Witness Name: Professor Ronan Lyons

Statement No.:

Exhibits:

Dated: 14/11/2023

**UK COVID-19 INQUIRY**

---

**WITNESS STATEMENT OF Professor Ronan Lyons**

---

I, Professor Ronan Lyons, will say as follows: -

1. I am an academic health researcher with a background in medicine, public health, data science and epidemiology.

2. I offered to help Welsh Government and NHS Wales to better understand the evolution of the COVID-19 pandemic in Wales and the effectiveness of counter-measures through the use of de-identified linked data held in the Secure Anonymised Information Linkage (SAIL) research system.

3. I was a member of the Welsh Government COVID-19 Technical Advisory Group (TAG).

4. My team conducted many different analyses of the situation which were presented to TAG and published in peer-reviewed scientific journals.

5. I have expressed my opinions on how the business of TAG was conducted and the strengths and limitations of the data available in the comments I have made below, in response to the questions posed to me by the enquiry team.

**Biography/background**

6. I am Professor Ronan Lyons, Professor of Public Health at Swansea University and Honorary Consultant in Public Health at Public Health Wales NHS Trust. I trained in medicine at Trinity College Dublin, qualifying in 1983, and worked in hospitals for a number of years before undertaking a Master's in Public Health degree in 1987/8. I undertook specialist public health training In Ireland and Wales and took up a consultant post with the then West Glamorgan Health Authority in 1992. I entered academia in 1998 (University of Wales Medical School, Cardiff) and since then have also held an honorary role with public health agencies, currently Public Health Wales NHS Trust. I moved to Swansea University in 2005 when the new medical school opened.

7. I hold the following degrees and awards: BA, MB, BCh, BAO, DCH, DipStat, MPH, FFPHMI, FFPH, MD, FLSW, OBE, FMedSci, MAE.

8. I co-created the Secure Anonymised Information Linkage – www.saildatabank.com) with my colleague Professor David Ford in 2008. SAIL, as it is known, is a privacy protecting system which allows de-identified data to be linked to answer research questions about issues where there is potential for public good. It is part of the national research infrastructure for Wales. All projects have to be approved by the independent Information Governance Review Panel (IGRP) which includes members of the general public. For the last 16 years I have used data from SAIL to answer a wide range of scientific questions. SAIL utilises the Secure eResearch Platform (SeRP) which is a set of technologies and services to facilitate remote access to data whilst protecting the data and preserving privacy. SeRP is led by Professor Simon Thompson at Swansea University. It provides data linkage and remote analysis services to many organisations in the UK and internationally. The SAIL/SeRP capacity in Wales led to Wales leading a large share of COVID-19 studies in data compiled by Health Data Research UK on 02/11/21 (Exhibit INQ000328556).

9. Health and Care Research Wales is the health research funding arm of the Welsh Government. Like its sister organisation in England (National Institute for Health Research – NIHR) it holds competitions for research activities and fellowships. I was awarded Senior Faculty status in 2013 and am supported by an annual award.

10. When Health Data Research UK (HDRUK) was founded in 2018 I led the combined Wales and Northern Ireland site and was the UK lead for public health research. I held this role until mid-2021 when plans for funding for the second five year of 2023-2028 were taking place, as I planned to retire in that period.

11. My role in Public Health Wales NHS Trust is to provide academic support for specialist registrars and help with their research, notably on linked cancer data and with the vaccine effectiveness group.

12. I also chair the Natural Language Processing (NLP) stream of advanced analytics for Digital Health and Care Wales. This involves bringing together NHS and academic staff with interests and experience in NLP to stimulate projects. NLP is key to a modern, efficient NHS as human coding capacity cannot keep up with the detailed information held in a wide variety of NHS records.

**Membership of Welsh Government COVID-19 Technical Advisory Group (TAG)**

13. On the 17/03/2020 I wrote to CMO Wales, Dr Frank Atherton (now Sir Frank) to offer the services of my HDRUK funded group to utilise the linked data in SAIL to supplement work being undertaken by Welsh Government and NHS Wales organisations to better understand who were most vulnerable to infection and to evaluate interventions. He replied positively that day with an invitation to join the Welsh Government COVID-19 Technical Advisory Group (TAG), chaired by Rob Orford, Chief Scientific Advisor (Health). I submitted the first application of link multiple datasets on the Welsh population to answer COVID questions on the 31/03/20 building on an existing MRC funded study of multi-morbidity in this population.

