HESonline™
HOSPITAL EPISODE STATISTICS

NHS
The Information Centre
for health and social care

HES User Guide

Contents

1 What is HES? 3
1.1 HES data sets 3
1.2 HES records 3

2 Background information 4
2.1 Data years 4
2.2 Clinical classifications 4
2.3 Episodes and spells 4
2.4 Type of admission 4
2.5 Unfinished records 4
2.6 Appointments 5
2.7 Sensitive data items 5
2.8 Linking patient records 5
2.9 Monthly data 6

3 Integrity issues 7
3.1 Year-on-year analysis 7
3.2 Data quality 7

4 Data that could identify individuals 8
4.1 Deaths 8
4.2 Purchasing/commissioning organisation code 8
4.3 Geographical 8

5 Accessing HES data 9
5.1 Downloadable tables 9
5.2 Self Service 9
5.3 Tabulations and extracts 9
5.4 HES Interrogation System 9

6 Publishing HES data 10
6.1 Publication restrictions 10

7 Further information 11

Appendix A1: Footnotes to be used in output tables 12

Appendix A2: Additional admitted patient care (Inpatient) footnotes to be used in output tables 15

Appendix A3: Additional outpatient footnotes to be used in output tables 21

Appendix A4: Additional accident and emergency footnotes to be used in output tables 24

1 What is HES?

Hospital Episode Statistics (HES) is a record-level data warehouse managed by The NHS Information Centre for health and social care (The NHS IC). It is mainly populated by taking extracts from routine data flows exchanged between healthcare providers and commissioners via the Secondary Uses Service (SUS).

The NHS Information Centre administers the HES Services on behalf of the Secretary of State for Health. The processing of the HES data warehouse has been outsourced to the HES Team in Northgate Information Solutions (Northgate).

1.1 HES data sets

HES currently comprises three main datasets:

- Admitted patient, which includes inpatients and day cases
- Outpatient, which was labelled as 'experimental data' until 2005-06
- Accident and emergency (A&E), which is labelled as 'experimental data' for 2007-08.

For the admitted patient dataset, cumulative, monthly extracts are taken, followed by an annual refresh of the entire financial year's data. For each financial year there are approximately 16 million records (episodes of care) in the HES data warehouse, representing all NHS-funded admitted patient care, and private care within NHS hospitals in England. (Records for regular day and night attendees from 2002-03 onwards are available on request.)

The outpatient dataset was published in 2006 for 2003-04 and 2004-05 activity, and annually from then on. Outpatient attendances comprise some 60 million records per year.

The accident and emergency dataset is available for 2007-08 as experimental. There are approximately 12 million accident and emergency attendances for 2007-08.

1.2 HES records

Each HES record (episode of care) can contain more than 50 pieces of information, collected directly by hospital providers or derived by the HES team. The type of information collected in HES for each episode of care includes:

- facts about the patient treated, eg date of birth
- details of where they were treated, eg the NHS trust, PCT (primary care trust) or independent sector hospital (at home or abroad)
- administrative details, including admission and discharge dates
- clinical details of diagnoses/treatments.

You can find details of the fields held in HES in the HES Data Dictionaries [http://www.hesonline.nhs.uk/Ease/servlet/ContentServer?siteID=1937&categoryID=289].

2 Background information

2.1 Data years

HES data is based on financial years (1 April xx to 31 March xy). The data currently available for analysis is:

- 1989-90 to 2007-08 (admitted patients)
- 2003-04 to 2007-08 (outpatients)
- 2007-08 (accident and emergency).

2.2 Clinical classifications

HES uses the World Health Organization's ICD (International Classification of Diseases and Related Health Problems) to record diagnosis information. The OPCS-4 (Office of Population, Censuses and Surveys: Classification of Interventions and Procedures, 4th Revision) classification is used to record details of any procedures or interventions performed, eg hip replacements.

2.3 Episodes and spells

It is important to remember that each admitted patient record is for a continuous period of care (episode) administered within a particular consultant specialty at a single hospital provider. Therefore, if a patient is transferred to another consultant or to a different provider during a spell of treatment (the total time a patient is in hospital, from admission to discharge) a new record is generated. This means that not all stays in hospital will be represented by a single HES record; in about eight per cent of cases, the spell of treatment will generate more than one record.

It is also quite common for a patient to undergo two or more separate spells of inpatient treatment during a HES data year. Each spell will result in a separate record or records. These details are important because where a count of individual patients is required it will be necessary to link spells together.

Order of episodes in a spell

The Episode Order (Epiorder) field shows you the order that episodes occur in a spell. The HES record detailing the first episode within a spell of treatment (the 'admission episode') will have an Epiorder value of '1'.

2.4 Type of admission

The Patient Classification (Classpat) field shows the type of admission. Ordinary admissions have a Classpat value of '1' and day cases have a Classpat value of '2'.

Regular day or night attenders (Classpat=3 or 4) are available in recent HES years (from 2002-03), but are generally excluded from analysis.

Less frequently used within the HES dataset is Classpat=5 (mothers using delivery facilities only).

2.5 Unfinished records

The Episode Status (Epistat) field shows whether the patient was still undergoing treatment on the last day of the HES data year (31/03/yyyy). Such cases, of which there are about 250,000 per data year, give rise to what is termed an unfinished record. These records don't contain any useful clinical data as it's only added when the episode is finished (an important consideration when requesting extracts or tabulations).

Unfinished records have an Epistat value of '1'. Most other records, ie those detailing episodes which ended within the data year, have an Epistat value of '3'. (The value 2 is no longer used).

2.6 Appointments

Each outpatient record represents a single appointment at a consultant clinic in a single hospital provider. Patients may see the same consultant, or members of the consultant's team, on more than one occasion during the course of a single year. They may have appointments at more than one consultant clinic in the same or different specialties, or in different providers, for the same or different conditions.

