

Contents

Page

Executive summary		4
Introduction		7
Annexes		
Annex A	Summary of changes since HESA 2002-03 derived statistics (HEFCE 2004/10)	14
Annex B	Comparison of HESES03 and the HESES03 re-creation	16
Annex C	Comparison of the HESES03 re-creation and the HESES03 re-creation based on cost centre sector norms for subjects	21
Annex D	Comparison of RAS03 and the RAS03 re-creation	28
Annex E	Research degree rates of qualification	33
Annex F	HEFCE statistical publications	35
Annex G	Obtaining data from the HEFCE extranet	37
Annex H	Guidance for action and implementation plans	39
Annex I	Submitting amendments to HESA data	46
Annex J	Submitting overrides to the algorithms	53
Annex K	List of abbreviations	59

Technical appendices (available on the web)

Appendix 1	HESES03 re-creation algorithms	
Appendix 2	Troubleshooting the differences between HESES03 and the HESES03 re-creation	
Appendix 3	Problems of fit with the HESES03 re-creation algorithms	
Appendix 4	HESES03 re-creation based on cost centre sector norms for subjects algorithms	
Appendix 5	Troubleshooting the differences between the HESES03 re-creation and the HESES03 re-creation based on cost centre sector norms for subjects	
Appendix 6	Problems of fit with the HESES03 re-creation based on cost centre sector norms for subjects algorithms	
Appendix 7	RAS03 re-creation algorithms	
Appendix 8	Troubleshooting the differences between RAS03 and the RAS03 re-creation	
Appendix 9	Problems of fit with the RAS03 re-creation algorithms	
Appendix 10	Research degree rates of qualification algorithms	
Appendix 11	HEFCE statistical publications algorithms	
Appendix 12	Example action and implementation plans	

2003-04 statistics derived from HESA data for monitoring and allocation of funding

To Heads of HEFCE-funded higher education institutions

Of interest to those responsible for Student data, Funding, Audit, Research

Reference 2004/47

Publication date 17 December 2004

Enquiries to For enquiries regarding the HESES03 comparison contact:

Anne Southworth tel 0117 931 7415

Scott Chu tel 0117 931 7365

Matthew Davies tel 0117 931 7373

e-mail hesa_heses_stats@hefce.ac.uk

For enquiries regarding the RAS03 comparison contact:

Hannah Wood tel 0117 931 7395

e-mail hesa_heses_stats@hefce.ac.uk

For enquiries regarding the use of HESA data to monitor achievement of threshold standards for research degree programmes, contact:

Ben Grassby tel 0117 931 7260

e-mail hesa_heses_stats@hefce.ac.uk

For enquiries regarding the cost centre sector norms comparison contact:

Anne Southworth tel 0117 931 7415

Scott Chu tel 0117 931 7365

Matthew Davies tel 0117 931 7373

e-mail hesa_heses_stats@hefce.ac.uk

For enquiries regarding the use of HESA data in HEFCE publications contact:

Lisa Readdy tel 0117 931 7492

e-mail hesa_heses_stats@hefce.ac.uk

Executive summary

Purpose

1. This document describes:
 - a. How we will use 2003-04 Higher Education Statistics Agency (HESA) student data to monitor returns made to HEFCE.
 - b. The responses required to these monitoring processes.
 - c. How we will use 2003-04 HESA student data to monitor achievement against threshold standards for research degree programmes.
 - d. How we will use 2003-04 HESA student data in HEFCE statistical publications.
2. This document is divided into the following sections:
 - a. The comparison of Higher Education Students Early Statistics Survey 2003-04 (HESES03) with HESA 2003-04 student data.
 - b. The comparison of cost centre assignments with cost centre sector norms for subjects.
 - c. The comparison of Research Activity Survey 2003 (RAS03) with HESA 2003-04 student data.
 - d. The use of HESA 2003-04 student data to monitor achievement against threshold standards for research degree programmes.
 - e. The use of HESA student data in HEFCE statistical publications.

Information on how we intend to use HESA data to inform funding, including widening participation (WP) allocations, will be announced separately in early 2005.

Key points

Data quality

3. We are confident that this exercise improves the data quality of both HESA and HEFCE returns. It also increases our understanding of data quality issues that relate to these returns.

Funding monitoring

4. The exercise is conducted in two interrelated but distinct parts. The first is the process of reconciling, explaining and amending the data up until the point where institutions are in a position to sign off a re-creation as a reasonable reflection of the outturn position for the year. The second part, which follows sign off, is the consideration of funding adjustments made, and the appeals process.
5. Our funding allocations are informed by the data provided by institutions. If we find, either through reconciliations with HESA data, or any data audit, that data do not reflect the outturn position for the year, and that this has resulted in institutions receiving incorrect funding allocations, then we will adjust their funding accordingly. (This is subject to the appeals process and the availability of our funds.)
6. Any funding adjustments arising from the reconciliation of a re-creation of HESES03 from 2003-04 HESA student data (the HESES03 re-creation) with HESES03 or from the comparison of cost centre assignments with the sector norms for subjects, are likely to affect the funding previously announced for 2003-04 and all subsequent years.
7. Any funding adjustments arising from the reconciliation of RAS03 with a re-creation of RAS03 from 2003-04 HESA student data (the RAS re-creation) are likely to affect the funding previously announced for 2004-05.
8. In some cases the funding adjustments may be significant. Therefore it is important for institutions to ensure that sufficient time and resources are allocated to allow the exercise to be completed accurately and promptly. If institutions have not signed off their re-creations by the deadlines given below, then we will implement any reductions to 2005-06 grant that we expect to arise, pending completion of the reconciliation process. This is an interim measure to avoid grant adjustments accumulating to the point at which they become difficult for institutions to manage. The deadlines are **27 May 2005** for institutions selected to respond to the comparison of HESES03

and the HESES03 re-creation; and **6 July 2005** for institutions selected to respond to the comparison of the HESES03 re-creation and the HESES03 re-creation based on cost centre sector norms for subjects.

Annexes and appendices

9. The annexes to this publication describe how we will use HESA data for this exercise. The web-only appendices to this publication contain technical descriptions of the algorithms we will use. The appendices are on the HEFCE web-site www.hefce.ac.uk, with this document under Publications.

Action required

Funding monitoring

10. We will write to heads of institutions, copied to HESES and RAS contacts, on 17 December 2004 specifying whether a response is required to any part of the exercise.

11. Where a response is required, action and implementation plan(s) must be sent to Ben Grassby by **11 February 2005**. The final deadline for receipt of amendments to HESA data and overrides to derived fields detailed in the action and implementation plan(s) is **29 March 2005**.

HESA data for verification and publication

12. Where institutions wish to correct HESA data that will be used to monitor achievement against threshold standards for research degree programmes, they should submit amendments by **29 March 2005** to Ben Grassby.

13. Where institutions wish to correct HESA data that will be used in HEFCE statistical publications, they should submit amendments by **29 March 2005** to Lisa Readdy. We will assume institutions are content for their HESA data to be published if they do not submit amendments by this date.

Timetable

14. The following timetable details the critical deadlines for this exercise.

17 December 2004	Issue of letter to all institutions, requesting response to exercise where appropriate
11 February 2005	Deadline for receipt of final action and implementation plan(s) produced by each institution required to respond
29 March 2005	Final deadline for receipt of amendments to the HESA data and overrides to derived fields as detailed in action and implementation plan(s)
29 March 2005	Final deadline for receipt of amendments which affect research degree rates of qualification
29 March 2005	Final deadline for receipt of amendments which affect HEFCE statistical publications
27 May 2005	For institutions asked to respond to the comparison of the HESES03 and HESES03 re-creation: deadline for confirmation that the HESES03 re-creation reasonably reflects the outturn position for 2003-04 to avoid interim grant adjustments (see paragraphs 38 - 43 of the Introduction)
6 July 2005	For institutions asked to respond to the comparison of the HESES03 re-creation and the HESES03 re-creation based on cost centre sector norm for subjects: deadline for confirmation that the HESES03 re-creation reasonably reflects the outturn position for 2003-04 to avoid interim grant adjustments

15. Table 1 summarises the response required for each of the comparisons, along with the possible causes of differences.

Table 1 **Response process for institutions required to respond**

Comparison causing selection	Differences to explain in action and implementation plan	Possible causes of differences
HESES03 and the HESES03 re-creation	All differences between HESES03 and the HESES03 re-creation	<p>Errors in HESA 2003-04 student data</p> <p>Errors/estimation discrepancies in HESES03</p> <p>Problems of fit with the HESES03 re-creation algorithms</p>
HESES03 re-creation and the HESES03 re-creation based on cost centre sector norms for subjects	All differences between the HESES03 re-creation and the HESES03 re-creation based on cost centre sector norms for subjects, to include any differences between the HESES03 and the HESES03 re-creation	<p>Errors in the HESA 2003-04 student data</p> <p>Errors/estimation discrepancies in HESES03</p> <p>Errors in the assignment of staff to academic departments</p> <p>Errors in the assignment of academic departments to cost centres</p> <p>Problems of fit with the HESES03 re-creation algorithms</p> <p>Problems of fit with the HEFCE cost centre sector norms for subjects mapping</p>
RAS03 and the RAS03 re-creation	All differences between RAS03 and the RAS03 re-creation	<p>Errors in the HESA 2003-04 student data</p> <p>Errors in RAS03</p> <p>Problems of fit with the RAS03 re-creation algorithms</p>

Introduction

16. This document describes how we will use 2003-04 HESA student data to monitor returns made to HEFCE. It also details the action required where either a response is requested or an institution wishes to correct errors in its HESA data. Descriptions are given of how we use 2003-04 HESA student data to monitor achievement against threshold standards for research degree programmes and in statistical publications.

Annual data returns

17. HESES and RAS data are used to determine the funding allocations made for teaching and research. HESES is used both to monitor the year's teaching funding allocation and to determine the teaching funding allocation for the following year. RAS is used to determine the research funding allocation for the following year. HESA student data are used to:

- a. Monitor HESES, RAS and the assignment of activity to cost centres through the re-creation of HESES and RAS returns. If we find, either through an institution's response to our reconciliations using HESA data, or any other method of assurance or data audit, that the HESES or RAS submission does not reflect the final outturn position for the year and that incorrect funding allocations have occurred as a result, then we will adjust the HEI's funding accordingly (subject to the appeals process and the availability of our funds.)
- b. Inform funding allocations where the necessary information is not collected on HESES (for example, qualification on entry and age data for determining the widening participation allocations).

18. Our monitoring processes are applied consistently to all institutions. We receive HESA student data 12 months after the equivalent year's HESES and RAS returns. We expect all institutions to have used the HESES and RAS re-creations generated by the '2003-04 statistics derived from HESA data: Guide to HEFCE web facility'

(HEFCE 2004/29) to review their HESES and RAS returns before submitting their HESA returns.

Monitoring funding

HESES03

19. HESA 2003-04 student data will be used to monitor HESES03. A re-creation of HESES03 is generated from HESA 2003-04 student data using the methods detailed in Annex B.

20. We employ thresholds to select which institutions must respond to this comparison. These thresholds are set in terms of the funding differences arising between the HESES03 and the HESES03 re-creation. The selection process represents a risk assessment. Primarily, this assessment is intended to identify those institutions whose data differences are most likely to have a material effect on their funding allocations.

21. Each institution that is required to make a response must provide an action and implementation plan. The plan must:

- a. Explain each constituent cause of difference between HESES03 and the HESES03 re-creation.
- b. Contain an estimate of the contribution that each constituent cause of difference makes to the discrepancy in terms of student numbers and FTEs and, where appropriate, in terms of contract range holdback, funds to be held back or funds due back.
- c. Detail all actions required to reconcile the two data sources broken down by each cause of difference.
- d. Detail the system or process changes that will be implemented to ensure that similar discrepancies do not recur.

22. After both the institution and HEFCE are content that the discrepancies between the two data sources are explained, we will request confirmation from the institution that the HESES03 re-creation reasonably reflects the outturn position for 2003-04.

23. Once confirmation has been asked for and received we will generate both a RAS03 re-creation and a HESES03 re-creation based on cost centre sector norms for subjects, to incorporate any amendments that have been made to HESA 2003-04 student data. We will request a further response where the selection thresholds are exceeded for the comparisons of HESA 2003-04 student data with either RAS03 or the HESES03 re-creation based on cost centre sector norms for subjects.

24. Once confirmation has been asked for and received for all comparisons where a response is required, the HESES03 re-creation will supersede HESES03, and any consequent grant adjustments will be calculated and made (subject to the appeals process and the availability of our funds). The thresholds we use to select institutions must not be interpreted as being the minimum grant adjustments that we might effect. The latter are set out in the relevant grant adjustments publication; for example for 2003-04 this is 'HEFCE grant adjustments 2003-04' (HEFCE 2003/24).

Assignment of activity to cost centres

25. HESA 2003-04 student data will be used to monitor the assignment, by institutions, of activity to cost centres, and subsequently price groups. A re-creation of HESES03 is generated from HESA 2003-04 student data using the methods detailed in Annex B. For subjects where the total student FTE assigned across the subject is greater than 100, this re-creation is compared to a re-creation of HESES03 that is generated from HESA 2003-04 student data using the methods detailed in Annex B, except that activity is assigned to cost centres based on sector norms for subjects using the methods described in Annex C.

26. Thresholds are used to select institutions to respond to this comparison. These thresholds are set in terms of funding differences arising from the HESES03 re-creation and the HESES03 re-creation based on cost centre sector norms for subjects. The selection process represents a risk assessment. Primarily, this assessment is intended to identify those institutions whose data differences are most

likely to have a material effect on their funding allocations.

27. Each institution required to make a response will be asked to fill in an action plan, which we will provide on 17 December and will make available on the HEFCE extranet. This will contain a description of subjects that are assigned to a cost centre that differs from the sector norm cost centre, and for which we require an explanation. Institutions will also be asked to provide an action and implementation plan to explain the differences in HESES and the HESES re-creation comparison. The plans should:

- a. Explain each constituent cause of difference between the HESES03 re-creation and the HESES03 re-creation based on cost centre sector norms for subjects. That is, they should explain why the subjects that are listed on the action plan are assigned to a cost centre that differs from the sector norm cost centre.
- b. Explain each constituent cause of difference between HESES03 and the HESES03 re-creation.
- c. Contain an estimate of the contribution that each constituent cause of difference between the HESES03 re-creation and the HESES03 re-creation based on cost centre sector norms for subjects makes to the discrepancy, in terms of student numbers and FTEs, staff numbers and FTEs.
- d. Contain an estimation of the contribution that each constituent cause of difference between HESES and the HESES03 re-creation makes to the discrepancy, in terms of student numbers and FTEs and, where appropriate, in terms of contract range holdback, funds to be held back or funds due back.
- e. Detail all actions required to reconcile the HESES03 re-creation and the HESES03 re-creation based on cost centre sector norms for subjects.
- f. Detail all actions required to reconcile HESES03 and the HESES03 re-creation broken down by each cause of difference.

- g. Detail the system or process changes that will be implemented to ensure that similar discrepancies do not recur.

28. After both the institution and HEFCE are satisfied that the discrepancies are explained between the HESES03 re-creation and the HESES03 re-creation based on cost centre sector norms for subjects, and also between HESES03 and the HESES03 re-creation, we will ask for confirmation that the HESES03 re-creation reasonably represents the outturn position for 2003-04.

29. Once confirmation has been asked for and received we will generate a RAS03 re-creation to incorporate any amendments made to HESA 2003-04 student data. We will request a response where the selection threshold is exceeded for the comparison of HESA 2003-04 student data with RAS03.

30. Once confirmation has been asked for and received for all comparisons where a response is required, the HESES03 re-creation will supersede HESES03, and any resulting grant adjustments will be calculated and implemented (subject to the appeals process and the availability of our funds). The thresholds we use to select institutions should not be interpreted as being the minimum grant adjustments that we might effect. For 2003-04 these are set out in 'HEFCE grant adjustments 2003-04' (HEFCE 2003/24).

RAS03

31. HESA 2003-04 student data will be used to monitor forms R1a, R1b, R2a and R2b of RAS03. A re-creation of RAS03 is generated from HESA 2003-04 student data using the methods detailed in Annex D.

32. Thresholds are used to select institutions to respond to this comparison. These thresholds are set in terms of differences of total FTEs for eight groupings of broadly similar Units of Assessment (UoAs) between the RAS03 and the RAS03 re-creation. The selection process represents a risk assessment. Primarily, this assessment is intended to identify those institutions whose data differences

are most likely to have a material effect on their funding allocations.

33. Each institution required to make a response will be asked to provide an action and implementation plan. The plan must:

- a. Explain each constituent cause of difference between RAS03 and the RAS03 re-creation.
- b. Contain an estimate of the contribution that each constituent cause of difference makes to the discrepancy in terms of student numbers and FTEs.
- c. Detail all actions required to reconcile the two data sources, broken down by each cause of difference.
- d. Detail the system or process changes that will be implemented to ensure that similar discrepancies do not recur.

34. After both the institution and HEFCE are content that the discrepancies between the two data sources are explained, we will ask for confirmation that the RAS03 re-creation reasonably represents the outturn position for 2003-04.

