

Friday, December 2, 2022

Update: December 2022 CPRD Aurum release

Dear Colleague

As part of CPRD's database build process, data quality assurance checks are carried out that have flagged a potential anomaly within the CPRD Aurum December 2022 database. We have therefore taken the decision to delay the CPRD Aurum release whilst an investigation takes place.

We apologise for any inconvenience this may cause and will keep you updated with any progress.

Best regards
CPRD

Monday, December 5, 2022

Update: December 2022 CPRD Aurum release

Dear Colleague

This follows from our earlier message that the CPRD Aurum December 2022 build release has been cancelled due to potential data quality issues.

We have found that the issues identified during preparation of the December 2022 Aurum build also affect the September, October and November 2022 builds. The main observation is a noticeable reduction in the numbers of records, in the DrugIssue and Observation tables over recent months. It appears to be connected to the recent migration of our data flows from the [REDACTED] legacy system to their new cloud-based [REDACTED] Analytics platform, and seems to affect most practices and individuals in CPRD Aurum. **Please do not use builds 202209, 202210 or 202211 until further notice.** Any further CPRD Aurum builds will be delayed until we have identified an explanation for this observation and resolved any issues (as needed).

The May 2022 CPRD Aurum build (202205) is unaffected; please use this build for all studies until further notice.

If you are partway through analysing a large dataset drawn from a more recent build, please contact us at CPRD Enquiries enquiries@cprd.com. We can provide a file of the last collection dates of practices (from the May 2022 CPRD Aurum build), which can be used to safely truncate the data.

CPRD GOLD is not affected.

We apologise for the inconvenience, are grateful for your patience, and are working to resolve these issues as quickly as possible. If you have any questions, please email us at enquiries@cprd.com.

Best regards
CPRD

Thursday, December 15, 2022

Update: CPRD Aurum Recommendations

Dear Colleague

As you know, data quality issues have been identified which affect the September, October and November 2022 builds of CPRD Aurum. As a reminder, **please do not use builds 202209, 202210 or 202211 until further notice** as a precautionary measure. To avoid

accidental access to these builds by users who may have missed these notices, we will be removing these CPRD Aurum builds from the data access portal. Instead, we will be making the CPRD Aurum May 2022 build available on the data access portal by the end of next week. If you need a redelivery of linked data because of these data quality issues, there will be no additional redelivery charge. Please share this update with all CPRD users in your institution.

CPRD GOLD is not affected.

We apologise for the inconvenience, are grateful for your patience, and are working to resolve these issues as quickly as possible. We will continue to send you regular updates. In the meantime, if you have any questions or you wish to discuss the impact of these data quality issues on your studies, please email us at enquiries@cprd.com.

Best regards

CPRD

Tuesday, December 20, 2022

Update: CPRD Aurum Recommendations

Dear Colleague

I wanted to send a personal message to you about the CPRD Aurum data quality issues with an update on where we are and to thank you for your patience and support during this time. Please share this message with all CPRD users in your institution.

As you know, data quality issues have been identified which affect CPRD Aurum (see previous messages and Q&A below). While these issues are being resolved, the newly restored CPRD Aurum May 2022 release should be used for your studies. This will be available on the data access portal by the end of this week (23/12).

CPRD now recommends that any client who has extracted CPRD Aurum data since Thursday 1 September 2022 from any version of the database (including May 2022, April 2022 etc.) should replace this with a re-extraction from the newly restored CPRD Aurum May 2022 database.

Potential scenarios:

1. CPRD Aurum data from May 2022 or a previous build was used in your study. The last Define/Extract used to access the data was run prior to September 2022. Your data is not impacted, no further action required.
2. CPRD Aurum data from May 2022 or a previous build was used in your study. Define/Extract has been used to access the data from Thursday 1 September 2022 onwards. For reproducibility, we recommend using the newly restored CPRD Aurum May 2022 full data extract as a replacement for your study data.
3. CPRD Aurum data from September 2022 or a later build was used in your study. We recommend using the CPRD Aurum May 2022 full data extract as a replacement for your study data.

