

Office for
Students



Consultation on Data Futures and data collection

Analysis of responses to part one of the
consultation and decisions

Reference OfS 2022.21

Enquiries to dfa@officeforstudents.org.uk

Publication date 12 May 2022

Contents

Summary	3
What we were consulting on	3
The consultation	4
Analysis of responses	4
Final decisions	13
Matters to which we have had regard	15
Annex A: Responses to Part one: Approach to in-year student data collection	17
Overview	17
Comments on approach 1: Two individualised student data collections a year with reduced data requirements in the first data return	17
Comments on approach 2: Cumulative in-year individualised student data collection with potential for differential reporting	28
Additional suggestions	32
Comments on approach 3: Changing the timing of collection	34
Comments on approach 4: Discrete in-year individualised student data collection three times a year	38

The Office for Students is the independent regulator for higher education in England. We aim to ensure that every student, whatever their background, has a fulfilling experience of higher education that enriches their lives and careers.

Our four regulatory objectives

All students, from all backgrounds, and with the ability and desire to undertake higher education:

- are supported to access, succeed in, and progress from, higher education
- receive a high quality academic experience, and their interests are protected while they study or in the event of provider, campus or course closure
- are able to progress into employment or further study, and their qualifications hold their value over time
- receive value for money.

Summary

What we were consulting on

1. Our consultation sought views about how the Office for Students (OfS) could ensure it has the data required to perform its functions and how that data could be collected in a way that is proportionate.
2. The focus of the consultation was on how the Higher Education Statistics Agency (HESA), as the designated data body in England, takes forward the Data Futures programme of work designed to provide in-year student data.¹ In particular we consulted on the nature, frequency and timing of data collections by HESA to enable us to better achieve our regulatory objectives. We also consulted on a number of other areas we had identified where the amount of data collected by HESA or the OfS could be reduced with limited impact on the performance of the OfS's functions.
3. The consultation had three parts.

Part one: Approach to in-year student data collection. In this section we set out the background to the Data Futures programme. We described the current higher education data landscape and methods of student data collection. We proposed three alternative approaches to collecting individualised student data, all of which involve returning some individualised student data during the academic year to reduce the delay between the majority of students starting their courses and having data on them. We also included the current Data Futures approach as a fourth, discounted proposal.

- a. Approach 1: Two individualised student data collections a year with reduced data requirements in the first data return.
- b. Approach 2: Cumulative in-year individualised student data collection with potential for differential reporting.
- c. Approach 3: Changing the timing of collection.
- d. Approach 4: Discrete in-year individualised student data collection three times a year.

4. **Part two: Other changes to data collection.** In this section we set out a number of proposals where we believe improvements to efficiency could be made, including the removal of a small number of student and staff data items that we no longer have a compelling reason to collect, possible clarification of data quality expectations, and changes to provider profile data.
5. **Part three: Use of linked and third-party data.** In this section we set out our proposals to continue to make use of data linking between individualised student data and data from third parties, and to continue to explore the potential of applying data science techniques to data collected through social media and other online sources.

¹ See <https://www.hesa.ac.uk/innovation/data-futures>.

The consultation

6. The consultation was published on the OfS website on 16 December 2021 and the deadline for responses was noon on 3 February 2022.
7. Respondents were invited to share their views on the consultation by submitting written responses to an online survey containing 14 questions. Questions one to five related to part one of the consultation, and questions 13 and 14 were relevant to all parts of the consultation. These questions were:

Question 1: Do you have any comments about our preferred Approach 1: Two individualised student data collections a year with reduced data requirements in the first data return?

Question 2: Do you have any comments about Approach 2: Cumulative in-year individualised student data collection with differential reporting by provider risk?

Question 3: Do you have any comments about Approach 3: Changing the timing of collection?

Question 4: Do you have any comments about our discounted Approach 4: Discrete in-year individualised student data collection three times a year?

Question 5: Please rank the approaches in order of preference. What are the reasons for your preference, particularly the approach you have ranked in first place?

Question 13: Are there aspects of the proposals you found unclear? If so, please specify which, and tell us why.

Question 14: In your view, are there ways in which the objectives of this consultation could be delivered more efficiently or effectively than proposed here?

8. We received 150 responses to the consultation: 144 via the online survey tool; six responses were submitted by email after the deadline. We considered all responses that we received.
9. The responses came mainly from providers (123, including 11 outside England), sector representative bodies (six), and data users (six). We received three submissions from software providers. There were 13 anonymous responses.

Analysis of responses

10. We undertook a qualitative analysis of the feedback that we received. In this document we identify and discuss the most significant issues raised by respondents in relation to **part one of the consultation** as we are conscious of the desire from providers to have clarity on the overall approach to Data Futures as early as possible.
11. The responses we received to part two and part three of the consultation were more mixed. We plan to make decisions and publish the analysis of these responses before the end of July 2022.

Overarching themes from the analysis of part one

12. Annex A summarises the responses we received in relation to each of the four approaches to in-year student data collection set out in the consultation and outlines our response to these issues. In this section, we discuss some of the overarching themes that emerged through the analysis of comments from part one of the consultation.
13. Respondents generally tended to agree with the advantages and disadvantages for each approach as set out in the consultation. The overwhelming majority (87 per cent) of respondents who expressed a preference for an approach were in favour of our preferred approach (approach 1: two individualised student data collections a year with reduced data requirements in the first data return) with approach 2 being favoured by six per cent of respondents and the other approaches roughly evenly spread.
14. Of those respondents who identified issues with the preferred approach, the reasons given were:
 - issues with changing approach now when effort had been invested in the current Data Futures approach²
 - a misunderstanding about how the reduced data requirements in the first data return would work and that it would add complexity
 - that for some providers only having two returns would lead to delays in reporting significant cohorts
 - concerns about the timing of the first data collection.
15. The points made in relation to the preferred approach often translated into advantages or preferences for the other approaches. Many of the respondents who expressed a concern about any change from the current Data Futures approach did so on the basis that significant investment had already been made in it. These respondents often did not recognise the similarity between the preferred approach and the current approach such that we would expect the majority of the investment already made in most providers to carry forward into the preferred approach.
16. While supportive of the proposed approach, many respondents noted that burden would only be reduced if it was possible to use the data collected to define the National Student Survey (NSS) target list and to replace the Higher Education Students Early Statistics Survey (HESES) return, both of which were currently activities that were seen as burdensome and were often undertaken by the same teams within providers. Similar points were made in respect of the medical and dental survey, although these tended to be more limited reflecting the smaller number of providers required to make this return. Some respondents noted that further efficiencies could be gained if the in-year return could replace, or more closely align with, the HESA Initial Teacher Training (ITT) return.
17. Many respondents flagged that the core team responsible for student data returns was also responsible for a number of other activities that occur in the autumn. Respondents noted that

² The current approach is the Data Futures approach that HESA had been working towards prior to this review and corresponded to approach 4 in the consultation, which we discounted.

the teams undertaking the work in providers were typically small and that resourcing an additional activity during the autumn would be a challenge for current teams.

18. Concerns about the peak of work in the autumn were often linked to the proposed census date of 15 November and interactions with the Christmas period. In addition to points about a peak in work, respondents also noted that a significant number of students would still be in the process of registering at this point, for example where student finance confirmation was delayed. There were also concerns that many postgraduate exam boards are held around this time and that this could lead to inaccurate or missing awards data for these students. A number of respondents also suggested that 15 November was not a date commonly used for reporting in providers who often use month ends.
19. There were a variety of other dates proposed by respondents; some of these suggested a modest change to 30 November or 1 December, while others suggested a much later date in order to allow time to capture January student intakes. Respondents often noted that the quality assurance and sign-off process for data would likely encompass the Christmas holiday period, during which many providers were closed, which would require an extended sign-off period.
20. A number of respondents noted that, currently, UCAS data to support the HESA return is not available in time to support data returned in November. To make an earlier return of data practical, UCAS would need to bring forward delivery of its data.
21. Irrespective of the approach adopted, respondents raised issues in relation to the transition year, when providers would be required to make the first return using the Data Futures data model closely followed by the first in-year return. These issues were considered most significant with approach 3, where the transition year would require a different collection covering 15 months or a separate three month return.³ For the other approaches, the autumn of 2023 would create a particular challenge as providers would be making their first returns using the Data Futures data model and then making their first in-year returns in a very short time period. Respondents took the view that this could create a peak in work with no opportunities to learn lessons from the first Data Futures return. Many respondents were in favour of delaying the first collection of in-year data until 2024, although some of these recognised that this delayed delivering the benefits.
22. A common point about the preferred approach was the difference in data requirements between the in-year and the end-of-year return. These points broadly fell into two distinct categories. The first set of points related to a misunderstanding about the proposals whereby some respondents had assumed that this would require a different schema for the two returns: this was not our proposal. The second set of points was more fundamental and surrounded the complexity of data requirements being linked to events, such as the end of the student course session, despite this being a feature of the current Data Futures approach. There were also requests for clarification about precisely which items would be included at which point, as this affects the burden of the returns. This reflected feedback elsewhere that more detail on all of the proposals would be welcome.

³ Approach 3: Changing the timing of collection.