Outpatient appointments are defined by two fields, First Attendance and Attended (or did not attend). A new field, Attendtype, has been created by combining these. If you are interested in attendance data you should look at records with Attendtype values of 1 (first attendance), 2 (subsequent attendance) or 3 (attendance but first or subsequent unknown).

Non-attendances

If you are interested in non-attendance rates you may wish to only include records from those trusts recording non-attendances (these records have not been returned by all trusts).

For DNA (did not attend) data you should only look at records with Attendtype values of 4, 5 or 6 (first, subsequent or not known).

Cancellations

Requests for patient cancellations should be restricted to Attendtype in the range 7 to 9 (first, subsequent, not known). Requests for hospital cancellations should be restricted to Attendtype in the range 10 to 12 (first, subsequent, not known).

2.7 Sensitive data items

Sensitive data items, such as date of birth, NHS Number and patient postcode are not routinely available to users of HES data. You should consider using anonymised equivalents, where available. Requests to use sensitive fields, rather than anonymised versions, must be approved by the Database Monitoring Sub Group (DMSG). For more information please see the HES Protocol or HES Extract Pack (available from HESonline).

Descriptions of all of the data items held within the HES data set are described within the HES Interrogation System and the HES Data Dictionaries [http://www.hesonline.nhs.uk/Ease/servlet/ContentServer?siteID=1937&categoryID=289].

2.8 Linking patient records

PSEUDO_HESID is a data field in HES. It was established to give each patient in the HES data warehouse an identifier that is unique to them and HES. This was implemented to prevent much of the need for the disclosure of potential identifiers purely for matching purposes, although they will still be needed for matching with data outside of HES, eg Office for National Statistics (ONS) mortality data, which are being systematically added to HES.

The PSEUDO_HESID can be used to link together records for a single patient from 1997-98 onwards, and has the potential to link admitted patient to outpatient records from 2002-03 onwards and to A&E records from 2007-08 onwards.

The PSEUDO_HESID succeeds the HESID and has been developed to bring it into line with general encryption standards under the guidance of the Ethics and Confidentiality Committee (ECC, formerly known as PIAG) of the National Information Governance Board. The algorithm which derives the PSEUDO_HESID has been enhanced to match as many patients as possible, even in cases in which data quality is poor.

To meet ECC's standards on data sharing, customers who request PSEUDO_HESID as part of an extract will receive a field called EXTRACT_HESID which is unique to the customer.

2.9 Monthly data

Monthly data is generally available to users of the HES Interrogation System and to some other users on request (see the 'Accessing HES data' section).

Potential users of monthly data should be aware of the following features of the data:

- They are not usually manually cleaned
- The latest month will always be incomplete as the monthly data is taken before all providers have to submit their data to SUS (ie at the PbR reconciliation date)
- Monthly data are taken cumulatively so earlier months' data are improved with each successive month
- Month 12 data (which will also contain the earlier eleven months) cannot be used to draw firm conclusions about annual activity, as it will be updated by the HES Annual Refresh.

Considering the points above, all monthly data must be clearly labelled 'provisional' and carry the appropriate footnote (see appendices). Due consideration must be given to fitness for purpose; the data is used at the users own risk.

More information about monthly HES data can be found in section 3.5 of the HES Protocol on HESonline.

3 Integrity issues

3.1 Year-on-year analysis

In general, the quality of HES data has been improving over time. This means that an analysis produced using 1989-90 admitted patient data (the oldest available) is not strictly comparable with a similar analysis run against data for later years. You must bear this in mind when producing a time series analysis, as small year-on-year changes may well be a product of shortfalls in the earlier years and should not automatically be interpreted as trends in treatment practice or activity.

There have been several other changes that also need to be taken into consideration. For example:

- There have been several changes in the organisation of the NHS. Some changes are structural, eg transfer of Directly Managed Units (DMUs) into hospital trusts, abolition of Regional Health Authorities (RHAs) and subsequent Regional Offices (ROs). Other changes relate to the geographic boundaries defining these structures.
- The clinical classification used to record diagnoses changed in April 1995-96 when the World Health Organization's ICD-10 classification of diseases and related health problems started to be used for coding diagnoses in place of the ICD-9 classification (used from 1989-90 to 1994-95).
- The clinical classification used to record procedures and interventions changed in 2006-07, when the OPCS-4.3 classification was introduced in place of OPCS 4-2 (used from 1989-90 to 1994-95). The OPCS-4.4 classification was introduced in 2007-08.
- Data items in the HES data warehouse have changed over the years (eg 'ethnic category' has been added and 'category of patient' is no longer used).
- The values within data items can change over time.

Further information can be found in the descriptions of data items within the HES Interrogation System and the HES Data Dictionaries [http://www.hesonline.nhs.uk/Ease/servlet/ContentServer?siteID=1937&categoryID=289].

These points should be shared with anyone using the data and appropriate warnings included in any analysis (see appendices).

3.2 Data quality

Poor quality data

Poor quality is not a reason for withholding data, unless it is completely unfit for purpose. There are areas in HES where, unfortunately, shortfalls are normal; maternity and psychiatric data, for example. Warnings to this effect should always be given and specifically included in any analysis (see appendices). If you are in doubt about the advisability of issuing 'suspect' data consult the HES team.

4 Data that could identify individuals

This section gives an overview of some instances where HES data could identify individuals. For more specific information and guidance please see the HES Protocol.

4.1 Deaths

Deaths analysis can be sourced from the HES dataset itself or from ONS death registration records supplied by ONS for linkage with HES. The DISMETH (Discharge method) and DISDATE (Discharge date) fields in HES are limited in that they do not capture deaths occurring post discharge from hospital. This can cause incompleteness or bias in some comparative analyses. Linking HES with the ONS registered deaths data addresses this issue.