14. I was a member of the Welsh Government COVID-19 Technical Advisory Group (TAG) from 20/03/20 to when it was stood down. I attended many meetings. Please see attached excel sheet (Exhibit INQ000328577) from Welsh Government on attendances. Attendances were not always recorded on their system. I attended the vast majority of the meetings.

15. I also attended several meetings of the Children and Education subgroup: 11/05/20, 28/05/20, and 04/06/20.

16. I was a member of the COVID-19 Research and Analysis Group: Adult Care Homes which met on 16/09/20. We conducted and published research on the risks of COVID-19 in care homes and on the uptake and effectiveness of vaccination (further details below).

17. I was also a member of the First Minister's Covid-19 Black, Asian and Minority Ethnic (BAME) Advisory Group and its Risk Assessment Subgroup and attended meetings on the 29/04/20 and 08/07/20. I presented on work we were doing on COVID-19 and ethnicity. I was a member of a subgroup developing a healthcare workforce risk assessment tool led by Prof Keshav Singhal. I was not a member of the group looking at the socio-economic impact of COVID led by Prof Emmanuel Ogbonna.

18. I attended one meeting of the SAGE Social Care Working Group in May 2020. Key staff from SAIL were invited to present on the 20th May 2020 to talk about the linked data available in Wales. It was decided that Dr Richard Fry (Swansea University and Rule 9 respondent) would represent Wales on that group.

19. I also contributed to the Academy of Medical Sciences report *Preparing for a Challenging Winter 2021/22* – See: https://acmedsci.ac.uk/file-download/51353957.

20. I didn't keep copies of the Agenda or Reports as these were available online on the Objective Connect TAG website, managed by Welsh Government. These should be available to the enquiry.

21. For most of the first year of COVID-19 I was working seven days a week but it was impossible to attend all of the many groups and sub-groups as these were created.

22. I was not involved in any WhatsApp or other messaging groups with Welsh Ministers, senior advisors or senior civil servants. All correspondence was through the Welsh Government organised Microsoft Teams meetings and e-mails. The documents are held on the Welsh Government COVID-19 TAG Objective Connect store.

**COVID-19 publications and reports provided to the Welsh Government COVID-19 Technical Advisory Group (TAG)**

23. To date, I have co-authored 73 research papers on the COVID pandemic in the UK. These are included in Appendix A (Exhibits INQ000328556, INQ000328585,

INQ000328596, INQ000328607, INQ000328618, INQ000328629, INQ000328640, INQ000328648, INQ000328659, INQ000328557, INQ000328566, INQ000328575, INQ000328578, INQ000328579, INQ000328580, INQ000328581, INQ000328582, INQ000328583, INQ000328584, INQ000328586, INQ000328587, INQ000328588, INQ000328589, INQ000328590, INQ000328591, INQ000328592, INQ000328593, INQ000328594, INQ000328595, INQ000328597, INQ000328598, INQ000328599, INQ000328600, INQ000328601, INQ000328602, INQ000328603, INQ000328604, INQ000328605, INQ000328606, INQ000328608, INQ000328609, INQ000328610, INQ000328611, INQ000328612, INQ000328613, INQ000328614, INQ000328615, INQ000328616, INQ000328617, INQ000328619, INQ000328620, INQ000328621, INQ000328622, INQ000328623, INQ000328624, INQ000328625, INQ000328626, INQ000328627, INQ000328628, INQ000328630, INQ000328631, INQ000328632, INQ000328633, INQ000328634, INQ000328635, INQ000328636, INQ000328637, INQ000328638, INQ000328639, INQ000328641, INQ000328642, INQ000328643, INQ000328644). Some are updates on earlier research which were published as pre-prints to make the details available to the scientific world as early as possible. The majority are based on linked Welsh data available through the SAIL system following the Welsh population after 01/01/2020 and comparing this with the historical situation from 2016-2019. This work was primarily funded by the HDRUK Wales grant and a specific grant from the Medical Research Council and National Institute for Health Research: Lyons RA, Williams C, Gravenor M, Brophy S, Akbari A, Thompson S, Fry R, Hollinghurst J, Cross L, Morgan A, Lyons J, Lucini B, Connor T, Jolles S, Taylor C, Cottrell S, Pacchiarini N, Bull M, Jones G, Diggle P, Davies J, Davies C. Controlling COVID19 through enhanced population surveillance and intervention (Con-COV): a platform approach UKRI-MRC, Aug 2020 to Mar 2022, MR/V028367/1.

