Consortium for Reliability and Reproducibility - CoRR

Xi-Nian Zuo (Ph.D. in Applied Mathematics)
Key Laboratory of Behavioral Science
Chinese Academy of Sciences
Highlights

• Concept of Test-Retest Reliability
• A Meta-Reliability Analysis
• ‘Optimal’ rfMRI Protocol: Challenges
• CoRR: Consortium for Reliability and Repeatability
• CCNP: Chinese Color Nest Program
Why Most Published Research Findings Are False

John P.A. Ioannidis

Over half of psychology studies fail reproducibility test

Largest replication study to date casts doubt on many published positive results.

Monya Baker

27 August 2015
Puzzlingly High Correlations in fMRI Studies of Emotion, Personality, and Social Cognition

Edward Vul, Christine Harris, Piotr Winkielman, & Harold Pashler

Voodoo Correlation? Reliability!

Power failure: why small sample size undermines the reliability of neuroscience

Katherine S. Button, John P. A. Ioannidis, Claire Mokrysz, Brian A. Nosek, Jonathan Flint, Emma S. J. Robinson and Marcus R. Munafò
Test-Retest Reliability

\[
ICC = \frac{MSb - MSw}{MSb + MSw}
\]

\[
A' = A + n(A)
\]

\[
B' = B + n(B)
\]

\[
r(A', B') = r(A, B) \sqrt{ICC(A)ICC(B)}
\]

http://en.wikipedia.org/wiki/Correction_for_attenuation

Zuo and Xing, 2014, Neurosci Biobehav Rev
Reliability: Individual Differences

\[ ICC = \frac{MSb - MSw}{MSb + MSw} \]

\[ A' = A + n(A) \]

\[ B' = B + n(B) \]

\[ r(A', B') = r(A, B) \sqrt{ICC(A)ICC(B)} \]
Validity and reliability are numbers between 0 and 1 (or 0% and 100%) are not “yes” or “no” answers.

Validity can never be greater than reliability. It is possible to have a diagnosis that is 100% reliable and 0% valid, but it is not possible to have a diagnosis with low reliability that has high validity. It is possible to increase the reliability of the diagnosis at the cost of validity.
Reliability correction for functional connectivity: Theory and implementation

Sophia Mueller\textsuperscript{1,2,3}, Danhong Wang\textsuperscript{1}, Michael D. Fox\textsuperscript{1,4,5}, Ruiqi Pan\textsuperscript{1,6}, Jie Lu\textsuperscript{6}, Kuncheng Li\textsuperscript{6}, Wei Sun\textsuperscript{7,*}, Randy L. Buckner\textsuperscript{1,2,8} and Hesheng Liu\textsuperscript{1,*}

Article first published online: 20 AUG 2015
DOI: 10.1002/hbm.22947
© 2015 Wiley Periodicals, Inc.

Human Brain Mapping
Early View (Online Version of Record published before inclusion in an issue)

\[ r(A,B) = \frac{r'(A,B)}{\sqrt{ICC(A)ICC(B)}} \]
A Meta-Reliability Analysis

Zuo and Xing, 2014, Neurosci Biobehav Rev
Toward Reliable Functional Connectomes: Reliable ICCs

Zuo and Xing, 2014, Neurosci Biobehav Rev
HCP Data Releases

Updated Mar 26, 2014: Volume-smoothed Task fMRI analysis packages have been released. See humanconnectome.org/data for all data updates.

HCP Open Access Data  |  ✔ Data Use Terms Accepted (View Terms)
Publicly released HCP datasets include high-resolution MR and MEG scans from healthy adults. Behavioral and physiological data is available for each released subject. For Connectome Workbench users, we are also making group average datasets available. More information: HCP Data Release Documentation

HCP MR Data (Q3 Release)  |  Updated Jan 28, 2014

Explore HCP Subjects  🌐 Where should I start?

