Digital Health and Care Wales

Data Quality Status Report
APC Data Set
2022-23

Document Version	1.0

Status	Approved

Document author:	Hannah Roberts
Approved by	Daniel Hughes
Date approved:	26/07/2023
Review date:	July 2024

TŶ GLAN-YR-AFON 21 Heol Ddwyreiniol Y Bont-Faen, Caerdydd CF11 9AD

TŶ GLAN-YR-AFON 21 Cowbridge Road East, Cardiff CF11 9AD

TABLE OF CONTENTS

1 Data Set

1.1 TITLE

Admitted Patient Care data set (APC ds)

1.2 SPONSOR

Welsh Government (WG)

1.3 IMPLEMENTATION DATE

1st April 1999

1.4 CHANGE HISTORY

See NHS Wales Data Dictionary[1]

1.5 DATA SET PURPOSE

The APC data set is the principal source of secondary use data for hospital admissions. The data set is made up of all episodes, whether these be Finished Consultant Episodes (FCEs) or unfinished episodes. Any activity which is not undertaken by a consultant should be excluded. 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 Digital Health and Care Wales 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.

[1] Admitted Patient Care Data Set (APC Ds) (wales.nhs.uk)

2 Document Purpose

2.1 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.

2.2 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.

2.3 FREQUENCY

The document is issued annually to accompany annual publications and data releases.

2.4 INFORMATION SOURCE

The aim is to describe the quality of the data held centrally in the Digital Health and Care Wales 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[2] (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)[3].

2.5 SCOPE

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

- **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 where the Admission Date is earlier than the Date of 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** 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

[2] WPAS was formerly called Myrddin

[3] All new DSCNs are published on the Digital Health and Care Wales Data Standards website via: Change Notices - Digital Health and Care Wales (nhs.wales)

valid values.

Further information about these dimensions can be found on the Digital Health and Care Wales Data Quality website[4].

3 Data Quality Standards

3.1 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[5]. Data Validity checks have been in operation since April 2010.
- Some data items are interdependent. For example, a patient's date of birth must not be after their attendance 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.

3.2 REGULAR MONITORING

Data Validity and Consistency performance is monitored monthly. 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 Digital Health and Care Wales Data Quality website[6]. The reports themselves are published on the corresponding intranet site[7].

[4] Data Quality - Digital Health and Care Wales (nhs.wales)
[5] WRDS (wales.nhs.uk) (accessible to NHS Wales users only)
[6] Data Quality Standards - Digital Health and Care Wales (nhs.wales)
[7] Data Quality Reports (sharepoint.com) (accessible to NHS Wales users only)

4 Data Set Quality Status

4.1 DATA VALIDITY

Regular monitoring and provider cooperation means that data validity is generally high. Issues causing percentages to fall considerably below the target (>4%) are explained and resolved by the health boards where resource and system constraints permit.

Referrer Code is below target at 6 out of the 8 health boards. For the majority of the health boards, a large percentage of the invalid Referrer Code records are as a result of the use of pooled GP codes as referrer or Locum Doctor activity. The GP updates are not up to date due to staff resource, however Aneurin Bevan are participating in a pilot for the GP updates to be automated. The aim is to be able to turn the warning message on WPAS to red to enforce the user to choose the defined GP. For Cwm Taf, a significant volume of blank referrer codes have been submitted.

Consultant Code is below target for Aneurin Bevan and Powys. A number of the invalid Consultant Code records are caused by nurse specialists/practitioner codes. Work is being carried out to ensure activity is reported against a defined clinician and increasing the scope to include nurse activity, as valid, is on the IQI agenda.

The validity targets for **Main Specialty (consultant)** and **Specialty of Treatment Code** continue to be met every year by all organisations apart from Powys. The records that are causing this low validity, relate mainly 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 and the code '620 (GP Other)' was retired in April 2015 as per DSCN 2014 / 07[8]. The Specialty of Treatment codes for general practice ceased to be valid in April 2016 as per DSCN 2014 / 08[9]. This issue has been highlighted in Powys as it accounts for a high proportion of their overall activity.

Principal Diagnosis is below target for Cardiff & Vale and Aneurin Bevan. Both health boards have submitted a significant volume of blank Principal Diagnosis codes. Overall coding completeness for coded episodes in APC is 91.15% for 2022/23, leaving 88,871 inpatient episodes uncoded.

A copy of the annual Data Validity report for 2022-23 is shown in Appendix A.

4.2 DATA CONSISTENCY

In general, data consistency compliance is good. Issues causing percentages to fall considerably below the target (>4%) are explained and resolved by the health boards where resource and system constraints permit.

The indicators that are showing low percentages are all indicators 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. For example, the low consistency of **Discharge Method v Specialty (of Treatment)** relates to 60 records across all Health Boards with 8 inconsistent episodes and **Discharge Method vs. Discharge Date & Date of Birth (i.e., Age)** relates to 3 records across all Health Boards with 3 inconsistent episodes. Similarly, **Primary Diagnosis Code vs. Admission Date & Birth Date [i.e. Age]** relates to 11 records for Aneurin Bevan

[8] DSCN 2014/07 - Main Specialty (Consultant)
[9] DSCN 2014/08 - Treatment Function Code

with 2 inconsistent episodes.

