A Simple and Objective Method for Reproducible Resting State Network (RSN) Detection in fMRI

Gautam V. Pendse^{1*}, David Borsook^{1,2}, Lino Becerra^{1,2}

1 P.A.I.N Group, Imaging and Analysis Group (IMAG), McLean Hospital, Harvard Medical School, Belmont, Massachusetts, United States of America, **2** A. A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Charlestown, Massachusetts, United States of America

Abstract

Spatial Independent Component Analysis (ICA) decomposes the time by space functional MRI (fMRI) matrix into a set of 1-D basis time courses and their associated 3-D spatial maps that are optimized for mutual independence. When applied to resting state fMRI (rsfMRI), ICA produces several spatial independent components (ICs) that seem to have biological relevance - the so-called resting state networks (RSNs). The ICA problem is well posed when the true data generating process follows a linear mixture of ICs model in terms of the identifiability of the mixing matrix. However, the contrast function used for promoting mutual independence in ICA is dependent on the finite amount of observed data and is potentially non-convex with multiple local minima. Hence, each run of ICA could produce potentially different IC estimates even for the same data. One technique to deal with this run-to-run variability of ICA was proposed by [1] in their algorithm RAICAR which allows for the selection of only those ICs that have a high run-to-run reproducibility. We propose an enhancement to the original RAICAR algorithm that enables us to assign reproducibility *p*-values to each IC and allows for an objective assessment of both within subject and across subjects reproducibility. We call the resulting algorithm RAICAR-N (N stands for null hypothesis test), and we have applied it to publicly available human rsfMRI data (<http://www.nitrc.org>). Our reproducibility analyses indicated that many of the published RSNs in rsfMRI literature are highly reproducible. However, we found several other RSNs that are highly reproducible but not frequently listed in the literature.

Citation: Pendse GV, Borsook D, Becerra L (2011) A Simple and Objective Method for Reproducible Resting State Network (RSN) Detection in fMRI. PLoS ONE 6(12): e27594. doi:10.1371/journal.pone.0027594

Editor: Wang Zhan, University of California San Francisco, United States of America

Received: June 28, 2011; **Accepted:** October 20, 2011; **Published:** December 12, 2011

Copyright: © 2011 Pendse et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by the National Institutes of Health (NIH) under the grants K24NS064050 (to DB) and R01NS065051 (to DB). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: gpendse@mclean.harvard.edu

Introduction

Independent component analysis (ICA) [2–5] models the observed data as a linear combination of a set of statistically independent and unobservable sources [6]. First proposed the application of ICA to the analysis of functional magnetic resonance imaging (fMRI) data. Subsequently, ICA has been applied to fMRI both as an exploratory tool for the purpose of identifying task related components [6] as well as a signal clean up tool for the purpose of removing artifacts from the fMRI data [7]. Recently, it has been shown that ICA applied to resting state fMRI (rsfMRI) in healthy subjects reveals a set of biologically meaningful spatial maps of independent components (ICs) that are consistent across subjects - the so called resting state networks (RSNs) [8]. Hence, there is a considerable interest in applying ICA to rsfMRI data in order to define the set of RSNs that characterize a particular group of human subjects, a disease, or a pharmacological effect.

Several variants of the linear ICA model have been applied to fMRI data including square ICA (with equal number of sources and sensors) [9], non-square ICA (with more sensors than sources) [6], and non-square ICA with additive Gaussian noise (noisy ICA) [10]. All of these models are well known in the ICA literature [2,3,5,11]. Since the other ICA models are specializations of the noisy ICA model, we will assume a noisy ICA model henceforth.

Remarkably, the ICA estimation problem is well posed in terms of the identifiability of the mixing matrix given several non-Gaussian and at most 1 Gaussian source in the overall linear mixture [3,12–14]. In the presence of more than 1 Gaussian source, such as in noisy ICA, the mixing matrix corresponding to the non-Gaussian part of the linear mixture is identifiable (upto permutation and scaling). In addition, the source distributions are uniquely identifiable (upto permutation and scaling) given a noisy ICA model with a particular Gaussian co-variance structure, for example, the isotropic diagonal co-variance. For details, see section 2.1.2.

While these uniqueness results are reassuring, a number of practical difficulties prevent the reliable estimation of ICs on real data. These difficulties include (1) true data not describable by an ICA model, (2) ICA contrast function approximations, (3) multiple local minima in the ICA contrast function, (4) confounding Gaussian noise and (5) model order overestimation. See section 2.1.3 for more details. A consequence of these difficulties is that multiple ICA runs on the same data or different subsets of the data produce different estimates of the IC realizations.

One technique to account for this run-to-run variability in ICA was proposed by [15] in their algorithm ICASSO. Using repeated runs of ICA with bootstrapped data using various initial conditions, ICASSO clusters ICs across ICA runs using agglomerative hierarchical clustering and also helps in visualizing the estimated ICs. The logic is that reliable ICs will show up in almost

all ICA runs and thus will form a tight cluster well separated from the rest. [16] proposed a technique similar to ICASSO called self-organizing group ICA (sogICA) which allows for clustering of ICs via hierarchical clustering in across subject ICA runs. When applied to multiple ICA runs across subjects, ICASSO does not restrict the IC clusters to contain only 1 IC from each subject per ICA run. In contrast, sogICA allows the user to select the minimum number of subjects for a “group representative” IC cluster containing distinct subjects. By labelling each ICA run as a different “subject” sogICA can also be applied to analyze multiple ICA runs across subjects.

Similar in spirit to ICASSO and sogICA [1], proposed an intuitive approach called RAICAR (Ranking and Averaging Independent Component Analysis by Reproducibility) for reproducibility analysis of estimated ICs. The basic idea in RAICAR is to select only those ICs as “interesting” or “stable” which show a high run-to-run “reproducibility”. RAICAR uses simple and automated spatial cross-correlation matrix based IC alignment, which has been shown to be more accurate compared to ICASSO [1]. RAICAR is applicable to both within subject as well as across subjects reproducibility analysis.

A few limitations of ICASSO, sogICA and RAICAR are worth noting:

- ICASSO requires the user to select the number of IC clusters and is inapplicable without modification for across subjects analysis of ICA runs since the IC clusters are not restricted to contain only 1 IC per ICA run.
- sogICA requires the user to select the minimum number of subjects for a “group representative” cluster and also a cutoff on within cluster distances.
- RAICAR uses an arbitrary threshold on the reproducibility indices selected “by eye” or set at an arbitrary value, such as 50% of the maximum reproducibility value.

We propose a simple extension to RAICAR that avoids subjective user decisions and allows for an automatic reproducibility cutoff. The reproducibility indices calculated in RAICAR differ in magnitude significantly depending on whether the input to RAICAR:

- (a) is generated using multiple ICA runs on the same data
- (b) comes from multiple ICA runs on varying data sets (e.g. between and across subject runs)

See Figure 1 for an illustration of this effect. Obviously, the reproducibility indices are much lower in case (b) since we account for both within subject and between subjects variability in estimating ICs. Case (b) is also of great interest from a practical point of view since we are often interested in making statements about a group of subjects. Hence, it is clear that a cutoff on RAICAR reproducibility values for the purposes of selecting the “highly reproducible” components should be data dependent. In this work,

1. We propose a modification of the original RAICAR algorithm by introducing an explicit “null” model of no reproducibility.
2. We use this “null” model to automatically generate p -values for each IC via simulation. This allows for an objective cutoff specification for extracting reproducible ICs (e.g. reproducible at $p < 0.05$) within and across subjects. We call the resulting algorithm RAICAR-N (N stands for “null” hypothesis test).
3. We validate RAICAR-N by applying it to publicly available human rsfMRI data.