35. Once confirmation has been asked for and received we will generate both a HESES03 re-creation and a HESES03 re-creation based on cost centre sector norms for subjects, to incorporate any amendments made to HESA 2003-04 student data. We will request a further response where the thresholds are exceeded for the comparisons of HESA 2003-04 student data with either HESES03 or the HESES03 re-creation based on cost centre sector norms for subjects.

36. Once confirmation has been asked for and received for all comparisons where a response is required, the RAS03 re-creation will supersede RAS03, and any consequent grant adjustments will be made (subject to the appeals process and the availability of funds).

Risk assessment

37. The necessarily complex process of explaining and resolving differences between data sources places a considerable burden on institutions and

HEFCE. To ensure this burden is both manageable and appropriate, the selection process represents a risk assessment. Primarily, this assessment is intended to identify those institutions whose data differences are most likely to have a material effect on their funding allocations. General improvements in data quality have enabled us to lower these threshold criteria in recent years.

Enacting grant adjustments – interim adjustments

38. The monitoring process can take many months to complete. In some cases in the past, by the time that confirmation is received that a HESES re-creation reasonably reflects the outturn position for the given year, the consequential grant adjustments have affected funding allocations over a four-year period. We recognise that this can be difficult for institutions to manage. Therefore, to reduce the risk of grant repayments accumulating to the point where they become difficult to manage, we will reduce monthly grant payments for institutions in the circumstances set out below.

HESES03 and the HESES03 re-creation comparison

39. We will reduce monthly grant payments for institutions where:

- a. We have requested a response to the HESES03 and HESES03 re-creation; and
- b. We have not asked for, or we have asked for and not received, confirmation that the HESES03 re-creation reasonably reflects the outturn position for 2003-04 by **27 May 2005**; and
- c. The grant adjustment for 2005-06 that would result from the HESES03 re-creation position shown on **27 May 2005** would represent a reduction in the teaching funding allocation for 2005-06.

40. In these circumstances we will effect the change to 2005-06 grant by using our own reasonable estimates, as at **27 May 2005**, of the final outturn position, reflecting the current HESES03 re-creation. The reduction in 2005-06

grant payments would be effected through the institution's standard monthly grant payment profile.

HESES03 re-creation and the HESES03 re-creation based on cost centre sector norms for subjects comparison

41. We will reduce monthly grant payments for institutions where:

- a. We have requested a response to the HESES03 re-creation and HESES03 re-creation based on cost centre sector norms for subjects; and
- b. We have not asked for, or we have asked for and not received, confirmation that the HESES03 re-creation reasonably reflects the outturn position for 2003-04 by **6 July 2005**; and
- c. The grant adjustment for 2005-06 that would result from the HESES03 re-creation and/or HESES03 re-creation based on cost centre sector norms for subjects shown on **6 July 2005** would represent a reduction in the teaching funding allocation for 2005-06.

42. In these circumstances we will effect the change to 2005-06 grant by using our own estimates, as at **6 July 2005**, of the final outturn position, reflecting the current HESES03 re-creation and/or the HESES03 re-creation based on cost centre sector norms for subjects. The reduction in 2005-06 grant payments would be effected through the institution's standard monthly grant payment profile.

43. Subject to the appeals process and the availability of our funds, we will make further grant adjustments, both for 2005-06 and for previous years as appropriate, once we ask for and receive confirmation that the HESES03 re-creation reasonably represents the outturn position for 2003-04.

Grant adjustments for institutions not required to respond

44. We do not gain assurance, through this exercise, regarding the reliability of either the HESES03 re-creation or the RAS03 re-creation for

institutions that have not been required to respond to these exercises. For such institutions we would not expect to adjust teaching or research funding allocations respectively based on these re-creations.

Further monitoring

45. We may audit data for institutions that are unable to provide acceptable explanations for the causes of discrepancies in any of the comparisons.

46. Notwithstanding the thresholds, we may also ask for further information from any institution in respect of any of the comparisons. This may result ultimately in adjustments to grant, where appropriate.

Funding allocations

Teaching funding allocations

47. A review is currently taking place to assess the relative costs of different types of provision in the cost centre for sports science and leisure studies. This was announced in 'Funding method for teaching from 2004-05: Outcomes of consultation' (HEFCE 2004/24). If the review determines that sports science and leisure studies activity should be reassigned to price groups, then HESA 2003-04 data will be used to incorporate any changes to the mainstream teaching funding allocation for 2005-06.

48. A letter will be sent to institutions early next year to inform them about the outcome of the review. If necessary the summaries of the HESA data used to inform the reassignments, along with details of the methodology used, will be made available to institutions early next year.

Widening participation funding allocation

49. It is likely that HESA 2003-04 student data will be used to inform the following WP funding allocations for 2005-06:

- widening access for full-time and part-time students
- widening access for disabled students
- improving retention.

Due to the effects of the introduction of the new UCAS tariff point system, which replaces the previous A-level points system, the 2005-06 WP funding allocations are being reviewed in February 2005. They are due to be published during March 2005. Institutions will be informed of the methodology that will be used by 1 April 2005, along with summaries of the HESA data used to inform the allocations. A timetable for submitting amendments to 2003-04 HESA data to inform WP funding allocations for 2005-06 will also be provided at this time.

Research degree rates of qualification

50. In our consultation paper 'Improving standards in postgraduate degree programmes' (HEFCE 2003/23) and the subsequent circular letter 'Postgraduate research degree programmes: minimum standards and funding' (HEFCE Circular Letter 18/2004) we made proposals on establishing minimum standards for postgraduate research degree programmes (RDPs). We envisage that all HEIs receiving HEFCE funding for research degree programmes should in due course be required to comply with appropriate minimum standards as a condition of their grant (see 'Review of research funding method', HEFCE 2003/38). In February 2004 it was announced that HEFCE will monitor the time that students take to obtain their qualifications through HESA returns. We will identify HEIs whose returns indicate a low proportion of students qualifying within a given time frame compared to the average in the sector. This will require robust figures covering a number of years. When there is cause for concern, and having established that this is not due to technical problems in data collection or relevant local circumstances we will enquire into the quality of provision at the institution and will ask QAA to follow this up initially.

51. The methodology used to monitor the time that students take to obtain their qualifications is described in Annex E.

HEFCE publications

Regional statistics

52. We intend to publish statistics derived from HESA 2003-04 student data for the following during 2005:

- students registered at one institution and taught by another institution
- campuses
- distance learning
- provision by location.

These data may be used for regional analysis and to inform policy decisions.

53. Details of the methods we intend to use to derive these statistics are given in Annex F.

HEFCE web facility for 2003-04 statistics derived from HESA data

54. On 6 August 2004 we made available the HEFCE web facility for 2003-04 statistics derived from HESA data (see HEFCE 2004/29). This facility is designed both to assist institutions in returning accurate data to HESA and to identify discrepancies between forecasting in HESES03 and the outturn position for 2003-04. We believe that the introduction of the web facility in 2002 and 2003 contributed to an improvement in data quality in the 2001-02 and 2002-03 HESA student data, and we expect a similar effect this year.

HEFCE- recognised funding consortia

55. For the lead institution of a HEFCE-recognised funding consortium, the HESES03 re-creations will incorporate data supplied by each consortium member. Data for provision in further education colleges (FECs) included in the consortium will be sourced from the ILR July 2003-04 student data submitted to the Learning and Skills Council (LSC). The algorithms used to generate HESES03 re-creation data for such FECs will be published in '2003-04 statistics derived from ILR data for the monitoring and allocation of funding in FECs', which we expect to issue in February 2005. We

will not receive ILR July 2003-04 student data from the LSC until late January 2005, therefore the timetable given on page 5 will differ for HEFCE-recognised consortia.

Next steps

56. We will write to heads of institutions, copied to HESES and RAS contacts, by 17 December 2004 explaining whether a response is required to this exercise. Enclosed with this letter will be the following sets of outputs:

- a. Output 1 – Comparison of HESES03 and the HESES03 re-creation.
- b. Output 2 – Comparison of the HESES03 re-creation and the HESES03 recreation based on cost centre sector norms for subjects.
- c. Output 3 – Comparison of RAS03 and the RAS03 re-creation: Forms R1a, R1b, R2a and R2b.
- d. Output 4 – Research degree rates of qualification (data that will be used to monitor the time that students take to obtain their qualifications).
- e. Output 5: – HEFCE statistical publications.

These data will also be available on the HEFCE extranet.

Guidance

HEFCE contact

57. Each institution required to make a response to this exercise has been assigned a HEFCE contact. This contact will be able to provide guidance during the response process. Details of the contact will be provided in a letter sent by Ben Grassby on 17 December 2004.

Action and implementation plans

58. Guidance for producing action and implementation plan(s) is given in Annex H. Example action and implementation plans are also included in Appendix 12.

Troubleshooting

59. Appendices 2, 5 and 8 will assist with the identification of the causes of discrepancies between the 2003-04 HESA student data and the HESES03, HESES03 based on cost centre sector norms for subjects and RAS03 comparisons respectively.

Supplementary data

60. Files can be accessed from the HEFCE extranet with details of how each student was classified in the re-creations, research degree rates of qualification and HEFCE published statistics. Details of how to access these files are in Annex G.

FAQs

61. Frequently asked questions (FAQs) for this exercise can be found on the HEFCE web-site under Learning & teaching/Data collection. We encourage institutions to refer to the FAQs for guidance in the first instance. We will only use our e-mail list of HESES or RAS contacts to notify institutions of significant changes or updates.

SAS code

62. We use the SAS programming language to generate all the derived statistics described in this publication. The SAS code we use to do this can be found on the HEFCE web-site under Learning & teaching/Data collection.

Comments

63. All institutions are invited to comment on any of the methods described in this publication. Comments should be sent to Ben Grassby.

Annexes and appendices

64. The annexes to this publication describe how we will use HESA data to monitor and allocate funding for higher education institutions. The web-only appendices to this publication contain technical descriptions of the algorithms we use to monitor and allocate funding. The appendices are on the HEFCE web-site with this document under Publications.

65. Notification of any grant adjustments will normally take approximately six weeks. Institutions will be given four weeks from notification of grant adjustments to submit any appeals for mitigation. Institutions will be informed of the outcome of any appeal and the final grant adjustments following consideration by the HEFCE chief executive.

Annex A

Summary of changes since HESA 2002-03 derived statistics (HEFCE 2004/10)

Purpose

1. This annex describes the changes that have been made to the monitoring of HESA returns and guidance since the release of 'HESA 2002-03 derived statistics for funding allocations and monitoring' (HEFCE 2004/10).
2. The name of the exercise has changed to '2003-04 statistics derived from HESA data for monitoring and allocation of funding' to emphasise the fact that HEFCE, not HESA, is responsible for the exercise.

New launch date

3. We have launched this exercise earlier than previously. This will allow institutions to respond throughout January and February rather than later in the year, when they have other conflicting commitments such as preparing for the HESA return.

Teaching and WP funding allocations

4. HESA 2003-04 data may be used to reassign sports science and leisure studies activity reported in HESES03 to price groups for the 2005-06 teaching funding allocation. The data will be made available separately early next year.
5. Due to the effects of introduction of the new tariff point system, the 2005-06 WP funding allocations are being reviewed in February 2005 and they are due to be published during March 2005.

Funding for research

6. A detailed explanation of the algorithms used to derive the funding for research using HESA 2003-04 student data has been added to the explanation of the RAS03 re-creation algorithms (Appendix 7).

RAS subject to UoA mapping

7. Additional information is provided in Appendix 7 to explain the methodology used to generate the mapping of UoAs to subjects.

Grant adjustment report

8. An explanation of the cover sheet has been included within the explanation of the algorithms used to derive the grant adjustment report, standard resource table and assumed fee income table in Appendix 1.

Additional cost centre information

9. Additional information is provided in Annex C to explain how the sector norm cost centre mapping was calculated.
10. A list of subjects, where we require explanation as to why the mapping of subject to cost centre differs from the sector norm mapping of subject to cost centre, is provided in the form of an action plan. This plan will be distributed to institutions asked to respond to the exercise on 17 December and will also be made available on the HEFCE extranet.
11. Details of the percentage of institutions mapping subjects to the sector norm cost centre are provided in an Excel file, on the HEFCE web-site under Learning & teaching/Data collection.

Submitting overrides to derived fields

12. A procedure for submitting override files has been set up that is similar to the procedure for submitting amendment files. This procedure is explained in Annex J.

Amendment check list

13. An amendment check list that contains the file name(s) of the amendment files that have been processed, the name of the HESA field(s) and the number of records that have been amended has been added to the zipped archive that is available from the HEFCE extranet. Details of how to access this file are provided in Annex G.

Re-ordering of fields within the individualised file

14. Fields within the individualised files have been re-ordered so that fields that are needed to re-create the tables are near the start of the file.

After 2005 we are committed to not changing the order.

Individualised files

15. For ease of use, we have divided the individualised file into separate files for each output generated by the re-creation. Each output's individualised files contain all the necessary data to generate the output. Details of the new files and how to access them can be found in Annex G.

Example action and implementation plans

16. Example action and implementation plans used to be provided on the HEFCE web-site under Learning & teaching/Data collection. These are now in Appendix 12.

Changes to algorithms

CDPRP, CMPRP, PCDRP, PCMPRP

17. These fields have been removed because they do not contain any additional information to the price group fields.

HESNHS

18. We have derived a new field that separately identifies the three different groups of students that are eligible for NHS bursaries. The HESNHS field is described in paragraph 35 of Appendix 1.

PRGA

19. The algorithm for assigning undergraduate clinical dentistry students to price group A has changed to include those students on the final four years (instead of final three and a third) of a registrable dental qualification. The algorithm is given in paragraphs 56-60 of Appendix 1.

HESEXCL

20. There has been a refinement where, for students taught wholly outside the UK, HESEXCL=16 where LOCSDY=7 and institution specific approval has not been given. The algorithm is given in paragraph 72 of Appendix 1.

HESCOMP

21. There has been a refinement to account for students who have not left but have failed to complete. The algorithm is given in paragraph 70 of Appendix 1.

PBLLEV

22. The algorithm for deriving PBLLEV has been revised to reflect the new population and incorporate further sub-degree levels of study. The algorithm is given in paragraph 26 of Appendix 11.

PBLSTU

23. The algorithm for deriving PBLSTU has been revised to reflect the new population and to take into account the revision of PBLLEV. The algorithm is given in paragraph 33 of Appendix 11.

PBLLOAD

24. The algorithm for deriving PBLLOAD has been revised to reflect the new population and to take into account the revision of PBLLEV. The algorithm is given in paragraph 34 of Appendix 11.

Annex B

Comparison of HESES03 and the HESES03 re-creation

Purpose

1. This annex details the process of making a response, where one is required, to the comparison of HESES03 and the HESES03 re-creation that has been generated from HESA 2003-04 student data. This annex also specifies the thresholds we have used to select institutions for response, based on discrepancies between their HESES03 and HESES03 re-creation. Where an institution's data leads to discrepancies that exceed any of the thresholds, we require a full response through an

action and implementation plan. Guidance for completion of an acceptable action and implementation plan is provided in Annex H.

HESES03 re-creation tables

2. The HESES03 re-creation tables and HESES03 tables can be accessed from the HEFCE extranet. Annex G describes how to access the Excel workbook (HESR03XXXX.xls-where XXXX denotes the HESA institution identifier). The workbook contains the following worksheets:

Page number	Worksheet*	Title
1	CoverSheet	Title page containing information about the thresholds and the difference in terms of contract range holdback, medical and dental holdback, funding conditional upon delivery of growth and undetermined completion status students between the HESES03 and the HESES03 re-creation
2	Summary	Summary comparison of HESES03 and the HESES03 re-creation
3	SummaryPG	Summary comparison of price group activity between HESES03 and the HESES03 re-creation
4	Excl	Students excluded from the HESES03 re-creation
5	FTS	HESES03 re-creation Table 1a: Full-time and sandwich years of programme of study
6	MED	HESES03 re-creation Table 1b: Medical and dental full-time and sandwich years of programme of study
7	SWOUT	HESES03 re-creation Table 2: Sandwich year-out years of programme of study
8	PT	HESES03 re-creation Table 3: Part-time years of programme of study and load
9	FEE	HESES03 re-creation Table 4: Home and EC fees
10	HBK	HESES03 re-creation grant adjustments
11	STD	HESES03 re-creation recalculation of standard resource
12	F03	HESES03 re-creation recalculation of assumed fee income
13	hFTS	HESES03 Table 1a: Full-time and sandwich years of programme of study
14	hMED	HESES03 Table 1b: Medical and dental full-time and sandwich years of programme of study
15	hSWOUT	HESES03 Table 2: Sandwich year-out years of programme of study
16	hPT	HESES03 Table 3: Part-time years of programme of study and load
17	hFEE	HESES03 Table 4: Home and EC fees
18	hHBK	HESES03 grant adjustments
19	hSTD	HESES03 recalculation of standard resource
20	hF03	HESES03 recalculation of assumed fee income

* For worksheet reference see spreadsheet tabs

3. The information contained in the HESES03 re-creation tables can be re-built from the individualised file which we provide. The file (HESR03XXXX.ind) contains details of how each student was classified in the re-creation. A full description of the data contained in the individualised file is given in Appendix 1.

Comparison

4. We derive a HESES03 re-creation, and hence an individualised file, by applying the algorithms detailed in Appendix 1 to HESA 2003-04 student data.

