



Data Resource Profile

Data Resource Profile: Hospital Episode Statistics Admitted Patient Care (HES APC)

Annie Herbert,^{1,2} Linda Wijlaars,¹ Ania Zylbersztejn,^{1,3}
David Cromwell^{4,5} and Pia Hardelid^{1,6*}

¹Population, Policy and Practice Programme, UCL Institute of Child Health, ²Department of Behavioural Science and Health, UCL Institute of Epidemiology and Healthcare, ³Farr Institute of Health Informatics Research, University College London, London, UK, ⁴Department of Health Services Research and Policy, London School of Hygiene & Tropical Medicine, London, UK, ⁵Clinical Effectiveness Unit, Royal College of Surgeons of England, London, UK and ⁶Department of Primary Care and Population Health, University College London, London, UK

*Corresponding author. Population, Policy and Practice Programme, UCL Great Ormond Street Institute of Child Health, 30 Guildford Street, London WC1N 1EH, UK. E-mail: p.hardelid@ucl.ac.uk

Editorial decision 16 January 2017; accepted 23 January 2017

Data resource basics

Scope

Hospital Episode Statistics Admitted Patient Care (HES APC) data are collected on all admissions to National Health Service (NHS) hospitals in England. HES APC also covers admissions to independent sector providers (private or charitable hospitals) paid for by the NHS.¹ It is estimated that 98–99% of hospital activity in England is funded by the NHS.² A hospital admission includes any secondary care-based activity that requires a hospital bed, thus including both emergency and planned admissions, day cases, births and associated deliveries. HES APC does not cover accident and emergency (A&E, emergency department) attendances or outpatient bookings; these data are held in separate HES databases. All HES databases are collated and curated by NHS Digital (previously the Health and Social Care Information Centre). In the financial year 2014/15 (April to March), 18 731 987 hospital episodes from 451 different NHS hospital trusts (known as ‘providers’) were recorded in HES APC.³

Purpose of data collection

The need for national data collection on hospital activity to inform management and planning of services was first recognized in the early 1980s by a Department of Health working group.⁴ Following these recommendations, a national programme was progressively rolled out, starting in 1987 and obtaining continual national coverage by (financial year) 1989/90.⁵ Since 2004/05, HES APC has also served as the basis for ‘Payment by Results’ (PbR), a pay-for-performance system of secondary care reimbursement in the NHS internal market.⁶

Structure

HES APC data files are structured according to financial years. Each row in HES APC indicates a ‘Finished Consultant Episode’ (FCE). An FCE represents a continuous period of care under one consultant, and each is specified with a start and an end date. Episodes are labelled as ‘finished’ and entered in HES APC according to the financial year in which they end. Consequently, episodes that

start in one financial year and end in another will be classified as unfinished in the starting financial year, and finished in the ending financial year. Unfinished episodes need to be removed before analysis to prevent double counting.

A hospital admission in HES APC is referred to as a 'spell', defined as an uninterrupted inpatient stay at one hospital. A spell may include several FCEs if the patient was seen by multiple consultants during the same stay, but does not include transfers between hospitals. If a patient is transferred to a different hospital, a new spell begins.

In order to identify and measure continuous hospital stays, which include transfers to other hospitals, continuous inpatient spells (CIPs) need to be derived. Although CIP identifiers are not provided in standard HES APC extracts, methods for linking FCEs into CIPs are available,⁷ including that recommended by NHS Digital.⁸

Research uses

HES APC has been frequently used for research and service evaluation, due to its universal coverage, long period of data collection and the ability to follow individuals over time. HES APC offers the opportunity to estimate population-based admission and procedure rates by condition and type of procedure, compare hospital performance and create hospital-based cohorts for short- or long-term follow-up. Since HES APC covers all births in NHS hospitals, representing 97.3% of births in England,⁹ it is also possible to create nationally representative birth cohorts.

Processing cycle and frequency of data collection

Upon discharge from the care of a particular consultant, the treating clinician completes a discharge summary for the patient of diagnoses made and procedures carried out during that FCE (where procedures include surgery, diagnostic imaging, ventilation and infusion/transfusion therapy). Discharge summaries are forwarded to a clinical coding department in the hospital, who enter the information onto the local electronic patient information database. Clinical coders undergo nationally accredited training programmes and follow standardized rules for translating information on discharge summaries into clinical codes.^{10,11}

Every month, data are extracted from local hospital databases to the Secondary User Service (SUS), a national data warehouse housed within NHS Digital.¹² Data from the SUS are extracted both for purposes of hospital reimbursement under PbR, and separately to create a provisional monthly HES extract. NHS Digital carry out basic data checks and cleaning, add geographical fields based on patient postcodes, and attach pseudonymized patient

identifiers ('HESIDs') to each episode.^{13,14} At the end of each financial year, NHS Digital allow hospitals one further data submission to HES (the 'Annual Refresh'), after which a provisional annual HES extract is produced for final review by hospitals. Once the Annual Refresh has been checked, a final annual HES dataset is made available.¹²

Linkage within HES APC

From 1997/98 onwards (when patients' NHS numbers became a mandated return from hospitals), HES APC episodes have been linked longitudinally to the same patient by tagging episodes with the HESID. This alphanumeric variable allows patient follow-up, yet avoids the need for supplying patient identifiers to researchers. The methods used to generate the HESID have been described elsewhere.¹⁵ Each HES APC extract contains a unique set of HESIDs to reduce the risk of individual disclosure through merging separate data extracts supplied to different research teams.

