Appendix 2

Troubleshooting the differences between HESES03 and the HESES03 re-creation

Purpose

1. This appendix aims to help institutions identify the cause of any discrepancies between their HESA student data and HESES03 return. It is expected that institutions will have worked through this appendix and consulted the web-based FAQ page on the HEFCE web-site under Learning & teaching/Data collection before seeking assistance from HEFCE on resolving discrepancies.

Using this appendix

2. Figure 13 provides a systematic method for identifying at what point discrepancies between the returns occur. The subsequent paragraphs give possible causes for each discrepancy. These causes can be grouped into two categories:

- errors in completing specific fields on the HESA return (addressed in this appendix)
- problems of fit with the HESES03 re-creation algorithms (addressed in Appendix 3).

3. The diagnostic diagram in Figure 13 can be used to help identify errors in completing specific fields on the HESA return.

4. The match between HESES and HESA data is unlikely to be exact, due to estimates made when returning HESES and approximations made in the re-creation algorithms (see Appendix 3 for further details). Therefore, when using the diagnostic diagram we expect institutions to exercise their own judgement to decide when small differences between the two data sources are not significant. However, institutions need to be aware that small differences may accumulate and become significant. When the cause of a significant difference cannot be determined, it may be necessary to backtrack to find the root of the problem.

5. The description in this appendix is aimed at identifying weaknesses in the HESA data rather than HESES. Therefore it is possible to follow the diagnostic diagram given in Figure 13 without resolving the discrepancies if they are due to errors in HESES03.

6. Throughout this appendix, fields taken from the HESA return or derived as part of the re-creation are shown in capitals using the names given in Tables 3 and 4 of Appendix 1.

Using the individualised file

7. When working through this appendix it is necessary to use the individualised file HESR03XXXX.ind, where XXXX is the HESA identifier for the institution. Full details of how to access this file are given in Annex G. This will show the allocation of students to cells within the tables and, where relevant, details of why they were excluded. 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.

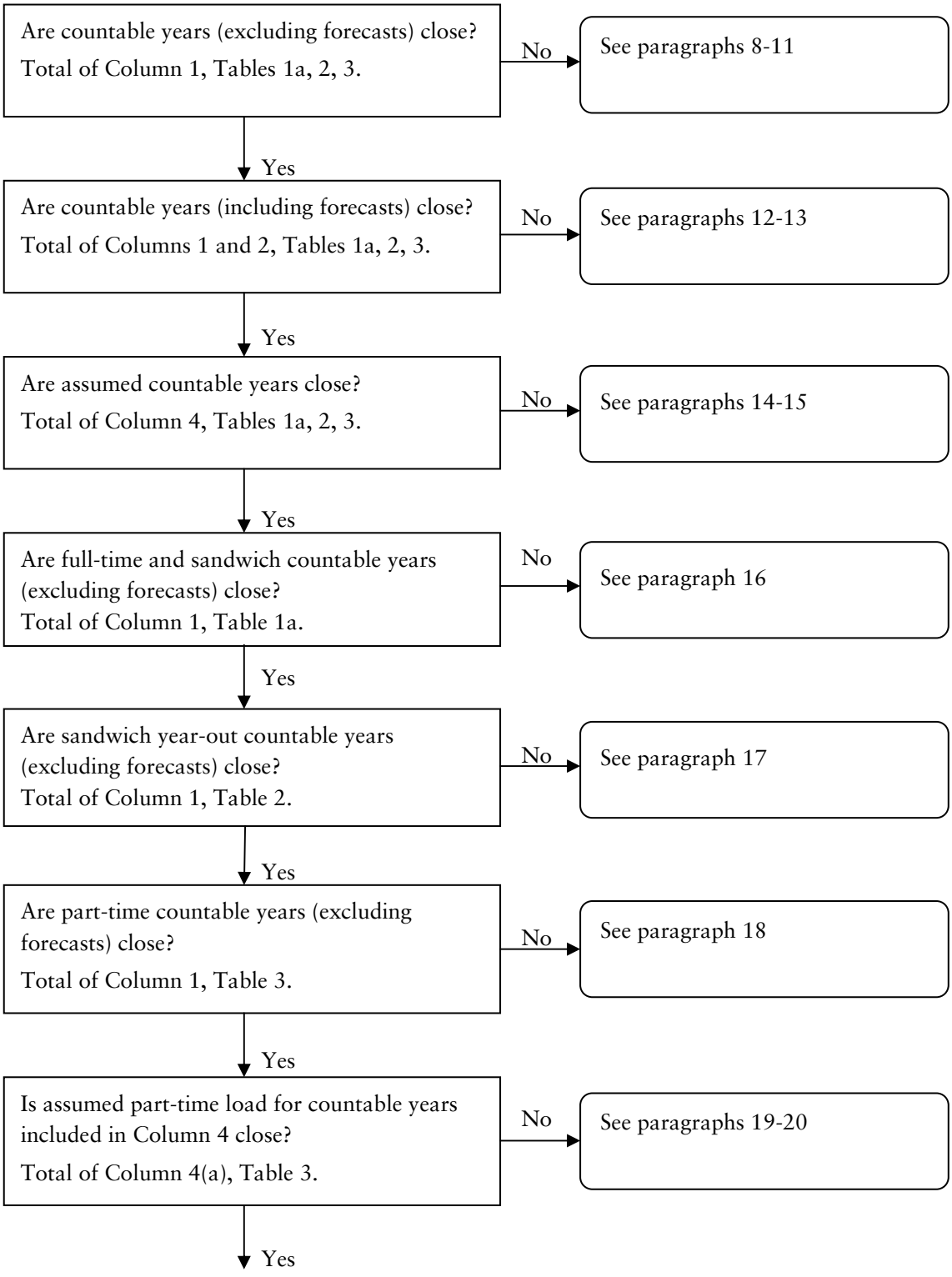
c. Select <Filter> from the <Data> menu and then <Autofilter>.