5. We compare the HESES03 re-creation to HESES03. This comparison takes place after the 2003-04 student data have been finalised with HESA.

6. We re-calculate a grant adjustment report for the HESES03 re-creation by applying the same formulae that were used to calculate the grant adjustment report for HESES03.

7. We select institutions to explain discrepancies between their HESES03 and HESES03 re-creation using a comparison of the grant adjustment reports derived from each return. Notwithstanding the thresholds, we may also ask for further information from any institution in respect of this comparison. This may result ultimately in adjustments to grant, where appropriate.

Selection of institutions required to respond

8. We will require a full, timely and detailed response from institutions where any of the following thresholds are exceeded:

- a. The difference between contract range holdback for HESES03 and the HESES03 re-creation exceeds £250,000, or 5 per cent of total recurrent teaching funding for 2003-04.
- b. The difference in any net grant adjustment relating to funding conditional upon delivery of growth between HESES03 and the HESES03 re-creation exceeds £150,000.

c. The difference between holdback for under-recruitment against the medical and dental contract FTE for HESES03 and the HESES03 re-creation exceeds £100,000.

d. More than 700 students are identified with undetermined completion status and there is a difference in total recurrent teaching funding exceeding £200,000, between HESES03 and the HESES03 re-creation.

9. In calculating the grant adjustment reports we have ignored any appeals for mitigation. Therefore, the grant adjustment report derived from HESES03 may differ from the final grant adjustment report notified for 2003-04. Before making adjustments to an institution's funding as a result of this exercise, we will take into account any previously agreed mitigation. We have adopted this approach to allow us to apply consistent monitoring procedures to all institutions, irrespective of individual circumstances that have affected previously announced funding allocations.

Action required

10. We will write to heads of institutions, copied to HESES contacts, on 17 December 2004 specifying whether a response is required.

11. Where we require a response, an action and implementation plan must be sent by **11 February 2005** to Ben Grassby, detailing how the institution will reconcile the two data sources. Guidance for submitting an action and implementation plan is included in Annex H. The final deadline for receipt by HEFCE of amendments to HESA data and overrides to derived fields is **29 March 2005**.

Action and implementation plan

12. Each institution required to make a response will be asked to provide an action and implementation plan. The plan should:

- a. Explain each constituent cause of difference between HESES03 and the HESES03 re-creation.

- b. Contain an estimate of the contribution that each constituent cause of difference makes to the discrepancy in terms of student numbers and FTEs and, where appropriate, contract range holdback, funds to be held back or funds due back.
- c. Detail all actions required to reconcile the two data sources, broken down by each cause of difference.
- d. Detail the system or process changes that will be implemented to ensure that similar problems do not recur.

13. If institutions do not provide satisfactory explanations for discrepancies, or do not respond according to the given timescales, we may carry out further investigations. This may include visits to institutions by us or our agents, in order to gain assurances concerning one or more of the following:

- the reliability of data returns
- the methodologies used to compile data returns
- the ability to respond in a full and timely manner to this exercise.

14. In order to gain these assurances we may need to collect or review data as part of these visits.

15. Paragraph 28a of the Financial Memorandum (HEFCE 2003/54) provides for the cost of such investigations to be deducted from institutions' grant.

16. We expect the explanations that institutions provide for discrepancies between the two data sources to fall into one or more of the following three categories:

- errors in HESA 2003-04 student data
- errors/estimation discrepancies in HESES03
- problems of fit with the HESES03 re-creation algorithms.

17. The action and implementation plan must specify where, and to what extent, each of these categories contributes to the overall discrepancy.

Errors in HESA data

18. If we find, either through reconciliations with HESA data, or any data audit, that the HESES submission does not reflect the final outturn position for the year, and that this has resulted in institutions receiving incorrect funding allocations, the HESES03 re-creation will supersede HESES03, and any consequent grant adjustments will be made (subject to the appeals process and the availability of our funds). Therefore it may be necessary for an institution to submit amendments to its HESA data to ensure they reasonably reflect the outturn position for 2003-04. See paragraphs 20-21 below for details about how to submit amendments to HESA data.

19. The procedures for the quality assurance of HESA data must take place before an institution signs off the HESA data as correct. Any amendments accepted after this point to recalculate funding must be seen as exceptional, and not as a part of quality assurance procedures.

20. Where errors are found in HESA data we require institutions to submit amendments. Institutions are strongly encouraged to submit amendments before the deadline of **29 March 2005** in order to ensure that, if required, any additional amendments can be submitted within this timeframe.

21. Amendments must follow the specification described in Annex I. This is essential in order to establish an audit trail of data changes, and to ensure that amendments are processed in a timely and accurate manner.

22. While we recognise that HESA returns are necessarily complicated, and that errors may occur in these returns, we expect that use by institutions of the HEFCE web facility for 2003-04 statistics derived from HESA data will keep the number of amendments to a minimum. (See HEFCE 2004/29.)

23. We may carry out a further investigation where amendments contradict our understanding of the broad characteristics of activity at an institution.

24. Amendments to HESA data will be incorporated in future HEFCE statistical publications and analyses. We provide HESA with all amendments to HESA data.

Errors/estimation discrepancies in HESES03 data

25. If we find, either through reconciliations with HESA data, or any data audit, that the HESES submission does not reflect the final outturn position for the year, and this is due to errors/estimation discrepancies in the HESES return, then the HESES03 re-creation will supersede HESES03, and any consequent grant adjustments will be made (subject to the appeals process and the availability of our funds). Therefore it will not be necessary for institutions to submit corrections to their HESES03.

Problems of fit with the HESES03 re-creation algorithms

26. We do not expect that problems of fit with the HEFCE algorithms will fully explain the discrepancies to which institutions are required to respond. However, where a problem of fit between our algorithms and HESES03 definitions contributes to a discrepancy, evidence will be required of where the problem occurs, and its impact, with details on the action and implementation plan. Appendix 3 details all known problems of fit with the HESES03 re-creation algorithms.

27. Where problems of fit are identified we require institutions to submit an override file. Institutions are strongly encouraged to submit overrides prior to the deadline of **29 March 2005** in order to ensure that, if required, any additional overrides and amendments can be submitted within this timeframe.

28. Overrides must follow the specification described in Annex J. This is essential in order to establish an audit trail of data changes, and to ensure that overrides are applied in a timely and accurate manner.

Further action

29. Amendments to HESA data and overrides to derived fields will be used to update the HESES03 re-creation. Once all amendments/overrides have been processed, and we are content that all discrepancies between the HESES03 return and the re-creation have been reasonably explained, we will ask the institution to confirm:

- that the HESES03 re-creation reasonably reflects the outturn position for 2003-04
- the accuracy of the amendments to HESA data.

30. A confirmation form will be sent to institutions that have made HESA amendments asking them to confirm that the amendments in the data file(s) are correct, and that the HESES03 re-creation reasonably reflects the outturn position for 2003-04. If no amendments to HESA data have been made, institutions should provide a letter confirming that the HESES03 re-creation reasonably reflects the outturn position for 2003-04.

31. If, after processing all amendments/overrides, we are not content that all discrepancies between the HESES03 return and the HESES03 re-creation have been reasonably explained, we will ask the institution to submit a further action and implementation plan to explain any remaining discrepancies between the two data sources. It is likely that we will visit institutions to discuss remaining discrepancies.

32. Once confirmation has been asked for and received we will generate both a RAS03 re-creation and a HESES03 re-creation based on cost centre sector norms for subjects, to incorporate any amendments that have been made to HESA 2003-04 student data. We will request a further response where the selection thresholds for the comparisons of HESA 2003-04 student data with either RAS03 or the HESES03 re-creation based on cost centre sector norms for subjects are exceeded.

Guidance

HEFCE contact

33. Each institution required to make a response to this exercise has been assigned a HEFCE contact. This contact will be able to provide guidance during the response process. We will provide information to institutions about their contact in a letter that we will send on 17 December 2004.

Action and implementation plan

34. Guidance for producing an action and implementation plan is given in Annex H. An example action and implementation plan can also be found in Appendix 12.

Troubleshooting

35. Appendix 2 will assist with identifying the causes of discrepancies between HESES03 and the HESES03 re-creation.

FAQs

36. FAQs for this exercise can be found on the HEFCE web-site under Learning & teaching/Data collection. We encourage institutions to refer to the FAQs for guidance in the first instance. We will only use our e-mail list of HESES contacts to notify institutions of significant changes or updates.

SAS code

37. We use the SAS programming language to generate the HESES03 re-creation. The SAS code we use to do this is on the HEFCE web-site under Learning & teaching/Data collection.

Comments

38. All institutions are invited to comment on the algorithms described in Appendix 1, and to suggest how they can be improved. Comments should be sent to Ben Grassby.

Deadline for responses

39. Action and implementation plans must arrive no later than **11 February 2005** and be sent to:

Ben Grassby
HEFCE
Northavon House
Analytical Services Group
Coldharbour Lane
BRISTOL
BS16 1QD

40. The final deadline for receipt of amendments to HESA data and overrides to derived fields detailed in the action and implementation plan is **29 March 2005**.

Annex C

Comparison of the HESES03 re-creation and the HESES03 re-creation based on cost centre sector norms for subjects

Purpose

1. This annex details the process of making a response, where one is required, to the comparison of a re-creation of HESES03 generated from HESA 2003-04 student data, and a re-creation of HESES03 that is generated from HESA 2003-04 student data using the assignment of activity to cost centres based on sector norms for subjects (HESES03 re-creation based on sector norms for subjects) for subjects where the total student FTE assigned across the subject is greater than 100.
2. This annex also specifies the threshold we have used to select institutions required to make a response to the exercise based upon discrepancies between their cost centre assignments and those calculated to be the sector norms.
3. Where an institution's return of cost centres on the HESA 2003-04 student data exceeds the threshold, we require a full response to be made through an action and implementation plan. This should explain discrepancies between the HESES03 re-creation and the HESES03 re-creation based on cost centre norms for subjects, and also between HESES03 and the HESES03 re-creation. Guidance

for completion of acceptable action and implementation plans is provided in Annex H.

Background

4. Our guidance is that student load should be allocated to cost centres based on the cost centre of the member of staff most directly associated with it. In general, academic staff should be assigned to the cost centre that best represents the majority of their academic activity.
5. Further guidance on assigning academic departments to cost centres is contained within 'Assigning departments to academic cost centres: 2003-04' (HEFCE Circular Letter 25/2003). This can be downloaded from the HEFCE web-site under Publications.

Tables for HESES03 re-creation based on cost centre sector norms for subjects

6. The tables for the HESES03 re-creation based on cost centre sector norms for subjects can be accessed from the HEFCE extranet. Annex G describes how to access the Excel workbook (SNCC03XXXX.xls-where XXXX denotes the HESA institution identifier). The workbook contains the following worksheets:

Page number	Worksheet*	Title
1	CoverSheet	Title page containing information about the thresholds and the difference in terms of contract range holdback between the HESES03 re-creation and the HESES03 re-creation based on cost centre sector norms for subjects
2	snSummary	Summary comparison of HESES03 re-creation and the HESES03 re-creation based on cost centre sector norms for subjects
3	snSummaryPG	Summary comparison of price group activity between HESES03 re-creation and the HESES03 re-creation based on cost centre sector norms for subjects
4	snExcl	Students excluded from the HESES03 re-creation based on cost centre sector norms for subjects
5	snFTS	HESES03 re-creation based on cost centre sector norms for subjects Table 1a: Full-time and sandwich years of programme of study
6	snMED	HESES03 re-creation based on cost centre sector norms for subjects Table 1b: Medical and dental full-time and sandwich years of programme of study
7	snSWOUT	HESES03 re-creation based on cost centre sector norms for subjects Table 2: Sandwich year-out years of programme of study
8	snPT	HESES03 re-creation based on cost centre sector norms for subjects Table 3:

		Part-time years of programme of study
9	snFEE	HESES03 re-creation based on cost centre sector norms for subjects Table 4: Home and EC fees
10	snHBK	HESES03 re-creation based on cost centre sector norms for subjects grant adjustments
11	snSTD	HESES03 re-creation based on cost centre sector norms for subjects recalculated standard resource
12	snF03	HESES03 re-creation based on cost centre sector norms for subjects recalculated assumed fee income
13	FTS	HESES03 re-creation Table 1a: Full-time and sandwich years of programme of study
14	MED	HESES03 re-creation Table 1b: Medical and dental full-time and sandwich years of programme of study
15	SWOUT	HESES03 re-creation Table 2: Sandwich year-out years of programme of study
16	PT	HESES03 re-creation Table 3: Part-time years of programme of study
17	FEE	HESES03 re-creation Table 4: Home and EC fees
18	HBK	HESES03 re-creation grant adjustments
19	STD	HESES03 re-creation recalculated standard resource
20	F03	HESES03 re-creation recalculated assumed fee income

*For worksheet reference see spreadsheet tabs

7. The information in the tables for the HESES03 re-creation based on cost centre sector norms for subjects can be re-built from the individualised file which we provide. The file (SNCC03XXXX.ind) contains details of how each student was classified in the re-creation. A full description of the data in the individualised file is given in Appendix 4.

Sector norm cost centre assignments

8. We generated a 'sector norm' mapping of subject activity to cost centres using HESA 2003-04 student data. To do this, we took the cost centre which most institutions map the subject to. This was calculated as follows:

- The FTE of each subject was calculated.
- The FTE of each subject greater than 50 was selected.
- The cost centre mapped to each subject in each institution that has the largest FTE was selected.
- The number of institutions that share the same mapping of cost centre for each subject was

counted.

- The cost centre for each subject that the largest number of institutions share was chosen.
- Details of the percentage of institutions mapping subjects to the sector norm cost centre are provided in an Excel file on the HEFCE web-site under Learning & teaching/Data collection.

Comparison

10. We derive a HESES03 re-creation, and hence the individualised file, by applying the algorithms detailed in Appendix 1 to HESA 2003-04 student data.

11. We also derive a HESES03 re-creation based on cost centre sector norms for subjects. The HESES03 re-creation based on cost centre sector norms for subjects is generated by applying the algorithms detailed in Appendix 1 to HESA 2003-04 student data, except that cost centres, and consequently price groups, are assigned using the sector norm cost centre mapping in Appendix 4.

12. We calculate a grant adjustment report for the HESES03 re-creation and the HESES03 re-creation based on cost centre sector norms for subjects by

applying the same formulae that were used to calculate the grant adjustment report for HESES03.

13. We select institutions to explain discrepancies between their cost centre assignments and the sector norm cost centre assignments for subjects using a comparison of the grant adjustment reports derived from each return. Notwithstanding the thresholds, we may also ask for further information from any institution in respect of any part of this comparison. This may result ultimately in adjustments to grant where appropriate.

14. During this comparison, we will incorporate previous decisions by HEFCE regarding the mapping at individual institutions of subject activity to cost centres.

Selection of institutions required to respond

15. We will require a full, timely and detailed response from institutions where the following threshold is exceeded:

- a. The difference in holdback for exceeding the contract range between the HESES03 re-creation and the HESES03 re-creation, based on sector norm assignments of subject activity to price groups, exceeds £700,000.

16. In calculating the grant adjustment reports we have ignored any appeals for mitigation. Before making adjustments to an institution's funding as a result of this exercise, we will take into account any previously agreed mitigation. We have adopted this approach to allow us to apply consistent monitoring procedures to all institutions, irrespective of individual circumstances that have affected previously announced funding allocations.

Action required

17. We will write to heads of institutions, copied to HESES contacts, on 17 December 2004 specifying whether a response is required.

18. Where we require a response, action and implementation plan(s) must be sent by **11 February 2005** to Ben Grassby detailing how the institution will reconcile the data sources.

Guidance for submitting action and implementation plans is included in Annex H. The final deadline for receipt by HEFCE of amendments to HESA data and overrides to derived fields is **29 March 2005**.

Action and implementation plan

19. Each institution required to make a response will be asked to fill in an action plan, which we will provide on 17 December and will make available on the HEFCE extranet. This will list the subjects that are assigned to a cost centre that differs from the sector norm cost centre, for which we require an explanation.

20. Institutions will also be asked to provide an action and implementation plan to explain the differences in HESES and the HESES re-creation comparison. The plans should:

- a. Explain each constituent cause of difference between the HESES03 re-creation and the HESES03 re-creation based on cost centre sector norms for subjects: that is, explain the subjects that are listed on the action plan that are assigned to a cost centre that differs from the sector norm cost centre.
- b. Explain each constituent cause of difference between HESES03 and the HESES03 re-creation.
- c. Contain an estimate of the contribution that each constituent cause of difference between the HESES03 re-creation and the HESES03 re-creation based on cost centre sector norms for subjects makes to the discrepancy – in terms of student numbers and FTEs, and staff numbers and FTEs.
- d. Contain an estimation of the contribution that each constituent cause of difference between HESES03 and the HESES03 re-creation makes to the discrepancy in terms of student numbers and FTEs and, where appropriate, contract range holdback, funds to be held back or funds due back.
- e. Detail all actions required to reconcile the HESES03 re-creation and the HESES03

re-creation based on cost centre sector norms for subjects.

