

**Awarding grades for the June 2020  
examination series:  
Qualifications Wales-regulated AS  
Methods report**

August 2020

## Contents

Introduction.....	3
Background .....	3
Key assessment principles .....	3
Sources of evidence for calculating grades .....	6
Marks/grades already obtained by candidates for assessments already completed as part of the qualification and historical data about qualification functioning.....	6
Internal assessment grades for work completed but not yet externally moderated or verified .....	7
Centre assessment grades (CAG) and rank orders .....	7
Analysis of centre assessment grades .....	9
Centre data and information .....	10
Prior attainment data .....	11
Grade calculation approaches .....	12
Direct centre performance (DCP) approach with no prior performance.....	12
Direct centre performance (DCP) approach with mean GCSE prior attainment and private candidate adjustments .....	13
Direct centre performance (DCP) approach with mean GCSE prior attainment, private candidate and small centre cohort adjustments.....	14
Testing results and final model selection .....	15
The final approach .....	16
Process stages .....	16
Data inputs .....	17
Run model .....	17
Initial calculation stage.....	17
Grade distribution and adjustment stage .....	18
Final grade allocation stage (including small centres adjustment) .....	18
Centres with no historical data in the subject .....	18
Slotting-in stage .....	19
Decision-making group.....	19
Responsible Officer sign-off .....	20
Qualifications Wales review.....	21
Final grades awarded.....	21

# Introduction

## Background

Welsh Government announced the cancellation of summer 2020 examinations on 18<sup>th</sup> March, due to the ongoing COVID-19 public health crisis<sup>1</sup>. Following this announcement, Welsh Government issued Qualifications Wales with a direction that the summer 2020 cohort of GCSE, AS, A-level and Welsh Baccalaureate Skills Challenge Certificate learners should be issued with results this summer. The approach adopted had to be fair and robust, based on centres' judgements of their learners' attainment in each subject, and standardised using a range of other evidence<sup>2</sup>. 'Standardisation', in this context, is a process involving the use of statistical models to calculate grades.

Following a public consultation, Qualifications Wales confirmed their expectation that the models will produce grade outcomes for this summer that are *broadly similar* to previous years, and that candidates across the cohort should be awarded a set of grades that are, overall, a fair reflection of what they would have received had they sat exams. All candidates, including private candidates where possible, will receive a calculated grade.

This document sets out recommendations for the approach to calculating grades for **AS** candidates. All AS qualifications were awarded in the summer 2019 examination series, meaning that there are no new qualifications being awarded for the first time in summer 2020. AS qualifications regulated by Ofqual but designated by Qualifications Wales for use in Wales are not covered by this technical report.

## Key assessment principles

In selecting approaches to calculating grades for learners, we have been guided by the key principles of assessment – validity, reliability, fairness, manageability, and comparability.

Qualifications Wales defines **validity** as follows:

*“the extent to which the assessment tests the things it is supposed to assess. The use(s) of the outcome(s) of an assessment is/are valid if supported by evidence and theory. The evaluation of validity involves the development of a clear argument to support the proposed interpretation of the outcomes and the intended uses of the assessment. The validity argument should be built on statements of the proposed interpretation and supporting evidence collected from all stages of the assessment process”.*

---

<sup>1</sup> Welsh Government (2020). *Cabinet Statement: Written Statement: Written Statement on Summer Examinations 2020*. <https://gov.wales/written-statement-written-statement-summer-examinations-2020>

<sup>2</sup> Welsh Government (2020). *Letter to Qualifications Wales from Minister for Education*, 6<sup>th</sup> April, <https://gov.wales/gcse-and-level-cancellations-letter>

This report is a form of validation argument. Ordinarily, validation arguments make claims about each feature or stage of the assessment process, to justify the judgements that we wish to make about a candidate's ability in the subject domain of interest<sup>3</sup>. As the grade calculation process is intended to replace an assessment process, here the validation argument is a justification of decisions made at each stage of calculation to date. The aim is therefore to demonstrate that – as far as is possible – the grades issued to candidates this summer will be as fit-for-purpose as the grades issued to candidates in any other examination series.

For general qualifications, the main purpose is to give a measure of learners' attainment that supports their future progression into work or further education<sup>4</sup>. This purpose has not changed. Under the heading of validity, therefore, the grade calculation process can be evaluated using the following principles.

Firstly, the **accuracy** of the approach used to calculate grades must be maximised. As we cannot know what grade outcome each learner would have achieved in this series, we need to use historical data to inform the choice of a model. In testing, we are likely to prefer models which most often correctly predict the grade that learners achieved via examinations, in a normal examination series. The choice of approach may differ for different qualifications, depending on what information is deemed to be most effective in correctly predicting outcomes.

Linked to accuracy is the need to deliver **fairness** for learners. In this examination series, we must ensure that any **bias** in the process of calculating grades against candidates with common attributes is minimised. Common attributes may include those covered by statutory equality duties, such as age, sex, race and disability; but also other considerations, such as socio-economic status, size of centre, and the language medium of entry. Considerations of bias are limited by data availability and quality, including data from previous examination series that can be used for comparison.

So that the goal of facilitating appropriate progression for learners is secured, grades must be **comparable** in meaning with previous series. Comparability of outcomes has two purposes in Qualifications Wales' definition: to ensure that fair comparison can be made about the attainment of learners with grades from different examination series and qualifications; and that outcomes can be used as a measure of standards in a subject over time<sup>5</sup>. In a normal series, the latter would be ensured via an awarding process and the principle of comparable outcomes, so that candidates are not advantaged or disadvantaged by any variation in challenge in the examination series in which they complete their qualification<sup>6</sup>. In this series, outcomes are expected to be 'broadly similar to previous series'. How this is achieved in grade calculation depends to some extent on the approach taken. This summer, some may also choose to compare approaches to calculating grades across qualifications and/or nations, to evaluate the extent to which grades are seen as comparable.