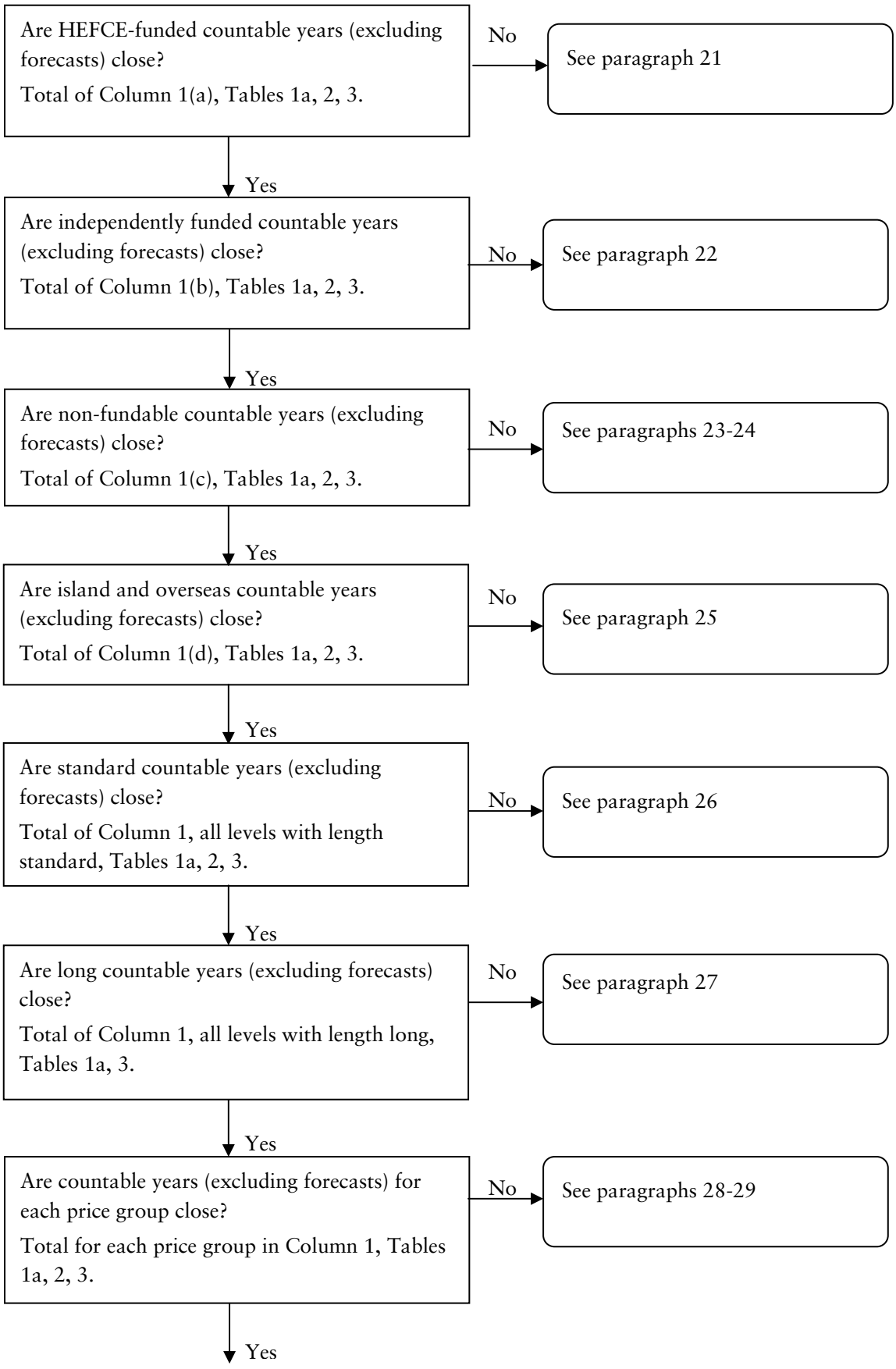
d. Click on the arrow in the column containing the data which you want to filter.

