

HES Accident & Emergency (A&E) to Admitted Patient Care (APC) Linkage Methodology

Purpose of this document

This document contains details of the methodology used to link HES A&E to HES APC.

Version history			
Version	Date issued	Brief summary of change	Change owner
1.0	19/01/11	First release	The NHS IC

For more information on the status of this document, please contact The NHS Information Centre for health and social care:

Tel: 0845 300 6016

Email: enquiries@ic.nhs.uk

Website: www.ic.nhs.uk

Contents

Summary	4
Introduction	5
Number of linked records expected	6
Accident and Emergency (A&E) and Admitted Patient Care (APC) linkage algorithm	8
Generic Condition 1: HESID in A&E is the same as HESID in APC data	8
Generic Condition 2: A&E discharge date is the same as APC admission date	9
Generic Condition 3: Episode Order equals 1	10
Records that don't have a one-to-one linkage	12
Relationship type - One A&E attendance to many APC episodes	12
Relationship type - Many A&E attendances to one APC episode	13
Relationship type - Many A&E attendances to many APC episodes.....	15
Linkage Quality	21
Quality rank	21
Quality rule 1	21
Quality rule 2	22
Quality rule 3	22
Quality rule 4	22
Number of records linked	23
Accessing HES	24
Feedback	25
Appendix 1 – Glossary of terms	26

Summary

Accident and Emergency (A&E) Hospital Episode Statistics (HES) data has been linked to HES Admitted Patient Care (APC) data to allow wider analysis of patients admitted to hospital following an A&E attendance. This paper considers the methodology developed for linking A&E attendances to APC episodes, using HES 2008-09 data to illustrate the linkage algorithm. The linkage on 2008-09 data resulted in 2.8 million of the 13.8 million A&E records (2008-09) being linked to an APC admission episode.

The linkage is presented as experimental statistics and initial findings are being made available to promote and highlight uses of this data, and to allow users to provide feedback on the algorithm. A&E fields are available as a tail to the HES APC universe for users of the HES interrogation system.

The linkage allows the patient's pathway to be followed and provides additional information beyond what is available from the standalone datasets. For example, for linked records we can identify what time admitted patients arrived at A&E and also look at the more detailed clinical coding available from the APC data.

Introduction

Accident and Emergency (A&E) and Admitted Patient Care (APC) are two datasets from the Hospital Episode Statistics (HES) suite of datasets. They allow interesting analysis as standalone datasets, however, extra analytical value can be attained from the datasets by linking them. The linkage allows the patient's pathway to be followed and provides additional information beyond what is obtainable from the standalone datasets.

The linkage algorithm has evolved from research and development undertaken on the two datasets. The algorithm is primarily based on fields that:

- identify the patient,
- report the outcome of the A&E attendance, and
- report the method of admission.

Further, the basic principles behind the linkage method are as follows:

- Records are linked where the unique patient ID (pseudo HESID¹) is the same in both A&E and APC data, and the A&E discharge date² and APC admission dates are the same. For multi-episode spells the admission date used for the linkage comes from the first episode in a spell.
- When a patient is linked more than once, in cases where the patient has had multiple A&E attendances and/or APC admissions, further algorithm rules are applied to find the most appropriate one-to-one A&E to APC link.

2008-09 A&E and APC HES datasets are used in this document to illustrate the linkage process. The sections in this document demonstrate how a linkage algorithm was achieved using the two different datasets.

The linkage is presented as experimental statistics and initial findings are being made available to promote and highlight uses of this, data and to allow users to provide feedback on the algorithm.

¹ Pseudo HESID is a unique patient identifier used in the suite of HES products. Further information is available on [HESonline](http://www.hesonline.nhs.uk/Ease/servlet/ContentServer?siteID=1937&categoryID=330) [http://www.hesonline.nhs.uk/Ease/servlet/ContentServer?siteID=1937&categoryID=330].

² A&E Discharge Date is calculated by the linkage algorithm. It uses the Arrival Date, Arrival Time and Duration to Departure to calculate an accurate discharge date.

Number of linked records expected

The A&E attendance disposal field reports what happens to the patient following their attendance. Table 3.1, below, shows a breakdown of 2008-09 A&E HES attendances by Attendance Disposal.

