Title: Tracking the spatial diffusion of influenza and norovirus using telehealth data: a spatio-temporal analysis of syndromic data.

Authors:

Duncan L Cooper (duncan.cooper@bradford.nhs.uk)
Gillian E Smith (gillian.smith@hpa.org.uk)
Martyn Regan (martyn.regan@hpa.org.uk)
Shirley Large (Shirley.Large@nhsdirect.nhs.uk)
Peter P Groenewegen (p.groenewegen@nivel.nl)

Version: 3 Date: 1 May 2008

Author's response to reviews: see over
Dear Dr da Silva

**RE: Tracking the spatial diffusion of influenza and norovirus using telehealth data.**

Thank you for your interest in our article.

In your response you offer us the opportunity to re-write the discussion section focussing more on a general medical audience (for publication in BMC Medicine). We accept this offer and have rewritten the discussion by dividing it into two parts. The first part is written for a general medical audience, the second part focuses on the strengths and weaknesses of the current spatial analyses.

The revised manuscript has also been submitted online. Please also see below a point by point response to the reviewer’s comments.

Due to these changes the discussion section in the revised manuscript is slightly longer than before, but with less content about the spatial analyses and statistical algorithm used (as suggested). All changes to the text are highlighted.

We hope you will consider this revised manuscript for BMC Medicine. All authors have authorised these changes.

Yours sincerely,

Duncan Cooper
Reviewer's report

Title: Tracking the spatial diffusion of influenza and norovirus using telehealth data: a spatio-temporal analysis of syndromic data.

Version: 1 Date: 20 March 2008

Reviewer 3: John Brownstein

****************************************************************************************

Reviewer's report:

This manuscript presents essentially a descriptive analysis of the use of telehealth data for the purpose of monitoring the spatial spread of both influenza and norovirus. While this should not be considered a definitive evaluation, the presentation of the data and its application are highly novel and warrant publication. To date, there have been very few analyses showing the potential of syndromic data for monitoring fine scale geographic patterns. This is a very exciting line of research.

While I am highly enthusiastic about this study, I do have a few concerns that the authors should consider addressing.

Minor Essential Revisions

1-The recommendation for the use SaTscan should be removed. Since the authors did not compare other spatial statistic methods, they have not evaluated the true utility of SaTscan. Overall, I would avoid considering this study an evaluation. It is more of a description of the use of a method to describe patterns of spread in real-time. This in itself is very useful and has real public health value. A study that evaluates methods would be a welcome follow-up.

Response: The word ‘recommendation’ has been removed and the text reworded so that SaTScan may be considered for future surveillance in the UK. The test has been revised as follows (page 20 - paragraph 1): “It [SaTScan] may
therefore be considered for future national surveillance initiatives e.g. the proposed hospital based surveillance and 2012 Olympic planning in the UK.”

The authors accept this study is not an evaluation of different methods or data sources so have reworded the following sections of text to emphasise the ‘descriptive’ nature of the work: Page 2 - paragraph 1, page 8 – paragraph 1, page 14 – paragraph 1, page 21 - lines 1-2.

2-Because of the obvious difficulty in validating whether these data were truly reflective patterns of specific viral spread, the authors should be cautious about how they interpret these results.(though the authors are clear about this limitation in the discussion, other areas of the paper including the abstract may be misleading). Perhaps referring to this limitation in the Introduction would reduce confusion.

Response: We agree that this point needs to be reinforced. The following text has been added to the conclusions sections within the discussion and abstract: “In interpreting these results, care must be taken to consider other infectious and non-infectious causes of fever and vomiting.” The following line has been reworded within the introduction (page 6 – paragraph 2): “In this article we have used syndromic data about fever and vomiting calls, of which there are many infectious and non-infectious causes.”

3-Further, have the authors considered the impact of RSV on their data? Including national estimates of the RSV impact would be very helpful at least. If the epidemics do not overlap, this would eliminate RSV as a confounder in the data (though the use of only patients over 5 years of age is a good start).