1.1 Notation

- The set of real numbers will be denoted by $\mathbb{R}$. Scalars variables and functions will be denoted in a non-bold font (e.g., σ^2, L, p or Ψ, f). Vectors will be denoted in a bold font (except Greek letters) using lower case letters (e.g., $\mathbf{y}, \mu, \eta$). Matrices will be denoted in bold font using upper case letters (e.g., $\mathbf{A}, \mathbf{\Sigma}, \mathbf{W}$). The transpose of a matrix $\mathbf{A}$ will be denoted by $\mathbf{A}^T$ and its inverse will be denoted by $\mathbf{A}^{-1}$. $\mathbf{I}_p$ will denote the $p \times p$ identity matrix and $\mathbf{0}$ will denote a vector or matrix of all zeros whose size should be clear from context. $\binom{N}{L}$ is the number of ways of choosing L objects from N objects when order does not matter.
- The j th component of vector $\mathbf{t}_i$ will be denoted by t_{ij} whereas the j th component of vector $\mathbf{t}$ will be denoted by t_j . The element (i, j) of matrix $\mathbf{G}$ will be denoted by $G(i, j)$ or G_{ij} . Estimates of variables will be denoted by putting a hat on top of the variable symbol. For example, an estimate of $\mathbf{s}$ will be denoted by $\hat{\mathbf{s}}$.
- If $\mathbf{x}$ is a random vector with a multivariate Normal distribution with mean $\mathbf{m}$ and covariance $\mathbf{\Sigma}$ then we will denote this distribution by $\mathcal{N}(\mathbf{x}|\mu, \mathbf{\Sigma})$. The joint density of vector $\mathbf{s}$ will be denoted by $\mathbf{p}_s(\mathbf{s})$ whereas the marginal density of s_i will be denoted as $p_{s_i}(s_i)$. $\mathbf{E}[f(\mathbf{s}, \mathbf{h})]$ denotes the expectation of $f(\mathbf{s}, \eta)$ with respect to both random variables $\mathbf{s}$ and η .

Methods

The organization of this article revolves around the following sequence of questions, which ultimately lead to the development of RAICAR-N:

1. Why is a reproducibility assessment necessary in ICA analysis? In order to answer this question, we cover the fundamentals of ICA including identifiability issues in sections 2.1 and 2.2.
2. How does the original RAICAR algorithm assess reproducibility? The answer to this question in section 2.3 will set up the stage for RAICAR-N.
3. How does RAICAR-N permit calculation of reproducibility p -values? In section 2.4, we describe the RAICAR-N “null” model and a simulation based approach for assigning p -values to ICs.
4. How to promote diversity in group ICA runs given a limited number of subjects when using RAICAR-N and how to display the non-Gaussian spatial structure in estimated ICs? These issues are covered in section 2.5 and 2.6.
5. How can RAICAR-N be extended for between group comparison of ICs and how does it compare to other approaches in the literature? This question is addressed in section 4.4.

2.1 ICA background

In this section, we provide a brief introduction to ICA along with a discussion of associated issues related to model order selection, identifiability and run-to-run variability. The noisy ICA model assumes that observed data $\mathbf{y}$ is generated as a linear combination of unobservable independent sources confounded with Gaussian noise:

$$\mathbf{y} = \mu + \mathbf{A}\mathbf{s} + \eta \quad (2.1)$$

Figure 1. Figure illustrates the variation in normalized reproducibility from RAICAR depending on whether the input to RAICAR is (a) Multiple ICA runs on single subject data or (b) Multiple ICA runs across subjects. Notice that the normalized reproducibility is much lower for across subjects analysis compared to within subject analysis.
doi:10.1371/journal.pone.0027594.g001

In this model,

$\mathbf{y} = p \times 1$ observed signal vector

$\boldsymbol{\mu} = p \times 1$ mean vector

$\mathbf{A} = p \times q$ mixing matrix with $p > q$ (more sensors than sources) and rank q

$\boldsymbol{\eta} = p \times 1$ Gaussian noise vector with density $\mathcal{N}(\boldsymbol{\eta} | \mathbf{0}, \boldsymbol{\Sigma})$ (2.2)

$\mathbf{s} = q \times 1$ vector of independent random variables (the ICs)

with $\mathbf{E}(\mathbf{s}\mathbf{s}^T) = \mathbf{D}$ (diagonal) and $\mathbf{E}(\mathbf{s}) = \mathbf{0}$

and with $\mathbf{s}$ and $\mathbf{h}$ independent

If the marginal density of the i th source s_i is $p_{s_i}(s_i)$ then the joint source density $\mathbf{p}_s(\mathbf{s})$ factorizes as $\prod_{i=1}^q p_{s_i}(s_i)$ because of the independence assumption but is otherwise assumed to be unknown. Also, since the elements of $\mathbf{s}$ are independent their co-variance matrix $\mathbf{D}$ is diagonal. The set of variables $\mathcal{F} = \{\boldsymbol{\mu}, \mathbf{A}, \mathbf{D}, \boldsymbol{\Sigma}\}$ represents the unknown parameters in the noisy ICA model. Before discussing the identifiability of model 2.1, we briefly discuss the choice of model order or the assumed number of ICs q .

2.1.1 Estimating the model order q . Rigorous estimation of the model order q in noisy ICA is difficult as the IC densities $p_{s_i}(s_i)$ are unknown. This means that $\mathbf{p}(\mathbf{y}|q, \mathcal{F})$, the marginal density of the observed data given the model order and the ICA

parameters cannot be derived in closed form (by integrating out the ICs) without making additional assumptions on the form of IC densities. Consequently, standard model selection criteria such as Bayes information criterion (BIC) [17] cannot be easily applied to the noisy ICA model to estimate q . One solution is to use a factorial mixture of Gaussians (MOG) joint source density model as in [5], and use the analytical expression for $\mathbf{p}(\mathbf{y}|q, \mathcal{F})$ in conjunction with BIC. This solution is quite general in terms of allowing for an arbitrary Gaussian noise co-variance $\boldsymbol{\Sigma}$, but maximizing $\mathbf{p}(\mathbf{y}|q, \mathcal{F})$ with respect to $\mathcal{F}$ becomes computationally intractable using an expectation maximization (EM) algorithm for $q > 13$ ICs [5]. Another rigorous non-parametric approach for estimating q that is applicable to the noisy ICA model with isotropic diagonal Gaussian noise co-variance i.e., with $\boldsymbol{\Sigma} = \sigma^2 \mathbf{I}_p$ is the random matrix theory based sequential hypothesis testing approach of [18]. To the best of our knowledge, these are the only 2 rigorous approaches for estimating q in the noisy ICA model.

Approximate approaches for estimating q commonly used in fMRI literature (e.g., [10]) consist of first relaxing the isotropic diagonal noisy ICA model (with $\boldsymbol{\Sigma} = \sigma^2 \mathbf{I}_p$) into a probabilistic PCA (PPCA) model of [19] where the source densities are assumed to be Gaussian i.e., where $\mathbf{p}_s(\mathbf{s}) = \mathcal{N}(\mathbf{s} | \mathbf{0}, \mathbf{I}_q)$. When using the PPCA model, it becomes possible to integrate out the Gaussian sources to get an expression for $\mathbf{p}(\mathbf{y}|q, \mathcal{F})$ that can be analytically maximized [19]. Subsequently, methods such as BIC can be applied to estimate q . Alternative approaches for estimating q in the PPCA

model consist of the Bayesian model selection of [20], or in data-rich situations such as fMRI, even the standard technique of cross-validation [21].

From a biological point of view, it has been argued [22] that the number of extracted ICs simply reflect the various *equally valid* views of the human functional neurobiology - smaller number of ICs represent a coarse view while a larger number of ICs represent a more fine grained view. However, it is worth noting that from a statistical point of view, over-specification of q will lead to over-fitting of the ICA model, which might render the estimated ICs less generalizable across subjects. On the other hand, under-specification of q will result in incomplete IC separation. Both of these scenarios are undesirable.

2.1.2 Identifiability of the noisy ICA model. To what extent is the noisy linear ICA model identifiable? Consider a potentially different decomposition of the noisy ICA model 2.1:

$$\mathbf{y} = \mu_1 + \mathbf{A}_1 \mathbf{s}_1 + \eta_1 \quad (2.3)$$

where

$\mathbf{y} = p \times 1$ observed signal vector

$\mu_1 = p \times 1$ mean vector

$\mathbf{A}_1 = p \times q$ mixing matrix with $p > q$ (more sensors than sources) and rank q

$\eta_1 = p \times 1$ Gaussian noise vector with density $\mathcal{N}(\eta_1 | 0, \Sigma_1)$ (2.4)

$\mathbf{s}_1 = q \times 1$ vector of independent random variables (the ICs)

with $\mathbf{E}(\mathbf{s}\mathbf{s}^T) = \mathbf{D}_1$ (diagonal) and $\mathbf{E}(\mathbf{s}) = \mathbf{0}$

and with $\mathbf{s}_1$ and $\mathbf{h}_1$ independent

What can be said about the equivalence between the parameterizations in 2.1 and 2.3?

Identifiability of μ : Equating the expectations of the right hand side of 2.3 and 2.1 and noting that $\mathbf{s}, \mathbf{h}, \mathbf{s}_1, \mathbf{h}_1$ have mean 0 we get:

$$\mu_1 = \mu \quad (2.5)$$

Thus the mean vector μ is exactly identifiable.