23. A number of respondents expressed support for the cumulative collection model embedded in approach 2 and suggested that this could be combined with the two collections per year model of approach 1.⁴ Respondents who favoured the cumulative collection model, either as described in approach 2 or as an alternative way of delivering approach 1, generally noted that this model would allow for more seamless correction of data errors which may be particularly beneficial during the transition year.
24. A number of respondents, however, took the view that a cumulative collection model would lead to the same data being returned more than once which would lead to increased burden. Some respondents also noted that:
- a cumulative collection model might also allow for a simpler implementation of reduced in-year data requirements as they could be applied to all students
 - this is a more significant change from the current approach than approach 1 and would therefore create additional burden; some of these respondents noted that if this had been adopted from the start it may have been more attractive
 - the value that could be extracted from data with partial coverage of providers was lower, meaning that for those providers required to make all returns the value to burden ratio would be lower
 - this is similar to the approach that is used for the Individualised Learner Record (ILR).
25. Support for approach 3 was very limited. A number of respondents did note that it represented an interesting solution but would need a significant amount of work to understand all the implications of such a radical shift. Most respondents felt that the benefits of reducing the number of returns were more than offset by the complexity of the single return, often citing the complexity they face with students whose activity does not fit their standard patterns, and a concern that this would then apply to the vast majority of their students. Points were also made about how future-proofed this approach would be, as it cannot be extended if more frequent data is required in future. A number of respondents agreed with our assessment that it would create issues with comparability with staff and Higher Education Business and Community Interaction (HE-BCI) survey data. Respondents noted that there would be similar comparability issues for finance data for those providers that have financial years aligned to the 1 August to 31 July academic year.
26. Respondents who favoured approach 4 over approach 1 often raised issues about how sustainable approach 1 would be in the long term, especially in the context of Lifelong Loan Entitlement (LLE).⁵ It was also noted that this approach was most neutral in terms of patterns of student recruitment throughout the year and therefore levelled the playing field for providers that operate different academic cycles. Respondents often recognised that, in the short term, approach 1 would be less burdensome as it required fewer returns, but they were concerned that, with an increase in flexible provision, it was possible that a new approach to data collection would become necessary in the near future and that the cost of change would outweigh any short-term savings from only having two returns. Some respondents with large January intakes or recruitment throughout the year considered that the preferred approach

⁴ Approach 2: Cumulative in-year individualised student data collection with potential for differential reporting.

⁵ Approach 4: Discrete in-year individualised student data collection three times a year (current Data Futures approach).

could lead to misleading data about their courses because a reasonable proportion of the new intake would be missed.

27. Questions were raised about how data errors would be handled under all proposals. Respondents noted that data quality naturally tends to improve over time and that, depending on the approach that is taken, querying data changes could generate significant burden for providers. Some respondents were of the view that it may not be possible to distinguish between genuine changes in data and the correction of errors. There was some concern that a proliferation of dates or reason for change fields, as had been seen in previous versions of Data Futures, may be introduced to the specification at a later stage to aid in the understanding of changes in data. Respondents also considered that it would be important for the OfS, and other data users, to show an understanding of the achieved data quality in using the in-year returns.
28. Respondents generally welcomed the focus on reducing burden and the reduction in burden that would come from the preferred approach. However, a number of respondents suggested that the focus on the OfS's use of the data diminished the original Higher Education Data and Information Improvement Programme (HEDIIP) vision of the HESA return serving the needs of a wide range of stakeholders. This could ultimately lead to an increase in the burden as respondents were concerned that very few non-OfS uses would be covered by the in-year return.

Our response

29. In response to points about changing the approach to the in-year collection of data now, when significant effort has already been invested in delivering the current Data Futures approach, we recognise that any change will incur some costs. However, we note that our preferred approach is very similar to the current option and so most preparation already undertaken by providers is likely to be transferable to the preferred option.
30. One of the drivers for selecting approach 1 as the preferred option in the consultation was its similarity to the current Data Futures approach which will minimise the need for re-work. We have also taken the view that the approach adopted will be in place for many years, so the cost of change is likely to be outweighed by the benefits of a different approach.
31. On the suggestion that burden would only be reduced if it was possible to use the data collected to replace the Higher Education Students Early Statistics Survey (HESES) return, we have committed to including proposals to remove the need for HESES in any future funding review. We will also cease collection of the medical and dental survey as this can be derived from the in-year returns.
32. We recognise the peak in work for providers during the autumn and its impact on the core team responsible for student data returns and will seek to mitigate the impact wherever possible. We note that providers are able to make decisions about how and where to allocate resources to ensure that they are able to satisfy all regulatory requirements.
33. In response to points about the peak of work in the autumn linked to the proposed census date of 15 November and interactions with the Christmas period, we recognise the challenge of finding an optimal census date for the in-year return. There is unlikely to be any approach or timing that fully aligns with the systems and approaches across the diverse range of providers

that we regulate. However, the primary challenge relates to the use of the first return in the year to generate the NSS target list. The ability to remove the NSS additions and removals process was seen as a key benefit of the proposals by many respondents.

34. Working with colleagues in the devolved administrations we have carefully considered the implications for the timing and response rates for the NSS of moving the census date from 15 November to 1 December with the collection closing in the third week of January. This would mean the NSS could not start before the third week in February. We know that high response rates and availability of data before the summer break are important to providers in using the NSS for their own purposes. If we are to ensure providers have NSS data before the summer break the ability to extend the fieldwork beyond the end of April is limited and reducing the fieldwork to 11 or 12 weeks is likely to reduce response rates by around 3-5 percentage points, although changes to the survey process could mitigate this.
35. We have not yet made a final decision about the end of the first reference period. Given the points made by providers, we are minded to end the first reference period on 1 December with data signed off by the third week of January. However, given the significant impact that this would have on the timing of the NSS, we intend to consult on this as part of our consultation on the future of the NSS in the summer. We expect to consult on two options:
 - a. The first reference period would end on 15 November with data signed off by providers no later than 23 December, which would allow the NSS survey window to run from mid-January to the end of April as now.
 - b. The first reference period would end on 1 December with data signed off by providers no later than the third week in January, which would allow the NSS survey window to run from late February to the end of April. This may lead to a small reduction in the response rate for the survey.
36. On the issue of delivery of UCAS data to support the HESA return, UCAS has committed to exploring bringing forward delivery of UCAS STAR(*)J data to support providers in making in-year data returns.
37. We have listened to the points made in relation to the peak in work in autumn 2023 as providers make their first returns using the Data Futures data model and then make their first in-year returns in a very short time period. We have therefore decided to delay the first collection of in-year data until 2024-25 to allow providers to learn lessons from the first return using the Data Futures model. We also expect that this delay will reduce the peak of work during the autumn of the transition year as providers will benefit from the experience of making the year end return in the previous year. We do not anticipate making any further changes to the date from which in-year data will be required.
38. In response to points about differences in data requirements between the in-year and the end-of-year return, we can confirm that the data model will be the same for both returns, although the data requirements will differ between returns. If providers wish, they will be able to return all elements in each collection, including those that are only required at the end of the student course session. We will confirm details of the data items that will not be required until the end of the student course session by the end of July 2022, as part of our decisions about parts two and three of the consultation.

39. We considered the suggestion to combine the cumulative collection model set out in approach 2, with a requirement for all providers to make two returns, to be a useful hybrid of approaches 1 and 2. We consider the primary driver for a cumulative approach to be the benefit that returns can be omitted; without this benefit, cumulative collection has the potential to increase burden through requiring data to be returned more than once. We do not consider the benefits of cumulative collection where all returns are mandatory outweigh the issues of returning data multiple times. We also note that this would require more significant systems change for providers as it is a more significant departure from the current approach.
40. There were no comments about approach 3 that led us to reconsider whether this option should be taken forward. The alternative options that were suggested would not have significantly addressed the points made by respondents about the complexity of the approach as the reporting year under this approach would still cut across academic structures in different providers.
41. We considered the points made by respondents who favoured approach 4 about how sustainable approach 1 would be in the long term, especially in the context of the LLE. The original Data Futures approach was designed to ensure that data on non-traditional courses was no more lagged than for traditional courses. However, while non-traditional courses are growing, and are likely to grow further through the LLE, the majority of students still start courses in the autumn. It is not clear that the proposed LLE and increased flexible and modular provision will require a move to three data collections per year. Given that there is currently no clear requirement for data more than twice a year, we do not think it is proportionate to impose a burden on providers now when two data returns are sufficient.
42. We recognise the possibility that, as more providers offer more flexible provision, we may need to increase the number of returns. If that were to happen the reduction in burden in the short term may be outweighed by the costs of two changes. While we recognise the cost of change, we view moving to two collections per year to be proportionate given the current landscape.
43. In response to points made about how data errors would be handled under all proposals, we consider that providers should ensure data is correct at the point of collection, although we recognise that providers will undertake some data cleansing at a later point. We do not intend to explicitly differentiate between data correction and updates. However, we would expect to question providers where there is a material level of change which may indicate that an earlier return was not fit for purpose. Where errors in returns are widespread and material, we may require that returns are corrected. We do not expect to require corrections to earlier returns where errors are not widespread or material.

Comments about the consultation process

44. Some respondents made points about the timing of the consultation in relation to the planned implementation of Data Futures. Some respondents considered that the consultation should have occurred much earlier in the process as they took the view that some work may have been wasted preparing for the now discounted option. However, a number of respondents raising this issue appeared to view the differences between approach 1 and 4 as more fundamental than they are in practice. One respondent took the view that considering the amount of work already undertaken, it was not helpful to signal that approach 4 had already

been discounted before consultation had taken place. However, many respondents supported our decision to discount this approach.

