

***NHS WALES INFORMATICS  
SERVICE***

**DATA QUALITY STATUS REPORT**

**ADMITTED PATIENT CARE  
DATA SET**

Version: 1.0

Date: 2<sup>nd</sup> November 2020

## Data Set

### Title

Admitted Patient Care data set (APC ds)

### Sponsor

Welsh Government (WG)

### Implementation Date

1<sup>st</sup> April 1999

### Change History

See NHS Wales Data Dictionary<sup>1</sup>

### Data Set Purpose

The APC data set is the principal source of secondary use data for hospital admissions. Each record in the data set describes a Finished Consultant Episode (FCE). The data set has a wide range of uses including:

- Service improvement
- Hospital capacity planning
- Budget planning
- Financial costing
- Performance monitoring
- Public health surveillance

This data is submitted by each provider organisation to the NHS Wales Informatics Service's Information Services Division (ISD), where the data is loaded into the national database. At the end of each year, the data for that year is 'frozen' to ensure that National Statistics for that year remains unaffected by changes to the database as a result of data resubmissions.

---

<sup>1</sup> <http://www.datadictionary.wales.nhs.uk/WordDocuments/admittedpatientcaredatasetapcds.htm>

## Document Purpose

### Function

ISD provides a central data processing, analysis and publishing service for NHS Wales. A key element of this process is to ensure that the data being processed is of suitable quality to maintain the integrity of the database which, in turn, enables the reporting of meaningful health information.

This document describes a range of data quality issues affecting this data set.

### Audience

This document acts as a situation report for the Sponsor as well as an information resource for other stakeholders who base decisions on the accuracy of this data.

### Frequency

The document is issued annually to accompany all formal annual publications.

### Information Source

The aim is to describe the quality of the data held centrally in the NHS Wales Informatics Service national database. The Data Quality and Data Acquisitions teams within ISD are in regular contact with the health boards who supply this data, as well as the WPAS<sup>2</sup> (Welsh Patient Administration System) development team to ensure that the data being entered on hospital systems and extracted from them comply with the standards set out in the NHS Wales Data Dictionary and relevant Data Set Change Notices (DSCNs)<sup>3</sup>.

### Scope

The following set of data quality dimensions are covered in this report:

- **Data Validity.** The term ‘data validity’ refers to whether the submitted data has been provided in the agreed format and, where applicable, whether it is populated with a nationally-agreed value.
- **Data Consistency** refers to whether related data items within the same data set are consistent with one another. For example, a record that indicates a male patient has given birth should be considered inconsistent and would require investigation.
- **Data Timeliness** is simply a measure of whether the data file was submitted in accordance with national timescales.
- **Data Completeness.** In this case, this is a measure of the ratio of records submitted : records loaded.

These are fundamental to the quality of the data which is submitted and, in turn, processed through to the national database.

Aside from the data quality dimensions listed above, this document does not seek to review the accuracy of the data reported via the APC ds – i.e. whether reported activity is a true reflection of the activity being carried out within NHS Wales organisations.

Note also that nationally defined default or bucket codes are permitted and are therefore classed as valid values.

---

<sup>2</sup> WPAS was formerly called Myrddin

<sup>3</sup> All new DSCNs are published on the NHS Wales Informatics Service Data Standards website via: <http://www.nwisinformationstandards.wales.nhs.uk/change-notices>

Further information about these dimensions can be found on the NHS Wales Informatics Service Data Quality website<sup>4</sup>.

## Data Quality Standards

### Validation At Source Service (VASS) Checks

This data set is used for high profile National Statistics where a high level of quality assurance is required. VASS provides an online resource for submitting organisations to check the quality of their data before formally submitting it to ISD to be processed through to the national database. VASS is comprised of 3 main types of data quality checks as described below:

- **Data Load checks** are used to protect the integrity of the database by identifying invalid values within a record. If a data load error is triggered, the whole record is rejected by the system, preventing it from being processed through to the national database. The fact that load errors prevent records from being loaded means that these are often reviewed and resubmitted immediately. While this has been a successful method of maintaining the quality of this data set, it is reliant on the cooperation of the data provider in reviewing these errors promptly.
- **A Data Validity check** tests whether the recorded entry within the associated database field is a valid national value. These national values are defined in the NHS Wales Data Dictionary and lists of codes are available from the Welsh Reference Data Service<sup>5</sup>. Data Validity checks have been in operation since April 2008.
- Some data items are interdependent. For example, a patient's date of birth must not be after their activity date. Relationships between data items are checked using **Data Consistency checks**. These were introduced for APC in April 2009.

