

Wales Summer 2020 Data Requirements



**GCE, GCSE and Welsh Baccalaureate Skills
Challenge Certificate qualifications**

Contents

Section 1	3
a. Summer 2020 requirements.....	3
b. Summer 2020 process	4
Section 2	5
Generating Statistical predictions.....	5
GCE A level.....	6
GCE AS.....	7
GCSE.....	8
Welsh Baccalaureate.....	9
Aims for the issue of qualification grades.....	10
Section 3	11
Standardisation of centre assessment grades.....	11
Section 4	12
Reporting.....	12
Section 5	16
Sending data.....	16
Annex A	17
Annex Ai	17
Data specification for historical learner data.....	17
Annex Aii	25
Data specification for learner prior attainment data for GCSE learners.....	25
Annex B	29
Data specification for summer 2020 learner data.....	29
Annex C	36
Aims underpinning the statistical models and issue of grades summer 2020	36
Annex D	37
Principles for using the secure data sharing site.....	37

Section 1

a. Summer 2020 Requirements

The following procedures outline the process that must be followed and the data that must be provided by WJEC during summer 2020 for each qualification forming part of these requirements.

This is a **Regulatory Document** under **Condition B7** of the Standard Conditions of Recognition¹ and the relevant conditions of The Extraordinary Regulatory Framework: General Qualifications – COVID-19 Conditions and Requirements.

Condition GQCov3.2(a)(i) allows us to specify requirements and guidance in relation to the calculation of results to be issued for GQ Qualifications in summer 2020.

This document, together with the annexes to it, sets out our requirements for the purposes of Condition GQCov3.2(a)(i).

An awarding body must comply with these requirements in calculating each result for a GQ Qualification¹ that it issues under Condition GQCov3.1.

Under Condition GQCov1.6, where there is any inconsistency between the requirements set under Condition GQCov3.2(a)(i) and a requirement elsewhere in our regulatory framework, an awarding body must comply with the former.

For clarity, and without prejudice to the generality of Condition GQCov1.6, in respect of results issued under Condition GQCov3.1, an awarding body is not required to comply with existing requirements published under the following Conditions, save where those requirements set out the relevant specified levels of attainment (grades) to be used –

- a) in relation to post-reform GCSE and GCE Qualifications, Condition ASC.Q1, and
- b) in relation to Welsh Baccalaureate Skills Challenge Certificate, Condition WB.M1.

¹ <https://qualificationswales.org/english/publications/standard-conditions-of-recognition/>

Condition GQCov3.2(a)(i) allows us to supplement the requirements set out in this document and its annexes with particular specifications. This allows for the possibility that, because of the unique situation for results in summer 2020, some technical amendments may need to be made with respect to these requirements as they are worked through by awarding bodies.

Throughout these requirements, reference is made to a 'subject'. By this is meant a particular GQ Qualification (e.g. GCSE English language), or a particular GQ Qualification awarded with specific endorsements (e.g. A level design and technology: fashion and textiles) as set out in Annex A.

This Regulatory Document applies to Wales' reformed GCE, GCSE and Welsh Baccalaureate Skills Challenge Certificate qualifications, and Wales specific legacy qualifications. Procedures for pre-reform GCE and GCSE qualifications which are awarded on a three-country basis are available in the [Summer 2020 requirements](#) document produced by Ofqual on behalf of the three regulators.

Grades should be issued at qualification level only. The rules set out in the regulatory document 'Requirements for setting specified levels of attainment for GCE and GCSE qualifications'² would be suspended for summer 2020.

b. Summer 2020 process

In respect of a GQ Qualification which it makes available, WJEC must follow the process set out below in order to calculate each result issued under Condition GQCov3.1 –

1. WJEC must work with other awarding bodies to collectively construct a historical learner level dataset according to the requirements in Annex A. The historical dataset must include data from all relevant qualifications in the relevant series and year.
2. WJEC must construct a collective summer 2020 learner level dataset according to the requirements in Annex B. The learner level dataset must include data from all awarding bodies offering the relevant qualifications in summer 2020.

² <https://www.qualificationswales.org/english/publications/requirements-for-setting-specified-levels-of-attainment-for-gce-and-gcse-qualificationspdf/>

3. WJEC must generate statistical predictions for the relevant subject, as specified in Section 2.
4. Applying the approved standardisation process, as referred to in section 3, WJEC must use:
 - a. the historical learner data set out in Annex A,
 - b. the summer 2020 learner level data set out in Annex B, and
 - c. the statistical predictions as specified in Section 2,to standardise the information provided by Centres under Condition GQCov3.2.
5. WJEC must provide Qualifications Wales with the information specified in Section 4 and according to the timetable specified.

