ROYAL STATISTICAL SOCIETY

DATA | EVIDENCE | DECISIONS

RSS alerts Ofqual to the statistical issues relating to exam grading and assessment in 2020

Sharon Witherspoon, RSS Vice President for Education and Statistical Literacy, sent the following email[1] to Ofqual (and the Department for Education), on 9th April 2020:

Sent to Ofqual

I am writing as the Royal Statistical Society Vice President for Education and Statistical Literacy, about the plans for estimating student grades as a result of the impact of COVID 19.

I should say at the outset that the RSS knows how difficult this decision must have been. As statisticians we are aware of the enormous ramifications of all the social changes brought about by COVID 19; the outcomes for students whose examinations cannot be taken to the normal timetable is only one of these.

I am aware too that some (including some RSS Fellows) may feel that all students should just defer their exams. That would have brought some advantages in the form of being a known process, but also some disadvantages, not least in delaying or derailing opportunities at a time when getting back to normal will be an important aspiration for many, and when regimes for studying have been disrupted. Against that is, of course, the issue of the complexity and uncertainty of any statistical estimation methodology. It is good to know that the option of taking exams is still open to students.

All this to say that the Royal Statistical Society is keen to be constructively helpful in extraordinarily difficult times. We appreciate the gravity and urgency of the situation, and the importance of this issue for students, schools and teachers. We understand too that this requires a solution at pace, and what might be ideal if time were less pressed may well not be possible under these circumstances. These are objectively difficult decisions by any standards.

The RSS thinks that involving some outside statisticians in a formative way may be helpful as the detailed methodologies both for collecting and for evaluating data are put in place. We believe this would be of benefit to students and schools, but also help ensure the most robust choices given the constraints (on data and time) and reduce the possibility of significant outside argument over the estimation methods used. We know students will have the option to sit exams, but in order to reduce individual and 'group' appeals over the methodology, we believe that having some RSS Fellows and others on an advisory group discussing the options in the light of the detailed data available and the analysis options could be helpful. **I return to that issue at the end of this email**.

Just to illustrate the range of statistical issues that will arise, I include some examples, as raised by Fellows in our internal discussions. This is not an exhaustive list of statistical issues, and we are intensely aware that we do not know many of the details about the available data or the statistical options that might arise in the light of those details. But taken

---

[1] Some extraneous text is not included

together, we hope you feel they give a concrete sense of why we think some deeper continuing statistical involvement by the RSS and other statistical methods might be useful as DfE and Ofqual statisticians and education experts move to define the exact methodology to be used. We focus here on the issue of A-level grades, because they are probably more consequential for longer-term life chances, but many of the same issues will arise in estimating GCSE grades.

Issues with data collection/ quality

- As we understand it, there are 3 types of predicted grades at A-level: one based on prior attainment; the one the teacher considers the student will attain; and the one that teachers enter on UCAS forms. There is quite a bit of information about the 'generous' nature of this last, and we presume this is not one being used.

- That said, there is still quite a bit of variation in the reliability of different teachers in these predicted grades; past evidence may be available from schools about this for most teachers (but not new ones), and one issue is whether taking account of this and any systematic bias is feasible or helpful.

- Another issue relates to more general questions about the degree of certainty in predicting grades and/ or rankings. This may be particularly important for students who are not predicted to get top or bottom grades, or who are not in the top ranks or in the bottom; middle grades and middle rankings are likely to be subject to more uncertainty (both in the sense of true uncertainty, and potentially systematic bias). There may be ways of collecting predictions with uncertainty estimates or to consider models to consider degrees of uncertainty. Or it may be possible to use (the individual distribution of) GCSE attainment to illuminate uncertainty, or to use some sort of sampling estimate (based on Centres) of possible uncertainty. But in any case, this is another issue that we think should be discussed in the light of the detailed options, about which we understand you will not have time to have a full public consultation. We understand that being transparent about this may mean that students whose marks are more likely to be more uncertain may choose to sit exams, but we believe that transparency means that not only will individuals have more informed options, but it helps all of society consider how to fairly acknowledge uncertainty in these unique circumstances.

- A final issue may arise if there is more than one teacher per student studying a subject, and how this will be handled, or where, within a Centre, a number of teachers teach the same subject (with different degrees of accuracy or bias in estimating student marks) and give results for the students whose grade attainment they are individually responsible for marking. Again, there are possible ways to take account of this statistically.

Issues with subsequent statistical modelling of anticipated grades

In addition to the statistical issues related to the individual estimates, other statistical issues will arise in the statistical adjustment to produce final estimates. Again, I should emphasise that we don't claim the issues below are exhaustive, but as illustrations, they may be useful.

- There are a set of issues to do with how to use evidence about the previous attainment of schools/ Centres. There will be uncertainty about this, especially if schools or colleges have taken steps to improve previous attainment, but also due to natural volatility naturally, and these may be particularly important in the case of exams with a small number of students per school.

- There are broader questions (some of which may have some prior empirical evidence) about whether there is evidence (either generally or for particular subjects) about whether particular groups of students tend systematically to do better or worse than their predicted grades / coursework attainment. For instance, we are aware of some evidence that boys, for instance, do better in exams than in course-work and this may affect predicted grades, and some Fellows have asked whether there is evidence about bias in predicting results for students from a BAME background.

These are of course just illustrations of the issues we know that you face. But the Royal Statistical Society is ready to help if that would be useful. **We believe that having an engaged expert advisory group would be of use as you make decisions, and would be preferable to a purely internal process.**

If you decide that having an advisory group would be useful, we can suggest now some people you might wish to include.

In particular, we would suggest the following initially. The two RSS Fellows we have suggested have discussed this issue with us, and we know they are willing to help. There may be others we can help provide links to, but we hope the suggestions below provide a starting point:

- Professor Guy Nason, Imperial College; RSS Fellow. Professor Nason has long been active in debates around the use of statistics in education, including in higher education, and serves already on a DfE advisory group regarding higher education statistics.

- Professor Paula Williamson, University of Liverpool; RSS Fellow. Although Professor Williamson is an eminent medical statistician, many of the same statistical issues (to do with uncertainty and institutional and individual estimations which may involve multi-level considerations) arise.

There are of course in additions, statistical experts who are not RSS Fellows.

I am happy to discuss these issues with you at any time.

With all good wishes
Sharon Witherspoon
Vice President, Education and Statistical Literacy at the Royal Statistical Society