

August 2004/29

**Core funding/operations**

Allocation of places and funds

This report is for information and guidance

This document is a guide to the HEFCE's web facility which generates several types of analysis from Higher Education Statistics Agency (HESA) student data. Use of the web facility will help institutions to return accurate data to HESA, prepare for the 2005 National Student Survey, and to identify errors and forecasting discrepancies in HEFCE funding returns.

# 2003-04 statistics derived from HESA data

## Guide to HEFCE web facility

<b>Executive summary</b>		3
<b>Introduction</b>		5
<b>Annexes</b>		
Annex A	Summary of changes since the HESES02 re-creation web facility	9
Annex B	Using the HEFCE web facility for 2003-04 statistics derived from HESA data	11
Annex C	Comparison of HESES03 and the HESES03 re-creation	16
Annex D	Comparison of the HESES03 re-creation and the HESES03 re-creation based on cost centre sector norms for subjects	18
Annex E	National Student Survey 2005 – provisional target list	21
Annex F	Comparison of RAS03 and the RAS03 re-creation	23
Annex G	Derived statistics used to inform the 2005-06 widening participation funding allocations	25
Annex H	Derived statistics used to inform any 2005-06 transfer of sports science activity	27
Annex I	HEFCE statistical publications	28
Annex J	Research degree rates of qualification *	
Annex K	HESES04 non-completion toolkit	31
Annex L	Invalid postcode list	32
Annex M	List of abbreviations	33
<b>Technical appendices</b>		
Appendix 1	HESES03 re-creation algorithms	34
Appendix 2	Troubleshooting the differences between HESES03 and the HESES03 re-creation	60
Appendix 3	Problems of fit with the HESES03 re-creation algorithms	70
Appendix 4	HESES03 re-creation based on cost centre sector norms for subjects algorithms	72
Appendix 5	Troubleshooting the differences between the HESES03 re-creation and the HESES03 re-creation based on cost centre sector norms for subjects	78
Appendix 6	Problems of fit with the HESES03 re-creation based on cost centre sector norms for subjects algorithms	80
Appendix 7	National Student Survey 2005 - provisional target list algorithms	81
Appendix 8	RAS03 re-creation algorithms	86
Appendix 9	Troubleshooting the differences between RAS03 and the RAS03 re-creation	97
Appendix 10	Problems of fit with the RAS03 re-creation algorithms	102
Appendix 11	Widening participation derived statistics algorithms	104
Appendix 12	Sports science derived statistics algorithms	110
Appendix 13	HEFCE statistical publications algorithms	113
Appendix 14	Research degree rates of qualification algorithms *	
Appendix 15	HESES04 non-completion toolkit algorithms	128
Appendix 16	Invalid postcode list description	129

\* Annex J and Appendix 14 are not yet available and will be put on the HEFCE web-site in the near future.

# 2003-04 statistics derived from HESA data

## Guide to HEFCE web facility

<b>To</b>	Heads of HEFCE-funded higher education institutions Heads of universities in Northern Ireland
<b>Of interest to those responsible for</b>	Student data, Audit, Finance
<b>Reference</b>	2004/29
<b>Publication date</b>	August 2004
<b>Enquiries to</b>	Matthew Davies tel 0117 931 7373 Gemma Harper tel 0117 931 7338 e-mail hesa_heses_stats@hefce.ac.uk

### Executive summary

#### Purpose

1. This document is a guide to HEFCE's web facility which institutions can use to check their Higher Education Statistics Agency (HESA) 2003-04 student data. It generates the following derived statistics:
  - a. Higher Education Students Early Statistics Survey 2003-04 (HESES03) re-creation.
  - b. HESES03 re-creation based on cost centre sector norms for subjects.
  - c. Provisional target list of students that will be included in the 2005 National Student Survey (NSS05).
  - d. Research Activity Survey 2003 (RAS03) re-creation.
  - e. Derived statistics that are likely to be used to inform the 2005-06 widening participation (WP) funding allocation.
  - f. Derived statistics that are likely to be used to identify sports science and leisure studies activity for funding from 2005-06.
  - g. Derived statistics that will be used in HEFCE statistical publications.
  - h. Research degree rates of qualification (to be included in a later release of this publication).
  - i. Higher Education Student Early Statistics Survey 2004-05 (HESES04) non-completion toolkit.
  - j. List of invalid postcodes.

## Key points

2. The use of the web facility will assist institutions in:
  - a. Returning accurate data to HESA.
  - b. Identifying discrepancies between forecasts in HESES03 and the outturn position for 2003-04.
  - c. Identifying errors in HEFCE funding returns.
  - d. Preparing for the first 2005 National Student Survey.
  - e. Estimating non-completion rates for HESES04.
3. This document, together with the accompanying technical appendices provides:
  - a. Guidance for using the web facility.
  - b. Algorithms for generating the derived statistics.
  - c. Troubleshooting guidance for the comparison of HEFCE funding returns with the re-creations.
  - d. Details of problems of fit with algorithms.
  - e. Guidance for using the HESES04 non-completion toolkit.
4. The name of the facility has changed to the 'HEFCE web facility for 2003-04 statistics derived from HESA data' to reflect that it now generates derived statistics other than the HESES and RAS re-creations. We will refer to it as a web facility throughout this guidance. The derived statistics that are new to this web facility will be:
  - a. HESES03 re-creation based on cost centre sector norms for subjects.
  - b. Provisional target list of students that will be included in the NSS05.
  - c. Derived statistics that are likely to be used to identify sports science and leisure studies activity for funding from 2005-06.
  - d. Research degree rates of qualification.
  - e. List of invalid postcodes.

## Relationship with 'Derived statistics from HESA data'

5. Each year we publish 'Statistics derived from HESA data for monitoring and allocation of funding' in which we monitor institutions' HESES and RAS returns using HESA student data. If we find, either through comparisons with HESA data or any data audit, that erroneous data have resulted in institutions receiving incorrect funding allocations then we will adjust their funding accordingly (subject to the appeals process and the availability of funds). Any funding adjustments following the comparison of HESA 2003-04 student data with HESES03 are likely to affect the funding previously announced for 2003-04 and all subsequent years. Any funding adjustments arising from the comparison of RAS03 and the RAS03 re-creation are likely to affect the funding previously announced for 2004-05. Whilst the web facility is provided to complement the monitoring exercise it does not replace it.

## Introduction

6. This document provides guidance notes on using the web facility and its outputs. The primary purpose of the web facility is to assist institutions in returning accurate data to HESA. The web facility gives institutions an opportunity to identify, and therefore rectify, any errors in fields that affect the outputs generated by the web facility prior to submitting final data to HESA. In the future, we expect that submission of individualised student data to HESA by institutions will be conditional upon prior use of the web facility.

7. Use of this facility, prior to submission of the HESA return, is strongly encouraged by both the HESA and HEFCE as both organisations regard it as an essential element of all institutions' data quality processes. The HEFCE web facility for 2003-04 statistics derived from HESA data is a separate utility to the HESA data collection system.

8. Typically institutions submit their data to the web facility and retrieve the resultant derived statistics several times before all identified errors are remedied. Therefore, we will not restrict institutions' use of the web facility. However, as resources at HEFCE are limited, at times when there is high demand, response times may be slower.

