

Benchmarking methodology

1. The [benchmarking methodology](#) we are proposing to adopt is similar to that which was used by HESA in the UK HE performance indicators (PIs). Put simply, the benchmarking methodology used by HESA is a form of direct standardisation designed to remove the effects of differences in learner populations when comparing providers.
2. The population for each part of the tertiary education sector is divided into a large number of mutually exclusive categories defined by combinations of the benchmarking factors, such as subject of study and entry qualification group. Depending on the number of benchmark factors and the measure to be benchmarked, this can result in several hundred or several thousand categories.
3. Benchmark factors should be selected where there is evidence that they:
 - Have a clear and measurable influence on the outcome being assessed;
 - Are largely outside the provider's direct control;
 - Are available, reliable, and consistently measured across all providers; and
 - Improve the fairness and comparability of performance measures without obscuring genuine differences in provider performance.
4. For each measure being benchmarked, the number and mix of benchmark factors should be carefully considered. Factors should be selected based on their relevance to the outcome being measured and their ability to explain variation between providers. At the same time, the benchmarking methodology should avoid creating an excessive number of categories, as this can result in small or empty groups and reduce the statistical robustness of the benchmark.
5. Including too many factors, or factors with a large number of categories, can diminish the usefulness of the benchmark by overcomplicating the methodology and producing unstable or unreliable results. Consequently, different measures may require different benchmark factors, reflecting the underlying influences and characteristics of each outcome being measured.
6. When choosing the groupings within each factor, which in turn form the categories used in the benchmarking methodology, the groupings should represent distinct and meaningful populations based on the available data. The categories should be as homogeneous as possible, so that learners within a category have similar characteristics and are expected to have similar outcomes. This helps ensure that differences between observed and benchmarked outcomes reflect genuine differences in performance, rather than variation within the benchmark categories themselves.
7. The tertiary education sector is made up of different provider groups (sixth forms, further education, apprenticeships, adult community learning and higher education). Within each provider group, an outcome measure is calculated for each benchmark category. For example, the proportion of learners continuing to the second year of their course might be calculated.

8. For each provider, the number of learners in each benchmark category is identified. These learner counts are then combined with the corresponding outcome measures for the provider group to which the provider belongs. The benchmark is calculated by applying the provider group's outcome measures to the provider's distribution of learners across the benchmark categories:

$$\text{Benchmark} = \sum (\text{number of provider learners in each benchmark category} \times \text{provider group outcome measure for that benchmark category}).$$
9. This produces the outcome that would be expected for the provider given its mix of learners if those learners experienced the average outcomes of providers in the same provider group. An example of this calculation can be found at the end of this Annex.
10. The provider's overall measure can then be compared with the benchmark, and the resulting difference (provider measure minus benchmark) measures the extent to which performance differs from that expected given the provider's learner profile. A positive difference indicates outcomes above expectation, while a negative difference indicates outcomes below expectation, subject to statistical significance testing of the difference.
11. The difference between a provider's measure and benchmark is assessed using an estimated standard deviation. A provider is flagged only when the difference is:
 - greater than three percentage points; and
 - greater than three standard deviations.
12. This identifies differences that are substantial both statistically and practically. Providers performing significantly above benchmark receive a "+" marker, while those significantly below benchmark receive a "-" marker.
13. Benchmarks should not be interpreted as targets. They represent average expected values and may vary from year to year as sector-wide performance changes. Their purpose is to indicate the level of performance that might be expected for a provider after accounting for the factors included in the benchmarking methodology. Where a provider's result differs from its benchmark, this may reflect the provider's performance, but it could also be influenced by factors that are not captured within the benchmark.
14. Results for providers with relatively small learner populations should be treated with caution. In such cases, measures can be subject to considerable year-on-year variation, making them less stable and potentially less reliable indicators of underlying performance.
15. Benchmarking will take into account the differing contexts within which different parts of the tertiary education sector in Wales operates. Therefore, we propose that measures for each HE provider will be benchmarked against data for UK HE providers and, as noted above, we are exploring the possibility of the benchmarks being calculated by HESA. We propose that measures for each provider within the four other sectors, school sixth forms, further education, apprenticeships and ACL,

will be benchmarked against data for all Wales providers within the same sector as that provider, and will be calculated by Medr.

16. Benchmark factors currently being explored for HE include qualifications on entry, subject studied, mode of study, level of study, age, ethnicity, sex, and deprivation. To measure deprivation for HE measures, we propose using HESA's [Socioeconomic Index for Small Areas \(SEISA\)](#), which can be used to classify learners according to whether they are domiciled in areas of relative deprivation within the UK.
17. For the UK Performance Indicators (PIs), HESA used qualifications on entry and subject studied as core benchmark factors, with additional factors included where there was evidence that they were relevant to the specific indicator being benchmarked. We propose adopting a similar approach.

Example of calculating a benchmark

The measure to be benchmarked is the proportion of entrants continuing to the second year of their course. We wish to calculate the benchmark for this measure for provider X, which has 3,500 entrants. The benchmark factors are subject of study and qualifications on entry. Subjects are divided into two groups and entry qualifications into two groups.

	Subject group 1	Subject group 2
Entry qualifications group 1	A	B
Entry qualifications group 2	C	D

This results in the following four categories

- A Entrants with entry qualifications 1, studying subject 1
- B Entrants with entry qualifications 1, studying subject 2
- C Entrants with entry qualifications 2, studying subject 1
- D Entrants with entry qualifications 2, studying subject 2

There are 42,000 entrants in the sector, and the overall number, and the number continuing to the second year for each of the above categories is provided in the table below (column a and column b). The benchmark for provider X is then calculated by following the steps in the bullets and outlined in the table below:

- In column c, for each category, we calculate the proportion of entrants in the sector who continue to the second year of their course, (column b / column a).
- In column e, for each category, the number of entrants at provider X (column d) is multiplied by the sector proportion (column c). Column e therefore represents the expected number continuing to the second year if provider X had the same proportion as the sector continuing in each of the A, B, C and D categories.
- Finally, the benchmark for provider X is calculated by dividing the sum of column e by the sum of column d. This represents the proportion that would be

expected to continue to the second year if the provider had the same proportion as the sector continuing in each of the A, B, C and D categories.

	Sector entrants	Sector entrants continuing to second year of their course	% sector entrants continuing to second year of their course	Provider X entrants	column d * column c
	a	b	c	d	e
A	9,159	7,693	84%	805	676.2
B	14,664	13,344	91%	1,190	1,082.9
C	12,813	11,916	93%	1,015	944.0
D	5,364	4,344	81%	490	396.9
Total	42,000	37,297	n/a	3,500	3,100.0
Provider X benchmark (total column e / total column d)					88.6%