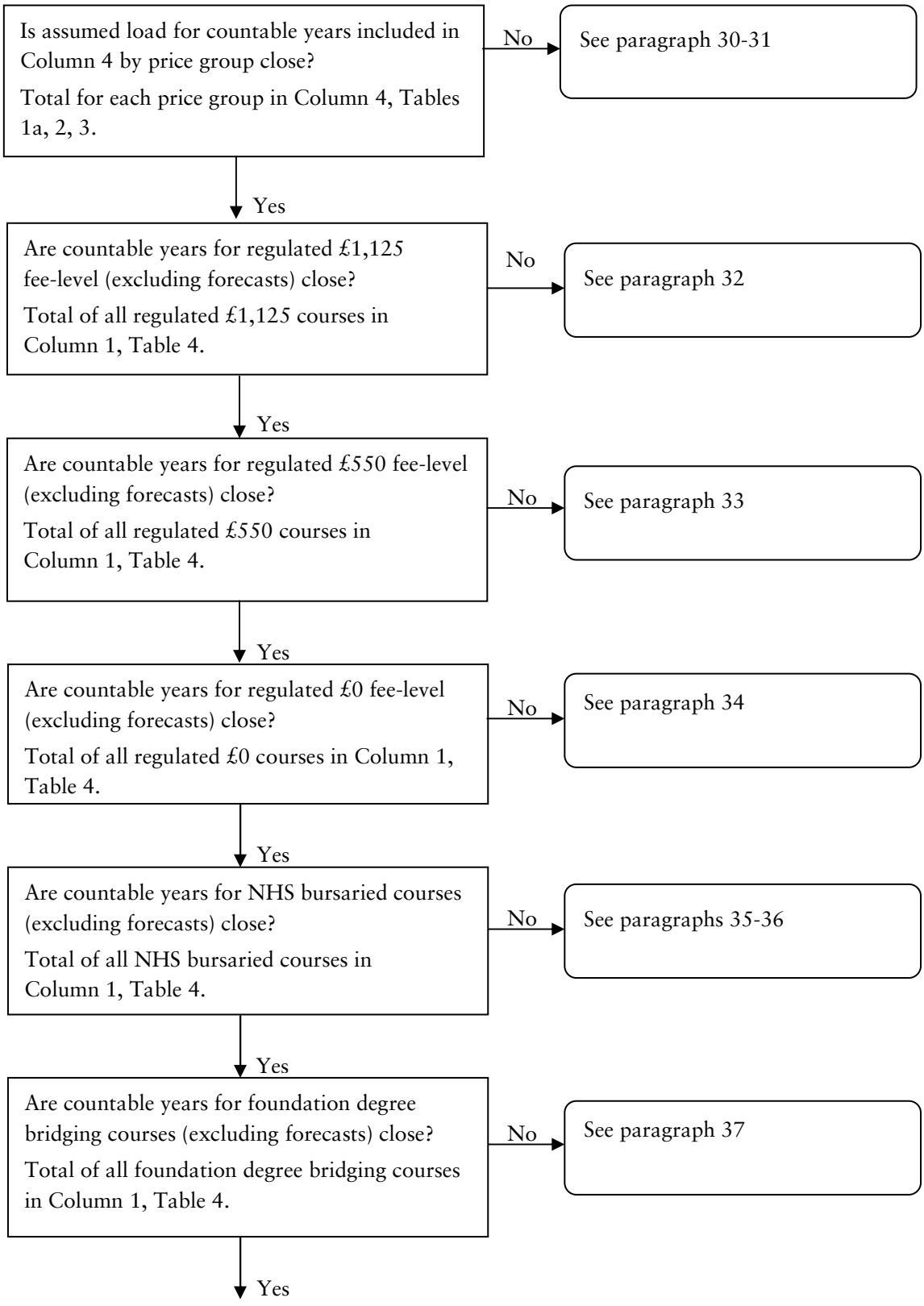
e. Either select a specific value or select <Custom> to apply a comparison operator other than equality.

