

GMC Assessment collection – FAQs

How will we submit data for this collection?

In this year of the collection, your return will consist of your existing results or exam board spreadsheets (along with any supplementary materials, as discussed in another FAQ below). Colleagues at the GMC will review and process these files to populate the STUDENT_ASSESSMENT table described in:

https://www.hesa.ac.uk/collection/c21055/data_items

The [data collection system](#) will be open from early September until November 30 2022, and access to this will be managed by your institution's nominated HESA contact. Further information on how this system works, along with other details regarding submissions for this collection, can be found in [this e-learning module](#).

How will these data be processed once collected?

The GMC team will load the submitted spreadsheets into the UKMED database held on GMC servers. We will derive some variables, including Z-scores within year, school and assessment and a classification of assessments into groups, such as "first year written exam". The approach to classifying assessment and calculating Zscores was decided with statistical input from the UKMED research sub-group.

We will provide each school with a load report that will include the following information:

- For each assessment, the proportion of cases in the main HESA Student records that have a record for that assessment
- Details of derived variables
- Exception report – details of any cases that did not load, for example linkage failures

How will these data be used?

These data are being collected for inclusion in the [UK Medical Education Database](#), which provides collated data on the performance of UK medical students and trainee doctors across their education and careers, for research and planning purposes.

UKMED data is accessible only via a rigorous application and review process, with data being provided in the form of research extracts tailored to meet the specific requirements of each application. Information on data security and anonymity can be found in another FAQ below.

Data will be used to produce an interactive Tableau report looking at the award gap in assessments UK-wide and within schools, this will show differences in Z-scores by protected characteristics such as ethnicity. This will be published under password to schools for comment and review. We would then expect to publish a key-findings report on the award gap within UK medical schools. At some point in the future the report may be included with the other GMC progression reports - [online reporting tool](#).

How will anonymity and data protection be ensured?

Any data released through UKMED are subject to strict controls on how the data is accessed, including procedures to prevent identification of individuals. In brief, approved researchers are provided with a pseudonymised dataset which can only be accessed through a Safe Haven based at the University of Dundee. This ensures that users are not able to identify individuals within the data or use the data for other purposes. In addition to these technical arrangements, data users are also required to sign a legal agreement with the General Medical Council, which sets out acceptable use of the data. Research outputs are also reviewed against a set of statistical disclosure controls.

Further information can be found on the [UKMED FAQs webpage](#) and full details are available in the [UKMED research process document](#).

There are data items included in the collection data model that are not normally included in our results spreadsheets. How can we submit these?

You will be able to upload any number of files to the data collection system, so do not be concerned if all data items are not consolidated in one file. Providing supplemental files is acceptable, but please ensure they are clearly labelled, or explained in a readme file – this will assist us in collating the final dataset. GMC colleagues will review submitted files for completeness and follow up with queries if necessary.

There is important contextual information about our data that we think you should know when processing our return. How can we provide this to you?

Providers are encouraged to include commentary on their submission in a readme file accompanying their submission. Any information that will aid us in interpreting and understanding your submission will be gratefully received.

Can we include formative assessments in our submission?

Formative assessments are outside of the scope of this collection, which is focused on summative assessments that have an impact on student progression. However, we will include formative assessments in the dataset at the request of medical schools, if they think that having these assessments in the dataset could be of use to them in future.

If you would like to include formative assessments in your submission, please let our colleagues at the GMC know by emailing HESA_Collection@gmc-uk.org. They will then confirm the details of this with you. In either case, please submit your files as they are – there is no need to edit submitted spreadsheets to remove formative assessments if they are not included. GMC will only take the data relating to assessments specified in the ASSESSMENT table you have provided.

Why are you interested in adjustments students receive for extenuating circumstances such as disabilities?

There is ongoing research interest in specific learning disabilities, and knowing which students received extra time or other adjustments is an important data point for that research. It is difficult to interpret the relationship between Specific Learning Disability (SLD) and student performance without knowing if the student had extra time. It is with this in mind that we are requesting information on reasonable adjustments received by individual students – this will complement data we receive on SLDs from the main HESA student record which includes the field [DISABLE](#)

Will the GMC and HESA review the collection processes?

At the end of the year's collection cycle we will seek feedback from medical schools on the collection process, and we will take this into account when planning the subsequent year's collection.

How will this collection evolve over the next few years?