Disposal Code	Disposal Description	A&E Attendances
01	Admitted to a hospital bed /became a lodged patient of the same healthcare provider	2,979,331
02	Discharged – follow-up treatment to be provided by general practitioner	2,447,969
03	Discharged - did not require any follow-up treatment	5,473,967
04	Referred to A&E clinic	459,584
05	Referred to fracture clinic	563,927
06	Referred to other outpatient clinic	550,190
07	Transferred to other healthcare provider	315,084
10	Died in department	23,102
11	Referred to other healthcare professional	330,468
12	Left department before being treated	449,924
13	Left department having refused treatment	51,705
14	Other	122,455
99		26,366
Total		13,794,072

The 'A&E Attendance Disposal' field can be used to provide an estimate for volumes of A&E attendances resulting in an admission to hospital. A&E attendances that result in an admission into hospital should be coded with '01'. A&E attendances can also result in patients being transferred and admitted to another provider, these are coded with A&E disposal code '07'.

In 2008-09 A&E HES data, there were a total of 3,294,415 A&E attendances reported as being either admitted into a hospital bed of the same provider, or transferred to another provider. As the A&E to APC HES linkage algorithm looks to find the best APC episode match for all records in the A&E HES data that result in an admission to hospital, we would expect the number of linked records to match most closely to this number. However, data quality varies across providers, resulting in A&E attendances with codes other than '01' and '07' matching to records in the APC data, and A&E attendances with either '01' or '07' not finding a link to an APC record.

The APC admission method field reports where the patient was admitted from. Table 3.2 shows a breakdown of APC episodes³ by admission method.

³ APC episodes where the 'Episode Order' is reported as the first episode (equal to '01'). This is to restrict to only the first episode in multi-episode spells.

Table 3.2: 2008-09 APC episodes broken down by admission method

Admission Method Code	Admission Method Description	APC Episodes ⁴
11	Waiting list	2,815,583
12	Booked	2,406,401
13	Planned	3,002,171
21	Accident and emergency or dental casualty department of the healthcare provider	3,308,384
22	General practitioner: after a request for immediate admission has been made direct to a hospital provider, ie not through a bed bureau, by a general practitioner or deputy	978,763
23	Bed bureau	93,457
24	Consultant clinic, of this or another healthcare provider	153,959
28	Other means Examples: - admitted from the accident and emergency department of another provider where they had not been admitted - transfer of an admitted patient from another hospital provider in an emergency	482,826
31	Admitted ante-partum	1,217,440
32	Admitted post-partum	35,118
81	The birth of a baby in this healthcare provider	176,347
82	Baby born outside the healthcare provider except when born at home as intended	644,262
83	Transfer of any admitted patient from other Hospital Provider other than in an emergency	5,248
99		21,734
Total		15,341,693

The admission method field in the APC dataset can also be used to provide an estimate for volumes of admissions to hospital via A&E. For 2008-09 APC HES data, there were a total of 3,791,210⁵ episodes reported as being admitted via an A&E or dental casualty department, or admitted from another provider's A&E department.

It would be expected that the number of A&E attendances that are reported to be admitted is similar to the reported admissions via A&E in the APC dataset. There is a difference of 496,795 records between the A&E hospital admissions figure of 3,294,415 (as reported in Table 3.1) and the APC admissions via A&E figure of 3,791,210. This highlights some potential coverage issues with the A&E data as well as possible quality issues with the values of one or both of these fields.

The A&E to APC HES linkage algorithm is developed to find the best possible APC episode match for all records in the A&E HES data that result in an admission.

⁴ APC episodes where the 'Episode Order' is reported as the first episode (equal to '01'). This is to restrict to only the first episode in multi-episode spells

⁵ 3,791,210 is the number of APC episodes where the 'Episode Order' is reported as the first episode (equal to '01'). This is to restrict to only the first episode in multi-episode spells

Accident and Emergency (A&E) and Admitted Patient Care (APC) linkage algorithm

This section outlines the creation of a linkage between the A&E and APC HES datasets using 2008-09 data. First, a set of generic conditions are applied to all records followed by conditions specific to the relationship types generated by applying the generic conditions.

The following relationships are found when linking the two datasets together using HESID only:

1. One A&E attendance to one APC episode.
2. One A&E attendance record to many APC episodes – this is when a patient has more than one admission. The linkage algorithm tries to find the first episode that occurs following the A&E attendance from the many episodes.
3. Many A&E attendances to one APC episode – this is when a patient has attended A&E more than once but has only one APC admission. The data also shows that in some cases more than one A&E attendance is showing an admission, but only one admission can be found within the APC data. In this case the algorithm tries to find the most appropriate A&E attendance to join to the one admission record found.
4. Many A&E attendances to many APC episodes – this is when a patient has multiple A&E attendances as well as multiple admissions. Here, the algorithm finds A&E attendances that are recorded as resulting in admission and then tries to find the most appropriate admission to join to from the multiple admissions available.