Queries concerning deaths should always be handled with caution and discussed with a member of the HES team, where necessary. Especially sensitive are requests for counts of deaths by operative procedure and/or diagnosis for individual NHS trusts.

HES deaths from whichever source data should not be used to infer the quality of care provided by NHS trusts; many factors not included within the HES dataset will have an effect on the care of the patient. If HES deaths data are used as part of such comparative analysis users should ensure that the relevant footnote (see Appendices A1-3 of this document) is included.

Cause of death data cannot be obtained from HES at the moment; refer to the mortality statistics [http://www.statistics.gov.uk/CCI/nscl.asp?ID=6444] produced by ONS from death certification, or to the Compendium of Health Statistics [http://www.ohecompendium.org] produced by the Office of Health Economics.

4.2 Purchasing/commissioning organisation code

The Purchaser Code (Purcode) field shows which organisation paid for or commissioned a patient's treatment. There are six main purchaser types on the system; their prevalence varies with each data year:

1. Health authorities
2. GP fund holders
3. Primary care groups
4. Primary care trusts
5. Private patients
6. The Department of Health.

4.3 Geographical

Small numbers at a small geographical area (combined with local knowledge) could allow the identification of individuals. For example:

- The fields Currward, Resladst, Ward 81/91, Oaward and Postcode Sector could generate small numbers for small geographical areas, leading to the identification of individual patients
- Low-level analyses might result in small numbers of episodes within a particular specialty in a single hospital provider, leading to the identification of individual consultants or individual patients.

To prevent the identification of individuals, you must follow the instructions set out in the HES Protocol for the publication of small numbers.

5 Accessing HES data

You can access HES data in a variety of ways:

- Download tables of data
- Use the Self Service area
- Request a tabulations or an extract
- Apply for access to the HES Interrogation System.

5.1 Downloadable tables

You can download a selection of tables containing the most commonly requested national-level information from HES inpatient and outpatient data, such as primary diagnosis, and procedures and interventions data. The tables are available from the 'Freely available data' area of the HESonline website [http://www.hesonline.nhs.uk].

5.2 Self Service

The Self Service area of HESonline offers you the opportunity to create your own tables of HES inpatient data for diagnosis, operation and HRG data. Unlike the downloadable tables, in Self Service you can access data from national through to trust level.

5.3 Tabulations and extracts

Tabulations present aggregate (or average) results of a query, from a single figure up to complex or nest cross-tabulations. Whereas extracts contain record-level data for selected cases, showing selected information for each case; an extract is ideal if you need many pieces of information or want sensitive data.

Any individual or organisation can request a tabulation or extract, subject to HES terms and conditions and the Data Protection Act 1998. Before making a request, we suggest you first make sure that what you require is not already available in the 'Freely available data' or 'Self Service' areas of HESonline.

What will it cost?

In most cases, The NHS Information Centre will charge a fee to recover the costs of processing your request. The exact fee will depend on the size and complexity of your request (see the HES Service Charter). There will be no charge for unsuccessful requests.

How do I make a request?

For more information about the request process and the relevant forms please see the HES Tabulation and Extract Packs in the 'Request a tailor-made report' area of HESonline.

5.4 HES Interrogation System

The HES Interrogation System gives direct access to a selection of HES inpatient and outpatient data held in the HES Data Warehouse, allowing you to create your own queries at any time and as often as you like.

To request access to the Interrogation System you (or your organisation) will need to be a member of the wider NHS Family or related public service organisations. For more information about costs and restrictions please see the 'Criteria for HES access' document, which is available from the FAQs area of the HESonline website [http://www.hesonline.nhs.uk], or contact the NHS Information Centre Contact Centre (0845 300 6016, enquiries@ic.nhs.uk).

6 Publishing HES data

The NHS Information Centre has the right to refuse to grant you permission to publish information arising from HES data on the grounds of data protection. Any reports, papers or statistical tables that are published or released to other organisations must not identify individuals or enable individuals to be identified (see the 'Publication restrictions' section below).

You should give the Information Governance Team in the NHS Information Centre notice of your intention to publish HES data and, where possible, a copy of the published work. Any published work containing HES data must acknowledge 'Hospital Episode Statistics (The NHS Information Centre for health and social care)' as the source.

6.1 Publication restrictions

If you intend to release or publish HES data that include small numbers (figures based on between one and five cases), then you should consider the potential for identification of patients or practitioners.

Generally, national figures or those based on broader geographical areas wouldn't potentially disclose any sensitive information. But for narrower geographical areas, such as provider, primary care trust, local authority district (LAD) or ward, you would need to do the following:

1. Suppress small numbers

Values between one and five must be suppressed. If you need to suppress only one cell, suppress at least one other component cell (the next smallest) to avoid calculation of suppressed value from the totals. You should replace these values with '*' and add a note, '* (asterisk) in this table means a figure between one and five'. (Zeros need not be suppressed.) You can set up a spreadsheet formula, or macro, to make these changes.

2. Aggregate data

As an alternative to suppressing values between one and five, consider a higher level of aggregation for one or more items. For example, you could:

- move from provider to health authority of treatment
- move from diagnosis at 4-character level to the 3-character level
- group ages into bands
- make existing age bands broader. (This option will be the best choice if a large number of cells are affected by the suppression rule.)

3. Add anonymous reference codes

You should consider providing the data at the requested low level, if essential for purpose, but anonymising the level of aggregation, ie replacing identifying codes or labels with arbitrary reference numbers.

Further details are in the HES Protocol (available from the HESonline website).

7 Further information

You can find more information about HES data, including data quality reports, process information, data and articles on the HESonline website [http://www.hesonline.nhs.uk].

If you have any questions not answered by the HES User Guide please don't hesitate to let us know by using the 'Contact us' facility on the website.