45. Respondents were keen that the outcomes of the consultation should be announced as soon as possible, and details of the approach finalised in order to allow them time to plan.
46. As we set out in the consultation document, we decided to review the Data Futures proposals as part of our commitment to ensuring that the requirements we place on providers are proportionate. This followed suggestions from some providers that the increased frequency of data collections proposed in the Data Futures programme could result in a substantial increase in administrative burden. We deliberately delayed consultation on these issues as our approach to the frequency of data collection is to a large extent dependent on the nature of the data we require to perform our functions, including:
 - our approach to the regulation of student outcomes, and other aspects of quality
 - the Teaching Excellence and Student Outcomes Framework (TEF)
 - the National Student Survey (NSS).
47. However, a few respondents also made points about the limited overlap between this consultation and the three OfS consultations on the new approach to regulating student outcomes, the TEF, and student outcome and experience data indicators, which they stated made it difficult to understand the wider context. Another respondent considered that it would have been more helpful if the data model had been confirmed before these other consultations were launched.
48. In deciding on the timing of the consultation, we sought to balance the need to decide on an approach to data collection for the 2023-24 academic year to allow providers and HESA time to prepare, and the need for visibility about the nature of the data required for producing student outcome and experience data indicators for the regulation of student outcomes and the TEF.
49. A small number of those who responded made points about the amount of time available to provide their response, noting that the consultation had coincided with the Christmas period. In recognition of the impact this timing could have on respondents, we extended the consultation period to seven weeks and also considered any late responses. We consider that this period was sufficient to enable those wishing to respond to do so. In addition, the roundtable discussions we held with stakeholders in July 2021 helped us to understand more about the relative burden of different types of student data collections and different activities required to complete a student data return, prior to publishing our consultation. We used the roundtable discussions with stakeholders to assist in the development of the proposals consulted on in December 2021.
50. One respondent suggested that the consultation did not provide sufficient information about why HESA data did not meet the needs of third parties, including professional, statutory, and regulatory bodies (PSRBs). Another response suggested that the consultation focused only on removing the data burden within OfS collections and in-year data that prioritised the OfS's regulatory needs, rather than those of other bodies which require in-year student data. As set out in paragraph 32 of the consultation, HESA's experience in designing the Data Futures

model, and feedback from providers, is that adapting the Data Futures programme to meet the needs of a wide range of stakeholders is likely to add considerably to the burden of data collection. We therefore considered that in-year collection of individualised student data through Data Futures needs to be justified predominantly on the basis of the needs of the current users of HESA data, rather than any potential future uses by PSRBs. Once the Data Futures model is established and delivering in-year individualised student data, we would expect it to be possible to extend the model to meet the needs of a wider range of stakeholders. Any such extensions would be optional for providers.

51. One respondent suggested that we had not included data protection legislation in the list of matters to which we had had regard in developing our proposals. We agree that data protection legislation is relevant to considering the appropriateness of the proposals, due to the extent of collection, processing and sharing of personal data of students. This is something we considered but did not list in Annex F of the consultation. Data protection considerations have been central to the development of the Data Futures programme to this point. Many of the data protection issues are the same now as under the original Data Futures programme; increasing the frequency of data collection does not fundamentally increase the scope of the data required. The proposals we consulted on in part two of the consultation were specifically addressing the data minimisation principle in data protection legislation.
52. Two respondents suggested that it would have been helpful if the consultation had been part of a joint consultation with the devolved nations to help speed up the process and to ensure unintended consequences of any decisions taken by the OfS as part of this consultation are understood, particularly in relation to the changes to the number of collections, quality assurance and the onward use of data. We could not fetter the OfS's or the devolved administrations' discretion in decision-making about data requirements and so while we have been engaging with colleagues in the devolved administrations as we have developed our proposals, we considered it appropriate to consult on the approach as it relates to England only.

Comments outside the scope of the consultation

53. We received a very small number of comments that raised issues we considered to be outside the scope of this consultation. For example, some respondents considered that statutory and other data returns which were not included in this consultation place a disproportionate burden on small providers with limited staff capacity and ability to invest in and manage complex and expensive data systems. The consultation explicitly recognised the burden that individualised student returns place on small providers and we take the view that reducing the number of returns is likely to disproportionately benefit those providers with less sophisticated student record systems. Additionally, it was suggested that a consultation on data quality rules and margins of error should be conducted. There were also points made by a data user in relation to delivery due to current long delays in receiving HESA data.

Final decisions

Approach to in-year student data collection

The Data Futures data model will be implemented as planned in 2022-23, with the year-end collection of student data signed off by providers in autumn 2023 submitted on the basis of the new data model.

In relation to the collection of in-year student data, we have decided to proceed with approach 1 in our consultation – two individualised student data collections a year with reduced data requirements in the first data return. This approach is conceptually, and technically, very similar to the current Data Futures approach.

We have made the following decisions about the implementation of this approach:

Transition year

Taking into account the points made by providers about the peak of work in autumn 2023, and that a transition to in-year data in 2023-24 would not allow providers to learn lessons from the first year of Data Futures, we have decided to delay the implementation of in-year collection until 2024-25. There will therefore be a year-end collection in 2023-24 that mirrors the year-end collection in 2022-23.

In delaying the first in-year collection, we expect HESA and providers to focus on ensuring that the first collection of in-year data in 2024-25 is of high quality and delivered to time so that we can rely on using the data for our regulatory purposes. We do not expect any further delay to the collection of in-year data.

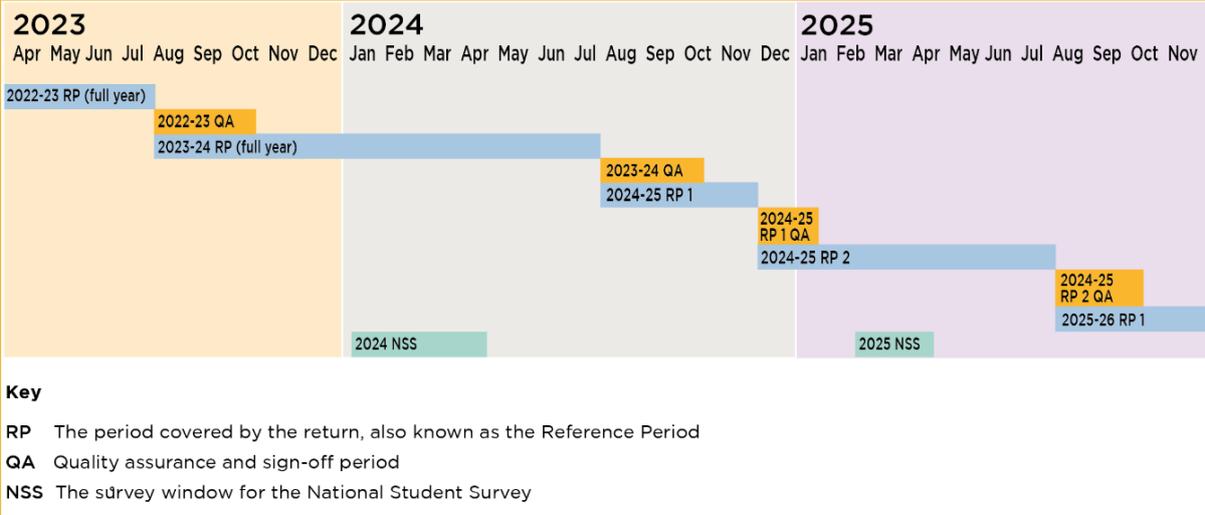
Data requirements and timing

The Data Futures model will be the same for both returns and, in light of the points made by respondents about the complexity of data requirements being linked to events, providers will be able to choose to return all data items in both returns for all students if this reduces burden for them.

We will confirm details of the data items that will be included in the first in-year return and which items will not be required for new students until the second return at the end of the student course session by the end of July 2022 as part of our decisions about parts two and three of the consultation.

We have not yet made a final decision about the end of the first reference period. Given the points made by respondents, we are minded to end the first reference period on 1 December with data signed off by the third week of January (see Figure 1). However, given the significant impact that this would have on the timing of the National Student Survey, we intend to consult on this as part of our consultation on the future of the NSS in summer 2022.

Figure 1: Summary of reference periods for the implementation of approach 1



Removal of HESES and the medical and dental student survey

Respondents considered the removal of HESES for providers in the Approved (fee cap) category to be a significant benefit to in-year individualised data return. Removing HESES would require a change to the OfS’s recurrent funding streams as it is used as the volume measure in all of them. When we have transitioned to in-year individualised student data collection we would expect to use this as the primary volume measure in our funding allocations. Using individualised data would mean that student number forecasts later in the year were not included; we would instead be likely to take account of the volume of students who started study after the census date in the previous academic year, so that funding was still informed by student data covering a full 12-month period.

We anticipate consulting on proposals to reform our funding approach. Where this is the case, we will include proposals that would allow us to remove HESES. We expect to launch this consultation in the next academic year.

Removal of the medical and dental survey would not require any material changes to our monitoring of medical and dental numbers. We will therefore cease collection of this survey in 2024-25.

NSS additions and removals process

We will simplify the generation of the NSS target list so that there is no need for the current NSS additions and removals process. We expect there to be a small number of cases where changes to the target list generated from the first reference period are allowed, for example, if a student has died or is seriously ill. Given the very low number of such changes expected, we will only require providers to flag students to be excluded when providing contact details.

Ability to update data

We expect providers to ensure data is correct at the point of collection. However, we accept that, in practice, data quality may improve over time, particularly as new approaches to data collection are embedded. Therefore, the quality assurance mechanisms will need to allow for some data correction.

While accepting that data quality could be expected to improve over time, we will still require that every return is fit for purpose. We will provide more information about how the OfS will judge whether a return is fit for purpose in our response to part 2 of the consultation.

Data needs of other stakeholders

Once the Data Futures model is established, and we have transitioned to in-year individualised student data collection, we expect that HESA, working with providers and PSRBs, will want to consider how the wider needs of stakeholders can be met. Any additional data requirements would need to fit within the timings of returns required by the OfS and would not be mandated by the OfS.

We have engaged with colleagues in the devolved administrations in forming our proposals and sought to align with their requirements where appropriate and possible, recognising the value many users place on UK-wide data. However, we do not judge that it would be proportionate to impose significant additional burdens on English providers solely to deliver UK-wide alignment.