A&EKey and Epikey are unique identifiers for each record. These fields are used to identify a one-to-one link. Ideally, the linkage algorithm should leave one A&E record matching to one APC record, ie one A&EKey to one Epikey.

Generic condition 1: HESID in A&E is the same as HESID in APC data

In 2008-09 HES data, there are:

- 13,794,072 A&E HES attendances with 9,349,040 distinct HESIDs, and
- 17,434,446 APC HES episodes with 8,473,905 distinct HESIDs.

HESID, known as 'pseudo HESID'⁶ is a unique patient identifier derived within HES.

There are 19,962,102 records in total when A&E and APC data is linked only on HESID. Records are multiplied where the HESID exists more than once, eg where a patient attends A&E twice and has three APC episodes, and only one of these episodes is an admission from A&E. Here, six records are generated when the link is only based on the HESID, as illustrated in the tables 4.1, 4.1.2 and 4.1.3, on the following pages. Further conditions need to be applied to match the A&E attendance to the correct APC episode.

A&E HESID	A&E Discharge Date	A&E Attendance Disposal
0001	01/11/2008	01
0001	05/11/2008	02

⁶ Pseudo HESID is a unique patient identifier used to allow the same patient to be safely and securely traced in all HES datasets across multiple years. Further information is available on [HESonline](http://www.hesonline.nhs.uk/Ease/servlet/ContentServer?siteID=1937&categoryID=330) [http://www.hesonline.nhs.uk/Ease/servlet/ContentServer?siteID=1937&categoryID=330].

APC HESID	Admission Date	Admission method
0001	01/11/2008	21
0001	04/03/2009	11
0001	25/03/2009	11

A&E HESID	A&E Discharge Date	A&E Attendance Disposal	APC HESID	Admission Date	Admission method
0001	01/11/2008	01	0001	01/11/2008	21
0001	01/11/2008	01	0001	04/03/2009	11
0001	01/11/2008	01	0001	25/03/2009	11
0001	05/11/2008	02	0001	01/11/2008	21
0001	05/11/2008	02	0001	04/03/2009	11
0001	05/11/2008	02	0001	25/03/2009	11

Table 4.1.3 illustrates how you end up with additional records when linking A&E and APC records only on HESID. The linkage algorithm applies further rules to join the A&E attendance to a single admission record, ie a one-to-one link.

Generic condition 2: A&E discharge date is the same as APC admission date

The second condition flags all records where the A&E attendance discharge date is the same as the APC admission date.

In 2008-9, there are 4,076,196 linked records when the following two conditions are applied:

1. HESID in A&E is the same as HESID in APC (AE.HESID=IP.HESID)
2. The A&E discharge date and APC admission dates are the same (AE.DEPDATE=IP.ADMIDATE).

The calculated A&E discharge date is derived by adding the 'Duration to Departure' to the 'A&E Arrival Date and Time'. Where there is no 'Duration to Departure Time' available, then the 'Discharge Date' is assumed to be the same as the 'Arrival Date'. 2.5% (348,717) of all A&E attendances (13,794,072) have no 'Duration to Departure Time' in the 2008-09 dataset.

Applying the two conditions discounts any APC episodes that did not occur on the same day as the A&E attendance, treating any admissions with a different 'Admission Date' to the calculated A&E attendance 'Discharge Date' as a non-A&E related admission. Similarly, any A&E attendances with a different calculated 'Discharge Date' to the 'Admission Date' are also excluded at this stage.

These conditions could therefore exclude records where the provider has submitted an incorrect 'APC Admission Date', 'A&E Arrival Date and Time' or 'A&E Duration Time'.

Application of these two conditions results in 2,894,353 unique A&E keys and 3,977,438 unique episode keys (Epikey), suggesting the final number of linked records will be 2,894,353 or less.

Generic condition 3: Episode order equals 1

A further condition is applied to restrict only those APC episodes where the 'Episode Order' is equal to 1. This condition has been applied to restrict the link to only the first episode when patients have been admitted to hospital.

Applying the following conditions reduces the number of linked records to 2,932,966:

1. HESID in A&E is the same as HESID in APC (AE.HESID=IP.HESID)
2. The A&E discharge date and APC admission dates are the same (AE.DEPPDATE=IP.ADMIDATE)
3. The APC Episode Order is equal to 1.