Records for the rest of Great Britain

The appropriate sections of the Northern Ireland, Scotland and Wales Offices collect data for NHS hospitals in Northern Ireland, Scotland and Wales. The NHS Information Centre's Contact Centre can give you the relevant contact details (telephone 0845 3006016 or email enquiries@ic.nhs.uk).

Appendix A1: Footnotes to be used in output tables

1. Small numbers

Reference	Area	
1a	Small numbers	To protect patient confidentiality, figures between 1 and 5 have been suppressed and replaced with "*" (an asterisk). Where it was possible to identify numbers from the total due to a single suppressed number in a row or column, an additional number (the next smallest) has been suppressed.
1b	Small numbers disclosure	The data supplied include small numbers of cases for some cells, for your organisation's internal use only. Please note that these must not be published or disclosed without masking figures based on between 1 and 5 cases, if these represent local data such as for a PCT or trust. If in doubt, please refer to the HES Protocol or contact the HES team.

2. Pre-release and provisional data

Reference	Area	
2a	Pre-release data (RESTRICTED STATISTICS)	You are reminded that these are official statistics to which you have privileged access in advance of release. Such access is carefully controlled and is provided for management, quality assurance and briefing purposes only. Release into the public domain or any public comment on these statistics prior to official publication would undermine the integrity of official statistics and conflict with legislation and central Government regulations on official statistics, notably the Statistics and Registration Act 2007 and Pre-release Access to Official Statistics Order 2008. Any accidental or wrongful release should be reported immediately and may lead to an inquiry. Wrongful release includes indications of the content, including descriptions such as "favourable" or "unfavourable". If in doubt you should consult Andy Sutherland or Belinda South who can advise. Please prevent inappropriate use by treating this information as restricted, refrain from passing information on to others who have not been given prior access and use it only for the purposes for which it has been provided.
2b	RESTRICTED STATISTICS - Latest Month of Monthly Extract	You are reminded that these are official statistics to which you have privileged access. Such access is carefully controlled and is provided for NHS and Public Body management purposes only. Release into the public domain or any public comment on these statistics would undermine the integrity of official statistics and conflict with legislation and central Government regulations on official statistics, notably the Statistics and Registration Act 2007. Any accidental or wrongful release should be reported immediately and may lead to an inquiry. Wrongful release includes indications of the content, including descriptions such as "favourable" or "unfavourable". If in doubt you should consult hes.questions@ic.nhs.uk for advice. Please prevent inappropriate use by treating this information as restricted, refrain from passing information on to others who have not been given prior access and use it only for the purposes for which it has been provided.

2c	Provisional data	The data is provisional and may be incomplete or contain errors for which no adjustments have yet been made. Counts produced from provisional data are likely to be lower than those generated for the same period in the final dataset. This shortfall will be most pronounced in the final month of the latest period, ie November from the (month 9) April to November extract. It is also probable that clinical data are not complete, which may in particular affect the last two months of any given period. There may also be a variety of errors due to coding inconsistencies that have not yet been investigated and corrected.

3. Specialty

Reference	Area	
3a	Consultant Main Specialty	This defines the specialty under which the consultant responsible for the care of the patient at that time is registered. Care is needed when analysing HES data by specialty, or by groups of specialties (such as "acute"). Trusts have different ways of managing specialties and attributing codes so it is better to analyse by specific diagnoses, operations or other patient or service information.
3b	Consultant Treatment Specialty	This defines the specialty under which the consultant responsible for care of the patient is working, which may be different to the main specialty under which the consultant is registered. Care is needed when analysing HES data by specialty, or by groups of specialties (such as "acute"). Trusts have different ways of managing specialties and attributing codes so it is better to analyse by specific diagnoses, operations or other recorded information.

4. Miscellaneous

Reference	Area	
4a	Copyright	The data supplied is covered by copyright. Copyright © 2008, Hospital Episodes Statistics (HES): The NHS Information Centre for health and social care. All rights reserved. This work remains the sole and exclusive property of The NHS Information Centre for health and social care, and may only be reproduced where there is explicit reference to the ownership of Health and Social Care Information Centre.
4b	Patient counts	Patient counts are based on the unique patient identifier, HESID. This identifier is derived based on a patient's date of birth, postcode, sex, local patient identifier and NHS number, using a standard algorithm. Where data are incomplete, HESID might erroneously link episodes or fail to recognise episodes for the same patient. Care is therefore needed, especially where the data includes duplicate records. The patient count cannot be summed across a table where patients may have episodes in more than one cell.
4c	Quality of care	Data derived from Hospital Episode Statistics (HES) cannot be used in isolation to evaluate the quality of care provided by NHS trusts or clinical teams. There are many factors that can affect the outcome of treatment and it is beyond the scope of HES to adequately record and reflect all of these.
4d	Treatment centres	HES now includes data from treatment centres (TCs), which are dedicated units that offer pre-booked day and short-stay surgery and diagnostic procedures in certain specialties. Treatment centres are run either by NHS trusts, primary care trusts (PCTs) or the independent sector, commissioned by the NHS. Normally, if data is tabulated by healthcare provider, the figure for

		an NHS trust gives the activity of all the sites as one aggregated figure. However, in the case of those with embedded treatment centres, this data is quoted separately. In these cases, '-X' is appended to the code for the rest of the trust, to remind users that the figures are for all sites of the trust excluding the treatment centres. Please note that although HES now includes data from TCs, the quality of returns are such that it may not be as complete as possible. Some NHS trusts have not registered their treatment centre as a separate site, and it is therefore not possible to identify their activity separately. Data from some independent sector providers, where the onus for arrangement of dataflows is on the commissioner, may also be missing. Therefore, care must be taken when using this data as the counts may be lower than true figures.
4e	Ungrossed data	Figures have not been adjusted for shortfalls in data (ie the data are ungrossed).
4f	Ethnicity	Ethnicity data on HES should be used with care and may not yet be robust enough to support analysis of ethnic differences. Ethnic group was collected from 1 April 1995 to 31 March 2002 and Ethnic category, using the definitions in the 2001 census, from 1 April 2002. Patients are asked to select their category from a standard list, and some decline to do this. The incompleteness of data collected centrally (as a by-product of local information systems) and issues around the quality of ethnic coding may mean that the data are not yet fit for routine analysis, but remain useful for addressing issues of data quality.