9. Institutions should include adequate time within their timetable to allow them to make full use of the web facility without jeopardising HESA quality arrangements or timetables.

10. Since its introduction two years ago nearly 120 higher education institutions have utilised the web facility. A review of the web facility has confirmed that it will be an essential element of most institutions' HESA data quality processes this year.

11. We believe that the introduction of the web facility has contributed to an improvement in HESA student data returns. As a result of this improved data quality we were able to significantly reduce the selection thresholds for the 'HESA 2002-03 derived statistics for the monitoring and allocation of funding' exercise (HEFCE 2004/10). Over the last two years, out of the institutions

selected for the statistics derived from HESA data exercise, there have been a disproportionate number of institutions that did not make use of the web facility. We expect to be able to reduce the selection thresholds further for the 2003-04 statistics derived from HESA data exercise as a result of expected improvements in data quality in the 2003-04 HESA student return.

12. Changes to the web facility since 'Guide to HESES02 re-creation web facility' (HEFCE 2003/39) are described in Annex A. Annex B describes how to access and use the web facility.

13. The web facility generates five types of output. These are:

- re-creations of HEFCE funding returns
- NSS target lists
- derived statistics used to inform HEFCE funding
- derived statistics used in HEFCE statistical publications
- data quality tools.

### Re-creations of HEFCE funding returns

#### HESES03 re-creation

14. The web facility generates a HESES03 re-creation from HESA 2003-04 student data, including the calculation of formulaic adjustments to grant, using the same algorithms that we intend to use in the 2003-04 statistics derived from HESA data exercise. The HESES03 re-creation is generated using the methods described in Annex C.

15. We advise that institutions use this opportunity to identify the cause of all discrepancies between their HESES03 return and the HESES03 re-creation, so that where errors in HESA data are a cause, these can be corrected before submission to HESA. The removal of such errors in HESA data reduces the likelihood of selection for the 2003-04 statistics derived from HESA data exercise.

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

16. The web facility will also generate, from HESA 2003-04 student data, a HESES03 re-creation that is based on a sector norm mapping of subjects to cost centres. This re-creation will include the calculation of formulaic adjustments to grant using the same algorithms that we intend to use in the 2003-04 statistics derived from HESA data exercise, except that the sector norm mapping of subjects to cost centres will be revised on receipt of full HESA 2003-04 student data. The web facility uses the cost centre sector norms generated from HESA 2002-03 student data, and we do not expect the sector norms for 2003-04 to differ significantly from them. The HESES03 re-creation based on cost centre sector norms for subjects is generated using the methods described in Annex D.

17. We advise that institutions use this opportunity to determine the cause of discrepancies between their HESES03 re-creation and the HESES03 re-creation based on cost centre sector norms for subjects, so that where errors in HESA data are a cause, these can be corrected before submission to HESA. Removal of errors in HESA data reduces the likelihood of selection to the 2003-04 statistics derived from HESA data exercise.

### **RAS03 re-creation**

18. The web facility will also generate, from HESA 2003-04 student data, a RAS03 re-creation, including a funding summary, using the same algorithms that we intend to use in the 2003-04 statistics derived from HESA data exercise. The RAS03 re-creation is generated using the methods described in Annex F.

19. We advise that institutions use this opportunity to determine the cause of discrepancies between their RAS03 and the RAS03 re-creation, so that where errors in HESA data are a cause, these can be corrected before submission to HESA. Removal of errors in HESA data reduces the likelihood of selection to the 2003-04 statistics derived from HESA data exercise.

### **NSS target lists**

20. In 'National Student Survey 2005: Consultation' (HEFCE 2004/22) we announced our intention to allow institutions to view a provisional list of students to be included in NSS05. This provisional list will be generated by the web facility. The methods used to generate the list are described in Annex E.

21. Institutions are encouraged to use these provisional lists to start preparing the contact details for students. We will not be able to provide definitive target lists until we have received final data from HESA in mid-December, and completed lists will need to be returned to the agency in early January. To help institutions meet this deadline we will retain the final copy of the target list generated by the web facility so that we can provide details of the changes in coverage between institutions' final use of the web facility and the generation of lists from signed off HESA data. We do not expect to make any changes to the algorithms for generating target lists, therefore any changes will arise from changes to institutions' data.

### **Derived statistics used to inform HEFCE funding**

#### **2005-06 WP funding allocations**

22. HESA 2003-04 student data are likely to be used to inform the following widening access funding allocations for 2005-06:

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

The derived statistics used to inform the 2005-06 WP funding allocations are likely to be generated using the methods described in Annex G.

23. The methods that will be used to calculate the improving retention allocations based upon qualifications on entry and age will be determined at a later date.

## **2005-06 funding for sports science and leisure studies**

24. In 'Funding method for teaching from 2004-05: Outcomes of consultation' (HEFCE 2004/24) we announced a review to assess the relative cost of different types of provision in the sports science and leisure studies cost centre. If the outcome of this review identifies the need for a transfer of sports science activity to different price groups than the current assignments, the derived statistics used to inform a 2005-06 transfer of sports science activity will be generated using the methods described in Annex H.

## **Research degree rates of qualification**

25. In our consultation paper 'Improving standards in postgraduate degree programmes' (HEFCE 2003/23) we made proposals for developing threshold standards and an underlying framework of good practice for postgraduate research degree programmes. 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)). Following these consultations the HE funding bodies are now discussing with the Quality Assurance Agency for HE (QAA) the possibility of implementing threshold standards through a revised version of the QAA code of practice, informed and supplemented by appropriate statistical data. We hope to make an announcement on this in the near future so that the new arrangements for threshold standards can be put in place from 2005-06. In the meantime, we anticipate that the statistical data to be taken into account within these arrangements will include a new analysis of research degree qualification rates. Construction of this indicator will require the availability of robust figures covering a number of years.

26. The research degree rates of qualification are likely to be generated using the methods that will be described in Annex J in a later release of this publication.

## **Derived statistics used in HEFCE statistical publications**

27. The web facility will generate derived statistics that will be used in HEFCE statistical publications, including the following, to be produced during 2005:

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

These data may be used to inform the allocation of any growth in foundation degree places.

28. The derived statistics used in HEFCE statistical publications during 2005 will be generated using the methods described in Annex I.

## **Data quality tools**

### **HESES04 non-completion toolkit**

29. The web facility will generate a spreadsheet that is designed to assist institutions in estimating non-completion rates for HESES04. A detailed explanation of how to use the HESES04 non-completion toolkit is given in Annex K.

### **Invalid postcode list**

30. To help institutions return accurate data to HESA we will provide a list of students that have postcodes that do not appear on the Post Office list of valid postcodes. Further details about the fields used in checking the validity of the postcodes are contained in the All Fields Postcode Directory (AFPD). Further details about the AFPD are available from the National Statistics web-site, [www.statistics.gov.uk](http://www.statistics.gov.uk). The invalid postcode list will be generated using the methods described in Annex L.