---

<sup>3</sup> Paul Newton (2017). *An approach to understanding validation arguments*. Coventry: Ofqual.

<sup>4</sup> Welsh Government (2020). *Letter to Qualifications Wales from Minister for Education*, 6<sup>th</sup> April, <https://gov.wales/gcse-and-level-cancellations-letter>

<sup>5</sup> Qualifications Wales (2020). Standard Conditions of Recognition. <https://www.qualificationswales.org/english/publications/standard-conditions-of-recognition/>

<sup>6</sup> Qualifications Wales, *A Closer Look at the Comparable Outcomes Approach*. <https://www.qualificationswales.org/media/4806/comparable-outcomes-approach.pdf>

The grade calculation process must also be **reliable**. Reliability usually relates to consistency – that each stage of the process would result in the same outcome if repeated. In a normal series, issues such as grade classification accuracy, quality of marking, and grade boundary checks are important aspects of reliability. When calculating grades, other aspects come to the fore, which relate closely to quality assurance and quality control:

- *Modelling considerations*: independent checks of model functioning.
- *Candidate-level considerations*: ensuring that candidates can only receive grades that they could reasonably have received in a normal series; accounting for likely resit improvement, early and multiple entry and banked assessment marks/grades; ensuring that candidate grade profiles across all subjects are similar to those seen in previous series.
- *Centre-level considerations*: plausibility checks of centre rank order, based on prior assessment data; reviews of centre outcome stability relative to previous series; accounting for known partnership arrangements between centres.

End-to-end, the grade calculation process should also be as **manageable** as possible for candidates, centres and those responsible for producing and quality-assuring the grades. Any reliable grade calculation process will place burdens on each of these groups, just as an examination series does. An unmanageable or overly burdensome process puts the intended purpose of calculating grades at risk, however. Models which are overly complicated or require too many data inputs may also create risks: parsimonious models have an advantage both in risk reduction but also in being more easily understood by learners, centres and wider users of qualifications and grades.

At all times, the recommendation and selection of the approach for calculating grades requires a balancing of factors relating to these key principles, but always focused on achieving the aims set out in the Welsh Government's direction and the requirements set out by Qualifications Wales.

## Sources of evidence for calculating grades

A range of evidence is required for calculating and quality-assuring grades. Types of evidence can be grouped into six categories<sup>7</sup>. Under each category, the sources of evidence available for Qualifications Wales-regulated AS qualifications have been reviewed to establish their potential value.

### **Marks/grades already obtained by candidates for assessments already completed as part of the qualification and historical data about qualification functioning**

Where performance on one unit (for which a mark has been banked) is predictive of performance in another unit, it may be appropriate to either calculate marks or grades for assessments that candidates were unable to complete in summer 2020, or for the qualification as a whole. Alternatively, the evidence could be used to quality assure centre assessment grades and rank orders.

Qualifications Wales-regulated AS qualifications are based on a unitised structure, with unit assessments available in the summer examination series each year. A majority of candidates study for the qualification in one year, taking the AS units in their first year before (if so desired) taking 'A2' units in the second year in order to achieve an A level grade. AS units are weighted at 40% of the A-level qualification. Achieved marks are converted to a uniform mark scale (UMS) for each unit via the awarding process in each examination series.

Unlike at A-level, there is not a significant volume of quality-assured banked evidence from previous assessments that could be used to inform the calculation of final grades. This also means that information on how performances at unit level relate to each other is of limited use in calculated grades. One possible exception to this are resitters, for whom marks will be banked from a previous series.

---

<sup>7</sup> This framework is derived from Ofqual (2020). [Exceptional arrangements for assessment and grading in 2020: consultation on the assessment and grading of vocational, technical and other general qualifications.](#)

## Internal assessment grades for work completed but not yet externally moderated or verified

Non-examination assessment is a key element of many AS qualifications. Before centres in Wales were formally closed in March, candidates may have completed some or all of the assessment, and in some cases, centres may have begun to mark and internally standardise completed work. Some centre moderation visits had taken place. Many candidates had not completed their assessments, however, and WJEC was not able to visit all centres prior to their closing or to fully quality assure the work of all visiting moderators.

Consequently, in order not to disadvantage candidates, and because the marks are not quality-assured, WJEC confirmed that provisional marks completed prior to the closure of centres in awarding grades would not be used<sup>8</sup>.

Similarly, some oral and practical examiner visits also took place prior to schools closing; however, not all visits had taken place and therefore not all centres have been assessed. In order not to disadvantage any candidates, WJEC did not progress with the quality assurance processes for centres whose candidates had been marked. As these marks have not been quality assured, WJEC did not use these in awarding grades either.

## Centre assessment grades (CAG) and rank orders

The Welsh Government direction requires that centres' judgements on candidates' attainment are used within the grade calculation process. As WJEC has not collected grade estimates for some years, there is no recent data to analyse to assess the accuracy of centre predictions. Including this data in the grade calculation process is important to validity as, in the absence of assessment performance evidence, teachers are uniquely placed to consider the attainment of their learners. There is evidence that teachers are better able to judge the attainment of their learners *relative* to each other than using an absolute estimate of the grades they will ultimately achieve (an *absolute* judgement), as teacher grade estimates tend to be positively biased<sup>9</sup>. For AS qualifications in Wales, there will not usually be prior unit marks or grades that can be used to inform judgements.