Table 4.3.1 – Number of records linked when the three generic conditions are applied

Description	Number of records
A&E HESID = APC HESID	19,962,102
A&E HESID = APC HESID, and A&E Discharge Date - APC Admission Date	4,076,196
A&E HESID = APC HESID, and A&E Discharge Date - APC Admission Date, and APC Episode Number = 1	2,932,966
Split by:	-
Number of one-to-one links	2,716,931
Other	216,035

Table 4.3.1 shows that the vast majority (93% [2,716,931]) of all (2,932,966) linked records based on the three conditions listed above are a one-to-one match, leaving 7% (216,035) of records where further conditions need to be applied to find a one-to-one linkage.

Table 4.3.2, on the following page, provides a breakdown of records that don't have a one-to-one link. Further conditions need to be applied to these records to establish a one-to-one link. Table 4.3.2 also highlights with an asterisk (*) symbol, the number of one-to-one records expected after applying further conditions, i.e. there are 57,619 records linked as one A&E to many admissions, and of these 28,546 are unique A&E attendances, which should be the total number of linked records after applying the additional filters.

For the many-to-many records, the final number of records will result in 10,587 or less, as not all unique A&E records will have an episode linked to it.

Table 4.3.2 - Breakdown of each relationship type after applying three generic conditions		
Relationship type	Number of records	% of linkage
One A&E to one APC	2,716,931	92.63%
A&EKey	2,716,931*	92.63%
Epikey	2,716,931	92.63%
One A&E to many APC	57,619	1.96%
A&EKey	28,546*	0.97%
Epikey	57,619	1.96%
Many A&E to one APC	136,685	4.66%
A&EKey	136,685	4.66%
Epikey	67,299*	2.29%
Many A&E to many APC	21,731	0.74%
A&EKey	10,756	0.37%
Epikey	10,587*	0.36%
Total	2,932,966	100%
Possible total one-to-one	2,823,363*	96.26%

Records that don't have a one-to-one linkage

This section lists the conditions applied to records that don't have a straight one-to-one link after the application of the three generic conditions.

All non-one-to-one records are uniquely linked using a scoring methodology, where the algorithm scores every record for each condition it meets. Records with the highest score, suggesting the strongest link are then flagged as the one-to-one link.

Relationship type - One A&E attendance to many APC episodes

The following conditions are applied to records which have matched one A&E attendance to many APC admissions after the three generic conditions are applied.

A score is given to every record that meets each of these conditions, with the highest scoring record flagged as the final one-to-one linked record.

1. ((APC Admission Method is equal to '21'⁷)
and
(A&E Attendance Disposal is equal to '01'⁸)
and
(APC Provider code is the same as A&E Provider code))
Score = 16 (applied if this criteria is met)
2. ((APC Admission Method is equal to '28'⁹)
and
(A&E Attendance Disposal is equal to '07'¹⁰)
and
(APC Provider code is different to the A&E Provider code))
Score = 16 (applied if this criteria is met)
3. ((APC Admission Method is equal to '21')
and
(A&E Attendance Disposal is equal to '01'))
or
((APC Admission Method is equal to '28')
and
(A&E Attendance Disposal is equal to '07'))
Score = 8 (applied if this criteria is met)
4. (APC record with the earliest Discharge date)
Score = 4 (applied if this criteria is met)
5. (The most complete [populated] record)
Score = 2 (applied if this criteria is met)
6. (Record with the smallest Epikey [unique episode key])
Score = 1 (applied if this criteria is met)

Table 4.4.1, on the following page, demonstrates an example of one A&E attendance linking to three APC admissions after applying the three generic conditions. Here it can be seen that one record has been identified as the most appropriate link having the highest score of 31, which shows it meets more conditions than the other two records.

⁷ Admission Method of '21' is where the patient was admitted from the accident and emergency or dental casualty department of the healthcare provider.

⁸ Attendance Disposal of '01' is where the patient was admitted to hospital bed / became a lodged patient of the same healthcare provider.

⁹ Admission Method of '28' is where the admission was an emergency (other means), including patients who arrive via the A&E department of another healthcare provider.

¹⁰ Attendance Disposal of '07' is where the patient was transferred to other healthcare provider.

