HCP Data Release Updates: Known Issues and Planned fixes

The WU-Minn Human Connectome Project (HCP) releases imaging data collected on subjects on a regular basis. Details on what data is included in each release are available in the Reference Manual for that release on the HCP documentation page. On this wiki page, we document known updates, data issues, bugs, and known missing files in the data distributed with each release (including those distributed in Connectome in a box for each release) and plans to fix or patch each bug. Date added is listed at the beginning of entries.

1200 Subjects Release (March 1, 2017)

Major Issues

- *5/4/2020:* **Subject 169444 preprocessed data from several 7T fMRI runs unusable**

  The following runs from 169444 are unusable due to major acquisition artifacts and instability. The following runs in the released data are affected and should not be used:

  rfMRI_REST1_7T_PA

  tfMRI_RETBAR1_7T_AP

  tfMRI_RETBAR2_7T_PA

  tfMRI_RETCCW_7T_AP

  tfMRI_RETCON_7T_PA

  tfMRI_RETCW_7T_PA

  tfMRI_RETEXP_7T_AP

  Data from the remaining 7T fMRI runs for this subject (and all other 3T and 7T data) are usable.

  No planned Fix: These data are released and we do not plan to remove them from ConnectomeDB or AWS S3. We recommend excluding these runs from your analyses.

- *3/9/2020:* **3T Functional Preprocessing Error of all 3T "RL" fMRI runs in 25 Subjects (7 are also 7T Subjects)**

  3T Functional MRI data from a batch of 25 subjects were incorrectly processed, with all "RL" runs (right-left phase encoding) being processed with the wrong phase encoding direction. Not only are the preprocessing results for the RL functional runs themselves incorrect, but the MSM-All registration for these subjects, which is dependent on the correct preprocessing outputs, is also incorrect. Therefore any analyses using the MSM-All aligned data from these subjects would be affected. Further, 7 of these subjects also have 7T data, so analyses of 7T data that use MSM-All aligned data would be affected as well (since the MSM-All registration applied to the 7T data was based on the 3T data). No Test-Retest subjects were affected.

  The 3T "LR" (left to right phase encoding) runs for these subjects were processed correctly, but we recommend using both runs in analyses to mitigate phase encoding bias. The MSM-Sulc data for the 3T LR functional runs and 7T functional data are not misaligned, but we don't recommend using MSM-Sulc aligned data for cross subject analyses.

  Recommendations:

  Due to the incorrect preprocessing of half the runs, we recommend excluding these subjects from any further analyses of the 3T functional data.

  7T preprocessed Functional data (all runs) were processed correctly and can be used. However, due to the non-optimally aligned MSM-All surfaces, if you are performing cross-subject analyses using the 7T functional data, your options are to either: (1) exclude these subjects from your analysis, (2) confirm that the MSM-All registration in these subjects is still reasonable*, or (3) use the MSM-Sulc registered surfaces instead for all subjects (not recommended).

  3T and 7T preprocessed Diffusion data for these subjects was not affected by the functional preprocessing error and can be used. However, due to the non-optimally aligned MSM-All surfaces, if you are performing cross-subject diffusion analyses involving surfaces, your options are to either: (1) exclude these subjects from your analysis, (2) confirm that the MSM-All registration in these subjects is still reasonable*, or (3) use the MSM-Sulc registered surfaces instead for all subjects (not recommended).

  *One may determine that the MSM-All registration is reasonable in the affected subjects by:

  - Comparing affected subjects' MSMAll myelin maps `${StudyFolder}/${Subject}/MNINonLinear/${Subject}.MyelinMap_BC_MSMAll.164k_fs_LR.dscalar` to the group template available for download in https://github.com/Washington-University/HCPpipelines/blob/master/global/templates/MSMAll/Q1-Q6_RelatedParcellation210.MyelinMap_BC_MSMAll_2_d41_WRN_DeDrift.32k_fs_LR.dscalar.nii

  - Looking at the affected subjects' MSMAll areal distortion maps `${StudyFolder}/${Subject}/MNINonLinear/${Subject}.ArealDistortion_MSMAll.164k_fs_LR.dscalar.nii`

  Affected subjects, with 7T subjects noted:

```
103010
113417
116423
120010
121719
127226
130114    7T
143830
169040    7T
185038
189652
202820
204218
329844
385046    7T
401422    7T
462139
469961
644246    7T
688569
723141
908860
943862    7T
969476
971160    7T
```

No planned Fix: These data are released and we do not plan to reprocess/fix or remove them from ConnectomeDB or AWS S3. We recommend excluding these subjects from your analyses of 3T or 7T functional data.

- *3/2/2020:* **Subject 168139 preprocessed data from several tfMRI runs unusable**

  Subject 168139 has an anatomical anomaly, as noted in HCP Subjects with Identified Quality Control Issues (QC_Issue measure codes explained)*, a large posterior fossa subarachnoid cyst. The cyst was large enough in this case to cause the EPI to T1w alignment to fail severely in the fMRI Volume part of functional preprocessing, and causes all subsequent preprocessing results to be unusable. The following runs in the released data are affected and should not be used:*

  tfMRI_EMOTION_LR

  tfMRI_GAMBLING_LR

  tfMRI_GAMBLING_RL

  tfMRI_LANGUAGE_LR

  tfMRI_LANGUAGE_RL

  tfMRI_MOTOR_LR

  tfMRI_MOTOR_RL

  tfMRI_RELATIONAL_LR

  tfMRI_RELATIONAL_RL

  tfMRI_SOCIAL_LR

  tfMRI_SOCIAL_RL

  tfMRI_WM_LR

  tfMRI_WM_RL

  Data from the remaining fMRI runs for this subject (all resting state and tfMRI_EMOTION_RL) are usable.

  No planned Fix: These data are released and we do not plan to remove them from ConnectomeDB or AWS S3. We recommend excluding these runs from your analyses.

- *7/25/2019:* **Subjects 196952, 748662, 809252 Unusable tfMRI runs**

  For a variety of reasons, some fMRI runs for particular subjects were released in the S1200 data release, but should not be used. The following runs are affected and should not be used:

  196952   tfMRI_WM_LR

  748662   tfMRI_SOCIAL_RL

  809252   tfMRI_SOCIAL_RL

  Data from the remaining fMRI runs for these subjects (all resting state and other tfMRI runs) are usable.

- *7/25/2019:* **tfMRI Runs with Incomplete brain coverage for subjects 144428, 168139, 186545, 192237, 223929, 320826, 644044, 822244, 870861, and 947668**

The following released runs display incomplete brain coverage due to misplaced acquisition Field of View (FOV). We recommend exclusion of these data from analyses.

144428	*tfMRI_MOTOR_RL*
144428	*tfMRI_GAMBLING_LR*
144428	*tfMRI_MOTOR_LR*
168139	*tfMRI_EMOTION_LR*
186545	*tfMRI_SOCIAL_LR*
186545	*tfMRI_SOCIAL_RL*
186545	*tfMRI_RELATIONAL_LR*
186545	*tfMRI_RELATIONAL_RL*
186545	*tfMRI_EMOTION_LR*
186545	*tfMRI_EMOTION_RL*
192237	*tfMRI_MOTOR_RL*
223929	*tfMRI_RELATIONAL_LR*
320826	*tfMRI_EMOTION_RL*
644044	*tfMRI_RELATIONAL_LR*
644044	*tfMRI_RELATIONAL_RL*
644044	*tfMRI_EMOTION_LR*
644044	*tfMRI_EMOTION_RL*
822244	*tfMRI_GAMBLING_LR*
822244	*tfMRI_MOTOR_RL*
870861	*tfMRI_GAMBLING_LR*
870861	*tfMRI_MOTOR_LR*
870861	*tfMRI_MOTOR_RL*
947668	*tfMRI_GAMBLING_LR*
947668	*tfMRI_GAMBLING_RL*
947668	*tfMRI_MOTOR_RL*

Data from the remaining fMRI runs for these subjects (all resting state and other tfMRI runs) are usable.