Centres' judgements have therefore been gathered in two forms.

- A *centre assessment grade*, based on what teachers would expect each candidate to achieve for each qualification, representing a fair, reasonable, and carefully considered judgement of the most likely grade that might be achieved in normal circumstances.
- A *rank order position* for each candidate within each grade. Centres were permitted to 'tie' candidates on a single rank, based on a scale dependent on the size of the centre's total entry. This is shown in *Table 1*. The scale was developed using a combination of statistical analysis and stakeholder feedback, based on the following principles:
  - we should not expect centres to be able to rank their learners to a greater degree of differentiation than the examination system would normally provide. In a normal examination

---

<sup>8</sup> WJEC (2020). [Coronavirus FAQs - Non-Examination Assessment \(NEA\)](#).

<sup>9</sup> Tim Gill, Methods used by teachers to predict final A-level grades for their students. *Research Matters: a Cambridge Assessment publication*, 28 (Autumn 2019), pp.33-42

series, centres will often see pairs or small groups of candidates achieve the same mark and grade;

- it will be necessary to limit the number of ties per candidate, so that the grade calculation process can function effectively – in particular to ensure that outcomes are broadly similar to previous series;
- based on analysis of data from previous examination series, there is a relationship between the size of the candidature and the number of other learners a candidate may share a result with within a centre. It may therefore be appropriate to allow more ties per candidate when a centre’s cohort is larger.
- there was no need to extend the number of ties permitted beyond 15, even for the largest centres.

**Table 1: Rules for allowing tied rank order positions**

Maximum group rank size: candidates can be tied into groups of...	Number of candidates in cohort (all ages)	Maximum proportion at bottom of grade rank
No ties permitted	0-49	n/a
2	50-99	4.00%
3	100-149	3.00%
4	150-199	2.67%
5	200-249	2.50%
6	250-299	2.40%
7	300-349	2.33%
8	350-399	2.29%
9	400-449	2.25%
10	450+	2.22%

All candidates were ranked on a single rank order, regardless of age, for each subject. For AS, this information was gathered from centres at qualification level, between 1st June and 12<sup>th</sup> June 2020. Centres were given general guidance on how to collate and submit the required information<sup>10</sup>, as well as subject-specific guidance on how to produce the grades and rank orders. This professional judgement was derived from evidence held within the centre (learner work or evidence of learner work), which had been reviewed by subject teachers/tutors/assessors and relevant heads of department. Private candidates were included in centres’ submissions where centres believed that they had seen sufficient evidence to make a reliable judgement. In some cases, centres were unable to do this for particular private candidates and therefore centre assessment grades and rank orders were not collected for these candidates.

Heads of centres were required to confirm that their grades and rank orders had been checked for accuracy, and represented a fair, objective and professional judgement of the grades that their learners would have been most likely to achieve in a normal series.

---

<sup>10</sup> Qualifications Wales (2020). Summer 2020 grades for GCSEs, AS and A levels, and Skills Challenge Certificate (SCC) Information for Centres on the submission of Centre Assessment Grades <https://qualificationswales.org/media/5973/information-for-centres-on-the-submission-of-centre-assessment-grades-version-2-18-may-2020.pdf>



### Analysis of centre assessment grades

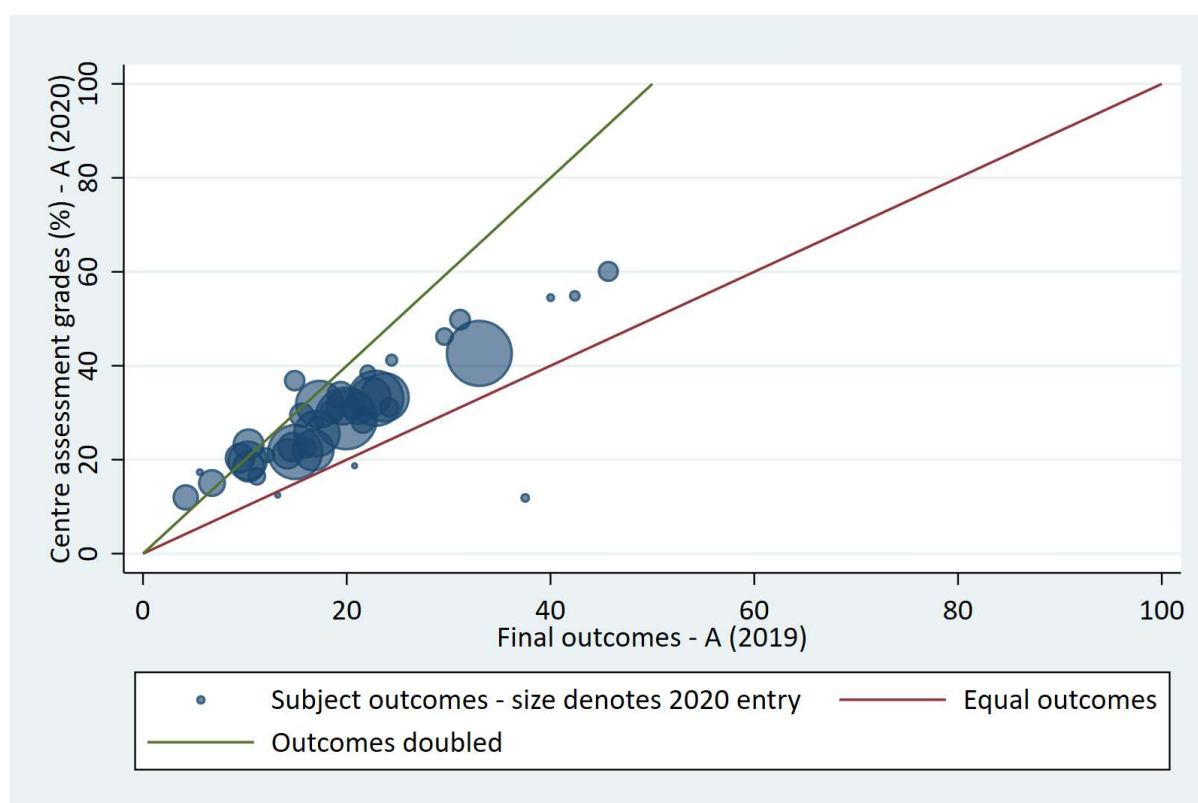
An analysis of the centre assessment grades provided to WJEC, shown in *Table 2*, indicates higher outcomes than summer 2019 outcomes were at the end of the period for review of marking and moderation.

