Reviewer's report

Title: Incidence and risk factors for influenza-like-illness in the UK: Online surveillance using Flusurvey

Version: 2 Date: 7 October 2013

Reviewer: James McCaw

Reviewer's report:

I have only minor queries.

1. Participant characteristics are collected upon "initial registration". After 4 years (and assuming some people remain enrolled over an extended period), how do you monitor (if at all) for and so account for any changes in status over the period of enrolment - e.g. vaccination, onset of disease cofactors (e.g. asthma), changes in lifestyle (e.g. public transport usage, employment).

2. Typo or phrase requiring some clarification: "..., further details are available upon request."

3. Tense issues: Under "Case definitions", the ECDC definition of ILI is present tense ("requires"), while the "fever" definition is in past tense ("required"). Furthermore, these two definitions are not precisely defined as, at least to me, there is an ambiguous use of "or" and "and", particularly in light of the comment that ILI^{fever} is a subset of ILI^{ECDC}.

4. I'm curious as to the rationale for the rule for choosing which of multiple surveys completed on a single day to keep. Why does meeting the ECDC definition trump the context of the other (earlier or later) report from the same day?

5. Somewhat related to 1. above, the decision to exclude the first symptom survey of all participants seems reasonable, but when is the first symptom survey completed? At the start of an influenza season, or at the start of first-ever enrolment? I suspect I may just be confused about when enrolment may first occur. A clarification in the Methods would help.

6. Table 2: Typo in heading of 4th column (unbalanced right brace)
7. Tables 3 and 4: "n/a" seems odd. The data is unavailable, not "not applicable".

8. While it is clear that with the significant sample size the differences between, for example, females (49.6%) and males (39.7%) are statistically significant, it would be helpful to present CIs in the text. This format for presentation will help when it comes to, say, the Northern Ireland listings where there are only 64 participants and so it is not at all obvious that the quoted difference is statistically significant.

9. Figure 1 is clear and helpful. However, a claim is made with respect to it that the 19-45 group follows the <18 group by one week with the very strong statement: "If the under 18 year olds had a peak, the 19-45 year olds would have a peak the following week." Even a cursory glance at the figure reveals that this is not actually the case. At 16th December it is true. At 20th January it is two weeks, not one until a small local peak in the 19-45 group occurs. If there is a true and robust week-delayed correlation please do the proper analysis to show this. If there is just a visual hint in the data, amend the text accordingly.

10. Table 6: Typo in row name "36-44" should be "35-44".

11. The association with female sex is interesting. McCaw et al (BMC ID 12:345 2012), in a study on transmission, also noted an affect (of borderline statistical significance but with a reasonably high point-estimate for effect size) of sex, even after accounting for presence of children. The Discussion therein, while not of direct relevant here due to the transmission context, provides some relevant and potentially interesting references on risk perception and pathogenesis which the authors may like to follow up to further inform their Discussion.

12. Strongly tied to point 9. above, the "lagged effect" claim for older age groups is again made. Please emphasise the by-eye nature of this statement or strengthen by conducting a proper analysis. Also, the group is listed as "25-45 year olds", which I presume should be "19-45 year olds".

13. Can the authors' examine the literature to find examples from community based (say, GP) surveys that do include virological sample analysis. Presuming they do exist, what do they indicate regarding the congruence or
otherwise of symptoms in test-positive and test-negative (for influenza) ILI samples? How does this information, if available, help with the interpretation of the current study?

Level of interest: An article of importance in its field

Quality of written English: Acceptable

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

Declaration of competing interests:
I declare that I have no competing interests