**One Wales initiative**

24. As time progressed more and more questions were being asked around the data held in the SAIL databank. The controlling COVID through data linkage (ConCOV) cohort of linked data on the Welsh population was developed with MRC funding to answer questions and increase the number of scientific groups with access to the data. Updating the cohort monthly with new entries and exits (moved out of Wales or died) and new datasets took considerable time and expertise. It made no sense to ask each interested group to set up a new cohort in SAIL and replicate these activities. Initially, the groups were listed as ConCOV groups. This later became known as the SAIL One

Wales approach. Appendix B Contains a list of the groups accessing the Con-COV data as part of the One Wales approach.

25. Those working on the data were provided with a list of the priority questions raised by TAG (see attached spreadsheet Exhibits INQ000328668 and INQ000328559) but were also free to conduct any research they felt relevant. Many were specialty specific and they tended to focus on their areas of interest e.g. mental health, critical care, cancer. Work in progress and published reports were fed to TAG.

26. However, a number of the groups produced little output. Manipulating and making sense of complex data requires many highly skilled analysts, proficient in understanding data structures, ontologies (coding systems), and the use of SQL, R and other computer languages. Many of the groups lacked these skill sets.

27. Some groups continue to work on the data as scientific questions continue to arise. It is important that such work continues to inform planning for future pandemics.

28. As the pandemic evolved new scientific questions arose. We received additional funding from Health and Care Wales to support extensions of our work: Lyons RA. Undertaking scientific analysis of policy relevant questions to support the work of the Welsh Government Covid-19 Evidence Centre. Health and Care Research Wales, Mar 2021 to May 2022.

29. Administrative Data Research Wales (Economic and Social Research Council) funded staff were also freed up staff to work on analysis.

30. We received additional specific funding to evaluate the impact of shielding on vulnerable people: Snooks H, Watkins A, Lyons RA, John A, Sewell B, Porter A, Evans B, Lyons J, Akbari A, Bailey R, Surman A, Edwards A, Carson-Stevens A, Dale J. Effects of shielding for Vulnerable people during COVID-19 pandemic on health outcomes, costs and Immunity including those with cancer: quasi-experimenTal Evaluation (EVITE Immunity). MRC/National Core Study: Immunity (NCSi4P), Nov 2020 – Mar 2022.

31. Most of the initial research was based on trying to get a better understanding of which groups were being infected by the SARSCoV-2 virus, which ones became ill and required hospitalisation, critical care or died. Most of our studies were on the general

population but we also conducted studies into the impact of the pandemic on selected populations: those in residential care, those on the shielding list, health and social care workers, and the school population.

32. As treatments and vaccinations became available we turned our attention to how effective these were and their safety profiles. The results were fed to TAG prior to publication and onwards to SAGE and other groups. As time progressed many of the analyses were conducted in collaboration with similar groups working in England, Scotland and Northern Ireland.

33. I made quite a few presentations on our research to TAG, its subgroups, the research community and public groups during the pandemic.

34. Appendix C contains a list of presentations (Exhibits INQ000328568, INQ000328569, and INQ000328572) made to TAG on COVID-19 analyses covering transmission in schools and impact on cancer diagnoses. Appendix D contains presentations (Exhibits INQ000328645, INQ000328646, INQ000328647, INQ000328649, INQ000328650, INQ000328651, INQ000328652, INQ000328653, INQ000328654, INQ000328655) made to other groups, including to the BBC and members of the public. I also answered questions at a Sports Wales online meeting.

**Views on TAG**

35. The Welsh Government TAG was set up in March 2020 and consisted of a number of senior civil servants who comprised the Technical Advisory Cell (TAC) reporting to Welsh Government and those with external experts from the NHS and academia who were invited to participate, based on their expertise. Terms of Reference were published by Welsh Government. As time went by and the pandemic evolved the breadth of expertise on TAG grew substantially as new members joined. The vast majority were from Wales but there were others from England. I cannot comment on the procedures that were in place for the commissioning of advice or whether Welsh Governments commissioned advice separately from TAG.

36. At the beginning of the pandemic there was a huge demand for operational information from many different sectors which I understand was difficult to manage, particularly for PHW. However, I was not involved in the production of operational data from the NHS and hence cannot comment in detail.