Appendix A2: Additional admitted patient care (Inpatient) footnotes to be used in output tables

5. Episodes and spells

Reference	Area	
5a	Finished Consultant Episode (FCE)	A finished consultant episode (FCE) is defined as a continuous period of admitted patient care under one consultant within one healthcare provider. FCEs are counted against the year in which they end. Please note that the figures do not represent the number of different patients, as a person may have more than one episode of care within the same stay in hospital or in different stays in the same year.
5b	Finished Provider Spell	A finished provider spell is defined as a period of inpatient care within one health care provider that commences with the patient's admission and ends on their date of discharge. It may comprise one or more finished consultant episodes (FCEs). Finished provider spells are counted against the year in which they end. Please note that the figures do not represent the number of patients, as a person may have more than one spell in the same or different hospitals within the year.
5c	Episode duration	Episode duration is calculated as the difference in days between the episode start date and the episode end date, where both are given. Episode duration is based on finished consultant episodes and only applies to ordinary admissions, ie day cases are excluded (unless otherwise stated).

6. Admissions

Reference	Area	
6a	Finished admission episodes	A finished admission episode is the first period of inpatient care under one consultant within one healthcare provider. Finished admission episodes are counted against the year in which the admission episode finishes. Please note that admissions do not represent the number of inpatients, as a person may have more than one admission within the year.
6b	Finished in-year admissions	A finished in-year admission episode is the first period of inpatient care under one consultant within one healthcare provider, excluding admissions beginning before 1 April at the start of the data year. Finished in-year admission episodes are counted against the year in which the admission episode starts and finishes. Therefore admission episodes that start and finish in different years are not counted – for most conditions this represents a small proportion of the total. Please note that admissions do not represent the number of inpatients, as a person may have more than one admission within the year.
6c	In-year admissions	An in-year admission episode is the first period of inpatient care under one consultant within one healthcare provider, excluding admissions beginning before 1 April at the start of the data year. Periods of care ongoing at the end of the data year (unfinished admission episodes) are included. Please note that admissions do not represent the number of inpatients, as a person may have more than one admission within the year.

7. Discharges

Reference	Area	
7a	Discharges	A discharge episode is the last episode during a hospital stay (a spell), where the patient is discharged from the hospital (this includes transfer to another hospital).
7b	Deaths	Hospital Episode Statistics (HES) data cannot be used to determine the cause of death of a patient while in hospital. Deaths recorded on the HES database may be analysed by the main diagnosis for which the patient was being treated during their stay in hospital, which may not necessarily be the underlying cause of death. For example, a patient admitted for a hernia operation (with a primary diagnosis of hernia) may die from an unrelated a heart attack. The Office for National Statistics (ONS) collects information on the cause of death, wherever it occurs, based on the death certificate and should be the source of data for analyses on cause of death.

8. Diagnoses

Reference	Area	
8a	Primary diagnosis	The primary diagnosis is the first of up to 20 (14 from 2002-03 to 2006-07 and 7 prior to 2002-03) diagnosis fields in the Hospital Episode Statistics (HES) data set and provides the main reason why the patient was admitted to hospital. ***Pre 2007-08 release:*** *The primary diagnosis is the first of up to 14 (7 prior to 2002-03) diagnosis fields in the Hospital Episode Statistics (HES) data set and provides the main reason why the patient was in hospital.*
8b	Secondary diagnoses	As well as the primary diagnosis, there are up to 19 (13 from 2002-03 to 2006-07 and 6 prior to 2002-03) secondary diagnosis fields in Hospital Episode Statistics (HES) that show other diagnoses relevant to the episode of care. ***Pre 2007-08 release:*** *As well as the primary diagnosis, there are up to 13 (6 prior to 2002-03) secondary diagnosis fields in Hospital Episode Statistics (HES) that show other diagnoses relevant to the episode of care.*
8c	Number of episodes in which the patient had a named primary diagnosis	These figures represent the number of episodes where the diagnosis was recorded in the primary diagnosis field in a Hospital Episode Statistics (HES) record.
8d	Number of episodes in which the patient had a named primary or secondary diagnosis	These figures represent the number of episodes where the diagnosis was recorded in any of the 20 (14 from 2002-03 to 2006-07 and 7 prior to 2002-03) primary and secondary diagnosis fields in a Hospital Episode Statistics (HES) record. Each episode is only counted once in each count, even if the diagnosis is recorded in more than one diagnosis field of the record. ***Pre 2007-08 release:*** *These figures represent the number of episodes where the diagnosis was recorded in any of the 14 (7 prior to 2002-03) primary and secondary diagnosis fields in a Hospital Episode Statistics (HES) record. Each episode is only counted once in each count, even if the diagnosis is recorded in more than one diagnosis field of the record.*

