General Details	
Dataset Title:	Hospital Admissions: Summary Statistics, 2003/04
Domain(s):	Health and Care
Time Period of Dataset(s):	1 April 2003 – 31 March 2004 (Financial Year)
Geographic Coverage:	England
Lowest Area Output:	Middle Layer Super Output Area (MSOA)
Supplier:	NHS Information Centre for Health and Social Care, formerly Health and Social Care Information Centre (HSCIC)
Department:	HES
National Statistics Data?	Not National Statistics – this information based on administrative data does not comply fully with National Statistics
No. of Variables (excluding area names and codes):	9

Scope and Purpose

Hospital Episodes Statistics (HES) is a data warehouse containing records for episodes of admitted patient care to NHS hospitals in England. It includes information on the patient, when and where they were treated, diagnoses and operations, type of admission and various supplementary and derived data. The information has many uses and supports a wide variety of work within the Department of Health, other government departments, the NHS, Public Health Observatories, medical research and the general public.

Finished Admission Episode (FAE) counts from HES have been produced at England, Government Office Region (GOR), Local Authority (LA) and MSOA level for the 2003/04 financial year, and information covers the following specified diagnosis, cause and operative procedures:

1. Diagnosis - Coronary Heart Disease
2. Diagnosis – Cerebrovascular Disease (including Stroke)
3. Diagnosis - Cancer (excluding non-melanoma skin cancer
4. External Cause - Falls (basic accidental falls)
5. Operation – Coronary Artery Bypass Graft (CABG) and Percutaneous Transluminal Coronary Angioplasty (PTCA) (Heart)
6. Operation - Hip Replacement (Total hip replacement)
7. Operation - Knee Replacement (Total knee replacement)
8. Operation - Cataract

Please see Appendix A for further details on the counts included in this dataset.

The data presented here focus upon 8 specific diagnoses and procedure groups which generally have a higher prevalence amongst older members of the population. Data is also provided for ‘All admission episodes’ in 2003/04 regardless of diagnoses or procedure.

Cancer, coronary heart disease (CHD) and cerebrovascular disease are major causes of death (accounting for around two-thirds of all deaths) and ill-health. There is a strong social gradient in these diseases, with incidence and mortality higher among more deprived groups. The Public Health White Paper “Choosing Health – Making Healthier Choices Easier” sets out the importance of giving a priority to tackling health inequalities. The Department of Health has set Public Service Agreement targets to reduce mortality rates from cancer and from cardiovascular disease (covering both CHD and cerebrovascular disease), with a requirement that faster progress should be made in deprived areas with poor health.

Administrative Procedures - Background Information

Hospital Episode Statistics (HES) provides information about the admitted patient treatment delivered by NHS hospitals in England. It includes private patients treated in NHS hospitals, patients who were resident outside of England and care delivered by the independent sector funded by the NHS. This includes day case surgery. Data for out-patient appointments are also collected in HES but not included in this data. NHS Trusts submit their data to the Secondary Uses Service (SUS) for onward transmission to commissioners and HES is built from a national, annual extract of these records. The data are then cleaned, validated and new fields derived. (For further information, please see the 'Validation and Quality Assurance' section).

The HES data provided here is for Finished Admission Episodes (FAE).

Concepts and Definitions

Finished Admission Episode (FAE):

A finished admission episode is the first period of in-patient care under one consultant within one healthcare provider. A FAE is counted irrespective of diagnosis or treatment, but mothers giving birth in hospital, babies delivered in hospital and patients formally detained under the provisions of mental health legislation or long-term (over one year) psychiatric patients have been excluded.

It is not uncommon for a given person to be admitted to hospital more than once during a year. It is for this reason, that the number of FAE's shown in the dataset may exceed the number of individual persons treated within the year.

In order to assist in the estimation of disease incidence resulting in hospitalisation, and also the analysis of clinical practice, the average number of episodes attributable to individual persons has been approximated by analysing patient records submitted during the 2003/04 financial year.

Coronary Heart Disease (CHD) – 1.50 in-patient admissions per person annually (e.g. every 15 admission episodes with a CHD diagnosis will, on average, be generated as a result of 10 individuals admitted with CHD).