Response: The authors have considered this. A previously published study estimated the impact of RSV and other pathogens on the volume of a range of respiratory syndromes reported to NHS Direct (see: Cooper DL, Smith GE, Edmunds WJ, Joseph C, Gerard E, George RC. The contribution of respiratory pathogens to the seasonality of NHS Direct calls. J Infect 2007; 55:240-248). The following short paragraph has been added (page 18 – paragraph 2):

“Common respiratory virus’s, such as respiratory syncytial virus (RSV), may have contributed to the rise in fever calls (although RSV predominantly effects children less than 5 years). Previous work estimating the contribution of a range of respiratory pathogens to NHS Direct calls found that RSV was responsible for approximately 15% of annual NHS Direct ‘cough’ calls. Although a significant relationship was found between the seasonal variation in fever calls and influenza, no relationship was observed with RSV, parainfluenza or rhinovirus. “
Reviewer's report
Title: Tracking the spatial diffusion of influenza and norovirus using telehealth data: a spatio-temporal analysis of syndromic data.
Version: 1 Date: 12 March 2008
Reviewer 1: Alain-Jacques Valleron
Reviewer's report:

This paper pursues 3 different related aims at the same time: demonstrate that telehealth data from the NHS can be used for syndromic surveillance; describe the performance of an algorithm (SatScan) to represent the dynamics of outbreaks; describe seasonal epidemics of ILI and norovirus during 2005-2006. The good performance of the system (data + algorithm) is visually obvious: to me, Fig 3 and 4 speak by themselves. The quantitative assessment of the performance of the system has yet to be done. In the case of ILI, the data of the Royal College of GP provide a reference to which the system can be compared, and more should be done (e.g. present AMOC curves). In the case of novovirus, there is no obvious gold standard that can be used to judge the sensitivity, specificity and timeliness of the proposed system. I do not think, neither, that the results presented support the conclusion that the scan statistic is “recommended” to analyse the telehealth data (no other methods are compared), even if obviously it worked well.

Response: This point has been addressed above. SaTScan is no longer recommended but instead suggested a means of analysing syndromic data in the UK.

Major compulsory revisions

I would recommend that the authors add a section with detailed quantitative assessment of he performance of the Telehealth/ saTScan system versus the RCGP data analysed with –say- a Serfling technique. It would be much more informative than the Fig 5.
Response: The aim of this paper was to describe the geographical origin and diffusion of rises in NHS Direct calls rather than report in detail purely temporal associations of the data. By doing this we would add to the aims, complexity and length of the paper.

However, the authors have conducted previous quantitative assessment of Telehealth and RCGP data (see Cooper DL, Verlander NQ, Elliot AJ, Joseph CA, Smith GE. Can syndromic thresholds provide early warning of national influenza outbreaks? J Public Health (in press)). In addition we have published results of applying a serfling-type model to telehealth (see reference 16 in the text). This model incorporated seasonal, weekend, bank holiday, and secular trend terms to estimate the seasonal variation in expected telehealth call volumes for different syndromes. Also, of relevance is a paper (Doroshenko A, Cooper D, Smith GE, Gerard E, Chinemana F, Verlander NQ. Nicoll A. Evaluation of Syndromic Surveillance Based on NHS Direct Derived Data in England and Wales. MMWR 2005; 54 (Suppl): 117-122.) that used a cross correlogram and fitted time-series models to study the relationship between telehealth and clinical data (RCGP). Not all these references were not included in the manuscript to avoid excessive self-referencing.

Discretionary revisions
- I am not convinced that Fig. 6 brings any information for the reasons explained in the text (page 14, last 3 lines). It could be deleted, unless the authors provide a statistical analysis of the data presented.

Response: The authors agree with this comment and have removed Figure 6 from the manuscript.

- I suggest putting Tables 2 and 3 in a supplementary section.

Response: In our experience those working in public health departments, and those researching syndromic surveillance (key audiences for this paper) expect to see details of the frequency, length and case numbers of reported clusters (Tables 2 and 3). These tables provide the data that support the maps in figures 3 and 4. We have therefore left these tables in the main text. However, if the editors wish to place Tables 2 and 3 in an appendix this can be done.
Reviewer’s report
Title: Tracking the spatial diffusion of influenza and norovirus using telehealth data: a spatio-temporal analysis of syndromic data.
Version: 1 Date: 12 March 2008
Reviewer 2: F C Tsui
Reviewer's report:
Reviewer's report

- Major Compulsory Revisions
1. The main issue with the study is the lack of definition of outbreak period based on the gold standard data (laboratory reports) and the quantitative measure of timeliness of the call centre data when comparing with the gold standard data. The authors used non-scientific term “rise” as a way to measure timeliness, which is troublesome to the reviewer.

