HCP Subjects with Identified Quality Control Issues (QC_Issue measure codes explained)

Subjects with Issues noted in the HCP quality control process

As of the S1200 release, a subject data measure (QC_Issue) has been added to ConnectomeDB to flag subjects with notable issues found in the HCP Quality Control process as described for each issue code below. The issues are notable, but were not considered severe enough to exclude the subject’s imaging data from release. We are sharing these findings of issues to alert users and allow them to decide whether to include or exclude these subjects in their own analyses.

The QC_Issue measure allows users to filter the data for subjects with or without these issues in the ConnectomeDB Subject Dashboard or in the CSV spreadsheet containing individual subject data that is downloadable from the site.

To view the issue codes in the ConnectomeDB Subject Dashboard, click the “Add Tab” dropdown at the top right of the section listing subjects and choose “QC Issues”. To filter on subjects with issue codes, click the “Add New Filter” button, select the “QC Issues” category, “Quality Control Issue Noted” assessment, “Coded QC Issue” attribute, and “IS NOT EMPTY” operator (see image). At the time of the S1200 Release, there are 157 subjects with imaging data available that have one or more QC_issue codes.

For some issues (e.g., scan and surface QC issues), more subject specific information, including affected run information for issue D, is available here on our QC Issues wiki page.

Note on the HCP QC process:

We explicitly reviewed the outputs of the HCP Structural Pipelines on all subjects and ascertained that all included subjects had at least reasonable (or better) cortical surface models. (Examples of some of the issue categories are on the QC Issues wiki page). However, a comprehensive QC of all modalities was not feasible, and HCP does not claim to have identified all QC issues for all HCP subjects and modalities. In particular, in the interest of providing data, fMRI and dMRI scans were only very rarely excluded for motion, and we have opted to release fMRI and dMRI scans collected in periods of known (albeit intermittent) problems with the head coil, which resulted in some temporal instabilities in acquisition.
Therefore, the absence of an issues code for a given subject does not imply that there are no issues with the data for this subject. We encourage users to alert us to issues that they discover. We will then update the QC_Issues field and wiki appropriately for the benefit of other users.

### Issue code A: Anatomical anomalies

All HCP subject structural scans and FreeSurfer segmentation and surface outputs from the structural pipeline were subject to a standard quality control process that involved manual viewing and rating of scan quality and anatomical abnormalities by an experienced rater (see Appendix 4: Post Session 1 Quality Control SOP, pp. 17-24). Brain anomalies evident in T1w and T2w scans were noted and further reviewed by a radiologist. Subjects with major radiologic anomalies that were expected to substantially impact brain connectivity were removed from the study, and their imaging and behavioral data are not included in the released data.

Some subjects had focal anomalies that are considered as normal variants or benign findings. We have released their data, but because of their altered anatomy, using data from these subjects may affect some analyses. These subjects are flagged with the issue code A in the QC_Issue measure.

A description and images of the location and kind of anomaly for each of these subjects are detailed in the following spreadsheet of Anatomical QC Findings: Anatomical-Issues-Subjects-xlsxmatch.xlsx and in the accompanying PDF of powerpoint slides showing Anatomical QC Findings: Anatomical-Issues-Subjects-xlsxmatch.pdf.

**Main HCP Issue A Subjects:**

<table>
<thead>
<tr>
<th>Subject ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>433839</td>
</tr>
</tbody>
</table>

**Retest HCP Issue A Subjects:**

**Issue code B: Segmentation and Surface QC**

All brain surfaces and brain structure segmentation results generated by FreeSurfer as part of the HCP structural pipeline were also inspected by a rater for any notable issues see Appendix 4: Post Session 1 Quality Control SOP, pp. 25-28). If there were major segmentation or surface errors, we reran the pipeline to attempt to fix the issues in an automated way (due to the cohort size of the HCP and limited staff resources, hand editing of segmentations and surfaces was not attempted on any subjects).