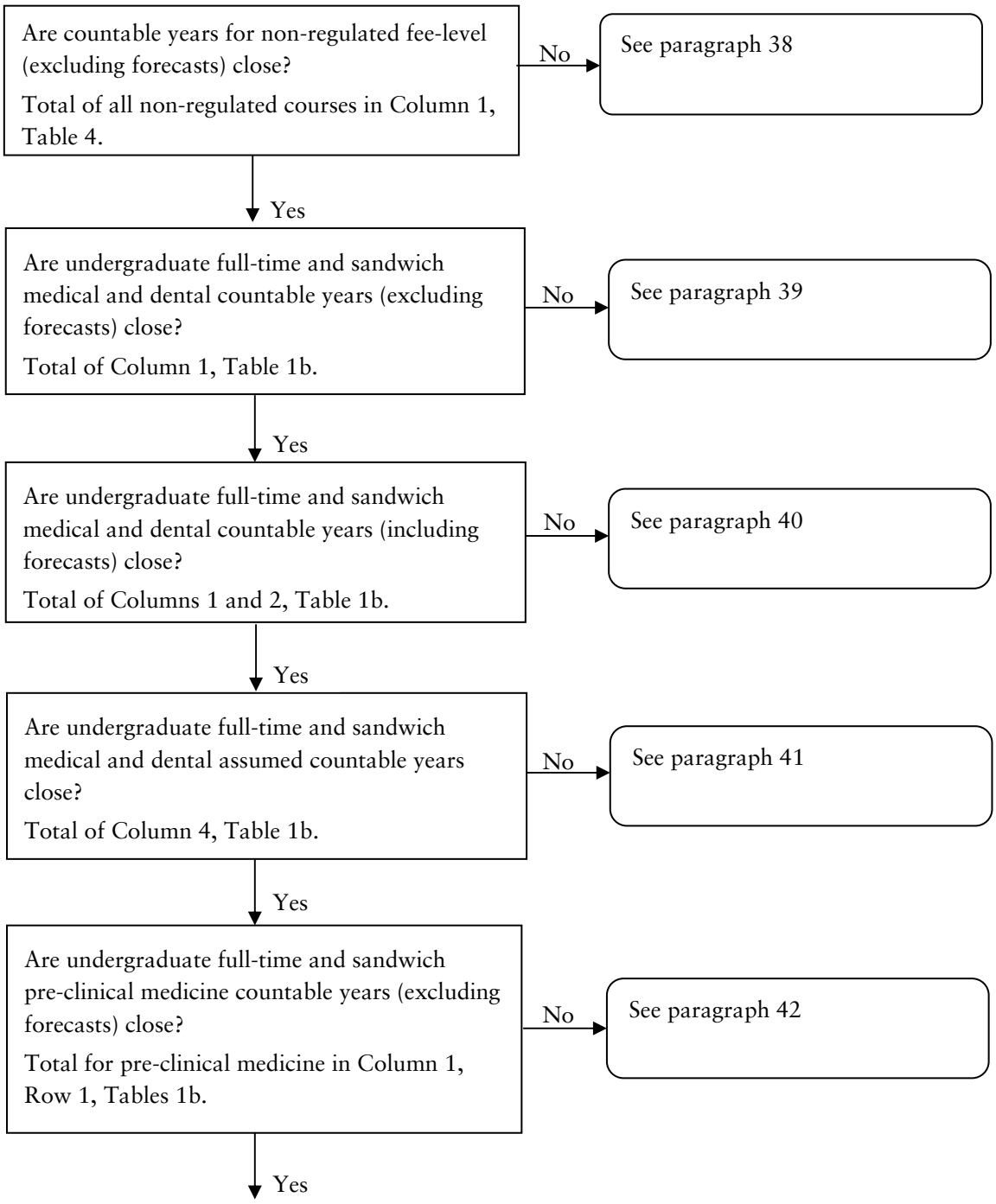
f. To select records using multiple fields, repeat steps d and e for each field.