Linkage to other datasets

HES APC data can be linked to other datasets held by NHS Digital, including HES A&E attendances (from 2007/08), HES Outpatient appointments (from 2003/04), adult critical care (from 2008/09), diagnostic imaging data (covering all radiology procedures from 2012/13), the Mental Health Services Dataset (for all adult community and outpatient mental health care contacts from 2006/07) and Patient Reported Outcome Measures (pre- and postoperative questionnaires filled out by patients undergoing knee or hip replacements, varicose vein surgery or groin hernia repair from 2009/10). Secondary users can link these datasets because the same HESID algorithm is applied to each dataset.

HES APC is also routinely linked to a number of external datasets. The Clinical Practice Research Datalink,¹⁶ a large UK primary care database, is linked to HES APC on a monthly basis. HES APC is linked to dates and causes of non-hospital deaths from the Register of Deaths in England and Wales held by the Office for National Statistics (for deaths registered since 1 January 1998), also on a monthly basis.¹⁷ Only deaths of patients recorded in HES APC are available through this linkage (i.e. deaths of persons who have not had a hospital admission since April 1997 are not included).

NHS Digital also provides a trusted third-party bespoke linkage service, through which secondary users can request that HES APC data be linked to other external datasets. For example, both national disease registries (such as the

National Joint Registry¹⁸ and the UK Renal Registry¹⁹) and well-established cohort studies including Whitehall II²⁰ and the Hertfordshire Cohort Study²¹ have been linked to HES APC. Secondary users need to obtain the appropriate approvals to enable these linkages.

Measures

Clinical and patient data

HES APC provides detailed clinical, demographic and organizational information for each FCE (see Table 1), with 270 variables available in the core dataset. Apart from data on diagnoses and procedures, HES APC contains information on dates of admission, operations and discharge, admission method (e.g. emergency or planned), care provider and many geographical variables mapped from a patient's postcode. The local health geographies and hospital providers in England have changed several times since 1997, and thus care needs to be taken to ensure continuity when carrying out local or provider level analyses that use HES APC data covering many years.

Socioeconomic status is measured by the Index of Multiple Deprivation 2004 (IMD), a small area-based indicator constructed from several different measures of deprivation.²² IMD is measured at Lower Super Output Area (LSOA) level, where an LSOA contains between 400 and 1200 households.²³ Individual-level measures of socioeconomic status (e.g. education level or income) are not available. Detailed information on variables available, specific cleaning rules and coding used are available in the HES APC Data Dictionary provided by NHS Digital.²⁴

Diagnoses are coded using the International Classification of Diseases version 10 (ICD-10).²⁵ ICD-9 was used between April 1989 and March 1995. The

number of diagnosis fields has increased over time: since April 2007, each FCE can have up to 20 ICD-10 codes entered (up from 7 codes before April 2002 and 14 in April 2002–March 2007). Each FCE has one primary diagnosis, which accounts for the majority of the length of stay of the FCE. The other diagnoses are referred to as comorbidities. According to NHS Digital cleaning rules, each FCE must have at least one primary diagnosis, although it may be recorded as unknown (ICD-10 code R69).

Operations and other interventions are coded using a UK-specific system, the Office of Population Censuses and Surveys Classification of Interventions and Procedures (OPCS, currently version 4.7).²⁶ This has evolved over time as new techniques and technologies have been introduced. A history of versions in use is available from the NHS Digital coding standards website.²⁶ Each FCE may have up to 24 operations recorded (up from 4 before April 2002 and 12 in April 2002–March 2007), but procedure fields are left empty if patient management did not require an intervention covered by OPCS (e.g. where the primary treatment was a drug regimen or observation). A primary procedure is selected for each FCE as that which is the most resource-intensive, but a procedure may be described using more than one code to indicate surgical approach, anatomical location and side of procedure (e.g. stent placed under radiological control in femoral artery of left leg). Dates are also entered for each procedure.

Birth and delivery information

Each birth event in HES APC generates at least two FCEs: one delivery episode and one or more birth episodes. Each delivery and birth episode includes an additional 'maternity tail', with detailed fields including the baby's

Table 1. Selection of key data fields available for each finished consultant episode (FCE) in HES APC data²²

Patient	Admission/FCE	Clinical	Geography	Provider/organisational	Maternity/birth (only in maternity tail)
HESID	Episode start date	Diagnoses (up to 20)	Government office region	Care provider (hospital)	Gestational age
Age at admission	Episode end date	Operations (up to 24)	Local authority	General practice of patient	Number of previous births
Age at discharge	Date of admission	Operation dates (up to 24)	Clinical commissioning group		Birth weight
Sex	Date of discharge		Index of multiple deprivation (IMD) 2004 rank, deciles and domains		Maternal age
Ethnic group	Admission method (e.g. - planned, emergency, birth)	Consultant specialty (admitting and treating consultant)			Mode of delivery
	Discharge method				Baby number (for multiple births)
	Admission source				
	Discharge destination				
	Waiting time (from date of decision to admit to date of admission)				

birthweight, gestational age, birth order (for multiple births), mode of delivery and maternal age (Table 1). The maternity tail is based on information entered via local maternity databases. Unlike the diagnostic and procedure fields, the maternity tail data fields use HES-specific categories rather than standardized classifications, and it is not a mandated return to NHS Digital. This leads to large variations in data completeness and quality.^{27,28} It is not possible to directly link a mother and a baby in HES APC; that is, the mother's HESID is not copied to the baby's birth record. However, linkage between mother and baby is possible using probabilistic methods.²⁹

Hospital use in England

Both numbers and rates of hospital admissions have increased during the period of HES APC data collection

(Figure 1), particularly among older adults (aged 60-74 and 75+). Between 1998/99 and 2014/15, the overall FCE rate has increased by 40% from 24.5 per 100 person-years to 34.3 per 100 person-years, with the steepest increase (73.0%) in adults aged 75+.