These checks are reviewed and updated as necessary.

### Regular Monitoring

Data Validity and Consistency performance is monitored on a monthly basis. The **Data Quality Standards** that each data provider must adhere to are defined by sets of indicators and nationally-agreed targets. These are based on the aforementioned VASS checks. Data Validity and Consistency reports are used to measure compliance with these standards

Further information on Data Quality Standards and how the quality of data is monitored can be found on the NHS Wales Informatics Service Data Quality internet site<sup>6</sup>. The reports themselves are published on the corresponding intranet site<sup>7</sup>.

## Data Set Quality Status

### Data Validity

Regular monitoring and provider cooperation means that data validity is generally high. A copy of the annual Data Validity report for 2019-20 is shown in [Appendix A](#). Issues causing percentages to fall considerably below the target (>4%) are explained and resolved by the health boards where resource and system constraints permit.

---

<sup>4</sup> <http://www.nwisinformationstandards.wales.nhs.uk/about-data-quality>

<sup>5</sup> <http://wrds.wales.nhs.uk> (accessible to NHS Wales users only)

<sup>6</sup> <http://www.nwisinformationstandards.wales.nhs.uk/data-quality-standards>

<sup>7</sup> <http://www.nwisinformationstandards.wales.nhs.uk/data-quality> (accessible to NHS Wales users only)

Cwm Taf Morgannwg did not meet the validity target for **Referrer Code**. 97% of the invalid values related to episodes that took place in the Princess of Wales Hospital, and the vast majority of those contained blank values. In April 2019, the geographical boundary between Abertawe Bro Morgannwg LHB and Cwm Taf LHB changed to create Swansea Bay LHB and Cwm Taf Morgannwg LHB. Princess of Wales Hospital moved from Abertawe Bro Morgannwg LHB to Cwm Taf Morgannwg LHB as part of that change, and this has resulted in some data quality issues.

The validity targets for **Main Specialty (Consultant)** and **Speciality of Treatment Code** are being met every year by all organisations apart from Powys. The records that are causing this low validity mainly relate to activity where a general practitioner is responsible for the patient during their inpatient stay and/or where the patient is treated under the specialty of general practice. The main specialty code relating to GPs changed when the list of values for this data item was revised in April 2015 as per DSCN 2014 / 07<sup>8</sup> and the specialty of treatment codes for general practice ceased to be valid in April 2016 as per DSCN 2014/08<sup>9</sup>. Although other organisations have also submitted such activity, this issue has been highlighted in Powys as it accounts for a high proportion of their overall activity. An upgrade to the WPAS system is needed before the validity of these data items will improve.

The validity target for **Consultant Code** was not being formally monitored in 2019/20 as the check was amended during 2018/19 and again during 2019/20.

## Data Consistency

A full breakdown of Data Consistency performance is shown in [Appendix B](#). Data Consistency compliance is good in general.

The consistency target for **Decision to Admit Date v Waiting List Date** is being met at each organisation apart from Velindre. As they do not record an inpatient waiting list, they do not enter these dates into CANISC therefore both dates are being calculated. Although Decision to Admit Date is calculated as a later date than Waiting List Date, the calculations cannot be amended as no further development is taking place on the CANISC system. Although this is not a new issue, it was not highlighted until this new check was introduced in early 2018.

The consistency target for **Referrer Code v Referring Organisation Code** is being met at each organisation apart from Cwm Taf Morgannwg. This is due to the number of referrals made by clinicians who are not listed within reference data as working for the Health Board and relates to the geographical boundary changes made to Health Boards in April 2019.

Most of the other indicators that are showing low values are ones where the denominator in the calculation is a subset of the total number of records. As the calculations are based on a relatively small number of records, the percentages can be somewhat deceptive. In Aneurin Bevan for example, the low consistency of **Primary Diagnosis Code v Admission Date & Birth Date [i.e. Age]** relates to only 39 records; in Cardiff & Vale that for **Discharge Method v Specialty (of Treatment)** relates to only 14 records and in Velindre that for **Admission Method v Source of Admission** relates to only 9 records.