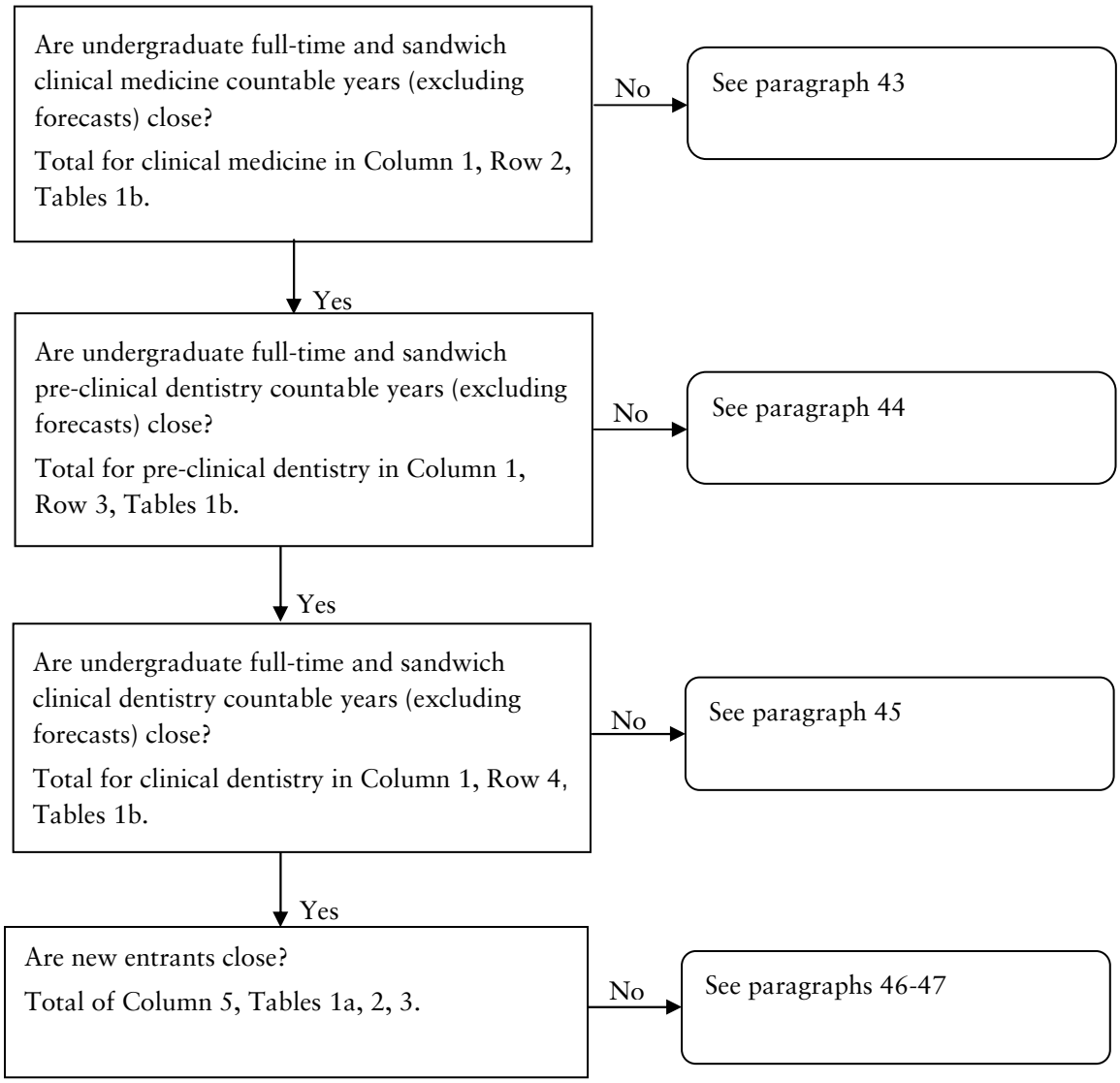
Figure 13 **Diagnostic flowchart**




Are HEFCE-funded countable years (excluding forecasts) close?
Total of Column 1(a), Tables 1a, 2, 3.
No
See paragraph 21
Yes
Are independently funded countable years (excluding forecasts) close?
Total of Column 1(b), Tables 1a, 2, 3.
No
See paragraph 22
Yes
Are non-fundable countable years (excluding forecasts) close?
Total of Column 1(c), Tables 1a, 2, 3.
No
See paragraphs 23-24
Yes
Are island and overseas countable years (excluding forecasts) close?
Total of Column 1(d), Tables 1a, 2, 3.
No
See paragraph 25
Yes
Are standard countable years (excluding forecasts) close?
Total of Column 1, all levels with length standard, Tables 1a, 2, 3.
No
See paragraph 26
Yes
Are long countable years (excluding forecasts) close?
Total of Column 1, all levels with length long, Tables 1a, 3.
No
See paragraph 27
Yes
Are countable years (excluding forecasts) for each price group close?
Total for each price group in Column 1, Tables 1a, 2, 3.
No
See paragraphs 28-29
Yes


Is assumed load for countable years included in Column 4 by price group close?
Total for each price group in Column 4, Tables 1a, 2, 3.
No
See paragraph 30-31
Yes
Are countable years for regulated £1,125 fee-level (excluding forecasts) close?
Total of all regulated £1,125 courses in Column 1, Table 4.
No
See paragraph 32
Yes
Are countable years for regulated £550 fee-level (excluding forecasts) close?
Total of all regulated £550 courses in Column 1, Table 4.
No
See paragraph 33
Yes
Are countable years for regulated £0 fee-level (excluding forecasts) close?
Total of all regulated £0 courses in Column 1, Table 4.
No
See paragraph 34
Yes
Are countable years for NHS bursaried courses (excluding forecasts) close?
Total of all NHS bursaried courses in Column 1, Table 4.
No
See paragraphs 35-36
Yes
Are countable years for foundation degree bridging courses (excluding forecasts) close?
Total of all foundation degree bridging courses in Column 1, Table 4.
No
See paragraph 37
Yes