- f. Detail all actions required to reconcile HESES03 and the HESES03 re-creation broken down by each cause of difference.
 - g. Detail the system or process changes that will be implemented to ensure that similar discrepancies do not recur.
21. The subjects listed on the action plan will not include:
- a. Subjects that are mapped to cost centres whose price group does not differ from the price group for the sector norm; or
 - b. Subjects where the total student FTE assigned across the subject is less than 100.
22. If the staff FTE is less than 20 we also do not require further explanation. If the staff FTE is less than 20 please state this clearly on the action plan. See Appendix 6 for further details about the problems of fit.
23. It must be noted that differences between HESES03 and the HESES03 re-creation must be explained, in addition to differences between the HESES03 re-creation and the HESES03 re-creation based on cost centre sector norms for subjects, regardless of whether the thresholds described in paragraph 8 of Annex B have been exceeded.
24. If institutions do not provide satisfactory explanations for discrepancies, or do not respond within the given timescales, we may carry out further investigations. This may include visits to institutions by us or our agents, in order to gain assurances concerning one or more of the following:
- the reliability of data returns
 - the methodologies used to compile data returns
 - the ability to respond in a full and timely manner to this exercise.
25. In order to gain these assurances we may need to collect or review data as part of these visits.
26. Paragraph 28a of the Financial Memorandum (HEFCE 2003/54) provides for the cost of such

investigations to be deducted from institutions' grant.

27. We expect explanations provided by institutions for discrepancies between the HESES03 re-creation and the HESES03 re-creation based on cost centre sector norms for subjects to fall into one or more of the following six categories:

- errors in HESA 2003-04 student data
- errors/estimation discrepancies in HESES03
- errors in the assignment of staff to academic departments
- errors in the assignment of academic departments to cost centres
- problems of fit with the HESES03 re-creation algorithms
- problems of fit with the HEFCE mapping of cost centre sector norms for subjects.

28. The action and implementation plan must specify where, and to what extent, each of the categories in paragraph 27 contributes to the overall discrepancy between the HESES03 re-creation and the HESES03 re-creation based on cost centre sector norms for subjects. The action and implementation plan must also specify where, and to what extent, each of the categories in paragraph 16 of Annex B contributes to the overall discrepancy between HESES03 and the HESES03 re-creation.

Errors in HESA data

29. The procedures for the quality assurance of HESA data must take place before an institution signs off the HESA data as correct. Any amendments accepted after this point to recalculate funding must be seen as exceptional, and not as a part of quality assurance procedures.

30. Where errors are found in HESA data we require institutions to submit amendments. Institutions are encouraged to submit amendments before the deadline of **29 March 2005**, in order to ensure that, if required, any additional amendments are submitted within this timeframe.

31. Amendments must follow the specification described in Annex I. This is essential in order to establish an audit trail of data changes, and to ensure that amendments are processed in a timely and accurate manner.

32. While we recognise that HESA returns are necessarily complicated, and that errors may occur in these returns, we expect that use by institutions of the HEFCE web facility for 2003-04 statistics derived from HESA data will keep the number of amendments to a minimum. (See HEFCE 2004/29.)

33. We may also carry out a further investigation where amendments contradict our understanding of the broad characteristics of activity at an institution.

34. Amendments to HESA student data will be incorporated in future HEFCE statistical publications and analyses. We provide HESA with all amendments to HESA data.

Errors in the assignment of staff to academic departments

35. Where errors are found in the assignment of staff to academic departments, amendments should be made to HESA student cost centre data for activity taught by the relevant staff. The HESES03 re-creation will supersede HESES03, and any consequent grant adjustments will be made (subject to the appeals process and the availability of our funds). Therefore it may be necessary for an institution to submit amendments to its HESA data to ensure they reasonably reflect the outturn position for 2003-04. See paragraphs 30 and 31 for details about how to submit amendments to HESA data.

Errors in the assignment of academic departments to cost centres

36. Where errors are found in the assignment of academic departments to cost centres, amendments should be made to HESA student cost centre data for activity taught by the staff in the academic department. The HESES03 re-creation will supersede HESES03, and any consequent grant adjustments will be made (subject to the appeals process and the availability of our funds).

Therefore it may be necessary for an institution to submit amendments to its HESA data to ensure they reasonably reflect the outturn position for 2003-04. See paragraphs 30 and 31 for details about how to submit amendments to HESA data.

Problems of fit with the HESES03 re-creation algorithms

37. We do not expect that problems of fit with the HEFCE algorithms will fully explain the discrepancies which institutions are required to respond to. However, where a problem of fit between our algorithms and HESES03 definitions contributes to a discrepancy, evidence will be required of where the problem occurs, and its impact, with details on the action and implementation plan. Appendix 3 lists all known problems of fit with the HESES03 re-creation algorithms.

38. Where problems of fit are identified we require institutions to submit an override file. Institutions are strongly encouraged to submit overrides before the deadline of **29 March 2005**, in order to ensure that, if required, any additional overrides and amendments can be submitted within this timeframe.

39. Overrides must follow the specification described in Annex J. This is essential in order to establish an audit trail of data changes, and to ensure that overrides are applied in a timely and accurate manner.

Problems of fit with the HEFCE mapping of cost centre sector norms for subjects

40. Where problems of fit are identified between the assignment of subject activity to cost centres and our sector norm mapping of subject activity to cost centres, once we have been informed and agree to the problem of fit, we will construct an override file. This will be implemented when the institution has approved the file.

Further action

41. Amendments to HESA data and overrides to derived fields will be used to update the HESES03 re-creation. Once all amendments/overrides have been processed and we are content that all discrepancies between HESES03, the HESES03 re-creation and the HESES03 re-creation based on cost centre sector norms for subjects have been reasonably explained, we will ask the institution to confirm:

- that the HESES03 re-creation reasonably reflects the outturn position for 2003-04
- the accuracy of the amendments to HESA data.

42. A confirmation form will be sent to institutions that have made HESA amendments, asking them to confirm that the amendments contained in the data file(s) are correct and that the HESES03 re-creation reasonably reflects the outturn position for 2003-04. If no amendments to HESA data have been made, institutions should provide a letter confirming that the HESES03 re-creation reasonably reflects the outturn position for 2003-04.

43. If after processing all amendments/overrides we are not content that all discrepancies between HESES03, the HESES03 re-creation and the HESES03 re-creation based on cost centre sector norms for subjects have been reasonably explained, we will ask the institution to submit further action and implementation plan(s) to explain any remaining discrepancies between the data sources. It is likely that we will visit institutions to discuss any remaining discrepancies.

44. Once confirmation has been asked for and received we will generate a RAS03 re-creation to incorporate any amendments made to HESA 2003-04 student data. We will request a response where the selection thresholds for the comparison of HESA 2003-04 student data with RAS03 are exceeded.

Guidance

HEFCE contact

45. Each institution required to make a response to this exercise has been assigned a HEFCE contact. This contact will be able to provide guidance during the response process. We will provide information to institutions about their contact in a letter that we will send on 17 December 2004.

Action and implementation plan

46. Guidance for producing action and implementation plan(s) is given in Annex H. Example action and implementation plans can also be found in Appendix 12.

FAQs

47. FAQs for this exercise can be found on the HEFCE web-site under Learning & teaching/Data collection. We encourage institutions to refer to the FAQs for guidance in the first instance. We will only use our e-mail list of HESES contacts to notify institutions of significant changes or updates.

SAS code

48. We use the SAS programming language to generate the HESES03 re-creation and the HESES03 re-creation based on cost centre sector norms for subjects. The SAS code we use to do this can be found on the HEFCE web-site under Learning & teaching/Data collection.

Comments

49. All institutions are invited to comment on the methods described in Appendix 4, and to suggest how they can be improved. Comments should be sent to Ben Grassby.

Deadline for responses

50. Action and implementation plans must arrive no later than **11 February 2005** and be sent to:

Ben Grassby
Analytical Services Group
HEFCE
Northavon House
Coldharbour Lane
BRISTOL
BS16 1QD

51. The final deadline for receipt of amendments to HESA data and overrides to derived fields detailed in the action and implementation plan is **29 March 2005**.

Annex D

Comparison of RAS03 and the RAS03 re-creation

Purpose

1. This annex details the process of making a response, where one is required, to the comparison of RAS03 and the RAS03 re-creation that has been derived from HESA 2003-04 student data. This annex also specifies the threshold we have used to select institutions for response, based on discrepancies between their RAS03 and RAS03 re-creation. Where an institution's data leads to discrepancies which exceed the threshold, we require a full response through an action and implementation plan. Guidance for completing an

acceptable action and implementation plan is provided in Annex H.

HESES03 re-creation tables

2. The RAS03 re-creation tables and RAS03 tables can be accessed from the HEFCE extranet. Annex G describes how to access the Excel workbook (RASR03XXXX.xls-where XXXX denotes the HESA institution identifier). The workbook contains the following worksheets:

Page number	Worksheet*	Title
1	CoverSheet	Title page containing information about the thresholds and the difference in terms of the sum of the absolute difference between RAS03 and the RAS03 re-creation over the eight subject groups which exceed 300 FTEs compared to the initial position
2	R1a	RAS03 re-creation Form R1a: All full-time research students by year of programme
3	R1b	RAS03 re-creation Form R1b: All part-time research students by year of programme
5	R2a	RAS03 re-creation Form R2a: Fundable Home and EC fee paying full-time research students by year of programme
6	R2b	RAS03 re-creation Form R2b: Fundable Home and EC fee paying part-time research students by year of programme
8	Funds	RAS03 re-creation calculation of quality related research (QR) funding for 2004-05 using HESA 2003-04 student data and other HEFCE research funding data
10	Summary	Summary comparison of RAS03 and the RAS03 re-creation
11	Summary by subject groupings	Summary comparison of RAS03 and the RAS03 re-creation by broad subject group
12	rR1A	RAS03 Form R1a: All full-time research students by year of programme
13	rR1B	RAS03 Form R1b: All part-time research students by year of programme
15	rR2A	RAS03 Form R2a: Fundable Home and EC fee paying full-time research students by year of programme
16	rR2B	RAS03 Form R2b: Fundable Home and EC fee-paying part-time research students by year of programme
18	rFunds	RAS03 HEFCE data for quality-related (QR) funding for 2004-05 using RAS 2003 and other HEFCE research funding data table
20	Broad subject groups	Mapping of units of assessment to broad subject groups used for comparison tables

*For worksheet reference see spreadsheet tabs

3. The information contained in the RAS03 re-creation tables can be re-built from the individualised file which we provide. The file (RASR03XXXX.ind) contains details of how each student was classified in the re-creation. A full description of the data in the individualised file is given in Appendix 7.

Comparison

4. We derive a RAS03 re-creation of forms R1a, R1b, R2a and R2b, and hence the individualised file, by applying the algorithms detailed in Appendix 7 to the HESA 2003-04 student data.

5. We compare the RAS03 re-creation to RAS03. This comparison takes place after the 2003-04 student data have been finalised with HESA.

6. We re-calculate the 2004-05 research grant from the RAS03 re-creation by applying the same formulae that were used to calculate it from RAS03. Further information about research funding is provided in Appendix 7.

7. We assign each UoA to one of eight subject groups. The assignments of UoAs to subject groups are given in paragraph 3 of Appendix 7.

8. We select institutions to explain discrepancies between their RAS03 and RAS03 re-creation where there are significant differences in FTEs over the eight subject groups between RAS03 and the RAS03 re-creation. Notwithstanding the thresholds, we may also ask for further information from any institution in respect of this comparison.

Selection of institutions required to respond

9. We will require a full, timely and detailed response from institutions where the following threshold is exceeded:

The sum of the absolute difference between RAS03 and the RAS03 re-creation over the eight subject groups exceeds 300 FTEs.

Action required

10. We will write to heads of institutions, copied to RAS contacts, on 17 December 2004 specifying whether a response is required.

11. Where we require a response, an action and implementation plan must be sent by **11 February 2005** to Ben Grassby, detailing how the institution will reconcile the two data sources. Guidance for submitting an action and implementation plan is included in Annex H. The final deadline for receipt by HEFCE of amendments to HESA data and overrides to derived fields is **29 March 2005**.

Action and implementation plan

12. Each institution required to make a response will be asked to provide an action and implementation plan. The plan must:

- a. Explain each constituent cause of difference between RAS03 and the RAS03 re-creation.
- b. Estimate the contribution that each constituent cause of difference makes to the discrepancy.
- c. Detail all actions required to reconcile the two data sources broken down by each cause of difference.
- d. Detail how system or process changes will be implemented to ensure that similar problems do not recur.

13. If institutions do not provide satisfactory explanations for discrepancies, or do not respond within the given timescales, we may carry out further investigations. This may include visits to institutions by us or our agents, in order to gain assurances concerning one or more of the following:

- the reliability of data returns
- the methodologies used to compile data returns
- the ability to respond in a full and timely manner to this exercise.

14. In order to gain these assurances we may need to collect or review data as part of these visits.

15. Paragraph 28a of the Financial Memorandum (HEFCE 2003/54) provides for the cost of such investigations to be deducted from institutions' grant.

16. We expect explanations provided by institutions for discrepancies between the two data sources to fall into one or more of the following three categories:

- errors in HESA 2003-04 student data
- errors in RAS03
- problems of fit with the RAS03 re-creation algorithms.

17. The action and implementation plan must specify where, and to what extent, each of these categories contributes to the overall discrepancy.

Errors in HESA data

18. Where errors are found in RAS03, the RAS03 re-creation will supersede RAS03, and any consequent grant adjustments will be made (subject to the appeals process and the availability of our funds). Therefore it may be necessary for an institution to submit amendments to its HESA data to ensure they reasonably reflect the outturn position for 2003-04. See paragraphs 20 and 21 below for details about how to submit amendments to HESA data.

19. The procedures for the quality assurance of HESA data must take place before an institution signs off the HESA data as correct. Any amendments accepted after this point to recalculate funding must be seen as exceptional, and not as a part of quality assurance procedures.

20. Where errors are found in HESA data we require institutions to submit amendments. Institutions are strongly encouraged to submit amendments before the deadline of **29 March 2005**, in order to ensure that, if required, any additional amendments are submitted within this timeframe.

21. Amendments must follow the specification described in Annex I. This is essential in order to establish an audit trail of data changes, and to

ensure that amendments are processed in a timely and accurate manner.

22. We may carry out a further investigation where amendments contradict our understanding of the broad characteristics of activity at an institution.

23. Amendments to HESA data will be incorporated in future HEFCE statistical publications and analyses. We provide the HESA with all amendments to HESA data.

Errors in RAS03 data

24. Where errors are found in RAS03, the RAS03 re-creation will supersede RAS03, and any consequent grant adjustments will be made (subject to the appeals process and the availability of our funds). Therefore it will not be necessary for institutions to submit corrections to their RAS03.

Problems of fit with the RAS03 re-creation algorithms

25. We do not expect that problems of fit with the HEFCE algorithms will fully explain the discrepancies which institutions are required to respond to. However, where a problem of fit between our algorithms and RAS03 definitions contributes to a discrepancy, evidence will be required of where the problem occurs, and its impact, with details on the action and implementation plan. Appendix 9 gives all known problems of fit with the RAS03 re-creation algorithms.

26. Where problems of fit are identified we require institutions to submit an override file. Institutions are strongly encouraged to submit overrides before the deadline of **29 March 2005**, in order to ensure that, if required, any additional overrides and amendments can be submitted within this timeframe.

27. Overrides must follow the specification described in Annex J. This is essential in order to establish an audit trail of data changes, and to ensure that overrides are applied in a timely and accurate manner.

28. Where problems of fit are identified with the mapping of subjects to UoAs, once we have been informed and agree to the problem of fit, we will construct an override file. This will be implemented when the institution has approved the file.

Further action

29. Amendments to HESA data and overrides to derived fields will be used to update the RAS03 re-creation. Once all amendments/overrides have been processed and we are content that all discrepancies between the RAS03 return and the re-creation have been reasonably explained, we will ask the institution to confirm:

- that the RAS03 re-creation reasonably reflects the outturn position for 2003-04
- the accuracy of the amendments to HESA data.

30. A confirmation form will be sent to institutions that have made HESA amendments, asking them to confirm that the amendments in the data file(s) are correct and that the RAS03 re-creation reasonably reflects the outturn position for 2003-04. If no amendments to HESA data have been made, institutions should provide a letter confirming that the RAS03 re-creation reasonably reflects the outturn position for 2003-04.

31. If after processing all amendments/overrides we are not content that all discrepancies between the RAS03 return and the RAS03 re-creation have been reasonably explained, we will ask the institution to submit a further action and implementation plan to explain any remaining discrepancies between the two data sources.

32. Once confirmation has been asked for and received we will generate both a HESES03 re-creation and a HESES03 re-creation based on cost centre sector norms for subjects to incorporate any amendments made to HESA 2003-04 student data. We will request a further response where the threshold for the comparisons of HESA 2003-04 student data with either HESES03 or the HESES03 re-creation based on cost centre sector norms for subjects are exceeded.

Guidance

HEFCE contact

33. Hannah Wood is the assigned HEFCE contact for all institutions that are required to make a response to the comparison of RAS03 and the RAS03 re-creation exercise. She will provide guidance during the response process.

Action and implementation plan

34. Guidance for producing an action and implementation plan is given in Annex H. An example action and implementation plan is also provided in Appendix 12.

Troubleshooting

35. Appendix 8 will assist with identifying the causes of discrepancies between RAS03 and the RAS03 re-creation.

FAQs

36. FAQs for this exercise can be found on the HEFCE web-site under Learning & teaching/Data collection. We encourage institutions to refer to the FAQs for guidance in the first instance. We will only use our e mail list of RAS contacts to notify institutions of significant changes or updates.