Subjects with focal segmentation and surface errors in their structural pipeline outputs are flagged with the issue code B for the QC_Issue measure.

A description and images of the location and kind of error for each of these subjects are detailed in the following Spreadsheet of Segmentation /Surface QC Findings: Segmentation-Issues-Subjects-xlsxmatch.xlsx and in the accompanying PDF of powerpoint slides showing Segmentation /Surface QC Findings: Segmentation-Issues-Subjects-xlsxmatch.pdf.

**Main HCP Issue B Subjects:**

<table>
<thead>
<tr>
<th>Subject ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>433839</td>
</tr>
</tbody>
</table>

**Retest HCP Issue B Subjects:**

**Issue code C: Some of subject’s data acquired during a period of instabilities in the head coil**

Data for some subjects was collected during discrete periods in which there were intermittent problems with the temporal stability of the head coil. Subjects collected during these periods of temporal instability are flagged with the issue code C for the QC_Issue measure and a list of these subjects is also available below.

Since spatially localized temporal instability is particularly problematic for analysis of resting state data, we evaluated network matrices ("netmats") derived from both the “minimally preprocessed” (MPP) data and the FIX-ICA cleaned data, for all rfMRI runs of all subjects scanned during these periods. As expected, many of the netmats from the MPP rfMRI data acquired during these periods contained severe artifacts. Indeed, these periods are a primary contributor to the subjects flagged with Issue Code D (see below). However, the resting-state netmats following the FIX cleaning process were reasonable in all cases. Most of the time, FIX automatically removed the temporal instability, as intended. However, in a small proportion of scans manual reclassification of some of the FIX-ICA components from ‘signal’ to ‘noise’ was necessary to appropriately remove artifact from the FIX-denoised rfMRI scans. (Subjects for which manual reclassification was necessary in one or more rfMRI scans are indicated via Issue Code E, see below).

The situation is less clear for the tfMRI and dMRI data of these ‘temporal instability’ subjects. We have not explicitly evaluated the tfMRI and dMRI data of these subjects, but both are very likely to contain artifacts arising from the temporal instability problems with the head coil. Currently, FIX-denoising has not been applied to the tfMRI data. It is possible that the impact of the temporal instability may be ameliorated to some degree in the released HCP Task Analyses by the model-constrained nature of these analyses. (However, any analysis of the MPP tfMRI data from those subjects using a "resting-state" like approach is almost certainly problematic). As for the dMRI data, it is possible that the "outlier correction" feature in ‘eddy’ will yield appropriately cleaned data for subjects scanned during these periods, but this remains to be investigated.

In summary, (1) the FIX-cleaned rfMRI scans from these subjects is fine to use; (2) the MPP rfMRI data from most of these subjects is particularly problematic, and is provided mostly with the intent that it will serve as a useful test-bed for testing denoising algorithms; and (3) use caution and extra review if you decide to use the tfMRI and dMRI data from these subjects.

**Main HCP Issue C Subjects:**

<table>
<thead>
<tr>
<th>Subject ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>433839</td>
</tr>
</tbody>
</table>
RestScansWithProminentArtifactInMPP_Codes_D.xlsx

reclassification of some FIX components (Code E) are detailed in the following Spreadsheet:

A list of the rfMRI scans and subjects exhibiting MPP artifact that was cleaned automatically by FIX (Code D) and cleaned after manual processing of some FIX components was performed to ensure appropriate denoising.

As expected, nearly all of the subjects affected with Issue Code C are also flagged with Issue Code D. However, Issue Code D includes additional subjects for whom prominent artifacts were empirically observed in their MPP rfMRI data, outside of known windows of head coil instability problems.

As noted above (for Issue Code C) we evaluated network matrices (“netmats”) derived from both the MPP and the FIX-ICA cleaned data, for all rfMRI runs associated with Issue Code D. Again, the resting-state netmats following the FIX cleaning process were reasonable in all cases. (As above, if necessary, manual reclassification of some FIX components was performed to ensure appropriate denoising).