Section 2

a. Generating statistical predictions

WJEC must generate predictions for all grades using the summer 2020 dataset generated under Annex B and the outcome matrices generated as in section 2f. Predictions must be generated for learners for whom there is a valid match as defined in Annex A (referred to in this Annex as 'matched learners'). Predictions must be generated for learners in Wales only.

WJEC must generate predictions for their specific cohort of matched learners for these subjects.

b. GCE A Level

b.i Predictions for reformed A level qualifications (phase 1 – first awarded in 2017)

WJEC must create for each specification a prediction for the Wales cohort of 18-year-old matched cash-in learners. The prediction must be based on the average of the national³ relationship between:

1. A level outcomes in that subject for 18-year-old learners in 2017 and GCSE outcomes for those learners in 2015.
2. A level outcomes in that subject for 18-year-old learners in 2018 and GCSE outcomes for those learners in 2016.

b.ii Predictions for reformed A level qualifications (phase 2 –first awarded in 2018)

WJEC must create for each specification a prediction for the Wales cohort of 18-year-old matched cash-in learners. The prediction must be based on the national relationship between:

1. A level outcomes in that subject for 18-year-old learners in 2018 and GCSE outcomes for those learners in 2016.
2. A level outcomes in that subject for 18-year-old learners in 2019 and GCSE outcomes for those learners in 2017.

b.iii Predictions for reformed A level qualifications (phase 3 –first awarded in 2019)

For all subjects WJEC must create a prediction for the Wales cohort of 18-year-old matched cash-in learners. The prediction must be based on the average of the national relationship A level outcomes in that subject for 18-year-old learners in 2019 and the GCSE outcomes for those learners in 2017.

³ National here, and throughout, means the prediction must be based on the average (weighted by entry size) of the three-country relationship applied to the Wales cohort.

b.iv Predictions for A level Health and Social Care, ICT and Applied ICT

WJEC must create for the Wales cohort of 18-year-old matched cash-in learners. The prediction must be based on the A level outcomes in that subject for 18-year-old learners in 2019 and the GCSE outcomes for those learners in 2017.

c. GCE AS Level

c.i Predictions for reformed AS qualifications (phase 1 –first awarded in 2016)

WJEC must create for each specification a prediction for the Wales cohort of 17-year-old matched cash-in learners. The prediction must be based on the average of the national relationship between:

1. AS outcomes in that subject for 17-year-old learners in 2016 and GCSE outcomes for those learners in 2015.
2. AS outcomes in that subject for 17-year-old learners in 2017 and GCSE outcomes for those learners in 2016.

c.ii.Predictions for reformed AS qualifications (phase 2 –first awarded in 2017)

WJEC must create for each specification a prediction for the Wales cohort of 17-year-old matched cash-in learners. The prediction must be based on the average of the national relationship between:

1. AS outcomes in that subject for 17-year-old learners in 2017, and GCSE outcomes for those learners in 2016.
2. AS outcomes in that subject for 17-year-old learners in 2018, and GCSE outcomes for those learners in 2017.

c.iii.Predictions for reformed AS qualifications (phase 3 –first awarded in 2018) and ICT and Applied ICT

WJEC must create for each specification a prediction for the Wales cohort of 17-year-old matched cash-in learners. The prediction must be based on the national relationship between:

1. AS outcomes in that subject for 17-year-old learners in 2018 and GCSE outcomes for those learners in 2017.
2. AS outcomes in that subject for 17-year-old learners in 2019 and GCSE outcomes for those learners in 2018.

c.iv.Predictions for AS Health and Social Care

WJEC must create a prediction for the Wales cohort of 17-year-old matched cash-in learners. The prediction must be based on the average of the national relationship between AS outcomes in that subject for 17-year-old learners in 2019 and GCSE outcomes for those learners in 2018.

d.GCSE qualifications

d.i.Predictions for reformed GCSE qualifications (phase 1 – first awarded in 2017)

WJEC must create for each specification predictions for the Wales cohort of 16-year-old matched cash-in learners. The predictions must be based on common centres analysis, taking the average of:

1. Best grade⁴ outcomes of the 16-year-old learners in the 2016/17 cohort.
2. Best grade⁵ outcomes of the 16-year-old learners in the 2017/18 cohort.

d.ii.Predictions for reformed GCSE qualifications (phase 2 –first awarded in 2018)

WJEC must create for each specification its own prediction for the Wales cohort of 16-year-old matched learners. The predictions must be based on common centre analysis taking the average of the relationship between:

1. GCSE outcomes in that subject for 16-year-old learners in 2018.
2. GCSE outcomes in that subject for 16-year-old learners in 2019.

⁴ Best grade outcomes may have been achieved from any series entered by the candidate.