Table 4.4.1 – An example of a match between one A&E attendance and many APC admissions

	A&EKey	Epikey	A&E Discharge Date	APC Discharge Date	APC Admission Method	A&E Disposal Code	A&E Provider Code	APC Provider Code	APC Data Completion	Scoring	Conditions met	One-to-one link
1	10001	20001	01/01/2009	01/01/2009	21	01	P01	P01	10	31	1, 3, 4, 5 & 6	Y
2	10001	20002	01/01/2009	03/09/2009	13	03	P02	P02	5	-	-	-
3	10001	20003	01/01/2009	02/02/2009	28	07	P01	P01	5	8	3	-

Table 4.4.2 - Breakdown of one-to-many records in the 2008-09 dataset that match each condition

Condition (one A&E to many APC)	Number of match records
APC Admission Method = 21 and A&E Attendance Disposal is equal to '01' and APC Provider Code = A&E Provider code	31,639
APC Admission Method = 28 and A&E Attendance Disposal is equal to '07' and APC Provider Code <> A&E Provider code	491
(APC Admission Method is equal to '21' and A&E Attendance Disposal is equal to '01') or (APC Admission Method is equal to '28' and A&E Attendance Disposal is equal to '07')	34,589
APC record with the earliest discharge date	31,382
Most complete/populated record	33,770
Record with smallest Epikey (unique episode key)	28,546
Number of one-to-one records	28,546

A total of 28,546 records were linked as one-to-one matches after running through the criteria. This total is the same as the number of unique A&E keys that matched one A&E and many APC admissions.

Relationship type - Many A&E attendances to one APC episode

The following conditions are applied where many A&E attendances have matched to one APC admission. Again, these conditions are applied following the three generic conditions.

Each of the following six conditions are applied to every record that has linked many A&E attendances to one APC episode. A score is applied to records that meet each condition and the record with the highest score is then flagged as the valid one-to-one link.

1. ((APC Admission Method is equal to '21') and (A&E Attendance Disposal is equal to '01') and (A&E Provider code is the same as APC Provider code))
Score = 16 (applied if this criteria is met)

2. ((APC Admission Method is equal to '28')
and
(A&E Attendance Disposal is equal to '07')
and
(A&E Provider code is different to the APC Provider code))
Score = 16 (applied if this criteria is met)
3. ((APC Admission Method is equal to '21')
and
(A&E Attendance Disposal is equal to '01'))
or
((APC Admission Method is equal to '28')
and
(A&E Attendance Disposal is equal to '07'))
Score = 8 (applied if this criteria is met)
4. (Latest A&E Arrival Time and Date)
Score = 4 (applied if this criteria is met)
5. (The most complete (populated) record)
Score = 2 (applied if this criteria is met)
6. (Record with the highest A&EKey [unique A&E attendance key])
Score = 1 (applied if this criteria is met)

Table 4.4.3 shows an example of where many A&E attendance records are linked to one APC admission. Record one is flagged to be the most appropriate one-to-one link as it has achieved the highest score.

Table 4.4.3: Example where many A&E attendance records are linked to one APC admission, record one is flagged to be the valid one-to-one record.

	A&EKey	Epikey	A&E arrival date & time	APC Admission Method	A&E Disposal Code	A&E Provider Code	APC Provider Code	APC Data Completion	Scoring	Conditions met	Record flagged as one-to-one link
1	10003	20001	01/01/2009 10:05	21	01	P01	P01	10	27	1, 3, 5 & 6	Y
2	10002	20001	01/01/2009	28	07	P01	P02	5	24	2 & 3	-
3	10001	20001	01/01/2009	28	03	P02	P01	5	4	4	-

Table 4.4.4 – Breakdown of many-to-one records in the 2008-09 dataset that match each condition

Condition (many A&E to ne APC)	Number of match records
APC Admission Method = 21 and A&E Attendance Disposal is equal to '01' and A&E Provider Code = APC Provider code	62,791
APC Admission Method = 28 and A&E Attendance Disposal is equal to '07' and A&E Provider Code <> APC Provider code	1,060
(APC Admission Method is equal to '21' and A&E Attendance Disposal is equal to '01') or (APC Admission Method is equal to '28' and A&E Attendance Disposal is equal to '07')	64,705
Latest A&E Arrival Date/Time	68,961
Most complete/populated record	82,507
Record with highest A&E Key (unique A&EKey)	67,299
Number of one-to-one records	67,299

The application of these conditions linked a total of 67,299 records, which is the same number of unique APC keys that matched many A&E attendance records to one APC admission.

Relationship type - Many A&E attendances to many APC episodes

The following conditions are applied to records that have many A&E attendances matched to many APC episodes. These conditions follow the application of the three generic conditions listed earlier in the document.