Matters to which we have had regard

54. In reaching our decisions we have had regard to our general duties as set out in section 2 of the Higher Education and Research Act 2017. General duty (g) is particularly relevant to these decisions, which relates to best regulatory practice, including the principles that regulatory activities should be transparent, accountable, proportionate and consistent. In reaching our decisions we have balanced the value that will be derived from having consistent data from all providers against the burden that a consistent approach may place on some providers and concluded that the value outweighs the additional burden. We consider this to be an appropriate and proportionate response.
55. We have given weight to duty (g) in reaching our decision to cease the collection of the medical and dental student survey and the NSS additions and removals process, both of which were seen as burdens by respondents to the consultation.
56. We have also considered the duty in section 64(8)(b) that the OfS must have regard to the desirability of reducing the burdens on such providers relating to the collection of information. We take the view that our decisions are consistent with this.
57. The introduction of an in-year individualised data return will create an environment where other users of data will potentially be able to access more timely data, which we know to have been one of the key barriers to the use of HESA data. We expect the designated data body to continue to work with other data users to enhance the collection of data to meet their needs in future.
58. We are required to have regard to the Regulators' Code in which 1.1 and 1.2 have prompted us to consider the burdens that our activities place on regulated entities. This has been central to our considerations throughout the formulation of these proposals and our decisions following consultation.

59. As an official statistics producer, we have had regard to the Code of Practice for Statistics. The Code of Practice for Statistics aims to ensure that the statistics produced by the government and public sector bodies are trustworthy (impartial and free from political influence), high quality and of public value and that effective governance structures are in place to protect transparency and accountability. The approach we have adopted will provide earlier public data on the number and nature of students in higher education which would enhance the value of the statistics produced.
60. Under section 149 of the Equality Act 2010, the OfS must have due regard to the public sector equality duty. This requires the OfS to have due regard to eliminating unlawful discrimination, foster good relations between different groups and take steps to advance equality of opportunity. Requiring the collection of data about individual students and their characteristics ensures we are able to continue to understand the impact of our regulation on those with protected characteristics and the performance of the higher education sector in relation to those students. We expect that having detailed data available earlier will give us, and others, a more timely view of these matters.
61. We have considered the impact of these decisions on the privacy of data subjects. These decisions do not significantly alter the amount of data we, or the designated data body, will hold about individual students. However, our decision to reduce the number of returns does reduce the number of times that data is transmitted, which supports the data minimisation principle. The decision to cease the NSS additions and removals process reduces the volume of data that the NSS survey contractor will process about students who should not be part of the survey, which has a positive impact on privacy.

Annex A: Responses to Part one: Approach to in-year student data collection

Overview

1. In this annex we have summarised the responses we received in relation to each of the four approaches set out in the consultation. In some cases, respondents focused the substance of their response on one approach, although a number of the comments made in relation to one approach could be equally applicable to any of the approaches.
2. The most common point made in relation to all approaches, raised by approximately 45 per cent of respondents, related to a request for further information about the removal of HESES and the medical and dental survey, and the significant increase in burden that would be experienced by providers if they were required to complete in-year Data Futures returns in addition to HESES.
3. A similar proportion of respondents made points relating to the timing of data returns and sign-off. Respondents raised the potential for peaks in work in late autumn and around Christmas, associated with submitting data for the first reference period.
4. One in five respondents expressed concerns across all approaches about the proposed implementation timescale, in particular that the autumn of 2023 would create a particular challenge as providers would be making their first returns using the Data Futures data model and then making their first in-year returns in a very short time period. Many respondents were in favour of delaying the first collection of in-year data until 2024, although some of these recognised that this delayed delivering the benefits.
5. We have set out further details relating to these points and our response to them in the section on comments on approach 1.

Comments on approach 1: Two individualised student data collections a year with reduced data requirements in the first data return

6. In this section we have summarised the responses we received in relation to question 1 about approach 1.
7. Approach 1 would reduce the frequency of collecting individualised data from the current Data Futures approach of three times a year to twice a year.
8. The way that the return is structured would mean that, for students starting their courses in the autumn, the first in-year data return would require a reduced number of data items. The second return would have more comprehensive data collection for most students although there would still be reduced requirements for students who start near the end of the academic year.
9. Our expectation is that the first data return would capture activity during the autumn term and would be the same for each provider. We would expect the second return to cover activity in the remainder of the academic year, with data submitted over the summer and early autumn.

10. Approach 1 would use the student course session concept within the Data Futures model. This would require data at the end of the student's year rather than the end of the reporting year, thereby ensuring that students only need to be returned in a period where they are active. This would avoid the need for different specifications for the two returns while in practice reducing the data requirements in the first return for the majority of students.
11. This was our preferred approach, as it will allow us to deliver most of the benefits of in-year data, while allowing providers more time to assure the quality of data that is not required at the start of the year.
12. The overwhelming majority of respondents (87 per cent) who expressed a preference for an approach ranked approach 1 as their preferred approach. A further 12 per cent of respondents ranked this approach second. Only one respondent ranked this as their least preferred option.
13. Of those respondents who ranked this approach first, over 40 per cent cited the reduction in burden conferred by this approach as the reason for their preference. The reduction in frequency of data returns and the requirements for the first return were seen as a welcome reduction in burden compared with the original Data Futures model. The removal of HESSES and the medical and dental survey was specifically mentioned by 40 per cent of respondents as the reason they supported this proposal, or an essential requirement to ensure this approach achieved its aim of reducing burden.
14. The most commonly cited issue specifically about approach 1, mentioned by a little under a third of respondents, was about the timing of data returns.
15. Comments received from respondents on approach 1 can be broken down into seven general themes:
 - a. Comments about burden and complexity.
 - b. Comments on data requirements and timings.
 - c. Comments about data accuracy and utility.
 - d. Comments on implementation.
 - e. Comments about alignment with existing systems.
 - f. Comments about alignment with other returns.
 - g. Comments about flexibility.
16. There were also several requests for further information, particularly in relation to a desire to understand the precise detail of each return. Given that the detail of the returns is likely to evolve over time we do not consider this detail was necessary to inform responses about the broad approach. Several respondents made additional suggestions in relation to this approach.

Burden and complexity

17. As mentioned in paragraph 13, over 40 per cent of respondents who ranked this approach first did so because they considered that it would result in a reduction in burden. Respondents

considered that the reduction in frequency of data returns from three to two per year, and the reduced data requirements for the first return, represented a welcome reduction in burden compared with the original Data Futures model.

18. A small number of respondents felt this approach was a good compromise in terms of providing an early view of student data while reducing administrative requirements. A few respondents also suggested that by reducing the number of times data was submitted, approach 1 would reduce the number of repetitive data quality queries across the collection cycle. This would be particularly beneficial to providers with students or courses less likely to follow the dominant academic year model or mode of study.
19. Some respondents suggested that the similarity of the proposed approach to the original HESA Data Futures collection model provided reassurance that it would be a workable model. This was confirmed by one of the major software providers, Tribal, indicating its support for this approach on the basis that it is conceptually and technically very similar to the current Data Futures approach and due to the flexibility built into the current software solution around reporting periods would not be a significant change.
20. The use of in-year student data to replace the HESES return was specifically mentioned by 40 per cent of respondents either as the reason they supported this proposal, or an essential requirement to ensure this approach achieved its aim of reducing burden. Similar points were made in respect of the medical and dental survey, although these tended to be more limited reflecting the smaller number of providers required to make this return. Some respondents noted that further efficiencies could be gained if the in-year return could replace, or more closely align with, the HESA Initial Teacher Training (ITT) return.
21. Respondents were clear that we should commit to the removal of HESES and the medical and dental survey, preferably with an agreed published timetable. Many respondents flagged that the failure to remove HESES and the medical and dental survey would result in a significant increase in burden if providers were required to complete the first Data Futures return at the same time.
22. Respondents noted that the teams undertaking the work in providers were typically small and that resourcing an additional activity during the autumn would be a challenge for current teams. The removal of HESES was seen by one respondent as a way of implementing Data Futures without incurring additional staffing costs.
23. There were, however, some points made about the potential for this approach to be burdensome, complex and costly. These generally broke down into specific themes:
 - a. Additional complexity of having different data requirements in each of the two reference periods, and the additional risk to data quality and burden of additional validation required to ensure the two returns meet OfS expectations.
 - b. Increased burden if module data needed to be returned.
 - c. The potential for peaks in work, particularly in late autumn and around Christmas, associated with submitting data for the first reference period concurrently with completing the final reference submission from the previous HESA year, HESA staff

record, aggregate offshore record, ITT in-year record, Health Education England (HEE) returns, major ILR returns and HESES census in the same period.

- d. Increased burden compared to the current retrospective approach to returning HESA data, though less than the current proposed Data Futures approach. Continued uncertainty around the finalised process, coverage and frequency of the Data Futures approach is adding to this burden.

OfS response

As we have set out in our final decisions, the data model will be the same for both returns. Providers will be able to return all elements in each collection, including those that are only required at the end of the student course session.

We have committed to including proposals to remove the need for HESES in any future funding review. We will cease collection of the medical and dental survey as this can be derived from the in-year returns.

We recognise the peak in work during the autumn and will seek to mitigate the impact wherever possible.

Data requirements and timings

24. The most commonly raised points about approach 1, raised by just under a third of respondents, related to the timing of data returns and sign-off, particularly of the first reference period.
25. Broadly, the points raised by respondents reflected the following themes:
 - a. The availability of enrolment and postgraduate masters' award data at the 15 November census date. Respondents noted that a significant number of students would still be in the process of registering at this point, for example where student finance confirmation was delayed. In addition, respondents highlighted that many postgraduate exam boards are held around this time and that this could lead to inaccurate or missing awards data for these students.
 - b. Current UCAS data for HESA Transaction STAR(*)J is not available in time to support data returned in November. To make the proposed earlier return of data practical, UCAS would need to bring forward delivery of their data.
 - c. That 15 November was not a date commonly used for reporting in providers who often use month ends.
 - d. Discussion of the challenges of managing a collection and submission during the autumn were linked to the proposed census date of 15 November and interactions with the Christmas period. Respondents often noted that the quality assurance and sign-off process would likely encompass the Christmas holiday period, during which many providers were closed, which would require an extended sign-off period.