⁵ Best grade outcomes may have been achieved from any series entered by the candidate.

d.iii. Predictions for reformed GCSE qualifications (phase 3 – first awarded for the first time in 2019)

WJEC must create for each specification its own prediction for the Wales cohort of 16-year-old matched learners. The predictions must be based on the data from common centre analysis of the outcomes achieved in that subject by 16-year-old learners from summer 2019.

d.iv. Predictions for reformed GCSE short course Physical Education (phase 3 – first awarded in 2018)

For GCSE Physical Education WJEC must create a prediction for the Wales cohort of 16-year-old matched learners. Predictions must be based on data from common centres in 2019.

d.v. Predictions for GCSE ICT (full course and short course)

WJEC must create for the full course specification its own prediction for the Wales cohort of 16-year-old matched learners. The predictions must be based on the data from common centre analysis of the outcomes achieved in that subject by 16-year-old learners from summer 2019.

d.vi. Predictions for GCSE Health & Social Care, Home Economics: Child Development

WJEC must create for each single award specification its own prediction for the Wales cohort of 16-year-old matched learners. The predictions must be based on the data from common centre analysis of the outcomes achieved in that subject by 16-year-old learners from summer 2019.

e. Welsh Baccalaureate Skills Challenge Certificates - Foundation (post-16) Skills Challenge Certificate, National (post-16) Skills Challenge Certificate, Foundation/National Skills Challenge Certificate and Advanced Skills Challenge Certificate

For the Advanced Skills Challenge Certificate WJEC must create predictions for the Wales cohort of 18-year-old matched cash-in learners. The predictions must be

based on the national relationship between Skills Challenge outcomes for 18-year-old learners in summer 19 and the GCSE results for those learners in 2017.

WJEC should also note the Principles for issue of Welsh Baccalaureate (Skills Challenge Certificates) in Appendix 3.

f. Outcome matrices

Where applicable, WJEC must generate outcome matrices using the historical learner level datasets as specified in Annex A, and the outcome matrix code specified in the Summer 2020 subject lists. Outcome matrices must exclude partial absences. Outcome matrices must only include learners in Wales.

Each outcome matrix must map the relationship between prior attainment and results for the relevant subject in the relevant reference year for learners for whom there is a valid match as defined in Annex A (referred to in this Annex as 'matched learners').

In generating outcome matrices for GCSE Qualifications, WJEC must not include data with respect to learners registered with independent or selective centres.

g. Aims for the issue of qualification grades

WJEC should note the Aims for the issue of qualification grades for reformed GCSE, AS and A level, and Skills Challenge Certificate qualifications in Appendix C.

Section 3

Standardisation of centre assessment grades

a. Standardisation procedure

WJEC must use the procedures approved by Qualifications Wales for calculation of grades for GCSE, GCE AS and GCE A level qualifications and Welsh Baccalaureate Skills Challenge Certificates.

Technical details for the standardisation procedures will be published by WJEC:

- In respect of GCE AS and A levels – Thursday 13 August.
- In respect of GCSE – Thursday 20 August.

b. Private candidates

An exceptional method should be applied to those private candidates who are not able to be provided with a centre assessment grade and rank order. This will apply to private candidates who are eligible, as laid out in the Private Candidate policy statement⁶ and in accordance with Conditions GQCov 3.1, GQCov3.5 and GQCov3.6. Details of the method will be published, along with technical details of the standardisation process.

c. Small cohorts

Some allowances in the standardisation procedures will be made for small cohorts, since less statistical information may be available.

⁶ <https://www.qualificationswales.org/media/6184/private-candidate-policy-statement.pdf>

Section 4

Reporting

a. GCE A level reporting

WJEC must report outcomes for matched learners and outcomes for all learners for each subject at all grades, using the templates provided and according to the reporting schedule in Table 1.

WJEC must provide Qualifications Wales with a schedule of dates to show when they expect to submit A level award outcomes and provide this to Qualifications Wales by **Monday 22 June 2020**. Submission should normally be two days after the provisional sign-off of the award (for example, if the provisional sign-off takes place on a Monday the outcomes should be submitted on the following Thursday), but there may be exceptions, for example when the outcomes for a suite of specifications need to be reviewed together before the awards are finalised.

Template A must be updated, according to the schedule provided to Qualifications Wales, to show A level results to date.

In a few cases WJEC may need to carry out further analyses post-award⁷ in order to establish the security of the outcomes. If there is to be a delay in submission, WJEC must notify Qualifications Wales as soon as possible and by the date shown on the schedule.

b. GCE AS level reporting

WJEC must report outcomes for matched learners and outcomes for all learners for each subject at all grades, using the templates provided and according to the reporting schedule in Table 1.