## **Data confidentiality**

31. The web facility is primarily a tool for institutions. Therefore, HEFCE staff will only access data relating to national student survey target lists (see Annex E) and will not access any

other data submitted to, or derived by, the web facility except where an institution explicitly gives permission. Only those fields listed in the technical appendices will be retained on our system. We will remove individualised files and any outputs from the HEFCE extranet by 1 May 2005, and the submitted data files will be deleted immediately after the outputs have been generated. Details about our commitments to handling personal data can be found on the HEFCE web site, [www.hefce.ac.uk](http://www.hefce.ac.uk), under About us/Standards of service/How we do it – our commitments/Release guidelines for personal data.

### **HEFCE-recognised funding consortia**

32. Later this year, we expect to make available a Higher Education in Further Education Students Survey 2003-04 (HEIFES03) re-creation web facility to further education colleges which will allow them to generate a HEIFES03 re-creation from their Individual Learner Record (ILR). We encourage member colleges of HEFCE-recognised funding consortia to use this facility prior to submission of the ILR July 2003-04 return to the Learning and Skills Council (LSC).

### **Guidance**

#### **FAQs**

33. Frequently asked questions (FAQs) regarding the web facility 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.

#### **Comments**

34. All institutions are invited to comment on any of the methods described in this publication. Comments should be sent to Ben Grassby ([b.grassby@hefce.ac.uk](mailto:b.grassby@hefce.ac.uk)).

35. Comments or queries regarding the invalid postcode lists should be sent to Anne Southworth

([a.southworth@hefce.ac.uk](mailto:a.southworth@hefce.ac.uk)), as a review is due to be carried out in November.

### **Annexes and technical appendices**

36. The annexes to this publication provide guidance on using the web facility and its outputs. The technical appendices describe the algorithms used to generate the derived statistics.



# Annex A

## Summary of changes since the HESES02 re-creation web facility

### Purpose

1. This annex describes the changes that have been made since the release of the equivalent publication 'Guide to HESES02 re-creation web facility' (HEFCE 2003/39) last year.
2. The name of the facility has changed to the 'HEFCE web facility for 2003-04 statistics derived from HESA data' to reflect that it now generates more than simply the HESES re-creation.
3. The web facility is only available via the HEFCE extranet at <https://extranet.hefce.ac.uk>. Full details of how to access and use the web facility can be found in Annex B.

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

4. In addition to the HESES03 re-creation, the web facility will generate a HESES03 re-creation based on cost centre sector norms for subjects which will be derived for HESA 2003-04 student data. As part of next year's 2003-04 statistics derived from HESA data exercise, we intend to use cost centre sector norms generated from HESA 2003-04 student data. The web facility uses the cost centre sector norms generated from HESA 2002-03 student data, however we do not expect the sector norms for 2003-04 to change significantly. A detailed explanation of the comparison is given in Annex D.

### Provisional NSS target lists

5. A provisional list of students that will be included in NSS05 will be generated by the web facility. An explanation of the methods we will use is given in Annex E.

### Funding for sports science and leisure studies

6. Subject to the outcome of review, we may transfer sports science and leisure studies activity to different price groups than the current assignments. An explanation of the methods we

will use to identify this activity is given in Annex H.

### Widening participation

7. The methods used to generate the derived statistics used to inform widening participation (WP) funding allocations have fundamentally changed.
8. An explanation of the methods we are likely to use to inform the 2005-06 widening access funding allocations is described in Annex G.
9. The methods that will be used to calculate the improving retention allocations based upon qualifications on entry and age will be determined at a later date.

### HEFCE statistical publications

10. The basis for deriving statistics for HEFCE publications has fundamentally changed. We now base the population definition on the HESA population definition rather than the HESES population definition. Annex I contains details of the changes.
11. Currently the web facility serves as the main mechanism for institutions to verify the derived statistics that will be used in HEFCE statistical publications during 2005.
12. Students that are registered at one institution and taught by another institution are described as 'franchised'.

### Research degree rates of qualification

13. Provisional research degree rates of qualification will be generated by the web facility. An explanation of the methods we will use will be described in Annex J.

### Invalid postcode list

14. A list of students with invalid postcodes will be generated by the web facility. An explanation of the methods we will use is given in Annex L.

## **Grant adjustment report**

15. A detailed explanation of the algorithms used to derive the grant adjustment report, standard resource table and assumed fee income table using HESA 2003-04 student data has been added to the explanation of the HESES03 re-creation algorithms contained in Appendix 1.

## **Technical appendices**

16. The technical descriptions of the methods we have used to generate the outputs have been separated into appendices that accompany the publication. The appendices are available, along with the rest of this publication, on the HEFCE web-site under Publications.

## **Individualised files**

17. 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 B.

## **Version control and table page numbers**

18. For ease of use we have added version numbers and table page numbers to outputs generated by the web facility. The version number will increment by one following each successful submission of data to the web facility.

## **Changes to algorithms**

19. The changes to algorithms listed here are those algorithms that have changed since 'HESA 2002-03 derived statistics for the monitoring and allocation of funding' (HEFCE 2004/10).

### **HESFEELV**

20. The algorithm for deriving HESFEELV has changed as a result of improvements to the algorithm for identifying NHS bursaried courses. The algorithm is given in paragraph 36 of Appendix 1.

### **HESNHS**

21. 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 detailed in paragraph 35 of Appendix 1.

### **PRGA**

22. 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 paragraph 59 of Appendix 1.

### **RASYEAR**

23. The algorithm for deriving RASYEAR has been revised to take into account students on non-standard academic years. The algorithm is given in paragraphs 12 and 13 of Appendix 8.

### **PBLLEV**

24. 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 16.

### **PBLSTU**

25. 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 16.

### **PBLLOAD**

26. 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 16.

# Annex B

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

### Purpose

1. This annex contains detailed instructions on how to use the web facility.

### Data preparation

2. The web facility can process data files structured in the formats described in the HESA 2003-04 individualised student coding manual. This includes both fixed-width and comma-separated formats, as well as both the student/module and combined records.

3. While every effort has been made to ensure the web facility is robust, it may not function reliably when the input files contain invalid data. In order to ensure reliable outputs are generated, institutions should correct errors identified by the HESA Validation kits before submission. HESA's validation kits can be found at [www.hesa.ac.uk](http://www.hesa.ac.uk) under Data Collection/Download Validation Kits.

4. In addition HESA publishes a number of commit-level validation checks. If data fail these checks, the web facility may generate unreliable results. In particular, if the combination of HUSID, INSTID and NUMHUS is not unique for any records for students with higher education qualification aims, then the re-creation will generate incorrect results.

### Format of files


5. In your institution's HESA file(s), all fields, for example RECID, should retain their leading zeros, otherwise the web facility will not be able to read the information correctly. You may find if you save your files in Excel you could lose the leading zeros. To save the file(s) with leading zeros we recommend the use of a text editor such as Notepad.

6. If your institution returns student HESA data using the student and module structure, both the student and module files should be submitted to generate the re-creation.