**Table 2: Centre assessment grades in summer 2020, compared with final results in summer 2019 (Wales 17-year-olds only)**

Subject	A	B	C	D	E	U	n
Summer 2019 (final)	19.58	38.59	59.75	76.84	88.66	100.00	30788
Summer 2020 (CAG)	29.55	56.25	80.86	92.99	98.50	100.00	34273

As *Figure 1* shows, the difference is reflected across most subjects at Grade A. A similar pattern is seen at all other grades.

**Figure 1: Centre assessment grades in summer 2020, compared with final results in summer 2019, by subject (Wales 17-year-olds only) – proportion of candidates achieving a grade A**



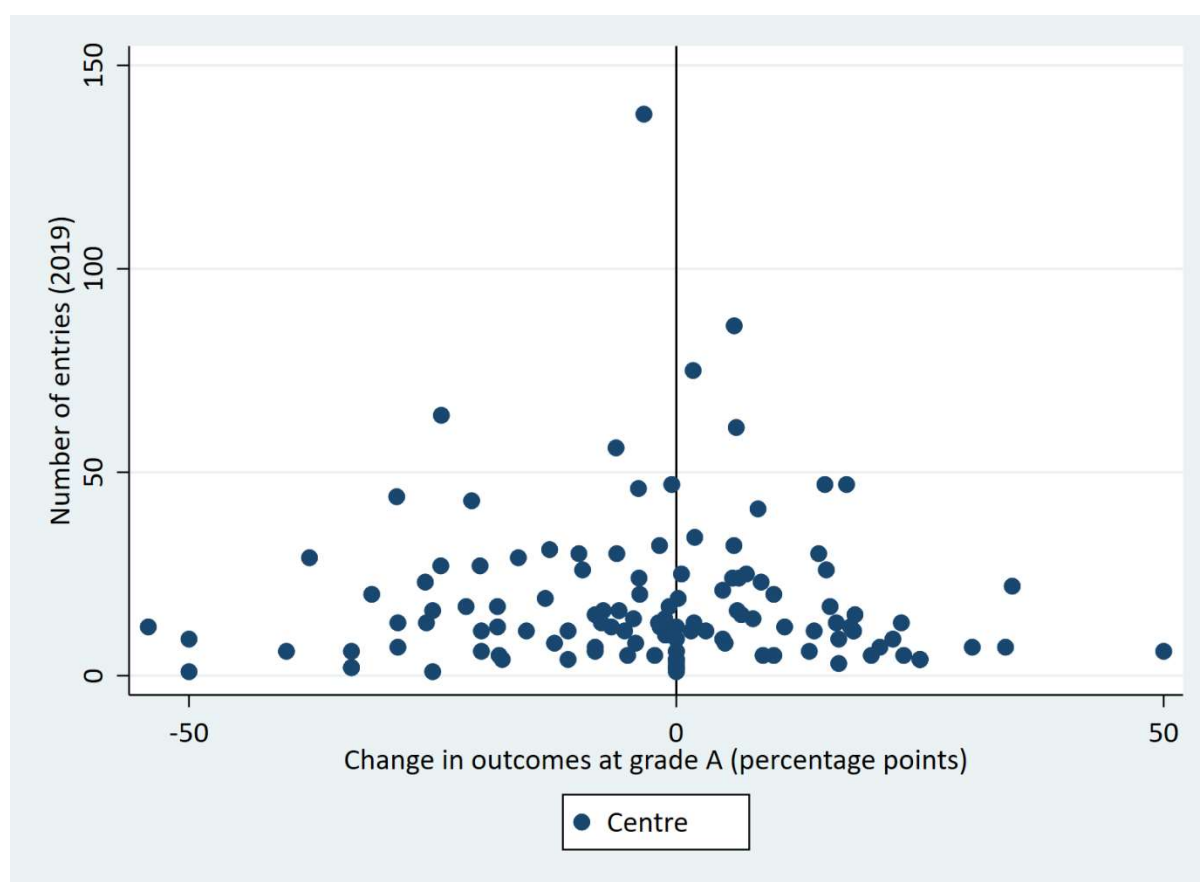
It is understandable that centre assessment grades are positive, as no specific methodology was provided, thus suggesting that such information should be treated with care when calculating grades. Given the volume of information that was used by teachers in the process of determining rank orders, the rank orders in particular can be considered to be reliable.