9. Operations/Procedures/Interventions

Reference	Area	
9a	Main procedure	The main procedure is the first recorded procedure or intervention in the Hospital Episode Statistics (HES) data set and is usually the most resource intensive procedure or intervention performed during the episode. It is appropriate to use main procedure when looking at admission details, (eg time waited), but a more complete count of episodes with a particular procedure is obtained by looking at the main and the secondary procedure.
9b	Secondary procedure	As well as the main operative procedure, there are up to 23 (11 from 2002-03 to 2006-07 and 3 prior to 2002-03) secondary operative procedure fields in Hospital Episode Statistics (HES) that show secondary or additional procedures performed on the patient during the episode of care. ***Pre 2007-08 release:*** *As well as the main operative procedure, there are up to 11 and 3 prior to 2002-03 secondary operative procedure fields in Hospital Episode Statistics (HES) that show secondary or additional procedures performed on the patient during the episode of care.*
9c	Number of episodes with a (named) main procedure	These figures represent the number of episodes where the procedure (or intervention) was recorded in the main operative procedure field in a Hospital Episode Statistics (HES) record. Please note that more procedures are carried out than the number of episodes with a main procedure. For example, patients under going a 'cataract operation' would tend to have at least two procedures – removal of the faulty lens and the fitting of a new one – counted in a single episode. The number of episodes with any main procedure include '&' (not known), and for OPCS 4.3/4.4 (from 2006-07) codes A01–X97, or for OPCS 4.2 (prior to 2006-07) codes A01–X59.
9d	Number of episodes with a (named) main or secondary procedure	These figures represent the number of episodes where the procedure (or intervention) was recorded in any of the 24 (12 from 2002-03 to 2006-07 and 4 prior to 2002-03) operative procedure fields in a Hospital Episode Statistics (HES) record. A record is only included once in each count, even if the procedure is recorded in more than one operative procedure field of the record. Please note that more procedures are carried out than episodes with a main or secondary procedure. For example, patients under going a 'cataract operation' would tend to have at least two procedures – removal of the faulty lens and the fitting of a new one – counted in a single episode. ***Pre 2007-08 release:*** *These figures represent the number of episodes where the procedure (or intervention) was recorded in any of the 12 operative procedure fields (4 prior to 2002-03) in a Hospital Episode Statistics (HES) record. A record is only included once in each count, even if the procedure is recorded in more than one operative procedure field of the record. Please note that more procedures are carried out than episodes with a main or secondary procedure. For example, patients under going a 'cataract operation' would tend to have at least two procedures – removal of the faulty lens and the fitting of a new one – counted in a single episode.*

9e	Total number of (named) procedures	These figures represent the total number of (named) procedures recorded in any of the 24 (12 from 2002-03 to 2006-07 and 4 prior to 2002-03) operative procedure fields in the Hospital Episode Statistics (HES) data set. Therefore, if a procedure is recorded in more than one operative procedure field during an episode, all procedures are counted. ***Pre 2007-08 release:*** *These figures represent the total number of (named) procedures recorded in any of the 12 (4 prior to 2002-03) operative procedure fields in the Hospital Episode Statistics (HES) data set. Therefore, if a procedure is recorded in more than one operative procedure field during an episode, all procedures are counted.*
9f	Changes to coding classifications: OPCS-4	Operative procedure codes were revised for 2006-07 and 2007-08. The 2007-08 data uses OPCS 4.4 codes, 2006-07 data uses OPCS 4.3 codes, data prior to 2006-07 uses OPCS 4.2 codes. All codes that were in OPCS 4.2 remain in later OPCS 4 versions, however the introduction of OPCS 4.3 codes enable the recording of interventions and procedures which were not possible in OPCS 4.2. In particular, OPCS 4.3 and OPCS 4.4 codes include high cost drugs and diagnostic imaging, testing and rehabilitation. Some activity may have been coded under different codes in OPCS 4.2. These changes need to be borne in mind when analysing time series and may explain some apparent variations over time. Please note that care needs to be taken in using the newer codes as some providers of data were unable to start using the new codes at the beginning of each datayear. More information about OPCS 4 changes is on the Connecting for Health website (www.connectingforhealth.nhs.uk).

10. Data quality

Reference	Area	
10a	Data quality	Hospital Episode Statistics (HES) are compiled from data sent by more than 300 NHS trusts and primary care trusts (PCTs) in England. Data is also received from a number of independent sector organisations for activity commissioned by the English NHS. The NHS Information Centre for health and social care liaises closely with these organisations to encourage submission of complete and valid data and seeks to minimise inaccuracies and the effect of missing and invalid data via HES processes. While this brings about improvement over time, some shortcomings remain.
10b	PCT/SHA data quality	PCT and SHA data was added to historic data years in the HES database using 2002-03 boundaries, as a one-off exercise in 2004. The quality of the data on PCT of treatment and SHA of treatment is poor in 1996-97, 1997-98 and 1998-99, with over a third of all finished episodes having missing values in these years. Data quality of PCT of GP practice and SHA of GP practice in 1997-98 and 1998-99 is also poor, with a high proportion missing values where practices changed or ceased to exist. There is less change in completeness of the residence-based fields over time, where the majority of unknown values are due to missing postcodes on birth episodes. Users of time series analysis including these years need to be aware of these issues in their interpretation of the data.
10c	Assessing growth through time	HES figures are available from 1989-90 onwards. The quality and coverage of the data have improved over time. These

		improvements in information submitted by the NHS have been particularly marked in the earlier years and need to be borne in mind when analysing time series. Some of the increase in figures for later years (particularly 2006-07 onwards) may be due to the improvement in the coverage of independent sector activity. Changes in NHS practice also need to be borne in mind when analysing time series. For example, a number of procedures may now be undertaken in outpatient settings and may no longer be accounted for in the HES data. This may account for any reductions in activity over time.

11. Grossed data

Reference	Area	
11a	Grossing	Grossing Figures to 2002-03 are grossed-up to compensate for estimated gaps in coverage. Due to improved data coverage thereafter, figures from 2003-04 are not adjusted.
11b	Ungrossed data	Figures have not been adjusted for shortfalls in the data, ie the data are ungrossed.