7. The web facility can process zip files, provided that the contents are in the format specified by HESA.

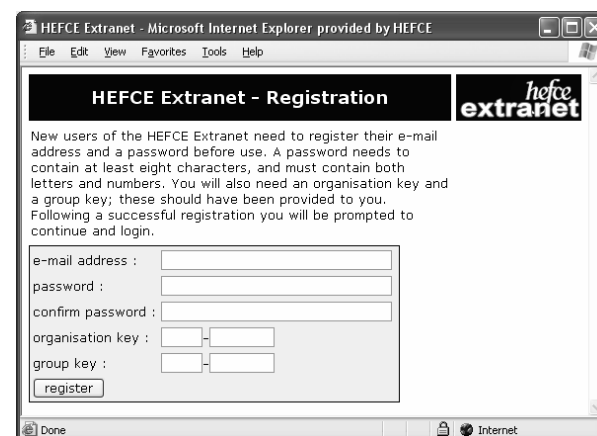
### Access to the facility

8. The web facility can be accessed via the HEFCE extranet at <https://extranet.hefce.ac.uk>. On viewing the screen below you will be prompted to either log-in or register for access to the HEFCE extranet as a new user.



### Register

9. New users of the HEFCE extranet need to register an e-mail address and extranet password before use. Select 'register' and enter your e-mail address, a password, confirmation of your password, your institution's 'organisation key' and the HEFCE web facility for 2003-04 statistics derived from HESA data 'group key', then click 'register'. The organisation key and group key are supplied in Annex A of the 'HEFCE web facility for 2003-04 statistics derived from HESA data' letter, which was sent to the head of your institution during the week of 2 August 2004. Following a successful registration you will be prompted to continue and log-in.



## Log-in

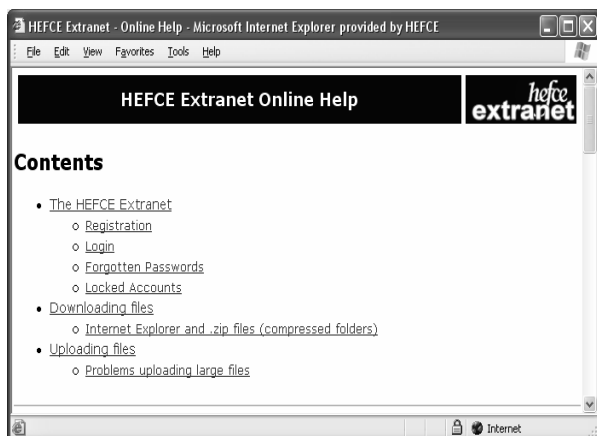
10. Once registered, users will need to log-in before each session. To do this enter your e-mail address and your extranet password, chosen during registration, then click 'Login'.

11. If you have used the HEFCE extranet for other HEFCE returns you will be required to log-in and join the HEFCE web facility for 2003-04 statistics derived from HESA data group. Follow the login procedure by entering your e-mail address and password and under 'Applications', click 'Join a group'. Enter the group key supplied in Annex A of the 'HEFCE web facility for 2003-04 statistics derived from HESA data' and select 'join group'.

12. We have recently changed each institution's organisation key. If you registered in the past but your account has expired, you will be required to re-register using the organisation key referred to in paragraph 9.

## Help

13. An information page can be accessed by clicking the 'online help' link, located above the log-in box, or on the right of the 'available resources' page, under 'Links'.



## Forgotten password

14. If you forget your password, select 'request a new one' on the log-in page. Enter your e-mail address in the space provided and click 'request password'. A new password will be generated and sent to you. You will be required to change this

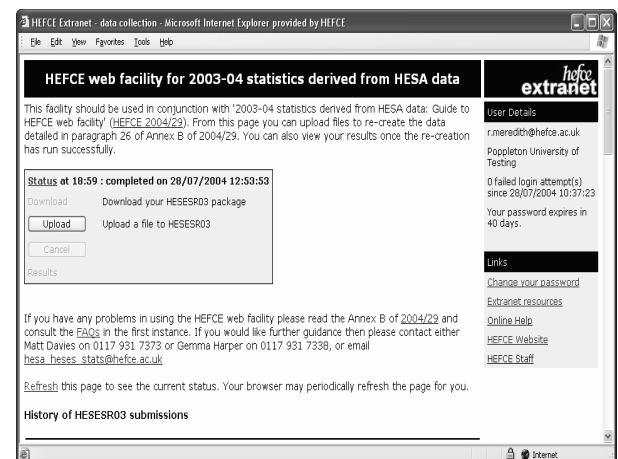
password before you can access the HEFCE extranet.

15. You have three attempts to log-in before your account will be locked. If you use all three attempts and still do not have access to the web facility, please contact Matthew Davies or Gemma Harper (e-mail [hesa\\_heses\\_stats@hefce.ac.uk](mailto:hesa_heses_stats@hefce.ac.uk)).

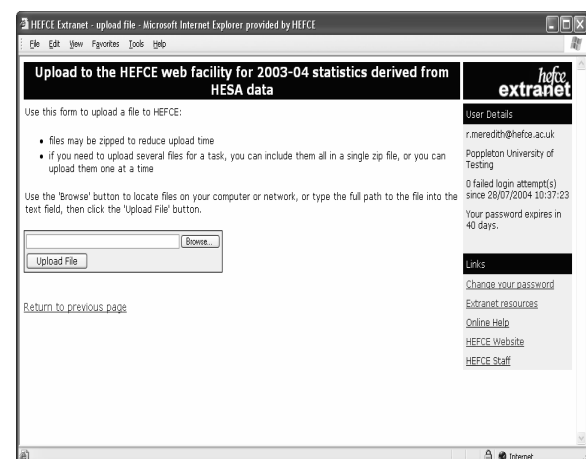
## Available resources

16. After verifying the username and password, your browser will be directed to the 'available resources' page.

17. From the 'Applications' list click on the 'HEFCE web facility for 2003-04 statistics derived from HESA data' link.



18. This option will allow you to upload HESA file(s), one at a time. Use the 'Browse' button to reach the location of the data file and once you have located the correct file, select 'Upload File'. The browse facility does not allow selection of multiple files.



Continue to add files until all HESA files have been uploaded. All files should be uploaded by repeating the process in paragraphs 17 and 18.

### Submit File(s)

19. Once all of the files have been added, select 'Submit Files'. This places the file(s) added in the previous step, into the processing queue. You can leave the extranet while the derived statistics are being generated. To do this you can exit the internet in the usual way, or select a new address in the address bar. You may be prompted to log-in again when you return to the page, but no information will be lost.

20. Each time you submit a collection of files for re-creation, you should refresh the page using either the icon in the tool bar or the link on the page. Refreshing the page ensures that the 'History of HESESR03 submissions' reflects the current stage of processing as the page may not update automatically.

21. Our systems only allow one submission to be processed at a time; therefore if more than one institution submits data, the submissions will be placed in a queue. A submission by an average sized institution will typically take 15 minutes to process (once it reaches the front of the queue). The total time taken to complete the re-creation, including waiting time, will depend on the number of institutions in the queue.

22. If the submitted files are waiting in the queue it is possible to cancel the submission by selecting 'Cancel'.

23. If you do not use the extranet for 60 minutes, you will be automatically logged out of the system. To regain entry, select any option and re-enter your e-mail address and extranet password.

### Output

24. Once the files have been processed, you can access the outputs by selecting 'Results'.

25. If the outputs were not generated, the reason for failure will be given, where the reason is known, under 'History of HESESR03 submissions'. If the outputs were not generated because an

unspecified error occurred, this will typically be a result of invalid HESA data.