## Centre data and information

Centres' outcomes in previous series are a useful source of evidence in calculating and quality-assuring grades. Overall, we might expect that the pattern of variation in centre outcomes between series should be similar between 2018 and 2019, and between 2019 and 2020. At centre level, outcomes contain information about the performance of learners over time and the value-added provided by the centre. Centre outcomes may also differ by centre cohort size, with greater variation in outcomes amongst centres with small cohorts each year, or where cohorts differ markedly in size each year.

Figure 2 shows the centre-level variation at grade A for AS Chemistry between 2018 and 2019. Smaller centre cohorts have outcomes which vary widely from series to series but, in general, centres with larger cohorts have more stable outcomes. Candidates' performance or attainment may differ, however, or the strength of a centre's cohort may differ from series to series – so some variation is to be expected.

Figure 2: Centre-level variation at grade A, for AS Chemistry, between 2018 and 2019



### **Prior attainment data**

Candidates' mean GCSE scores (based on candidates' grades in the summer of the academic year they turn 16) have been used as part of the awarding process for GCE qualifications for many years<sup>11</sup>. At candidate level, there is a moderately strong relationship ( $R^2=0.559$ ) overall between Wales 17-year-olds' performance at AS (expressed as UMS) and their mean GCSE decile; at qualification level, this varies between 0.37 and 0.84. Prior attainment data can give an indication of the strength of a cohort at centre or qualification level from series to series.

---

<sup>11</sup> *cf.* Benton, T. (2015). Can we do better than using 'mean GCSE grade' to predict future outcomes? An evaluation of Generalised Boosting Models. *Oxford Review of Education*, 41:5, pp.587-607.

## Grade calculation approaches

In this section, the key approaches developed for calculating Wales AS grades this summer are described. In order to tie the approaches with the analysis presented in this document, the outline of the procedure behind each of the methods is exemplified based on the modelling exercises undertaken using historical data, the results of which are presented in the next section of the report. Incorporating real examination outcomes served as the most reliable way of verifying the robustness of the methods, although it can never account for any issues of quality relating to centre assessment grades and rank order positions.

Based on the experience of modelling A-level outcomes, combined with the preliminary modelling exercise undertaken by Ofqual and the other examination boards for general qualifications in England, the Mark Based Regression (MBR) method was ruled out as not being suitable for estimating AS grades on the basis that it produced less accurate grades than other approaches. In addition, as most candidates take all AS units in a single year, the use of normalised scores (the BUCP method proposed for A-level) was not considered to be a viable option given the absence of high volumes of banked assessment evidence. WJEC therefore concluded that the most appropriate methods for standardising AS grades are based on the 'direct centre performance' approach'. Three different variations of the method were created with the goal of finding the most accurate approach.

These three methods are all broadly based on the approach being developed by Ofqual for the general qualifications they regulate

### Direct centre performance (DCP) approach with no prior performance

This is very close to a 'common centre analysis'. The underlying assumption for this approach is that the percentage of candidates achieving any given grade or better in a centre in 2020 (2019 for testing purposes) would be the same as the percentage who achieved that grade in the centre between 2017, 2018 and 2019 (2016, 2017 and 2018 for testing purposes). The only addition to this principle is that, as proposed by Benton and Bramley (2020)<sup>12</sup>, each centre has one extra "historic candidate" added who received the same outcome as the national rate.

The historic percentage at each grade is calculated and then an extra candidate is added to represent the national rates. This extra candidate helps pull the outcomes slightly closer to the national outcomes. For example, if the national percentage getting an A grade was 10% and historically only one candidate had sat the subject between 2016 and 2018 in a particular centre, our prediction for the number of A grades at that centre would either be.

- If the historic candidate had an A =  $(100\% + 10\%)/2 = 55\%$
- If the historic candidate did not get an A =  $(0\% + 10\%)/2 = 5\%$

---

<sup>12</sup> Benton, T. and T. Bramley (2020), *Estimating grades for candidates in GCSEs and A-levels in summer 2020 – Cambridge Assessment suggested approach*. Paper presented to Ofqual Standards and Technical Implementation Group (STIG), March 2020.

Using this method, a prediction at each grade is calculated for each centre. This is used to decide the basket of grades that each centre is going to receive. Once that basket is decided the grades are allocated according to the rank order provided by the centre.

Finally, candidates have an imputed score calculated based on their ranking and their initial standardised grade. If the overall outcomes from the process are different to the required grade distribution (for testing purposes, the final outcomes for 2019 were used as the target grade distribution) then the candidates closest to the imputed score boundaries are moved in order to hit the prediction. For example, if the proportion of candidates achieving getting a grade C (or higher) was lower than expected then the top D grade candidates (based on the imputed score) would be moved up one grade.

### **Direct centre performance (DCP) approach with mean GCSE prior attainment and private candidate adjustments**

This model is very similar to the 'direct centre performance' approach tested for calculated grades at A-level in Wales and is the same as the method proposed by Benton and Bramley.

In the same way as the model without prior attainment the outcomes from previous series (2016, 2017 and 2018 for the purpose of testing), centres' historic performance is calculated for each subject.

The centre performance projections are then adjusted according to the difference in the prior attainment profile of the prior attainment 'matched' 17-year-old candidates at the centre, in the current and historical examination series. The prior attainment measure used is the mean GCSE score used to generate statistical predictions for grade outcomes in normal examination series. The difference in the prior attainment profiles is essentially the difference between two subject predictions for the centre. The weighting of the prior attainment difference adjustment is dependent on the proportion of the centre subject entry cohort that comprises matched 17-year-olds. The proportions of matched candidates in the historical and current examination series is calculated for each centre, and the lowest of these two values is used.