- 1 Subject  🔄 Download Now  Explore [BETA]
- 10 Unrelated Subjects  🔄 Download Now  Explore [BETA]
- 40 Unrelated Subjects  🔄 Download Now  Explore [BETA]
- 80 Unrelated Subjects  🔄 Download Now  Explore [BETA]
- All Subjects  🔄 Order Connectome In A Box  Explore [BETA]
RECOMMENDATION

In light of the review on previous findings and discussions in prior, we recommend the use of functional connectomics at the current stage with:

- The instruction of eyes-open with fixation during RFMRI scans.
- The setting of a longer duration (6–15 min) for RFMRI scans.
- The acceleration of EPI sequences for high tempo-spatial resolution.
- The strategy of two-dimensional surface-based processing and analysis.
- The use of ICA, ReHo and VMHC functional connectivity metrics.
- The examination of default, control and attention networks.
‘Optimal’ rfMRI Protocols: Challenges

- ‘Resting State’ Settings
- Physiological Noises
- In-scanner Head Motion
- Registration Variation
- Global Factors (Unknown Others)

Zuo and Xing, 2014, Neurosci Biobehav Rev
Zuo et al., 2013, NeuroImage; Birn et al., 2013, Neuroimage; Laumann et al., 2015, Neuron
Motion Correction

Power et al., 2012; 2013;2014;2015; Zuo et al., 2013, Neurolmage
Motion and Registration

Zuo and Xing, 2014, Neurosci Biobehav Rev
Global Factors

Standardizing the intrinsic brain: Towards robust measurement of inter-individual variation in 1000 functional connectomes

Chao-Gan Yan a,b,c, R. Cameron Craddock a,b, Xi-Nian Zuo d, Yu-Feng Zang e, Michael P. Milham a,b,*

a Nathan Kline Institute for Psychiatric Research, Orangeburg, NY, USA
b Center for the Developing Brain, Child Mind Institute, New York, NY, USA
c The Phyllis Green and Randolph Cowen Institute for Pediatric Neuroscience, New York University Child Study Center, New York, NY, USA
d Key Laboratory of Behavioral Science, Laboratory for Functional Connectome and Development, Magnetic Resonance Imaging Research Center, Institute of Psychology, Chinese Academy of Sciences, Beijing, China
e Center for Cognition and Brain Disorders, Hangzhou Normal University, Hangzhou, China

<table>
<thead>
<tr>
<th>Category</th>
<th>Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Acquisition-related variations</td>
<td>Scanner make and model (Friedman and Glover, 2006b), sequence type (spiral vs. echo planar; single-echo vs. multi-echo) (Klarhofer et al., 2002), parallel vs. conventional acquisition (Feinberg et al., 2010; Lin et al., 2005), coil type (surface vs. volume, number of channels, orientation), repetition time, number of repetitions, flip angle, echo time, and acquisition volume (field of view, voxel size, slice thickness/gaps, slice prescription) (Friedman and Glover, 2006a)</td>
</tr>
<tr>
<td>2. Experimental-related variations</td>
<td>Participant instructions (Hartstra et al., 2011), eyes-open/eyes-closed (Yan et al., 2009; Yang et al., 2007), visual displays, experiment duration (Fang et al., 2007; Van Dijk et al., 2010)</td>
</tr>
<tr>
<td>3. Environment-related variations</td>
<td>Sound attenuation measures (Cho et al., 1998; Elliott et al., 1999), attempts to improve participant comfort during scans (e.g., music, videos) (Cullen et al., 2009), head-motion restraint techniques (e.g., vacuum pad, foam pad, bite-bar, plaster cast head holder) (Edward et al., 2000; Menon et al., 1997), room temperature and moisture (Vanhoupt et al., 2006).</td>
</tr>
<tr>
<td>4. Participant-related variations</td>
<td>Circadian cycle (Shannon et al., 2013), prandial (Haase et al., 2009), caffeine (Rack-Gomer et al., 2009), and nicotine status (Tanabe et al., 2011), sleepiness/arousal (Horovitz et al., 2008), sleep deprivation (Samann et al., 2010), scanner anxiety (de Bie et al., 2010), and menstrual cycle status (for women) (Protopopescu et al., 2005)</td>
</tr>
</tbody>
</table>
Global Mean Signal

Global Signal Regression (Individual Level)

Global Connectivity Correction (Group Level)

Jiang et al., 2015, Trans Psychiatry
Functional Smoothing

Jiang et al., 2015, Brain Struct Func
Finding the Elusive Psychiatric “Lesion” With 21st-Century Neuroanatomy: A Note of Caution

Daniel R. Weinberger, M.D., Eugenia Radulescu, M.D.