SAS code

37. We use the SAS programming language to generate the RAS03 re-creation. The SAS code we use to do this can be found on the HEFCE web-site under Learning & teaching/Data collection.

Comments

38. All institutions are invited to comment on the algorithms described in Appendix 7, and to suggest how they can be improved. Comments should be sent to Ben Grassby.

Deadline for responses

39. Action and implementation plans must arrive no later than **11 February 2005** and be sent to:

Ben Grassby
HEFCE
Northavon House
Analytical Services Group
Coldharbour Lane
BRISTOL
BS16 1QD

40. The final deadline for receipt of amendments to HESA data and overrides to derived fields detailed in the action and implementation plan is **29 March 2005**.

Annex E

Research degree rates of qualification

Purpose

1. This annex provides details of how we will use HESA student data to measure and report information relating to the rates of qualification obtained by students undertaking research degree programmes. In addition we will produce context statistics on the number of students that are active but have not qualified.
2. We are committed to improving the standards of the postgraduate research degree programmes which we support. As part of this work, we will monitor at institutional level the proportion of postgraduate research students qualifying within a particular time period.
3. Further information about our requirements for postgraduate research degree programmes is given in 'Postgraduate research degree programmes: minimum standards and funding' (HEFCE Circular Letter 18/2004).
4. The approach detailed below relies on the ability to follow cohorts of students over a number of years. We intend to use the HUSID x INSTID pair to track students between years. It is critical therefore that HUSIDs are maintained between years.
5. The production of the derived statistics offers institutions the opportunity to check these data at an early stage to ensure they are fit for purpose in the future.
6. As well as checking data as part of this proposed approach, institutions are reminded that HESA provides extensive HUSID, INSTID, NUMHUS (HIN) checking documentation as part of its annual data collection cycle, and institutions are encouraged to use this during data collection.

Research degree rates of qualification tables

7. The research degree rates of qualification data can be accessed from the HEFCE extranet. Annex G describes how to access the Excel workbook (RDQR03XXXX.xls – where XXXX denotes the HESA institution identifier).

8. The information contained in the research degree rates of qualification tables can be re-built from the individualised file which we provide. The file (RDQR03XXXX.ind) contains details of how each student was classified in the re-creation. A full description of the data contained in the individualised file is given in Appendix 10.

Research degree rates of qualification population

9. For each year, starting with 1995-96, we have identified the population of students that started aiming for a postgraduate research degree (MPhil, PhD and comparable qualifications). Each population has been tracked across academic years, up to 2003-04, in order to derive cumulative qualification rates across time up to 2003-04. We removed students that left in the first year of the qualification aim. Students who die are excluded in the year that they die.
10. For each year the numbers and rates have been grouped into the following categories:
 - full-time expected to qualify in less than or equal to three years
 - full-time expected to qualify in greater than three years
 - part-time students
 - students that switched modes during the qualification aim.
11. The research degree rate of qualification was calculated by dividing the number of qualifiers by the number of PhD students. The activity statistic was calculated by dividing the number of active PhD students by the number of PhD students that are not qualifiers.

PhD students

12. PhD students, for this purpose, are defined as students in the population that started aiming for a research degree in each given year, who have been on a doctoral programme (not MPhil or other short research-based masters degree) at the point of calculating the rate.

Qualifiers

13. Qualifiers are defined as those PhD students who were awarded a PhD or similar research degree (not MPhil or other short research-based masters degree) at the point of calculating the rate of qualification.

Active PhD students

14. Active PhD students are defined as those PhD students who were not awarded a PhD or similar research degree, and who were actively pursuing studies at the point of calculating the activity statistic in the academic year.

Annex F

HEFCE statistical publications

Purpose

1. This annex describes the data we intend to publish or use to inform policy decisions, during 2005, generated from the HESA 2003-04 student data.
2. We will generate data from HESA 2003-04 student data covering four broad areas:
 - students registered at one institution and taught by another institution (also referred to as 'franchised' students)
 - campus data
 - distance learning data
 - provision by location.

These data may be used for regional analysis and to inform policy decisions.

3. Full details of the algorithms used in generating the data for publication are given in Appendix 11. Where appropriate we have included derived fields on the individualised student data file, to allow institutions to identify and better understand our derivations made from the data.

Publication tables

4. The publication data can be accessed from the HEFCE extranet. Annex G describes how to access the Excel workbook (PUB03XXXX.xls-where XXXX denotes the HESA institution identifier).
5. The information contained in the published tables can be re-built from the individualised file which we provide. The file (PUB03XXXX.ind) contains details of how each student was classified in the re-creation. A full description of the data contained in the individualised file is given in Appendix 11.

Students registered at one institution and taught by another institution

6. In HEFCE 2004/36 we published details of students where we have derived that they are registered at one institution and taught by another institution, using HESA 2002-03 student data. We intend to generate the equivalent data for 2003-04 to enable institutions to check that it is accurate

and suitable for publication. We expect to publish this information during 2005.

7. Where students are registered at one institution and taught by more than two institutions, only two teaching institutions, other than the registering institution, are shown in the tables.

Campus data

8. Campus data returned on the HESA 2003-04 student record will be used in regional analysis. We are providing summary campus information to enable institutions to check that it is accurate and suitable for publication. In particular, institutions must ensure campuses that are a significant distance apart are separately identified on the HESA student record. Campuses are defined as being a significant distance apart if they are in different cities or are over 25 miles apart. We encourage institutions to identify separate campuses wherever they are geographically distinct. Where institutions have merged since 31 July 2002 we have created a separate entry for each campus of the merged institution. We would expect institutions to continue to make this distinction in future years.

Distance learning data

9. We intend to publish data on distance learning in the 'Learning in the region' page of the annual HEFCE regional profiles publication. This page shows the extent to which the number of students registered at institutions in the region differs from those that are learning in the region. We assume that distance learners study in their region of domicile. We wish to publish both the numbers registered in the region but domiciled outside it, and those that are registered outside the region but domiciled inside it.

Provision by location

10. A database of provision by location, derived from the HESA student record and the LSC's individualised student record (ISR), was used for regional analyses. It is likely that we will continue to use similar data in the future. We expect to publish this information during 2005. Therefore,

it is essential that we have accurate data about where students are taught.

11. The location where students are taught is derived from either the data described in paragraphs 6 and 7 above or the campus at which the student is taught, as described in paragraph 8 above. Where neither campus information, nor information about students that are registered at one institution and taught by another, is given, students are assumed to be taught at the institution's main campus.

12. The following geographical categories based on the teaching institution's postcode were included in the database: region, county, county/unitary authority, local education authority and the LSC area.

13. Students who study via distance learning are not included in the database of provision by location.

14. The database also contains details of the students' mode and level of study as well as broadly grouped subjects of study.

Errors in HESA data

15. The procedures for the quality assurance of HESA data must take place before an institution signs off the HESA data as correct. Any amendments accepted after this point to recalculate funding must be seen as exceptional, and not as part of quality assurance procedures.

16. If institutions wish to correct the HESA 2003-04 student data which affect HEFCE statistical publications they must submit amendments by **29 March 2005** to Lisa Readdy.

17. Amendments must follow the specification described in Annex I. It is essential that amendments are processed in a timely and accurate manner.

18. We may carry out a further investigation where amendments contradict our understanding of the characteristics of activity at an institution.

19. Amendments to HESA data will be incorporated in future HEFCE statistical

publications and analysis. We provide HESA with all amendments to HESA data.

Deadline for responses

20. If institutions are concerned about the suitability for publication of data on franchised students, campuses, distance learning, or provision by location, they should contact:

Lisa Readdy
HEFCE
Northavon House
Analytical Services Group
Coldharbour Lane
BRISTOL
BS16 1QD

by **29 March 2005**.

Annex G

Obtaining data from the HEFCE extranet

1. Data files should be accessed from the HEFCE extranet at <https://extranet.hefce.ac.uk>.
2. Where amendments/overrides are made to individualised data, the version of the re-creation held on our extranet will be overwritten with an updated version once these amendments/overrides have been incorporated. Therefore, if institutions wish to retain intermediate versions of the re-creation, they will need to make copies of the files on their own system.

Registering a new account

3. New users of the HEFCE extranet will first need to register an e-mail address and extranet password. This can be done by clicking the 'Register' link on the login screen. In order to register, you will require an 'organisation key' and a 'group key', details of which have been provided in the letter sent by Ben Grassby on 17 December 2004.
4. Once registered, you should be able to log in by entering the e-mail address you used during registration, and the password that you created.

Existing users of the extranet

5. If you have used the HEFCE extranet for other HEFCE returns, you will be required to log in and join the group for '2003-04 statistics derived from HESA data'. Follow the log-in procedure by entering your e-mail address and password. You will be directed to a page for 'HEFCE extranet – available resources'; under 'Applications', click 'Join a group'. Enter the group key supplied in Annex A of the letter entitled '2003-04 statistics derived from HESA data', sent on 17 December 2004, and select 'Join group'.
6. If you have registered in the past but your account has expired, you will be required to refresh your account using the organisation key referred to in paragraph 3.

Athens Single Sign On Account

7. You can also log in to the HEFCE extranet using Athens Single Sign On account (if this is available at your institution).
 - a. Follow the 'Log in via Athens SSO' link on the login page.
 - b. Log in to Athens as normal (if you have not already done so).
 - c. When Athens has authenticated you, your browser will be directed to the 'HEFCE extranet – available resources page' where institutions will have access to their output files.
 - d. You will be required to join the group for '2003-04 statistics derived from HESA data'. Under 'Applications', click 'Join a group'. Enter the group key supplied in Annex A of the letter titled '2003-04 statistics derived from HESA data', sent on 17 December 2004, and select 'Join group'.

Accessing the output files

8. After verifying the e-mail address and password, your browser will be directed to the 'HEFCE extranet – available resources' page where institutions will have access to their output files.
9. Under the 'Folders' heading there will be a link to '2003-04 statistics derived from HESA data'. If this link is not visible, it is possible that you do not have the appropriate access. To obtain this, you will need the appropriate group key (see paragraph 3 above for further details). Clicking on the link to '2003-04 statistics derived from HESA data' will cause a similar link to appear under the 'Files' heading. Click this second link to start the download of a zipped archive containing the following output files:
 - HESR03XXXX.xls. This is an Excel workbook containing the HESES03 re-creation tables.
 - SNCC03XXXX.xls. This is an Excel workbook containing the tables for the

HESES03 re-creation based on cost centre sector norms for subjects.

- RASR03XXXX.xls. This is an Excel workbook containing the RAS03 re-creation tables.
- RDQR03XXXX.xls. This is an Excel workbook containing the research degree rates of qualification.
- PUB03XXXX.xls. This is an Excel workbook containing the derived statistics that will be used in HEFCE statistical publications.
- Amendcheck.xls. This is an Excel worksheet containing the file names of amendment files that have been processed, the name of the HESA field(s) and the number of records that have been amended.
- HESR03XXXX.ind. This is a comma-separated file containing the derived fields that generate the HESES03 re-creation.
- SNCC03XXXX.ind. This is a comma-separated file containing the derived fields that generate the HESES03 re-creation based on cost centre sector norms for subjects.
- RASR03XXXX.ind. This is a comma-separated file containing the derived fields that generate the RAS03 re-creation.
- RDQR03XXXX.ind. This is a comma-separated file containing the derived fields that generate the research degree rates of qualification.
- PUB03XXXX.ind. This is a comma-separated file containing the derived fields that generate the derived statistics that will be used in HEFCE statistical publications.

In each case, XXXX is the institutional identifier.

10. For further information on zipped files, click on the 'online help' link, located above the login box, or on the right of the page when you have successfully logged in.

11. Institutions are reminded that the individualised data are covered by the Data Protection Act. In order for these data to be

accessible to someone, they need to have both the 'organisation key' and the appropriate 'group key' for the data. You must not pass these keys on to unauthorised personnel.

Using the individualised files

12. All records with a value in a specific field can be found in the following way:

- a. Open an individualised file, from the list in Microsoft Excel and click <File>, <Open>. You will need to specify 'All Files' in the 'Files of type' box before the individualised file will appear in the file listing. Once you have selected the file, the 'Text Import Wizard' will appear. Ensure that 'Delimited' is selected near the top of the window, then click 'Next'. On the next page, uncheck 'Tab' and check 'Comma'. Click 'Finish' to open the file.
- b. Select the row containing the field headings.
- c. Select <Filter> from the <Data> menu and then <Autofilter>.
- d. Click on the arrow in the column containing the data which you want to filter.
- e. Either select a specific value or select <Custom> to apply a comparison operator other than equality.
- f. To select records using multiple fields, repeat steps d and e for each field.

13. The tables are best viewed in Microsoft Excel. When you open the individualised file, some fields (such as HUSID) may look like '9.64E+08'. To obtain the 13-digit code as returned to HESA follow these steps:

- a. Highlight the column containing the field you wish to format.
- b. Click on <Format>, then <Cells>.
- c. In the <Number> tab, select the <Custom> category.
- d. In the <Type> text box, enter 13 zeros.
- e. Click <OK> to get back to the spreadsheet.

Annex H

Guidance for action and implementation plans

Purpose

1. All institutions required to respond to this exercise must submit action and implementation plan(s). This annex describes the information that we require within the action and implementation plan(s).
2. Example action and implementation plans can be found in Appendix 12.
3. An action plan will be provided on 17 December and will be made available on the HEFCE extranet for those institutions required to respond to the comparison of HESES03 and HESES03 based on cost centre sector norms for subjects. This will contain a list of subjects that are assigned to a cost centre that differs from the sector norm cost centre, for which we require an explanation. Templates for the HESES03 and the HESES03 re-creation comparison, and for RAS03 and RAS03 re-creation comparison, are contained in Figures 1 and 2 respectively.
4. The action and implementation plan(s) provide us with assurance of an institution's ability to identify and remedy areas and causes of discrepancy between two data sources. To do so, they must contain the following:
 - a. Identification of all areas of difference between the data sources. For the HESES03 re-creation and HESES03 re-creation based on cost centre sector norms for subjects, the areas of difference will be the differences in the mapping of subject to cost centre compared to the sector norm mapping from subject to cost centre.
 - b. Identification of causes of difference between the data sources which have resulted in the areas of difference. The causes must be fully considered and detailed appropriately.
 - c. A list of all HESA fields requiring amendments and derived fields requiring overrides for each area of difference (where appropriate).
 - d. A date for submitting amendments to HESA data or overrides to derived fields (where

appropriate), within the deadlines given for the comparison.

- e. Estimates of the contribution of the discrepancy to the overall discrepancy. The sum of each cause of difference must amount to the overall discrepancy between two data sources.
- f. An implementation plan of how changes to systems or processes will eliminate the likelihood of similar errors recurring in future returns. The implementation plan must specify the date by which each change will be made.

Action plan

Identification of the areas of difference

5. We advise institutions to use the troubleshooting guides in Appendices 2, 5 and 8, which will help identify the specific areas of difference. We expect institutions to exercise their own judgement to decide when small differences between the two data sources do not warrant inclusion within the action and implementation plan. However, institutions need to be aware that small differences may accumulate, and will reduce our confidence in the institution's ability to identify areas of discrepancy between two data sources.

Identification of the cause of difference

6. The action and implementation plan must fully identify the cause of each difference. The possible causes of discrepancy are described in detail below.

HESES03 and HESA 2003-04 student data

- a. Errors in HESA 2003-04 student data.
- b. Errors/estimation discrepancies in HESES03.
- c. Problems of fit with the HEFCE algorithms.

HESES03 re-creation and the HESES03 re-creation based on the cost centre sector norms for subjects

- a. Errors in HESA 2003-04 student data.
- b. Errors/estimation discrepancies in HESES03 data.
- c. Errors in the assignment of staff to academic departments.
- d. Errors in the assignment of academic departments to cost centres.
- f. Problems of fit with the HESES03 re-creation algorithms.
- g. Problems of fit with the HEFCE mapping of cost centre sector norms for subjects.

RAS03 and HESA 2003-04 student data

- a. Errors in HESA 2003-04 student data.
- b. Errors in RAS03.
- c. Problems of fit with the HEFCE algorithms.

HESA fields requiring amendment

7. Where errors in HESA data are the cause of a discrepancy between two data sources, a full list of the fields that require amendment must be identified by the institution and supplied within the action and implementation plan.

Derived fields requiring override

8. Where problems of fit with the HEFCE algorithms are the cause of a discrepancy between two data sources, a full list of the derived fields that require overrides must be identified by the institution and supplied within the action and implementation plan. For further information on which derived fields are affected by problems of fit, see Appendices 3, 6 and 9.

Date for submitting HESA amendments

9. If the cause of a discrepancy is erroneous HESA data, the action and implementation plan must indicate the date by which an amendment file to correct the data will be submitted. This file must be in the format and structure described in

Annex I. Typically the process of submitting and receiving feedback for an amendment file will take up to 10 working days. It is possible that more than one iteration of amendments will be necessary to entirely reconcile differences between data sources, therefore we strongly advise that amendment files are submitted early enough to ensure a full and adequate response can be made by the final deadline for receipt of amendments.