WJEC must provide Qualifications Wales with a schedule of dates to show when they expect to submit AS level award outcomes and provide this to Qualifications Wales by **Monday 22 June 2020**. Submission should normally be two days after

⁷ Awards refer to the Calculation of Grades Subject-Specific Decision-Making Group meeting which will be convened by WJEC for each qualification being awarded in the summer 2020 examination series. This applies throughout this section.

the provisional sign-off of the award (for example, if the provisional sign-off takes place on a Monday the outcomes should be submitted on the following Thursday), but there may be exceptions, for example when the outcomes for a suite of specifications need to be reviewed together before the awards are finalised.

Template B must be updated, according to the schedule provided to Qualifications Wales, to show AS level results to date.

In a few cases WJEC may need to carry out further analyses post-award in order to establish the security of the outcomes. If there is to be a delay in submission, WJEC must notify Qualifications Wales as soon as possible and by the date shown on the schedule.

c. GCSE reporting

For all GCSE full and short course specifications and GCSE double awards WJEC must report the prediction, outcomes for matched learners, outcomes for all learners, cut-scores and tiering data (for tiered subjects only) for each subjects at all grades using the templates provided and data specification(s).

For GCSE art and design, the outcomes from the suite of options must be reported as a whole.

For GCSE French, German and Spanish we would expect the outcome is positive relative to prediction to compensate for the demand of the subject⁸.

WJEC must provide Qualifications Wales with a schedule of dates to show when they expect to submit GCSE award outcomes and provide this to Qualifications Wales by **Monday 22 June 2020**. Submission should normally be on the two days after the provisional sign-off of the award (for example, if the provisional sign-off takes place on a Monday the outcomes should be submitted on the following Thursday), but there may be exceptions, for example when the outcomes for a suite of specifications need to be reviewed together before the awards are finalised.

WJEC must update **Templates C and D** according to the schedule provided to Qualifications Wales.

⁸ <https://qualificationswales.org/english/news/mfl-nov-2019/>

WJEC must report all of its June 2020 GCSE Mathematics and Mathematics – Numeracy results for re-sitting 16-year-olds compared with the results for previous best grade for the same learners, and all results for the Mathematics compared to the results for Mathematics-Numeracy for the same learners, using **Template E**.

In a few cases WJEC may need to carry out further analyses post-award in order to establish the security of the outcomes. If there is to be a delay in submission, WJEC must notify Qualifications Wales as soon as possible and by the date shown on the schedule.

d. Welsh Baccalaureate Skills Challenge Certificate reporting

WJEC must report all of its June 2020 Skills Challenge Certificate outcomes, using **Template F**.

e. Reporting concerns for GCE A level, GCE AS level, GCSE and Skills Challenge Certificates

WJEC must report any concerns with awarding outcomes. These must be reported as soon as possible after the award and sign-off by the Responsible Officer, together with an explanation, using **Template G**.

f. Reporting schedule

WJEC must submit data according to the schedule in Table 1.

	Deliverable	Data specification	Date (time?)
GCE A level	WJEC predictions, outcomes for matched learners, outcomes for all learners and cut-scores by subject		9am Tuesday 28 July
GCE AS	WJEC predictions, national outcomes for matched learners, national outcomes		9am

	for all learners and cut-scores for each subject		Tuesday 28 July
Advanced SCC	WJEC predictions, outcomes for matched learners, outcomes for all learners and cut-scores for each challenge		9am Tuesday 28 July
GCSE	WJEC predictions, outcomes for matched learners, outcomes for all learners and cut-scores for each subject		Noon Thursday 30 July
National and Foundation SCC	WJEC predictions, outcomes for matched learners, outcomes for all learners and cut-scores for each challenge		Noon Thursday 30 July

Section 5

Sending data

WJEC must confirm their data exchange contacts with Qualifications Wales before the process begins. All data are to be uploaded to the secure data sharing site (see Appendix 5).

All returns are subject to the final ratification of the awards by WJEC.

On receipt of the above information, Qualifications Wales:

1. will review the data on the basis of its published procedure.
2. may contact WJEC to ask for additional information or clarification and aim to do this within 24 hours of receipt.
3. will hold regular teleconferences with WJEC technical colleagues, as necessary, to review data and explore any implications.
4. will discuss outcomes with WJEC at the Standards meetings on **30 July 2020 (GCE and Advanced Welsh Baccalaureate)** and **3 August 2020 (GCSE and National and Foundation Welsh Baccalaureate)**.

Annex A

Annex Ai

Data specification for historical learner data

WJEC must collaborate with the other awarding bodies awarding the relevant subjects in summer 2020 to construct historical learner level datasets split into the following three files –

1. Level 2 qualification file relating to GCSE subjects.
2. Level 3 qualification file comprising GCE AS, GCE A level and extended project qualifications.