Please note that there will not be a December 2022 or January 2023 release of CPRD Aurum. CPRD GOLD is unaffected and the next release is due to take place on Monday 11 January 2023.

How is CPRD resolving the data quality issues?

CPRD receives two types of data collections for CPRD Aurum: a full data extract (containing all the data for a practice at a given point in time) and a daily delta extract (containing incremental changes only). After investigation, CPRD has determined that the full data extracts provide a more accurate representation of the data held at the practice than the deltas used previously. CPRD will now start a process of downloading full data extracts for

all practices. CPRD will also implement further data quality checks to ensure the accuracy of the data that is released.

How quickly will the data quality issues be resolved?

Due to the volumes of data involved, there will not be a December 2022 or January 2023 release of CPRD Aurum. Future releases will include incremental updates to an increasing subset of practices until all have had their full data reextracted. We will provide updates on the next release of CPRD Aurum in the new year.

Please contact us at CPRD Enquiries enquiries@cprd.com if you have any questions.

Once again, please accept my apologies for this inconvenience and thank you for your ongoing understanding while we work towards full recovery.

Best regards,

Puja Myles

Director, CPRD

Tuesday, January 17, 2023

Update: CPRD Aurum Data Quality Issues

Dear Colleague

As we work to resolve the data quality issues with CPRD Aurum, we wanted to provide an update on where we are and to thank you for your continued patience and support during this time. Please share this message with all CPRD users in your institution.

In our last message, we recommended that **any client who has extracted CPRD Aurum data since Thursday 1 September from any version of the database (including May 2022, April 2022, etc.) should replace this with a re-extraction from the newly restored CPRD Aurum May 2022 database.** To support you with this recommendation, please find some additional information about the issues below:

What is the level of inaccuracy in the September, October, and November 2022 releases?

As highlighted, the September, October, and November 2022 releases were based on data received following the migration of [REDACTED] data flows that occurred in the summer of 2022. The daily updates (deltas) used to create the releases were at least 50% smaller than we would expect given previous data flows – investigations have indicated that the missing data impacts all GP practices during this time period. We recommend using the newly restored CPRD Aurum May 2022 release as a replacement in any studies originally accessing CPRD Aurum from September 2022 onwards.

Why are CPRD Aurum releases prior to, and including, May 2022 impacted?

During the migration of data flows, a large volume of incorrect data was received from [REDACTED] that needed to be deleted from our database. Unfortunately, this deletion led to some records being removed from the database which should not have been. This means that if CPRD Aurum releases prior to, and including, May 2022 were accessed from Thursday 1 September 2022 onwards, they would not have had the complete set of records that were included at the original time of release. For reproducibility, we recommend using the newly restored CPRD Aurum May 2022 database as a replacement in any studies where Define/Extract was used to access the data from September onwards.

What patterns of missing data have been identified in these earlier releases?

Investigations repeating previous queries that had been run on earlier releases (up to and including May 2022) returned 2-35% fewer records than expected.

How quickly will the data quality issues be resolved?

The next release of CPRD Aurum will include updated data for 200 new practices added since the May 2022 release, whilst existing practices will have their data included as per the restored May 2022 release. We are in the process of developing a series of additional data quality checks we will apply prior to release, which we plan to share with you shortly. There will not be a release of CPRD Aurum data in February 2023 as the data processing is taking longer than originally anticipated. We are working towards a March 2023 release and will confirm the timing of the next release nearer the time once we have completed our data quality checks. Future releases will include incremental updates to an increasing subset of practices until all have had their full data reextracted.

Please contact us at CPRD Enquiries (enquiries@cprd.com) if you have any questions.

Once again, please accept our apologies for this inconvenience and for your ongoing understanding while we work towards full recovery.

Best regards

CPRD

Wednesday, February 1, 2023

Update: CPRD Aurum Data Quality Issues

Dear Colleague,

During the ongoing work to resolve the data quality issues with CPRD Aurum we wanted to provide a further update on where we are and to thank you for your continued patience and support. Please share this message with all CPRD users in your institution.

We are working towards a March 2023 release and will confirm the timing of the next release once we have completed our additional data quality checks.