Cerebrovascular Disease (including Stroke) – 1.26 in-patient admissions per person annually (e.g. every 13 admission episodes with a Cerebrovascular Disease (including Stroke) diagnosis will, on average, be generated as a result of 10 individuals admitted with a stroke).

Cancer (excluding non-melanoma skin cancer) – 3.20 in-patient admissions per person annually (e.g. every 32 admission episodes with a cancer diagnosis will, on average, be generated as a result of 10 individuals admitted with cancer).

Falls - 1.06 in-patient admissions per person annually (e.g. approximately every 11 admission epsiodes due to a fall will, on average, be generated as a result of 10 individuals being admitted for a fall).

Note: the ratios give an estimation of the number of finished admission episodes likely to be generated by each admitted patient during the year. This includes data for patients where their age and/or sex was unknown. Episode/person ratios have not been calculated for the operation groups, as they would be less meaningful. Note, however, that in the case of Cataract and Hip operations, it is not uncommon for the same person to undergo two such treatments during a year.

It is also necessary to point out that a patient could have a FAE for diagnoses or operations different to those listed. For example, they may be in hospital for a diagnosis other than a stroke, coronary heart attack or cancer, or may have had an operation for something other than a hip replacement, cataract treatment or an operation involving the heart. This is why the variable totals in this dataset will not sum up to the total FAE count.

All Diagnoses – count of episodes:
These figures represent a count of all FAE's where the diagnosis was mentioned in any of the 14 (7 prior to 2002-03) diagnosis fields in a HES record.

Primary Diagnosis:
The primary diagnosis is the first of these 14 diagnosis fields (7 prior to 2002-03) in the HES data set and provides the main reason why the patient was in hospital.

Secondary Diagnosis:
The remainder 13 of the diagnosis fields (6 prior to 2002-03)) are the secondary diagnosis fields in HES that show other diagnoses relevant to the episode of care.

All Operations – Count of episodes
These figures represent a count of all FAE's where the procedure was mentioned in any of the 12 (4 prior to 2002-03) operation fields in a HES record.

Main Operation:
The main operation is the first recorded operative procedure on in the HES data set and is usually the most resource intensive procedure performed during the episode.

Patient counts:
Patient counts are based on the unique patient identifier HESID. This identifier is derived based on patient's date of birth, postcode, sex, local patient identifier and NHS number, using an agreed algorithm.

Provider: This is the organisation that provided the admitted patient care, typically an NHS, or Primary Care, Trust.

Middle Super Output Area (MSOA)

Super Output Areas (SOA) are part of a new geographic hierarchy designed by the Office for National Statistics to improve the reporting of small area statistics in England and Wales. There are 3 layers to super output areas including lower, middle and higher super output areas. There are 6,780 Middle Layer SOAs' in England; 413 in Wales).Note that Neighbourhood Statistics also includes the Isles of Scilly as a single Middle Layer SOA. However, due to the fact that the area has fewer than 5000 residents it cannot be a true Middle Layer SOA and so is not included in these national totals. In fact the Isles of Scilly Middle Layer SOA is exactly the same as the single Lower Layer SOA on the islands. It has been treated as a pseudo-SOA, however, to give Middle Layer SOA datasets complete national coverage.

Grossing:
Figures published by HES may be adjusted to take account of shortfalls (or duplication) in the records received, and also for missing diagnoses and operations codes. This generally results in a slight uplifting of the national counts. However, grossing is generally not appropriate at a low geographic level and data for Neighbourhood Statistics are ungrossed.

Datayear:
Each HES database table contains records relating to a particular reporting year (the period 1st April to the 31st March). This is often referred to as the HES datayear.

Data Classifications

Standard Classifications used (if any):	All Diagnoses coded according to the International Classification of Diseases and Related Health Problems – Tenth Revision (ICD-10). More information on ICD-10 is available. All Operations coded according to the Classification of Surgical Operations and Procedures, Fourth Revision (OPCS-4.2). More information about OPCS-4 is available.
Further Details about Classifications:	**ICD-10** The following codes are included in this dataset: Codes in the range of I20 to I25 (Coronary Heart Disease). Codes in the range of I60 to I69 (Cerebrovascular Disease including Stroke). Codes in the range of C00 to C97, (excluding C44) (Cancer