37. My experience of involvement with TAG was very positive. The group worked extraordinarily hard in trying to understand the situation and provide advice on potential remediations to slow down the spread of infection and limit its impact on individuals, the NHS and wider society. I was impressed by the attitude taken by Dr Rob Orford and Fliss Benee, the co-convenors of TAG. When new aspects of the pandemic emerged they recruited additional experts to fill gaps. We had many presentations from a range of experts across the UK. I do not recall any limitations on the questions that could be asked of the group. The group worked well together and was open to robust questioning of assumptions and the introduction of new concepts and knowledge from external sources.

38. I was also particularly impressed by the professionalism and hard work of the more junior administrative staff in setting up meetings and document repositories, especially Richard Roberts who appeared to work 24/7 at the beginning of the pandemic. The contribution that such people make is rarely acknowledged.

**Sharing and linkage of health data**

39. The enquiry asked for an overview of the relevant legislation, guidance, policies, structures and organisations involved in supporting and regulating data sharing in Wales immediately prior to the beginning of the pandemic in January 2020.

40. Please see the attached legal summary which forms part of the SAIL Data Protection Impact Assessment (Exhibit INQ000328573).

41. My role was in supplying intelligence on the pandemic using the linked de-identified data held in the SAIL databank. I am aware that there are a number of routes to legally share data across the NHS and other sectors but I do not know enough to comment on the extent of this activity. Data will have been linked on patients within health boards, by DHCW and PHW and possibly by the Joint Biosecurity Centre (JBC). Welsh Government, like the UK Government, issued a 'COPI letter' to organisations providing health services, general practices and health boards providing them with authority to share personal data under the Health Service (Control of Patient Information) Regulations (2002). The aim of this was to facilitate the sharing of personal information between organisations to support the response to the pandemic.

42. Linkage of data is essential not only in a public health emergency but also as part of any information and evidence based provision of service. The rationale is quite simple. Outcomes are what matter to people (not getting infected, not getting very ill, not dying). Understanding the effectiveness of counter-measures requires linking exposure variables (infection, COVID disease, vaccination, treatment with anti-virals, etc) to personal characteristics (age, sex, deprivation, ethnicity, illnesses and treatments, etc) and then to outcome variables, using robustly designed studies. NHS datasets are generally designed to count activity for reporting to funders rather than study the effectiveness of care. Most datasets cover only a range of topics and require to be linked to understand what does and doesn't work.

43. The outcomes of linked data research, such as the QCOVID risk factor studies (Exhibits INQ000328556, INQ000328607, INQ000328640, INQ000328583, INQ000328591, INQ000328598), informed the development of national policies, for example, the prioritisation for vaccines. Operational systems within the NHS were then required to replicate as best they can the algorithms to implement the prioritisation.

44. Initially, we focussed on understanding the situation in Wales but as time went by the nature of questions changed, e.g. effectiveness of vaccines in population subgroups and possible adverse effects. Increasingly, we worked with groups across the UK to conduct analyses as quickly as possible to answer these questions.

**Adequacy of data available in Wales during the pandemic.**

45. A variety of data sources were used during the pandemic, a mixture of count data (also known as SitRep or Situation Report) such as number of people in hospital or ITU with COVID-19 and linked electronic health record data. SitRep data were collected daily and provided useful insights into trends but by their nature cannot tell what sorts of conditions led to admission or the severity of illness. Linked electronic records are required for such analyses. If Wales had had functioning electronic records throughout secondary (hospital) care SitRep data would not have been needed but could be pulled from the system. Some datasets were linked together within health boards, by DHCW and PHW but I have no great insight into what they did. There was also an organisation called the Joint Biosecurity Centre who were charged with the analysis of data across the UK but I do not know what analyses they undertook.

46. The data in SAIL, whilst not compete for all datasets, provided the best insight into patterns of transmission, the categories of individuals who tested positive for infection, became unwell or very ill and who, sadly, died. Provision of data to SAIL is voluntary. All health boards provide data as do DHCW. SAIL covers 86% of the Welsh population (82% of GP practices) and has 100% coverage from datasets supplied by DHCW and PHW.

47. SAIL links de-identified data at multiple levels, including at individual and household, care home and school populations. These linkages are much more extensive than elsewhere in the UK. This enabled early analyses on the impact of the pandemic on care homes and transmission patterns in school children (see Exhibits INQ000328585, INQ000328659, INQ000328575, INQ000328578, INQ000328580, INQ000328590, INQ000328593, INQ000328594, INQ000328600, INQ000328604, INQ000328619).

