderson

References

Andrich D, Van Schoubroeck L (1989). The General Health Questionnaire: a psychometric analysis using latent trait theory. *Psychological Medicine* **19**, 469–485. doi:10.1017/S0033291700012502.

Duncan-Jones P, Henderson S (1978). The use of a two-phase design in a prevalence survey. *Social Psychiatry* **13**, 231–237.

Grayson DA (1988). Two-group classification in latent trait theory: scores with monotone likelihood ratio. *Psychometrika* **53**, 383–392.

Kessler RC, Chiu WT, Demler O, Walters EE (2005). Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the national comorbidity survey replication. *Archives of General Psychiatry* **62**, 617–627. doi:10.1001/archpsyc.62.6.617.

Niculescu AB, Levey D, Le-Niculescu H, Niculescu E, Kurian SM, Salomon D (2015). Psychiatric blood biomarkers: avoiding jumping to premature negative or positive conclusions. *Molecular Psychiatry* **20**, 286–288. doi:10.1038/mp.2014.180. Published online 13 January 2015.

Slade T, Johnston A, Oakley Browne M, Andrews G, Whiteford H (2009). The 2007 National survey of mental health and wellbeing: methods and key findings. *Australian and New Zealand Journal of Psychiatry* **43**, 594–605. doi:10.1080/00048670902970882.