	excluding non-melanoma skin cancer). Codes in the range of W00 to W19 (Basic accidental falls). A full list of these codes is available. **OPCS-4.2** The following codes are included in this dataset: Codes in the range of K40 to K44 (Coronary Artery Bypass Graft). Codes from K49, K50.1 (Percutaneous Transluminal Coronary Angioplasty. Note: *K50.8* paired with L71.8 has also been included which reflects a percutaneous transluminal rotary blade angioplasty. Codes in the range of W37 to W39 (Total hip replacement). Codes in the range of W40 to W42 (Total knee replacement). Codes in the range of C71 to C75 (Cataract). A full list of these codes is available.
Edit and Imputation Procedures	
Figures have not been adjusted for shortfalls in data (i.e. the data are ungrossed).	
Validation and Quality Assurance	

HES go through several stages of data quality checks and measures to make the data as useable as possible. A number of checks are employed throughout the year to monitor the data and a dialogue exists between trusts and the HES data quality team to help resolve problems before the end of each financial year. HES cleaning processes are then applied to pick up anomalies in the data set.

- Manual cleaning
  This is a set of cleaning rules individually specified (usually at trust level), applied to the data to correct any anomalies. These are mainly to remove duplicates or reassign data appearing under the wrong trust code. Some cleans are specified by the trust, for example to correct data where a system problem has made it impossible for the trust to correct the data.

- Auto cleaning and derivation
  A set of rules are applied to the data to 'standardise' it into a format that is easy for analysts to use, eg. by overwriting inappropriate values that are seen to be inconsistent and replacing invalid values with a default code for 'not known'. At the same time, new fields are derived (eg. age of patient, duration of elective wait, length of stay).

Nevertheless, HES is reliant on trusts to provide complete, accurate and reliably coded data and some inaccuracies will remain where it is not possible to compensate for shortcomings. Further information on HES data quality can be found on the HES website.

There are a number of data quality indicators for HES data that are relevant to these tables, although they measure the validity rather than the accuracy of the data. These are listed below with their score in 2003/04:

- Record Linkage Component (examines the validity of date of birth, gender and the postcode entries in records): 99.6%

- Diagnosis Component (assesses the proportion of missing or invalid primary diagnosis codes): 97.5%

- Operation Codes Component (assesses the proportion of invalid operative procedure codes): 99.9%

Above results indicate that the data quality of the fields used to produce 2003/04 table were reliable.

Geographic Referencing

Please note: From 2002/03 all datasets have been referenced to the boundaries in place at the time of collection. Since there have been no boundary changes at Local Authority (LA) level in England that affect comparability, all datasets are presented here on the 2004 NeSS Geography Hierarchy. Publishing the full series on the same hierarchy will allow NeSS website users to make on-screen time series comparisons between datasets.

The Local Authority to which the FAE belonged was derived from the postcode of the patient, where given and where English (FAEs were listed in this dataset only if the patient had an English postcode. This occurred for around 99% of relevant cases).

Records with missing postcodes, or those not matched to the ONS geographic reference data postcode file for Neighbourhood Statistics, were excluded from the data. The criteria for a postcode "not matching" to the postcode file included instances where no postcode was recorded for a FAE, or where a postcode was not recognised by the look-up file (for example an incomplete postcode, or a postcode for a new building that has not yet been listed in the current look-up file).

In a small number of cases HES references the Middle Layer Super Output Area to a different Local Authority than the NeSS postcode directory. This affects less than 0.05% of total England cases, and, in the worse affected LA, less than 0.5% of cases. The data provided has been corrected in order to preserve comparability to the NeSS postcode directory. This does mean that some of the figures provided for a selected number of Local Authorities will not be comparable to other Hospital Episode Statistics at Local Authority level.

Data Quality

Relevance:	HES provides the only consistent, national database of all admitted patient care in NHS Trusts in England. Additional activity not covered includes Accident and Emergency attendances, admission of English patients to hospitals outside of England (unless NHS commissioned and recorded) and primary care.
Accuracy:	Records with 'not known' or 'not specified' sex values were excluded from the diagnosis and operations grouped data (this accounted for 0.012% of the Finished Admission Episodes). Records with 'not known' or 'missing' age values were excluded from the diagnosis and operations grouped data (this accounted for 0.012% of the Finished Admission Episodes). Note that 'All Finished Admission' episodes does include 'not known' or 'not specified' sex values as well as 'not known' or 'missing' age values.
Timeliness and Punctuality:	HES annual data are generally published around December for the previous financial year.
Accessibility and Clarity:	Once released, a limited number of freely available data tables are published on the HES website. These include tables which present data by 1) 'hospital provider'-the organisation that admitted the patient; 2) 'Strategic Health Authority (SHA) of residence'-the SHA area in which the patient lives; and 3) 'responsible Primary Care Trust (PCT)'-the PCT area with whom the patient is registered. Further tabulations and extracts of HES data are available for a