- **ASR_{Syndrome}_Pct scores are T-scores (not Age/Gender adjusted Percentiles)**

Age and Gender-adjusted T-scores for the Achenbach Self Report behavioral survey have been mislabeled/misdefined as Age and Gender adjusted Percentile scores as variable names and in Data Dictionary descriptions since these scores were released for the 500 Subjects data release. The affected scores are:

Current	Correct
ASR_Anxd_Pct	*ASR_Anxd_T*
ASR_Witd_Pct	*ASR_Witd_T*
ASR_Soma_Pct	*ASR_Soma_T*
ASR_Thot_Pct	*ASR_Thot_T*

ASR_Attn_Pct	ASR_Attn_T
ASR_Aggr_Pct	ASR_Aggr_T
ASR_Rule_Pct	ASR_Rule_T
ASR_Intr_Pct	ASR_Intr_T
DSM_Depr_Pct	DSM_Depr_T
DSM_Anxi_Pct	DSM_Anxi_T
DSM_Somp_Pct	DSM_Somp_T
DSM_Avoid_Pct	DSM_Avoid_T
DSM_Adh_Pct	DSM_Adh_T
DSM_Antis_Pct	DSM_Antis_T

Completed Fix: Labels and definitions of T-Scores for the ASR Syndrome and DSM-Oriented Scales were updated with the rerelease of the 7T HCP functional data in April 2018.

- **"Noise_Comp" NIH Toolbox Words in Noise variable data incomplete in ConnectomeDB**

  Between HCP data collection for subjects released in the 500 Subjects Release and 900 Subjects Release, NIH Toolbox changed how the Words In Noise (WIN) audition test was administered to subjects (through computer speakers in v1 to through earphones in v2). Due to how our internal database was recording the scores for WIN v1 and WIN v2 separately, only the v1 scores have been correctly released in ConnectomeDB under the "Noise_Comp" variable, the WIN Computed score. Subjects tested with WIN v2 (most subjects newly released in S900 and S1200 releases) erroneously have either no score listed or a score of -99. NIH Toolbox has verified that the v1 and v2 WIN scoring has been normed to be directly comparable, so we plan to combine the v1/v2 scores for release as the single "Noise_Comp" variable.

  Completed Fix: Words In Noise v2 scores were added to the "Noise_Comp" variable released with the rerelease of the 7T HCP functional data in April 2018.

- **Some EMOTION variables were swapped in ConnectomeDB "In-scanner performance data"**

  For many previously-released participants (specifically, subjects initially released up to and including the "500 Subjects Release"), some "In-scanner performance" EMOTION variables were swapped in behavioral data .csv files downloaded from ConnectomeDB (either from the quick downloads section of the S1200 project page or from the subject dashboard). Specifically, the Emotion_Task_Acc and Emotion_Task_Face_Acc variables were swapped, and the Emotion_Task_Median_RT and Emotion_Task_Face_Median_RT variables were swapped. The following data were not affected: .csv files for individual scan runs (obtained via Connectome-in-a-Box or downloaded .zip packages), and values for participants that were first released in the 900 Subjects Release and subsequent releases.

  Completed Fix: The EMOTION variables in ConnectomeDB .csv files were corrected as of July 28, 2017. The EMOTION task accuracy and reaction time variables for all affected participants are correct in spreadsheets downloaded after this date.

- **7T Preprocessed fMRI AP run data (Resting State, Retinotopy, and Movie-watching Task) unwarped using the incorrect direction in the fMRIVolume pipeline for all 7T Subjects**

  A bug in the 7T fMRIVolume pipeline caused all Anterior->Posterior (AP phase-encoding direction) runs in the 7T fMRI data on all subjects to be unwarped using the incorrect direction, causing registration and all subsequent processing results to be incorrect. Since the unwarping is a very early processing step, all AP run 7T fMRI data is affected, including all preprocessed and FIX-cleaned resting state, retinotopy, and movie-watching data. All 7T fMRI data should not be used. Therefore, we have made this data unavailable for download in ConnectomeDB until fixed by reprocessing. Unprocessed 7T fMRI and all 7T diffusion MRI data is unaffected by this issue and is still available for download.

  Completed Fix: All 7T fMRI data was reprocessed with the bug fixed and rereleased in April 2018.

- **Image reconstruction algorithm versions**

  Two versions of the image reconstruction algorithm were used for the fMRI data (both task and resting-state) over the course of the Young-Adult HCP: version "r177" for subjects scanned early (Q1 through mid-Q3) and version "r227" for subjects scanned after that (mid-Q3 onward). The reconstruction version has a notable signature on the fMRI data that can make a large difference in fMRI data analysis. Users wanting to use only data with the same reconstruction algorithm in their analyses can filter data on the ConnectomeDB dashboard by fMRI reconstruction version (in the Study Completion category, "fMRI_3T_ReconVrs"). See Ramifications of Image Reconstruction Version Differences for more details. On ConnectomeDB, we also provide separate Resources for Group Average Functional Connectivity ("dense connectomes") and Parcellation+Timeseries+Netmats ("PTN") files based on both all subjects ("r177+r227") and just the "r227" subjects.

- **Other identified QC Issues**

  With 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. 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. More details can be found here: HCP Subjects with Identified Quality Control Issues (QC_Issue measure codes explained)

- **Subjects without Field Maps for Structural scans**

  Nineteen subjects from the S1200 Release (including two originally released in S500+MEG2 and eleven originally released in the S900 Release) do not have gradient-recalled echo (GRE) field maps accompanying their structural (T1 and T2) scans.

Affected Subjects:

102614
111009
111514
115017
121416
130821
138332
179952
299760
300618
392750
406432
429040
633847
662551
679770
688569
693461
815247

The field maps for these subjects were not collected or are unusable. Field maps are not required for HCP structural processing (in contrast to fMRI processing, where field maps are required). However, the lack of field maps prevents us from performing readout distortion correction on the structurals and will have these two effects:

- ☐ *Some minor differences in the susceptibility regions (superior orbitofrontal and inferior temporal gyrus regions) of myelin maps produced by the HCP Structural pipeline.*

- ☐ *Tissue volumes in these regions may also be slightly different.*

Issue Cannot be Fixed

Minor Issues & Updates

- **Retest 341834 rfMRI_REST1_RL missing FIX results**

  FIX processing for the rfMRI_REST1_RL RETEST run of subject 341834 was never successfully completed so this data is not available in the release.

  No planned Fix. Due to other priorities, we will not be generating this data. Users needing this data can run the HCP FIX pipeline on the functional processing results.

- **Bug affecting lh.curv.pial and rh.curv.pial in Structural Extended Packages**

  In version 3.4.1 of the HCPpipelines (used to process the structural data for HCP-YA S1200 release), there was a bug in the FreeSurferH iResPial.sh script that caused the `lh.curv.pial` and `rh.curv.pial` files (curvature of the pial surface) to be written out incorrectly. Therefore those two files ***should not be used***.

  Instead, use the **{lh,rh}.curv.pial.T2.two** files that are also contained in the `{Subject_ID}/T1w/{Subject_ID}/surf` directory of the "Structural Extended Packages"

  No planned Fix. Due to the minor nature of the bug, that it can be easily worked around, and the age of ConnectomeDB, we do not plan to fix this in the released data.