### Accessing results

26. The results are made available as a zip archive (HESR03XXXX, where XXXX denotes the HESA institution identifier) containing the following files:

- HESR03XXXX.xls. This is an Excel workbook containing the HESES03 re-creation tables.
- SNCC03XXXX.xls. This is an Excel workbook containing the HESES03 re-creation based on cost centre sector norms for subjects tables.
- NSSP05XXXX.xls. This is an Excel workbook containing the provisional NSS05 target lists.
- RASR03XXXX.xls. This is an Excel workbook containing the RAS03 re-creation tables.
- WP03XXXX.xls. This is an Excel workbook containing the derived statistics likely to be used to inform 2005-06 WP funding allocations.
- CC3803XXXX.xls. This is an Excel workbook containing the derived statistics likely to be used to inform the 2005-06 transfer of sports science activity.
- PUB03XXXX.xls. This is an Excel workbook containing the derived statistics that will be used in HEFCE statistical publications.
- RDQR03XXXX.xls This is an Excel workbook containing the research degree rates of qualification. \*
- NCTK03XXXX.xls. This is the HESES04 non-completion toolkit.
- 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.

- NSSP05XXXXX.ind. This is a comma-separated file containing the derived fields that generate the provisional NSS05 target list.
- RASR03XXXXX.ind. This is a comma-separated file containing the derived fields that generate the RAS03 re-creation.
- WP03XXXXX.ind. This is a comma-separated file containing the derived fields used to generate the derived statistics that are likely to be used to inform the 2005-06 WP funding allocations.
- CC3803XXXXX.ind. This is a comma-separated file containing the derived fields that generate the derived statistics likely to be used to inform any transfer of 2005-06 sports science activity.
- PUB03XXXXX.ind. This is a comma-separated file containing the derived fields that generate the derived statistics that will be used in HEFCE statistical publications.
- RDQR03XXXXX.ind. This is a comma-separated file containing the derived fields that generate the research degree rates of qualification. \*
- POSTCXXXXX.ind. This is a comma-separated file containing the invalid postcode lists.

\* The research degree rates of qualification will be available at a later date. See the HEFCE web-site under Publications for further details.

### Using the individualised files

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

- a. Open an individualised file, from the list above 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.

### Using the output

28. 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.

These files will be overwritten each time new data are submitted. Please note that it is not possible for institutions to see one another's files.

29. At this point it may be useful to check whether the tables reflect the number of students you would expect to see. If it does not, please view the files uploaded and check that they are the files that you intended to submit. HEFCE is not responsible for the accuracy of the files submitted to the web facility.

### Technical issues

30. While there are no limits on the size of files that can be uploaded to the facility, it is common for individual institutions to have limits imposed by their network administrators. Attempts to upload a file beyond the size limit imposed within the institution are unlikely to generate a meaningful error message, it will however cause the browser to 'stall' or 'hang'.

31. If you are encountering difficulties uploading files due to file size or any other issue relating to connectivity we will normally be able to advise. We recommend that you involve your network administrator in such discussions.

32. Please note that the Internet Protocol (IP) address for this year's web facility has changed to 195.194.167.114.

# Annex C

## Comparison of HESES03 and the HESES03 re-creation

### Purpose

1. This annex describes the methods we use to generate a provisional HESES03 re-creation via the web facility. It also describes how we will use the comparison of HESES03 and the HESES03 re-creation in the 2003-04 statistics derived from HESA data exercise. The outputs are provisional because they are generated prior to submitting final data to HESA.

### Web facility outputs

2. The web facility generates a HESES03 re-creation. Annex B describes how to access the provisional HESES03 re-creation Excel workbook (HESR03XXXX.xls-where XXXX denotes the HESA institution identifier). The provisional HESES03 re-creation workbook contains the following worksheets:

Page number	Worksheet*	Description
1	CoverPage	Title page
2	FTS	Provisional HESES03 re-creation Table 1a: Full-time and sandwich years of programme of study
3	MED	Provisional HESES03 re-creation Table 1b: Medical and dental full-time and sandwich years of programme of study
4	SWOUT	Provisional HESES03 re-creation Table 2: Sandwich year-out years of programme of study
5	PT	Provisional HESES03 re-creation Table 3: Part-time years of programme of study
6	FEE	Provisional HESES03 re-creation Table 4: Home and EC fees
7	HBK	Provisional HESES03 re-creation grant adjustment report
8	STD	Provisional HESES03 re-creation standard resource table
9	F03	Provisional HESES03 re-creation assumed fee income table
10	Excl	Provisional HESES03 re-creation exclusion table
11	Summary	Provisional HESES03 re-creation comparison summary table
12	SummaryPG	Provisional HESES03 re-creation price group comparison summary table
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 adjustment report
19	hSTD	HESES03 standard resource table
20	hF03	HESES03 assumed fee income table

\* For worksheet reference see spreadsheet tabs

The workbook also contains a version number that will increment by one following each successful submission of data to the web facility, as well as contact details for any queries relating to data contained in the workbook.



3. The web facility will also generate a provisional HESES03 re-creation individualised file (HESR03XXXX.ind) to supplement the provisional HESES03 re-creation. The file 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.

## **Methods**

4. The algorithms used to generate the provisional HESES03 re-creation are described in Appendix 1.

5. Appendix 2 contains troubleshooting guidance on the causes of differences between HESES03 and the provisional HESES03 re-creation.

6. All known problems of fit with the provisional HESES03 re-creation algorithms are described in Appendix 3.

## **2003-04 statistics derived from HESA data exercise**

7. In early 2005 we will generate a HESES03 re-creation from 2003-04 student data submitted to the HESA, and compare it to HESES03. We intend to use the algorithms described in Appendix 1 to generate the HESES03 re-creation; however should we identify an error in our algorithms, changes will be made to correct the error.

8. We will select institutions to respond to the comparison on the basis of thresholds. The thresholds are likely to be set in terms of funding differences arising from HESES03 and the HESES03 re-creation. Notwithstanding the thresholds, we may also ask for further information from any institution in respect of the comparison.

9. Each institution requested to make a response will be asked to reconcile the two data sources. After reconciling the two data sources and confirming that the HESES03 re-creation reasonably reflects the outturn position for 2003-04, the HESES03 re-creation will supersede HESES03 and any consequent grant adjustments

will be made (subject to the appeals process and the availability of funds).

# Annex D

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

### Purpose

1. This annex describes the methods we use to generate a provisional HESES03 re-creation based on cost centre sector norms for subjects via the web facility. It also describes how we will use the comparison of HESES03, the HESES03 re-creation and the HESES03 re-creation based on cost centre sector norms for subjects in the 2003-04 statistics derived from HESA data exercise. The outputs are provisional because they are generated prior to submitting final data to HESA.