	variety of uses by researchers, public health bodies, etc. These data are available by NeSS geographical areas (including Local Authority and Middle Layer Super Output Area) if requested. Confidentiality restrictions apply.
Comparability:	The Hospital Admissions: Summary Statistics data series provided on Neighbourhood Statistics is currently available for 2002/03, 2003/04, 2004/05, 2005/06 and 2006/07. In the main, this data will be comparable to other years of HES data to be published on the Neighbourhood Statistics website. Other Hospital Episode Statistics are available on the Neighbourhood Statistics website including; age and sex breakdown information 'Hospital Admissions by Age and Sex'.
Coherence:	This dataset was referenced to the same versions of ICD-10 for diagnosis groups as the other HES datasets (also to be published on the Neighbourhood Statistics website). This dataset was referenced using OPCS4.2 for procedure groups. 2006/07 HES datasets published on the Neighbourhood Statistics website will reference OPCS4.3. This will not cause much impact between the datasets due to the broad categories included as variables within this dataset.

Disclosure Control

Every effort has been made by the NHS Information Centre to ensure the data are not disclosive. A statistical disclosure control software tool called Tau Argus has been used to protect potentially disclosive cells in this dataset. Any disclosive cells due to the availability of other Hospital Episode Statistics available on the Neighbourhood Statistics website have also been protected.

The Office for National Statistics carries out a number of checks to safeguard confidentiality. In accordance with standard procedures these datasets have been reviewed and approved for release.

Data have been suppressed in this dataset to protect both the confidentiality of individual information and the potential statistical instability caused by low counts.

Sources for Further Information or Advice

Tabulations covering a wide range of NHS in-patient activity, plus explanations of the content and further information on HES, may be downloaded from the Hospital Episode Statistics web pages .

Specific information about this dataset can be obtained from:
SUS/HES Analysis Team
The NHS Information Centre for Health and Social Care
Trevelyan Square
Boar Lane
Leeds
LS1 6AE

Telephone: 0845 300 6016
e-mail: hes.questions@ic.nhs.uk

Appendix A

Details of the information held on this dataset are outlined below.		
Variable Name	**Data Type**	**Variable Description**
All Finished Admission Episodes	Count	The total number of finished admissions to hospital during the period 1st April 2003 to 31st March 2004 (i.e. irrespective of diagnosis or treatment).
Coronary Heart Disease (CHD); Diagnosis	Count	FAE's where the mention in any of the diagnosis fields (including the primary diagnosis field) was Coronary Heart Disease
Cerebrovascular Disease (including stroke); Diagnosis	Count	FAE's where the mention in any of the diagnosis fields (including the primary diagnosis field) was Cerebrovascular disease (including stroke)
Cancer (excluding non-melanoma skin cancer); Diagnosis	Count	FAE's where the mention in any of the diagnosis fields (including the primary diagnosis field) was Cancer (excluding non-melanoma skin cancer).
Falls (basic accidental falls); External Cause	Count	FAE's where the External Cause mentioned in any of the diagnosis fields (including the primary diagnosis field) was a fall.
Coronary Artery Bypass Graft (CABG) and Percutaneous Transluminal Coronary Angioplasty (PTCA); Operation	Count	FAE's where the operation mentioned in any of the operative procedure fields (including the main operation field) was CABG or PTCA.
Hip Replacement; Operation	Count	FAE's where the operation mentioned in any of the operative procedure fields (including the main operative procedure field) was a total hip replacement.
Knee Replacement; Operation	Count	FAE's where the operation mentioned in any of the operative procedure fields (including the main operative procedure field) was a total knee replacement.
Cataract; Operation	Count	FAE's where the operation mentioned in any of the operative procedure fields (including the main operative procedure field) was of the cataract.