The next CPRD Aurum release will contain:

- (1) existing CPRD practices as in the May 2022 CPRD Aurum release, and
- (2) data collected up until January 2023 for circa 200 practices that are now contributing data to CPRD but were not present in the original May 2022 release.

Based on the issues from late 2022, additional checks will be performed to ensure that we have full clinical data for the new practices and that previous CPRD Aurum releases are unaffected. The checks will include:

- a comparison of the record counts per table in the raw data at CPRD against those in bulk [REDACTED] data;
- a check that the volume of data received by CPRD is within the expected range for practices based on previous unaffected builds;
- rerunning DEFINE and EXTRACT queries that were previously run in the unaffected May 2022 build for practices included in that build.

Please contact us at CPRD Enquiries (enquiries@cprd.com) if you have any questions.

Please accept our continued apologies once again for this inconvenience and for your ongoing understanding while we work towards full recovery.

Best regards

CPRD

Monday, April 3, 2023

April 2023 CPRD GOLD release and CPRD Aurum update

Dear Colleague

The April 2023 release of CPRD GOLD is now live on the Azure Data Access portal. Please find attached release notes that detail changes to the lookups, data and customer applications.

Updated lookup documents will need to be requested from CPRD Enquiries.

Due to the bank holidays on 1 May and 8 May, the next release of CPRD GOLD is due to take place on Monday 15 May 2023.

CPRD Aurum

The CPRD Aurum build intended for release in March 2023 is ready but there are known missing data issues in records relating to immunisations, allergies and referrals in the observation table. As a number of researchers have requested updated data for their studies, we have decided to delay this release until we receive further source data verification (SDV) checks from [REDACTED] to confirm that this is a localised issue. These issues will be rectified in subsequent builds.

We will make reports from these SDV checks available to nominated users along with the build so that researchers can make an informed decision about whether they should use this build or wait for the next build, based on specific study requirements.

We will be back in touch with further information about when we will release this build. We would like to thank you again for your patience and understanding while we are working on these data quality checks. Further information about our quality checks can be found on our [Data quality web page](#).

If you have any questions or feedback, or have any colleagues who wish to be added to our mailing list to receive these Release Notes, please contact CPRD Enquiries at enquiries@cprd.com.

Best regards

CPRD

Wednesday, April 19, 2023

Update: CPRD Aurum data quality

Dear Colleague

As noted in our 3 April communication, there are known issues with missing data in records relating to immunisations, allergies and referrals in the Observation table of the latest CPRD Aurum build.

Please see the table below for source data verification (SDV) checks completed by the CPRD Aurum data supplier ([REDACTED]) on data available in April 2023. Checks were based on all new practices in the planned CPRD Aurum March 2023 build, and a random sample of old practices included in the CPRD Aurum May 2022 build. Counts from both builds were

compared to counts in [REDACTED] Web (the online system where GPs enter patient healthcare data) with no time restrictions applied.

Checks confirm that data in the CPRD Aurum March 2023 Observation table on tests and the Drug Issue tables are reliable. Checks on the Consultation and Referral tables are ongoing, and we will provide an update on these and the timing of the release shortly.

Our initial tests confirm that for new practices (i.e., those newly contributing to CPRD Aurum since the migration to [REDACTED] [REDACTED] there was an issue with the flow of immunisations data to the CPRD Aurum build, such that there are very few recorded events. Our discussions with [REDACTED] and initial SDV checks suggest that the missing records are a localised issue affecting immunisation, allergies and referrals data in this build. These data are now held in separate tables in [REDACTED] [REDACTED] and we will be able to access these data, and rectify these issues in subsequent CPRD Aurum builds. Please note that whilst we receive the data via separate tables, CPRD will incorporate them into the existing CPRD Aurum structure in any future builds.

Assuming that there are no issues with the Consultation table, we will release the March 2023 version of the CPRD Aurum build. We will make the full report from these SDV checks available to nominated users, along with the build, so that researchers can make an informed decision about whether they should use this build or wait for the next build, based on specific study requirements.

We would like to thank you again for your patience and understanding while we are working on these data quality checks.