- e. Comments about the split of student activity between the two data collections for providers with non-traditional intakes spanning the November census date.
 - f. Suggestions that a later submission date would lead to a more stable data set and align better with the Scottish Funding Council Early Statistics return.
26. A variety of alternative dates for the first return were proposed by respondents. Some of these suggested a modest change to 30 November or 1 December, to align with established processes oriented to the HESES census and to allow the great majority of autumn awards to be conferred. Some respondents suggested the first data return sign-off should be delayed until the end of January to allow staff a break over Christmas and New Year. Other respondents suggested a much later date in order to allow time to capture January intakes.
27. Other comments about data requirements included suggestions that providers with large part-time cohorts would only be able to return accurate full-time equivalent (FTE) data at the end of the year, and that it is not meaningful to calculate and report FTE based on arbitrary reference and reporting periods. It was also suggested that it would not be meaningful to include the FUNDCOMP field as currently defined in the first submission for entrants.
28. There were requests for further information about which items would be included at which point as this affects the burden of the returns. One respondent stated that they would be happy to return two full data sets for a student if the data was available, rather than submitting a partial return and then a complete return.
29. One respondent was concerned about the burden that this approach would place on its existing student record system in terms of running student return processes alongside admissions, enrolment and re-enrolment. As student record systems can only run one return at a time, any delay in meeting deadlines for one will likely have a knock-on effect to subsequent returns and risk not meeting statutory requirements to return data.
30. Another respondent raised concerns that a complete dataset would not be available by mid-November as registration processes for new starters are not complete until the end of November. While in principle it might be possible to restructure some of these processes, the autumn term is the busiest period of the year so this would be costly and laborious.

OfS response

We recognise the challenge of finding an optimal census date for the in-year return and recognise the issues with 15 November. The primary challenge relates to the use of the first return in the year to generate the NSS target list. The removal of the NSS additions and removals process was seen as a key benefit by many respondents.

We have carefully considered the implications on the NSS timing and response rates of moving the reference date to 1 December. This would mean that the collection would close in the third week of January, and therefore the NSS could not start before the third week in February. We know that high response rates and availability of data before the summer break are important to providers in using the NSS for enhancement. If we are to ensure providers have data before the summer break, the ability to extend the fieldwork beyond the

end of April is limited and reducing the fieldwork to 11 or 12 weeks could reduce response rates by around 3-5 percentage points, although it may be possible to mitigate this.

We have not yet made a final decision about the end of the first reference period. Given the concerns raised by providers, we are minded to end the first reference period on 1 December with data signed off by the third week of January. However, given the significant impact that this would have on the timing of the National Student Survey, we intend to consult on this as part of our consultation on the future of the NSS in the summer.

UCAS has committed to exploring bringing forward delivery of UCAS STAR(*)J data to support providers in making in-year data returns.

We are carefully considering the detailed data requirements as part of our second stage response but note the issues with FTE and funding completion.

Data accuracy and utility

31. Several respondents commented on issues relating to data accuracy, data quality arrangements and the ability to utilise data. Many of these comments focused on the ability to update data between reference periods in order to correct errors. It was suggested that the lack of opportunity to make corrections to the earlier return is a weakness of this approach. The discrete model of data collection could also mean that students may be missed if they enrol late or if data is received too late, but they withdraw prior to the final collection. Respondents welcomed further clarification about how the data correction process would be managed if data is discovered to be incomplete or inaccurate after it is signed off.
32. Some respondents noted that it may not be possible to distinguish between genuine changes in data and the correction of errors, and that data improvement between periods could be viewed as data correction. A few respondents noted that there could be a validation burden of explaining changes between the two data returns, and the disproportionate impact on resource if providers were required to submit error reports and amendments in response to small and understandable changes.
33. A small number of respondents raised concerns about the ability of providers to achieve the expected data accuracy in the first data return if it is collected at an earlier census date. Respondents noted that it would be important for the OfS, and other data users, to show an understanding of the achieved data quality in using the in-year returns. Respondents took the view that, where regulatory decisions are made on the basis of this data, the data would need to be of high reliability.
34. Some respondents welcomed the ability to complete data quality work on a reduced first return in advance of the larger submission later in the year, and supported the differential approach to data quality set out in the consultation that would allow providers time to complete data quality work for students starting shortly before the end of a reference period. However, one respondent was concerned that the data standards that providers are expected to meet in making the HESSES return could be undermined if the quality expectations for the first HESA data return are lower.

35. One respondent raised concerns about the shortening of the period for responding to data quality queries, particularly as these could include queries raised about changes between the mid-year and the final return.

OfS response

We expect data to be correct at the point of collection, although we recognise that providers will undertake some data cleansing at a later point. We do not intend to explicitly differentiate between data correction and updates. However, we would expect to question providers where there is a material level of change which may indicate that an earlier return was not fit for purpose. Where errors in returns are widespread and material, we may require that returns are corrected. We do not expect to require corrections to earlier returns where errors are not widespread or material.

Implementation

36. Several respondents who preferred approach 1 cited the similarity of this approach to the existing Data Futures model as a reason for their support, and that this provided reassurance that it is a workable model. As many of the processes required to deliver the data are already in place, or currently being worked towards, this approach would be the easiest to implement.
37. One respondent suggested that even with the close alignment of this approach with the existing Data Futures model, providers would likely incur additional costs if they needed to make changes to existing software and internal processes, and train or recruit experts to work with these data sets.
38. Some respondents commented on the proposed implementation timescale, in particular, that the autumn of 2023 would create a particular challenge as providers would be making their first returns using the Data Futures data model and then making their first in-year returns in a very short time period. They further noted that there would be additional burden if a HESES return was also required during this timeframe. This could create a significant peak in work with no opportunities to learn lessons from the first Data Futures return. Many respondents were in favour of delaying the first collection of in-year data until 2024, although some of these recognised that this delayed delivering the benefits.

OfS response

We recognise that any change will incur some costs. One of the drivers for selecting approach 1 as the preferred option in the consultation was its similarity to the current approach which will minimise the need for re-work.

As set out in our final decisions, we have listened to the points made by respondents and have decided to delay the first collection of in-year data until 2024-25 to allow providers to learn lessons from the first return using the Data Futures model. We also expect the delay will reduce the peak of work during the autumn of the transition year as providers will benefit from the experience of making the year end return in the previous year.

Alignment with existing systems

39. A few respondents expressed concerns about the potential for a lack of alignment between this approach and existing student record systems. It was suggested that the first reference period does not align with dates used in other reporting. One provider raised a concern that the timing failed to take into account externally validated students who would need to be kept in student data systems beyond the natural cycle of teaching as there are often delays receiving results from awarding bodies.
40. Several respondents who supported the approach did so on the basis that it would align with existing student record systems and systems developed by software providers. A few providers suggested that the approach would be feasible with minor revisions to existing schedules and would allow records of students on non-standard length courses to be closed down in a timely way.

OfS response

There is unlikely to be any approach or timing that fully aligns with the systems and approaches across the diverse range of providers that we regulate. We consider that the mapping of internal data to external definitions is best done by providers who will understand their own delivery models.

Alignment with other returns

41. Many respondents supported this approach because of the perceived additional benefits, such as the potential removal of the HESES return, medical and dental survey and the additions and removals process for the NSS target list.
42. Both Health Education England (HEE) and the Office for National Statistics (ONS) considered that this approach could provide a timely snapshot of student numbers earlier in the year as well as end of year data. This would allow HEE to use data to inform timely decisions around education and workforce planning. The extent of the usefulness of the approach to the ONS would be dependent on the variables included in the first return.
43. However, some respondents raised concerns about the impact this approach could have on data returns for other stakeholders, particularly in cases of coinciding collections, relating to:
 - a. The continued data burden of providing data and commentary to PSRBs according to their specific definitions and schedule, and to international PSRBs in relation to accreditation and monitoring.
 - b. The lack of alignment with the Scottish Funding Council early statistics return which uses a census date of 1 December.
 - c. The importance of the timing of the first reference period in terms of the impact that first year students changing subjects may have on cost centres and price group information.
 - d. Whether, in the light of the removal of the HESES return, the in-year data return would provide sufficient data for the determination of the strategic priority grant.

OfS response

We welcome the potential use of Data Futures by HEE and the ONS to provide a timely snapshot of student numbers earlier in the year. We recognise the benefits that meeting the needs of these organisations can bring.

Where possible, we would hope that in the future PSRBs and others will be able to benefit from the use of the data collected by the designated data body. However, the needs of these bodies are diverse and it is unlikely that we would be able to determine data content and timing that would meet the needs of all of these bodies. Therefore, we think it is appropriate for us to ensure that our needs are met.

We have sought to align with requirements in the devolved administrations where possible, recognising the value many users place on UK-wide data. However, we do not judge that it would be proportionate to impose significant additional burdens on English providers solely to deliver UK-wide alignment.

Flexibility

44. There were mixed views on the flexibility that this approach would provide. Some respondents felt that it would be beneficial for providers with non-standard patterns of recruitment, while others were concerned that there may be a negative impact on the quality and accuracy of data returned by those providers. Other respondents were concerned about how this approach may impact on non-traditional provision and whether it would prove future-proof, particularly with regard to the move towards Lifelong Loan Entitlements (LLE). One respondent suggested that the assumption of the primacy of an autumn intake within this approach may favour traditional providers in terms of data usage by the OfS and league table compilers.