Identifiability of $\mathbf{A}$: A fundamental decomposition result states that the noisy ICA problem is well-posed in terms of the identifiability of the mixing matrix $\mathbf{A}$ upto permutation and scaling provided that the components of $\mathbf{s}$ are independent and **non-Gaussian** [3,12–14]. If Λ is a diagonal scaling matrix and $\mathbf{P}$ is a permutation matrix then the identifiability result can be stated as:

$$\mathbf{A}_1 = \mathbf{A} \Lambda \mathbf{P} \quad (2.6)$$

where 2.3 is another decomposition of $\mathbf{y}$ with $\mathbf{s}_1$ containing independent and non-Gaussian components. In other words, the mixing matrix $\mathbf{A}$ is identifiable upto permutation and scaling.

Identifiability of $\mathbf{D}$ and Σ : Equating the second moments of the right hand side of 2.3 and 2.1 and noting the equality of means 2.5 and the independence of $\mathbf{s}, \mathbf{h}$ and $\mathbf{s}_1, \mathbf{h}_1$ we get:

$$\mathbf{E}[(\mathbf{y} - \mu)(\mathbf{y} - \mu)^T] = \mathbf{A} \mathbf{D} \mathbf{A}^T + \Sigma = \mathbf{A}_1 \mathbf{D}_1 \mathbf{A}_1^T + \Sigma_1 \quad (2.7)$$

Let $\mathbf{W}$ be a $q \times p$ matrix and $\tilde{\mathbf{Q}}$ be a $p \times (p - q)$ orthogonal matrix such that:

$$\begin{aligned} \mathbf{W} &= (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \\ \tilde{\mathbf{Q}}^T \mathbf{A} &= \mathbf{0} \\ \tilde{\mathbf{Q}}^T \tilde{\mathbf{Q}} &= \mathbf{I}_{p-q} \end{aligned} \quad (2.8)$$

From 2.8 and 2.7 we get:

$$\mathbf{D} + \mathbf{W} \Sigma \mathbf{W}^T = \Lambda \mathbf{P} \mathbf{D}_1 \mathbf{P}^T \Lambda^T + \mathbf{W} \Sigma_1 \mathbf{W}^T \quad (2.9)$$

$$\tilde{\mathbf{Q}}^T \Sigma \tilde{\mathbf{Q}} = \tilde{\mathbf{Q}}^T \Sigma_1 \tilde{\mathbf{Q}}$$

Case 1: $\Sigma = \sigma^2 \mathbf{I}_p$ and $\Sigma_1 = \sigma_1^2 \mathbf{I}_p$. The second equation in 2.9 along with the orthogonality of $\tilde{\mathbf{Q}}$ gives $\sigma^2 = \sigma_1^2$ and thus $\Sigma = \Sigma_1$. If we fix the scaling of $\mathbf{A}_1$ by selecting $\Lambda^2 = \mathbf{I}_q$ then from the first equation in 2.9 we get:

$$\begin{aligned} \mathbf{D} &= \Lambda \mathbf{P} \mathbf{D}_1 \mathbf{P}^T \Lambda^T \\ &= \mathbf{P} \mathbf{D}_1 \mathbf{P}^T \Lambda^2 \quad (\mathbf{P} \mathbf{D}_1 \mathbf{P}^T \text{ is diagonal}) \\ &= \mathbf{P} \mathbf{D}_1 \mathbf{P}^T \end{aligned} \quad (2.10)$$

In other words, the noise co-variance $\Sigma = \sigma^2 \mathbf{I}_p$ is uniquely determined and for a fixed scaling $\Lambda^2 = \mathbf{I}_q$, the source variances $\mathbf{D}$ are also uniquely determined upto permutation.

Case 2: Σ and Σ_1 arbitrary positive definite matrices. Suppose $\mathbf{X}$ is a square matrix and let $\text{diag}(\mathbf{X})$ be the diagonal matrix obtained by setting the non-diagonal elements of $\mathbf{X}$ to 0 and similarly let $\text{offdiag}(\mathbf{X})$ be the matrix obtained by setting the diagonal elements of $\mathbf{X}$ to 0. The noise-covariance is partially identifiable by the following conditions:

$$\begin{aligned} \tilde{\mathbf{Q}}^T \Sigma \tilde{\mathbf{Q}} &= \tilde{\mathbf{Q}}^T \Sigma_1 \tilde{\mathbf{Q}} \\ \text{offdiag}(\mathbf{W} \Sigma \mathbf{W}^T) &= \text{offdiag}(\mathbf{W} \Sigma_1 \mathbf{W}^T) \end{aligned} \quad (2.11)$$

For a fixed scaling $\Lambda^2 = \mathbf{I}_q$, the sources variances $\mathbf{D}, \mathbf{D}_1$ are constrained by:

$$\mathbf{D} + \text{diag}(\mathbf{W} \Sigma \mathbf{W}^T) = \mathbf{P} \mathbf{D}_1 \mathbf{P}^T + \text{diag}(\mathbf{W} \Sigma_1 \mathbf{W}^T) \quad (2.12)$$

In general, the source variances $\mathbf{D}$ cannot be uniquely determined as noted in [14].

Identifiability of the distribution of $\mathbf{s}$: Is the distribution of the non-Gaussian components of $\mathbf{s}$ identifiable? From 2.1 and 2.3:

$$\mu + \mathbf{A} \mathbf{s} + \eta = \mu_1 + \mathbf{A}_1 \mathbf{s}_1 + \eta_1 \quad (2.13)$$

Substituting 2.5 and 2.6 in 2.13 we get:

$$\mathbf{A} \mathbf{s} + \eta = \mathbf{A} \Lambda \mathbf{P} \mathbf{s}_1 + \eta_1 \quad (2.14)$$

Left multiplying both sides by $\mathbf{W}$ from 2.8 we get:

$$\mathbf{s} + \mathbf{W} \mathbf{h} = \mathbf{L} \mathbf{P} \mathbf{s}_1 + \mathbf{W} \mathbf{h}_1 \quad (2.15)$$

Let $\Psi_s, \Psi_h, \Psi_{s_1}, \Psi_{s_1}$ be the characteristic functions of $\mathbf{s}, \mathbf{h}, \mathbf{s}_1$ and $\mathbf{h}_1$ respectively. Then

$$\begin{aligned}\Psi_s(\mathbf{t}) &= \mathbb{E}[\exp(it^T \mathbf{s})] & \Psi_\eta(\mathbf{t}) &= \mathbb{E}[\exp(it^T \boldsymbol{\eta})] \\ \Psi_{\mathbf{s}_1}(\mathbf{t}) &= \mathbb{E}[\exp(it^T \mathbf{s}_1)] & \Psi_{\eta_1}(\mathbf{t}) &= \mathbb{E}[\exp(it^T \boldsymbol{\eta}_1)]\end{aligned}\quad (2.16)$$

where $i = \sqrt{-1}$ and $\mathbf{t}$ is a vector of real numbers of length equal to that of the corresponding random vectors in 2.16. Using 2.15, we can write:

$$\mathbb{E}[\exp(it^T \{\mathbf{s} + \mathbf{W}\boldsymbol{\eta}\})] = \mathbb{E}[\exp(it^T \{\boldsymbol{\Lambda} \mathbf{P} \mathbf{s}_1 + \mathbf{W}\boldsymbol{\eta}_1\})] \text{ for all } \mathbf{t} \in \mathbb{R}^q \quad (2.17)$$

Noting the independence of $\mathbf{s}, \boldsymbol{\eta}$ and $\mathbf{s}_1, \boldsymbol{\eta}_1$:

$$\begin{aligned}\mathbb{E}[\exp(it^T \{\mathbf{s}\})] \mathbb{E}[\exp(it^T \{\mathbf{W}\boldsymbol{\eta}\})] \\ = \mathbb{E}[\exp(it^T \{\boldsymbol{\Lambda} \mathbf{P} \mathbf{s}_1\})] \mathbb{E}[\exp(it^T \{\mathbf{W}\boldsymbol{\eta}_1\})] \\ \Rightarrow \Psi_s(\mathbf{t}) \Psi_\eta(\mathbf{W}^T \mathbf{t}) \\ = \Psi_{\mathbf{s}_1}(\mathbf{P}^T \boldsymbol{\Lambda}^T \mathbf{t}) \Psi_{\eta_1}(\mathbf{W}^T \mathbf{t}) \text{ for all } \mathbf{t} \in \mathbb{R}^q\end{aligned}\quad (2.18)$$