- ***Missing tfMRI_WM_RL preprocessing files for Subject 668361***

  The preprocessing results for the tfMRI_WM_RL run are incomplete for subject 668361, despite our collecting complete tfMRI_WM_RL data for this subject.

  These directories/files are missing from the 668361/MNINonLinear/Results/tfMRI_WM_RL directory:

  /EVs
  /RibbonVolumeToSurfaceMapping
  Movement_*.txt files
  *TAB.txt files
  brainmask_fs.2.nii.gz
  tfMRI_WM_LR.nii.gz
  tfMRI_WM_LR_Atlas.dtseries.nii
  tfMRI_WM_LR_Jacobian.nii.gz
  tfMRI_WM_LR_SBRef.nii.gz
  tfMRI_WM_LR_hp200_s4_level1.fsf

  Planned Fix: The tfMRI_WM_RL run will be reprocessed and the missing files regenerated to update the 668361 WM tfMRI preproc package and the 668361 data in AWS S3.

- ****TAB.txt files missing in preprocessed task fMRI data and *Physio_log.txt files missing for ALL preprocessed rfMRI and tfMRI data for 16 subjects***

Due to a late stage error in the running of the functional preprocessing pipeline for some subjects newly released with the S1200, *TAB.txt and *Physio_log.txt files are missing in the {Subject_ID}/MNINonLinear/Results/tfMRI_{TASK}_{RUN} directories for ALL tasks completed and the *Physio_log.txt files are missing in the {Subject_ID}/MNINonLinear/Results/rfMRI_{REST?}_{RUN} directory.

The missing files do exist for these subjects in {Subject_ID}/unprocessed/3T/tfMRI_{TASK}_{RUN}/LINKED_DATA/EPRIME, {Subject_ID}/unprocessed/3T/tfMRI_{TASK}_{RUN}/LINKED_DATA/PHYSIO, and {Subject_ID}/rfMRI_{REST?}_{RUN}/LINKED_DATA/PHYSIO directories, so there was a problem with the transfer to the preprocessed directories.

Affected subjects:

116423
120010
121719
127226
130114
143830
202820
385046
401422
462139
469961
644246
723141
943862
969476
971160

Because this issue is isolated to subjects newly released with S1200, this DOES NOT affect earlier release data distributed by HCP.

Planned Fix: The missing *TAB.txt and *Physio_log.txt files will be transferred to the appropriate preprocessing {Subject_ID} /MNINonLinear/Results directories, the affected rfMRI and tfMRI preprocessing packages will be updated in ConnectomeDB, and AWS S3 will be updated.

Initial 7T fMRI Release (June 20, 2016)

Major Issues

- ***7T Preprocessed fMRI AP run data (Resting State, Retinotopy, and Movie-watching Task) unwarped using the incorrect direction in the fMRIVolume pipeline for all 7T Subjects***

    A bug in the 7T fMRIVolume pipeline caused all Anterior->Posterior (AP phase-encoding direction) runs in the 7T fMRI data on all subjects to be unwarped using the incorrect direction, causing registration and all subsequent processing results to be incorrect. Since the unwarping is a very early processing step, all AP run 7T fMRI data is affected, including all preprocessed and FIX-cleaned resting state, retinotopy, and movie-watching data. <u>All 7T fMRI data should not be used.</u> Therefore, we have made this data unavailable for download in ConnectomeDB until fixed by reprocessing. Unprocessed 7T fMRI and all 7T diffusion MRI data is unaffected by this issue and is still available for download.

    Completed Fix: All 7T fMRI data was reprocessed with the bug fixed and rereleased in April 2018.

900 Subjects Release (December 8, 2015)

Major Issues

- ***ASR_{Syndrome}_Pct scores are T-scores (not Age/Gender adjusted Percentiles)***

    Age and Gender-adjusted T-scores for the Achenbach Self Report behavioral survey have been mislabeled/misdefined as Age and Gender adjusted Percentile scores as variable names and in Data Dictionary descriptions since these scores were released for the 500 Subjects data release. The affected scores are:

Current	Correct
ASR_Anxd_Pct	ASR_Anxd_T
ASR_Witd_Pct	ASR_Witd_T
ASR_Soma_Pct	ASR_Soma_T
ASR_Thot_Pct	ASR_Thot_T
ASR_Attn_Pct	ASR_Attn_T
ASR_Aggr_Pct	ASR_Aggr_T

ASR_Rule_Pct	ASR_Rule_T
ASR_Intr_Pct	ASR_Intr_T
DSM_Depr_Pct	DSM_Depr_T
DSM_Anxi_Pct	DSM_Anxi_T
DSM_Somp_Pct	DSM_Somp_T
DSM_Avoid_Pct	DSM_Avoid_T
DSM_Adh_Pct	DSM_Adh_T
DSM_Antis_Pct	DSM_Antis_T

Completed Fix: Labels and definitions of T-Scores for the ASR Syndrome and DSM-Oriented Scales were updated with the rerelease of the 7T HCP functional data in April 2018.

- **"Noise_Comp" NIH Toolbox Words in Noise variable data incomplete in ConnectomeDB**

   Between HCP data collection for subjects released in the 500 Subjects Release and 900 Subjects Release, NIH Toolbox changed how the Words In Noise (WIN) audition test was administered to subjects (through computer speakers in v1 to through earphones in v2). Due to how our internal database was recording the scores for WIN v1 and WIN v2 separately, only the v1 scores have been correctly released in ConnectomeDB under the "Noise_Comp" variable, the WIN Computed score. Subjects tested with WIN v2 (most subjects newly released in S900 and S1200 releases) erroneously have either no score listed or a score of -99. NIH Toolbox has verified that the v1 and v2 WIN scoring has been normed to be directly comparable, so we plan to combine the v1/v2 scores for release as the single "Noise_Comp" variable.

   Completed Fix: Words In Noise v2 scores were added to the "Noise_Comp" variable released with the rerelease of the 7T HCP functional data in April 2018.

- **Some EMOTION variables were swapped in ConnectomeDB "In-scanner performance data"**

   For many previously-released participants (specifically, subjects initially released up to and including the "500 Subjects Release"), some "In-scanner performance" EMOTION variables were swapped in behavioral data .csv files downloaded from ConnectomeDB (either from the quick downloads section of the S900 project page or from the subject dashboard). Specifically, the Emotion_Task_Acc and Emotion_Task_Face_Acc variables were swapped, and the Emotion_Task_Median_RT and Emotion_Task_Face_Median_RT variables were swapped. The following data were not affected: .csv files for individual scan runs (in Connectome-in-a-Box or downloaded .zip packages), and values for participants that were first released in the 900 Subjects Release and 1200 Subjects Release.

   Completed Fix: The EMOTION variables in ConnectomeDB .csv files were corrected as of July 28, 2017. The EMOTION task accuracy and reaction time variables for all affected participants are correct in spreadsheets downloaded after this date.

- **Diffusion data removed for Subjects 130922 and 150423 [RESOLVED with S1200 Release]**

   Two subjects had some or all previously released diffusion scans which were deemed to be unusable and removed from the released data with the S1200 Release. For subject 130922, originally part of the Q3 release, the DWI_dir95_RL diffusion scan (only) was determined to be unusable. For subject150423, originally part of the Q1 release, only two diffusion scans were completed and that is not enough for useful analysis, therefore all diffusion data was removed.

   Completed Fix: As of the S1200 Release, the DWI_dir95_RL diffusion scan for subject 130922 and all diffusion data for subject 150423 were removed from the released data. The affected preprocessed dMRI data for these subjects from previous releases should not be used (whether obtained via Connectome in a Box or downloaded online).