Since HES APC covers all hospital admissions, infants and older adults (aged 65+) are over-represented in HES APC compared with the general population of England (Table 2).

Data resource use

Although no up-to-date bibliography of published research based on HES APC is curated by the data providers, a 2013 systematic review identified 148 articles using HES APC data published between 1989 and July 2011.³⁰ We carried out a subsequent search on PubMed on the 8 June 2016 using the search term 'Hospital Episode Statistics' for article

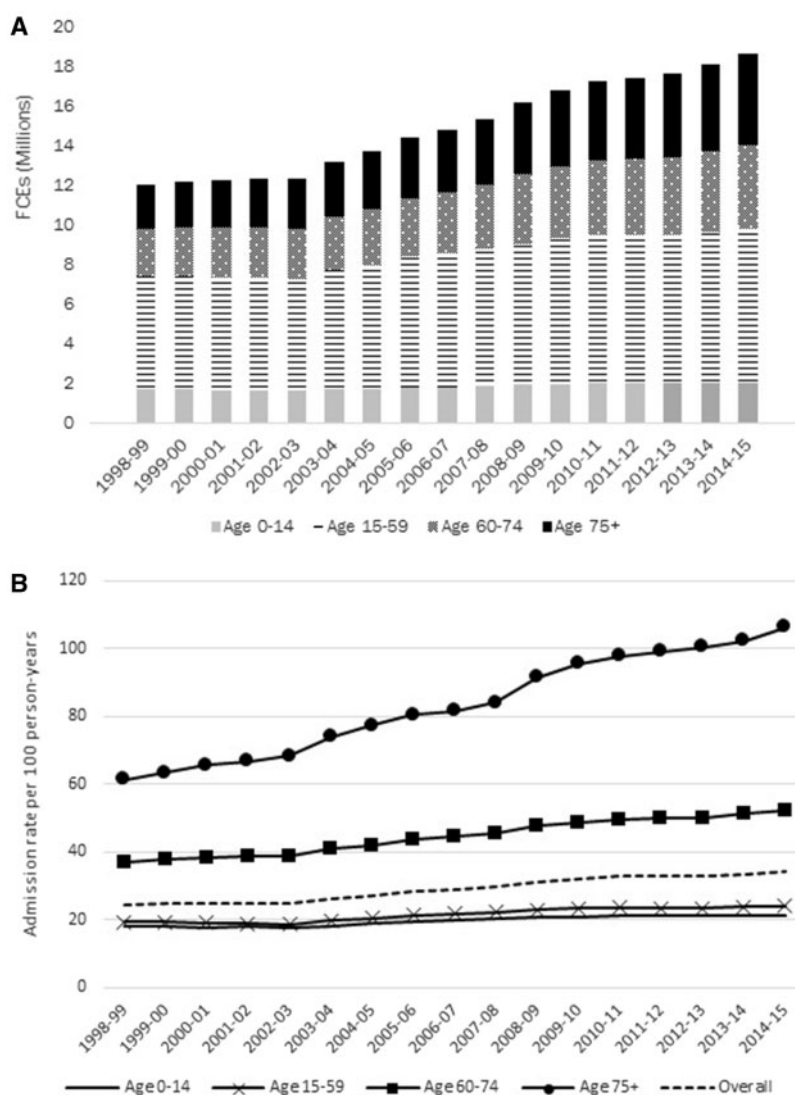


Figure 1. A) Number of finished consultant episodes (FCEs) by age group from financial years 1998/99 to 2014/15; and B) episode rates by age group per 100 person-years. Denominators for rates are based on mid-year population estimates for England⁷⁸.

Table 2. Demographic characteristics of HES APC patients compared with general population of England

Characteristic	HES APC ^a	England ^b
Finished consultant episodes	18731964	
Admissions	15892434	
Admission type		
Emergency	5615707 (30.0)	
Waiting list	6119234 (32.7)	
Planned	2154564 (11.5)	
Other	2002929 (10.7)	
Sex		
Male	8359362 (44.6)	26773200 (49.3)
Female	10370245 (55.4)	27543400 (50.7)
Gender unknown	2357 (0.01)	–
Age		
0 years	1013476 (5.4)	664183 (1.2)
1–4 years	454461 (2.4)	2766774 (5.1)
5–14 years	568902 (3.0)	6245420 (11.5)
15–24 years	1167439 (6.2)	6837371 (12.6)
25–34 years	1880715 (10.0)	7425591 (13.7)
35–44 years	1573273 (8.4)	7103408 (13.1)
45–54 years	1986116 (10.6)	7635651 (14.1)
55–64 years	2319214 (12.4)	6100512 (11.2)
65–74 years	3013044 (16.1)	5162873 (9.5)
75–84 years	2941250 (15.7)	3099319 (5.7)
85+ years	1711354 (9.1)	1275516 (2.3)
Missing	102720 (0.5)	

Numbers within parentheses represent proportions of FCEs (for HES APC) and proportions of persons (for England)

^aData source: HES APC 2014–15.³

^bONS 2014 mid-year population estimates.⁷⁵

abstracts published since July 2011. We identified 264 relevant publications where the primary analysis involved the use of HES APC data, and a further 130 papers where HES data had been linked to cohorts created in other datasets. The annual number of publications using HES APC data has increased from 2 in 1993³⁰ to 88 in 2015.