The datasets comprising these files must use the following data –

- a. For all phase 1 and 2 GCSE subjects, historical learner data from 2018 and 2019.
- b. For all phase 3 GCSE subjects, historical learner data from 2019.
- c. For all GCE AS and A level subjects historical learner data from 2017, 2018 and 2019.

The data used to generate the datasets must be based on mark/grade held by awarding bodies as at 23 May 2020. This means that it must reflect any changes to results made after they have been issued through, for example a review of marking or appeal.

The data used must relate to results from summer assessment series only and to subjects in the subject list (available separately).

Awarding body data based on learner-level mark/grade data as of 23 May 2020. Data to only include summer series results, except for phase 1 reformed Wales GCSEs for which November 2019 grades are included.

Variable Name	Variable Description	Source	Notes
ReportingDate	Date data submitted	Awarding body data	
ExamSeries	Exam series of results	Awarding body data	Month Year
Country	Country of centre	Generated from National Centre Number (NCN)	Based on NCN as below: Wales = 68000 to 68999 Scotland = 69000 to 70999 N Ireland = 71000 to 71999 Isle of Man = 45000 to 45999 Channel Isles = 67000 to 67999 Other = any other
QualificationLevel	Level of qualification	Awarding body data	
AwardingOrganisation	Name of Awarding organisation of qualification	Awarding body data	
QAN	Qualification number	Awarding body data	
ReformPhase	Phase of reform	Awarding body data	

JCQCode	JCQ subject classification code (broad subject area)	Subject list	
JCQTitle	JCQ subject classification Title (broad subject area)	Subject list	
IBTitle1	Inter-board subject code, broad subject area	Subject list	
IBTitle2	Inter-board subject code, fine subject area	Subject list	
OMCode	Outcome matrix number (used to group qualifications for predictions)	Subject list	
OMTitle	Outcome Matrix subject title	Subject list	
SpecificationCode	Awarding body code used for specification	Awarding body data	
SpecificationTitle	Title of specification	Awarding body data	
CertificationCode	Certification code for qualification	Awarding body data	

CentreNo	NCN centre number	Awarding body data	
CentreType	Type of centre	NCN Centre type data lookup file	<p>Centre type matched to NCN using the NCN register. Where centres are not included in the NCN register the centre type is missing</p> <p>Centre types are based on the following groups</p> <ul style="list-style-type: none"> 1 – Secondary Comprehensive or Middle (Community, Voluntary Aided Controlled) 2 – Secondary Selective (Community, Voluntary Aided Controlled) 3 – Secondary Modern (Community, Voluntary Aided Controlled) 4 – Secondary Comprehensive or Middle (Foundation) 5 – Secondary Selective (Foundation) 6 – Secondary Modern (Foundation) 7 – Independent 8 – FE Establishment 9 – Sixth Form College 10 – Tertiary College 11 – Other (including private candidates) 13 – City Academy 14 – Free Schools <p>Matched from NCN register by AQA using NCN</p>
UniqueCandidateIdentifier	Candidate ID	Awarding body data	
YearEndAge	Age of learner at end of year exams were taken	Calculated based on date of birth in the awarding body data	<p>Age calculated as of 31 August in the corresponding academic year for all learners except those in Northern Ireland</p> <p>Age calculated as of 30 June in the corresponding academic year for all learners in Northern Ireland</p>

Year GCSE sat when aged 16 [GCE data only]	Indicates the year the prior GCSE attainment measures come from	Obtained from prior GCSE attainment dataset	
NGCSEsTaken [GCE data only]	Number of GCSEs taken by learner at age 16	Obtained from prior GCSE attainment dataset	<p>NGCSEs and Mean GCSE calculated based on GCSE qualifications only (does not include level 1/level 2 certificates or international GCSEs)</p> <p>NCRA variables matched with awarding body data by AQA and duplicated by WJEC, using procedure specified in this Annex. Matching only undertaken for learners that took 3 or more GCSEs</p> <p>NCRA variables matched for learners aged 17 – 23 years old</p> <p>Normalised mean GCSE calculated (and duplicated) by Ofqual</p>
MeanGCSE [GCE data only]	Learners mean GCSE grade	Obtained from prior GCSE attainment dataset	
TenPointScaleFlag [GCE data only]	Flag to indicate if mean GCSE is on 10-point scale	Obtained from prior GCSE attainment dataset	
Normalised Mean GCSE_ENGLAND [GCE data only]	Mean GCSE score normalised by age within year for all learners who took GCSEs at England centres only	Obtained from prior GCSE attainment dataset	
NormalisedMeanGCSE_UK [GCE data only]	Mean GCSE score normalised by age within year for all learners who took GCSEs at UK centres	Obtained from prior GCSE attainment dataset	