Betsi Cadwaladr, Powys and Velindre have low consistency rates for **Discharge Method vs. Discharge Destination.** It is low for Betsi Cadwaladr and Powys because conflicting Discharge Destination codes '19 - Own Home', '51' and '55' relating to transfers and '21' relating to temporary residence have been submitted alongside Discharge Method code '8' relating to 'spell not finished'. Similarly, consistency is low for Velindre because these conflicting Discharge Destination codes have been submitted alongside Discharge Method code '4 – Patient died'.

Powys and Cwm Taf have low consistency rates for **Referrer Code vs. Referring Organisation Code** as the Consultant/Nurse codes that have been submitted are not registered against the submitted organisation code.

A copy of the annual Data Consistency report for 2022-23 is shown in Appendix B.

4.3 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 **5 late submissions** received in 2022-23, 2 for Betsi Cadwaladr in June and September 2022, 1 for Cardiff & Vale in June 2022, 1 for Cwm Taf in June 2022 and 1 for Velindre in April 2023.

In addition to monthly deadlines, there is an annual deadline for resubmissions (20th June 2023) which allows providers to improve the quality of their APC data before it is frozen. All Organisations submitted their data before the annual deadline.

4.4 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.

Out of the 8 health boards, only Powys (98.3%) and Velindre (98.2%) achieved the 98% Clinical Coding Completeness target for rolling 12 months' data. Hywel Dda (95%) were within 4% of the target and the remaining health boards, Aneurin Bevan (79.5%), Betsi Cadwaladr (91.9%), Cardiff & Vale (72.8%), Cwm Taf (78.8%), and Swansea (83.8%) were more than 4% below the target.

4.5 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.

4.6 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 Digital Health and Care Wales Data Quality Team on request*)*.

4.7 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.

4.8 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.

The Information Quality Improvement (IQI) initiative is a mechanism to improve the quality of information being used to support patient care and address issues of compliance with national data standards, as well as to review and develop the standards themselves. In turn, the Initiative aims to address deficiencies in the data and information being used to plan, inform, and monitor NHS Wales healthcare activity and services.

The initiative is responsible for driving forward these improvements by identifying national priorities, as well as developing policies and action plans designed to focus attention on the related information quality issues and finding solutions to the underlying causes of these problems. These groups comprise representation from a wide range of bodies which can effect change in this area.

IQI is therefore a coordinated national initiative which proactively seeks to tackle quality issues affecting strategically important information by promoting a collaborative approach which makes best use of scarce resources to drive effectual and worthwhile improvements.

More information about the initiative can be found by visiting Information Quality Improvement Initiative.

5 Appendix A: APC Data Validity Report 2022-23

Data Item	DATA VALIDITY STANDARD	All Welsh Providers	Aneurin Bevan University LHB	Betsi Cadwaladr University LHB	Cardiff & Vale University LHB	Cwm Taf Morgannwg ULHB	Hywel Dda University LHB	Powys Teaching LHB	Swansea Bay ULHB	Velindre NHS Trust
APC submission received by the 17th	-	-	✓	✓	✓	✓	✓	✓	✓	✓
Number of Records Loaded	-	1005539	203567	203327	142222	121166	112022	4260	142523	76452
Administrative Category	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Admission Date	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Admission Method	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Consultant Code	98%	96.5%	89.1%	✓	96.2%	✓	✓	69.2%	✓	✓
Date of Birth	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Decision to Admit Date	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Discharge Date	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Discharge Destination	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Discharge Method	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Episode Start Date	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Ethnic Group	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
HRG Code†††	95%									
Intended Management	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Last Episode in Spell Indicator	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Legal Status	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Local Health Board of Residence	95%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Main Specialty (consultant)	98%	✓	✓	94.9%	✓	96.5%	✓	77.2%	✓	✓
NHS Number	95%	✓	✓	✓	✓	✓	✓	✓	✓	✓
NHS Number Status Indicator	95%	✓	✓	✓	✓	✓	✓	✓	✓	✓

NHS Number Valid & Traced	95%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Patient Classification	95%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Postcode††††	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Principal Diagnosis†	95%	91.8%	85.0%	✓	77.4%	✓	✓	✓	91.5%	✓
Principal Procedure Code†/††	95%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Principal Procedure Date	95%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Referrer Code	98%	65.1%	25.1%	67.3%	✓	77.5%	74.5%	53.8%	50.3%	97.8%
Registered GP Practice Code	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Sex	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Site Code (of Treatment)	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Source of Admission	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Specialty of Treatment Code	98%	✓	✓	✓	✓	✓	✓	76.0%	✓	✓
Waiting List Date	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓

6 Appendix B: APC Data Consistency Report 2022-23

Data Consistency Check	DATA CONSISTENCY STANDARD	All Welsh Providers	Aneurin Bevan University LHB	Betsi Cadwaladr University LHB	Cardiff & Vale University LHB	Cwm Taf Morgannwg ULHB	Hywel Dda University LHB	Powys Teaching LHB	Swansea Bay ULHB	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%	✓	✓	✓	96.4%	✓	✓	✓	✓	95.0%
Decision to Admit Date vs. Admission Date	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Decision to Admit Date vs. Waiting List Date	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Discharge Method vs. Discharge Date & Date of Birth [i.e. Age]*	98%	0.0%	n/a	0.0%	n/a	n/a	n/a	n/a	n/a	n/a
Discharge Method vs. Discharge Destination*	98%	93.9%	✓	78.1%	✓	✓	97.5%	93.8%	✓	90.0%
Discharge Method vs. Specialty (of Treatment)*	98%	86.7%	✓	57.1%	91.4%	0.0%	0.0%	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%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Patient Classification vs. Discharge Date & Admission Date [i.e. Length of Stay]*	95%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Postcode vs. Local Health Board of Residence**	95%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Primary Diagnosis Code vs. Admission Date & Birth Date [i.e. Age]†*	95%	✓	81.8%	✓	✓	✓	✓	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%	97.6%	✓	✓	✓	89.3%	✓	93.9%	✓	✓

Specialty (of Treatment) vs. Sex*	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Waiting List Date vs. Admission Date	98%	✓	✓	✓	✓	✓	✓	✓	✓	✓
Waiting List Date vs. Admission Method	98%	✓	✓	✓	✓	✓	✓	✓	✓	97.1%

7 Appendix C: Additional Data Quality Issues

Issue	Impact	Proposed Resolution	Benefit	Status
Assessment Unit (AU) Activity	Inconsistency in approaches to recording assessment activity across Wales.	A national review is being undertaken to consider an appropriate approach to the future recording and reporting of AU activity. For the purposes of financial costing only, an alternative approach to identify 'short stay' emergency activity (based on episode length) is being used by the WG Financial Information Strategy in the interim.	Availability of consistent data relating to assessment activity.	Workshop taken place July 2023 to review the draft new definition of Assessment Units. It has been agreed within these workshops that the terminology will be changed to 'Acute Clinical Units'. Definition to be shared with Health Boards for final feedback.
Source of Admission / Discharge Destination	Inconsistency in values recorded in Source of Admission and Discharge Destination, particularly in records relating to transfers, causing difficulties in tracking patient journeys, and deriving provider spell data.	A consultation with health boards in September 2013 revealed little appetite for changes in national definitions. Compliance with data quality standards continue to be monitored with issues being addressed on a case-by-case basis.	Improvements in the accuracy of these data would allow for stricter logic in scripts used to derive provider spells and greater accuracy in related analyses.	New values were implemented April 2019. Analysis has shown mixed uptake of new values and large use of retired values, issues remain.
Nurse led activity	Non consultant activity is not currently recorded in the dataset giving an incomplete picture of inpatient activity.	Incorporate nurse led activity into the APC data.	Give a clearer picture of activity carried out.	Changes to increase APC scope to go via standards assurance process. Follow up with health boards and document requirements and timescales for the inclusion of the activity into data submissions, documented feedback to be provided back to Welsh Government.

Issue	Impact	Proposed Resolution	Benefit	Status
Treatment Function Code	620 has been invalid since 2015/16 as a Treatment Function code. 4 of 7 health boards submit monthly data under this value.	Address TFC value set to enable capture of GP led data.	Allow Organisations to record GP led data within the dataset using valid reference data.	PAS standardisation work on hold. Not currently being reviewed.
Admission Method	Missing the ability to record certain types of admission along with several outdated values that are still active.	To review usage of the existing values and enact changes to existing data standard.	Improve Data Quality.	Impact Assessments conducted, proposed changes to be put into draft DSCN to go via the assurance process.
Pooled GPs	The use of a Pooled GP value as a referrer code in monthly data quality reporting flags entries as invalid during the VASS checks.	PAS team to collate a new GP file to provide more up to date reference data which will then in turn reduce the number of records mapped to GP Pooled list.	Improve Data Quality.	PAS team are looking at collating a new GP file. Swansea Bay contacts in the process of updating GP file, BCU to be completed first then will move onto other HBs to implement changes so GP data will be correct moving forward.
Site Code of Treatment	Issue raised via BCU on the usage of the site code of treatment field and what processes are in place for creation of new sites.	Agree a standardized approach for the handing of data 'Not a Hospital Site'.	Give a clearer picture of activity carried out.	Look at the data behind 'Not a Hospital Site' submissions and establish if this should be retired and if reference data should be used for all unique locations.
Diagnosis codes	Issues with the overall quality and completeness of coding with specific problems in Dermatology, Midwifery and Highly Sensitive Conditions. Users of the APC ds data are making decisions based on an incomplete picture which risks poor outcomes.	Health boards to address issues regarding low quality and completeness of coding.	Improved quality and completeness of coding data.	Ongoing feedback provided to health boards along with access to dashboards to allow them to resolve the issues.