Published studies using HES APC data have covered a diverse range of topics. They have explored the incidence of conditions across regions and over time.^{31,32} They have also examined cross-sectional and longitudinal patterns of treatment by organization,³³ including comparing NHS and privately contracted providers³⁴ or regions,^{35,36} both from descriptive and analytical perspectives. Regional comparisons have included evaluating the impact of clinical evidence³⁷ or guidelines³⁸ as well as health care policies.³⁹ They have examined the outcome of medical as well as surgical therapies (such as survival,⁴⁰ short-term postoperative mortality,⁴¹ complications,⁴² reoperation⁴³ and hospital readmissions⁴⁴), with some seeking to identify factors that are associated with these outcomes, in terms of both patient characteristics^{45,46} and organizational factors such as surgical volume⁴⁷ or day of week.⁴⁸ Methodological studies

include creating coding frameworks,²⁸ applying comorbidity scores,⁴⁹ developing risk prediction models⁵⁰ and using look-back methods to impute missing data items.⁵¹

Many high profile routinely produced reports on the quality of secondary care are based on HES APC data. These include hospital mortality monitoring reports produced by NHS Digital⁵² and commercial organizations,⁵³ and research reports by independent think-tanks⁵⁴ and Royal Medical Colleges.⁵⁵

Strengths and weaknesses

Coverage

The key strength of the HES APC database is its universal coverage, which provides an unselected sample of hospital episodes. The large size of HES APC makes it possible to precisely estimate admission rates and capture outcomes for rare conditions, including congenital anomalies or specific cancers.

Longitudinal linkage

Another strength is the possibility to longitudinally link patients using the HESID, allowing for the creation of HES-based cohort studies if a suitable inception date can be identified. The long period of data collection of HES (currently up to 19 years) allows long-term follow-up of admitted patients, which has allowed the development of risk prediction models for distal outcomes.⁴⁴

Standardized coding

ICD-10 coding of clinical diagnoses offers the opportunity to use HES APC for international comparisons of secondary care use. Since ICD-10 is used in hospital administrative data across the UK, Europe, Canada, Australia and New Zealand, HES APC has been used to assess the impact of differential health policy between NHS systems and internationally.^{56–58} International studies using HES APC include cross-country comparisons of the incidence of neonatal abstinence syndrome⁵⁹ and non-small cell lung cancer.⁶⁰ Nonetheless, international comparisons are challenging due to differences between countries in admission thresholds, organization of care provision, and whether secondary care is free at point of use or requires health insurance or other payment.

HES APC episodes are readily linked to information on costs of care, due to the ability to match each episode to a Healthcare Resource Group, and hence a unit cost.⁶¹ This makes HES APC an important data resource for health economics.^{62–64}

Coding variation

One of the key challenges in interpreting HES APC is the reliance on diagnostic and procedure codes for identifying study participants and outcomes. Despite centrally issued coding rules, clinical coders rely on the quality and detail of completed discharge summaries to enter data consistently. Consequently, diagnostic coding practices vary between hospitals, particularly for comorbidities.⁶⁵

Since the roll-out of PbR, financial incentives now exist for hospitals to improve coding depth in order to ensure accurate reimbursement. This has led to an increase in the number of diagnostic codes used and improvements in coding accuracy.^{7,66} The introduction of PbR therefore poses challenges for interpreting time-series studies using HES APC data, and care must be taken to not overinterpret results identifying increasing complexity of cases admitted.⁷

Sensitivity to admission thresholds

Since HES APC covers only admitted patients, it is sensitive to variation between hospitals or over time in admission thresholds. The introduction of the four-hour waiting target in A&E departments in 2004 has been suggested as a contributing factor for the increase in rates of emergency admissions in children during the 2000s.^{67,68} Changes in thresholds for emergency admissions can be examined using linked HES A&E data;⁶⁹ however, variation in admission thresholds for planned procedures cannot readily be determined using HES datasets.

Missing data

Although age, sex and clinical characteristics are well completed in HES APC (see Table 2), data on ethnicity are not. Ethnicity has been a mandated return for all NHS contacts since 1991. Although ethnicity recording has improved over time, the proportion of patients with a known ethnicity recorded was still only 85% in 2011, up from 41% in 1997.⁷⁰

Further, there is a high proportion of missing data in the maternity tail fields (see Figure 2). Postcodes were not extracted from the SUS for birth episodes prior to 2013/14, which means earlier birth episodes cannot be mapped to geographical variables, including the Index of Multiple Deprivation (IMD).⁷¹ As an example, completeness of the IMD decile variable for singleton birth episodes in 2012/13 was 7.8%, compared with 81.9% in 2013/14.

Quality of internal linkage

The HESID linkage algorithm relies heavily on the accurate recording of NHS number across all hospital episodes to avoid missed matches (FCEs that have failed to link to a patient). Consequently, there is a substantial proportion of missed matches in HES APC. A recent estimate puts the HESID missed-match rate at 4%,⁷² leading to an underestimation of readmission rates by 3.8%. NHS numbers were not provided at birth until 2002, meaning that linkage within HES APC and to other HES and external datasets is not reliable for births before 2002/3.⁷³