Now $\boldsymbol{\eta}$ and $\boldsymbol{\eta}_1$ are multivariate Gaussian random vectors both with mean 0 and co-variance matrix $\boldsymbol{\Sigma}$ and $\boldsymbol{\Sigma}_1$ respectively. Hence, their characteristic functions are given by [23,24]:

$$\Psi_\eta(\mathbf{W}^T \mathbf{t}) = \exp\left(-\frac{1}{2} \mathbf{t}^T \mathbf{W} \boldsymbol{\Sigma} \mathbf{W}^T \mathbf{t}\right) \text{ for all } \mathbf{t} \in \mathbb{R}^q \quad (2.19)$$

$$\Psi_{\eta_1}(\mathbf{W}^T \mathbf{t}) = \exp\left(-\frac{1}{2} \mathbf{t}^T \mathbf{W} \boldsymbol{\Sigma}_1 \mathbf{W}^T \mathbf{t}\right) \text{ for all } \mathbf{t} \in \mathbb{R}^q$$

Claim 2.1 A sufficient condition for identifiability upto permutation and scaling of the non-Gaussian distributions in $\mathbf{s}$ given two different parameterizations in 2.1 and 2.3 is:

$$\text{diag}(\mathbf{W} \boldsymbol{\Sigma} \mathbf{W}^T) = \text{diag}(\mathbf{W} \boldsymbol{\Sigma}_1 \mathbf{W}^T) \quad (2.20)$$

Proof. From 2.20 and 2.11, we get:

$$\mathbf{W} \boldsymbol{\Sigma} \mathbf{W}^T = \mathbf{W} \boldsymbol{\Sigma}_1 \mathbf{W}^T \quad (2.21)$$

Thus from 2.19,

$$\Psi_\eta(\mathbf{W}^T \mathbf{t}) = \Psi_{\eta_1}(\mathbf{W}^T \mathbf{t}) \text{ for all } \mathbf{t} \in \mathbb{R}^q \quad (2.22)$$

From 2.19, $\Psi_\eta(\mathbf{W}^T \mathbf{t})$ and $\Psi_{\eta_1}(\mathbf{W}^T \mathbf{t})$ are not equal to 0 for any finite $\mathbf{t}$, therefore, from 2.22 and 2.18 we get:

$$\Psi_s(\mathbf{t}) = \Psi_{\mathbf{s}_1}(\mathbf{P}^T \boldsymbol{\Lambda}^T \mathbf{t}) \text{ for all } \mathbf{t} \in \mathbb{R}^q \quad (2.23)$$

Note that $\boldsymbol{\Lambda}$ is a diagonal scaling matrix with entries $\lambda_1, \lambda_2, \dots, \lambda_q$ on the diagonal and $\mathbf{P}$ is a permutation matrix. Thus,

$$\mathbf{P}^T \boldsymbol{\Lambda}^T \mathbf{t} = \begin{pmatrix} \lambda_{i_1} t_{i_1} \\ \lambda_{i_2} t_{i_2} \\ \vdots \\ \lambda_{i_q} t_{i_q} \end{pmatrix} \quad (2.24)$$

where $i_1, i_2, \dots, i_q$ is some permutation of integers $1, 2, \dots, q$. Suppose $\Psi_{s(j)}$ is the characteristic function of the j th component of $\mathbf{s}$ and $\Psi_{s_1(j)}$ is the characteristic function of the j th component of $\mathbf{s}_1$. Since the components of $\mathbf{s}$ and $\mathbf{s}_1$ are independent by assumption, the joint characteristic functions $\Psi(\mathbf{s})$ and $\Psi(\mathbf{s}_1)$ factorize:

$$\Psi_s(\mathbf{t}) = \Psi_{s(1)}(t_1) \Psi_{s(2)}(t_2) \dots \Psi_{s(j)}(t_j) \dots \Psi_{s(q)}(t_q) \quad (2.25)$$

$$\begin{aligned}\Psi_{\mathbf{s}_1}(\mathbf{P}^T \boldsymbol{\Lambda}^T \mathbf{t}) &= \Psi_{s_1(1)}(\lambda_{i_1} t_{i_1}) \Psi_{s_1(2)}(\lambda_{i_2} t_{i_2}) \\ &\dots \Psi_{s_1(j)}(\lambda_{i_j} t_{i_j}) \dots \Psi_{s_1(q)}(\lambda_{i_q} t_{i_q})\end{aligned}$$

From 2.25 and 2.23

$$\begin{aligned}\Psi_{s(1)}(t_1) \dots \Psi_{s(j)}(t_j) \dots \Psi_{s(q)}(t_q) \\ = \Psi_{s_1(1)}(\lambda_{i_1} t_{i_1}) \dots \Psi_{s_1(j)}(\lambda_{i_j} t_{i_j}) \dots \Psi_{s_1(q)}(\lambda_{i_q} t_{i_q})\end{aligned}\quad (2.26)$$

All characteristic functions satisfy [23,24]:

$$\begin{aligned}\Psi_{s(k)}(0) &= 1 \\ \Psi_{s_1(k)}(0) &= 1 \text{ for all } k\end{aligned}\quad (2.27)$$

Since $i_1, i_2, \dots, i_q$ is simply a permutation of integers $1, 2, \dots, q$, there exists a j such that $i_j = 1$. Then set $t_2 = 0, t_3 = 0, \dots, t_q = 0$ in 2.26. Then 2.27 and 2.26 imply:

$$\Psi_{s(1)}(t_1) = \Psi_{s_1(j)}(\lambda_{i_j} t_{i_j}) = \Psi_{s_1(j)}(\lambda_1 t_1) \text{ for all } t_1 \in \mathbb{R} \quad (2.28)$$

Select the scaling matrix as $\boldsymbol{\Lambda}^2 = \mathbf{I}_q$ and thus $\boldsymbol{\Lambda}$ is a diagonal matrix with elements ± 1 on the diagonal. Thus $\lambda_1 = \pm 1$ and 2.28 can be re-written as:

$$\Psi_{s(1)}(t_1) = \Psi_{s_1(j)}(\pm t_1) \text{ for all } t_1 \in \mathbb{R} \quad (2.29)$$

Therefore,

$$\Psi_{s(1)}(t_1) = \Psi_{s_1(j)}(t_1) \text{ for all } t_1 \in \mathbb{R} \quad (2.30)$$

or

$$\Psi_{s(1)}(t_1) = \Psi_{s_1(j)}(-t_1) = \Psi_{-s_1(j)}(t_1) \text{ for all } t_1 \in \mathbb{R}$$

Hence the characteristic function of the 1st component of $\mathbf{s}$ is identical to the characteristic function of the (possibly sign-flipped) j th component of $\mathbf{s}_1$. Since characteristic functions uniquely characterize a probability distribution [23], the distribution of $\mathbf{s}(1)$ and $\pm \mathbf{s}_1(j)$ is identical. Next, by setting $t_1 = 0, t_3 = 0, \dots, t_q = 0$, we can find a distribution from $\mathbf{s}_1$ that matches the 2nd component $\mathbf{s}(2)$ of $\mathbf{s}$. Proceeding in a similar fashion, it is clear that the distribution of each component of $\mathbf{s}$ is uniquely identifiable upto sign flips for the choice $\boldsymbol{\Lambda}^2 = \mathbf{I}_q$. For a general $\boldsymbol{\Lambda}$, the source distributions are uniquely identifiable upto permutation and (possibly negative) scaling, as claimed.

While the source distributions might not be uniquely identifiable for arbitrary co-variance matrices $\mathbf{S}$, they are indeed uniquely identifiable upto permutation and scaling for the noisy ICA model with isotropic Gaussian noise co-variance. For more general conditions that guarantee uniqueness of source distributions, please see [25,26].

Corollary 2.2 If $\mathbf{S} = \sigma^2 \mathbf{I}_p$ and $\mathbf{S}_1 = \sigma_1^2 \mathbf{I}_p$, then the source distributions are uniquely identifiable upto sign flips for $\Lambda^2 = \mathbf{I}_q$.

Proof Suppose $\Sigma = \sigma^2 \mathbf{I}_p$ and $\Sigma_1 = \sigma_1^2 \mathbf{I}_p$. Then from 2.9 $\Sigma = \Sigma_1$ and thus $\text{diag}(\mathbf{W}\Sigma\mathbf{W}^T) = \text{diag}(\mathbf{W}\Sigma_1\mathbf{W}^T)$. The corollary then follows from Claim 2.1.

Corollary 2.3 If $\mathbf{D} = \mathbf{D}_1 = \mathbf{I}_q$, then the source distributions are uniquely identifiable upto sign flips for $\Lambda^2 = \mathbf{I}_q$.