Date for submitting overrides

10. If the cause of a discrepancy is a problem of fit with the HEFCE algorithms, the action and implementation plan must indicate the date by which an override file will be submitted. This must be in the format and structure described in Annex J.

Estimate of contribution to discrepancy

11. The action and implementation plan must include an estimate of the contribution each cause of difference makes to the discrepancy in terms of student numbers and FTEs (and staff numbers and FTEs for the comparison of the HESES03 re-creation and the HESES03 re-creation based on cost centre sector norms for subjects), and, where appropriate, contract range holdback, funds to be held back or funds due back. Where the sum of the contributions to the discrepancy does not account for the whole discrepancy, our confidence in the institution's ability to identify areas of discrepancy between two data sources will be reduced.

Implementation plan

12. Errors in data usually arise from deficiencies in the systems and processes used for ensuring data quality and consistency on the institution's student record system. For each difference caused by erroneous data, an implementation plan describing a change to systems or processes to eliminate the likelihood of similar errors recurring must be included. Each implementation plan must also include a date by which the changes will be implemented.

Further action

13. Where an action and implementation plan does not fulfil our expectations, in that we are unable to gain assurance from it about the institution's ability to identify and remedy areas and causes of discrepancy, we will take appropriate action to seek this assurance from the institution.

Figure 1 **Action plan: HESES03 and the HESES03 re-creation**

Institution name:

HESA code:

Area(s), cause(s), contribution to discrepancy and date[†] for correction of differences

Reference number	Area of difference (eg Column 1 full-time undergraduates HEFCE-funded price group D)	Cause of difference (eg HESES, HESA algorithm)	HESA fields requiring amendment or requiring an override (eg 623 changes to FUNDCODE) [†]	Date for submitting HESA amendment or override [†]	Estimate of contribution to discrepancy		
					Funding adjustment (eg £300,000 funds due back)	Student numbers	FTE

[†] Where appropriate

Figure 2 **Action plan: RAS03 and the RAS03 re-creation**

Institution name:

HESA code:

Area(s), cause(s), contribution to discrepancy and date[†] for correction of differences

Reference number	Area of difference (eg full-time postgraduate research student totals on form R1A)	Cause of difference (eg RAS, HESA algorithm)	HESA fields requiring amendment or requiring an override (eg 60 changes to FUNDLEV) [†]	Date for submitting HESA amendment or override [†]	Estimate of contribution to discrepancy	
					Student numbers	FTE

[†] Where appropriate

Figure 3 **Action plan: HESES03 re-creation and the HESES03 re-creation based on cost centre sector norms for subjects**

Institution name:

HESA code:

Area(s), cause(s), contribution to discrepancy and date[†] for correction of differences

Reference number	Subject where the price group that the cost centre is assigned to differs from the price group of the sector norm cost centre	Cause of difference (eg HESA, staff*, academic departments [‡] , mapping)	HESA fields requiring amendment or requiring an override (eg COSTCN01-03) [†]	Date for submitting HESA amendment or override [†]	Estimate of contribution to discrepancy			
					Student numbers	Student FTE	Staff numbers	Staff FTE
	**				**			

* Assignment of staff to academic departments

‡ Assignment of academic departments to cost centres

† Where appropriate

** This information is provided in the action plan that we will distribute on 17 December 2004

Figure 4 **Implementation plan**

Institution name: _____

HESA code: _____

Area of difference: _____

Change to system or process: _____

Date of implementation: _____

Signed: _____

Name (please print): _____

Position in organisation: _____

Date: _____

Annex I

Submitting amendments to HESA data

Purpose

1. This annex details the data structure and format for amendments to HESA data. Institutions must only supply HESA amendment files using the file structure and format detailed within this annex.
2. There are three types of amendment file, which perform the following actions:
 - change file – corrects field values for records on the HESA return
 - addition file – adds records omitted from the HESA return
 - deletion file – deletes records incorrectly included on the HESA return.

Examples of these three types of amendment file can be found at the end of this annex (see Figures 5-7), with a summary of the information we require for each type of amendment file (see Table 2).

3. These specifications are necessary to ensure we can process amendments to HESA data in a timely and accurate manner. We will require institutions to re-submit amendment files that differ, either in structure or format, to the specifications detailed in this annex.

Format and naming

4. Amendments to HESA data must be sent as a comma-separated file via the HEFCE extranet or as an e-mail attachment sent to hesa_heses_stats@hefce.ac.uk. We will also accept amendments on a 3.5" floppy disk or CD-ROM. Amendment files must be given a file name in the form `vvvXXXXn.amd`, where:

- `vvv` is the amendment prefix – that is, `chg` for change files, `add` for addition files and `del` for deletion files
- `XXXX` is the HESA institution identifier for the institution
- `n` is a sequential number starting at 1.

Structure

5. The structure of an amendment file depends on the type of amendments being submitted. Details of the structure of change, addition and deletion files are given in paragraphs 6, 11 and 14 respectively. Each amendment file must only contain records for one HESA record type (combined, student or module).

Change file

6. This amendment type allows values of fields to be corrected at the individual record level in our copy of HESA data. Typically, change files correct a small number of fields which contain incorrect values.

7. The amendment prefix for a change file is `chg`. For example the first set of amendments in a change file must be submitted in the form `chgXXXX1.amd`, and a second set of different amendments must be submitted in the form `chgXXXX2.amd`, where `XXXX` is the HESA institution identifier for the institution.

8. Each record must contain complete data for all fields included in the amendment file, even if a particular field remains unchanged in some cases.

9. Only HESA fields should be included in amendment files. Where changes to derived fields are required, the underlying HESA fields must be changed. For example, if `HESLEVEL` is incorrect, `FUNDLEV` would need to be changed. For further information on which HESA fields to change to achieve a given effect, see Appendices 1 and 7.

10. In addition to the records being changed the file must contain a header in the following form:

line 1 – amendment reference in the form `chgXXXXn` where `XXXX` is the HESA institution identifier and `n` is a sequential number starting at 1; this will be the file name with the `.amd` file extension removed.

line 2 – creation date of the amendment file in the form `ddmmyyyy`. For example `01032005` for a file created on 1 March 2005.

line 3 – a brief description of the correction and the reference number of the area of difference on the action plan that it rectifies. For example ‘Changes to HESA fields DATELEFT and RSNLEAVE’ - reference number 2.

line 4 – this line must contain the word CHANGE.

line 5 – the HESA fields used to uniquely identify records on the amendment file, comma-separated. For example: RECID, HUSID, OWNSTU, OWNPSD, NUMHUS

line 6 – the HESA fields being changed, comma-separated. For example: DATELEFT, RSNLEAVE

line 7 – number of records contained in the amendment file, excluding the first 12 lines of header information.

line 8 – the HESA field used to calculate the file check sum (see paragraph 22 for an explanation of the file check sum).

line 9 – file check sum.

lines 10 and 11 – these lines can contain any notes the institution wishes to include.

line 12 – HESA fields included in the amendment file. These fields must appear in the same order as each row of the data and must be comma-separated on one line. For example: RECID,HUSID,OWNSTU,OWNPSD,NUMHUS,QUALAIM,DATELEFT,RSNLEAVE

line 13 – amendment data must begin on this line.

end of file – there must be a single blank line following the final record in the amendment file.

Addition file

11. This amendment type allows full records that were omitted from the initial HESA return to be added to our copy of HESA data. Each amendment file must only contain additional records for one record type, for example combined, student or module. In this case all HESA fields for that record type must be included in the amendment file even if they are blank.

12. The amendment prefix for an addition file is add. For example, the first set of amendments submitted in an addition file must be submitted in the form addXXXX1.amd, and a second set of different amendments must be submitted in the form addXXXX2.amd, where XXXX is the HESA institution identifier for the institution.

13. Institutions must return the file containing a header in the following form:

line 1 – amendment reference in the form amdXXXXn where XXXX is the HESA institution identifier and n is a sequential number starting at 1; this will be the same as the file name with the ‘.amd’ file extension removed.

line 2 – creation date of the amendment file in the form ddmmYYYY, for example 01032005 for a file created on 1 March 2005.

line 3 – a brief description of the additional records and the reference number of the area of difference on the action plan that it rectifies. For example ‘Nursing courses omitted from original HESA return – reference number 2’.

line 4 – this line must contain the word ADD.

line 5 – this line must be left blank.

line 6 – this line must be left blank.

line 7 – number of records contained in the amendment file, excluding the first 12 lines of header information.

line 8 – the HESA field used to calculate the file check sum (see paragraph 22 for an explanation of the file check sum).

line 9 – file check sum.

lines 10 and 11 – these lines can contain any notes the institution wishes to include.

line 12 – a list of all HESA fields in the HESA record type. These fields must appear in the same order as each row of the data and must be comma-separated.

line 13 – amendment data must begin on this line.

end of file – there must be a single blank line following the final record in the amendment file.

Deletion file

14. This amendment type allows records incorrectly submitted to HESA to be removed from our copy of HESA data.

15. The amendment prefix for a deletion file is del. For example, a first set of amendments submitted in a deletion file must be submitted in the form delXXXX1.amd, and a second set of different amendments must be submitted in the form delXXXX2.amd, where XXXX is the HESA institution identifier for the institution.

16. Institutions must return the file containing a header in the following form:

line 1 – amendment reference in the form delXXXXn where XXXX is the HESA institution identifier and n is a sequential number starting at 1; this will be the same as the file name with the ‘.amd’ file extension removed.

line 2 – creation date of the amendment file in the form ddmmYYYY, for example 01032005 for a file created on 1 March 2005.

line 3 – a brief description of the correction and the reference number of the area of difference on the action plan that it rectifies. For example ‘Non-credit bearing students incorrectly included in HESA data - reference number 6’.

line 4 – this line must contain the word DELETE.

line 5 – HESA fields that uniquely identify records on the amendment file, comma-separated. For example:

RECID,HUSID,OWNSTU,OWNPSD,NUMHUS

line 6 – this line must be left blank.

line 7 – number of records contained in the amendment file, excluding the first 12 lines of header information.

line 8 – the HESA field used to calculate the file check sum (see paragraph 22 for an explanation of the file check sum).

line 9 – file check sum.

lines 10 and 11 – these lines can contain any notes the institution wishes to include.

line 12 – HESA fields included in the amendment file. These fields must appear in the same order as each row of the data and must be comma-separated, on one line. For example:
RECID,HUSID,OWNSTU,OWNPSD,
NUMHUS,QUALAIM

Note that QUALAIM has been included for the file check sum.

line 13 – amendment data must begin on this line.

end of file – there must be a single blank line following the final record in the amendment file.

Uniquely identifying records

17. To enable us to link change or deletion files to our master HESA data set, we must be able to identify each record on the amendment file in such a way that it uniquely identifies the record on the HESA return. The field, or combination of fields, enabling us to achieve this must be listed, comma-separated, on line 5 of the change or deletion file.

18. We recommend institutions to use the following combination of five fields to uniquely identify HESA combined or student records:

- Field 1, RECID
- Field 4, HUSID
- Field 149/134, OWNSTU
- Field 150/135, OWNPSD
- Field 151/136, NUMHUS.

19. We recommend institutions to use the following combination of two fields to uniquely identify HESA module records:

- Field 1, RECID
- Field 5, MODID.

Validation

20. We use HESA’s validation software to ensure that all amendment files are valid and do not cause validation failures elsewhere in our master data sets. We will ask institutions to re-submit amendment files if validation failure occurs.

Saving files

21. Saving amendment files in Microsoft Excel usually results in the loss of leading zeros and the corruption of very large values into exponential form (for example, 9.91E+12). We recommend that amendment files are viewed and saved using a text editor, for example Notepad.

Check sum

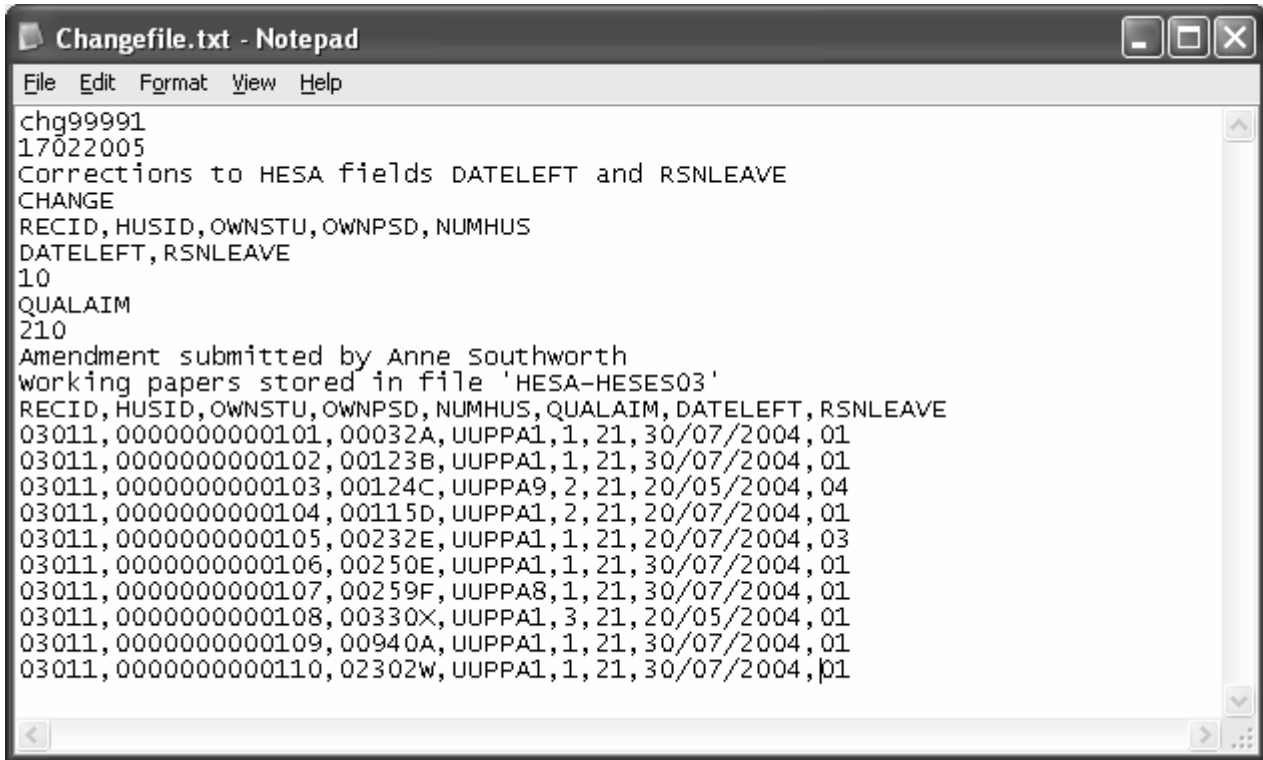
22. To ensure amendment files have not been corrupted during transit, we will check that the sum of values in this field match the sum returned on line 9. If an amendment file does not contain any numeric fields suitable for calculating a check sum, an additional field from the appropriate HESA record must be included solely for this purpose, for example QUALAIM. Numeric fields that contain values greater than 20,000 (approximately) are unsuitable for calculating the check sum.

Outcome

23. When we receive a valid amendment file in the structure and format detailed in this annex, we will aim to provide feedback within five working days. Institutions will be notified by e-mail what further action is required following incorporation, and when the revised re-creation tables and individualised file will be available from the HEFCE extranet.