12. Admitted patient care: miscellaneous

Reference	Area	
12a	Bed occupancy	Beddays of finished episodes and beddays of finished spells include days of bed occupancy during previous years, eg a patient discharged in 2003-04 may have been admitted during 2002-03. Conversely, beddays within the year includes only those days falling between 1 April and 31 March of the data year (including unfinished episodes, unless otherwise stated).
12b	Cause code	The cause code is a supplementary code that indicates the nature of any external cause of injury, poisoning or other adverse effects. The field within HES counts only the first external cause code which is coded within the episode.
12c	Consultant code	The consultant code is an eight-character code, usually comprising the GMC code of the lead consultant responsible prefixed with 'C' (or other code and prefix if lead is not a consultant). The code is inputted by the provider trust, which is responsible for ensuring it is correctly attributed. Prior to 1 April 1999, the consultant code was validated using an algorithm based on an integral check digit, whereby the eighth digit was overwritten by 1 (valid code) or 0 (invalid code). Valid codes for 1997-98 and 1998-99 are now being reconstructed to undo this overwriting, to ensure consistency of valid codes over time. From 1 April 1999, the eighth character was no longer a check digit (for newly registered codes) so no such validation was possible.
12d	Length of stay (duration of episode)/ Length of stay (duration of spell)	Length of stay (LOS) is calculated as the difference in days between the admission date and the episode end date (duration of episode) or discharge date (duration of spell), where both dates are given. LOS is based on hospital stays and only applies to ordinary admissions, ie day cases are excluded (unless otherwise stated). Information relating to LOS figures, including discharge method/destination, diagnoses and any operative procedures, is based only on the final episode of the spell.

12e	Time waited (days)	Time waited statistics from Hospital Episode Statistics (HES) are not the same as the published waiting list statistics. HES provides counts and time waited for all patients admitted to hospital within a given period, whereas the published waiting list statistics count those waiting for treatment on a specific date and how long they have been on the waiting list. Also, HES calculates the time waited as the difference between the admission and decision to admit dates. Unlike published waiting list statistics, this is not adjusted for self-deferrals or periods of medical/social suspension.
12f	Source statement	Source: Hospital Episode Statistics (HES), The NHS Information Centre for health and social care
12g	Inpatients	Inpatients are defined as patients who are admitted to hospital and occupy a bed, including both admissions where an overnight stay is planned and day cases.
12h	Activity included	Activity in English NHS Hospitals and English NHS commissioned activity in the independent sector

Appendix A3: Additional outpatient footnotes to be used in output tables

13. Attendance		
Reference	Area	
13a	Attended or did not attend	This indicates whether or not a patient attended an appointment. If the patient did not attend an indication is provided of whether or not advanced warning was given. Analysis of 'Did not attend' and 'cancellations' must be restricted to only the trusts which return data.
13b	Attendance Type	Attendance Type identifies if the patient attended an appointment, and if it was first or subsequent attendance, or if the appointment was cancelled or the patient did not attended. This field is complete for over 99% of appointments.
13c	Outcome of Attendance	This records the outcome of an outpatient attendance. The field is very well completed, however there are still around 20 trusts which do not use 'discharged' as an outcome so analysis on this must be restricted.
13d	Medical staff type	This gives details about the type of care professional dealing with the patient during a consultant outpatient attendance or nurse or midwife contact. The coding classification changed from 1 and 2 (consultant and member of consultant firm) to 3 and 4 (lead care professional and member of care professional team) between 2003-04 and 2004-05. All four codes feature in both years. Analysis must be restricted to organisations with high level of validity as under 60% of records are currently valid.
13e	First Attendance	This indicates whether a patient is making a first attendance or a follow-up attendance and whether the consultation was face-to-face or via telephone/telemedicine consultation.

14. Referral details		
Reference	Area	
14a	Priority Type	This details the priority of a request for services; in the case of services to be provided by the consultant, it is as assessed by or on behalf of the consultant. Please note that the opinions of what is classed as a priority will vary from consultant to consultant, and service to service.
14b	Service Type Requested	This describes the terms of reference for the referral request. This field is fairly complete but analysis should still be restricted to organisations with high levels of coding.
14c	Source of Referral	This identifies the source of referral for each outpatient consultant episode. The referral may or may not be initiated by the responsible consultant depending on the circumstances. There is a very high level of validity but there is evidence of variable provider practice in recording this for subsequent ("follow-up") attendances. 36 of 162 organisations with inpatient dental specialties were not recording GDP as a referral source. Analysis is restricted to first attendances, unless local provider practice for subsequent attendances is known. Analysis should also be restricted to GDP referrals for relevant organisations who record this.

15. Time waited

Reference	Area	
15a	Time waited	Time waited statistics from Hospital Episode Statistics (HES) are not the same as the published waiting list statistics. Waiting times are usually only calculated for first attendances of patients referred by GPs and dentists (see below). It is unclear whether the data collected has any relevance to subsequent attenders. Analysis of the data has revealed high (up to 100%) percentages of zero day waits for some providers suggesting poor data recording. Waits also 'bunch' at every 7th day suggesting that clocks are not starting with the receipt of a referral letter, but rather with a booking occurring in clinic. HES calculates the time waited as the difference between the referral request received and appointment (whether attended or not) dates.
15b	Waiting time indicator	This is an indicator to show whether and how a time waited figure has been calculated. This can be used to filter out records that have no time waited figure calculated yet and can show the reason for this.
15c	Waiting time weeks	This shows the number of completed weeks waited and can be used as a comparison between the QM08 aggregate returns. Waits are grouped into completed weeks (the basis upon which waiting times are recorded in aggregate returns). So a wait of 6 days is recorded as 0 weeks and 13 days as 1 week.

16. Measures

Reference	Area	
16a	Appointment Count	This provides a count of the number of planned/booked appointments for outpatients. The database is constructed of one row per appointment that was made, whether it was attended or not.