OfS response

The original Data Futures approach was designed to ensure that data on non-traditional courses was no more lagged than for traditional courses. However, while non-traditional courses are growing, and are likely to grow further through the LLE, the majority of students still start courses in the autumn. It would be disproportionate to require additional reporting now on the basis that it may be required in the future. While we recognise the cost of change, we view moving to two collections per year is proportionate given the current landscape.

Additional suggestions

45. Several respondents made additional suggestions in relation to this approach, including that:
- HESA should consider constructing an output which incorporates the two composite populations in order to facilitate year-on-year reporting
 - only a single annual data return should be required in the first year or for a two-year transitional period

- the OfS conducts analysis to assess the quality of the data made through the early returns process
- data from the first return could also be used for the HESA ITT return
- an opportunity to input on further detail would be welcomed.

46. Alternative approaches proposed included:

- A reduced first HESA return followed by a revamped HESES return closely aligned with HESA.
- A cumulative approach (as in approach 2) but with a reduced return in reference period 1.
- A phased approach of approach 1 with a review after three years, and a long-term aim of adopting approach 2.
- Cutting down the in-year submission and instead having a cumulative full return at the end of the academic year.

Further information requested about the proposal

47. Approximately 25 per cent of respondents sought further information about this approach. The majority of these wanted additional detail about the data specification for the returns, in particular the first return, and the timetable for implementing the approach.

48. While we recognise that additional detail may have helped in better understanding precisely how the approach would be implemented in the first year, the overarching aim of part one of the consultation was to enable us to provide HESA with an overall approach to take forward the Data Futures programme of work at the earliest opportunity. It is important to acknowledge that we expect the agreed Data Futures approach to remain relevant for many years, during which time the detail of the data specification is likely to change. The overarching approach therefore needs to work irrespective of the detailed data specification, and we therefore judge that the level of detail provided was appropriate.

49. Part two of the consultation set out potential changes to the data specification of the Data Futures collection, of which question 6 sought comments on these proposals. Notwithstanding, in their response to question 1, several respondents requested further clarity about the data requirements for this approach, including information about:

- the in-year reporting requirements for entry qualification data, placement data and student characteristics
- the impact of the proposals on students in non-traditional provision with non-standard patterns of recruitment, such as those on roll-on-roll-off programmes, those taking modular continuing professional development provision, or early leavers.
- the level of data quality tolerance we would apply to the data returns.

50. We will publish further information about data specifications in our response to part two of the consultation.

51. A few respondents sought further information about the timing of the data return reference periods, sign-off, and the submission dates in the transition year. Our final decisions provide this information.
52. Several respondents requested further information about the removal of HESES as part of this approach, including whether there would be any changes to the funding approach should it be removed or a requirement for additional information to capture forecasts and funding completion expectations. Other respondents requested further information about how this approach would impact the ability to correct or update data. We have addressed both of these issues in our final decisions.
53. Further requests for information about approach 1 focused on the impact of the proposals on other stakeholder returns, how this approach would avoid an unmanageable peak in work, how data error would be distinguished from data updates, and the impacts of this approach on specifications and software providers. All of these points have been covered in our response to comments on approach 1.

Comments on approach 2: Cumulative in-year individualised student data collection with potential for differential reporting

54. In this section we have summarised the responses we received in relation to question 2 about approach 2.
55. Approach 2 cumulatively builds up collection of student data over a year, with three reference periods covering August to November, August to March and August to July. Students who complete or leave part way through the academic year would continue to be returned until the end of the academic year. The final reference period in the cumulative approach would be a complete data return for the entire academic year (1 August to 31 July).
56. This approach allows some flexibility over which of the two in-year returns a provider makes. It would be possible for the OfS not to require one or more of the in-year returns as a provider would return all the data at the end of the year. This flexibility means that it may be possible to have different in-year returns for different providers to match the main dates in their academic calendar. It would also be possible, once the new in-year collection approach was established, to vary collection frequency, but not timing, to reflect the OfS's view of regulatory risk for a provider.
57. The majority of respondents (approximately 70 per cent) who commented on approach 2 raised concerns about the proposal. The most frequently cited concerns were about increased burden and complexity. Only nine respondents ranked this approach as their preferred option, while 11 respondents ranked this approach as their least preferred option.
58. Comments from respondents on approach 2 could be broken down into seven general themes:
 - a. Comments about burden and complexity.
 - b. Comments about data accuracy and utility.
 - c. Comments on data requirements and timings.
 - d. Comments on implementation.
 - e. Comments about alignment with existing systems.
 - f. Comments about alignment with other returns.
 - g. Comments about flexibility.
59. There were also several requests for further information, as well as additional suggestions covering alternative approaches.

Burden and complexity

60. Overall, approximately half of respondents considered that this approach would result in increased burden. This included suggestions that:

- a. Making a second submission covering activity in August to March could be disproportionate in terms of the workload required and the benefits from it.
- b. Greater time would be spent undertaking quality assurance of data and the cumulative approach would result in an increased amount of redundant data.
- c. Staff and operational costs of making three returns may increase, particularly for smaller providers.
- d. Submitting a full set of data for the first reference period could have significant resourcing implications given that the period coincides with other statutory collections including: HESA staff, HESA aggregate offshore, ITT in-year, HESES.
- e. This approach deviated from the original Data Futures model in a way which would limit the usefulness of existing preparation by providers, for example in terms of technology, systems and processes.

61. It was also suggested by some respondents that this approach was overly complex. Comments of this nature included:

- a. The approach would have few advantages if only a small number of students commenced their studies after the first collection.
- b. Differential reporting may limit the utility of the data across the sector.
- c. The inclusion of students in multiple reference periods within the reporting year would lead to redundancy of data and unnecessary complexity.
- d. Providers with a substantial number of students on non-standard intakes, whose scheduled completion dates are spread across the year, could expect to return a high proportion of records at the later two data collection points that had already been returned in a previous dataset.
- e. Small institutions may struggle to meet the proposed requirements.

62. Some respondents did highlight this approach as potentially reducing burden, but they were a small minority. Their reasons included:

- a. It would enable the removal of HESES.
- b. Although three discrete collections might not be completely necessary currently, there would be less upheaval in the sector in the future if the timetable of data collection remained the same, but the process changed.
- c. As the returns would be drawn from the student record system, any data that was in scope for the reference period could be automatically selected, as currently happens with ITT census updates and the ILR.

Data accuracy and utility

63. There were mixed views about how this approach would impact on data accuracy and the utility of the data reported. While some respondents highlighted that they felt the approach

would allow for greater timeliness of data and greater accuracy overall, there was disagreement about the approach to dealing with data errors and the utility of the data.

64. A few respondents expressed support for the cumulative model of data collection, rather than the discrete model, citing the ability to more easily correct and update data. In particular, respondents supported the creation of opportunities to ensure continuous improvements in data quality and to simplify the data amendments process to correct data errors (which may be particularly beneficial during the transition year), and the potential to reduce pressure on providers to get all data complete and accurate for the first data return.
65. However, a number of respondents expressed concern that a cumulative collection model would lead to the same data being returned more than once which would lead to redundancy of earlier data and an increased chance of data errors creeping in to returns.
66. Multiple respondents queried how data error would be distinguished from data updates. Concerns were focused on the requirement for a significant degree of checking to assure data quality, and the risk of weakening internal messaging being used to drive improvements in data quality at the point of capture and processing. Supporters of the proposal felt that the approach to distinguishing data error from data updates should be praised, including stating that in-year continuity checking should not create a barrier to implementation.
67. Several respondents felt that this approach would create challenges when comparing and benchmarking data using this proposal, particularly across providers. If different providers were to submit at different reference periods, the data would not form a comparable dataset for any of the optional reference periods, making it difficult to gain a view on institutional performance compared with benchmarkable data, and diminishing parity between data collection for all providers. The ONS raised concerns about the potential for inconsistencies to arise in statistics if differential reporting arrangements were agreed for providers. The data available from providers returning three data collections a year would be more timely at any point in the year compared with those providers returning one or two data collections.
68. The reduced value that could be extracted from data with partial coverage of providers could mean that, for those providers required to make all returns, the value to burden ratio would be lower.

Data requirements and timings

69. A few respondents raised concerns about the data requirements inherent in this approach, including:
 - a. The potential for duplicated returns for students with more than one course instance. For example, the record of an undergraduate student who completes in December would be entirely separate to the record for that student if they started a postgraduate course in January.
 - b. The idea of changing the frequency of collection based on regulatory risk which could prove impractical in light of the lead times required to ensure both systems and staffing are in place to deliver returns.

Implementation

70. A few respondents raised concerns about the implementation of this approach, in particular that:
- a. There would be insufficient time to implement the changes this approach proposed as it represents a considerable change from the current Data Futures project. As such, this approach would require a significant investment of time and resources.
 - b. It would be unfeasible and burdensome to quality assure data in a limited timescale such as a transition year.

Alignment with existing systems

71. Several respondents raised concerns about how this approach would align with existing systems. Respondents discussed concerns relating to specification and software changes, as well as challenges in using this proposal with existing student record systems. The approach was not one which had been envisioned or specified to student record system software providers and would require significant redevelopment work. While software providers might in theory be able to support the revised requirements, this would be too late to be viable. The Data Futures programme is about to begin beta testing, so this approach would effectively mean starting the project again.

Alignment with other returns

72. A few respondents had concerns about how this approach would align with other stakeholder returns, in particular the data requests from PSRBs which were likely to continue on the basis that they collect data using slightly different definitions and to different schedules. Other statutory data collections due in the first reference period included the HESA staff record, HESA aggregate offshore, ITT in-year and HESES.
73. A couple of respondents were more positive about how this approach would align with other returns, including the suggestion that it would create the potential for greater alignment between higher education and further education as it aligns with the ILR more effectively than the other approaches. The ONS was broadly supportive of this approach on the basis that three data collections would provide them with more timely statistics, but raised concerns about consistency of statistics if differential reporting arrangements were agreed.