- **Image reconstruction algorithm versions**

   Two versions of the image reconstruction algorithm applied to dMRI and fMRI data have been used in HCP to date: version r177 for subjects scanned in Q1 through mid-Q3, version r227 for subjects scanned mid-Q3 and after. We were able to retroactively recon all dMRI data using the newer r227 version. However, for the fMRI data, this was not possible and data reconstructed with both r177 and r227 versions remain in ConnectomeDB. The reconstruction version makes a notable signature on the data that can make a large difference in fMRI data analysis. Users wanting to use only data with the same reconstruction algorithm applied in their analyses can filter data on the ConnectomeDB dashboard by fMRI reconstruction version (in the Study Completion category). See Ramifications of Image Reconstruction Version Differences for more details.

- **Intermittent orbitofrontal artifact affecting a limited number of rfMRI and tfMRI scans [CAPTURED as part of the "QC_Issues" code with the S1200 Release]**

   An artifact (characterized by a striped spatial pattern in image intensity around orbitofrontal susceptibility areas) caused by intermittent technical problems with the 32 channel head coil has been found in a limited number of fMRI scan session (24 rfMRI scans in 18 subjects; 6 tfMRI scans in 6 subjects) that were obtained during the summer of 2013. We have opted to release these data for public use with the expectation that investigators will review the data, before and after their chosen pre-processing steps, to decide whether this artifact might affect their data analysis. See Intermittent orbitofrontal artifact affecting some fMRI scans for details, including a list of affected subjects /scans and example images.

Minor Issues & Updates

- ***Missing tfMRI_WM_RL preprocessing files for Subject 668361***

  The preprocessing results for the tfMRI_WM_RL run are incomplete for subject 668361, despite our collecting complete tfMRI_WM_RL data for this subject.
  These directories/files are missing from the 668361/MNINonLinear/Results/tfMRI_WM_RL directory:

  /EVs
  /RibbonVolumeToSurfaceMapping
  Movement_*.txt files
  *TAB.txt files
  brainmask_fs.2.nii.gz
  tfMRI_WM_LR.nii.gz
  tfMRI_WM_LR_Atlas.dtseries.nii
  tfMRI_WM_LR_Jacobian.nii.gz
  tfMRI_WM_LR_SBRef.nii.gz
  tfMRI_WM_LR_hp200_s4_level1.fsf

  Planned Fix: The tfMRI_WM_RL run will be reprocessed and the missing files regenerated to update the 668361 WM tfMRI preproc package and the 668361 data in AWS S3.

- ***Imaging data Released for Subjects 121719 & 169040 (Originally released in S900 as Behavior only) [RESOLVED with S1200 Release]***

  The T1w and T2w structural data quality was reassessed and subjects 121719 & 169040 were deemed of sufficient quality to release all available imaging data with the S1200 Release.

  Completed Fix: All available imaging data for subjects 121719 & 169040 was released with the S1200 Release.

- ***Subjects without Field Maps for Structural scans***

  Thirteen subjects from the S900 Release (including two originally released in the S500+MEG2 Release) do not have gradient-recalled echo (GRE) field maps accompanying their structural (T1 and T2) scans.

  Affected Subjects:

  111009
  111514
  115017
  121416
  130821
  179952
  300618
  392750
  406432
  429040
  633847
  662551
  679770
  693461

  The field maps for these subjects were not collected or are unusable. Field maps are not required for HCP *structural* processing (in contrast to fMRI processing, where field maps *are* required). However, the lack of field maps prevents us from performing readout distortion correction on the structurals and will have these two effects:

  - Some minor differences in the susceptibility regions (superior orbitofrontal and inferior temporal gyrus regions) of myelin maps produced by the HCP Structural pipeline.
  - Tissue volumes in these regions may also be slightly different.

  We anticipate that a limited number of subjects that will be included in upcoming HCP releases will also have this issue. Those subjects will be identified on this Issues wiki in the appropriate release section.

500 Subjects Release (June 6, 2014) + MEG2 (November 26, 2014)

Major Issues

- ***ASR_{Syndrome}_Pct scores are T-scores (not Age/Gender adjusted Percentiles)***

  Age and Gender-adjusted T-scores for the Achenbach Self Report behavioral survey have been mislabeled/misdefined as Age and Gender adjusted Percentile scores as variable names and in Data Dictionary descriptions since these scores were released for the 500 Subjects data release. The affected scores are:

Current	Correct
ASR_Anxd_Pct	ASR_Anxd_T
ASR_Witd_Pct	ASR_Witd_T

ASR_Soma_Pct	ASR_Soma_T
ASR_Thot_Pct	ASR_Thot_T
ASR_Attn_Pct	ASR_Attn_T
ASR_Aggr_Pct	ASR_Aggr_T
ASR_Rule_Pct	ASR_Rule_T
ASR_Intr_Pct	ASR_Intr_T
DSM_Depr_Pct	DSM_Depr_T
DSM_Anxi_Pct	DSM_Anxi_T
DSM_Somp_Pct	DSM_Somp_T
DSM_Avoid_Pct	DSM_Avoid_T
DSM_Adh_Pct	DSM_Adh_T
DSM_Antis_Pct	DSM_Antis_T

Completed Fix: Labels and definitions of T-Scores for the ASR Syndrome and DSM-Oriented Scales were updated with the rerelease of the 7T HCP functional data in April 2018.

- **Some EMOTION variables were swapped in ConnectomeDB "In-scanner performance data"**

  For many previously-released participants (specifically, subjects initially released up to and including the "500 Subjects Release"), some "In-scanner performance" EMOTION variables were swapped in behavioral data .csv files downloaded from ConnectomeDB (either from the quick downloads section of the S500 project page or from the subject dashboard). Specifically the Emotion_Task_Acc and Emotion_Task_Face_Acc variables were swapped, and the Emotion_Task_Median_RT and Emotion_Task_Face_Median_RT variables were swapped. The following data were not affected: .csv files for individual scan runs (in Connectome-in-a-Box or downloaded .zip packages), and values for participants that were first released in the 900 Subjects Release and 1200 Subjects Release. (**)

  (**) Technically, the small set of 15 subjects that were first released in the "S500+MEG2" Release were also not affected by this problem.

  Completed Fix: The EMOTION variables in ConnectomeDB .csv files were corrected as of July 28, 2017. The EMOTION task accuracy and reaction time variables for all participants are correct in spreadsheets downloaded after this date.

- **Intermittent orbitofrontal artifact affecting a limited number of rfMRI and tfMRI scans**

  An artifact (characterized by a striped spatial pattern in image intensity around orbitofrontal susceptibility areas) caused by intermittent technical problems with the 32 channel head coil has been found in a limited number of fMRI scan session (24 rfMRI scans in 18 subjects; 6 tfMRI scans in 6 subjects) that were obtained during the summer of 2013. We have opted to release these data for public use with the expectation that investigators will review the data, before and after their chosen pre-processing steps, to decide whether this artifact might affect their data analysis. See Intermittent orbitofrontal artifact affecting some fMRI scans for details, including a list of affected subjects /scans and example images.

- **Image reconstruction algorithm versions**

  Two versions of the image reconstruction algorithm applied to dMRI and fMRI data have been used in HCP to date: version r177 for subjects scanned in Q1 through mid-Q3, version r227 for subjects scanned mid-Q3 and after. We were able to retroactively recon all dMRI data using the newer r227 version. However, for the fMRI data, this was not possible and data reconstructed with both r177 and r227 versions remain in ConnectomeDB. The reconstruction version makes a notable signature on the data that can make a large difference in fMRI data analysis. Users wanting to use only data with the same reconstruction algorithm applied in their analyses can filter data on the ConnectomeDB dashboard by fMRI reconstruction version (in the Study Completion category). See Ramifications of Image Reconstruction Version Differences for more details.