### Web facility outputs

2. The web facility generates a HESES03 re-creation based on cost centre sector norms for subjects. Annex B describes how to access the provisional HESES03 re-creation based on cost centre sector norms for subjects Excel workbook (SNCC03XXXX.xls - where XXXX denotes the HESA institution identifier). The provisional HESES03 re-creation based on cost centre sector norms for subjects workbook contains the following worksheets:

Page number	Worksheet*	Description
1	CoverPage	Title page
2	snFTS	Provisional HESES03 re-creation based on cost centre sector norms for subjects Table 1a: Full-time and sandwich years of programme of study
3	snMED	Provisional 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
4	snSWOUT	Provisional HESES03 re-creation based on cost centre sector norms for subjects Table 2: Sandwich year-out years of programme of study
5	snPT	Provisional HESES03 re-creation based on cost centre sector norms for subjects Table 3: Part-time years of programme of study
6	snFEE	Provisional HESES03 re-creation based on cost centre sector norms for subjects Table 4: Home and EC fees
7	snHBK	Provisional HESES03 re-creation based on cost centre sector norms for subjects grant adjustment report
8	snSTD	Provisional HESES03 re-creation based on cost centre sector norms for subjects standard resource table
9	snF03	Provisional HESES03 re-creation based on cost centre sector norms for subjects assumed fee income table
10	snExcl	Provisional HESES03 re-creation based on cost centre sector norms for subjects exclusion table
11	snSummary	Provisional HESES03 re-creation based on cost centre sector norms for subjects comparison summary table
12	snSummaryPG	Provisional HESES03 re-creation based on cost centre sector norms for subjects price group comparison summary table
13	FTS	Provisional HESES03 re-creation Table 1a: Full-time and sandwich years of programme of study
14	MED	Provisional HESES03 re-creation Table 1b: Medical and dental full-time and sandwich

		years of programme of study
15	SWOUT	Provisional HESES03 re-creation Table 2: Sandwich year-out years of programme of study
16	PT	Provisional HESES03 re-creation Table 3: Part-time years of programme of study
17	FEE	Provisional HESES03 re-creation Table 4: Home and EC fees
18	HBK	Provisional HESES03 re-creation grant adjustment report
19	STD	Provisional HESES03 re-creation standard resource table
20	F03	Provisional HESES03 re-creation assumed fee income table

\* For worksheet reference see spreadsheet tabs

The workbook also contains a version number that will increment by one following each successful submission of data to the web facility, as well as contact details for any queries relating to data contained in the workbook.

3. The web facility will also generate a provisional HESES03 re-creation based on cost centre sector norms for subjects individualised file (SNCC03XXXX.ind) to supplement the provisional HESES03 re-creation based on cost centre sector norms for subjects. The file 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 4.

## Methods

4. The algorithms used to generate the provisional HESES03 re-creation based on cost centre sector norms for subjects are described in Appendix 4.

5. Appendix 5 contains troubleshooting guidance to identify the causes of differences between HESES03, the provisional HESES03 re-creation and the provisional HESES03 re-creation based on cost centre sector norms for subjects, especially where they are attributable to errors in HESA data.

6. All known problems of fit with the provisional HESES03 re-creation based on cost centre sector norms for subjects algorithms are described in Appendix 6.

## 2003-04 statistics derived from HESA data exercise

7. In early 2005 we will generate a HESES03 re-creation based on cost centre sector norms for subjects from 2003-04 student data submitted to the HESA and compare it to the HESES03 re-creation. We intend to use the algorithms described in Appendix 4 to generate the HESES03 re-creation based on cost centre sector norms for subjects, except that the sector norm mapping of subjects to cost centres will be revised on receipt of full HESA 2003-04 student data. The web facility uses the cost centre sector norms generated from HESA 2002-03 student data, and we do not expect the sector norms for 2003-04 to differ significantly from them. If we identify an error in our algorithms however, they will be corrected.

8. We will select institutions to respond to the comparison on the basis of thresholds. The thresholds are likely to be 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. Notwithstanding the threshold, we may also ask for further information from any institution in respect of the comparison.

9. Each institution requested to make a response will be asked to reconcile their HESES03, HESES03 re-creation and the HESES03 re-creation based on cost centre sector norms for subjects. After reconciling all three and confirming that the

HESES03 re-creation reasonably represents the outturn position for 2003-04, the HESES03 re-creation will supersede HESES03 and any consequent grant adjustments will be made (subject to the appeals process and the availability of funds).

# Annex E

## National Student Survey 2005 – provisional target list

### Purpose

1. This annex describes how we will use HESA 2003-04 student data submitted over the web facility to provide HEIs in England and Northern Ireland (the Welsh funding council is proposing the same for HEIs in Wales) with a provisional list of students that should be included in the 2005 National Student Survey (NSS05). The definitive list will be confirmed in December 2004 based on final HESA 2003-04 student data.
2. This annex also describes how we will use institutions' HESA 2003-04 student data to generate an individualised student data file that identifies, by reason, the students that have been excluded from the provisional NSS05 target list.
3. The algorithms used to generate the list are explained in Appendix 7.
4. Our proposals for conducting the first NSS in 2005 are set out in 'National Student Survey 2005: Consultation' (HEFCE 2004/22). More information will be available in October 2004.
5. Institutions are encouraged to use the provisional target lists to start preparing the contact details for students. We will not be able to provide definitive target lists until we have received final data from HESA in mid-December; completed lists will need to be returned to the agency in early January. To help institutions meet this deadline we will retain the final copy of the target list generated by the web facility so that we can provide details of the changes in coverage between institutions' final use of the web facility and the generation of lists from signed off HESA data. We do not expect to make any changes to the algorithms for generating target lists, therefore any changes will arise from changes to institutions' data.
6. HEFCE staff will access data relating to national student survey target lists. Further details about the data confidentiality agreement can be found in paragraph 31 of the Introduction of the main text.

### Background

7. The NSS is a new survey that has been introduced to provide information to prospective students and their advisors who are making choices relating to what and where to study. This information is to be provided together with other information relating to teaching quality. It also aims to contribute to public accountability, and to provide information that will help institutions to enhance teaching quality, by supplementing internal feedback mechanisms.

### Web facility outputs

8. The web facility generates the provisional NSS05 target lists. Annex B describes how to access the provisional NSS05 target lists (NSSP05XXXX.xls – where XXXX denotes the HESA institution identifier).

### Proposed scope of students to be included

9. For the 2005 survey, we propose that the population of students that will be surveyed comprises all full-time and part-time undergraduate students registered at HEIs in England and Northern Ireland, who are at or near the point of completion of their programmes of study. We propose the selection of students as follows:
  - a. Where the course length is known, students are included if the course is expected to end in 2004-05 (that is, they are expected to be in their final year of study).
  - b. Where the course length is not known, students are included in the survey in their fourth year of study or if they are recorded as leaving within 2004-05.

Thus not all students will be in their final year, but will have had significant experience at the institution (at least one year FTE over four years).

10. We expect all students that are registered at one institution and taught by another institution, international students, and independently-funded students to be included in the population. The

Department of Health is currently considering whether students on NHS-funded courses should also be included.

11. The following students will be excluded from the population surveyed:

- a. Students on programmes that do not lead to undergraduate qualifications or credits.
- b. Students on short courses (of less than one FTE) and on low intensity courses.
- c. Students at FECs directly funded by HEFCE or FECs that are part of a HEFCE-recognised funding consortium.
- d. Students on courses funded by the Teacher Training Agency.

# Annex F

## Comparison of RAS03 and the RAS03 re-creation

### Purpose

1. This annex describes the methods we use to generate a provisional RAS03 re-creation via the web facility. It also describes how we will use the comparison of RAS03 and the RAS03 re-creation in the 2003-04 statistics derived from HESA data exercise. The outputs are provisional because they are generated prior to submitting final data to HESA.

