Office for Statistics Regulation
1 Drummond Gate
London SW1V 2QQ

I&S
regulation@statistics.gov.uk
osr.statisticsauthority.gov.uk
@statsregulation

Ed Humpherson, Director General for Regulation

Scott Heald
Head of Profession for Statistics
Public Health Scotland
(By email)

14 January 2021

Dear Scott

Presentation of findings from the Discharges from NHSScotland hospitals to care homes report

On 28 October, Public Health Scotland (PHS) published the report 'Discharges from NHSScotland hospitals to care homes between 1 March and 31 May 2020'. This was commissioned by the Cabinet Secretary for Health and Sport on 18 August and you collaborated with experts from the University of Glasgow and University of Edinburgh in the production of this report.

As you are already aware, we have received correspondence which includes a query around the presentation of the findings in the report. I would like to thank you and your team for your prompt and positive engagement with us on this.

This is a management information publication and we are encouraged to see that many of the principles of the Code of Practice for Statistics have been adopted by you in the publication of this report. For example, the publication date of the report was pre-announced via the PHS statistical release calendar, and a pre-release access list was published. A comprehensive methodology note was published which includes the data sources used, the level of quality assurance and data completeness. This is of particular importance, given the data challenges you faced for the research.

We understand that the methods used were based upon a similar study by Public Health Wales. We have reviewed the communication of findings in consultation with method experts from the Office for National Statistics (ONS). We consider the report is clear, thorough and transparent about the limitations and quality of the dataset and we recognise that the data and analyses undertaken were complex.

When reporting statistical analyses, the uncertainty of estimates and confidence intervals should be communicated in a measured and clear way. Whilst we understand that communicating complex statistical messaging can be challenging, we consider there are

lessons to be learnt from this case and expect the improvement points (Annex) to be addressed in the presentation of any future statistical analysis of the care home data.

Yours sincerely

Personal Data

Ed Humpherson

Director General for Regulation

Annex - Feedback for Public Health Scotland on the report on "Discharges from NHSScotland Hospitals to Care Homes between 1 March and 31 May 2020"

Communicating uncertainty

Often in analyses and reports of this nature, the main challenge is how best to express the strength of evidence and how to communicate statistical uncertainty. Whilst we recognise the difficulties in expressing this clearly, we consider there are some sections of the report that could be confusing to readers.

The hazard ratios in tables 10 and 11 show that while there is a statistical relationship between hospital discharge and care home outbreak in the univariate analyses, this relationship ceases to be significant in the adjusted models. This is clearly explained in the report, but some of the discussion of the uncertainty around the estimates led to the report feeling a little inconsistent in its messaging. For example, on page 39, where the report outlines that the estimated risk of hospital discharge is not statistically significant, and then proceeds to detail a best estimate of risk figure. Greater clarity and consistency with explanations would assist the reader to understand the findings of the statistical modelling.

Specific feedback on analysis of associations between any hospital discharge and outbreak (table 10)

The adjusted hazard ratio when looking at discharge compared to no discharge is 1.21 with confidence intervals of 0.94-1.54. Although this is not statistically significant, the fact that the lower confidence interval is close to 1 means that this is marginal for this level of confidence. The section on interpreting table 10 states clearly and in bold that "hospital discharge was not statistically significantly associated with care home outbreaks (adjusted HR 1.21)".

The conclusion section for this table then goes on to acknowledge that *"the best estimate of the hazard ratio for hospital discharge is >1 and the confidence interval in the adjusted analysis is relatively wide. We therefore cannot statistically exclude the presence of a small risk from hospital discharge".*

While it is good to see this discussion of uncertainty, this sentence feels quite technical and perhaps harder for a less experienced user to understand. It might also have been helpful to include this point in the section on interpreting table 10, alongside the statement of the non-significant finding, rather than in the conclusion. The way the information is presented in the report gives too much emphasis to the non-significant finding and not enough to the uncertainty. Presenting all of the information together (rather than under separate "interpreting table 10" and "conclusion" headers) would allow for a more balanced overall discussion of the statistical finding and the uncertainty around it.

Specific feedback on analysis of associations between different types of hospital discharge and outbreaks (table 11)

When looking at the different types of discharge, we see adjusted hazard ratios of 1.00 for tested negative, 1.27 for untested and 1.45 for tested positive. Although the confidence intervals again suggest these findings are not significant, the observed 'dose-response' pattern in the adjusted hazard ratios is consistent with a causal relationship between positivity and outbreak. Given the sensitivity of the care home setting during this pandemic, and the likely uses of the evidence from this analysis, some users may have benefited from additional discussion of this in the report.

Similar to table 10, the statement made in the section on interpreting the findings for table 11 is that none of the differences are statistically significant. The conclusion for table 11 then goes on to state: *"The analysis does not find statistical evidence that hospital discharges of any kind were associated with care home outbreaks. However, our certainty about the three types of hospital discharge defined by testing status varies."*

The first sentence could be read as not finding evidence for each and all of the types of discharge, but that is inconsistent with the second sentence.

The conclusion section for table 11 then goes on to provide some explanation of the uncertainty by stating that there may be a small risk where a person was untested and a moderate to large risk where they tested positive. This explanation was clearer and easier to understand than that for table 10. Again though, it would have been clearer if both the non-significant finding and the uncertainty around it were discussed in the section on interpreting the findings, rather than under separate headings.

Confidence intervals

Some of the confidence intervals around the estimates were very wide. In the report, there is some explanation around the uncertainty of wide confidence intervals and how caution is needed in interpreting them. In the section on table 11, however, you draw attention to the upper bound of a confidence interval of 374% and point to this as evidence of an association between a positive test and an outbreak. This seems more of a reflection of the sample size and the lack of statistical power – having less data would tend to make the upper bound of the confidence interval higher, but would not mean that the evidence was stronger. Given the importance of communicating uncertainty clearly, we encourage you to be careful in your discussion of confidence intervals and be clear where these indicate that we should be less confident overall in the robustness of the findings.

Impact of methodological choice on findings

A clear rationale is provided in the report as to why the estimates need to be adjusted for care home size: *"Larger homes will receive more discharges, and are also more likely to be services for older people, provide nursing care and be privately owned. These relationships between different care home characteristics means that simple comparison by single characteristics may be misleading."*

However, it is possible that there may be collinearity between hospital discharges and care home size, and the degree to which the model estimates may be affected by this was not covered in the report. It would be helpful for users to have more detailed explanation of how the possible relationship between discharges and care home size might have impacted on the findings.

Use of charts

The report makes good use of tables to display both univariate and adjusted hazard ratios along with their associated confidence intervals. This is very helpful and clear but of more use to a more experienced reader. Using charts in addition to the tables would enable you to present the information in a more visual way that might be easier for less expert users to understand.

Links to further helpful information

The following resources available on the GSS website may be of interest for any future presentation of the care home data.

- Making analytical publications accessible
- Communicating quality, uncertainty and change
- Introduction to data visualisation (also contains links to several other useful resources)

The Good Practice Team will also carry out reviews of statistical publications. Their page on communicating statistics on the GSS website provides further information on this.