The widespread use of MRI has led to a wealth of structural and functional anatomical findings in patients with diverse psychiatric disorders that may represent insights into pathobiology. However, recent technical reports indicate that data from popular MRI research—particularly structural MRI, resting-state functional MRI, and diffusion tensor imaging—are highly sensitive to common artifacts (e.g., head motion and breathing effects) that may dominate the results. Because these and other important confounders of MRI data (e.g., smoking, body weight, metabolic variations, medical comorbidities, psychoactive drugs, alcohol use, mental state) tend to vary systematically between patient and control groups, the evidence that findings are neurobiologically meaningful is inconclusive and may represent artifacts or epiphenomena of uncertain value. The authors caution that uncritically accepting from study to study findings that may represent fallacies of all sorts carries the risk of misinforming practitioners and patients about biological abnormalities underlying psychiatric illness.

*AJP in Advance* (doi: 10.1176/appi.ajp.2015.15060753)
BIG RETEST DATA: CORR LONGITUDINAL DATA
CoRR: Consortium for Reliability and Repeatability

- Project Director: Xi-Nian Zuo; Founders: Xi-Nian Zuo, Michael P. Milham

CoRR Concept

Goal of CoRR

The overarching goal of CoRR is to create an open science resource for the imaging community that will facilitate the assessment of test-retest reliability and reproducibility for functional and structural connectomics. In order to accomplish this, we will aggregate resting state fMRI (R-fMRI) and diffusion imaging data across laboratories around the world, and share the data via the International Neuroimaging Data-sharing Initiative (INDI) to enable the:

1. Establishment of test-retest reliability and reproducibility for commonly used MR-based connectome metrics
2. Determination of the range of variation in the reliability and reproducibility of these metrics across imaging sites and retest study designs
3. Creation of a standard/benchmark test-retest dataset for the evaluation of novel metrics

Data Aggregation

Contributors: Any laboratory willing to openly share multimodal imaging datasets (including an R-fMRI scan and a corresponding anatomical image at a minimum) with at least one retest occasion. Institutional IRB/ethical committee approval or waiver (see below) is required prior to contribution of data. MRI data: Our primary focus is on R-fMRI data, with a secondary focus on diffusion imaging data. While we encourage sharing data with minimal movement, we are placing no exclusion criteria for motion. This decision was based on the realizations that: 1) there is no consensus on acceptable criteria for movement in functional MRI or diffusion imaging data, 2) high motion datasets are essential to the determination of the impact of motion on reliability, and 3) new approaches continue to be developed to account for movement artifacts. We also encourage submission of data from other modalities (e.g., ASL) or experimental paradigms (e.g., task data) when available for the same participants for whom R-fMRI data are being provided.

Total 36 imaging groups from 18 sites
> 1600 participants, > 10,000 images
One-Year Test-Retest

Total 243 Subjects

Test

Retest

(PI: Jiang Qiu)
Six-Month Test-Retest

Total 23 Subjects (4 Retests/Subject)
Total 30 Subjects (9 Retests/Subject)
8AM ~ 8PM: 3 scans per hour

day1 ... day3

day1 ... day3

day1 ... day3

Test | Three-Day Test-Retest | Retest

Total 2 Subjects (44 Retests/Subject)

(PI: Xi-Nian Zuo)
An open science resource for establishing reliability and reproducibility in functional connectomics