- **Physiological Log file timing errors for all data released in S500 and earlier [RESOLVED with S900 Release]**

  In preparation for the HCP 900 subjects data release, we identified previously unknown timing errors in the in the processed physiological log files (e.g., rfMRI_REST1_RL_Physio_log.txt). These errors have resulted in variable offsets in timing, sometimes amounting to hundreds of milliseconds. The timing errors affected rsfMRI, tfMRI, and pulse oximetry data that was collected during diffusion MRI scans. Due to limitations in computational resources, it was not possible to correct the physio files in the packages being released. Therefore, the physio log files in those packages should not be used.

  Completed Fix: The corrected processed physiological log files were released for all subjects with the 900 Subjects Release in December 2015.

- **Subject 142626 removed from ConnectomeDB [RESOLVED with S900 Release]**

  Subject 142626, released in the 500 Subjects Release (June 2014), has been found to have the same identity as another subject in the HCP study. Thus, we have removed all data for subject 142626 from ConnectomeDB. For any ongoing analyses, we recommend that if possible you exclude subject 142626 from your analyses.

Note: we have not removed the duplicate subject's dataset, which is comparable in quality to other released subjects, from ConnectomeDB or from ongoing HCP analyses. However, for reasons of confidentiality, we are electing to not reveal the subject number for the duplicate subject. If the subject number of the duplicate subject is found by other investigators using HCP data, we request that you to keep this information confidential.

Subject 142626 was originally released in the 500 Subjects Release and have been available in ConnectomeDB and through Connectome in a Box (CinaB) and Amazon S3 since then.

Completed Fix: All data from subject 142626 has been removed from ConnectomeDB, CinaB, and Amazon S3 as of November 6, 2015 and will no longer be distributed by HCP.

- **Diffusion data for subject 103515 corrupted [RESOLVED with S900 Release]**

  A problem with the middle stages of the v3.1 (500 Subjects) preprocessing of diffusion data released for subject 103515 has been found. The file size for the Diffusion data for this subject is only about 400Mb, which is much smaller than the normal size, and should not be used. (In particular, while the NIFTI header indicates the correct dimensions, frames 83 onward are corrupted in that they contain a single intensity throughout the brain).

  Completed Fix: Diffusion data for subject 103515 was completely reprocessed and the fixed version was available as part of the 900 Subjects Release. The preprocessed dMRI for this subject from the 500 Subjects release should not be used (whether obtained via Connectome in a Box or downloaded online). The reprocessed files will be included in 900 Subjects Connectome in a Box versions and in the 900 Subjects update to the HCP Amazon S3 bucket.

- **SOCIAL_LR tfMRI data for subject 748662 corrupted [RESOLVED with S900 Release]**

  A significant amount of head motion in the SBRef scan collected for the SOCIAL_LR task for S500 subject 748662 corrupted data for the entire scan. Therefore, the S500 released data for SOCIAL_LR of 748662 (including the unprocessed, minimally preprocessed, and level 2, cross run analyzed data) <u>should not be used</u>.
  We are in the process of removing the corrupt data from ConnectomeDB. SOCIAL tfMRI data for this subject downloaded from June 6, 2014-October 31, 2014 (or distributed in Connectome in a Box drives dated 20140805 or earlier) will contain the corrupted data and should not be used.

  Completed Fix: The corrupt SOCIAL_LR tfMRI data for subject 748662 has been permanently removed from ConnectomeDB as of 10/31/2014. These files need to be removed from 500 Subjects Connectome in a Box dated 20140805 or earlier. These files have been removed in 500 Subjects Connectome in a Box versions dated 20141203 and later.

- **Diffusion data for subject 173132 corrupted [RESOLVED with S900 Release]**

  A problem with the early stages of the v3.1 (500 Subjects) preprocessing of diffusion data released for subject 173132 has been found. Diffusion data for this subject downloaded from June 6, 2014-October 31, 2014 (or distributed in Connectome in a Box drives dated 20140805 or earlier) should not be used.

  Completed Fix: Diffusion data for subject 173132 has been completely reprocessed and the fixed version is available in ConnectomeDB as of 10/31/14. These files will need to be replaced for 500 Subjects Connectome in a Box dated 20140805 or earlier. The reprocessed files are included in 500 Subjects Connectome in a Box versions dated 20141203 and later.

Minor Issues and Updates

- **Subjects without Field Maps for Structural scans: 111514 & 662551**
  Two subjects from the 500 Subjects + MEG2 release were released without gradient-recalled echo (GRE) field maps accompanying their structural (T1 and T2) scans. The field maps for these subjects were not collected or are unusable. Field maps are not required for HCP *structural* processing (in contrast to fMRI processing, where field maps *are* required). However, the lack of field maps prevents us from performing readout distortion correction on the structurals and will have these two effects:
    - Some minor differences in the susceptibility regions (superior orbitofrontal and inferior temporal gyrus regions) of myelin maps produced by the HCP Structural pipeline.
    - Tissue volumes in these regions may also be slightly different.
  We anticipate that a limited number of subjects that will be included in upcoming HCP releases will also have this issue. Those subjects will be identified on this Issues wiki in the appropriate release section.

- **Co-Twins with different exact ages or age ranges**
  In the 500 Subjects Release there are **4** sets of twins who do not have the same exact ages in ConnectomeDB. Exact age is restricted data, but it is also used in binning subjects into age ranges, which is open access. The differences in the exact age of these 4 twin pairs is due to the fact that some twin pairs are interviewed and scanned months apart due to their individual schedules. In these 4 pairs, one twin happened to be interviewed in the months before their birthday and the other in the months after their birthday. Because twin status and co-twin IDs are restricted data, we can not be more specific here than to say that 4 twin pairs are affected.

  For all HCP participants, we collect the "exact age" information at the time of the SSAGA phone interview that precedes the scanning visit. It should be noted that sometimes months can elapse between the SSAGA and the scan visit, however "exact age" is always the participant's age at the time of the SSAGA interview, not the scan date. Users should be aware of this as they conduct their analyses.

- **Problems using CSV files downloaded from ConnectomeDB with some platforms**
  Users have reported issues using some CSV readers/importers (including MATLAB) and user-generated code to read the comma separated variable (CSV) files downloaded from ConnectomeDB that contain Behavioral and Demographic data on HCP subjects.

The issue is that some CSV importers ignore double-quotation marks (" "), which by CSV standard practice are used to handle single fields that contain commas or carriage returns. Our CSV files contain a small number of fields that contain commas within the field. These single fields are being split into multiple fields by CSV importers/code that ignore the standard, resulting in added and misaligned data columns. This issue affects all CSVs (both Open Access and Restricted Access versions) downloaded from ConnectomeDB from Q1 (March 2013) until November 2014.

Completed Fix: As of the MEG2 release (Nov 26, 2014) the commas were removed from the few data fields that contain these in the HCP dataset to fix this issue for future downloaded CSVs.

- **WM_Stats.csv files for subject 154835**
  A bug produced a short versions of 154835_WM_LR_Stats.csv and 154835_WM_RL_Stats.csv that were in ConnectomeDB from June 6, 2014 - August 13,2014. This also affects Connectome in a Box drives dated 20140805 or earlier.

  Completed Fix: We have replaced the short files with correct files of the same name. Users should re-download and replace the WM task unprocessed and preprocessed packages for this subject in order to obtain the corrected files. These files need to be replaced for 500 Subjects Connectome in a Box dated 20140805 or earlier.