Proof If $\mathbf{D} = \mathbf{D}_1 = \mathbf{I}_q$, then noting that $\mathbf{P}\mathbf{P}^T = \mathbf{I}_q$, we get $\mathbf{D} = \mathbf{P}\mathbf{D}_1\mathbf{P}^T$. Hence from 2.12, we get $\text{diag}(\mathbf{W}\Sigma\mathbf{W}^T) = \text{diag}(\mathbf{W}\Sigma_1\mathbf{W}^T)$. The corollary then follows from Claim 2.1.

2.1.3 Why is there a run-to-run variability in estimated ICs? From the discussion in section 2.1.2, it is clear that for a noisy ICA model with isotropic diagonal additive Gaussian noise co-variance:

1. The noisy ICA parameters $\mathcal{F} = \{\mu, \mathbf{A}, \mathbf{D}, \Sigma\}$ are uniquely identifiable up to permutation and scaling.
2. The source distributions in $\mathbf{s}$ are uniquely identifiable upto permutation and scaling.

While the above theoretical properties of ICA are reassuring, there are a number of practical difficulties that prevent the reliable estimation of ICs on real data:

1. **Validity of the ICA model:** The assumption that the observed real data is generated by an ICA model is only that - an “assumption”. If this assumption is not valid, then the uniqueness results do not hold anymore.
2. **Mutual information approximations:** From an information theoretic point of view, the ICA problem is solved by minimizing a contrast function which is an *approximation* to the mutual information [27] between the ICs that depends on the finite amount of observed data. Such an approximation is necessary, since we do not have access to the marginal source densities p_{s_i} . Different approximations to mutual information will lead to different objective functions and hence different solutions. This is one of the reasons why different ICA algorithms often produce different IC estimates even for the same data.
3. **Non-convexity of ICA objective functions:** The ICA contrast function is potentially non-convex and hence has multiple local minima. Since global minimization is a challenging problem by itself, most ICA algorithms will only converge to local minima of the ICA contrast function. The run-to-run variability of IC estimates will also depend on the number of local minima in a particular ICA contrast function.
4. **IC estimate corruption by Gaussian noise:** For noisy ICA, the IC realizations cannot be recovered exactly even if the true mixing matrix $\mathbf{A}$ and mean vector μ are known in 2.1. Commonly used estimators for recovering realization of ICs include the least squares [10] as well as the minimum mean square error (MMSE) [14]. Consider the least squares estimate $\hat{\mathbf{s}}$ of a realization of $\mathbf{s}$ based on $\mathbf{y}$:

$$\hat{\mathbf{s}} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T (\mathbf{y} - \mu) = \mathbf{s} + (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \eta \quad (2.31)$$

This means that even for known parameters, IC realization estimates $\hat{\mathbf{s}}$ will be corrupted by correlated Gaussian noise.

Hence using different subsets of the data under the true model will also lead to variability in estimated ICs.

5. **Over-fitting of the ICA model:** Over specification of the model order leads to the problem of over-fitting in ICA. As we describe below, this can lead to (1) the phenomenon of IC “splitting” and (2) an increase in the variance of the IC estimates.

1. IC “splitting”

Suppose that the true model order or the number of non-Gaussian sources in an ICA decomposition of $\mathbf{y}$ such as 2.1 is q . Then a fundamental result in [12, Theorem 1] states that for any other ICA decomposition of $\mathbf{y}$, the number of non-Gaussian sources remains the same while the number of Gaussian sources can change. In other words, $\mathbf{y}$ cannot have two different ICA decompositions containing different number of non-Gaussian sources.

In view of this fact, how can a model order q ICA decomposition containing q non-Gaussian sources be “split” into a $(q+1)$ ICA decomposition containing $(q+1)$ non-Gaussian sources when performing ICA estimation using an assumed model order of $(q+1)$? As we describe below, the order $(q+1)$ ICA decomposition is only an *approximation* to the order q ICA decomposition.

Let $\mathbf{a}_i$ be the i th column of $\mathbf{A}$ in 2.1. In the presence of noise, it might be possible to *approximate*:

$$\mathbf{a}_i s_i \approx \mathbf{a}_i^1 s_i^1 + \mathbf{a}_i^2 s_i^2 \quad (2.32)$$

Here:

- $\mathbf{a}_i s_i$ is the contribution of the i th non-Gaussian source s_i to the ICA model 2.1.
- s_i^1 and s_i^2 are independent non-Gaussian random variables that are also independent with respect to all non-Gaussian sources $s_j, j \neq i$ in 2.1.
- $\mathbf{a}_i^1$ and $\mathbf{a}_i^2$ are the basis time courses corresponding to s_i^1 and s_i^2 respectively.
- The time courses $\mathbf{a}_i^1$ and $\mathbf{a}_i^2$ look similar to each other.

Note that if $\mathbf{a}_i^1 = \mathbf{a}_i^2$, then 2.32 can be made into an equality by choosing $s_i = s_i^1 + s_i^2$. By replacing $\mathbf{a}_i s_i$ in 2.1 using 2.32, we arrive at an *approximate* model order $(q+1)$ decomposition of $\mathbf{y}$. In this decomposition, the component s_i from a model order q decomposition appears to be “split” into two sub-components: s_i^1 and s_i^2 .

2. Inflated variance of IC estimates

Overestimation of model order will lead to over-fitting of the mixing matrix $\mathbf{A}$. In other words, $\mathbf{A}$ could have several columns that are highly correlated with each other. This could happen as a result of IC “splitting” as discussed above. Now, for a given realization $\mathbf{s}$, the variance of $\hat{\mathbf{s}}$ is given by $\text{Var}(\hat{\mathbf{s}}) = \sigma^2 (\mathbf{A}^T \mathbf{A})^{-1}$ (for isotropic Gaussian co-variance). An increase in number of columns of $\mathbf{A}$ and the fact that many of them are highly correlated implies that the variability of IC estimates $\text{Var}(\hat{\mathbf{s}})$ is inflated.

In other words, running ICA multiple times on the same data or variations thereof with random initialization could produce different ICs.

2.2 ICA algorithms, single subject ICA and group ICA

In this section, we give a brief summary of how the ICA parameters are estimated in practice and also summarize the two most common modes of ICA application to fMRI data - single subject ICA (section 2.2.1) and temporal concatenation based group ICA (section 2.2.2).

Given several independent observations $\mathbf{y}$ as per the noisy ICA model 2.1, most ICA algorithms estimate the ICA parameters $\mathcal{F} = \{\mu, \mathbf{A}, \mathbf{D}, \mathbf{\Sigma}\}$ and the realizations of $\mathbf{s}$ in 2 steps. We only consider the case with $\mathbf{\Sigma} = \sigma^2 \mathbf{I}_p$, since as shown in section 2.1.2, the mixing matrix $\mathbf{A}$ and source distributions of $\mathbf{s}$ are identifiable upto permutation and scaling for this case.

1. First, the diagonal source co-variance is arbitrarily set as $\mathbf{D} = \mathbf{I}_q$. The mean vector $\mathbf{m}$ is estimated as $\mathbf{E}(\mathbf{y})$. Then, using PCA or PPCA [19], the mixing matrix $\mathbf{A}$ is estimated, upto an orthogonal rotation matrix $\mathbf{O}$, to be in a signal subspace which is spanned by the principal eigenvectors corresponding to the largest eigenvalues of the data co-variance matrix $\mathbf{E}[(\mathbf{y} - \mu)(\mathbf{y} - \mu)^T]$. The noise variance σ^2 is estimated in this step as well.
2. Next, an estimator $\hat{\mathbf{s}}$ for the source realizations is defined using techniques such as least squares or MMSE. The only unknown involved in these estimates is the orthogonal rotation matrix $\mathbf{O}$.
3. Finally, the non-Gaussianity of the empirical density of components of $\hat{\mathbf{s}}$ is optimized with respect to $\mathbf{O}$ using algorithms such as fixed point ICA [27,28].

For more details on noisy ICA estimation, please see [10] and for more details on ICA algorithms, please see [29].

2.2.1 Single subject ICA. How is ICA applied to single subject fMRI data? Suppose we are given a single subject fMRI scan which we rearrange as a $p \times n$ 2D matrix $\mathbf{Y}$ in which column i is the $p \times 1$ observed time-course $\mathbf{y}_i$ in the brain at voxel i . Observed time-courses $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n$ are considered to be n independent realizations of $\mathbf{y}$ as per the linear ICA model 2.1. Suppose $\hat{\mathbf{S}} = [\hat{\mathbf{s}}_1, \hat{\mathbf{s}}_2, \dots, \hat{\mathbf{s}}_n]$ is the $q \times n$ matrix containing the estimated source realizations at the n voxels. The j th row of $\hat{\mathbf{S}}$ is the j th IC. In other words, we decompose the time by space fMRI 2D matrix into a set of basis time-courses and a set of q 3D IC maps using ICA.