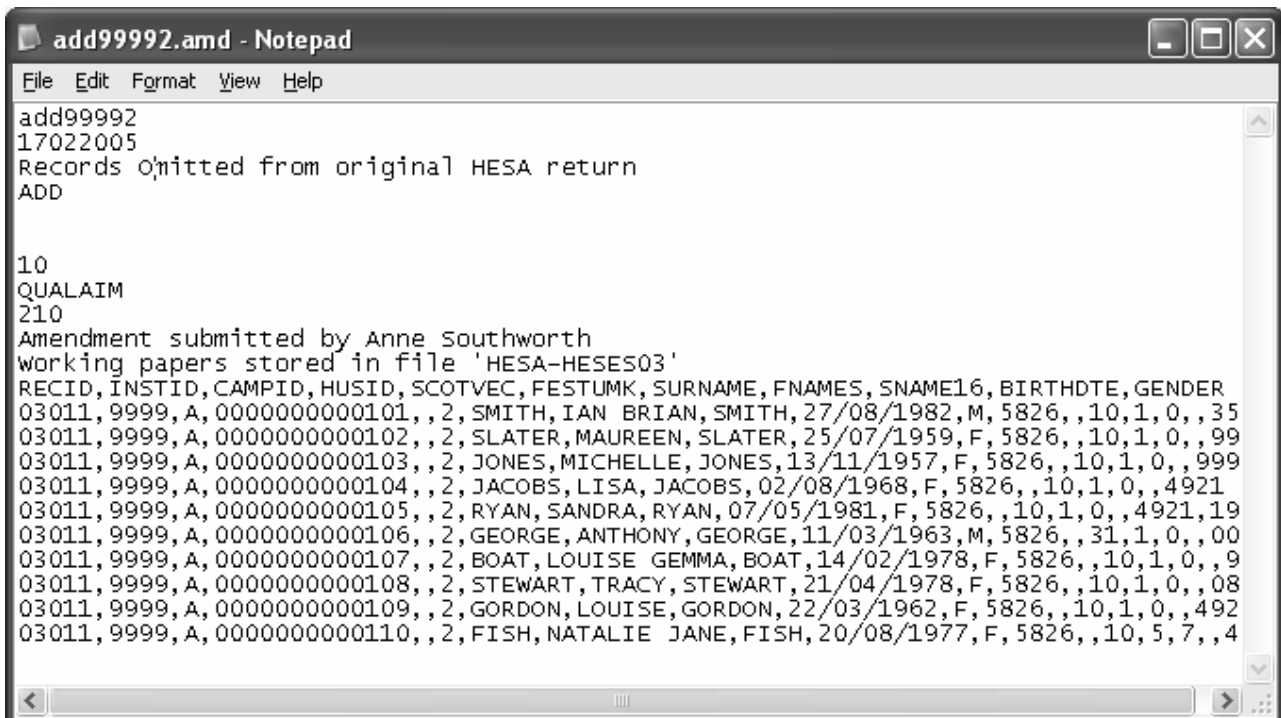
Examples of amendment files

Figure 5 **A change file**



```
Changefile.txt - Notepad
File Edit Format View Help
chg99991
17022005
Corrections to HESA fields DATELEFT and RSNLEAVE
CHANGE
RECID,HUSID,OWNSTU,OWNPSD,NUMHUS
DATELEFT,RSNLEAVE
10
QUALAIM
210
Amendment submitted by Anne Southworth
Working papers stored in file 'HESA-HESES03'
RECID,HUSID,OWNSTU,OWNPSD,NUMHUS,QUALAIM,DATELEFT,RSNLEAVE
03011,00000000000101,00032A,UUPPA1,1,21,30/07/2004,01
03011,00000000000102,00123B,UUPPA1,1,21,30/07/2004,01
03011,00000000000103,00124C,UUPPA9,2,21,20/05/2004,04
03011,00000000000104,00115D,UUPPA1,2,21,20/07/2004,01
03011,00000000000105,00232E,UUPPA1,1,21,20/07/2004,03
03011,00000000000106,00250E,UUPPA1,1,21,30/07/2004,01
03011,00000000000107,00259F,UUPPA8,1,21,30/07/2004,01
03011,00000000000108,00330X,UUPPA1,3,21,20/05/2004,01
03011,00000000000109,00940A,UUPPA1,1,21,30/07/2004,01
03011,00000000000110,02302W,UUPPA1,1,21,30/07/2004,01
```

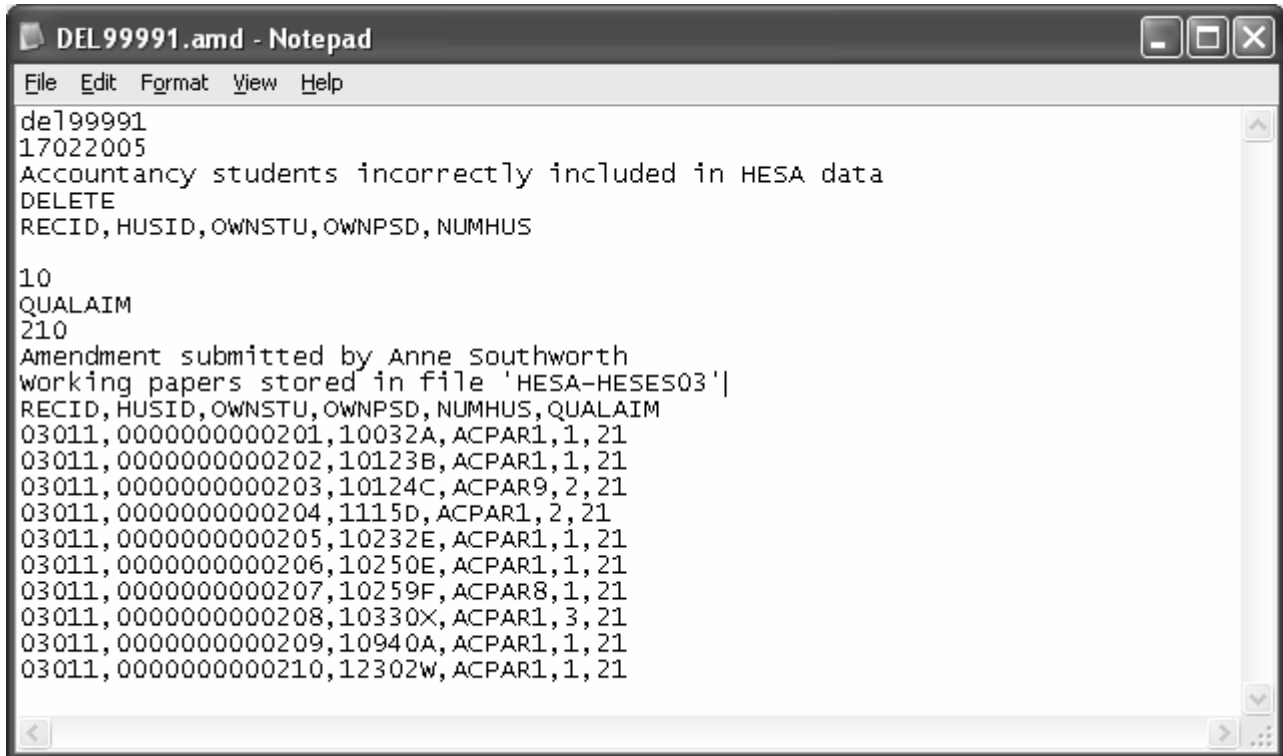
Figure 6 **An addition file**



```
add99992.amd - Notepad
File Edit Format View Help
add99992
17022005
Records omitted from original HESA return
ADD

10
QUALAIM
210
Amendment submitted by Anne Southworth
Working papers stored in file 'HESA-HESES03'
RECID,INSTID,CAMPID,HUSID,SCOTVEC,FESTUMK,SURNAME,FNAMES,SNAME16,BIRTHDTE,GENDER
03011,9999,A,00000000000101,,2,SMITH,IAN BRIAN,SMITH,27/08/1982,M,5826,,10,1,0,,35
03011,9999,A,00000000000102,,2,SLATER,MAUREEN,SLATER,25/07/1959,F,5826,,10,1,0,,99
03011,9999,A,00000000000103,,2,JONES,MICHELLE,JONES,13/11/1957,F,5826,,10,1,0,,999
03011,9999,A,00000000000104,,2,JACOBS,LISA,JACOBS,02/08/1968,F,5826,,10,1,0,,4921
03011,9999,A,00000000000105,,2,RYAN,SANDRA,RYAN,07/05/1981,F,5826,,10,1,0,,4921,19
03011,9999,A,00000000000106,,2,GEORGE,ANTHONY,GEORGE,11/03/1963,M,5826,,31,1,0,,00
03011,9999,A,00000000000107,,2,BOAT,LOUISE GEMMA,BOAT,14/02/1978,F,5826,,10,1,0,,9
03011,9999,A,00000000000108,,2,STEWART,TRACY,STEWART,21/04/1978,F,5826,,10,1,0,,08
03011,9999,A,00000000000109,,2,GORDON,LOUISE,GORDON,22/03/1962,F,5826,,10,1,0,,492
03011,9999,A,00000000000110,,2,FISH,NATALIE JANE,FISH,20/08/1977,F,5826,,10,5,7,,4
```

Figure 7 **A deletion file**



```
de199991
17022005
Accountancy students incorrectly included in HESA data
DELETE
RECID,HUSID,OWNSTU,OWNPSD,NUMHUS

10
QUALAIM
210
Amendment submitted by Anne Southworth
working papers stored in file 'HESA-HESES03'
RECID,HUSID,OWNSTU,OWNPSD,NUMHUS,QUALAIM
03011,00000000000201,10032A,ACPAR1,1,21
03011,00000000000202,10123B,ACPAR1,1,21
03011,00000000000203,10124C,ACPAR9,2,21
03011,00000000000204,1115D,ACPAR1,2,21
03011,00000000000205,10232E,ACPAR1,1,21
03011,00000000000206,10250E,ACPAR1,1,21
03011,00000000000207,10259F,ACPAR8,1,21
03011,00000000000208,10330X,ACPAR1,3,21
03011,00000000000209,10940A,ACPAR1,1,21
03011,00000000000210,12302W,ACPAR1,1,21
```

Table 2 **Summary of header information for amendment files**

Line	Change	Addition	Deletion
1	Amendment reference. For example 'chgXXXXn'	Amendment reference. For example 'addXXXXn'	Amendment reference. For example 'delXXXXn'
2	File creation date. For example 01032005	File creation date. For example 01032005	File creation date. For example 01032005
3	Brief description. For example 'Changes to HESA fields DATELEFT and RSNLEAVE'	Brief description. For example 'Records incorrectly omitted from original HESA return'	Brief description. For example 'Records incorrectly included in original HESA return'
4	CHANGE	ADD	DELETE
5	Unique identifying fields. For example: RECID,HUSID,OWNSTU, OWNPSD,NUMHUS	This line must be left blank	Unique identifying fields. For example: RECID,HUSID,OWNSTU, OWNPSD,NUMHUS
6	Fields to be corrected. For example: DATELEFT,FUNDCOMP	This line must be left blank	This line must be left blank
7	Number of records in the amendment file, excluding header	Number of records in the amendment file, excluding header	Number of records in the amendment file, excluding header
8	HESA field for check sum. Must be numeric	HESA field for check sum. Must be numeric	HESA field for check sum. Must be numeric
9	File check sum value	File check sum value	File check sum value
10	Any notes you wish to include	Any notes you wish to include	Any notes you wish to include
11	Any notes you wish to include	Any notes you wish to include	Any notes you wish to include
12	List of HESA fields in the same order as the data	List of HESA fields in the same order as the data	List of HESA fields in the same order as the data
13	The data must begin on this line	The data must begin on this line	The data must begin on this line
End	The final line in the file must be blank	The final line in the file must be blank	The final line in the file must be blank

Annex J

Submitting overrides to the algorithms

Background

1. An override file would only be applied where the data submitted to the HESA return is correct but there is a problem of fit with the HEFCE algorithms. Therefore changes need to be made to derived fields that generate the re-creations rather than to the underlying HESA data.
2. All known problems of fit with the HESES03 re-creation algorithms are described in Appendix 3. All known problems of fit with the HESES03 re-creation based on cost centre sector norms for subjects are described in Appendix 6. All known problems of fit with the RAS03 re-creation algorithms are described in Appendix 9.
3. The problem of fit that the override is rectifying should be stated clearly on the action plan. We will only apply overrides where we agree that they are appropriate, and in the case of an override to the sector norm cost centre mapping have made a decision based on evidence provided. Therefore we may seek further information where necessary. For example, in the case of the sector norm cost centre mapping we may seek further module information. It may not always be possible to determine whether an override is appropriate until we have examined the students affected. Therefore we may refuse to accept an override once submitted or will seek further clarification.

Purpose

4. This annex details the data structure and format for overrides to derived fields. Institutions must only supply override files using the file structure and format detailed within this annex.
5. Override files should contain the data structure and format which is described in paragraphs 8 to 11, with slight modifications for overrides provided in four special cases. The four special cases occur where overrides are made to the following derived fields:
 - a. Proportion of countable year in each price group (PRGA, PRGB, PRGC, PRGD, PRGMEDIA, PRGPSYCH, PRGITT, PRGINSET).

- b. Proportion of countable year in each sector norm price group (SNPRGA, SNPRGB, SNPRGC, SNPRGD, SNMEDIA, SNPSYCH, SNITT, SNINSET).
 - c. RAS UoAs (RASUOA1-3).
 - d. Proportion of time spent in each subject area, used to scale FTE (UOAP1-UOAP3).
6. The structure and format that the override file should contain for each of these cases is described in this annex. An example of a typical override file, with examples of files that should be generated for each of the four cases above, can be found at the end of this annex (see Figures 8-12).
 7. These specifications are necessary to ensure we can process overrides to derived fields in a timely and accurate manner. We will require institutions to re-submit override files that differ, either in structure or format, to the specifications detailed in this annex.

Format and naming

8. Overrides to derived fields must be sent as a comma-separated file in an e-mail attachment to hesa_heses_stats@hefce.ac.uk. We will also accept overrides on a 3.5" floppy disk or CD-ROM. Override files must be given a file name in the form `ovrXXXXn.amd`, where:
 - XXXX is the HESA institution identifier for the institution
 - n is a sequential number starting at 1.

For example, the first override file submitted would be called `ovrXXXX1.amd`, and the second would be called `ovrXXXX2.amd`.

Structure

9. Each record must contain complete data for all fields included in the override file, even if a particular field remains unchanged in some cases. For example it should contain the fields used to identify records (line 5).
10. Only derived fields should be included in the change line (line 6) for override files. For further information on which derived fields are affected by problems of fit, see Appendices 3, 6 and 9.

11. In addition to the records that contain derived fields being overwritten, the file must contain a header in the following form:

line 1 – override reference in the form ovrXXXXn where XXXX is the HESA institution identifier and n is a sequential number starting at 1; this will be the file name with the ‘.amd’ file extension removed.

line 2 – creation date of the override file in the form ddmmYYYY. For example 01032005 for a file created on 1 March 2005.

line 3 – a brief description of the override and the reference number of the area of difference on the action plan that it rectifies. For example ‘Overrides to derived field HESFEELV’ - reference number 3.

line 4 – this line must contain the words OVERRIDE, NORMAL, and either the word TEMPORARY or PERMANENT. If the override is temporary then the last academic year that it applies to should be entered. For example ‘OVERRIDE, NORMAL, TEMPORARY, 2004’ indicates that the override will be applied in academic year 2004-05 but not in 2005-06 onwards.

line 5 – the HESA fields used to identify records on the override file, comma-separated. For example: OWNPSD could be used to identify records on a course level or HUSID, NUMHUS, RECID can be used to identify records on a student level.

line 6 – the derived fields being changed, comma-separated. For example: HESFEELV.

line 7 – number of records contained in the file, excluding the first 12 lines of header information.

line 8 – the HESA/derived field used to calculate the file check sum (see paragraph 18 for an explanation of the file check sum).

line 9 – file check sum.

lines 10 and 11 – these lines can contain any notes the institution wishes to include.

line 12 – HESA/derived fields included in the override file. These fields must appear in the same

order as each row of the data and must be comma-separated. For example:
OWNPSD, HESFEELV on one line.

line 13 – override data must begin on this line.

end of file – there must be a single blank line following the final record in the override file.

Special cases

Proportion of countable year in each price group

12. If overrides are being applied to price groups then we require information about all of the price group fields PRGA, PRGB, PRGC, PRGD, PRGMEDIA, PRGPSYCH, PRGITT, PRGINSET to be provided (even if a particular price group is not being changed). Also the word PRICEGRP must be substituted in line 4 to replace the word NORMAL. An example of this file is given in Figure 9.

Proportion of countable year in each sector norm price group

13. If permission has been granted by HEFCE, and overrides are being applied to the sector norm price groups, then the override file should contain information about all of the sector norm price group fields SNPRGA, SNPRGB, SNPRGC, SNPRGD, SNMEDIA, SNPSYCH, SNITT, SNINSET (even if a particular sector norm price group is not being changed). We would advise that a field called SBJ that contains the JACS code of the subject of the area of study is used as the linking field. The word SNPRGRP must be substituted in line 4 to replace the word NORMAL. An example of this file is given in Figure 10. If required, we will construct an override file. This will be implemented when the institution has approved the file.

RAS UoA's

14. If permission has been granted by HEFCE, and overrides are being applied to RASUOA1-3, then we would advise that the subject of qualification aim fields SBJQA1-3 are used as the linking fields. The word RASUOA must be substituted in line 4

to replace the word NORMAL. An example of this file is given in Figure 11. If required, we will construct an override file. This will be implemented when the institution has approved the file.

Proportion of time spent in each subject area, used to scale FTE

15. If overrides are being applied to UOAP1, UOAP2, UOAP3 then we require information about all of these fields. We would advise that the subject of qualification aim fields SBJQA1-3 are used as the linking fields. An example of this file is given in Figure 12.

Identifying records

16. To enable us to link override files to our derived HESA dataset, we must be able to identify the records on the HESA return where the override should be applied. The field, or combination of fields, enabling us to achieve this must be listed, comma-separated, on line 5 of the override file.

Saving files

17. Saving override files in Microsoft Excel usually results in the loss of leading zeros and the corruption of very large values into exponential form (for example, 9.91E+12). We recommend that override files are viewed and saved using a text editor, for example Notepad.