- **prepare_level2_feat_analysis.sh script not included in tfMRI packages**
  The prepare_level2_feat_analysis.sh helper bash script for setting up the registration matrices/directory prior to running level 2 feat analysis was inadvertently omitted from all the minimally preprocessed data packages for the tfMRI data in the 500 Subjects release: e.g. ${StudyFolder}/${SubjectID}/MNINonLinear/Results/tfMRI_SOCIAL/prepare_level2_feat_analysis.sh

  Completed no-fix alternate: Instead of distributing this "legacy" script with the preprocessed tfMRI packages, we are distributing the script along with tfMRI analysis advice via the links below:

  Users who would like this script for current use can download it by clicking this link: prepare_level2_feat_analysis.sh

  Please also see the **Advice for FEAT Analysis of HCP task fMRI data** for more information on performing tfMRI analysis on HCP data.

  Note: Users can also download task analysis data packages containing the results of individual subject, cross-run ("level 2") fMRI grayordinates analysis in FSL for each HCP task at various smoothing levels. Analogous "level2" data packages for volume-based tfMRI processing are available at 4mm smoothing. Please see the HCP 500 Subjects Data Release Reference Manual for more details.

- **LANGUAGE_Stats.csv files for subjects 901442 and 713239**
  A bug produced blank versions of 901442_LANGUAGE_RL_Stats.csv and 713239_LANGUAGE_LR_Stats.csv that were in ConnectomeDB from June 6, 2014 - July 29,2014. This also affects Connectome in a Box drives dated 20140805 or earlier.

  Completed Fix: We have replaced the blank files with correct files of the same name. Users should re-download and replace the LANGUAGE task unprocessed and preprocessed packages for these two subjects in order to obtain the corrected files. These files need to be replaced for 500 Subjects Connectome in a Box dated 20140805 or earlier.

- **Color Vision/Eye updates: subjects** 125525, 126931, 140925, 154936, 180129, 211417, 355542, 390645, 424229, 586460, 599671, 644044, 729254, 733548
  The Color Vision and/or Eye tested restricted data for the following subjects was incorrect from June 6, 2014 - July 11, 2014. Restricted access users should check family structure fields in ConnectomeDB for more details.

  125525
  126931
  140925
  154936
  180129
  211417
  355542
  390645
  424229
  586460
  599671
  644044
  729254
  733548

  Completed Fix: We have updated the Color Vision/Eye tested information for these subjects in ConnectomeDB as of June 12, 2014. Restricted Access users that downloaded the restricted dataprior to that date should update their restricted data accordingly.

- **Race/Ethnicity update for subject 163836**
  The Race and Ethnicity restricted data for subject 163836 was incorrect from June 6, 2014 - July 10, 2014. Restricted access users should check these fields in ConnectomeDB for more details.

  Completed Fix: We have updated the race and ethnicity information for this subject in ConnectomeDB as of July 11, 2014. Restricted Access users that include this subject in their analyses will need to check the demographic data in DB and update downloaded restricted data accordingly.

- **Family Structure update for subject 256540**
  Some of the Restricted family structure data for subject 256540 was incorrect from June 6, 2014 - July 2, 2014. Restricted access users should check family structure fields in ConnectomeDB for more details.

Completed Fix: We have updated the family structure information for this subject in ConnectomeDB as of July 3, 2014. Restricted Access users that include this subject in their analyses will need to check the demographic data in DB and update downloaded restricted data accordingly.

- *Diffusion data reprocessed for subject 125525*
  A problem with the v3.1 (500 Subjects) preprocessing of diffusion data released for subject 125525 has been found. Diffusion data for this subject downloaded from June 6-June 30, 2014 should not be used. Diffusion data for this subject released in Q1-Q3 should not be affected by this problem.

  Completed Fix: Diffusion data for subject 125525 has been reprocessed and is now correct in ConnectomeDB as of July 1, 2014. Users should re-download the preprocessed dMRI data for this subject. 500 Subjects Connectome in a Box contains the updated data.

- *Days since last Menstruation data updates: subjects 547046, 135932, 510326, 422632 & 116322*
  The Days since last Menstruation restricted data for subjects 547046, 135932, 510326, 422632 & 116322 was incorrect from June 6, 2014 - June 11, 2014. Restricted access users should check this field in ConnectomeDB for more details.

  Completed Fix: We have updated the Days since last menstruation information for these subjects in ConnectomeDB as of June 12, 2014. Restricted Access users that downloaded the restricted data during this time period should update their restricted data accordingly.

- *Twin status/zygosity data updates for subjects 161630 & 163836*
  The Twin status and zygosity restricted data for subjects 161630 & 163836 was incorrect from June 6, 2014 - June 11, 2014. Restricted access users should check these fields in ConnectomeDB for more details.

  Completed Fix: We have updated the twin status/zygosity information for these subjects in ConnectomeDB as of June 12, 2014. Restricted Access users that downloaded the restricted data during this time period should update their restricted data accordingly.

Q3 Release (September 23, 2013; more dMRI on November 26, 2013, all Q1-Q3 diffusion available January 31, 2014)

Major Issues

- *Image reconstruction algorithm versions*

  Two versions of the image reconstruction algorithm applied to dMRI and fMRI data have been used in HCP to date: version r177 for subjects scanned in Q1 through mid-Q3, version r227 for subjects scanned mid-Q3 and after. We were able to retroactively recon all dMRI data using the newer r227 version. However, for the fMRI data, this was not possible and data reconstructed with both r177 and r227 versions remain in ConnectomeDB. The reconstruction version makes a notable signature on the data that can make a large difference in fMRI data analysis. Users wanting to use only data with the same reconstruction algorithm applied in their analyses can filter data on the ConnectomeDB dashboard by fMRI reconstruction version (in the Study Completion category). See Ramifications of Image Reconstruction Version Differences for more details.

- *Subjects 209733 and 528446 removed from ConnectomeDB*

  Second-level QC on all subjects has revealed two Q2 subjects with structural brain abnormalities, specifically gray matter heterotopia. The abnormalities significantly affect functional and structural connectivity in these subjects, beyond what is considered normal variation. Therefore, we recommend that data from subjects 209733 and 528446 NOT be used in analyses.

  Subjects 209733 and 528446 were originally released in Q2 and have been available in ConnectomeDB and through Connectome in a Box since then. Both subjects were included in the Unrelated 40, Unrelated 80, and Related 120 groups until the 500 subject release (June 6, 2014).

  Completed Fix: All data from subjects 209733 and 528446 have been removed from ConnectomeDB as of the 500 subject release and will no longer be distributed by HCP.

- *T1w_acpc_dc_restore_1.25.nii.gz file not included in preprocessed diffusion MRI data packages*

  The diffusion preprocessing output file containing the structural volume sampled at the same resolution as the diffusion data (necessary for analysis) was inadvertently left out of the dMRI data packages released for Q3:
  e.g. ${StudyFolder}/${SubjectID}/T1w/T1w_acpc_dc_restore_1.25.nii.gz

  Completed Fix: The dMRI data packages have been rebuilt with this file and as of 1/28/2014 are available via download from ConnectomeDB.

  Users can also generate this file for previously downloaded or Connectome in a Box (CINAB) data, using the following FSL commands:

  #Generate 1.25mm structural space for resampling the diffusion data into

  ${FSLDIR}/bin/flirt -interp spline -in T1w_acpc_dc_restore -ref T1w_acpc_dc_restore -applyisoxfm 1.25 -out T1w_acpc_dc_restore_1.25

  ${FSLDIR}/bin/applywarp --rel --interp=spline -i T1w_acpc_dc_restore -r T1w_acpc_dc_restore_1.25 -o T1w_acpc_dc_restore_1.25

Minor Issues and Updates

- **Subjects released in Q3 (or earlier) that were not re-released in the 500 Subjects Release**
Data for a handful of subjects were released in Q3, but were not re-released in the 500 Subjects Release, because of minor issues. Some subjects did not have gradient-echo fieldmaps to accompany the structurals. The GE fieldmaps are not mandatory for running the structurals through the HCP pipelines, but we didn't get the process for handling those exceptions automated in time for the 500 subject release. The affected subjects are:

    111009
    111514
    429040
    688569

    Functional data for one subject, 727654, needs to be reconstructed with the newer reconstruction method. Because of this, functional data for this subject _should not be used_ in analyses.