1. ((APC Admission Method is equal to '21')
and
(A&E Attendance Disposal is equal to '01')
and
(APC Provider code is the same as A&E Provider code))
Score = 16 (applied if this criteria is met)
2. ((APC Admission Method is equal to '28')
and
(A&E Attendance Disposal is equal to '07')
and
(APC Provider code is different to the A&E Provider code))
Score = 16 (applied if this criteria is met)
3. ((APC Admission Method is equal to '21')
and
(A&E Attendance Disposal is equal to '01'))
Score = 8 (applied if this criteria is met)
4. ((APC Admission Method is equal to '28')
and
(A&E Attendance Disposal is equal to '07'))
Score = 8 (applied if this criteria is met)
5. (Earliest A&E Arrival Time and Date)
Score= 4 (applied if this criteria is met)
6. (Earliest APC Discharge Date)
Score= 4 (applied if this criteria is met)

7. (The most complete [populated] record)
Score = 2 (applied if this criteria is met)
8. ((Record with the smallest A&EKey [unique A&E attendance key])
and
(Record with the smallest Epikey [unique episode key]))
Score = 1 (applied if this criteria is met)

Similar to the previous two relationship types, a link is only formed on records with the highest score. However, with the many-to-many links, it is quite possible to get a tie (two records scoring the same). Here, a link is formed on the record in the tie with the lowest unique record identifier. This is an arbitrary decision but ensures that a link is formed. Once the first link is formed, the algorithm then cycles back and tries to form the second link but this time excludes the unique A&EKey and Epikey already used. This cycle continues until all possible links are made.

The following table (Table 4.4.5) shows a worked example to illustrate this process of forming links between many-to-many records. The example uses a patient that has had many A&E attendances linked to many APC episodes based on the three generic conditions.

Table 4.4.5 – Example showing a patient that has many A&E attendances linked to many APC episodes

ID	A&EKey	Epikey	Score
101648972194609450	101648972	194609450	33
101648972194612957	101648972	194612957	32
101648972194611061	101648972	194611061	30
101649111194609450	101649111	194609450	29
101649553194609450	101649553	194609450	29
101649111194612957	101649111	194612957	28
101649553194612957	101649553	194612957	28
101649111194611061	101649111	194611061	26
101649553194611061	101649553	194611061	26
101648972199353967	101648972	199353967	8
101649111199353967	101649111	199353967	4
101649553199353967	101649553	199353967	4

Highlighted in grey, Table 4.4.5 contains three unique A&EKeys and four unique Epikeys, suggesting three records can be formed as valid one-to-one links. As this example contains only one record with the highest score (33) the linkage algorithm will flag this record to be the first link (see Table 4.4.6, on the following page). Both the A&EKey and Epikey from this link are excluded from forming the next two links.

Table 4.4.6 – Example showing a patient that has many A&E attendances linked to many APC episodes, with the first valid link flagged

ID	A&EKey	Epikey	Score
101648972194609450	101648972	194609450	33
101648972194612957	101648972	194612957	32
101648972194611061	101648972	194611061	30
101649111194609450	101649111	194609450	29
101649553194609450	101649553	194609450	29
101649111194612957	101649111	194612957	28
101649553194612957	101649553	194612957	28
101649111194611061	101649111	194611061	26
101649553194611061	101649553	194611061	26
101648972199353967	101648972	199353967	8
101649111199353967	101649111	199353967	4
101649553199353967	101649553	199353967	4

Once the first valid link is made, the algorithm then repeats the process looking for the next record with the highest score which does not have a previously used Epikey or A&EKey, and repeats the identifying and linking process.

In this example, there is a tie for the next link with two records having a score of 28. In this situation the tie is broken by linking on the record with the lowest ID, as demonstrated in Table 4.4.7.

Table 4.4.7 – Example of where a patient has many A&E attendances linked to many APC admissions, with second valid link flagged*

ID	A&EKey	Epikey	Score
101648972194609450	101648972	194609450	33
101648972194612957	101648972	194612957	32
101648972194611061	101648972	194611061	30
101649111194609450	101649111	194609450	29
101649553194609450	101649553	194609450	29
101649111194612957	101649111	194612957	28
101649553194612957	101649553	194612957	28
101649111194611061	101649111	194611061	26
101649553194611061	101649553	194611061	26

Table 4.4.7 continued

ID	A&EKey	Epikey	Score
101648972199353967	401648972	199353967	8
101649111199353967	101649111	199353967	4
101649553199353967	101649553	199353967	4

*Please note: records that can no longer be linked due to the A&EKey and/or Epikey already been used have been crossed out.