2.2.2 Group ICA. How is ICA applied to data from a group of subjects in fMRI? Suppose we collect fMRI images from m subjects. First, we register all subjects to a common space using a registration algorithm (e.g., affine registration). Next, we rearrange each of the fMRI scans into m 2D matrices $\mathbf{Y}_1 \dots \mathbf{Y}_m$, each of size $p \times n$. Column j in $\mathbf{Y}_i$ is the demeaned time-course observed at voxel location j for subject i . The matrices $\mathbf{Y}_1 \dots \mathbf{Y}_m$ are temporally concatenated to get a $pm \times n$ matrix $\mathbf{Z}$ as follows:

$$\mathbf{Z} = \begin{pmatrix} \mathbf{Y}_1 \\ \vdots \\ \mathbf{Y}_i \\ \vdots \\ \mathbf{Y}_m \end{pmatrix} \quad (2.33)$$

Column i of $\mathbf{Z}$ is the $pm \times 1$ vector $\mathbf{z}_i$ which is assumed to follow a linear ICA model 2.1. $\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_n$ are considered to be independent realizations of the model 2.1. Suppose $\hat{\mathbf{S}}_G = [\hat{\mathbf{s}}_1, \hat{\mathbf{s}}_2, \dots, \hat{\mathbf{s}}_n]$ is a $q \times n$ matrix containing the estimated source realizations at the n voxels. The j th row of $\hat{\mathbf{S}}_G$ is the j th group IC. In group ICA, the joined time-series across subjects is modeled using noisy linear ICA. In practice, $\mathbf{Y}_i$ is the PCA reduced data set for subject i . The PCA reduction is either done separately for each subject using subject specific data co-variance [9] or an average data co-variance across subjects [8]. The average co-variance approach requires each subject to have the same number of time points in fMRI scans.

2.3 The original RAICAR algorithm

In this section, we give a brief introduction to the RAICAR algorithm of [1]. Suppose we are given a data set which we decompose into n_C ICs using ICA (e.g., single subject or group ICA). Our goal is to assess which ICs consistently show up in multiple ICA runs i.e., the reproducibility of each of these n_C ICs. To that extent, we run the ICA algorithm K times. Suppose $\mathbf{x}_j^{(m)}$ is the $n \times 1$ vector (e.g. spatial ICA map re-arranged into a vector) of the j th IC from m th ICA run. Suppose $\mathbf{G}_{lm}$ is a $n_C \times n_C$ absolute spatial cross-correlation coefficient matrix between the ICs from runs l and m :

$$\mathbf{G}_{lm}(i,j) = |\text{corrcoef}(\mathbf{x}_i^{(l)}, \mathbf{x}_j^{(m)})| \quad (2.34)$$

where $|\cdot|$ denotes absolute value. $\mathbf{G}_{lm}(i,j)$ is the absolute spatial cross-correlation coefficient between IC i from run l and IC j from run m . The matrices $\mathbf{G}_{lm}$ are then arranged as elements of a $K \times K$ block-matrix $\mathbf{G}$ such that the l th row and m th column of $\mathbf{G}$ is $\mathbf{G}(l,m) = \mathbf{G}_{lm}$ (Figure 2). This block matrix $\mathbf{G}$ is the starting point for a RAICAR across-run component matching process.

Since ICs within a particular run cannot be matched to each other, the $n_C \times n_C$ matrices $\mathbf{G}(l,l), l=1 \dots K$ along the block-diagonal of $\mathbf{G}$ are set to 0 as shown in Figure 2 with a gray color. The following steps are involved in a RAICAR analysis:

1. Find the maximal element of $\mathbf{G}$. Suppose this maximum occurs in matrix $\mathbf{G}_{lm}$ at position (i,j) . Hence component i from run l matches component j from run m . Let us label this matched component by MC_1 (the first matched component).
2. Next, we attempt to find from each run s ($s \neq l$ and $s \neq m$) a component that matches with component MC_1 . Suppose element (a_s, j) is the maximal element in the j th column of $\mathbf{G}_{sm}$. Then component a_s is the best matching component from run s with the j th component from run m .

Similarly, suppose element (i, b_s) is the maximal element in the i th row of $\mathbf{G}_{ls}$. Then component b_s is the best matching component from run s with component i from run l . As noted in [1], in most cases $a_s = b_s$. However, it is possible that $a_s \neq b_s$. Hence the component number e_s matching MC_1 from run s is defined as follows:

$$e_s = \begin{cases} a_s & \text{if } \mathbf{G}_{sm}(a_s, j) \geq \mathbf{G}_{ls}(i, b_s), \\ b_s & \text{if } \mathbf{G}_{sm}(a_s, j) < \mathbf{G}_{ls}(i, b_s). \end{cases} \quad (2.35)$$

We would also like to remove component e_s of run s from further consideration during the matching process. To that extent, we zero out the e_s th row from $\mathbf{G}_{sr}, r=1 \dots K$ and the e_s th column from $\mathbf{G}_{rs}, r=1 \dots K$.

3. Once a matching component e_s has been found for all runs $s \neq l, m$, we also zero out the i th row from $\mathbf{G}_{lr}, r=1 \dots K$ and the i th column from $\mathbf{G}_{rl}, r=1 \dots K$. Similarly, we zero out the j th column from $\mathbf{G}_{mr}, r=1 \dots K$ and the j th row from $\mathbf{G}_{mr}, r=1 \dots K$. This eliminates component i from run l and component j from run m from further consideration during the matching process.
4. Steps 1–3 complete the matching process for one IC component across runs. These steps are repeated until n_C components are matched across the K runs. We label the matched component s as MC_s which contains a set of K matching ICs one from each of the K ICA runs.

Figure 2. Pictorial depiction of the original RAICAR algorithm [1]. The ICA algorithm is run K times with each run producing n_C ICs. $\mathbf{G}$ is a $K \times K$ block matrix with elements $\mathbf{G}(l, m) = \mathbf{G}_{lm}$ where $\mathbf{G}_{lm}$ is the $n_C \times n_C$ absolute spatial cross-correlation matrix between ICs from runs l and m . The numbered green circles indicate the sequence of steps in applying RAICAR to a given data set. Our definition of normalized reproducibility in box 7 averages un-thresholded correlation coefficients thereby avoiding the selection of a correlation coefficient threshold prior to averaging. doi:10.1371/journal.pone.0027594.g002

Suppose matched component s , MC_s consists of the matched ICs $\mathbf{x}_{i_1}^{(1)}, \mathbf{x}_{i_2}^{(2)}, \dots, \mathbf{x}_{i_K}^{(K)}$. Form the $K \times K$ cross-correlation matrix H_{MC_s} between the matched components in MC_s . The (a, b) th element of this matrix is simply:

$$H_{MC_s}(a, b) = |\text{corrcoef}(\mathbf{x}_{i_a}^{(a)}, \mathbf{x}_{i_b}^{(b)})| \quad (2.36)$$

The normalized reproducibility of MC_s is then defined as:

$$\text{Reproducibility}(MC_s) = \left(\frac{2}{(K-1)K} \right) \sum_{a=1}^K \sum_{b=a+1}^K H_{MC_s}(a, b) \quad (2.37)$$

The double sum in 2.37 is simply the sum of the upper triangular part of H_{MC_s} excluding the diagonal. The normalizing factor $\frac{(K-1)K}{2}$ is simply the maximum possible value of this sum.

Hence the normalized reproducibility satisfies: $\text{Reproducibility}(MC_s) \leq 1$.

Note that our definition of normalized reproducibility is slightly different from that in [1]. Whereas [1] averages the *thresholded* absolute correlation coefficients, we simply average the *un-thresholded* absolute correlation coefficients to compute reproducibility thereby avoiding the selection of a threshold on the absolute correlation coefficients.

2.4 The RAICAR-N enhancement

In this section, we describe how to compute reproducibility p -values for each matched component in RAICAR. Note that the RAICAR “component matching” process can be used to assess the reproducibility of *any* spatial component maps - not necessarily ICA maps. For instance, RAICAR can be used to assess the reproducibility of a set of PCA maps across subjects.

In order to generate reproducibility p -values for the matched component maps:

1. We need to determine the distribution of normalized reproducibility that we get from the RAICAR “component matching” process when the input to RAICAR represents a set of “non-reproducible component maps” across the K runs.
2. In addition, we would also like to preserve the *overall structure* seen in the *observed* sets of spatial component maps across the K runs when generating sets of “non-reproducible component maps” across the K runs.