---

<sup>8</sup> <http://www.nwisinformationstandards.wales.nhs.uk/opendoc/253025>

<sup>9</sup> <http://www.nwisinformationstandards.wales.nhs.uk/opendoc/258253>

## Data Timeliness

Issues with timeliness are rare due to an established process of file submission and sign-off via the NHS Wales Data Switching Service (NWDSS). The Data Acquisitions team issue reminders to data providers ahead of the monthly submission deadline and provide assistance with any VASS errors to reduce delays and minimise the probability of missed deadlines. There were only **1 late submission** in 2019-20 which equates to less than 1% of the total APC submissions received.

In addition to monthly deadlines, there is an annual deadline for resubmissions (17<sup>th</sup> June 2020) which allows providers to improve the quality of their data before it is frozen. All health boards met the deadline.

## Data Completeness

ISD data processing timescales must be adhered to in order to ensure compliance with reporting deadlines. If a monthly submission deadline is missed, the data cannot be processed until after the submission deadline for the following month. This can result in temporary data completeness issues. This does not affect the data used in annual reports as these are only run after files for the entire year have been received (and resubmitted where necessary) using the frozen data.

With the existence of Data Load checks there is an added risk of data completeness issues if invalid data is submitted. Although rejected records are generally reviewed and resubmitted before the data is loaded, if these are not corrected, the national database (and any reporting outputs) will contain incomplete data. This is not a significant issue at present as a relatively small number of records are rejected by the system each month and not loaded into the national database. Any instances where a high proportion of records are rejected are flagged up on Data Completeness reports. These are monitored by the Data Acquisitions team and issues are communicated to the submitting organisation immediately requesting that the data is resubmitted in time for the data to be processed.

Four organisations (BCU, Powys, Swansea Bay and Velindre) met both clinical coding completeness targets. C&V met the national standard of 95% coded within 3 months of episode end date but failed to achieve the 98% target for rolling 12 months' data, while the other three organisations (Aneurin Bevan, Cwm Taf Morgannwg and Hywel Dda) did not meet either target.

## Additional Issues

The data quality dimensions described above capture the major issues which can be easily monitored. However, there are some additional issues which, although not captured by regular monitoring, are highlighted to the Data Quality team on an ad hoc basis. The table in [Appendix C](#) describes the current position.

For further information regarding these issues, please contact the Data Quality team via [data.quality@wales.nhs.uk](mailto:data.quality@wales.nhs.uk).

## Quality Assurance

ISD follows a routine process to assure the quality of the data used in National Statistics. This process is described in the document *Data Quality Assurance – National Statistics (June 2014)*, which is available from the NHS Wales Informatics Service Data Quality Team on request).

## Impact on Reporting and Publishing

There are no major issues preventing this data from being used for reporting, providing that the recipient is made aware of the relevant issues described in this report.

## Overall Data Quality Status

Well established processes for submitting, checking and monitoring the quality of this data set means that the timeliness, completeness, validity and consistency of the data are generally good. These dimensions continue to be monitored on a regular basis to further improve quality.

Improvements to the other areas summarised in [Appendix C](#) are largely dependent on developments to operational systems or to the structure and scope of the data set itself.

WHC (2015) 027<sup>10</sup> was issued in June 2015 to introduce a national initiative to address the causes of poor information quality. The initiative, namely the Information Quality Improvement (IQI) initiative, has now been established and a work programme has also been developed along with detailed proposals for tackling the underlying causes of a number of issues affecting information quality.

More information about the initiative can be found by visiting [www.iqi.wales.nhs.uk](http://www.iqi.wales.nhs.uk).