In the first year of the collection, we agreed to undertake a 'soft launch', and as such we only collected data on assessments in the first and final years of your PMQ programmes. From the 2021/22 collection onwards, we will be collecting data on assessments in all years of your PMQ programmes.

We also envision changes in how data will be submitted in future years, that are in line with HESA's [Data Futures project](#). It is intended that the tables included in the [GMC Assessment data model](#) will be incorporated into the Data Futures model – further details are provided in the 'Resolved specification items' section of [this webpage](#). Prior to this, there will also be an intermediate step where data is collected using a standardised template. The proposed timescales for this are as follows:

- ▣ 2020/21 (collected Sep-Nov 2021): spreadsheet return, first and final years only
- ▣ 2021/22 (collected Sep-Nov 2022): spreadsheet return, all years of study
- ▣ 2022/23 (collected Sep-Nov 2023): spreadsheet return using standardised template
- ▣ 2023/24 (collected Sep-Nov 2024): spreadsheet return using standardised template
- ▣ 2024/25 – first potential collection of GMC Assessment data within Data Futures (to be confirmed)

HESA plan to consult with providers after the first and second years of collecting these data to gain feedback from providers and make any necessary improvements.

We have students who have had some of or all their attempts at an assessment nullified or deferred, for example as the result of an appeals process. As such, these attempts do not count against their academic record. How will these be handled?

Please ensure these attempts are clearly labelled or highlighted in your results spreadsheets. Based on the examples we have seen of exam board spreadsheets, these instances are noted in a pass/fail column. Where this isn't the case, please record any nullified attempts in a notes column.

We will record this information in the STUDENTRESULT column of the final dataset. In addition to the 'Pass' and 'Fail' outcomes, we will also use the labels 'Nullified' and 'Deferred' to highlight the cases described. In due course we will amend the entry [here](#) for STUDENTRESULT to note these additional values.

Will the GMC provide any reporting on this dataset?

In addition to this dataset being available to researchers via application for a UKMED project, the GMC also intends to provide some standard reporting. Details of this are still being confirmed. The first Tableau product from this data will be a visualisation reporting on assessment outcomes against protected characteristics and demographics. This will use z-scores of score relative to passmark and a grouping of assessments into broad categories (e.g. 'First year written exam', 'Final year clinical exam' – exact categories to be confirmed). We anticipate providing a first draft under password for comment in March 2022, with a second draft incorporating feedback in May 2022. The March 2022 release will only include data for final year exams as we will not be able to link data for year 1 assessment results until we receive the HESA 2020/21 records in April. The May 2022 release will include the results from year 1 and incorporate feedback received on the first draft.

Are assessments from intercalated years included in this collection?

No. We have agreed with schools that these will not be included, due to the variety and complexity of assessments possible across intercalated degrees, and the difficulty in obtaining this data from other schools or departments.

Should we include students at overseas affiliate schools in our return?

No. HESA does not yet collect person-level data on overseas students, so there would be no student or demographic data to link assessment data to for overseas students. This limits the usefulness of overseas data currently.

Please note that HESA is currently consulting on collecting person-level data on overseas students. If this does happen, in future years the HESA GMC Assessment record may change to include assessment data for overseas students. For now, however, any overseas data we do receive will be discounted and not included in the dataset.

How does this collection relate to GMC's Medical Licensing Assessment (MLA) project?

Starting in 2024-25, medical students will be required to pass the MLA prior to graduating from medical school (more information on the MLA [here](#)). The MLA will consist of two assessments, which will be included in this dataset – however, due to differences in how these assessments are delivered, the data will be collected differently:

- **Clinical and professional skills assessment (CPSA):** this component will consist of one or more of your existing clinical skills assessments, agreed with and quality assured by the GMC.

Data for CPSA assessments will be collected as part of the HESA GMC Assessment record. There is a data item in the ASSESSMENT table called MLACPSA – this allows schools to flag which assessments contribute to the CPSA. If you already know which assessments will comprise the CPSA at your school, please include this in this year's ASSESSMENT table.

- **Applied knowledge test (AKT):** this component will consist of an online exam delivered by the Medical Schools Council in partnership with a third-party provider.

The GMC will receive AKT data for all schools directly from the MSC, so this data will not need to be included in the HESA GMC Assessment return. Any AKT data we do receive from schools will be excluded to prevent duplication.