This process of linking records by excluding previously used A&EKeys and Epikeys continues until no further links can be made. In this example, the third and final A&EKey is yet to be linked and is available against two records. The linkage algorithm will again form a link using the record with the highest score, as shown in Table 4.4.8.

Table 4.4.8 – Example showing a patient that has multiple A&E attendances linked to many APC admissions, with the third valid link flagged*

ID	A&EKey	Epikey	Score
101648972194609450	401648972	194609450	33
101648972194612957	401648972	194612957	32
101648972194611061	401648972	194611061	30
101649111194609450	401649111	194609450	29
101649553194609450	401649553	194609450	29
101649111194612957	401649111	194612957	28
101649553194612957	401649553	194612957	28
101649111194611061	401649111	194611061	26
101649553194611061	101649553	194611061	26
101648972199353967	401648972	199353967	8
101649111199353967	401649111	199353967	4
101649553199353967	101649553	199353967	4

*Please note: records that can no longer be linked due to the A&EKey and/or Epikey having already been used have been crossed out. Records with a single strike are those where the A&EKey or the Epikey was used in creating the first linked record. Records with two strikes are those where the A&EKey or the Epikey was used to create the second linked record.

Table 4.4.9 - Breakdown of many-to-many linked records in the 2008-09 that match each condition

Condition (many A&E to many APC)	Number of match records
APC Admission Method = 21 and A&E Attendance Disposal = 01 and APC Provider Code = A&E Provider Code	13,656
APC Admission Method = 28 and A&E Attendance Disposal = 07 and APC Provider Code <> A&E Provider Code	88
APC Admission Method = 21 and A&E Attendance Disposal = 01	16,030
APC Admission Method = 28 and A&E Attendance Disposal = 07	105
Earliest A&E Arrival Date/Time	10,745
Earliest APC Discharge Date	12,642
Most complete/populated record	13,225
Record with smallest A&E Key (unique A&E key) and Record with smallest Epikey (unique episode key)	10,642
Number of one-to-one records	10,500

Unlike the previous two relationship types (one-to-many and many-to-one), the many-to-many link does not join all unique A&E or APC records. There are a range of potential many-to-many links (two to two, two to three, etc) that can exist after the three generic conditions are applied. Table 4.4.10, below, summarises the combinations that are found in the dataset.

Table 4.4.10 - The range of many-to-many links in the 2008-09 data

Count of Epikey	Count of A&EKey				Total
	2	3	4	5	
2	4,916	210	19	2	5,243
3	78	11	2		
4	4	1			
Total	4,998	222	21	2	

As we can see from the worked example (tables 4.4.5 to 4.4.8), the number of potential links is based on the lowest of the number of A&EKey and Epikey, eg when two A&EKeys are linked to three Epikeys, a maximum of two one-to-one links can be achieved. To this end, in order to calculate the number of actual links formed, each count must be multiplied by the smaller of the two reference numbers (A&EKey and Epikey) - please see Table 4.4.11, on the following page. The grand total (10,500) is the highest number of links that can theoretically be formed and is the actual result of the linkage algorithm.

Table 4.4.11 - Many-to-many links in the 2008-09 data expected

Count of Epikey	Count of A&EKey				
	2	3	4	5	
2	9,832	420	38	4	
3	156	33	6		
4	8	3			
Total	9,996	456	44	4	10,500

Linkage quality

This section considers the quality of the 2008-09 linked records.

Once a link is made between all possible A&E attendances and APC admission records, each record is then scored against a number of data quality conditions. Records that score poor in quality, ie rating of either three or four, are excluded from the final output of linked data, available to users of the HES interrogation system in Business Objects.

These quality checks are put into place to primarily exclude linked records where the A&E attendance is reported to be either 'dead on arrival' or 'died in department'.

Quality rank

A score of one to four is applied to each linked record, with a score of one meaning a very good link and a score of four to be a poor link.

These four quality rules are:

Quality rule 1

The first rule is based on fields that would suggest a strong link.

The majority, 85.8% (2,420,967) of all the linked records met the following conditions and are given a quality score of 1.

(APC Admission Method is equal to '21'¹¹
and
A&E Attendance Disposal is equal to '01'¹²
and
APC Provider code is the same as A&E Provider code)
OR
(APC Admission Method is equal to '28'¹³
and
A&E Attendance Disposal is equal to '07'¹⁴
and
APC Provider code is different to the A&E Provider code)
AND
(A&E Patient Group is not equal to '70'¹⁵)
OR
(A&E Disposal Code is not equal to '10'¹⁶)

¹¹ Admission Method of '21' is where the patient was admitted from the accident and emergency or dental casualty department of the healthcare provider.