Check sum

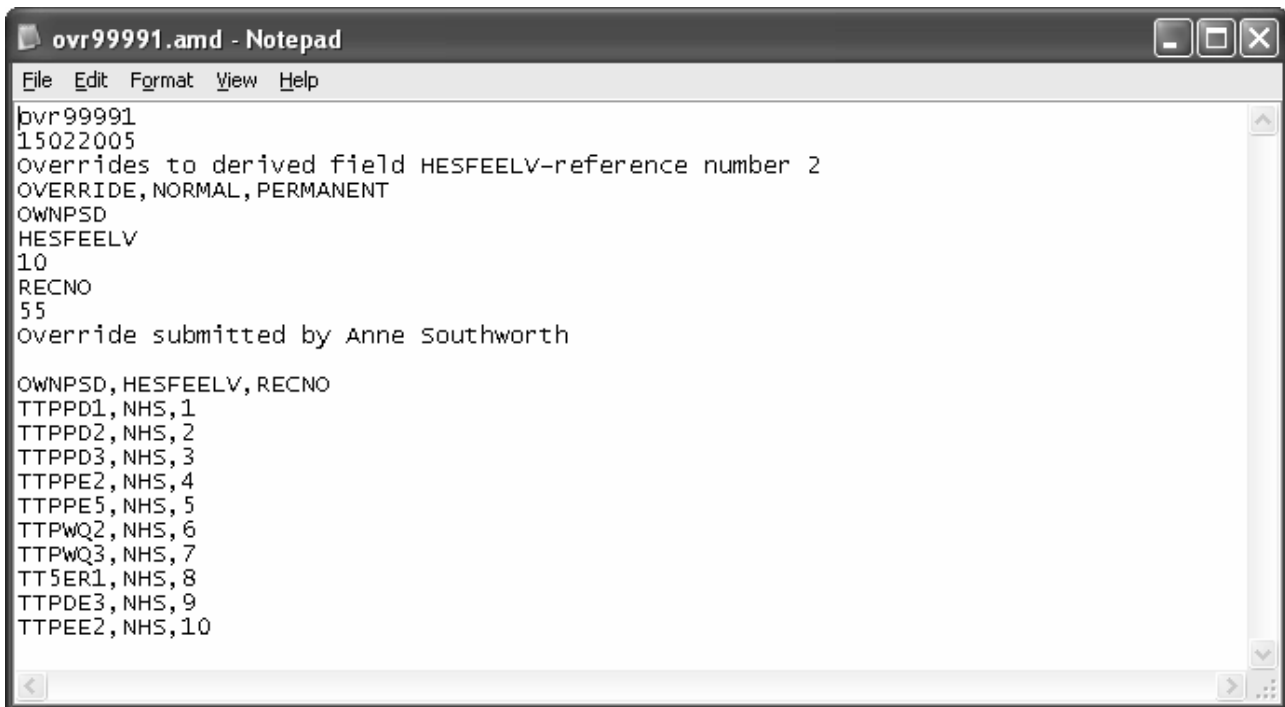
18. To ensure override files have not been corrupted during transit, we will check that the sum of values in this field match the value returned on line 9. If an override file does not contain any numeric fields suitable for calculating a check sum, an additional field from the appropriate HESA record must be included solely for this purpose, for example QUALAIM. Numeric fields that contain values greater than 20,000 (approximately) are unsuitable for calculating the check sum. If information is not being changed at the student level, then a sequential field called RECNO may be created for the purpose of the check sum. For example RECNO may contain 1, 2, 3, 4, 5 etc.

Outcome

19. When we receive a valid override file in the structure and format detailed in this annex, we will aim to provide feedback within five working days. Institutions will be notified by e-mail what further action is required following incorporation, and when the revised re-creation tables and individualised file will be available from the HEFCE extranet.

Examples of override files

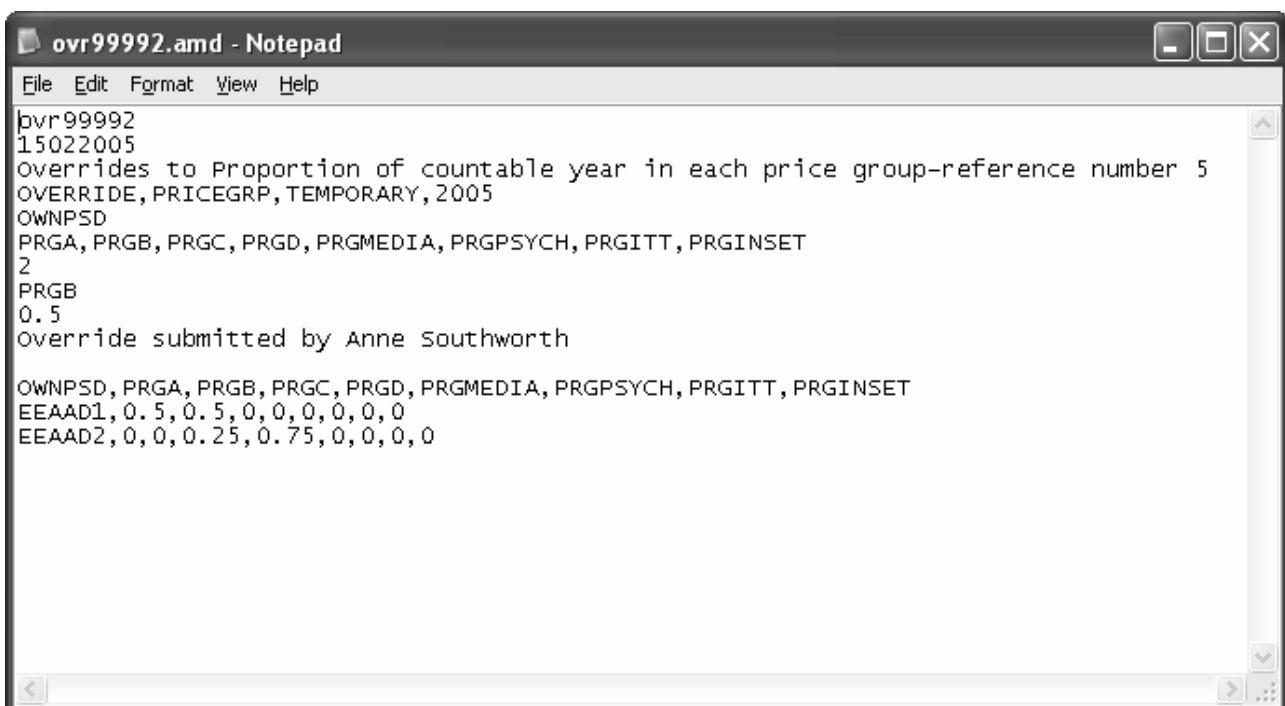
Figure 8 **A typical override file**



```
ovr99991
15022005
Overrides to derived field HESFEELV-reference number 2
OVERRIDE,NORMAL,PERMANENT
OWNPSD
HESFEELV
10
RECNO
55
Override submitted by Anne Southworth

OWNPSD, HESFEELV, RECNO
TTPPD1, NHS, 1
TTPPD2, NHS, 2
TTPPD3, NHS, 3
TTPPE2, NHS, 4
TTPPE5, NHS, 5
TTPWQ2, NHS, 6
TTPWQ3, NHS, 7
TT5ER1, NHS, 8
TTPDE3, NHS, 9
TTPPE2, NHS, 10
```

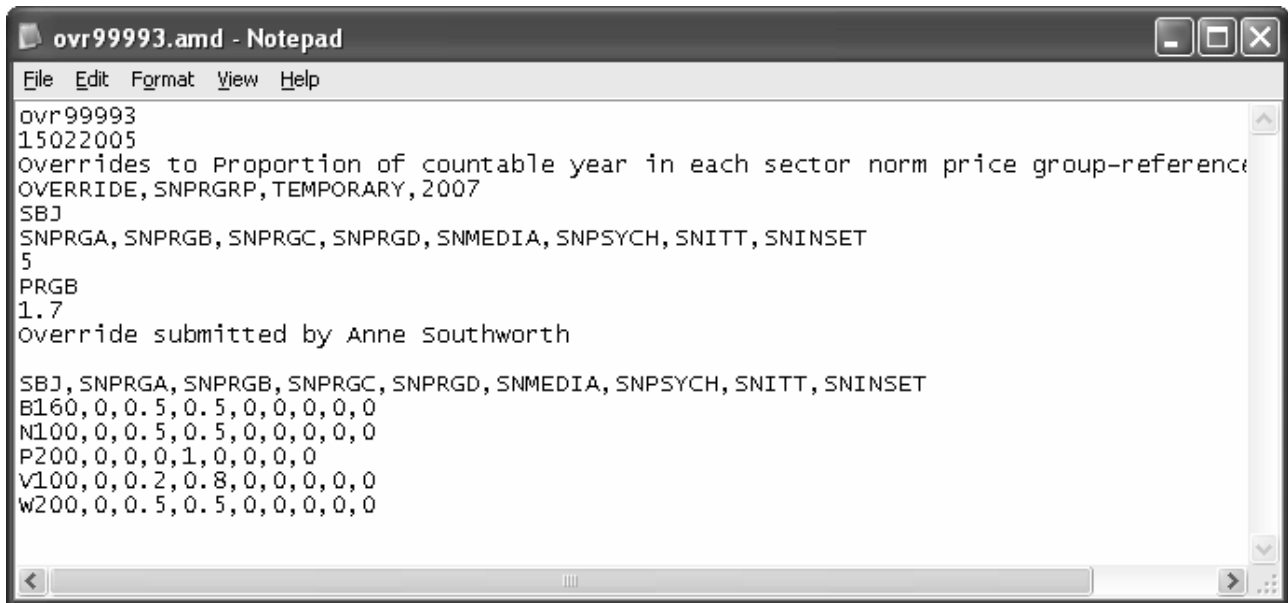
Figure 9 **Proportion of countable year in each price group file**



```
ovr99992
15022005
Overrides to Proportion of countable year in each price group-reference number 5
OVERRIDE, PRICEGRP, TEMPORARY, 2005
OWNPSD
PRGA, PRGB, PRGC, PRGD, PRGMEDIA, PRGPSYCH, PRGITT, PRGINSET
2
PRGB
0.5
Override submitted by Anne Southworth

OWNPSD, PRGA, PRGB, PRGC, PRGD, PRGMEDIA, PRGPSYCH, PRGITT, PRGINSET
EEAAD1, 0.5, 0.5, 0, 0, 0, 0, 0, 0
EEAAD2, 0, 0, 0.25, 0.75, 0, 0, 0, 0
```

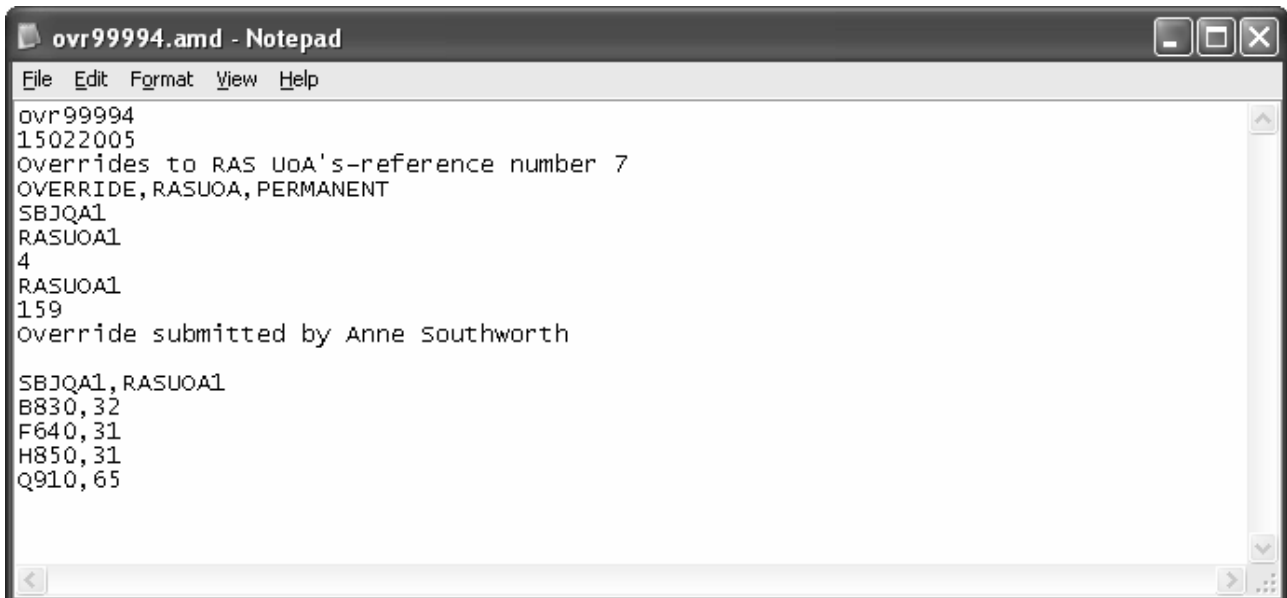

Figure 10 **Proportion of countable year in each sector norm price group file**



```
ovr99993
15022005
Overrides to Proportion of countable year in each sector norm price group-reference
OVERRIDE,SNPRGRP,TEMPORARY,2007
SBJ
SNPRGA,SNPRGB,SNPRGC,SNPRGD,SNMEDIA,SNPSYCH,SNITT,SNINSET
5
PRGB
1.7
Override submitted by Anne Southworth

SBJ,SNPRGA,SNPRGB,SNPRGC,SNPRGD,SNMEDIA,SNPSYCH,SNITT,SNINSET
B160,0,0.5,0.5,0,0,0,0,0
N100,0,0.5,0.5,0,0,0,0,0
P200,0,0,0,1,0,0,0,0
V100,0,0.2,0.8,0,0,0,0,0
W200,0,0.5,0.5,0,0,0,0,0
```

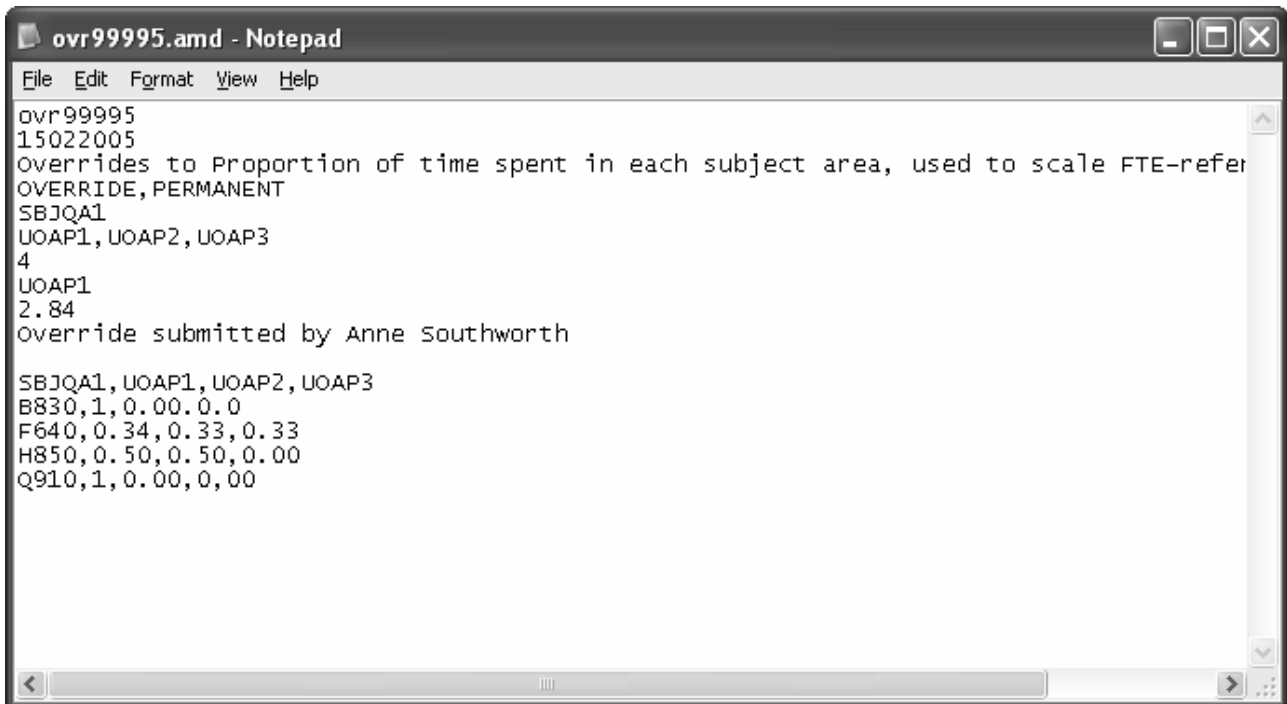
Figure 11 **RAS UoAs file**



```
ovr99994
15022005
Overrides to RAS UoA's-reference number 7
OVERRIDE,RASUOA,PERMANENT
SBJQA1
RASUOA1
4
RASUOA1
159
Override submitted by Anne Southworth

SBJQA1,RASUOA1
B830,32
F640,31
H850,31
Q910,65
```

Figure 12 **Proportion of time spent in each subject area, used to scale FTE**



```
ovr99995
15022005
Overrides to Proportion of time spent in each subject area, used to scale FTE-refer
OVERRIDE,PERMANENT
SBJQA1
UOAP1,UOAP2,UOAP3
4
UOAP1
2.84
Override submitted by Anne Southworth

SBJQA1,UOAP1,UOAP2,UOAP3
B830,1,0.00,0.0
F640,0.34,0.33,0.33
H850,0.50,0.50,0.00
Q910,1,0.00,0,00
```

Annex K

List of abbreviations

Acronym	Description
EC	European Community
FAQ	Frequently asked question
FE	Further education
FEC	Further education college
FTE	Full-time equivalent
HESA	Higher Education Statistics Agency
HESES	Higher Education Students Early Statistics
HIN	HUSID X INSTID X NUMHUS
ISR	Individualised Student Record
JACS	Joint Academic Coding System
LSC	Learning and Skills Council
RAS	Research Activity Survey
UCAS	Universities and Colleges Admissions Service
UoA	Unit of Assessment