Reviewer's report

Title: Internet-based surveillance of Influenza-like-illness in the UK during the 2009 H1N1 influenza pandemic

Version: 1 Date: 6 August 2010

Reviewer: Jouke van der Zee

Reviewer's report:

my major concern is that I do not get a good impression what the authors want to state or to prove with their paper.

They start stating that collecting information about (for example) flu via internet has several advantages over the traditional way of collecting this type of information, particularly the datacollection via GPs. The latter way of datacollection has the disadvantage that the information is collected from people who visit their GPs, probably missing cases who do not seek clinical attention. So there is a gap between the number of reported cases and the 'real' number in the open population. Internet questionnaire will bridge that gap. The other advantage is that due to the delay implied in seeking care internet based information is more timely than GP-based data.

The authors immediately admit that internet data may have their own bias (as not reaching the major target groups for flu: children under 5 and the frail elderly (over 80)) but statistical weighing techniques can overcome this problem.

This point of departure of the authors evokes the following comments: 1) why only mentioning advantages of internet based data collection and not a major disadvantage namely the lack of validity of the data (especially if one is interested in the specific type of flu virus)

2) one would expect that the authors would focus a) on the gap between the number of GP-based reported cases and the 'true' number of cases in the population and b) on the better timeliness of the internet information. In the end one could state whether the costs of this internet based flu-survey operation (staff, recruitment, datacollection and processing) would outweigh the benefits, by arguing that either the current GP based information service could be replaced by the internet survey or that the additional information collected by the internet survey is so valuable that it it should be continued.

The paper does not focus on this population/GP attendance gap (I'd very much appreciate if it would) and the internet data do not seem to be more timely than the GP data. Finally the validity problem (the big disadvantage of questionnaire based data) is not touched at all; if a policy making authority is interested in the type of flu virus circulating (as in case of a pandemic); the internet survey data are far too general

My conclusion from the paper is that the paper does not succeed in proving what the value of internet based surveys is in this case (pandemic flu surveillance) and
certainly not what it adds to the existing datacollection structure.

My proposal is, that both the introduction and the conclusion/discussion section should be made less one-sided (and pro-internet). I also propose that intermediate ways of datacollection (eg via NHS Direct in the UK) are included in the discussion.

So my advice is: accept after essential revions (and some minor less essential ones)

Minor compulsory revisions: 1) some typical UK features (like what is the Pandemic Flu Services and how does that influence the Flu reports) should be explained for a non-UK audience 2)please explain in more detail how the case assesment numbers of HPA are being calculated; it is a crucial element in the paper (it serves as a sort of gold standard) but it is not explained 3) I miss references to the use of NHS Direct data

ref 18 is strange (a handful of initials)

**Level of interest:** An article whose findings are important to those with closely related research interests

**Quality of written English:** Acceptable

**Statistical review:** Yes, and I have assessed the statistics in my report.

**Declaration of competing interests:**

I have no competing interests

jvdz