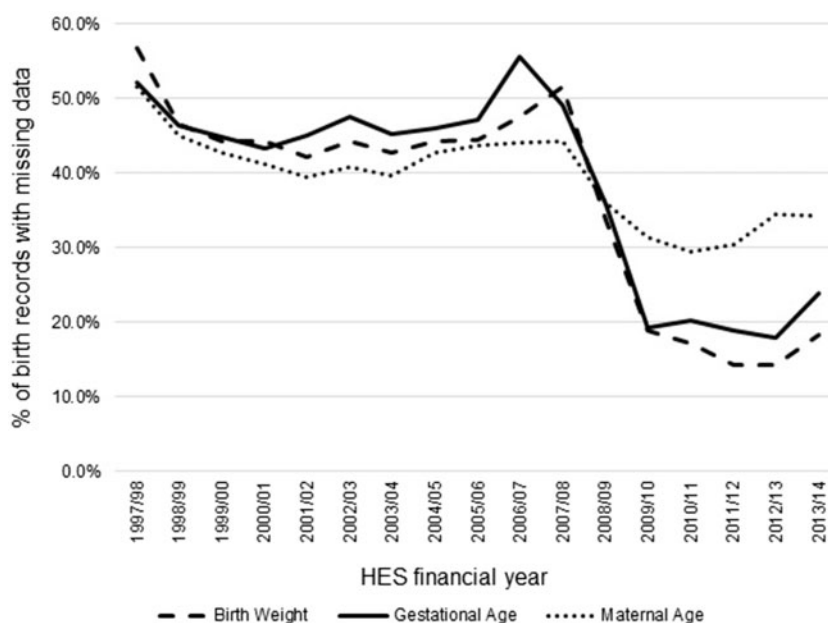


Figure 2. Proportion of birth records with missing data for selected variables in the maternity tail from financial years 1997/98 to 2013/14.

Scope limitations

HES APC covers higher dependency (HDU) or intensive care unit (ICU) periods, but it does not contain ‘flags’ to identify such stays, nor detailed information on level of care or HDU/ICU interventions. A separate HES dataset covers adult critical care from 2008/09,⁷⁴ whereas data relating to neonatal or paediatric intensive care are collected through systems external to NHS Digital.

Data on drugs prescribed through hospital pharmacies to inpatients are not available in HES APC. There is currently no national individual-level hospital prescribing database for England.

Opt-outs

Patients who do not wish their records to leave NHS Digital can lodge a ‘type 2 opt-out’ with their primary care practice.⁷⁵ From 29 April 2016, any records (including in previous financial years) relating to persons who have opted out in any NHS Digital dataset (including HES APC) will therefore be removed before supply to secondary users. Overall, for the 2014/15 HES APC annual extract, 2.3% of episodes will be removed, with substantial geographical variation in opt-out rates.⁷⁵

Data resource access

Access to HES APC data is provided by NHS Digital for the NHS, government, researchers and commercial health care bodies. Those requesting an extract of the data must show that their work will support health and social care and improve health.⁷⁶ Data cannot be released for solely commercial purposes.

Data are requested through the online Data Access Request Service (DARS). Applications are evaluated by the Data Access Advisory Group which check all data requests for patient-level data to evaluate whether there is an appropriate legal basis for data dissemination and that appropriate data security is in place. Details about HES applications and associated costs are available on the DARS website [<http://content.digital.nhs.uk/DARS>].

NHS Digital carries out audits to check that data users meet obligations regarding the terms and conditions of use, including disclosure control.⁷⁷

Funding

L.W. is supported by funding from the Department of Health Policy Research Programme through funding to the Policy Research Unit in the Health of Children, Young People and Families (grant reference number 109/0001). This is an independent report commissioned and funded by the Department of Health. The views expressed are not

Profile in a nutshell

- HES APC contains data on all admissions to National Health Service (NHS) hospitals in England, or to independent hospitals where the costs are met by the NHS. It was originally set up for purposes of management and planning of hospital services. Data are now also collected for purposes of reimbursing hospital activity.
- HES APC includes all hospital care episodes from the financial year 1989/90 onwards (1 April 1989–31 March 1990). Pseudonymized patient identifiers that allow for longitudinal follow-up of patients are available from 1997/98 onwards.
- HES APC data are entered from medical records by clinical coders in each hospital, according to national clinical coding standards. The database is collated and processed centrally by NHS Digital (previously the Health and Social Care Information Centre).
- Data fields exist for diagnoses, procedures, patient demographics (including ethnicity and area-level deprivation), admission and discharge dates, hospital and other variables.
- HES APC data can be linked to outpatient and emergency department attendances as well as datasets external to NHS Digital, including death registrations.
- Aggregate data are accessible via the NHS Digital website and individual-level data are available through the NHS Digital Data Access Request Service, subject to approval and a cost recovery charge.

necessarily those of the Department. A.Z.’s PhD studentship is supported by awards to establish the Farr Institute of Health Informatics Research, London, from the Medical Research Council, Arthritis Research UK, British Heart Foundation, Cancer Research UK, Chief Scientist’s Office, Economic and Social Research Council, Engineering and Physical Sciences Research Council, National Institute for Health Research, National Institute for Social Care and Health Research and Wellcome Trust (grant MR/K006584/1). P.H. is funded by a National Institute for Health Research postdoctoral fellowship (number PDF-2013-06-004). This article represents independent research funded by the National Institute for Health Research (NIHR). The views expressed are those of the authors and not those of the NHS, the NIHR or the Department of Health.

Conflict of interest: None declared.

Acknowledgement

We are grateful to Professor Ruth Gilbert for her helpful input regarding this paper.