In summary, (1) the FIX-cleaned rfMRI scans from subjects with Issue Code D is fine to use; and (2) the MPP rfMRI data from these scans is particularly problematic, and is provided mostly with the intent that it will serve as a useful test-bed for evaluating denoising algorithms.

A list of the rfMRI scans and subjects exhibiting MPP artifact that was cleaned automatically by FIX (Code D) and cleaned after manual reclassification of some FIX components (Code E) are detailed in the following Spreadsheet: RestScansWithProminentArtifactInMPP_Codes_D&E.xlsx

Main HCP Issue D Subjects:

101107 106016 108323 109830 111716 116726 118528 117237 120315 130467 134728 136126 136631 138130 145632 146129 151728 155635 158843 161630 162733 165941 166438 167238 176744 177241 178647 181232 186040 191841 194847 198249 203418 204218 206929 211215 211720 220832 231928 238033 257946 299765 308331 336841 339847 382242 386250 392750 433839 453441 468050 510326 558657 559457 567759 590047 599469 609143 614439 627549 628248 662551 706040 723578 738654 759725 779370 783462 786569 792766 815247 818455 818859 825553 880157 902242 943862 957974 962058 970764

Retest HCP Issue D Subjects:

115320 130518 135528 139839 146129 169343 172332 175439 204521 562345 662551 917255

Issue code E: Manual reclassification of some FIX-ICA components

As noted above for Issue Codes C and D, in a small number of instances the automatic labeling of ICA components as either “signal” or “noise” by FIX was incorrect, and as a result, the initial denoising by the automatic FIX-ICA process was insufficient, and prominent artifacts remained in the rfMRI scans after FIX. (This probably occurred because the set of scans used to train FIX early in the HCP project did not include some of the more rare/unique types of artifacts that we encountered over the course of the study). For these rfMRI scans, we manually reclassified one or more “signal” components as “noise” components and ran the ReApplyFix pipeline to appropriately clean these data. Note that for the manually reclassified datasets, we used the original MSMAAll registration rather than re-running MSMAAll after the reclassification. We do not expect a substantial difference in registration because MSMAAll relies on a multi-variate approach to resting state fMRI modeling (weighted regression) that is more robust than the univariate approaches (e.g. correlation) typically used to generate functional connectivity maps. In addition, for these subjects, the ReApplyFix (for both MSA-All and MSMAAll registrations) and RestingStateStats pipelines were run on the Washington University Center for High Performance Computer (WU-CHPC) Cluster #2, rather than the WU-CHPC Cluster #1, on which all other 3T processing pipelines were run for all S1200 subjects.

As noted above (for Issue Code C) we evaluated network matrices (“netmats”) derived from both the MPP and the FIX-ICA cleaned data, for all rfMRI runs associated with Issue Code D. Again, the resting-state netmats following the FIX cleaning process were reasonable in all cases. (As above, if necessary, manual reclassification of some FIX components was performed to ensure appropriate denoising).

In summary, (1) the FIX-cleaned rfMRI scans from subjects with Issue Code D is fine to use; and (2) the MPP rfMRI data from these scans is particularly problematic, and is provided mostly with the intent that it will serve as a useful test-bed for evaluating denoising algorithms.

A list of the rfMRI scans and subjects exhibiting MPP artifact that was cleaned automatically by FIX (Code D) and cleaned after manual reclassification of some FIX components (Code E) are detailed in the following Spreadsheet: RestScansWithProminentArtifactInMPP_Codes_D&E.xlsx

Main HCP Issue E Subjects (no Retest scans with this issue):

101107 111716 134728 136631 148436 162733 165941 178647 186040 204218 211720 336841 567759 614439 628248 763557 783462 815247 818455 818859 825553 943862
Reference