Year KS2 sat when aged 11 [GCSE data only]	Indicates the year the prior KS2 attainment measure comes from	Added based on the prior KS2 attainment dataset	
KS2Score [GCSE data only]	Learners Normalised KS2 score	Added based on the prior KS2 attainment dataset	Normalised mean KS2 score calculated by AQA KS2 score matched with awarding body data by AQA and duplicated by WJEC, using procedure specified in this Annex KS2 score matched for learners aged 15 – 18 year olds
PartialAbsence	Whether the learner was absent for some of their assessments within the qualification	Awarding body data	
SubjectLevelGrade	Grade learner was awarded on qualification	Awarding body data	Includes any post-results changes (i.e. grade is the final grade awarded to learners)
PrivateCandidate	Indicates if learner is a private candidate for that specification	Awarding body data	
Learner name		Awarding body data	
Learner sex		Awarding body data	
Learner DoB		Awarding body data	

Finding a match with prior GCSE attainment for level 3 qualifications

In respect of each Learner identified in the Level 3 qualification file WJEC must use the following process to find a match and determine if it should be classified as a valid match.

Name Processing

Names are processed in the following way:

- Convert full name to uppercase.
- Surname is everything before a delimiter of “.” or “,” or “;”
- Forename is everything after the delimiter (excluding any leading space(s) that follow the delimiter) until the next space, bracket, hyphen or comma
- Leading/trailing spaces and special characters are removed from surname and forename (any of ;',.,(){}[]+=-><?!/\\$@&%*)

Assigning a Match Level

Any match against the prior GCSE attainment dataset is graded according to the following levels:

Level Items Matched

- 1 UCI3, surname, forename, DoB, sex
- 2 Surname, forename, DoB, sex
- 3 UCI, surname, forename, DoB
- 4 UCI, surname, forename, sex
- 5 UCI, surname, DoB, sex
- 6 UCI, surname, forename
- 7 UCI, surname, forename, (where forename and surname are swapped)
- 8 UCI, surname, DoB

9 UCI, surname, sex

10 UCI, surname

11 UCI, forename, DoB, sex

12 Surname, forename, DoB

When matching learners with prior attainment, it is important to reduce the probability of false-positives (these are instances where a match is erroneously confirmed).

When applying this specific algorithm, a valid match is defined as one which satisfies any one of the matching criteria for Levels: 1, 2, 3, 4, 6 and 7.

Duplicates

Duplicates within any of the above levels are removed completely to improve confidence in the match. In practice, this will only occur at levels 2 or 11, as the UCI within the prior GCSE attainment database is unique.

Welsh Government matched KS3 attainment data.

KS3 data are matched by UPN (there are robust systems in place to ensure this number is unique and that it follows the same child as they move between schools).

Annex Aii

Data specification for learner prior attainment data for GCSE learners

WJEC must use learner prior attainment data, detailed below, provided by the Welsh Government to model GCSE outcomes.

Field	Source	Use
UPN (WJEC only)	KS4 (WED)	to identify unique learners
ULN (WJEC only)		to identify unique learners
UCI (WJEC only)		To identify unique learners
Name (WJEC only)		To identify unique learners
LEA Number		To allow outputs of statistical models to be considered at LEA level
Establishment Number		For use in modelling of results in centres and to allow outputs of modelling to be considered by centre
Centre Number (KS4)		For use in modelling of results in centres and to allow outputs of modelling to be considered by centre
Age		To identify/verify cohorts and for statistical use to understand relationships in assessment data
Date of birth (WJEC only current entries only)		To enable data linkage

Newbes		to use to help understand missing data and potentially as a filter for including/excluding learners from models should the value added be atypical
Gender		For data linkage and for equalities impact analysis
Exam number		ID for exam taken
Exam board		to filter data by board
Year taken		To help identify first vs last grades
Exam Level		Qualification info
Welsh Indicator		For statistical use to understand relationships in assessment data
QAN Code		allows easy id of qual including whether it is legacy or reformed version
Main Grade		provides qualification outcome data needed for statistical modelling of historical cohort data
Additional Grade		provides qualification outcome data needed for statistical modelling of historical cohort data
Consortium		To enable us to look at the outputs of standardisation models at consortium level
Welsh Baccaulaureate		Qualification outcome information needed for statistical modelling
Year		To identify cohorts and when grades where achieved
Ethnicity		For statistical use to understand relationships in assessment data for equalities impact analysis
FSM Eligible		For statistical use related to equalities impact analysis