Response: There is no quantitative laboratory measure of defining an outbreak period that is regularly used in England and Wales (as there is for clinical and telehealth data). However, those that define clinical thresholds within the UK Health Protection Agency (see reference 4 in text) have recommended that ‘2 consecutive weeks with more than 5 community laboratory reports’ be used as a working threshold for community influenza circulation. This threshold has been used in our paper and described in the text (page 13 - paragraph 2: “A rise in laboratory reports is defined as five or more laboratory report during two consecutive weeks [4]).”

2. The reviewer does not see the first fever call peak occurred prior to the ILI in North England in Figure 5. As a matter of fact, according to Figure 5, the ILI peak (around 50/2005) occurs prior to the peak of fever calls (around 51/2005).

Response: The authors interpret this statement to mean that the fever peak is in fact after the ILI peak, so telehealth data offer little added value. Figure 5 displays national NHS Direct fever calls against regional GP data. These national
fever data hide the earlier significant spatio-temporal rise in fever calls in Northern England demonstrated by Figure 3 (weeks 47-49). So although the national rise in telehealth calls may have happened at a similar time to a rise in GP consultations, the local rise in telehealth occurred first, demonstrating additional timeliness.

3. The authors use only one year data to demonstrate the usefulness of call centre data is not convincing. The reviewer suggests authors to use multiple years of data to demonstrate the timeliness.

At the time of the study only 12 months nationally geo-referenced call data were available. The authors accept this as a weakness of the work and have highlighted in the discussion the need to conduct the analyses during subsequent winters and with other NHS Direct syndromes (page 20 – paragraph 2):

“Further winters’ spatial analysis of respiratory syndromes is recommended to assess operational reliability of these results, and also to monitor influenza A strains associated with a high adult incidence.”

Quantitative analyses of multiple years national, but not local, call data have already been conducted, see:


- Minor Essential Revisions
1. Although authors stated that Influenza B infection predominantly affects school aged children, it’s not clear to the reviewer why authors chose only Influenza B instead of using both Influenza A and B (normally more Influenza A cases than Influenza B)?

Response: Data about Influenza A and influenza B laboratory confirmations were in fact used for this study and are displayed by week in Figure 5. We have re-emphasised this point clear in page 10 paragraph 2.

In the UK an influenza B ‘season’ occurs every three or four years when the predominating influenza strain is B rather than A. The 2005/2006 winter we
studied just happened to be an influenza B season. The reviewer quite rightly states that in most seasons influenza A predominates. This is why we included the following lines in the discussion (page 20 – paragraph 2): “We used fever calls about school aged children because we were studying a national influenza B outbreak. Spatial analyses of data about the NHS Direct ‘colds and flu’ syndrome is already used for national surveillance of influenza-like-illness in adults [16, 47]. Spatial analysis of these data is recommended to monitor influenza A strains associated with a high adult incidence.”

2. Why do the authors use weekly data instead of daily data? If the frequency of data collection is weekly, authors should state it.

Response: This is a valid point and has been addressed in the methods section (page 10) by adding the following line: “Although daily NHS Direct data were available we conducted weekly analyses to simplify presentation of data and align temporal comparisons with clinical and laboratory data which are available weekly”.

3. On Page 9 Line 8, “Also if during a specific week (e.g., the Christmas holiday when doctor’s surgeries are closed) all areas experience a doubling of vomiting call, no clusters will be identified.” The author should explain why it’s not the case where the entire country becomes a large cluster.

Response: The following sentences have been added to Stage 2 in the methods section (page 9 – paragraph 2): “For this method the number of observed calls in a cluster is compared to the expected number, assuming that the spatial and temporal locations of calls are independent of each other (no space-time interaction). At each time interval specified, a cluster is identified if a specific area has a higher proportion of excess cases than surrounding areas.

4. No Table 4 available but cited on page 12

This should have read Table 3 and has been corrected.