The predicted proportion of candidates at each grade within each centre is converted to a basket of grades for the centre, and grades are allocated accordingly to reflect the centre assessment rank order. Where required, allocation thresholds can be moved to bring overall outcomes into line with expected outcomes, as with the basic DCP method described in the previous section.

One final adjustment for this model is that private candidates are excluded from both the historical and current data. If they have a centre assessment grade and rank order position they are 'slotted into' the centre's grade distribution with the grade that is closest to their centre assessment grade, but is not more than the grade of the candidate ranked above them nor below the grade of the candidate ranked below them.

### **Direct centre performance (DCP) approach with mean GCSE prior attainment, private candidate and small centre-cohort adjustments**

This model is based on work that Ofqual undertook on how to account for small centre cohorts in their standardisation model. The model assumes that when the number of candidates being entered by a centre is small, the statistical information is weaker. The model therefore gives more weight to the centre assessment grades if a centre's cohort for a subject is small. In testing, it is important to note that the simulated centre assessment grade used was the actual grade achieved by the candidate grade received in 2019. Therefore, giving more weight to the centre assessment grade will always make the results more similar to the observed results in 2019, creating a bias in the test.

A harmonic mean of the total number of candidates entered for a qualification in the historical data, and the total number of candidates in the current series, is calculated. Any centre with a harmonic mean of 15 or higher has its grades calculated in exactly the same way as outlined for the model without the small centre adjustment. These centres are used as the basis for achieving the target grade distribution in the model.

At the end of the process, any centre with a harmonic mean of five or fewer entries has their candidates' grades changed to be the same as that of their centre assessment grades. Any centre with a harmonic mean between 5 and 15 gets a weighted average of their centre assessment grades and their projected grade. The outcomes for a centre with a harmonic mean of 6 would receive grades heavily weighted towards their CAGs; a centre with a harmonic mean of 14 would receive grades weighted towards the centre's standardised grade distribution.

## Testing results and final model selection

The models were tested for a representative cross-section of A-level subjects, covering mathematics, science technology; languages (including modern foreign languages); creative arts; humanities and social sciences. The testing provided strong evidence that the model should include an adjustment for candidates' prior attainment at GCSE. Additionally, the method was found to cause no substantial concerns over differences by either gender or age. The approach of removing private candidates with rank order positions from the initial model, and then slotting them into the rank order, was also accepted as the fairest way of accounting for these candidates without disadvantaging or advantaging the centres where they have entered.

The most difficult issue to address related to centres with small cohorts. In most cases, small cohorts are small in both the current year and historically. This means that we are modelling a standardised grade distribution based on weaker statistical evidence about those centres' true underlying long-term performance in a subject (due to centre-level variation) and even weaker statistical evidence about the relationship between outcomes and prior performance.

Our recommendation was to include the small cohorts adjustment within the final approach. Although at this point, WJEC was aware that the distribution of centre assessment grades was higher than the grade distribution in previous year, it was felt that including a relatively small inflationary effect on outcomes would be preferable to relying on the weak statistical evidence that the model alone provides in these cases. This risked introducing punitive adjustments on candidates in small centre-cohorts without strong statistical evidence. Crucially, the approach means that candidates in larger cohorts are not disadvantaged as outcomes for these centres can still be aligned to a target grade distribution that does not disadvantage them. Small cohorts are treated consistently regardless of the size of the centre's overall intake. We agreed to review the impact of this approach in the decision-making group meetings.

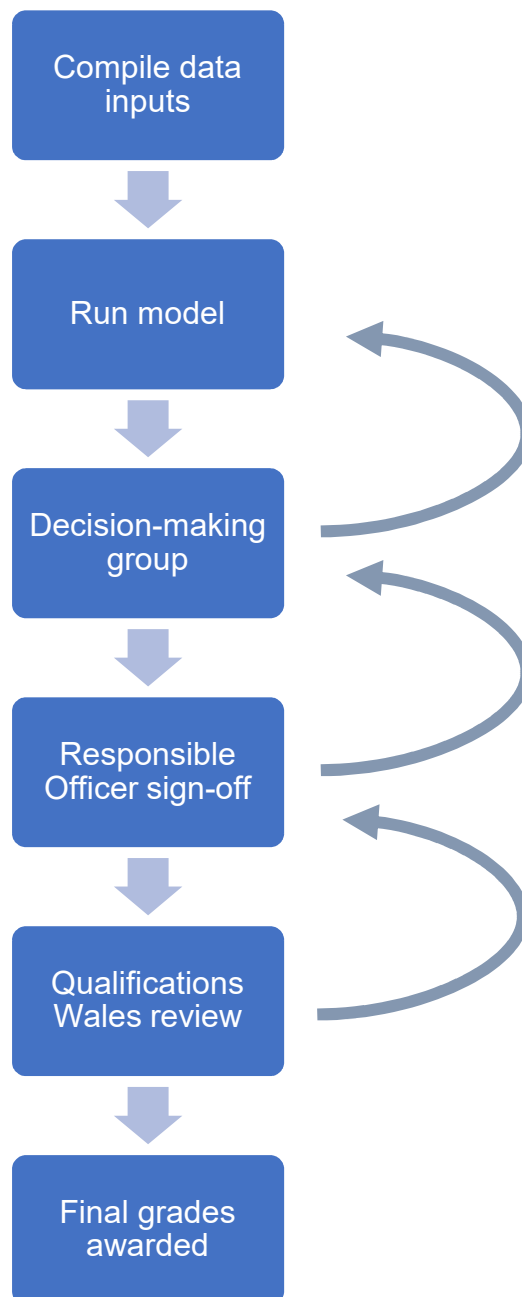
A common approach to calculating grades was recommended across the suite of qualifications for validity and manageability reasons, as the proposed approach produced good accuracy outcomes (compared with actual 2019 grades awarded) across all subjects. This approach was approved by Qualifications Wales in July 2020. Centre-level variation was also measured for the selected model; patterns of variation were generally in line with those seen from examination series to series, with relatively few outliers.