Xi-Nian Zuo¹,²,*, Jeffrey S. Anderson³, Pierre Bellec⁴, Rasmus M. Birn⁵, Bharat B. Biswal⁶, Janusch Blautzik⁷, John C.S. Breitner⁸, Randy L. Buckner⁹, Vince D. Calhoun¹⁰, F. Xavier Castellanos¹¹,¹², Antao Chen⁷, Bing Chen¹³, Jiangtao Chen¹, Xu Chen¹, Stanley J. Colcombe¹¹, William Courtney¹⁰, R. Cameron Craddock¹¹,¹⁴, Adriana Di Martino¹², Hao-Ming Dong¹¹,¹⁴, Xiaolan Fu¹⁴, Qi Yong Gong¹⁷, Krzysztof J. Gorgolewski¹¹, Ying Han¹⁹, Ye He¹,²⁵, Yong He¹⁰, Erica Ho¹¹,¹⁴, Avram Holmes¹⁴, Xiao-Hui Hou¹⁴,¹⁵, Jeremy Huckins¹², Tianzi Jiang¹³, Yi Jiang¹, William Kelley²², Clare Kelly²², Margaret King¹⁰, Stephen M. LaConte²⁴, Janet E. Lainhart⁵, Xu Lei², Hui-Jie Li³, Kaiming Li¹⁷, Kuncheng Li²⁵, Qixiang Lin²⁰, Dongqiang Liu¹³, Jia Liu²⁰, Xun Liu¹, Yijun Liu², Guangming Lu²⁶, Jie Lu²⁵, Beatriz Luna²⁷, Jing Luo²⁸, Daniel Lurie¹¹,¹⁴, Ying Mao²⁹, Daniel S. Margulies¹⁸, Andrew R. Mayer¹⁰, Thomas Meинд十七, Mary E. Meyerand³⁹, Weizhi Nan²,³,⁵, Jared A. Nielsen³, David O’Connor¹²,¹⁴, David Paulsen³⁷, Vivek Prabhakaran³¹, Zhigang Qi²⁵, Jiang Qiu², Chunhong Shao³², Zarrar Shehzad³³,¹⁴, Weijun Tang³³, Arno Villringer³⁴, Huijing Wang³⁵, Kai Wang³,¹⁵, Dongtao Wei¹, Gao-Xia Wei², Xu-Chu Weng³³, Xuehai Wu²⁹, Ting Xu³,¹³,¹⁴, Ning Yang¹,¹⁵, Zhi Yang³, Yu-Feng Zhang¹³, Lei Zhang¹,¹⁵, Qinglin Zhang³, Zhe Zhang¹,¹⁵, Zhiqiang Zhang²⁶, Ke Zhao¹, Zonglei Zhen²⁰, Yuan Zhou⁴, Xing-Ting Zhu⁴,¹⁵ & Michael P. Milham¹¹,¹⁴
### CCNP: RELIABILITY FOR DEVELOPING BRAINS

#### Structured 3-Cohort Longitudinal Design

<table>
<thead>
<tr>
<th>基线</th>
<th>追踪1</th>
<th>追踪2</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>7.25</td>
<td>8.5</td>
</tr>
<tr>
<td>7</td>
<td>8.25</td>
<td>9.5</td>
</tr>
<tr>
<td>8</td>
<td>9.25</td>
<td>10.5</td>
</tr>
<tr>
<td>9</td>
<td>10.25</td>
<td>11.5</td>
</tr>
<tr>
<td>10</td>
<td>11.25</td>
<td>12.5</td>
</tr>
<tr>
<td>11</td>
<td>12.25</td>
<td>13.5</td>
</tr>
<tr>
<td>12</td>
<td>13.25</td>
<td>14.5</td>
</tr>
<tr>
<td>13</td>
<td>14.25</td>
<td>15.5</td>
</tr>
<tr>
<td>14</td>
<td>15.25</td>
<td>16.5</td>
</tr>
<tr>
<td>15</td>
<td>16.25</td>
<td>17.5</td>
</tr>
<tr>
<td>16</td>
<td>17.25</td>
<td>18.5</td>
</tr>
<tr>
<td>17</td>
<td>18.25</td>
<td>19.5</td>
</tr>
<tr>
<td>18</td>
<td>19.25</td>
<td>20.5</td>
</tr>
</tbody>
</table>

- **192 Participants**
- **192*4 Scans**

#### International Collaboration NSFC Grant with Drs. Milham, Castellanos and Sporns

#### Local Collaboration with Drs. Chen Xu, Qiu Jiang and Chen Antao