References

- Health and Social Care Information Centre. *Hospital Episode Statistics (HES) Analysis Guide*. 2015 http://content.digital.nhs.uk/media/1592/HES-analysis-guide/pdf/HES_Analysis_Guide_Jan_2015.pdf (30 October 2016, date last accessed).
- National Audit Office. *Healthcare Across the UK: A Comparison of the NHS in England, Scotland, Wales and Northern Ireland*. 2012. <https://www.nao.org.uk/wp-content/uploads/2012/06/1213192.pdf> (16 June 2016, date last accessed).
- Health and Social Care Information Centre. *Hospital Episode Statistics, Admitted Patient Care 2014/15*. 2015. <http://content.digital.nhs.uk/catalogue/PUB19124> (27 October 2016, date last accessed).
- Black D. Data for management: the Korner Report. *Br Med J* 1982;**285**:1227–28.
- NHS Digital. *Hospital Episode Statistics*. 2016 <http://content.digital.nhs.uk/hes> (27 October 2016, date last accessed).
- Department of Health. *A Simple Guide to Payment by Results*. 2011 https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/213150/PbR-Simple-Guide-FINAL.pdf (18 March 2016, date last accessed).
- Hardelid P, Dattani N, Davey J, Pribramska I, Gilbert R. *Overview of Child Deaths in the Four UK Countries*. 2013. <http://www.rcpch.ac.uk/system/files/protected/page/CHR-UK%20MODULE%20B%20REVISED%20v2%2015112013.pdf> (3 March 2014, date last accessed).
- Health and Social Care Information Centre. *Methodology to Create Provider and CIP Spells from HES APC Data*. 2014. http://content.digital.nhs.uk/media/11859/Provider-Spells-Methodology/pdf/Spells_Methodology.pdf (27 October 2016, date last accessed).
- Office for National Statistics. *Characteristics of Birth 2, England and Wales*. 2015. <http://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/livebirths/bulletins/characteristicsofbirth2/2014-11-17> (28 February 2017, date last accessed).
- HSCIC Clinical Classifications Service. *National Clinical Coding Standards OPCS-4*. Leeds, UK: Health and Social Care Information Centre, 2016.
- HSCIC Clinical Classifications Service. *National Clinical Coding Standards ICD-10*. 5th edn. Leeds, UK: Health and Social Care Information Centre, 2016.
- Health and Social Care Information Centre. *The HES Processing Cycle and Data Quality*. 2014. http://content.digital.nhs.uk/media/1366/The-HES-processing-cycle-and-HES-data-quality/pdf/The_HES_Processing_Cycle_and_HES_Data_Quality_v3.pdf (27 October 2016, date last accessed).
- Health and Social Care Information Centre. *Cleaning Rules: Admitted Patient Care*. 2016. http://content.digital.nhs.uk/media/1367/HES-Hospital-Episode-Statistics-Inpatient-cleaning-rules/pdf/HES_APC_052015b.pdf (27 October 2016, date last accessed).
- Health and Social Care Information Centre. *Data Quality Checks Performed on SUS and HES Data*. 2014. http://content.digital.nhs.uk/media/13655/Data-quality-checks-performed-on-SUS-and-HES-data/pdf/Data_quality_checks_performed_on_SUS_and_HES_data_v2.pdf (27 October 2016, date last accessed).
- Health and Social Care Information Centre. *Methodology for Creation of the HES Patient ID (HESID)*. 2014. http://content.digital.nhs.uk/media/1370/HES-Hospital-Episode-Statistics-Replacement-of-the-HES-patient-ID/pdf/HESID_Methodology.pdf (27 October 2016, date last accessed).
- Herrett E, Gallagher AM, Bhaskaran K *et al*. Data Resource Profile: Clinical Practice Research Datalink (CPRD). *Int J Epidemiol* 2015;**44**:827–36.
- Health and Social Care Information Centre. *A Guide to Linked Mortality Data from Hospital Episode Statistics and the Office for National Statistics*. 2015. http://content.digital.nhs.uk/media/11668/HES-ONS-Mortality-Data-Guide/pdf/mortality_guide.pdf (27 October 2016, date last accessed).
- Stafford GH, Charman SC, Borroff MJ, Newell C, Tucker JK. Total hip replacement for the treatment of acute femoral neck fractures: results from the National Joint Registry of England and Wales at 3–5 years after surgery. *Ann R Coll Surg Engl* 2012;**94**:193–98.
- Fotheringham J, Jacques RM, Fogarty D, Tomson CR, El Nahas M, Campbell MJ. Variation in centre-specific survival in patients starting renal replacement therapy in England is explained by enhanced comorbidity information from hospitalization data. *Nephrol Dial Transplant*. 2014;**29**:422–30.
- Britton A, Milne B, Butler T *et al*. Validating self-reported strokes in a longitudinal UK cohort study (Whitehall II): Extracting information from hospital medical records versus the Hospital Episode Statistics database. *BMC Med Res Methodol* 2012;**12**:83.
- Simmonds SJ, Syddall HE, Walsh B *et al*. Understanding NHS hospital admissions in England: linkage of Hospital Episode Statistics to the Hertfordshire Cohort Study. *Age Ageing*. 2014;**43**:653–60.
- Office of the Deputy Prime Minister. *The English Indices of Deprivation 2004: Summary (Revised)*. 2004. <http://webarchive.nationalarchives.gov.uk/20100410180038/http://www.communities.gov.uk/documents/communities/pdf/131206.pdf> (27 October 2016, date last accessed).
- Office for National Statistics. *Super Output Area (SOA)*. 2011. <http://webarchive.nationalarchives.gov.uk/20160105160709/http://www.ons.gov.uk/ons/guide-method/geography/beginners-guide/census/super-output-areas-soas/index.html> (27 October 2016, date last accessed).
- Health and Social Care Information Centre. *HES Data Dictionary: Admitted Patient Care*. 2016. <http://content.digital.nhs.uk/media/19425/APC-DD-Final-Doc2/pdf/DD-APC-V7.pdf> (27 October 2016, date last accessed).
- World Health Organization. *International Statistical Classification of Diseases and Related Health Problems – 10th Revision*. Geneva: WHO, 2011.
- NHS Digital. *OPCS-4 Classification*. 2016. <http://systems.digital.nhs.uk/data/clinicalcoding/codingstandards/opcs4> (27 October 2016, date last accessed).
- Murray J, Saxena S, Modi N *et al*. Quality of routine hospital birth records and the feasibility of their use for creating birth cohorts. *J Public Health (Oxf)* 2013;**35**:298–307.
- Knight HE, Gurol-Urganci I, Mahmood TA *et al*. Evaluating maternity care using national administrative health datasets: how