## The final approach

### Process stages

Figure 3 sets out the approach taken to standardising grades for AS qualifications. Each stage is explained in more detail below. Note that several of the stages are iterative. Depending on the decisions made at each stage, a qualification could be referred back to a previous stage. This could be in order to consider additional analysis, adjust the target outcome set within the model, to check data inputs, or to undertake additional quality assurance prior to the completion of the process.

Figure 3: Stages of the standardisation approach





## Data inputs

Results data for Qualifications Wales-approved GCE qualifications was used as the first historical data input, in line with the requirements set down by Qualifications Wales in the Data Requirements for Summer 2020<sup>13</sup> document. This was combined with the centre assessment grades and rank order information provided by centres, as described previously in this report.

Prior attainment data for GCE qualifications is produced and quality assured annually by JCQ members for the purposes of setting and maintaining standards, on a three-country<sup>14</sup> basis, reflecting Qualifications Wales' Data Requirements. Candidates' mean GCSE scores are based on a common grade conversion which accounts for the fact that GCSE grade scales (9-1 for reformed Ofqual-regulated GCSES, A\*-G for legacy GCSEs and reformed GCSEs in Wales and Northern Ireland) differed between qualifications in 2019, when the June 2020 17-year-old AS cohort took most of their GCSEs. In the AS model, this information is used both within the model itself and for the purposes of producing statistical predictions of the grade distribution of matched 17-year-old candidates in each subject.

## Run model

Once all data was compiled and quality assured, the standardisation model was run.

### *Initial calculation stage*

- For each centre, a harmonic mean of the total number of candidates entered for a qualification in the historical data, and the total number of candidates in the current series, was calculated. Any centre with a harmonic mean of 15 or higher was included in the initial run of the model. These centres are used as the basis for achieving the target grade distribution in the model. Private candidates were excluded from that part of the model.
- The centre performance projections were then adjusted according to the difference in the prior attainment profile of the prior attainment 'matched' 17-year-old candidates at the centre, in the current and historical examination series. The weighting of the prior attainment difference adjustment is dependent on the proportion of the centre subject entry cohort that comprises matched 17-year-olds.
- The adjusted centre performance projections were then used to create a grade distribution for each centre, and grades are allocated to each candidate in the centre based on their rank order position.
- The initial standardised grade and rank order position were then used to derive an imputed score for each candidate.

---

<sup>13</sup> Qualifications Wales (2020). *Wales Summer 2020 Data Requirements. GCE, GCSE, Welsh Baccalaureate Skills Challenge Certificate qualifications*. <https://www.qualificationswales.org/english/publications/wales-summer-2020-data-requirements---gce-gcse-and-welsh-baccalaureate-skills-challenge-certificate-qualifications/>

<sup>14</sup> England, Wales and Northern Ireland.

### ***Grade distribution and adjustment stage***

'Cut scores' (the values at which scores are classified as representing one grade or another) are then set to assign all scores to a grade. It is possible to amend these values to bring overall outcomes closer to a predefined grade distribution. In the initial model run, the model parameters were set to award grades to a distribution based on principles agreed with Qualifications Wales.

- The cumulative proportion of the cohort achieving each grade should be higher than the cumulative proportion of candidates achieving that grade in 2019 after reviews of marking and moderation had been completed.
- The cumulative proportion of the 'matched 17-year-old' cohort (for whom mean GCSE prior attainment is available) achieving each grade should be higher than the statistical prediction defined by Qualifications Wales in the Data Requirements document for this series<sup>15</sup>.

At subsequent stages of the standardisation process, target outcomes were adjusted to reflect the considerations of the decision-making groups, as well as the Responsible Officer, Standards Officer and Qualifications Wales.

### ***Final grade allocation stage (including small centres adjustment)***

Once cut score values are set, final grade allocations were produced for each centre, based on the imputed score distribution for their candidates. Grades were then distributed to the candidates included in the calculation stage of the model, based on the rank order provided by each centre.

Small centres – any centre with a harmonic mean of five or fewer – received their candidates' centre assessment grades accepted as their final grades. Any centre with a harmonic mean between 5 and 15 received as its grade allocation a weighted average of their centre assessment grades and the projected grade distribution as calculated via the model. The outcomes for a centre with a harmonic mean of 6 would receive grades heavily weighted towards their CAGs; a centre with a harmonic mean of 14 would receive grades weighted towards the centre's standardised grade distribution.

In line with the recommendation made to Qualifications Wales, the thresholds at which small centre adjustments are made – 5 and 15 – were reviewed as part of the decision-making process. Following analysis of the impact of moving the thresholds downwards, WJEC and Qualifications Wales concluded that no change would be made to the thresholds.