    As a reminder, _do not mix preprocessed data from different data releases in your analyses_. The data in Q1, Q2+Q3, and the 500 Subjects Releases were run through different versions of the HCP pipelines, with differences significant enough that preprocessed data should not be compared across releases.

    Data for all these subjects should be part of the next "new subject" data release coming in Spring 2015.

- **Timing issue with tfMRI Working Memory EVs**
For all eight blocked Working Memory EVs, the onset of the block was set to the onset of the first trial, rather than the preceding cue as it should have been. The consequence was that the boxcar started 2.5s later and ended 2.5s later than it should have. In addition, the "Measure" variable names were slightly inconsistent with those of the other tasks.

    Completed Fix: We replaced the Working Memory EVs with new versions set to onset at the cue before the block, and last until the end of the final trial in the block. The "Measure" variable names were modified to be consistent with those of the other tasks. These changes are available as part of the March 2014 data patch.

- **Error in tfMRI Working Memory Stats files**
The accuracy for targets within individual blocks (e.g., 2BK FACE) was not correct. The accuracy reported for targets in most conditions was actually the value derived from nontarget trials. In addition, the "Measure" variable names were slightly inconsistent with those of the other tasks.

    Completed Fix: We corrected the Target accuracies in the stats files. The "Measure" variable names were modified to be consistent with those of the other tasks. These changes are available as part of the March 2014 data patch.

- **Improved formatting and additional variables in all tfMRI Stats files**

    Completed update: The following improvements to all tfMRI Stats files are available as part of the March 2014 data patch:

    - Change to csv format to make processing behavioral files simpler.
    - Consistency in variable naming and formatting across Stats files.
    - Attempt to provide additional performance metrics for the specific conditions being used in the imaging analyses.

- **Timing issue with HCP fMRI physiological monitoring data**
There was a bug in the timing of physiological monitoring data released during Q1 - Q3 is offset due to a roughly 500 ms delay in the onset of recording for the respiratory, pulse oximetry, and external trigger signal during fMRI runs. Because the onsets of the log files for each of these physiological modalities were not simultaneous as we expected them to be, the physiological data in the processed log files (e.g., tfMRI_WM_RL_Physio_log.txt) was also offset. More specifically, relative to the external trigger signal contained in the processed log files, the respiratory signal was delayed by roughly 1000 ms and the pulse-oximeter signal was delayed by roughly 1500 ms.

    Completed Fix: We re-aligned the onsets of the physiological signals with the external trigger signal in the processed log files for all Q1-Q3 physiological data that could be fixed. The realignment will also be applied to all future release subjects that are affected. The fix for the Q1-Q3 physiological data is available as part of the March 2014 data patch.

- **EMOTION_LR tfMRI data for Q3 subject 727654 has reconstruction artifact**
Images collected during the EMOTION_LR task for Q3 subject 727654 have artifacts related to the switchover in the HCP image reconstruction algorithm that was occurring at the time. Therefore the Q3 released data for EMOTION_LR of 727654 (including the unprocessed, minimally preprocessed, and level 2, cross run analyzed data) underline should not be used.

    Completed Fix: We reconstructed the images from the EMOTION_LR task for 727654, reran the preprocessing and analysis pipelines for this dataset, and have them available as of the 500 subject data release (June 2014).

- **_prepare_level2_feat_analysis.sh_ script not included in tfMRI packages**
The prepare_level2_feat_analysis.sh helper bash script for setting up the registration matrices/directory prior to running level 2 feat analysis was inadvertently omitted from all the minimally preprocessed data packages for the tfMRI data in Q3: e.g. ${StudyFolder}/${SubjectID}/MNINonLinear/Results/tfMRI_SOCIAL/prepare_level2_feat_analysis.sh

    Completed Fix: This script has been added back into the Q3 preprocessed tfMRI packages as part of the March 2014 data patch. As of the 500 Subjects release (June 2014), instead of distributing this "legacy" script with the preprocessed tfMRI packages, we are distributing the script along with tfMRI analysis advice via the links below:

    Users who would like this script for current use can download it by clicking this link: prepare_level2_feat_analysis.sh

    Please also see the Advice for FEAT Analysis of HCP task fMRI data for more information on performing tfMRI analysis on HCP data.

Note: With the Q3 and 500 Subjects release, users can also download task analysis data packages containing the results of individual subject, cross-run ("level 2") fMRI grayordinates analysis for each HCP task at various smoothing levels. Analogous "level2" data packages for volume-based tfMRI processing at 4mm smoothing were made available as of March 2014 and also were part of the 500 Subjects released. Please see the HCP Q3 Data Release Reference Manual or the HCP 500 Subjects Data Release Reference Manual for more details.

- ***missing values in tfMRI Movement_Regressors.txt for subjects 156233, 165849, and 366446***
  Some values are missing in the following files :

  ${StudyFolder}/156233/MNINonLinear/Results/tfMRI_GAMBLING_RL/Movement_Regressors.txt
  ${StudyFolder}/156233/MNINonLinear/Results/tfMRI_SOCIAL_LR/Movement_Regressors.txt
  ${StudyFolder}/165840/MNINonLinear/Results/tfMRI_WM_LR/Movement_Regressors.txt
  ${StudyFolder}/366446/MNINonLinear/Results/tfMRI_WM_RL/Movement_Regressors.txt

  Completed Fix: These files are corrected and released as part of the March 2014 data patch.

Q2 Release (June 14, 2013)

including Q1 + Q2 v2 preprocessed data

Users who have the Q1+Q2 data: A zipped archive is available containing all of the fixed files for all Q1+Q2 subjects and a python script for patching existing download or Connectome in a box data, which places the fixed files in the HCP directory structure (details in the HCP Q2 Data Release Reference Manual and the associated directory structure Appendix III). Download the Q1/Q2 data patch script here (requires ConnectomeDB login).

Major Issues

- ***Subjects 209733 and 528446 removed from ConnectomeDB***

  Second-level QC on all subjects has revealed two Q2 subjects with structural brain abnormalities, specifically gray matter heterotopia. The abnormalities significantly affect functional and structural connectivity in these subjects, beyond what is considered normal variation. Therefore, we recommend that data from subjects 209733 and 528446 NOT be used in analyses.

  Subjects 209733 and 528446 were originally released in Q2 and have been available in ConnectomeDB and through Connectome in a Box since then. Both subjects were included in the Unrelated 40, Unrelated 80, and Related 120 groups until the 500 subject release (June 6, 2014).

  Completed Fix: All data from subjects 209733 and 528446 were removed from ConnectomeDB as of the 500 subject release and will no longer be distributed by HCP.

- ***DICOM to NIFTI failure (incorrect number of volumes/lines) in tfMRI Relational data***

  For one Q1+Q2 subject, DICOM to NIFTI failed to output correct number of volumes or lines in some tfMRI Relational data files:
   ${StudyFolder}/138231/MNINonLinear/Results/tfMRI_RELATIONAL_LR/*

Subject/scan affected:	file type	Issue:
138231/tfMRI_RELATIONAL_LR	4D volume	volumes should be 232 instead of 106
138231/tfMRI_RELATIONAL_LR	all movement regressor files	lines should be 232 instead of 106
138231/tfMRI_RELATIONAL_LR	Physio_log.txt	triggers should be 232 instead of 105

  Surface files for this subject will also be affected.