¹² Attendance Disposal of '01' is where the patient was admitted to hospital bed / became a lodged patient of the same healthcare provider.

¹³ Admission Method of '28' is where the admission was an emergency (other means), including patients who arrive via the A&E department of another healthcare provider.

¹⁴ Attendance Disposal of '07' is where the patient was transferred to other healthcare provider.

¹⁵ A&E Patient Group of 70 is where the patient was 'brought in dead'.

¹⁶ A&E attendance with a Disposal Code of '10' were reported as 'died in department'

Quality rule 2

Records that meet the following set of conditions are given a score of two, suggesting the link is of good quality.

These conditions are only applied to records which have not met condition one.

(A&E Attendance Disposal is equal to '01')
OR
(A&E Attendance Disposal is equal to '07')
OR
(APC Admission Method is equal to '21')
OR
(APC Admission Method is equal to '28')
AND
(A&E Attendance Disposal is not equal to '10')
OR
(A&E Patient Group is not equal to '70')

A total of 367,409 (13.0%) of all linked records met this rule and are given a quality score of two.

Quality rule 3

Records that meet the third set of conditions are given a score of three. These conditions are only applied to one-to-one records that have not met either condition one or two.

((Have not met rule 1)
OR
(Have not met rule 2))
AND
((A&E Attendance Disposal is not equal to '10')
OR
(A&E Patient Group is not equal to '70'))

There were 34,283 (1.2%) of all linked records in 2008-9 that are assigned a quality score of 3.

Quality rule 4

The fourth set of conditions are based around patients that are reported to be either dead on arrival or died in A&E.

((A&E Attendance Disposal is equal to '10')
OR
(A&E Patient Group is equal to '70'))

There are 617 linked records in 2008-09 where the reported 'A&E Attendance Disposal' is '10', 'died in department' or the 'A&E Patient Group' is '70', 'brought in dead'. It is unlikely that these patients would result in an admission and therefore these links are considered as poor and given a quality score of four.

Number of records linked

In 2008-09 A&E HES data, there were a total of 3,294,415 patients reported as being either admitted into a hospital bed of the same provider or transferred to another provider, which is 23.9% of all A&E attendances. The linkage algorithm has linked a total of 2,823,276 A&E records to an APC admission, representing 20.5% of all A&E attendances.

A total of 34,900 (1.2%) are excluded from the final output in Business Objects due to poor data quality – these are linked records scoring either three or four on the quality checks.

Accessing HES

How to access A&E HES data

Freely available HES data, including this report, is accessible via [HESonline](http://www.hesonline.nhs.uk) [<http://www.hesonline.nhs.uk>].

Ad hoc tabulations and extracts based on experimental A&E and APC linked HES data are available on request, subject to agreement of terms and conditions of use. Users requiring such access should refer to the information in the 'Request a tailor-made report' area of [HESonline](http://www.hesonline.nhs.uk) [<http://www.hesonline.nhs.uk/Ease/servlet/ContentServer?siteID=1937&categoryID=1342>].

NHS organisations can also request direct access to the full APC HES data including linked A&E records via the HES Interrogation System. Please note that there are restrictions on who can access HES data in this way. For further information on this service, please contact the NHS Information Centre (0845 300 6016 or enquiries@ic.nhs.uk).

Feedback

The NHS Information Centre welcomes all feedback relating to any aspect of this publication.

In particular we would welcome feedback on the method used for linking A&E attendance records to APC admissions.

How to provide feedback

Feedback can be provided by either telephone or email:

Telephone: 0845 300 6016

Email: enquiries@ic.nhs.uk

Responsible statistician:

Simon Rhea, HES/SUS Development Section Head

Contact via enquiries@ic.nhs.uk or 0845 300 6016

Appendix 1 – Glossary of Terms

A&E	Accident and Emergency
A&E HES	<p>'Accident and Emergency Hospital Episodes Statistics' is a name given to the data set that contains data on individual A&E attendances. A&E HES is one of a number of data sets available from the family of HES products.</p> <p>More information on HES products is available on the HESonline website [http://www.hesonline.nhs.uk].</p>
A&EKey	A&EKey is a number unique to each A&E record
Epikey	Epikey is a number unique to each APC admission
APC	Admitted Patient Care
HES	Hospital Episode Statistics is a brand that holds a collection of data sets produced from regular CDS submissions, these data sets include, admitted patient care, outpatients and now A&E.
NHS IC	The NHS Information Centre for health and social care
NHS	National Health Service