48. The SAIL ConCOV cohorts (also known as C20, short for Cohort 2020 – see Exhibit INQ000328618) were based on an existing study – the Wales Multimorbidity Cohort (WMC), funded by UKRI- MRC. This enabled rapid replication of the research commissioned by the UK CMOs on identifying groups at risk of hospitalisation or death. The QCOVID study was based on data from English general practices and replicated and validated in Wales using the data in SAIL. Replication was not exact due to slight differences in data availability and a DHCW restriction on the use of sensitive codes in SAIL (such as HIV status, sexually transmitted disorders and terminations). Hence HIV status, which was included in the English and Scottish algorithms could not be used in Wales. As it turned out HIV status was not associated with an elevated risk of hospitalisation or death in the English analyses. These algorithms were used to identify groups prioritised for vaccination across the UK.

49. There are, however, multiple limitations of the system. The population denominator used in SAIL and by NHS Wales organisation are those listed on the Welsh Demographic Spine (WDS), maintained by DHCW. WDS is derived from GP registrations and hence does not include people not registered. The system is largely blind to such people.

50. The 2010 Equality Act specifies nine protected characteristics: age, disability, gender reassignment, marriage and civil partnership, pregnancy and maternity, race, religion or belief, sex, and sexual orientation.

51. NHS systems collect data on age, sex, morbidities, pregnancy and maternity which are suitable for analysis.

52. The quality of ethnicity data in the NHS is very poor, with large amounts of missing data and different classification used over time. This is shown in research we conducted as part of our work attempting to address ethnic risks in COVID (Exhibits INQ000328588, INQ000328598, INQ000328614, INQ000328655). Due to deficiencies in the NHS we sought permission from ONS to bring in the 2011 Census which contains such fields for those who were around at that time. This improved the proportion of the population who could be classified enormously and has since been used by PHW to report on vaccine uptake by ethnicity, responding to a request from the First Minister's Black, Asian and Minority Ethnic Advisory Group (Exhibit INQ000328588).

53. The NHS does not measure disability per se but conditions, many of which are disabling. There are two approaches to monitoring disability – the first is reliance on grouping medical conditions that should be disabling and the second on using variables in the Census on limiting long term conditions.

54. There are difficulties in identifying long COVID in health care records as reported by a recent Administrative Data Research Wales report (https://adrwales.org/wp-content/uploads/2023/08/Clinical-coding-and-capture-of-Long-COVID.pdf). To some extent this is due to delays in updating GP computer systems in Wales, which still uses READ codes. English practices use the modern SNOMED codes and long COVID was defined using SNOMED codes.

55. I had multiple discussions with CSA (Health) Rob Orford on the need to improve the linkage of data and found him very helpful and supportive. He contributed to the success of our requests for data from the NHS on Health Care Workers (Electronic Staff Record) and Social Care Staff – held by Social Care Wales in May 2020. He and Fliss Bennee were also supportive of bringing in the viral genomic data (collected by Public Health Wales) into SAIL to support analyses by viral variant but we never managed to actually bring the data in. This meant that we could not undertake analysis by viral variant and had to rely on categories based on when each variant was dominant.

56. Linkable data on reasons for hospitalisation was slow to turn around. Hospital coding of discharge data feeds into a system called the Patient Episode Database for Wales (PEDW) managed by DHCW. There is always a delay in coding and difficulties in recruiting and retaining coding staff. As a result usable hospital data are often 3-6 months in arrears.

57. Different datasets are held by different parts of the NHS organisation, such as health boards and trusts, DHCW and the Business Services Organisation depending on the primary purpose of each dataset.

58. On the positive side during the pandemic we were able to increase the frequency of access to coded data from GP systems in SAIL from the usual three monthly to monthly when additional funding was made available from Welsh Government to DHCW to change the contracts with the company that deals with the GP system suppliers. We also received daily feeds of new infections, deaths and other sources with others at weekly intervals (see Exhibit INQ000328574) which lists the data flows, their owners and date of first receipt).

**Primary care and care homes data**

59. Computerised primary care, individual record systems have been around for decades and are used by all Welsh general practices. There are only a few system suppliers with a lot of standardisation. Welsh systems still use the Read version 2 or 3 coding ontologies in contrast to those in England which have moved to SNOMED. In contrast social care, including care homes, does not have a standardised ontology to code conditions and activity. Systems are currently being rolled out. I am not aware of any electronic communication between GP and care home systems.

**Statement of Truth**

I believe that the facts stated in this witness statement are true. I understand that proceedings may be brought against anyone who makes, or causes to be made, a false statement in a document verified by a statement of truth without an honest belief of its truth.

**Signed**: Personal Data

**Dated**: 14/11/2023