Are countable years for non-regulated fee-level (excluding forecasts) close?
Total of all non-regulated courses in Column 1, Table 4.
No
See paragraph 38
Yes
Are undergraduate full-time and sandwich medical and dental countable years (excluding forecasts) close?
Total of Column 1, Table 1b.
No
See paragraph 39
Yes
Are undergraduate full-time and sandwich medical and dental countable years (including forecasts) close?
Total of Columns 1 and 2, Table 1b.
No
See paragraph 40
Yes
Are undergraduate full-time and sandwich medical and dental assumed countable years close?
Total of Column 4, Table 1b.
No
See paragraph 41
Yes
Are undergraduate full-time and sandwich pre-clinical medicine countable years (excluding forecasts) close?
Total for pre-clinical medicine in Column 1, Row 1, Tables 1b.
No
See paragraph 42
Yes


Are undergraduate full-time and sandwich clinical medicine countable years (excluding forecasts) close?
Total for clinical medicine in Column 1, Row 2, Tables 1b.
No
See paragraph 43
Yes
Are undergraduate full-time and sandwich pre-clinical dentistry countable years (excluding forecasts) close?
Total for pre-clinical dentistry in Column 1, Row 3, Tables 1b.
No
See paragraph 44
Yes
Are undergraduate full-time and sandwich clinical dentistry countable years (excluding forecasts) close?
Total for clinical dentistry in Column 1, Row 4, Tables 1b.
No
See paragraph 45
Yes
Are new entrants close?
Total of Column 5, Tables 1a, 2, 3.
No
See paragraphs 46-47

Countable years (excluding forecasts)

8. To identify countable years (excluding forecasts) from the individualised file select HESEXCL = 0 and HESREG = 1. The algorithms for deriving HESEXCL and HESREG are given in paragraphs 72-73 and 69 respectively of Appendix 1.

9. Exceptionally a student generates two countable years of programme of study on the HESES return (see paragraph 11 of Annex D of 'Higher Education Students Early Statistics Survey 2003-04' (HEFCE 2003/44) for further details). The field for identifying multiple years of programme of study is STUBID. The algorithms for deriving STUBID are given in paragraphs 31-34 of Appendix 1.

10. If our record of the method used to return FTE on the HESA student record is incorrect, differences are likely to arise (see paragraphs 7-13 of Annex D of 'Higher Education Students Early Statistics Survey 2003-04' (HEFCE 2003/44) for further details). This is likely to cause a discrepancy between the two data sources for records where any of the following are true:

- FTE_TYPE = 1 and TYPEYR ≠ 1
- FTE_TYPE = 2 and STULOAD = 0
- FTE_TYPE = 3 and COMDATE > 31 July 2003 and STULOAD = 0
- FTE_TYPE = 4 and DATELEFT > 31 July 2003 and FUNDCOMP ≠ 2 and STULOAD = 0

11. We make an assumption about students writing up a thesis or a similar piece of work. Details of this assumption are given in paragraph 13 of Appendix 3.

Countable years (including forecasts)

12. To identify countable years (including forecasts) from the individualised file select HESEXCL = 0. The algorithms for deriving HESEXCL are given in paragraphs 72-73 of Appendix 1.

13. There may be a variance as a result of forecasting countable years on HESES03.

Assumed countable years

14. To identify assumed countable years from the individualised file select HESCOL4 = 1. Forecast non-completions are identifiable by HESEXCL = 0 and HESCOMP = 3. The algorithms for deriving HESCOL4, HESEXCL and HESCOMP are given in paragraphs 85, 72-73 and 70 respectively of Appendix 1.

15. There may be a variance as a result of forecasting non-completions on HESES03.

Full-time and sandwich countable years (excluding forecasts)

16. To identify full-time and sandwich countable years (excluding forecasts) from the individualised file select HESEXCL = 0 and HESREG = 1 and HESMODE = FTS. The algorithms for deriving HESEXCL, HESREG and HESMODE are given in paragraphs 72-73, 69 and 22 respectively of Appendix 1.

Sandwich year-out countable years (excluding forecasts)

17. To identify sandwich year-out countable years (excluding forecasts) from the individualised file select HESEXCL = 0 and HESREG = 1 and HESMODE = SWOUT. The algorithms for deriving HESEXCL, HESREG and HESMODE are given in paragraphs 72-73, 69 and 22 respectively of Appendix 1.

Part-time countable years (excluding forecasts)

18. To identify part-time countable years (excluding forecasts) from the individualised file select HESEXCL = 0 and HESREG = 1 and HESMODE = PT. The algorithms for deriving HESEXCL, HESREG and HESMODE are given in paragraphs 72-73, 69 and 22 respectively of Appendix 1.

