

# HIGHER EDUCATION DATA WORKSHOP

**10 June 2008**



# Higher Education Data Workshop

## Outline

- Data requirements
- Data Efficiency Project
- HEIDI
- Data quality analysis
- HESA/EYM comparisons
- Timetable
- Presentation by Bangor University
- New audit process outcomes
- WAG – Analysis of Longitudinal DLHE data

# Data Requirements – HESA Student Record 2007/08

## Module information

- Returning partial completions on the MODOUT (Module Outcome) field of the Module Entity

# Data Requirements – HESA Student Record 2008/09

## Student on Module Entity

- New field – MODCOUNT
- Collects whether the module is countable or not under HESES/EYM rules
- Intended to flag those modules being done over and above maximum for the course
- Purpose: to avoid counting more than the maximum for the course so that data can be extracted from HESA using HESES criteria

# Data Requirements 2007/08 – Widening Access Premium

- New PI – uses POLAR2 to classify neighbourhoods into low participation
- Will develop new method for low affluence part of the widening access premium for 2009/10
- Will continue to use home postcode, taken from 2007/08 record

# Future use of HESA data for monitoring funding allocations

- New fields mean that information collected on HESES/EYM should be able to be extracted from HESA records in future
- Will collect EYM for at least one more year (2008/09) – future collections will depend on quality of HESA data returned in 2007/08 and 2008/09

# 2007/08 HESA Student Record

- Institutions have expressed concerns about the consequences of the new arrangements
- Data used in premiums will be extracted using the same criteria as before,
  - Fundable
  - Active
  - Studying for more than 10 credits
- Not expecting new structure to cause difficulties for data extraction
- Not using new information in the premium calculations

# 2007/08 HESA Student Record

- Also concerns about possibility of:
  - an HEI making a poor quality return or failing to submit data
  - system failure at HESA
- HEFCW do not intend to allow confirmations to be changed where institutions have submitted poor quality data
- System failure at HESA unlikely but will use alternative data for funding if HESA unable to supply data

# Data Efficiency Project

- Originally HEFCE/HESA/TDA project
- Purpose: ‘To research and obtain evidence on: the burdens that arise in institutions in meeting the data requirements of funding bodies and government departments through HESA returns; and on the operational barriers that need to be overcome within institutions in order to improve the efficiency and timeliness of the data submission process.’
- Workshop 13 June 2008

# HEIDI

- Less take up in Wales than rest of UK
- Continuously being developed by HESA
- HEIDI user group

# HEFCW External Stakeholders Survey

- HEIs invited to comment on HEFCW
- Some comments about data returns and audit (but anonymous!)
- Happy to discuss any particular concerns people have

# HESA Data Quality Analysis 2006/07 – 1

Areas considered are included because they:

- have significance in relation to funding
  - are areas receiving greater focus from policy teams or the Welsh Assembly Government / National Assembly for Wales
  - impact on equal opportunities analysis
  - relate to historic data problems
- No changes to methodology from 2005/06 analysis

# HESA Data Quality Analysis 2006/07 – 2

## Table 1 - Postcodes

- Postcode validity very high: 1% invalid or missing

## Table 2 - Ethnicity

- Small increase known ethnicity (95% to 96%)
- Decrease in enrolments with unknown ethnicity (3% to 2%)
- Ethnicity refused unchanged (2%)

# HESA Data Quality Analysis

## 2006/07 – 3

### Table 3 - Welsh subject provision

- Q500 Celtic studies used instead of Q560 Welsh or Q561 Welsh literature

### Table 4 – Proportion of module taught through Welsh

- very few enrolments on modules where the number of credits taught through Welsh is less than 2 but have increased

### Table 5 – Coding of non-standard academic years

- academic years that are potentially non-standard coded as standard (e.g. masters courses expected to last more than 40 weeks)

### Table 6 – Average FTE/credits per enrolment

# HESA Data Quality Analysis

## 2006/07 – 4

**Table 7 – Data linked by HUSID or HUSID and instance to 2005/06 data**

- 5% increase in data linked by HUSID and HUSID and instance
- Proportions of linked records with different date of birth, ethnicity, commencement date or postcode remain the same:

Percentage with different:	Birth Date	Ethnicity	Commencement Date	Domicile	Postcode
2005/06	0.2%	0.3%	1.5%	0.2%	0.5%
2006/07	0.2%	0.3%	1.5%	0.2%	0.6%

**Table 8 – Data linked to 2005/06 data using HEFCE linking**

- linked by using name, date of birth etc.
- 7% of linked records with different HUSID (8% 2005/06)
- Proportion with different institution remains around 5.8%

# HESA Data Quality Analysis 2006/07 – 5

## Table 9 – 2006/07 starters also present in 2005/06

- Some also present in 2005/06 – 1.3% overall
- proportions at individual institutions vary from 0% to 6% (0% to 2% in 2005/06)
- some large proportions for some mode/level combinations at a few institutions

# HESA Data Quality Analysis 2006/07 – 6

## Table 10 – Non-completion of current year

- 66% of all non-completing students left institution
- 15% dormant overall

## Table 11 – Non-completion – reasons for leaving

- 67% personal or other reasons
- 10% unknown
- 8% written off (25% 2005/06)