### Web facility outputs

2. The web facility generates a RAS03 re-creation. Annex B describes how to access the provisional RAS03 re-creation Excel workbook (RASR03XXXX.xls - where XXXX denotes the HESA institution identifier). The provisional RAS03 re-creation workbook contains the following worksheets:

Page number	Worksheet*	Description
1	CoverPage	Title page
2	R1a	Provisional RAS03 re-creation Form R1a: All full-time research students by year of programme
3	R1b	Provisional RAS03 re-creation Form R1b: All part-time research students by year of programme
5	R2a	Provisional RAS03 re-creation Form R2a: Fundable Home and EC fee paying full-time research students by year of programme
6	R2b	Provisional RAS03 re-creation Form R2b: Fundable Home and EC fee paying part-time research students by year of programme
8	Funds	Provisional 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 table
10	Summary	Provisional RAS03 re-creation summary table
11	Summary by subject groupings	Summary comparison of RAS03 and the provisional 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 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 tab.

The workbook also contains a version number that will increment by one following each successful submission of data to the extranet, as well as contact details for any queries relating to data contained in the workbook.

3. The web facility will also generate a provisional RAS03 re-creation individualised file (RASR03XXXX.ind) to supplement the provisional RAS03 re-creation. The file 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 8.

### **Methods**

4. The algorithms used to generate the provisional RAS03 re-creation are described in Appendix 8.

5. Appendix 9 contains troubleshooting guidance for the causes of differences between RAS03 and the provisional RAS03 re-creation, especially where they are attributable to errors in HESA data.

6. All known problems of fit with the provisional RAS03 re-creation algorithms are described in Appendix 10.

### **2003-04 statistics derived from HESA data exercise**

7. In early 2005 we will generate a RAS03 re-creation from 2003-04 student data submitted to the HESA and compare it to RAS03. We intend to use the algorithms described in Appendix 8 to generate the RAS03 re-creation, except that the assignment of units of assessment (UoAs) to subject groups will be revised on receipt of full HESA 2003-04 student data. In the absence of these data and for the purpose of the web facility we will use the assignment of UoAs to subject groups generated from HESA 2002-03 student data, however we do not expect the assignment to change significantly. Should we identify an error in our algorithms changes will be made to correct the error.

8. We will select institutions to respond to the comparison on the basis of thresholds. Initially, the thresholds are likely to be set in terms of differences in the FTEs at the broad subject group level arising from the RAS03 and the RAS03 re-creation. However, these FTE differences will also have consequent funding adjustments associated with them. Notwithstanding the thresholds, we may also ask for further information from any institution in respect of the comparison.

9. Each institution requested to make a response will be asked to reconcile the two data sources. After reconciling the two data sources and confirming that the RAS03 re-creation reasonably reflects the outturn position for 2003-04, 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).



# Annex G

## Derived statistics used to inform the 2005-06 widening participation funding allocations

### Purpose

1. This annex describes how we propose to use HESA student data to inform the widening participation (WP) allocations for 2005-06. The web facility offers institutions the opportunity to check and, where errors would affect these allocations under current proposals, correct their data prior to submission to HESA.
2. WP allocations consist of three separate elements: widening access for students from disadvantaged backgrounds, widening access for disabled students, and improving retention. Details of the algorithms used are given in Appendix 11.
3. HESA 2002-03 student data informed the following widening access allocations for 2004-05:
  - widening access for students from disadvantaged backgrounds
  - widening access for disabled students.
4. The methods that will be used to calculate the improving retention allocations based upon qualifications on entry and age have yet to be determined.

### Web facility outputs

5. The web facility generates the derived statistics likely to be used to inform 2005-06 WP funding allocations. Annex B describes how to access the derived statistics likely to be used to inform the 2005-06 WP funding allocations Excel workbook (WP03XXXX.xls – where XXXX denotes the HESA institution identifier).

### Widening access for students from disadvantaged backgrounds

#### Young and mature

6. For the widening access allocation we define a student as young if they are under 21 on entry to the programme of study, otherwise the student is defined as mature.

### Ward data

7. We use post office data to map student postcodes to census wards. Postcodes that are recognised as schools, prisons, hospitals and similar public institutions are excluded from this mapping. A full list of excluded postcodes can be found on the HEFCE web-site under Learning & teaching/Data collection/ HEFCE web facility for 2003-04 statistics derived from HESA data.

### Young participation by 1991 census ward

8. We use a rolling average of 1997-1999 child benefit data to identify the total number of 15 year-olds as our estimate of potential entrants to higher education in 1999-2002. This figure is then divided by the actual number of entrants in 1999-2002. The results can be found on the participation of local areas (POLAR) web-site. Details of how to access this site are given through the HEFCE web-site at [www.hefce.ac.uk/research/polar](http://www.hefce.ac.uk/research/polar).

### Educational attainment by 2001 census ward

9. We use 2001 census data to identify the percentage of 16-74 year-olds within each ward that have a qualification of at least degree level. The underlying data can be found on the ONS website, [www.nationalstatistics.gov.uk](http://www.nationalstatistics.gov.uk), under Neighbourhood/Statistics by subject/Education, Skills and Training/Qualifications and Students, 2001/2003 Administrative Areas - National to Ward level.

### Young full-time students

10. For young full-time, home-domiciled (English, Scottish, Welsh and Northern Irish) HEFCE-funded undergraduates, we make the following calculation:

- a. Assign students to wards excluding students with excluded postcodes.
- b. Assign a weighting to each student based on the participation rate of their ward. The weightings for the participation quintiles are as given in Table 1.

Table 1 **Weightings for participation quintiles**

Participation quintile	Weighting
1	2
2	1
3, 4, 5	0

**Mature full-time and part-time students**

11. We make the following calculation for part-time and mature full-time English and Welsh domiciled HEFCE-funded undergraduates:

- a. Assign students to wards excluding students with excluded postcodes.
- b. Assign a weighting to each student based on the educational attainment rate of their ward. The weightings for each quintile are given in Table 2.
- c. Assign a zero weight to students who have unknown entry qualifications or already have an equivalent or higher qualification than their current qualification aim. The hierarchy of qualification aims is outlined in paragraph 15 of Appendix 11.

Table 2 **Weightings for educational attainment quintiles**

Educational attainment quintiles	Weighting
1	2
2	1
3, 4, 5	0

**Full-time students**

12. We calculate the average weighting for full-time students, both young and mature, with non-excluded postcodes and apply this average weight to the London weighted assumed full-time and sandwich FTEs for 2005-06. Funding is distributed pro-rata to these weighted FTEs.

**Part-time students**

13. We calculate the average weightings for part-time undergraduates, and apply this average weight to the London weighted part-time assumed FTEs for 2004-05. Funding is distributed pro-rata to these weighted FTEs.

**Widening access for students with disabilities**

14. The disabled students allocation is calculated as follows:

- a. The proportion of home and EC students who were in receipt of the Disabled Students Allowance (DSA) is calculated.
- b. Each institution is assigned to a weighting quartile according to the proportion of students in receipt of the DSA. The quartile for each institution is moderated to ensure that no institution drops by more than one quartile compared to last year. The weightings are given in Table 3.
- c. The resulting weights are then applied to the London weighted assumed FTEs for 2005-06 to determine the institution's pro rata share of the funding available.

