

7th October 2020

To:
Secretary of State for Health and Social Care
Chief Executive of the National Health Service (NHS)

Covid-19 Data Collection Crisis - data error blamed on limitations of spreadsheet, the Association of Professional Healthcare Analysts (AphA) would like to support an independent review and learning

Dear Mr Hancock and Sir Stevens,

As we have learned on Monday, 5th October more than 50,000 potentially infectious people may have been missed by contact tracers after 15,841 positive tests were not included in the daily reported figures. A limit on Microsoft's Excel spreadsheet software caused the data error¹.

At a time where our country is yet again working to contain a spike in coronavirus cases, this delay in tracing contacts could have devastating consequences.

It has been reported that one lab had sent its daily test report to Public Health England (PHE) in the form of a CSV file, the simplest possible database format. That report was then loaded into Microsoft Excel, and the new tests at the bottom were added to the main database.

These new tests were truncated from the database, and thus missed, because while CSV files hold virtually limitless data, the older version of Microsoft Excel files are limited, when a CSV file longer than 65k rows is opened by Excel, the bottom rows are cut off. Once the lab had performed more than a million tests, it was only a matter of time before rows of data representing thousands of infected people were being missed by PHE.

Each row of data missed is one person going undetected by contact tracers. At current estimates, each infected person goes on to infect between 1.3 and 1.63² more people, so each row of data being missed is one row too many.

This is not an IT glitch and to describe it as such trivialises the issue, this is the symptom of a more serious, chronic failure of the systems and processes.

This is a systematic failure that led to an individual making a poor decision on the most appropriate method of collecting data. This can only be improved by long term investment in infrastructure and training.

The Association of Professional Healthcare Analysts (AphA) identified a number of strands that could have contributed to this event, we therefore make the following recommendations:

- We must judge the risks and understand the limitations of software systems to ensure that they are fit for purpose.
- We must apply appropriate analytical methods (eg multi-variate analysis, advanced forecasting, machine learning and Artificial Intelligence techniques) whilst ensuring supply of suitable tools for analysts and data scientists.
- An out of date version of Excel should not be used to transfer vital data sets in the real world.
- We must invest in a suitably skilled analytical workforce to demonstrate the importance of data to drive decision making in a 21st century health and care system³.
- Training is essential. For example, there is promise in the potential for Artificial Intelligence (AI) and data-driven technologies to improve health and care, but to realise this potential requires the Government to invest in training in Mathematics, not simply buying machines.
- The Government must fund and support programmes for data scientists and provide systematic incentives and financial backing to foster open source sharing.
- We must equip our analytical workforce with the right tools and techniques. For example, the relevant bodies must develop a national professional framework for health and care analysts and invest in analytical leadership eg appointing Chief Analytical Officers⁴.
- We must establish the difference between Information Technology (IT) and Information Management (IM).

- It is vital that the Government invests in analytical capability within the health and care system, and supports structured analytical programmes to develop a future ready workforce as outlined in the Topol Report (2019)⁵.
- We should look at this event not as a catalyst to blame or shame, but as an opportunity for learning and improvement.
- We should reflect and move forward to design a robust method of data collecting and reporting that is fit to manage the risk of coronavirus and protect the country, engaging with the analytical professionals to develop a rapid, high quality analytical solution.

About AphA – Who We Are

The Association of Professional Healthcare Analysts Network was established in 2012 to raise the profile of healthcare analysts and provide a professional support network. AphA provides a platform for its members to achieve professional registration status, and to drive up the quality and applicability of robust analytics as an aid to evidence-based decision making in a modern health and care system.

We would also like to raise awareness amongst health and care organisations of the importance of the roles of the membership in helping to strengthen the Health and Social Care strategy and vision by providing knowledge and expertise to support informed decision making.

AphA now has a large and growing membership of analytical workforce at all levels within the Health and Social care sector.

We urge the Government to take this opportunity to explore ways of improving practice, and to commission an independent review from analytical experts and professionals.

The Association of Professional Healthcare Analysts (AphA) are here to engage with the essential analytical issues and to provide recommendation and guidance and would be willing to support with a review and / or learning.

We look forward to hearing from you.

Yours truly,

Rony Arafin

AphA Chief Operating Officer
Assistant Director of Informatics
NHS Devon STP

Professor Martin Pitt

AphA President
Professor of Applied Healthcare Modelling and
Data Science - University of Exeter Medical School

Paul Stroner

AphA President Elect
Deputy Director NHS-R Community

Dr Marc Farr

Chief Analytical Officer - East Kent Hospitals
University NHS Foundation Trust
Founder - Beautiful Information

Ellen Coughlan

Programme Manager for Analytical Capability
The Health Foundation

1. <https://www.theguardian.com/world/2020/oct/05/england-covid-cases-error-unknown-how-many-contacts-not-traced-says-minister>
2. <https://www.gov.uk/guidance/the-r-number-in-the-uk>
3. Bringing NHS data analysis into the 21st century- <https://journals.sagepub.com/doi/full/10.1177/0141076820930666>
4. <https://www.health.org.uk/publications/reports/untapped-potential-investing-in-health-and-care-data-analytics>
5. <https://topol.hee.nhs.uk>