SENStatus		For statistical use related to equalities impact analysis
EAL		For statistical use related to equalities impact analysis
Year	Key Stage 3 assessments	Cohort identification
KS3 Centre ID		Centre linked to key stage 3 TA scores and test scores for the learner
Keystage		To identify key stage associated with assessment results
KSSubject		To identify subjects associated with assessment results
Method		To confirm assessment method
Result		To identify assessment results
NCYearActual		To identify/confirm cohorts
Eng TA	Key Stage 3 only	To create learner ability measure for standardising GCSE grades
Cym TA	Key Stage 3 only	To create learner ability measure for standardising GCSE grades
Mat TA	Key Stage 3 only	To create learner ability measure for standardising GCSE grades
Sci TA	Key Stage 3 only	To create learner ability measure for standardising GCSE grades
CSI TA	Key Stage 3 only	To create learner ability measure for standardising GCSE grades
ART	Key Stage 3 only	To create learner ability measure for standardising GCSE grades

DAT	Key Stage 3 only	To create learner ability measure for standardising GCSE grades
GEO	Key Stage 3 only	To create learner ability measure for standardising GCSE grades
HIS	Key Stage 3 only	To create learner ability measure for standardising GCSE grades
ICT	Key Stage 3 only	To create learner ability measure for standardising GCSE grades
MFL	Key Stage 3 only	To create learner ability measure for standardising GCSE grades
MUS	Key Stage 3 only	To create learner ability measure for standardising GCSE grades
PED	Key Stage 3 only	To create learner ability measure for standardising GCSE grades
WEL	Key Stage 3 only	To create learner ability measure for standardising GCSE grades
For year 9 only		
KSSubject		To identify test
Component		To identify test
TestPaperSat		To identify test
AgeStandardisedScore		To provide score for statistical modelling
Raw score		To provide score for statistical modelling
Maximum score		To provide score for statistical modelling

Annex B

Data specification for summer 2020 learner data

WJEC must collaborate with the other awarding bodies awarding the relevant subjects in summer 2020 to construct summer 2020 learner level datasets split into the following three files –

1. Level 2 qualification file comprising GCSE subjects.
2. Level 3 qualification file comprising GCE AS, GCE A level and extended project qualifications.

The data used to generate the datasets must be based on data held by awarding bodies, and prior attainment matching based on entries, as at 23 May 2020.

The data used must relate to Summer 2020 subjects list (available separately).

The data must be recorded in the form specified in the table below.

Variable Name	Variable Description	Source	Notes
Country	Country of centre	Calculated from first two digits of National Centre Number (NCN)	Based on NCN as below: UK = 10000 to 71999 Wales = 68000 to 68999 Scotland = 69000 to 70999 N Ireland = 71000 to 71999 Isle of Man = 45000 to 45999

			Channel Isles = 67000 to 67999 England = UK minus (Wales + Scotland + N Ireland + IoM + CI)
AwardingOrganisation	Name of Awarding organisation of qualification	Awarding body data	
UniqueCandidate Identifier	Candidate ID	Awarding body data	
YearEndAge	Age calculated from the year in which the learner certificated	Calculated based on the date-of-birth in the awarding body data	Age calculated as of 31 August in the corresponding academic year for all candidates except those in Northern Ireland Age calculated as of 30 June in the corresponding academic year for all learners in Northern
NGCSEsTaken [GCE data only]	Number of GCSEs taken by learner at age 16	Obtained from prior GCSE attainment dataset	Learners matched against prior GCSE attainment data when they were aged 16. All learners of ages: 17, 18, 19, 20, 21, 22 and 23 matched with prior GCSE attainment.
MeanGCSE [GCE data only]	Learners mean GCSE score	Obtained from prior GCSE attainment dataset	NGCSEs and Mean GCSE calculated based on GCSE qualifications only (does not include level 1/level 2 certificates or international GCSEs). Normalised mean GCSE calculated by Ofqual
TenPointScaleFlag [GCE data only]	Flag to indicate if mean GCSE is on 10-point scale	Calculated based on the specific prior GCSE attainment dataset.	