# HESA Data Quality Analysis 2006/07 – 7

## Table 12 – Welsh Fluency

- 66% of Welsh domiciled students with known Welsh speaker status (increase of 8% on 2005/06)

## Table 13 – Disability

- 93% have known disability status (as per 2005/06)
- 9 institutions have 100% known disability status

## Table 14 – Disabled Students' Allowance (DSA)

- DSA information not known for 18% of enrolled disabled students

# HESA Data Quality Analysis 2006/07 – 8

**Table 15 – Level of subject coding**

JACS Code	JACS Title	Level of Coding
F800	Physical and Terrestrial Geographical and Environmental Sciences	Principal subject - coded to two places
F830	Topography	Coded to three places
F831	Cartography	Coded to four places

- 70% of enrolments coded to principal subject level
- 30% of enrolments coded to three or more places
- at individual HEIs, the proportion of enrolments coded to three or more places varied between 14% and 45%

# COMPARISON OF EYM/HESA DATA

**Study carried out each year to replicate EYM returns using HESA Student and Module record.**

- Ongoing check of data quality of HESA data.
- Highlights any localised data errors.
- Identifies where HESSES/EYM guidance may need further information.

# **EYM/HESA 2006/07 – HE SECTOR**

**Total number of assumed completed credit values extracted from HESA compared to verified EYM data**

**In 2005/06 HESA 3% higher than EYM data**

**In 2006/07 HESA 1% higher than EYM data**

# EYM/HESA 2006/07 - HEIs

**The range of % differences between HESA and EYM total assumed completed credit values range from:**

**FT -5% to 3% (sector -1%)**

**SW -100% to 77% (sector 5%) (small numbers)**

**PT -24% to 47% (sector 8%)**

**Total -4% to 4% (sector 1%)**

# **EYM/HESA 2006/07 – OBSERVATIONS 1**

## **Credit values (Table 1a, 1b, 1c):**

- **Apportionment of modules to subject codes.**
- **Miscoding of sandwich year out on HESA.**
- **Non-fundable courses coded as fundable on HESA, not included on EYM.**

# **EYM/HESA 2006/07 – OBSERVATIONS 2**

## **ITT (QTS) (Table 2a):**

- **Miscoding of registrations gaining QTS.**
- **Registrations coded using general subject areas e.g. Other Modern Languages on HESA rather than French etc. on EYM.**

# **EYM/HESA 2006/07 – OBSERVATIONS 3**

## **Registrations (Tables 3 and 4):**

- **Improvement in inclusion of non-fundable and overseas students.**
- **PGR extracted from HESA closely matches EYM.**
- **Y coded courses still appearing on HESA, yet courses assigned to ASC on EYM.**

# **HESA 2006/07 – OBSERVATIONS**

## **Franchised Out**

### **06/07 HESA compared against 05/06 EYM**

- **Figures comparable in general.**
- **Some franchise to colleges in 05/06 EYM do not exist in 06/07 HESA.**
- **Registrations coded as fundable in HESA when look like should be non-fundable.**
- **FRNCHACT =2.**

# HESA/EYM – NEXT STEPS

## Analysis of 2007/08 HESA data against 2007/08 EYM data

- **Closer matched analysis possible.**
- **Identify any gaps in new guidance or initial errors on returning data to HESA.**
- **Results reported during 2009.**

# HESA data – ITT (QTS) 1

## Greater use of ITT (QTS) HESA data

- Internal policy and external interest.
- Recruitment against targets.
- Gender/age/ethnicity balance of ITT students/graduates.
- Bilingual nature of courses.
- Modelling students through ITT training and completion.
- Destination of leavers.

# HESA data – ITT (QTS) 2

## HESA fields (07/08 HESA Record)

- **COURSE.TTCID**  
(type of teacher training)
- **COURSE.BITTM**  
(bilingual nature of course)
- **INSTANCE.ITTPHSC**  
(phase/scope of teacher training aim)
- **INSTANCE.YEARPRG**  
(year of the course)

# 2008/09 Timetable

10 June	Workshop for institutions
13 June	Data Efficiency Project workshop
4 August	EYM return date
22 August	EYM verifications sent out
29 August	Return date for redistribution exercise
8 September	EYM verifications return date
Sept/October	Audit of data (selected institutions)
Mid September	Data submitted to HESA
Mid Sept/Oct	HESA data quality checking
Mid October	HESES survey issued
Mid November	HESES return date

# New Audit Process

- External audits for 2006/07 completed using HEFCW appointed external auditors
- 13 internal audit reports have been returned (deadline 30 May)
- Internal audit reports will be used to help determine who will be externally audited

# New Audit Process – Outcomes

## Common themes

- Lack of documentation of processes used to compile returns
- Lack of evidence that checks carried out
- Processes for calculating estimates not documented
- Risks not included on risk registers

# New Audit Process – Timetable for 2008/09

1 July 2008	Planning meeting with external auditors
2 July 2008	Informal notification of whether in the sample or not
Mid-July 2008	Letters to institutions about outcomes of internal audit and formal notification of external audits
Sept/Oct 2008	External audits carried out