Assumed part-time load for countable years included in Column 4

19. To identify assumed part-time countable years included in Column 4 from the individualised file select HESCOL4 = 1 and HESMODE = PT. To obtain the load for these countable years sum the values of HESESFTE. The algorithms for deriving HESCOL4, HESMODE and HESESFTE are given in paragraphs, 85, 22 and 51-52 respectively of Appendix 1.

20. We make an assumption about the load of part-time students. Details of this assumption are given in paragraphs 6-8 of Appendix 3.

HEFCE-funded countable years (excluding forecasts)

21. To identify HEFCE-funded countable years (excluding forecasts) from the individualised file select HESEXCL = 0 and HESREG = 1 and HESTYPE = HOMEF. The algorithms for deriving HESEXCL, HESREG and HESTYPE are given in paragraphs 72-73, 69 and 24 respectively of Appendix 1.

Independently funded countable years (excluding forecasts)

22. To identify independently funded countable years (excluding forecasts) from the individualised file select HESEXCL = 0 and HESREG = 1 and HESTYPE = HOMEIF. The algorithms for deriving HESEXCL, HESREG and HESTYPE are given in paragraphs 72-73, 69 and 24 respectively of Appendix 1.

Non-fundable countable years (excluding forecasts)

23. To identify non-fundable countable years (excluding forecasts) from the individualised file select HESEXCL = 0 and HESREG = 1 and HESTYPE = HOMENF. The algorithms for deriving HESEXCL, HESREG and HESTYPE are given in paragraphs 72-73, 69 and 24 respectively of Appendix 1.

24. We make an assumption about non-fundable students. Details of this assumption are given in paragraph 17 of Appendix 3.

Island and overseas countable years (excluding forecasts)

25. To identify island and overseas countable years (excluding forecasts) from the individualised file select HESEXCL = 0 and HESREG = 1 and HESTYPE = ISOV. The algorithms for deriving HESEXCL, HESREG and HESTYPE are given in paragraphs 72-73, 69 and 24 respectively of Appendix 1.

Standard countable years (excluding forecasts)

26. To identify standard countable years (excluding forecasts) from the individualised file select HESEXCL = 0 and HESREG = 1 and LENGTH = S. The algorithms for deriving HESEXCL, HESREG and LENGTH are given in paragraphs 72-73, 69 and 40 respectively of Appendix 1.

Long countable years (excluding forecasts)

27. To identify long countable years (excluding forecasts) from the individualised file select HESEXCL = 0 and HESREG = 1 and LENGTH = L. The algorithms for deriving HESEXCL, HESREG and LENGTH are given in paragraphs 72-73, 69 and 40 respectively of Appendix 1.

Countable years (excluding forecasts) by price group

28. To identify countable years (excluding forecasts) from the individualised file select HESEXCL = 0 and HESREG = 1. To obtain the proportion of activity in each price group sum the values of each of the price group fields (PRGA, PRGB, PRGC, PRGD, PRGMEDIA, PRGPSYCH, PRGITT, PRGINSET). The algorithms for deriving HESEXCL, HESREG and price groups are given in paragraphs 72-73, 69 and 56-60 respectively of Appendix 1.

29. We make assumptions about students on low-credit bearing courses. Details of this assumption are given in paragraphs 21-22 of Appendix 3.

Assumed load (FTE) for countable years included in Column 4 (and Column 4a for part-time) by price group

30. To identify assumed countable years included in Column 4 (and Column 4a for part-time) from the individualised file select HESCOL4 = 1. To obtain the load for these countable years multiply the values of each of the price group fields (PRGA, PRGB, PRGC, PRGD, PRGMEDIA, PRGPSYCH, PRGITT, PRGINSET) with HESESFTE, sum the values and divide by 100. The algorithms for deriving HESCOL4, HESESFTE and price groups are given in paragraphs, 85, 51-52 and 56-60 respectively of Appendix 1.

31. For the 'Summary comparison of price group activity between HESES03 re-creation and the HESES03' table, the Assumed load (FTE) is calculated by summing all values in Column 4 for Table 1a, half the values in Column 4 for Table 2 and all values in Column 4a for Table 3.

Regulated £1,125 fee-level countable years (excluding forecasts)

32. To identify regulated £1,125 fee-level countable years (excluding forecasts) from the individualised file select HESEXCL = 0, HESREG=1 and HESFEELV = 1125. The algorithms for deriving HESEXCL and HESFEELV are given in paragraphs 72-73 and 36 respectively of Appendix 1.

Regulated £550 fee-level countable years (excluding forecasts)

33. To identify regulated £550 fee-level countable years (excluding forecasts) from the individualised file select HESEXCL = 0, HESREG=1 and HESFEELV = 550. The algorithms for deriving HESEXCL and HESFEELV are given in paragraphs 72-73 and 36 respectively of Appendix 1.