---

<sup>10</sup> <http://howis.wales.nhs.uk/doclib/WHC2015027-e.pdf>

## Appendix A: APC Data Validity Report 2019-20

Data Item	DATA VALIDITY STANDARD	All Welsh Providers	Aneurin Bevan University LHB	Betsi Cadwaladr University LHB	Cardiff & Vale University LHB	Cwm Taf University LHB	Hywel Dda University LHB	Powys Teaching LHB	Swansea Bay University LHB	Velindre NHS Trust
APC submission received by the 17th	-	-	✓	✓	✓	✓	✓	✓	✓	✓
Number of Records Loaded	-	1165417	255486	226140	156541	156923	130681	4830	153199	81617
Administrative Category	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Admission Date	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Admission Method	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Consultant Code	98%	96.4%	90.9%	95.0%	96.7%	✓	✓	79.4%	✓	✓
Date of Birth	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Decision to Admit Date	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Discharge Date	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Discharge Destination	98%	✓	✓	✓	✓	✓	✓	97.3%	✓	✓
Discharge Method	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Duration of Elective Wait	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Episode Start Date	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Ethnic Group	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
HRG Code <sup>††</sup>	95%									
Intended Management	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Last Episode in Spell Indicator	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Legal Status	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Local Health Board of Residence	95%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Main Specialty (consultant)	98%	✓	✓	✓	✓	✓	✓	76.1%	✓	✓
NHS Number	95%	✓	✓	✓	✓	✓	✓	✓	✓	✓
NHS Number Status Indicator	95%	✓	✓	✓	✓	✓	✓	✓	✓	✓
NHS Number Valid & Traced	95%	✓	✓	91.4%	✓	✓	✓	✓	✓	✓



Data Item	DATA VALIDITY STANDARD	All Welsh Providers	Aneurin Bevan University LHB	Betsi Cadwaladr University LHB	Cardiff & Vale University LHB	Cwm Taf University LHB	Hywel Dda University LHB	Powys Teaching LHB	Swansea Bay University LHB	Velindre NHS Trust
Patient Classification	95%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Postcode <sup>††††</sup>	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Principal Diagnosis <sup>†</sup>	95%	94.1%	89.9%	✓	✓	85.9%	90.5%	✓	✓	✓
Principal Procedure Code <sup>†/††</sup>	95%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Principal Procedure Date	95%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Referrer Code	98%	93.1%	92.2%	94.4%	✓	77.9%	✓	97.0%	92.5%	✓
Registered GP Practice Code	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Sex	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Site Code (of Treatment)	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Source of Admission	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Specialty of Treatment Code	98%	✓	✓	✓	✓	✓	✓	75.1%	✓	✓
Waiting List Date	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓

† "Principal ICD Diagnosis", "Principal Procedure Code" and "HRG Code" will show as blank for the first 3 months of a new financial year.

†† "Principal Procedure Code" only reports the % validity of all records where a primary procedure code is present on an episode.

††† "HRG Code" (HRG v4) is not presently monitored for data validity as it is currently a derived field.

†††† "Postcode" may occasionally be incorrectly flagged as invalid due to issues with the postcode file received from ONS.

## Appendix B: APC Data Consistency Report 2019-20

Data Consistency Check	DATA CONSISTENCY STANDARD	All Welsh Providers	Aneurin Bevan University LHB	Betsi Cadwaladr University LHB	Cardiff & Vale University LHB	Cwm Taf University LHB	Hywel Dda University LHB	Powys Teaching LHB	Swansea Bay University LHB	Velindre NHS Trust
Admission Date vs. Date of Birth	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Admission Method vs. Intended Management	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Admission Method vs. Patient Classification	95%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Admission Method vs. Source of Admission*	98%	✓	✓	✓	94.9%	✓	✓	✓	✓	88.9%
Decision to Admit Date vs. Admission Date	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Decision to Admit Date vs. Waiting List Date	98%	83.3%	✓	✓	✓	✓	✓	✓	✓	0.1%
Discharge Method vs. Discharge Date & Date of Birth [i.e. Age]*	98%	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Discharge Method vs. Discharge Destination*	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Discharge Method vs. Specialty (of Treatment)*	98%	83.7%	83.3%	50.0%	78.6%	✓	n/a	✓	✓	n/a
Episode End Date vs. Admission Date	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Episode End Date vs. Discharge Date	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Episode End Date vs. Date of Birth	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Episode End Date vs. Episode Start Date	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Episode Start Date vs. Admission Date	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Episode Start Date vs. Discharge Date	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Episode Start Date vs. Date of Birth	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
HRG Code vs. Sex†*	95%	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Last Episode in Spell vs. Episode End Date & Discharge Date*	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Legal Status vs. Specialty (of Treatment)**	98%	✓	✓	✓	✓	✓	✓	✓	✓	n/a
Patient Classification vs. Discharge Date & Admission Date [i.e. Length of Stay]*	95%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Postcode vs. Local Health Board of Residence**	95%	✓	✓	✓	✓	✓	✓	✓	✓	90.0%
Primary Diagnosis Code vs. Admission Date & Birth Date [i.e. Age]†*	95%	✓	74.4%	✓	✓	✓	✓	n/a	n/a	n/a
Primary Diagnosis Code vs. Sex†*	95%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Primary Procedure Code vs. Sex†*	95%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Primary Procedure Date vs. Episode Start Date & Episode End Date	95%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Referrer Code vs. Referring Organisation Code	98%	✓	✓	✓	✓	86.4%	✓	✓	✓	✓