Table 3 **Weighting quartile according to the proportion of students in receipt if the DSA**

Participation rate	Weighting
A (lowest proportion)	1
B	2
C	3
D	4

# Annex H

## Derived statistics used to inform any 2005-06 transfer of sports science activity

### Purpose

1. This annex describes how we will use HESA 2003-04 student data to identify sports science and leisure studies activity for the 2005-06 teaching funding allocation. Details of the algorithms used to generate the data are given in Appendix 12.

### Web facility outputs

2. The web facility generates the derived statistics used to inform the 2005-06 transfer of sports science activity. Annex B describes how to access the derived statistics used to inform the 2005-06 transfer of sport science activity Excel workbook (CC3803XXXX.xls – where XXXX denotes the HESA institution identifier).

### Identifying sports science and leisure studies activity

3. The identification of sports science and leisure studies students is calculated as follows:

- a. For students generating HEFCE-funded countable years in Column 4 of the HESES03 re-creation, we identify the proportion of their year of programme of study attributed to sports science and leisure studies (cost centre 38).
- b. For each student, a sports science FTE is generated by applying the proportion derived in step a (above) to their FTE.
- c. Sports science FTE is summarised over mode, level and length for each institution.
- d. We identify the proportion of HESES03 re-creation HEFCE-funded FTE in Column 4a that is sports science FTE, for each combination of mode, price group, level and length.

# Annex I

## HEFCE statistical publications

### Purpose

1. This annex describes the data we intend to publish, or use to inform policy decisions, during 2005.
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’)
  - campus data
  - distance learning data
  - provision by location.

These data may also be used to inform any allocation of growth in foundation degree places.

3. Full details of the algorithms used in generating the data for publication are given in Appendix 13. Where appropriate we have included derived fields on the individualised student data file, to allow institutions to better understand the derivation of the data.
4. The basis for deriving statistics for HEFCE statistical publications has fundamentally changed since HEFCE 2003/39. We now base the population definition on the HESA population definition rather than the HESSES population definition. Therefore in Appendix 13 we detail the algorithms that will be used to generate published data in 2005 for each of the four categories above.

### Web facility outputs

5. The web facility generates the derived statistics that will be used in HEFCE statistical publications. Annex B describes how to access the derived statistics that will be used in HEFCE statistical publications Excel workbook (PUB03XXXX.xls - where XXXX denotes the HESA institution identifier).

### Students registered at one institution and taught by another institution

6. In HEFCE 2004/02 we published details of students registered at one institution and taught by another institution, using HESA 2001-02 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 should 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 campus for each campus of the merged institutions. 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, and 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. Last year 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. 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 the student is taught as described in paragraph 8 above. Where neither information about students that are registered at one institution and taught by another institution, or campus information 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, local authority district/unitary authority, local education authority and the LSC area.

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

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

# Annex J

## Research degree rates of qualification

This annex is not yet available.

# Annex K

## HESES04 non-completion toolkit

### Purpose

1. This annex contains details on how to use the non-completion toolkit.

### Description of the non-completion toolkit

2. The non-completion toolkit is a spreadsheet that allows institutions to generate estimated non-completion rates of HEFCE-funded students at a course and year level for 2003-04. The estimated values generated may be used by institutions to inform the forecasts of non-completions returned in HESES04. The facility enables the user to create estimated non-completion rates using a selection of fields that they regard influence the non-completion rate for their institution. Non-completion rates are derived from the submitted 2003-04 HESA student data.

### Reliability of estimated non-completion rates

3. Institutions should note that the toolkit is designed purely as an aid for institutions to generate their own estimations of non-completions in HESES04. We expect institutions to use their own judgement in assessing the applicability of non-completion rates derived from the toolkit.

4. Estimates may be unreliable if the numbers of students involved are small or if the underlying HESA data contains errors. Institutions should consult paragraphs 14-21, Annex D of HEFCE 2003/44 for definitive guidance on non-completions.

### Using the non-completion toolkit

5. Users should follow the instructions in Annex B to access the Excel file NCTK03XXXX.xls, then open it in Excel2000 or a later version of Excel. Due to the technology used the tool will not function correctly in earlier versions of Excel. The creation of tables uses Excel macros and these will need to be enabled to use the facility. Excel will prompt you to 'Enable macros'. The Excel workbook contains a worksheet called NCDData, which contains the raw data used to generate the tables, and a further sheet, Pivot Table, for

displaying the results. When you open the file you will be prompted to select the fields you want to summarise the non-completion rate by. We advise summarising the data by course in the first instance. To create the table click 'OK', the screen will close enabling you to view the table. To create another table click on 'Create a new table'. The previous table will be removed from the worksheet.

6. Non-completion rates based on small numbers of students may be unreliable. Therefore, if the calculated non-completion rate is based on a total population which is less than 100 it is coloured red and a note is placed in the cell indicating the unreliability of the rate. In these cases we advise using less detailed data in order to obtain a reliable estimate of the non-completion rate.

7. In order to save a table you have created, open the <File> menu and select <Save As> and select the folder you want to save the file in, then rename the file, and click 'Save'.

# Annex L

## Invalid postcode list

### Purpose

1. This annex provides details of the postcode validation data file, generated by the web facility, which identifies those students returned on the 2003-04 HESA student record with postcodes that are not contained within the Post Office list of valid postcodes.
2. A description of the fields used to generate this list are detailed in Appendix 16.

### Web facility outputs

3. The web facility generates the invalid postcode list. Annex B describes how to access the invalid postcode list comma-separated file (POSTCXXXX.ind -where XXXX denotes the HESA institution identifier).

### Comparison

4. The web facility generates a comma-separated file identifying each student in the HESA 2003-04 student return who is recorded with an invalid postcode as defined above. The invalid postcode is displayed alongside linking fields for the student. Details of how to access the file are provided in paragraph 26 of Annex B.
5. We have created a list of all valid postcodes using information contained in the All Fields Postcode Directory (AFPD) that is provided by the Office of National Statistics.
6. We compare this list with the postcode of each student returned on the HESA 2003-04 student record, and those students with postcodes not included in our list are identified in the postcode validation file.
7. The postcode validation file is provided to aid institutions identify invalid postcodes within the provisional HESA 2003-04 return, and allow them to correct these postcodes prior to submitting the finalised data to HESA.



# Annex M

## List of abbreviations

<b>Acronym</b>	<b>Description</b>
AFPD	All Fields Postcode Directory
ASN	Additional student numbers
CFTE	Contract full-time equivalent
DSA	Disabled Student's Allowance
EC	European Community
FAQ	Frequently asked question
FE	Further education
FEC	Further education college
FTE	Full-time equivalent
HE	Higher education
HEI	Higher education institution
HESA	Higher Education Statistics Agency
HESES	Higher Education Students Early Statistics
HIN	HUSID X INSTID X NUMHUS
INSET	In-service education for teachers
ILR	Individualised learner record
ITT	Initial teacher training
JACS	Joint Academic Coding System
LSC	Learning and Skills Council
NSS	National Student Survey
ONS	Office for National Statistics
PGCE	Postgraduate certificate of education
QTS	Qualified teacher status
RAS	Research Activity Survey
UoA	Unit of Assessment
USS	Universities Superannuation Scheme
WP	Widening participation