- are statistics affected by the quality of data on method of delivery?. *BMC Health Serv Res* 2013;13:200.
29. Harron K, Gilbert R, Cromwell D, van der Meulen J. Linking data for mothers and babies in de-identified electronic health data. *PLoS One* 2016;11:e0164667.
 30. Sinha S, Peach G, Poloniecki JD, Thompson MM, Holt PJ. Studies using English administrative data (Hospital Episode Statistics) to assess health-care outcomes – systematic review and recommendations for reporting. *Eur J Public Health* 2013;23:86–92.
 31. Green CA, Yeates D, Goldacre A *et al.* Admission to hospital for bronchiolitis in England: trends over five decades, geographical variation and association with perinatal characteristics and subsequent asthma. *Arch Dis Child* 2016;101:140–46.
 32. Shah V, Hall N, Goldacre MJ. Retinal detachment in England: database studies of trends over time and geographical variation. *Br J Ophthalmol* 2015;99:639–43.
 33. Bragg F, Cromwell DA, Edozien LC *et al.* Variation in rates of caesarean section among English NHS trusts after accounting for maternal and clinical risk: cross sectional study. *BMJ* 2010;341:c5065.
 34. Mason A, Street A, Verzulli R. Private sector treatment centres are treating less complex patients than the NHS. *J R Soc Med* 2010;103:322–31.
 35. Judge A, Welton NJ, Sandhu J, Ben-Shlomo Y. Equity in access to total joint replacement of the hip and knee in England: cross sectional study. *BMJ* 2010;341:c4092.
 36. Chou MR, Malik AN, Suleman M, Gray M, Yeates D, Goldacre MJ. Time trends over five decades, and recent geographical variation, in rates of childhood squint surgery in England. *Br J Ophthalmol* 2013;97:746–51.
 37. Gurol-Urganci I, Cromwell DA, Edozien LC, Onwere C, Mahmood TA, van der Meulen JH. The timing of elective caesarean delivery between 2000 and 2009 in England. *BMC Pregnancy Childbirth* 2011;11:43.
 38. Sheldon TA, Cullum N, Dawson D *et al.* What's the evidence that NICE guidance has been implemented? Results from a national evaluation using time series analysis, audit of patients' notes, and interviews. *BMJ*.2004;329:999.
 39. Liu A, Guzman Castillo M, Capewell S, Lucy J, O'Flaherty M. Reduction in myocardial infarction admissions in Liverpool after the smoking ban: potential socioeconomic implications for policymaking. *BMJ Open* 2013;3:e003307.
 40. Smolina K, Wright FL, Rayner M, Goldacre MJ. Long-term survival and recurrence after acute myocardial infarction in England, 2004 to 2010. *Circ Cardiovasc Qual Outcomes* 2012;5:532–40.
 41. Sinha S, Karthikesalingam A, Poloniecki JD, Thompson MM, Holt PJ. Inter-relationship of procedural mortality rates in vascular surgery in England: retrospective analysis of hospital episode statistics from 2005 to 2010. *Circ Cardiovasc Qual Outcomes* 2014;7:131–41.
 42. Gurol-Urganci I, Cromwell DA, Edozien LC *et al.* Third- and fourth-degree perineal tears among primiparous women in England between 2000 and 2012: time trends and risk factors. *BJOG* 2013;120:1516–25.
 43. Jeevan R, Cromwell DA, Trivella M *et al.* Reoperation rates after breast conserving surgery for breast cancer among women in England: retrospective study of hospital episode statistics. *BMJ* 2012;345:e4505.
 44. Herbert A, Gilbert R, González-Izquierdo A, Pitman A, Li L. 10-y risks of death and emergency re-admission in adolescents hospitalised with violent, drug- or alcohol-related, or self-inflicted injury: a population-based cohort study. *PLoS Med* 2015;12:e1001931.
 45. Wijlaars LP, Hardelid P, Woodman J, Allister J, Cheung R, Gilbert R. Who comes back with what: a retrospective database study on reasons for emergency readmission to hospital in children and young people in England. *Arch Dis Child* 2016;101:714–18.
 46. Aljuburi G, Laverty AA, Green SA, Pheko KJ, Bell D, Majeed A. Socio-economic deprivation and risk of emergency readmission and inpatient mortality in people with sickle cell disease in England: observational study. *J Public Health (Oxf)* 2013;35:510–17.
 47. Burns EM, Bottle A, Almoudaris AM *et al.* Hierarchical multi-level analysis of increased caseload volume and postoperative outcome after elective colorectal surgery. *Br J Surg* 2013;100:1531–38.
 48. Freemantle N, Ray D, McNulty D *et al.* Increased mortality associated with weekend hospital admission: a case for expanded seven day services?. *BMJ* 2015;351:h4596.
 49. Bottle A, Aylin P. Comorbidity scores for administrative data benefited from adaptation to local coding and diagnostic practices. *J Clin Epidemiol* 2011;64:1426–33.
 50. Soong J, Poots AJ, Scott S, Donald K, Bell D. Developing and validating a risk prediction model for acute care based on frailty syndromes. *BMJ Open* 2015;5:e008457.
 51. Cromwell DA, Knight HE, Gurol-Urganci I. Parity derived for pregnant women using historical administrative hospital data: accuracy varied among patient groups. *J Clin Epidemiol* 2014;67:578–85.
 52. NHS Digital. *Summary Hospital-level Mortality Indicator*. 2016. <http://content.digital.nhs.uk/SHMI> (27 October 2016, date last accessed).
 53. Dr Foster. *The Hospital Guide 2013*. 2013 <http://www.drfooster.com/updates/recent-publications/the-hospital-guide-2013/> (16 June 2016, date last accessed).
 54. King's Fund. *Variations in Health Care: The Good, the Bad and the Inexplicable*. 2011. http://www.kingsfund.org.uk/sites/files/kf/field/field_publication_file/Variations-in-health-care-good-bad-inexplicable-report-The-Kings-Fund-April-2011.pdf (20 May 2016, date last accessed).
 55. Royal College of Obstetricians and Gynaecologists. *Patterns of Maternity Care in English NHS Hospitals 2011/12*. 2013. https://www.rcog.org.uk/globalassets/documents/guidelines/research-audit/patterns-of-maternity-care-in-english-nhs-hospitals-2011-12_0.pdf (18 March 2016, date last accessed).
 56. Farrar S, Yi D, Sutton M, Chalkley M, Sussex J, Scott A. Has payment by results affected the way that English hospitals provide care? Difference-in-differences analysis. *BMJ* 2009;339:b3047.
 57. Gonzalez-Izquierdo A, Cortina-Borja M, Woodman J *et al.* Maltreatment or violence-related injury in children and adolescents admitted to the NHS: comparison of trends in England and Scotland between 2005 and 2011. *BMJ Open* 2014;4:e004474.