### ***Centres with no historical data in the subject***

In a small number of cases, a centre had no previous entries in the subject in years included in the historical data source. In these circumstances, in the absence of any other information on the centre's cohort, centre assessment grades were awarded as the final grades to candidates.

---

<sup>15</sup> Qualifications Wales (2020). *Wales Summer 2020 Data Requirements. GCE, GCSE, Welsh Baccalaureate Skills Challenge Certificate qualifications*. <https://www.qualificationswales.org/english/publications/wales-summer-2020-data-requirements---gce-gcse-and-welsh-baccalaureate-skills-challenge-certificate-qualifications/>

### **Slotting-in stage**

Candidates not included in the calculation stage of the model were 'slotted into' a grade according to their centre assessment grade (CAG) and rank order position, so that each of these candidates received the closest grade to their CAG which does not break the centre's rank order.

For example, for candidate X, if the candidate above X in the centre rank order received a grade B via the model, and the candidate below X receives an E, then if X's CAG is a B or better they will get a B; if it is E or worse they will get an E and otherwise they will be awarded their CAG on the basis that it falls between the grades for candidates ranked either side of X.

An analysis of outcomes, setting out aggregated entries and proposed grade distributions, was then prepared for the decision-making groups to consider.

### **Decision-making group**

A Calculation of Grades Subject-Specific Decision-Making Group was convened for each qualification being awarded in the summer 2020 examination series. Each group comprised:

- Responsible Officer (Director of Qualifications and Assessment Delivery) – Chair of meeting
- Standards Officer (Assistant Director (Standards, Processing and Research))
- Assistant Director(s) from the Qualifications and Assessment Delivery directorate
- Subject Officer(s) and/or domain leader(s) for the qualification(s) being discussed.

The purpose of the group was to review the approach taken to determining grades for each candidate entered for the qualification(s) being reviewed, and to decide if the approach is approved, or required further review or amendment. In making this decision, group members were asked to consider the following key principles.

- Validity and comparability – ensuring the grades issued in this series are fit-for-purpose, supporting appropriate progression for learners, that reflects their levels of attainment, and that are comparable in meaning to grades issues in previous series and by other awarding organisations where appropriate.
- Reliability – ensuring that the grade calculation process has sufficient quality controls and assurances in place.
- Fairness – that any biases in outcomes are minimised, so that learners with common attributes are not unreasonably adversely affected by the process of calculating grades.

Groups were either convened to meet remotely via Microsoft Teams, or virtually (in that no meeting was convened and feedback was gathered from group members for consideration by the Responsible Officer and Standards Officer once all group members had considered the proposed model and outcomes). Any personnel with a conflict of interest against any of the qualifications being discussed was asked to declare it before discussions commenced.

A Grading Partner from the Research & Standards team was responsible to preparing the statistical evidence for each qualification to inform the group, and then presented the key findings to the group. This evidence included:

- entries analysis: a breakdown of entries by age, gender, medium, country, centre type, private candidates, centre size, prior attainment, and other aspects where relevant.
- methods: an overview and justification of the statistical model used to calculate outcomes.
- grading outcomes: a comparison of cumulative grade distributions from the previous series, centre assessment grades (CAG), proposed final outcomes by the model. A breakdown by age, gender, centre type, and other groups was also provided where data was available.
- centre variation: including the change between cumulative CAG outcomes and cumulative proposed final grades at centre level for key grades.

Group members were then invited to provide their views on the approach and the rationale for the proposed outcomes, as well as any other concerns arising from the report.

### **Responsible Officer sign-off**

The Responsible Officer, accounting for the statistical evidence presented and the feedback from Group members, decided either to:

- to accept the recommended approach and outcomes in full
- to accept the recommended approach and/or outcomes, subject to a revision
- to refer the qualification for further review by the Research & Standards department.

If qualifications were referred for further review, or a revision was requested, further evidence was presented to the Responsible Officer for further consideration prior to sign-off. If necessary, decision-making group members were invited to give additional feedback once the review was completed.

## Qualifications Wales review

Grade outcomes were reported to Qualifications Wales as part of the Summer 2020 Data Requirements and discussed at regular bilateral Standards meetings between WJEC and Qualifications Wales, prior to final approval in August 2020. Once approval was received from Qualifications Wales, the final standardised grades were processed.

## Final grades awarded

The grade outputs from the model were then assigned into grading systems. Several additional quality assurance stages were applied prior to issuing results.

- In a small number of cases, candidates had enough banked unit scores to cash-in the qualification and no unit entries. These candidates were awarded a grade based on the banked evidence, and their centres were asked not to include them in their centre assessment grade submission.
- In some other cases, where candidates were resitting units and cashing-in the qualification the standardisation model produced a grade for candidates which their existing banked evidence would have exceeded. These candidates were awarded a grade based on their banked evidence. This meant the centre rank order was broken in some instances, by agreement with Qualifications Wales, to uphold the principle of fairness. No candidate received a lower grade because of this adjustment.
- Private candidates with no centre assessment grade or rank order position were eligible to receive a calculated grade where they met criteria set down by Qualifications Wales<sup>16</sup>. All of these candidates had banked marks for AS units previously. For units which candidates had planned to re-sit, an uplift to the UMS score was applied in line with that applied to A-level candidates who had intended to re-sit AS units this summer. For each unit, the best available UMS was used to aggregate the final mark, and the grade awarded on the basis of the standard UMS grade boundaries as set out in the qualification's specification.

---

<sup>16</sup> Qualifications Wales (2020). Private candidate policy statement.  
<https://www.qualificationswales.org/media/6184/private-candidate-policy-statement.pdf>