Hence for IC reproducibility assessment, we propose to use the original set of ICs across the K runs to generate the “non-reproducible component maps” across the K runs.

Suppose K ICA runs are submitted to RAICAR which gives us a $n_C \times 1$ vector of *observed* normalized reproducibility values $\text{Reproducibility}(MC_i), i = 1 \dots n_C$ - one for each IC. We propose to attach p -values for measuring the reproducibility of each IC in a data-driven fashion as follows:

1. First, we label the Kn_C ICs across the K runs using unique integers. In run 1, the ICs are labelled using integers $1, \dots, n_C$. In run 2, the ICs are labelled using integers $(n_C + 1), \dots, 2n_C$ and so on. In run K , the ICs are labelled using integers $(K-1)n_C + 1, \dots, Kn_C$.
2. Our “null” hypothesis is:

$$\mathbf{H}_0 : \text{None of the ICs are reproducible} \quad (2.38)$$

To do this, we randomly permute the integers $1, 2, \dots, Kn_C$ to get the permuted integers $p(1), \dots, p(Kn_C)$. Obviously $p(i) \neq p(j)$ if $i \neq j$.

3. The K sets “non-reproducible component runs under $\mathbf{H}_0$ ” are constructed by assigning components with labels:
 - $p(1), \dots, p(n_C)$ to run 1 under $\mathbf{H}_0$.
 - $p(n_C + 1), \dots, p(2n_C)$ to run 2 under $\mathbf{H}_0$
 - $p((K-1)n_C + 1), \dots, p(Kn_C)$ to run K under $\mathbf{H}_0$
4. After K runs have been generated under $\mathbf{H}_0$, we subject these to a RAICAR analysis. This gives us n_C values of normalized reproducibility, one for each matched component under $\mathbf{H}_0$.
5. Steps 1–4 are repeated R times to build up a *pooled* $Rn_C \times 1$ vector of normalized reproducibility $\text{Reproducibility}_{Null}$ under $\mathbf{H}_0$.
6. Finally, we assign a p -value for reproducibility to each matched IC across the K runs. The observed reproducibility for i th matched IC is $\text{Reproducibility}(MC_i)$ and its p -value is:

$$\text{Reproducibility}_{pval}(MC_i) = \frac{\{\text{no. of } \text{Reproducibility}_{Null} \geq \text{Reproducibility}(MC_i)\} + 1}{Rn_C + 1} \quad (2.39)$$

7. Only those components with $\text{Reproducibility}_{pval}(MC_i) < p_{crit}$ are considered to be significantly reproducible. We can use a fixed and objective value for p_{crit} such as 0.05. Note that this fixed cutoff is independent of the amount of variability in the input to RAICAR-N. Please see Figure 3 for a pictorial depiction of this process.

Figure 4 shows a schematic of the RAICAR-N analysis process.

2.5 How many subjects should be used per group ICA run in RAICAR-N?

The input to RAICAR-N can either be single subject ICA runs or group ICA runs across a set of subjects. Note that the individual subject ICA runs are spatially unconstrained whereas a group ICA spatially constrains the group ICs across a set of subjects. Hence the number of ICs that can be declared as significantly reproducible at the group level are usually more than those that can be declared significantly reproducible at the single subject level. Hence the following question is relevant:

Suppose we have a group of N subjects. We randomly select L subjects and form a single group of subjects. We repeat this process K times to get K groups of L subjects each of which is subjected to a group ICA analysis. Given the number of subjects N , how should we choose L and K ?

First, we discuss the choice of L . If $L = N$ then each of the K groups will contain the same N subjects and hence there will be no diversity in the K groups. We would like to control the amount of diversity in the K groups of L subjects. Consider any 2 subjects X and Y . The probability $P_{XY}(L)$ that both X and Y appear in a set of L randomly chosen subjects from N subjects is given by:

$$P_{XY}(L) = \frac{\binom{N-2}{L-2}}{\binom{N}{L}} \quad (2.40)$$

The expected number of times that X and Y appear together in sets of L subjects out of K independently drawn sets is:

$$E_{XY}(L) = K P_{XY}(L) \quad (2.41)$$

Ideally, we would like $E_{XY}(L)$ to be only a small fraction of K . Hence we impose the restriction:

$$E_{XY}(L) = K P_{XY}(L) \leq \alpha_{max} K \quad (2.42)$$

where α_{max} is a user defined constant such as $\alpha_{max} = 0.05$. This implies that the chosen value of L must satisfy:

$$P_{XY}(L) \leq \alpha_{max} \quad (2.43)$$

In practice, we choose the largest value of L that satisfies this inequality. As shown in Figure 5, if $N = 23$ and $\alpha_{max} = 0.05$ then the largest value of L that satisfies 2.43 is $L = 5$. The number of group ICA runs K should be as large as possible. From our experiments on real fMRI data we can roughly say that values of $K > 50$ give equivalent results.

2.6 How to display the estimated non-Gaussian spatial structure in ICA maps?

The ICs have been optimized for non-Gaussianity. However, there can be many types of non-Gaussian distributions. It has been empirically found that the non-Gaussian distributions of ICs found in fMRI data have the following structure:

8

Once we calculate Reproducibility(MC_1), Reproducibility(MC_2), ..., Reproducibility(MC_{n_C}), how can we calculate a cutoff to separate the high reproducibility ICs from the low reproducibility ICs?

unique integer associated with each component in each run

9

$1 \dots n_C$ run 1 $[n_C + 1] \dots 2n_C$ run 2 ... $[(K-1)n_C + 1] \dots Kn_C$ run K

291.432

10

H_0 : Null hypothesis is that none of the ICs are reproducible. Hence we can randomly label IC i from run l as IC d from run s . This random labelling produces one realization of ICs under H_0 .

random labelling

integers $1, 2, \dots, Kn_C$ $\xrightarrow{\text{random permutation of integers}} p(i) \neq p(j) \text{ if } i \neq j$ $\rightarrow$ permuted integers $p(1), p(2), \dots, p(Kn_C)$

11

$p(1) \dots p(n_C)$ run 1 | H_0 $p(n_C + 1) \dots p(2n_C)$ run 2 | H_0 ... $p((K-1)n_C + 1) \dots p(Kn_C)$ run K | H_0

reproducibility calculation for 1 realization of ICs under H_0

12

- We repeat the reproducibility calculation for R realizations of ICs (e.g. $R = 100$) under H_0 using the random labelling process described above.
- This gives us a distribution of "Reproducibility" values under H_0 . Each realization of H_0 produces n_C "null" reproducibility values.
- Hence for R realizations we get $R \times n_C$ values that define the null distribution for testing the observed reproducibility values Reproducibility(MC_i).
- Denote the vector of this "null" reproducibility values by Reproducibility_{Null}

Calculate p -values for Reproducibility

13

Figure 3. Pictorial depiction of the process for generating a "null" distribution in RAICAR-N. Our "null" hypothesis is: " H_0 : None of the ICs are reproducible. Hence, we can randomly label IC i from run l as IC d from run s ". Therefore we randomly split the Kn_C ICs across K runs into K parts and run the RAICAR algorithm on each set of randomly split ICs. This gives us a set of "null" reproducibility values which can be used to compute p -values for the observed reproducibility of ICs in the original RAICAR run. The green circles indicate the sequence of steps for generating the "null" distribution after the steps in Figure 2.

doi:10.1371/journal.pone.0027594.g003

Figure 4. Flowchart for a group ICA based RAICAR-N analysis. The N single subject data sets are first pre-processed and subsequently bootstrapped to create K groups, each group containing L distinct subjects. Each group of L subjects is submitted to a temporal concatenation group ICA analysis. The resulting IC maps (either raw ICs or ICs scaled by noise standard deviation) are subjected to a RAICAR analysis. The cross-realization cross correlation matrix (CRCM) is randomly permuted multiple times: $G \rightarrow G(g)$ where g is a random permutation of integers from $1, \dots, Kn_C$. The permuted CRCMs are subjected to a RAICAR analysis to generate a realization of reproducibility values under the “null” hypothesis. The computed “null” distribution of reproducibility values is used to assign p values to the observed reproducibility of the original RAICAR run. Finally, reproducible ICs are averaged using a random effects analysis and the resulting t -statistic images are subjected to Gamma_{neg}, Student t and Gamma_{pos} mixture modeling.
doi:10.1371/journal.pone.0027594.g004

1. A central Gaussian looking part and
2. A tail that extends out on either end of the Gaussian

It has been suggested in [10] that a Gaussian/Gamma mixture model can be fitted to this distribution and the Gamma components can be thought of as representatives of the non-Gaussian structure. We follow a similar approach:

1. The output of a RAICAR-N analysis is a set of spatial ICA maps (either z -transformed maps or raw maps) concatenated into a 4-D volume.
2. We do a voxelwise transformation to Normality using the voxelwise empirical cumulative distribution function as described in [30].