  Completed Fix: DICOM2NIFTI was rerun and files are fixed as of the Q3 release (September 2013).

  Users should either: 1) replace unprocessed and processed versions Q1+Q2 Relational data for subject 138231 by re-downloading the Relational packages (both unproc and preproc) for subject 138231, or 2) download and run the Q1/Q2 release patch script available here (requires ConnectomeDB login).

- ***tfMRI Social EV *.txt files incomplete***

  For some Q1+Q2 subjects, EV files for the Social tfMRI task were missing an entry for one or more blocks, and thus those EV files are inaccurate and incomplete. Specifically, there was a bug in the EV creation script used for the Q1 and Q2 releases for the SOCIAL task such that blocks were omitted from the EV file if the subject didn't respond during that block.

  e.g. ${StudyFolder}/${SubjectID}/MNINonLinear/Results/tfMRI_SOCIAL_RL/EVs/*.txt

  Subjects/affected directories with this bug: Q1+Q2 Subjects with Shortened tfMRI Social EV *.txt Files

  Completed Fix: These bugs have resolved for the Q3 release (September 2013).

Users should run the Q1/Q2 data patch script to correct all these incomplete files (download available here) (requires ConnectomeDB login).

Minor Issues

- ***task fMRI .fsf files parameters incorrect, or .fsf file missing***
  We ran a script checking the parameters of the tfMRI .fsf files from the Q2 release and found the following bugs for the indicated subjects:
  e.g. ${StudyFolder}/${SubjectID}/MNINonLinear/Results/tfMRI_WM_RL/*.fsf

Subject/scan affected:	File	Issue:
150423/tfMRI_GAMBLING_RL	tfMRI_GAMBLING_RL_hp200_s4_level1.fsf	npts should be 232 instead of 253
150423/tfMRI_RELATIONAL_LR	tfMRI_RELATIONAL_LR_hp200_s4_level1.fsf	npts should be 193 instead of 232
329440/tfMRI_GAMBLING_RL	tfMRI_GAMBLING_RL_hp200_s4_level1.fsf	npts should be 202 instead of 253
250427/tfMRI_WM_RL	tfMRI_WM_RL_hp200_s4_level1.fsf	file missing
250427/tfMRI_GAMBLING_RL	tfMRI_GAMBLING_RL_hp200_s4_level1.fsf	file missing

  Completed fix: These bugs have resolved for the Q3 release (September 2013).

  Users should run the Q1/Q2 data patch script to correct all these incomplete files (download available here) (requires ConnectomeDB login).

- ***Mislabelling of FreeSurfer cortical parcels in aparc.a2009s+aseg.nii.gz files***

  There is a bug in the Q1+Q2 aparc.a2009s+aseg.nii.gz files (both in the T1w/ and MNINonLinear/ directories as shown below) for all subjects, such that all of cortex is labelled as 14175, which is wm_rh_S_temporal_transverse. The current files are corrupted in every subject and not usable for any analyses.

  ${StudyFolder}/${SubjectID}/T1w/aparc.a2009s+aseg.nii.gz

  ${StudyFolder}/${SubjectID}/MNINonLinear/aparc.a2009s+aseg.nii.gz

  Completed fix: These aparc files have been regenerated in Q1/Q2 subjects from pipeline intermediates and all Q1/Q2 packages have been updated with this change for Q3. The bug has been fixed in the PostFreeSurfer pipeline and Q3 subjects' data should not have this issue.

  Users should run the Q1/Q2 data patch script to correct all these incomplete files (download available here) (requires ConnectomeDB login).

- ***Some empty vertices (zeroes) near medial wall in rfMRI dtseries data***

  In all minimally preprocessed fMRI data mapped to grayordinate surface vertices, e.g. 136833/MNINonLinear/Results/rfMRI_REST2_LR /rfMRI_REST2_LR_Atlas.dtseries.nii, there are some empty vertices (timeseries = 0s) very near the medial wall. This is because the FreeSurfer defined medial wall sometimes extends outside the atlas medial wall used to define the grayordinates space. Due to a bug in the PostFreeSurfer Pipeline, these vertices were set to zero because the affected subject's specific medial wall ROI was larger than the atlas medial wall ROI. This bug appears in approximately half the subjects' right and/or left hemispheres (too many to list here).

  This bug should not create a problem for most users because the empty vertices should always be located around the medial wall (where the data is not very good anyway). However, the zero values could cause some algorithms to crash if they can't cope with zeros in CIFTI/surface data.The tfMRI analysis pipeline is not affected by this bug and the issue is fixed in the analyzed task data. FIX is not negatively affected by this, but the bug remains in the fix cleaned timeseries.

  Planned Fix: This bug is not trivial to fix for all Q1, Q2, and Q3 subjects without running all these subjects through a new version of the minimally preprocessing pipelines (specifically the PostFreeSurfer pipeline, see Glasser et al. 2013 for details), which will take a substantial amount of computing time. To date (August 2013), we have not decided exactly when we will do this.

  Users who need to fix the files before we make the permanent fix, a relatively simple DIY fix of affected files is available using the command "-cifti-dilate" in **wb_command**.

Missing Files

- ***Missing fMRI Physio_log Files***
  Some Q1+Q2 subjects are missing physiological data (*_Physio_log.txt) for collected functional (rfMRI or tfMRI) sessions. The missing files were also documented in the Q2 Reference Manual, Appendix 8.

  e.g. ${StudyFolder}/${SubjectID}/MNINonLinear/Results/rfMRI_REST1_RL/rfMRI_REST1_RL_ Physio_log.txt

  Subjects missing files *_Physio_log.txt : Q1+Q2 Subjects Missing fMRI Physio_log Files

  Planned fix: In some of these cases where there is missing data, there was a problem with the data collection that is recoverable (e.g., in the physiological data, a gap in the timing between the trigger pulse and pulse/respiratory signal); in other cases, the data was not collected or is otherwise unrecoverable.

For those subjects whose physiological data is recoverable, the data is being extracted and will be released with a patch script to place them in the appropriate directory structure.

- ***Missing tfMRI Recognition Memory E-Prime Data Files***
  Some Q1+Q2 subjects are missing recognition memory (REC_run*_TAB.txt) tab-delimited versions of the E-Prime data files for analysis of the working memory task. The missing files were also documented in the Q2 Reference Manual, Appendix 8.

  e.g. ${StudyFolder}/${SubjectID}/MNINonLinear/Results/tfMRI_WM_RL/REC_run1_TAB.txt

  Subjects missing one or both REC_run1_TAB.txt, REC_run2_TAB.txt files: Q1+Q2 Subjects Missing tfMRI Recognition Memory E-Prime Data Files

  <u>Planned fix:</u> For those subjects whose recognition memory data is recoverable, the data is being extracted and will be released with a patch script to place them in the appropriate directory structure.

- ***Missing (Non-NIH Toolbox) Behavioral Data***
  In a limited number of Q1-Q3 subjects, responses to some behavioral and individual difference measures (specifically the tests we do in addition to the NIH Toolbox battery, i.e. Non-Toolbox tests) were not uploaded correctly when the tests were performed.
  Here is a list of the affected subjects: Q1+Q2 Subjects Missing (Non-NIH Toolbox) Behavioral Data

  <u>Near Completed fix:</u> For all but the last two subjects, 131924 and 677968, the data has been recovered and will be released as part of the Q3 release in the ConnectomeDB. For the two remaining subjects, we have sent paper forms for them to refill their responses to the questions asked at their subject visit, which we will add to ConnectomeDB once they are received.

  Users who have the Q1+Q2 data should re-download the fixed <u>behavioral data for all subjects</u> available here or on the ConnectomeDB splash page.