Reviewer’s report:

Title: Do people with risky behaviours participate in biomedical cohort studies?

Version: 1 Date: 15 September 2005

Reviewer: T Chandola

The paper is well written and the analysis is appropriate. However, the rationale for this paper is not evident to this reviewer. The background section of the abstract does not raise any research hypotheses; in fact, it describes the sampling methods and variables in analysis. The paper investigates the representativeness of a particular cohort, so what can one generalise from the analysis? It appears that the findings relate only to this particular cohort. If this is the case, the proper place for it would be in the context of another paper exploring substantive research hypotheses, and the analysis in this paper would then establish the representativeness of this particular cohort.

The lack of any research hypotheses was made clearer when reading the Tables- we don’t know the hypothesised direction of the associations. Are we supposed to expect fewer younger participants in the NWAHCS? Fewer low income participants? More women? The reader does not know whether the age difference in the NWAHCS and census (Table 1) is to be expected. This originates from the lack of any meaningful research hypotheses that can be generalised to other studies. The conclusion “People with risky behaviours participated in this health study in the same proportions as people without these risk factors” underlines the fact that these results cannot be generalised to other studies.

Other minor points:
1) Details of how the weights were calculated for the NWAHCS sample were not given
2) Page 10, line 4-5: “Their overall SELF-RATED health status is the same…”
3) Page 12, 1st para: The relevance of these details of how the study intends to reduce loss to follow-up is irrelevant to the Discussion. The Discussion should be in the context of how these findings can be seen in the context of other research. Such details are more appropriate in the methods section of a paper investigating substantive research questions.
4) Page 12, last para: Sentence beginning “Results of investigations of the initial samples…” is clumsy. Does this mean that there are few recently started cohort studies, that have been peer reviewed in terms of the representativeness of their samples? If this is what the authors mean, it would be an extremely contentious statement. The authors do not provide any evidence for this in the Introduction or Discussion. Most cohort studies I know of do publish peer reviewed details of the representativeness of their samples within the context of substantive research papers.
5) Tables 1-4: Write down the acronyms of the datasets analysed below the Table