17. Data quality

Reference	Area	
17a	Outpatient Data Quality	Outpatient Hospital Episode Statistics (HES) data were collected for the first time in 2003-04. It is not mandatory to code procedures on outpatient records and therefore only around 2% of records have completed clinical codes. We have no reliable existing data source to validate this data against, as Department of Health aggregate returns have never collected clinical codes; it is not clear how representative the figures are. The data represents a sample of outpatient attendances. Statistical estimates (such as median and 90th percentile waiting times for main operations) must be regarded as potentially unreliable until it is possible to assess the quality of local coding.
17b	Data quality	Hospital Episode Statistics (HES) are compiled from data sent by over 300 NHS trusts and primary care trusts (PCTs) in England. Data is also received from a number of independent sector organisations for activity commissioned by the English NHS. The NHS Information Centre for health and social care liaises closely with these organisations to encourage submission of complete and valid data and seeks to minimise inaccuracies and the effect of missing and invalid data via HES processes. While this brings

		about improvement over time, some shortcomings remain.

18. Outpatient miscellaneous

Reference	Area	
18a	Source statement	Source: Hospital Episode Statistics (HES); Outpatients, The NHS Information Centre for health and social care

Appendix A4: Additional accident and emergency footnotes to be used in output tables

19. Statistics		
Reference	Area	
19a	Total attendance	An attendance is a record for every patient that attends an A&E department, including a major A&E department, single specialty A&E departments, walk-in centres and minor injuries units. Any one patient can have multiple attendances, which may be in the same or different time period, for the same or different condition. An attendance can be split into whether the attendance was either a new or a follow-up attendance, using the 'Attendance Category' data field.
20. Data Quality		
Reference	Area	
20a	Data Quality	Hospital Episode Statistics (HES) are compiled from data sent by a number of NHS providers across England. The NHS Information Centre for Health and Social Care liaises closely with these organisations to encourage submission of complete and valid data and seek to minimise inaccuracies and the effect of missing and invalid data via HES processes. While this brings about improvement over time, some shortcomings remain. Accident and Emergency HES data is available for the year 2007-08, which covers attendances reported between April 2007 and March 2008. The A&E HES data for 2007-08 is the first record level national A&E attendance data to be available within HES. The current coverage and quality of A&E data in HES is poor and for this reason the dataset has been labelled as 'experimental'. Allowing access to this data will also help stimulate discussion and encourage trusts to improve quality for subsequent releases. The 2007-08 A&E HES publication addresses some of the key data quality and coverage issues. This report is available on HESonline [www.hesonline.nhs.uk]
21. Clinical		
Reference	Area	
21a	Diagnosis	Any diagnosis carried out during an attendance at an A&E department can be recorded using three different classifications of codes; Accident and Emergency diagnosis codes, ICD-10 and READ-5 codes. For more information on these, visit HESonline [www.hesonline.nhs.uk]. From April 2008, all providers are mandated to use the A&E classification of diagnosis codes. Analysis of diagnosis based on A&E HES data is produced using A&E diagnosis codes unless stated otherwise. Providers are able to submit unlimited number of diagnoses for each attendance, however, only the first 12 diagnoses are available in HES.

		Analysis on diagnoses in A&E HES is based on the primary diagnosis code submitted, unless stated otherwise. The coverage and quality of diagnosis data in 2007-08 A&E HES is poor and therefore great caution is needed before interpreting this in any way. Further information on the quality and coverage of diagnosis data is available in the 2007-08 A&E HES publication, which is available on HESonline [www.hesonline.nhs.uk]
21b	Investigation	Investigations can be requested to assist with any diagnosis during an A&E attendance. Not all attendances involve investigations. Providers are able to submit unlimited numbers of investigations for each attendance, however, A&E HES contains up to the first 12 investigations coded for each attendance. A&E investigation codes are used for reporting an investigation for an A&E attendance. For more information on A&E investigation, visit HESonline [www.hesonline.nhs.uk]. The coverage and quality of investigation data in 2007-08 A&E HES is poor and therefore great caution is needed before interpreting this in any way. Further information on the quality and coverage of investigation data is available in the 2007-08 A&E HES publication, which is available on HESonline [www.hesonline.nhs.uk].
21c	Treatment	Treatment is any intervention that takes place during an A&E attendance. For the financial year 2007-08, providers had the option of using one of three different treatment classifications of codes; A&E treatment codes, OPCS-4 and READ-5. For more information on these, visit HESonline [www.hesonline.nhs.uk]. From April 2008, all providers are mandated to use the A&E classification of treatment codes. Analysis of treatment based on A&E HES data is produced using A&E treatment codes unless stated otherwise. Providers are able to submit unlimited number of treatments for each attendance, however, only the first 12 treatment codes are available in HES. Analysis on treatment in A&E HES is based on the primary treatment code submitted, unless stated otherwise. The coverage and quality of treatment data available in 2007-08 A&E HES is poor and therefore great caution is needed before interpreting this in any way. Further information on the quality and coverage of treatment data is available in the 2007-08 A&E HES publication, which is available on HESonline [www.hesonline.nhs.uk].

22. A&E Miscellaneous

Reference	Area	
22a	Source statement	Source: Accident and Emergency Hospital Episode Statistics (A&E HES); The NHS Information Centre for Health and Social Care

Copyright re-use: HES

If you wish to re-use this information in any format for example in a presentation at an exhibition, publication or broadcast, please contact the Information Governance & Quality Directorate of the NHS Information Centre by completing the form below. This will be logged so that we can record the re-use of our data.

Ref (internal use only):

Name:

Email address:

Contact phone number:

Postal address:

What information would you like to re-use:

How will you be re-using this information:

- ☐ Broadcast/ Exhibit
- ☐ Presentation
- ☐ Web use
- ☐ Publish
- ☐ Other (please specify)

Once completed please send this form to the Information Governance & Quality Directorate using any of the means below:

Email: information.governance@ic.nhs.uk

Fax: 0113 2547299

Postal address: Data Access and Sharing Team
Information Governance & Quality Directorate
The NHS Information Centre
1 Trevelyan Square
Boar Lane
Leeds, LS1 6AE