Regulated £0 fee-level countable years (excluding forecasts)

34. To identify regulated £0 fee-level countable years (excluding forecasts) from the individualised file select HESEXCL = 0, HESREG=1 and HESFEELV = 0. The algorithms for deriving HESEXCL and HESFEELV are given in paragraphs 72-73 and 36 respectively of Appendix 1.

NHS bursaried courses countable years (excluding forecasts)

35. To identify students on NHS bursaried courses from the individualised file select HESEXCL = 0, HESREG=1 and HESFEELV = NHS. The algorithms for deriving HESEXCL and HESFEELV are given in paragraphs 72-73 and 36 respectively of Appendix 1.

36. We make assumptions about students in receipt of an NHS bursary. Details of these assumptions are given in paragraphs 10-11 of Appendix 3.

Foundation degree bridging courses countable years (excluding forecasts)

37. To identify students on foundation degree bridging courses from the individualised file select HESEXCL = 0, HESREG=1 and HESFEELV = FDBC. The algorithms for deriving HESEXCL and HESFEELV are given in paragraphs 72-73 and 36 respectively of Appendix 1.

Non-regulated fee-level countable years (excluding forecasts)

38. To identify non-regulated fee-level countable years (excluding forecasts) from the individualised file select HESEXCL = 0 and HESREG = 1 and HESFEELV = OTHER. The algorithms for deriving HESEXCL, HESREG and HESFEELV are given in paragraphs 72-73, 69 and 36 respectively of Appendix 1.

Undergraduate full-time and sandwich medical and dental countable years (excluding forecasts)

39. To identify undergraduate full-time and sandwich medical and dental countable years (excluding forecasts) from the individualised file select HESEXCL = 0 and HESREG = 1 and HESMED = 1. The algorithms for deriving HESEXCL, HESREG and HESMED are given in paragraphs 72-73, 69 and 61 respectively of Appendix 1.

Undergraduate full-time and sandwich medical and dental countable years (including forecasts)

40. To identify undergraduate full-time and sandwich medical and dental countable years (including forecasts) from the individualised file select HESEXCL = 0 and HESMED = 1. The algorithms for deriving HESEXCL and HESMED are given in paragraphs 72-73 and 61 respectively of Appendix 1.

Undergraduate full-time and sandwich medical and dental assumed countable years

41. To identify undergraduate full-time and sandwich medical and dental assumed countable years from the individualised file select HESCOL4 = 1 and HESMED = 1. The algorithms for deriving HESCOL4 and HESMED are given in paragraphs 85 and 61 respectively of Appendix 1.

Pre-clinical medical countable years (excluding forecasts)

42. To identify medical and dental countable years (excluding forecasts) from the individualised file select HESEXCL = 0 and HESREG = 1 and HESMED = 1. To obtain the proportion of activity in pre-clinical medicine select REGBODY=01 and sum the values of PRGA subtracted from 1. The algorithms for deriving HESEXCL, HESREG, HESMED and PRGA are given in paragraphs 72-73, 69, 61 and 56-60 respectively of Appendix 1.

Clinical medical countable years (excluding forecasts)

43. To identify medical and dental countable years (excluding forecasts) from the individualised file select HESEXCL = 0 and HESREG = 1 and HESMED = 1. To obtain the proportion of activity in clinical medicine select REGBODY=01 and sum the values for PRGA. The algorithms for deriving HESEXCL, HESREG, HESMED and PRGA are given in paragraphs 72-73, 69, 61 and 56-60 respectively of Appendix 1.

Pre-clinical dental countable years (excluding forecasts)

44. To identify medical and dental countable years (excluding forecasts) from the individualised file select HESEXCL = 0 and HESREG = 1 and HESMED = 1. To obtain the proportion of activity in pre-clinical dentistry select REGBODY=02 and sum the values for PRGA subtracted from 1. The algorithms for deriving HESEXCL, HESREG, HESMED and PRGA are given in paragraphs 72-73, 69, 61 and 56-60 respectively of Appendix 1.

Clinical dental countable years (excluding forecasts)

45. To identify medical and dental countable years (excluding forecasts) from the individualised file select HESEXCL = 0 and HESREG = 1 and HESMED = 1. To obtain the proportion of activity in clinical dentistry select REGBODY=02 and sum the values for PRGA. The algorithms for deriving HESEXCL, HESREG, HESMED and PRGA are given in paragraphs 72-73, 69, 61 and 56-60 respectively of Appendix 1.

New entrants

46. To identify new entrants from the individualised file select HESEXCL = 0 and YEARONE = 1. The algorithms for deriving HESEXCL and YEARONE are given in paragraphs 72-73 and 38-39 respectively of Appendix 1.

47. We make an assumption about the year of programme of study of new entrants. Details of this assumption are given in paragraph 24 of Appendix 3.