Flexibility

74. A couple of respondents questioned how this approach would work for providers with non-standard patterns of recruitment and multiple start dates, as this approach appeared to assume the precedence of the traditional autumn intake. One respondent was concerned that having a substantial proportion of students on non-standard intakes and deliveries, whose scheduled completion dates are spread across the year, would result in a high proportion of records needing to be returned at the later two data collection points that would already have been returned in a previous dataset.
75. However, other respondents liked the apparent flexibility of this approach, suggesting that it would enable providers to continue to offer flexible and mid-year entry points and would be effective for reporting data on non-traditional provision, making it easier to make timely returns on short course provision. Two respondents highlighted that this approach would help to align

data collection with a flexible delivery approach, future-proofing flexibility in reporting requirements needed to support continued innovative delivery across the sector.

76. Some respondents liked the potential flexibility of providers being able to skip collection points where there are no interim start dates; however, it was noted that the majority of providers would most likely end up making three returns under this approach.

Additional suggestions

77. A few respondents made additional suggestions in relation to this approach, including:
- a. Collecting data twice a year rather than three times a year.
 - b. Moving to this approach from an approach closer to approach 1 when circumstances necessitate it.
 - c. Combining the cumulative data collection approach with the two collections per year model of approach 1 as an alternative way of delivering approach 1.
 - d. Introducing flexibility on the number of returns required based on FTE size and onsite delivery.
78. Respondents identified the following as essential elements to the success of this approach:
- a. Limiting the cost to providers.
 - b. Requiring different quality standards for the different reference periods.

OfS response

The comments received largely aligned with the thinking set out in the consultation document which led to this not being the preferred option. We have noted the points made by providers and data users about the impact of differential reporting on the value that can be derived from the data.

The Higher Education and Research Act 2017 requires us to have regard to the principles of best regulatory practice, including that regulatory activities should be proportionate (see HERA section 2(1)(g)(a)). We therefore do not accept that we would be likely to require all three returns from providers even when the differences between the returns were limited. However, we recognise that by not being able to set out our decision making process in advance that these concerns exist.

We considered whether combining the cumulative collection model with a requirement for all providers to make two returns represented a useful hybrid of approaches 1 and 2. We consider the primary driver for a cumulative approach to be the benefit that returns can be omitted; without this benefit, cumulative collection has the potential to increase burden through requiring data to be returned more than once. We do not consider the benefits of cumulative collection where all returns are mandatory outweigh the issues of returning data multiple times. We also note that this would require more significant systems change as it is a more significant departure from the current approach.

We note that a number of providers were concerned about this approach because effort that had been expended on the current Data Futures approach would be wasted. We attached limited weight to these concerns as we have taken the view that the approach adopted will be in place for many years, so the cost of change is likely to be outweighed by the benefits of a different approach.

Further information requested

79. Several respondents requested clarification about the approach, including further information about what the data requirements would be for each return under this approach if it were to be taken forward, and what the data quality requirements would be for each of the reference periods. As set out in paragraph 48, we consider that the consultation contained sufficient information for respondents to understand our proposals in relation to the broad approach to in-year data collection. We do not consider a detailed description of individual data items to be necessary to allow meaningful consultation responses because we expect this to change over time. Other respondents asked whether penalties would be incurred if corrections were made in the second and third returns, and how we would distinguish data error from data updates.
80. A few respondents asked for further information about whether there would be two or three compulsory returns under this approach. They also asked what criteria we might use to determine whether a provider needed to make the second return if this was not compulsory, for example the minimum number of students necessitating a return in the second reference period. We set out our intention in paragraph 63 of the consultation document to explore the options for flexing the number of returns based on the regulatory risk for a provider.
81. Other respondents asked how this approach would align with non-standard patterns of recruitment, how it would impact on other stakeholder returns, why we did not propose a reduced return concept for this approach, how we would prevent the approach resulting in an increase in burden and how the transition year would work.

Comments on approach 3: Changing the timing of collection

82. In this section we have summarised the responses we received in relation to question 3 about approach 3.
83. One way to meet many of the needs for in-year data while retaining a single collection of individualised student data would be to change the timing of the reporting periods, so that instead of covering an academic year (1 August to 31 July) they cover a different period, for example 1 November to 31 October. Under this approach, the data from this single return would be available for use early in the next calendar year.
84. The majority of respondents (85 per cent) who commented on this approach raised concerns about the proposal. The most frequently cited concerns were about the complexity and additional burden this approach would lead to, the proposed timing of the return and the alignment with existing student record systems. Support for approach 3 was very limited: only four respondents ranked this approach as their preferred option; 64 respondents ranked this approach as their least preferred option; a further 54 respondents ranked this as their third option. Weighting the rankings, this approach was ranked fourth.
85. Comments from respondents on approach 3 could be broken down into seven general themes:
 - a. Comments about burden and complexity.
 - b. Comments about data accuracy and utility.
 - c. Comments on data requirements and timings.
 - d. Comments on implementation.
 - e. Comments about alignment with existing systems.
 - f. Comments about alignment with other returns.
 - g. Comments about flexibility.
86. There were also several requests for further information, as well as additional suggestions covering alternative approaches.

Burden and complexity

87. Many respondents expressed concerns about the potential for increased burden resulting from this proposal. Respondents cited the increased resourcing burden (staff, time, financial) of supporting the realignment of existing student data systems to the new reporting period, and of reviewing internal reporting requirements for comparability against other data sets collected in-line with the current academic year.
88. A few respondents noted that this approach would create a significant peak in work as the timing of the reporting period would coincide with the HESES return.

89. A small number of respondents raised the cost of the proposed approach as a result of changes to student record system software, or increased assurances required by external auditors.
90. Several respondents considered that this approach would not reduce burden compared to the currently proposed Data Futures model. However, a few respondents considered that the requirement for a single data collection per year could lead to a reduction in burden.
91. Many responses discussed the complexity of the single return approach. This included comments about the difficulty and unsustainability of administering the proposed approach, the complexity of applying this approach to students whose activity does not fit standard patterns, the complexity of the data checking process, and the potential for the cost to outweigh the benefits.

Data accuracy and utility

92. Several providers expressed concerns about the ability to benchmark and compare data should this proposal be adopted. It was suggested that submissions spanning two academic years would limit clarity about which year aspects of the HESA return referred to, and that comparisons with other data sets (including past data sets) would not be possible for both internal reporting and external comparison.
93. A number of respondents agreed with our assessment that it would create issues with comparability with HESA staff and HE-BCI data. Respondents noted that there would be similar comparability issues for finance data for those providers that have financial years aligned to the 1 August to 31 July academic year.

Data requirements and timings

94. Many respondents expressed concerns about the timing of data returns and sign-off. It was suggested that the October return would not fit with enrolment timings, and that postgraduate masters' completions may not be captured without moving academic processes, such as marking and examination boards. The effect of this might be that some postgraduate completions would need to be carried over to the subsequent year's data return, meaning that a one-year postgraduate student might need to be included in three different annual returns.
95. Respondents were concerned that it would be burdensome to manage students who would span multiple academic years and that moving away from the traditional academic year would be challenging. It was also questioned as to whether this approach would deliver the aims of Data Futures around the timeliness of sector data.

Implementation

96. A small number of respondents noted that there would be insufficient time to implement the changes set out in this proposed approach, including suggesting that it would require the updating of several internal processes to move returns to the new census dates. A similar number of respondents expressed concerns about the proposed approach to the transition year, discussing the potential for significant upheaval during that period.
97. One respondent suggested that this approach could have been their preferred option had there been more time to explore and establish a viable implementation solution. However, with the time available, they considered this option was not feasible.

Alignment with existing systems

98. Many respondents commented on the impact of this approach on specifications and software providers. It was suggested that this approach would have significant technical implications and require a high degree of system change and redevelopment, which could represent disproportionate burden.
99. Many respondents were concerned about how this proposed approach would align with existing student record systems. Respondents suggested that this approach would create unresolvable problems for planning and benchmarking, and that it would require major changes to student record systems which would likely result in additional costs to providers, software providers and HESA.

Alignment with other returns

100. A few respondents expressed concerns about the impact of this proposed approach on other stakeholder returns. This included suggestions that there would be an impact on the STAR(*)J data received from UCAS, as well as about alignment more generally with other returns, such as the HESA staff and finance returns.
101. A small number of respondents were concerned about the impact of this approach on the HESES return; however, others suggested that this approach could bring additional benefits, including to the populations for the NSS and Cohort D of the Graduate Outcomes survey.
102. The ONS considered that the proposed approach to move away from collecting data aligned with the academic year would be detrimental to their analysis and research, and additionally would not improve timeliness of HESA data.

Flexibility

103. A few respondents expressed concerns about how the proposed approach would impact providers with non-standard intake patterns, stating there may be an effect on data quality. There were also some concerns about how this proposal would impact non-traditional provision, including stating that it would require a large amount of development work.
104. One respondent supported this approach, suggesting it could benefit non-standard provision in terms of the timely reporting of qualifications such as nursing courses, masters' programmes and fast track awards.
105. Some respondents echoed the points raised in our consultation about how future-proof this approach would be as it cannot be extended if more frequent data is required in future, for example to reflect the increase in flexible learning which will be taking place over the next 10 years.

Additional suggestions

106. Several respondents made additional suggestions in relation to this approach, including:
 - a. That this approach would be improved by the removal of HESES.
 - b. An alternative approach could be to implement the timing of the reporting periods as set out in approach 3, with the addition of an in-year submission in March.