Clinical Events	Data Table	Practice Status	N Practices	Counts in [REDACTED] Web	Counts in CPRD Aurum	% Counts in CPRD Aurum & CPRD Web
Total Cholesterol	Observation	New	228	12,575,540	12,297,199	97.79 ^a
HbA1c	Observation	New	228	10,957,956	10,622,385	96.94 ^a
COVID-19	Immunisation	New	228	5,814,326	196	0.00 ^b
MMR	Immunisation	New	228	4,419,999	3,405	0.08 ^b
COVID-19 product	Drug Issue	New	228	4,735,699	4,674,063	98.70 ^a
MMR product	Drug Issue	New	228	22,405	22,178	98.99 ^a
Nuvaxovid product	Drug Issue	New	228	755	712	94.30 ^a
Allergy	Observation	New	228	TBD	TBD	TBD
Referral	Observation	New	228	TBD	TBD	TBD
Consultation	Consultation	New	228	TBD	TBD	TBD
Total Cholesterol	Observation	Old	226	16,198,410	15,269,089	94.26 ^c
HbA1c	Observation	Old	226	15,945,500	14,145,942	88.71 ^c
COVID-19	Immunisation	Old	226	8,650,634	7,339,187	84.84 ^c
MMR	Immunisation	Old	226	5,596,370	4,755,489	84.97 ^c
COVID-19 product	Drug Issue	Old	226	6,871,848	6,002,509	87.35 ^c
MMR product	Drug Issue	Old	226	24,507	22,217	90.66 ^c
Nuvaxovid product	Drug Issue	Old	226	1,460	1,202	82.33 ^c
Allergy	Observation	Old	226	TBD	TBD	TBD
Referral	Observation	Old	226	TBD	TBD	TBD
Consultation	Consultation	Old	226	TBD	TBD	TBD

^a Discrepancies due to national opt-outs being applied to the CPRD Aurum data and not to [REDACTED] Web data.

^b Discrepancies due to immunisation data not flowing to CPRD Aurum (this is a localised issue, which we can rectify in subsequent builds)

^c Discrepancies are due to opt-outs being applied to the CPRD Aurum data in addition to differences in temporal coverage (counts from CPRD Aurum are based on data up to May 2022, counts from [REDACTED] web are based on counts to April 2023)

Best regards

CPRD

Friday, April 28, 2023

March 2023 CPRD Aurum release

Dear Colleague

The March 2023 release of CPRD Aurum is now live on the Azure Data Access portal. Please find attached release notes that detail changes to the lookups, data and customer applications. Updated lookup documents will need to be requested from CPRD Enquiries. As noted in previous communications, ***there is limited additional temporal coverage in this database and known missing data issues in records relating to immunisations and referrals in the observation table.*** This CPRD Aurum release includes data from 1,720 practices, comprising:

- 228 new practices with data collected up to the end of January 2023 (currently contributing practices)
- 1,491 practices with the same data as the May 2022 release
- 1 additional practice with data to March 2022

The output from our source data verification (SDV) checks is available as an annex in the release notes (attached) for this release. **Please read these additional notes to support making an informed decision about whether to use this database or wait for the next update based on specific study requirements.**

We are now working on the next release of CPRD Aurum, which will include additional temporal coverage for a proportion of existing practices, as well as the missing immunisation and referrals data from the new practices included in this March 2023 release.

We would like to thank you again for your patience and understanding while we undertook these necessary data quality checks, and as we work towards restoring a full CPRD Aurum database.

If you have any questions or feedback, or have any colleagues who wish to be added to our mailing list to receive these Release Notes, please contact CPRD Enquiries at enquiries@cprd.com.

Best regards

CPRD

Wednesday, July 26, 2023

CPRD Aurum update

Dear Colleague

Due to ongoing validation and data quality checks the next release of CPRD Aurum is delayed and will not take place on Monday 31 July 2023 as previously anticipated.

We will be back in touch with further information about the timings for this next CPRD Aurum release.

We would like to thank you again for your patience and understanding while we work towards restoring a full CPRD Aurum database.

The next CPRD GOLD release is due to take place on Monday 7 August 2023.

Best regards

CPRD