<b>Data Consistency Check</b>	<b>DATA CONSISTENCY STANDARD</b>	<b>All Welsh Providers</b>	<b>Aneurin Bevan University LHB</b>	<b>Betsi Cadwaladr University LHB</b>	<b>Cardiff &amp; Vale University LHB</b>	<b>Cwm Taf University LHB</b>	<b>Hywel Dda University LHB</b>	<b>Powys Teaching LHB</b>	<b>Swansea Bay University LHB</b>	<b>Velindre NHS Trust</b>
Waiting List Date vs. Admission Date	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Waiting List Date vs. Admission Method	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Specialty (of Treatment) vs. Sex*	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓

The term "data consistency" refers to whether related data items within the same data set are inconsistent. For example, a record that indicates a male patient has given birth should be considered inconsistent and would require investigation).

† The national standard for clinical coding stipulates that primary diagnosis/procedure should be completed within 3 months of episode end date. However, these checks look at all submitted coding irrespective of the elapsed time since the episode end date.

\* Reported performance shows the percentage of consistent records for those records that contain a dependant value only. Please refer to the supporting documentation for a full breakdown of the dependencies for each check.

\*\* "Postcode" may occasionally be incorrectly flagged as invalid due to issues with the postcode file received from ONS.

\*\*\* This measure shows consistency of mental health related specialties only.

n/a = no relevant activity data submitted to test in relation to this check

## Appendix C: Additional data quality issues

Issue	Impact	Proposed Resolution	Benefit	Status
APC Weekly submission.	The COVID-19 pandemic has reinforced the need for closer to real-time data.	To change the scope of the Admitted Patient Care data set (APC ds) to include unfinished episodes and to increase the frequency of submission from monthly to weekly.	The move is to change APC to weekly submission and include unfinished episodes within the submission, this will provide data in a timelier manner and allow for a national picture of current activity on a weekly basis.	Changes implemented, frequency of reporting increased and increased scope of reporting. Further post implementation analysis and reviews to take place.
Gaps in data captured	Address gaps and inconsistency's in data not captured in current reporting.	Incorporate data not currently captured in the APC data, such as Nurse activity, therapies, outsourced activity.	Give a clearer picture of activity carried out.	On the Information Quality Improvement (IQI) agenda.
Treatment Function Code	Issues remain with the Treatment Function Codes set out in 2016, impacted the financial costing processes and data continuity in national statistics.	As a temporary fix, codes have been rolled back to the old Treatment Function Codes for costing.	Allows the recording of submitted codes as well as derived codes.	Issue remains on quality agenda, though a resolution is a long term objective.
Pooled GP Codes	Receiving generic GP data in place of a named GP, reference data sets do not hold lookups to these services and therefore all such records are flagged as invalid.	Submission of Pooled GP values as Registered GP are valid. Submission of Pooled GP for referrer code should be invalid, VASS check to remain as is.	Increased validity check performance.	On the Information Quality Improvement (IQI) agenda.
Discharge Ready Date & Estimated Discharge Date	Recording Medically fit for discharge data item has long been discussed on the agenda.	Produce DSCN to establish both fields into the dataset and issue to implement change.	Data items add value to gaining more real time information from the APC dataset.	Agreed Discharge ready date was more appropriate naming than medically fit for discharge. On the Information Quality Improvement (IQI) agenda.