58. World Health Organization. *WHOFIC Implementation Database*. <http://apps.who.int/classifications/implementation/> (20 May 2016, date last accessed).
59. Davies H, Gilbert R, Johnson K *et al*. Neonatal drug withdrawal syndrome: cross-country comparison using hospital administrative data in England, the USA, Western Australia and Ontario, Canada. *Arch Dis Child Fetal Neonatal Ed* 2016;**101**:F26–30.
60. McGuire A, Martin M, Lenz C, Sollano JA. Treatment cost of non-small cell lung cancer in three European countries: comparisons across France, Germany, and England using administrative databases. *J Med Econ* 2015;**18**:525–32.
61. Department of Health. *NHS Reference Costs*. 2015. <https://www.gov.uk/government/collections/nhs-reference-costs> (20 May 2016, date last accessed).
62. Meacock R, Doran T, Sutton M. What are the costs and benefits of providing comprehensive seven-day services for emergency hospital admissions?. *Health Econ* 2015;**24**:907–12.
63. Ashraf SQ, Burns EM, Jani A *et al*. The economic impact of anastomotic leakage after anterior resections in English NHS hospitals: are we adequately remunerating them?. *Colorectal Dis* 2013;**15**:e190–98.
64. Keeping ST, Tempest MJ, Stephens SJ, Carroll SM, Sangar VK. Penile cancer treatment costs in England. *BMC Public Health* 2015;**15**:1305.
65. Department of Health. *Payment By Results Data Assurance Framework: Key Findings From The 2012 /13 Programme*. 2013. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/262027/pbr_data_assurance_framework_key_find_2012-13.pdf (15 July 2016, date last accessed).
66. Burns EM, Rigby E, Mamidanna *et al*. Systematic review of discharge coding accuracy. *J Public Health (Oxf)* 2012;**34**:138–48.
67. Gill PJ, Goldacre MJ, Mant D *et al*. Increase in emergency admissions to hospital for children aged under 15 in England, 1999–2010: national database analysis. *Arch Dis Child* 2013;**98**:328–34.
68. Saxena S, Bottle A, Gilbert R, Sharland M. Increasing short-stay unplanned hospital admissions among children in England; time trends analysis '97-'06. *PloS One* 2009;**4**:e7484.
69. Meacock R, Anselmi L, Kristensen S, Doran T, Sutton M. Higher mortality rates amongst emergency patients admitted to hospital at weekends reflect a lower probability of admission. *J Health Serv Res Policy* 2016, May 6; pii: 1355819616649630. [Epub ahead of print.]
70. Mathur R, Bhaskaran K, Chaturvedi N *et al*. Completeness and usability of ethnicity data in UK-based primary care and hospital databases. *J Public Health (Oxf)* 2014;**36**:684–92.
71. Health and Social Care Information Centre. *HES 2012–13 Inpatient Data Quality Note*. 2013. <http://content.digital.nhs.uk/catalogue/PUB12744/nhs-mate-eng-2012-13-ip-qual-rep.pdf> (27 October 2016, date last accessed).
72. Hagger-Johnson G, Harron K, Fleming T *et al*. Data linkage errors in hospital administrative data when applying a pseudonymisation algorithm to paediatric intensive care records. *BMJ Open* 2015;**5**:e008118.
73. Moser K, Hilder L. Assessing quality of NHS Numbers for Babies data and providing gestational age statistics. *Health Stat Q* 2008;**Spring**:15–23.
74. Health and Social Care Information Centre. *Hospital Episode Statistics: Adult Critical Care in England - April 2014 to March 2015*. 2016. <http://content.digital.nhs.uk/catalogue/PUB19938/adul-crit-care-data-eng-apr-14-mar-15-rep.pdf> (27 October 2016, date last accessed).
75. NHS Digital. *Information on Type 2 Opt-outs*. 2016. <http://content.digital.nhs.uk/article/7092/Information-on-type-2-opt-outs> (27 October 2016, date last accessed).
76. Medical Research Council. *Obtaining Data from the Health and Social Care Information Centre (HSCIC) for Health Research – A Guide for Researchers*. 2016. <https://www.mrc.ac.uk/documents/pdf/obtaining-hscic-data-guidance-v020616/> (17 June 2016, date last accessed).
77. NHS Digital. *Data Access Advisory Group (DAAG)*. 2016. <http://content.digital.nhs.uk/daag> (27 October 2016, date last accessed).
78. Office for National Statistics. *Population Estimates*. 2016. <https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates> (15 July 2016, date last accessed).