- c. Moving the census date for a single return to 1 December to improve coverage and completeness.
- d. An alternative date in January or February could offer more complete data for new entrants and postgraduate masters' outcomes, but would mean waiting even longer for undergraduate outcomes to be confirmed than with the current data collection cycle.
- e. This approach could be reconsidered in the future once the Data Futures model is established.

OfS response

The comments received largely aligned with the thinking set out in the consultation document which led to this not being the preferred option. Respondents raised significant concerns about the complexity and burden that would be introduced through this approach. The alternative options that were suggested would not have significantly addressed the concerns raised by respondents about the complexity of the approach as they would still cut across existing student record systems and academic processes in providers.

Adding an additional return in March would negate one of the main benefits of this approach as it would introduce a second return while retaining the complexity of shifting the reporting year.

Further offsetting the reporting year from the academic year would delay the availability of data and critically would not allow data to be available to support the NSS which is one of the key uses of the in-year data.

There were no comments made in the responses that led us to reconsider whether this option should be taken forward.

Further information requested about the proposal

- 107. A few respondents noted that approach 3 represented an interesting solution but would need further work to understand all the implications of such a radical shift. In particular, further information was requested about how this proposal would impact on other stakeholder returns, and how it would impact on the HESES return.
- 108. A small number of respondents requested further information about the detailed data requirements of this proposed approach. This included asking whether returns would consider the 'intended STULOAD' or look at STULOAD retrospectively, and how FTE reporting would operate under this model. As set out in paragraph 48, we do not consider it would be appropriate to base the approach on the specific data requirements now as we expect these to evolve.

Comments on approach 4: Discrete in-year individualised student data collection three times a year

109. In this section we have summarised the responses we received in relation to question 4.
110. Approach 4 is the current Data Futures approach, which we originally considered as an alternative proposal but have now discounted.
111. The original plan for Data Futures was for HESA to make three discrete collections of in-year individualised student data. Under this model, providers would make three returns in each academic year covering the reference periods August to November, December to March, and April to July. Each return would cover all the students who are active during the period covered by the return, and each would operate independently, with no overlap in the reporting of activity between the returns.
112. Half of all respondents who commented on this approach said they supported the proposal to discount this approach. The most frequently cited concerns about approach 4 were about the increased burden and complexity of returning data three times a year. 67 respondents ranked this approach as their least preferred option and a further 35 respondents ranked this as their third option. By weighting the rankings, this approach was ranked third.
113. Support for proceeding with approach 4 was limited: only six respondents ranked this approach as their preferred option. Of those, the rationale cited was that this was the approach that providers and software suppliers had been working towards and processes, systems and resources had been focused on delivering this model. There was some concern that it would be short-sighted to discount this approach so close to the original launch period.
114. There were fewer comments about this approach compared with the other proposals, reflecting the fact that we had discounted it as a possibility within the consultation. Those comments received from respondents on approach 4 could be broken down into seven general themes:
 - a. Comments about burden and complexity.
 - b. Comments about data accuracy and utility.
 - c. Comments on data requirements and timings.
 - d. Comments on implementation.
 - e. Comments about alignment with existing systems.
 - f. Comments about alignment with other returns.
 - g. Comments about flexibility.
115. There was also one additional suggestion for an alternative approach.

Burden and complexity

116. Almost half of respondents commenting on approach 4 had concerns that proceeding with this approach would result in an increased or unmanageable burden, and as such many respondents welcomed the proposal to discount this approach.
117. Respondents considered that making three discrete data returns per year would be too burdensome, particularly for small providers. The main reasons cited for the increased burden related to the data quality and assurance processes required for all three data returns.
118. A few respondents also considered that the data quality assurance processes required to ensure consistency between the three independent data collections made approach 4 overly complex. A small number of respondents noted that the second data return would have minimal value, especially for providers with no or limited student intakes or completions between November and March.
119. A small number of respondents had concerns that this approach would result in a significant peak in work around Christmas for the first collection, particularly if HESES were to remain. Having three collections means there is very little downtime in the cycle – which is critical when departments need to do system changes, upgrades and improvements (IT and procedural).
120. A small number of respondents considered that approach 4 was manageable and that moving away from this approach would not result in reduced burden or that any reduction would not be significant enough to justify a change to approach at this stage.

Data accuracy and utility

121. As noted in paragraphs 117 and 118, respondents were concerned about the burden of the quality assurance processes required to return accurate data in each of the three reference periods.
122. Several respondents expressed concern about the complexity of updating records (either to improve data or to correct errors) when data is not automatically superseded. The process of amending data could be resource heavy if small changes of this type were widespread across the sector
123. One respondent was concerned that the approach would not distinguish between data error and data improvement between reference periods, with providers required to submit an error report for every item changed between reference periods. Another noted that this approach could represent an increased cost to providers if they were required to pay for amendments to the database for single changes between reference periods.
124. One respondent considered that data quality could be compromised with this approach because of the timing of the first return which would contain all data items.
125. One respondent had concerns about the ability to compare and benchmark data with this approach as there would be a lack of alignment with other institutional reporting periods. Another respondent was concerned that they would no longer have discrete data for the whole of the academic year.

126. One respondent felt that this approach would improve the quality of data returns as it would allow well defined processes and systems to be put in place.

Data requirements and timings

127. A small number of respondents expressed concerns about the timing of data returns and sign-off, in particular for the first data return. Respondents raised concerns about staffing issues around Christmas creating difficulties for the first return and the reduced time for collation and validation of data.
128. A similar number of respondents expressed concerns about the data requirements contained within this approach, highlighting the heavy data burden on providers for the first collection and concerns that for some providers, without a January intake of students, the data contained in the second return would likely be irrelevant. These respondents welcomed the proposal to discount this approach and move away from three collections a year.
129. A small number of respondents with a significant volume of student activity outside the standard academic year, considered that the timing of data returns in approach 4 would fit better with their academic model and recruitment cycles. The second reference period would align well for those providers with a large January intake of students.
130. One respondent noted that an advantage of approach 4 was that the data specification would be consistent throughout, and therefore easier to understand and maintain – the variable being the reporting window. Another respondent noted that approach 4 felt cleaner and more structured than other approaches.

Implementation

131. A small number of respondents noted that a significant amount of work had already been completed in preparing for this approach and, while the approach had its limitations, it would be difficult for providers and software suppliers to comply with any other option within the specified timeframe. Additionally, a small number of respondents expressed concerns about the impact of discounting this approach on software providers who had been developing Data Futures specifications in line with approach 4.
132. In their response to the consultation, the software provider Tribal highlighted that both approach 1 and approach 4 could be supported by the SITS HESA Data Futures component as it stands in the present. This reflects the fact that approach 1 is technically very similar to approach 4, and the current software solution is flexible in how it supports reporting periods.
133. One respondent had concerns about the challenges of adopting this approach in the first year of transition.
134. A small number of respondents were concerned that whatever approach was taken forward, all the devolved nations should be following the same model. We recognise the desirability of UK-wide data and have been engaging with colleagues in the devolved nations in developing our proposals. However, each nation will need to take decisions based on their own context.

Alignment with existing systems

135. One provider considered that approach 4 would work well for them, as their existing student record systems would align with the activity that needs to be returned under this approach.

136. Another respondent considered that the even split of data returns through the year would be beneficial as this makes scheduling the work simpler.
137. One respondent raised concerns that this approach would impact on their ability to produce internal statistical reporting for management purposes.

Alignment with other returns

138. The ONS supported approach 4, stating that they felt the timeliness of the data provided would bring benefits to stakeholders. The approach would maintain accuracy and increase frequency, providing a better idea of student information throughout the year, rather than at one point during the year.
139. One respondent felt that this approach was the only one that would maintain the original aims of Data Futures, to align HESA reporting requirements with those of other regulatory bodies.

Flexibility

140. A small number of respondents felt that this approach should not be discounted as it provided the greatest flexibility to accommodate non-traditional provision and the growth of more flexible learning, such as the future implementation of higher technical qualifications and the Lifelong Loan Entitlement.
141. Two respondents suggested that approach 4 seemed the most future-proof and responsive to changes in the wider higher education landscape, as well as best able to respond to the possibility of more shared data. Another respondent wanted to see a consistent and co-ordinated approach to data collection across post-18 regulators to better reflect the government's policy ambitions.

Additional suggestions

142. One respondent suggested that if we were to adopt the timetable of data returns outlined in approach 4, then a cumulative approach (as detailed in approach 2), where student data could be updated in each return, would be preferable and would ensure higher data quality in the final end of year return.

OfS response

The comments received largely aligned with the thinking set out in the consultation document which led to this being a discounted option. We recognise that for some providers there may be some benefit in having three returns of data per year. However, these benefits are likely to be small, so we do not consider that the benefits for these providers outweigh the significant additional costs for other providers where two returns of data are sufficient.

We consider that the suggestion to adopt a cumulative approach to data collection as part of approach 4 would have the potential to increase burden through requiring data to be returned more than once. The responses we received surrounding approach 2 mean that we do not consider the benefits of cumulative collection in enabling data to be updated in each return outweigh the issues of returning data multiple times.

We note the concerns raised by providers that, as this was the current approach, significant effort had been invested in delivering it. However, we note that our preferred approach is very similar to this option so most preparation is likely to transfer between the two options.

We recognise the risk that, as more providers offer more flexible provision, we may need to increase the number of returns. If that were to happen, the savings in burden reduction in the short term may be outweighed by the costs of two changes. However, it is not clear that the Lifelong Loan Entitlement proposals and increasingly flexible and modular provision will require a move to three collections per year. Given that there is no clear requirement for data more than twice a year, we do not think it is proportionate to impose a burden on providers now where two data returns are sufficient.



© The Office for Students copyright 2022

This publication is available under the Open Government Licence 3.0 except where it indicates that the copyright for images or text is owned elsewhere.

www.nationalarchives.gov.uk/doc/open-government-licence/version/3/