NormalisedMeanGCSE_England [GCE data only]	Mean GCSE score normalised by age within year for learners who took GCSEs at England centres only.	Obtained from prior GCSE attainment dataset	GCSE prior attainment variables matched with awarding body data using procedure specified in Annex A. Matching only undertaken for learners that took 3 or more GCSEs
NormalisedMeanGCSE_UK [GCE data only]	Mean GCSE score normalised by age within year for learners who took GCSEs at UK centres.	Obtained from prior GCSE attainment dataset	
Year GCSEsatWhenAged16 [GCE data only]	Indicates the year the prior GCSE attainment measures come from	Obtained from prior GCSE attainment dataset	
DoubleNormalisedKS3Score [GCSE data only]	Learners Double-Normalised KS3 score	Obtained from prior KS3 attainment dataset	Learners matched against prior attainment KS3 data. To calculate the double-normalised score the average Yr9 score for each learner is calculated. This average Yr9 score is then standardised within each year (subtract mean, then divide by standard deviation) to provide the normalised average Yr9 score. The average KS3 teacher assessment level is calculated for each learner. This average KS3 teacher assessment level is then standardised within each year and then standardised. Then the average of the standardised average Yr9 score and the standardised KS3 teacher

			assessment level is taken. This average is then standardised within year to provide the 'double-normalised' score.
AS_UMS	UMS score from AS units	WJEC data	The uniform mark scale (UMS) score for learners taking A level units.
GCSE_UMS	UMS score from GCSE units (GCSE Biology, Chemistry, English literature, Physics, Science (Double Award) only)	WJEC data	The uniform mark scale (UMS) score for learners taking unitised GCSE qualifications, for those units used as a basis for standardising grades.
Learner name		Awarding body data	
Learner Sex		Awarding body data	
Learner DoB		Awarding body data	
Entry Code	Awarding body code used for specification. This is equivalent to the SpecificationCode field in the historic datasets	Awarding body data	
Pearson tier-of-entry [GCSE data only]	Used for Pearson tiered GCSE specifications only	Awarding body data	

Entry Title	Title of specification	Awarding body data	
Learner ULN	Learner ID	Awarding body data	
Learner UPN	Learner ID	Awarding body data	
Learner number	Learner ID (4-digit number)	Awarding body data	
Private Candidate Flag	Indicates if learner is a private candidate for that specification	Awarding body data	
Centre number	NCN centre number	Awarding body data	
Centre Type	Type of centre	NCN centre type data lookup file	<p>Centre types are based on the following groups</p> <ul style="list-style-type: none"> 1 – Secondary Comprehensive or Middle (Community, Voluntary Aided Controlled) 2 – Secondary Selective (Community, Voluntary Aided Controlled) 3 – Secondary Modern (Community, Voluntary Aided Controlled) 4 – Secondary Comprehensive or Middle (Foundation) 5 – Secondary Selective (Foundation) 6 – Secondary Modern (Foundation) 7 – Independent 8 – FE Establishment 9 – Sixth Form College

			10 – Tertiary College 11 – Other (including private candidates) 13 – City Academy 14 – Free Schools Matched from NCN centre type data lookup file using NCN
Subject Grouping	Subject group mapping	JCQ subject mapping data	
Cross AO Flag	Flag to indicate whether a centre has split learner entries across AOs for a specific 'Subject Grouping'	Y/N	
Qualification Level	Level of qualification		
Type	Qualification Type		
Reform Phase	Phase of reform	JCQ subject mapping data	
JCQ Code	JCQ subject classification code (broad subject area)	JCQ subject mapping data	
JCQ Title	JCQ subject classification title (broad subject area)	JCQ subject mapping data	

IB Title1	Inter-board subject code (broad subject area)	JCQ subject mapping data	
IB Title 2	Inter-board subject code (fine subject area)	JCQ subject mapping data	
Outcome Matrix Code	Outcome matrix number	JCQ subject mapping data	
Outcome Matrix Title	Outcome matrix subject title	JCQ subject mapping data	

Annex C

Aims underpinning the statistical model and issuing of grades in summer 2020

The issue of GCSE, GCE AS and GCE A level, and Skills Challenge Certificate qualifications must be consistent with the four aims which underpin the statistical model and issuing of grades this summer.

The four aims are:

1. All learners for whom a qualification-level centre assessment grade and rank order are submitted will receive a grade.
2. National outcomes will be broadly similar to those in previous years to reduce the risk of unfairness for learners over time and maintain public confidence.
3. As far as possible the process for awarding grades will not systematically advantage or disadvantage learners, including those with characteristics protected by equalities legislation.
4. The standardisation model will use a range of evidence to calculate the likely grades that learners would have achieved, had they been able to complete their assessments this summer.

Annex D

Principles for using the secure data sharing site

1. The site is to be used for sensitive data relating to the data exchange.
2. Email should be used as usual where no sensitive data is included or attached.
3. In terms of version control, all templates must be uploaded using the date in the format YYYYMMDD, as a prefix, for example 20160701-WJEC-Template 1–08.30. Therefore, whenever a revised data template is uploaded, please create this as a new document using the date as a prefix and the time as a suffix where necessary.
4. Old versions of templates must not be deleted from the site.
5. WJEC users must not create extra folders within the data exchange library of the portal without consulting Qualifications Wales first.