3. Next, we submit the resulting 4-D volume to a voxelwise group analysis using ordinary least squares. The design matrix for group analysis depends on the question being considered. In our case, the design matrix was simply a single group average design.
4. The resulting t -statistic maps are subjected to Student t , Gamma_{pos} and Gamma_{neg} mixture modeling. The logic is that if the original ICA maps are pure Gaussian (i.e., have no interesting non-Gaussian structure) then the result of a group average analysis will be a pure Student t map which will be captured by a single Student t (i.e., the Gamma_{pos} and Gamma_{neg} will be driven to 0 class fractions). Hence the “null” hypothesis will be correctly accounted for.

Figure 5. Figure shows a plot of $P_{XY}(L)$ vs L for $N=23$ in blue. The red line shows the $\alpha_{max}=0.05$ cutoff. The largest value of L for which $P_{XY}(L) \leq 0.05$ is $L=5$.

doi:10.1371/journal.pone.0027594.g005

5. If the Gamma distributions have >0.5 posterior probability at some voxels then those voxels are displayed in color to indicate the presence of significant non-Gaussian structure over and above the background Student t distribution.

Examples of Student t , Gamma_{pos} and Gamma_{neg} mixture model fits are shown in Figure 6.

Results

3.1 Human rsfMRI data

rsfMRI data titled: Baltimore (Pekar, J.J./Mostofsky, S.H.; $n=23$ [8M/15F]; ages: 20–40; TR = 2.5; # slices = 47; # timepoints = 123), a part of the 1000 functional connectomes project, was downloaded from the Neuroimaging Informatics Tools and Resources Clearinghouse (NITRC): http://www.nitrc.org/projects/fcon_1000/.

3.2 Preprocessing

Data was analyzed using tools from the FMRIB software library (FSL: <http://www.fmrib.ox.ac.uk/fsl/>). Preprocessing steps included motion correction, brain extraction, spatial smoothing with an isotropic Gaussian kernel of 5 mm FWHM and 100 s high-pass temporal filtering. Spatial ICA was performed using a noisy ICA model as implemented in FSL MELODIC [10] in either single subject or multi-subject temporal concatenation mode also called group ICA. Please see section 2.2 for a brief summary of single subject ICA and group ICA. In each case, we fixed the model order of ICA at $q=40$ to be consistent with the model order range typically extracted in rsfMRI and fMRI [16,31]. For temporal concatenation based group ICA, single subject data was first affinely registered to the MNI 152 brain and subsequently

resampled to $4 \times 4 \times 4$ resolution (MNI $4 \times 4 \times 4$) to decrease computational load. Please see Figure 4 for a schematic of the RAICAR-N analysis process. In this work, we report across subject RAICAR-N analyses, but as shown in Figure 7, within subject ICA runs can also be entered into RAICAR-N.

3.3 RAICAR-N analysis with 1 ICA run per subject

Spatial ICA was run once for each of the $N=23$ subjects in their native space. The resulting set of ICA components across subjects were transformed to MNI $4 \times 4 \times 4$ space and were submitted to a RAICAR-N analysis. In all RAICAR-N analyses reported in this article, we used the z -transformed IC maps - which are basically the raw IC maps divided by a voxelwise estimate of noise standard deviation (named as `melodic_IC.nii.gz` in MELODIC). It is also possible to use the raw IC maps as inputs to RAICAR-N. ICA components were sorted according to their reproducibility and p -values were computed for each ICA component. Please see Figure 8.

We compared the reproducible RSNs from the single subject RAICAR-N analysis to the group RSN maps reported in literature [8]. Please see Figure 9.

To summarize, when single subject ICA runs are combined across subjects:

- We are able to declare 4 “standard” RSNs as significantly reproducible at a p -value <0.05 .
- There are 2 other “standard” RSNs that achieve a reproducibility p -value between 0.05 and 0.06.
- There are 2 other “non-standard” RSNs that are of interest: one achieves a p -value of 0.0125 and the other achieves a p -value of 0.05699.

Figure 6. Examples of displaying non-Gaussian spatial structure using a Student t , Γ_{pos} and Γ_{neg} mixture model. Notice how the Γ_{neg} density is driven to near 0 class fraction in the absence of significant negative non-Gaussian structure.
doi:10.1371/journal.pone.0027594.g006

3.4 RAICAR-N on random sets of 5 subjects - 50 group ICA runs

To promote diversity across the group ICA runs, as discussed in section 2.5, $L=5$ subjects were drawn at random from the group of $N=23$ subjects and submitted to a temporal concatenation based group ICA. This process was repeated $K=50$ times and the resulting set of 50 group ICA maps were submitted to a RAICAR-N analysis. ICA components were sorted according to their reproducibility and p -values were computed for each ICA component. Please see Figure 10.

We compared the reproducible RSNs from the single subject RAICAR-N analysis to the RSN maps reported in literature [8]. Please see Figure 11.

In summary, when 50 random 5 subject group ICA runs (from a population of 23 subjects) are combined using RAICAR-N:

- We are able to declare 8 “standard” RSNs as significantly reproducible at a p -value <0.05 .
- There are 6 other “non-standard” RSNs that can be declared as significantly reproducible at a p -value <0.05 .
- There is 1 other “non-standard” RSN that achieves a p -value of 0.05299.

3.5 RAICAR-N on random sets of 5 subjects - 100 group ICA runs

To promote diversity across the group ICA runs, as discussed in section 2.5, $L=5$ subjects were drawn at random from the group of $N=23$ subjects and submitted to a temporal concatenation based group ICA. This process was repeated $K=100$ times and the resulting set of 100 group ICA maps were submitted to a RAICAR-N analysis. ICA components were sorted according to

Figure 7. p -value cutoffs for within and across single subject analysis using RAICAR-N. This figure illustrates the intuitive fact that within subject ICA runs are much more reproducible compared to across subject ICA runs.
doi:10.1371/journal.pone.0027594.g007

Figure 8. Single subject rsfMRI ICA runs across 23 subjects were combined using a RAICAR-N analysis. Figure (a) shows the observed values of normalized reproducibility (bottom) as well as the “null” distribution of normalized reproducibility across $R = 100$ simulations (top). Figure (b) shows the p -values for each IC along with the 0.05 and 0.1 cutoff lines.
doi:10.1371/journal.pone.0027594.g008

their reproducibility and p -values were computed for each ICA component. Please see Figure 12.

We compared the reproducible RSNs from the single subject RAICAR-N analysis to the RSN maps reported in literature [8]. Please see Figure 13.

In summary, when 100 random 5 subject group ICA runs (from a population of 23 subjects) are combined using RAICAR-N:

- We are able to declare 8 “standard” RSNs as significantly reproducible at a p -value < 0.05 .
- There are 6 other “non-standard” RSNs that can be declared as significantly reproducible at a p -value < 0.05 .
- There is 1 other “non-standard” RSN that achieves a p -value of 0.05824.

Discussion

As discussed in section 2.1.2, in the noisy linear ICA model with isotropic diagonal Gaussian noise co-variance, for a given true model order, the mixing matrix and the source distributions are identifiable up to permutation and scaling. However, as pointed out in section 2.1.3, various factors prevent the convergence of ICA algorithms to unique IC estimates. These factors include ICA model not being the true data generating model, approximations to mutual information used in ICA algorithms, multiple local minima in ICA contrast functions, confounding Gaussian noise as well as variability due to model order over-estimation. A practical implication of these factors is that ICA algorithms converge to different IC estimates depending on how they are initialized and on the specific data used as input to ICA. Hence, there is a need for a rigorous assessment of reproducibility or generalizability of IC estimates. A set of reproducible ICs can then be used as ICA based characteristics of a particular group of subjects.

We proposed an extension to the original RAICAR algorithm for reproducibility assessment of ICs within or across subjects (Figure 7). The modified algorithm called RAICAR-N builds up a “null” distribution of normalized reproducibility values under a

random assignment of observed ICs across the K runs. This “null” distribution is used to compute reproducibility p -values for each observed matched component from RAICAR. An objective cutoff such as $p < 0.05$ can be used to detect “significantly reproducible” components. This avoids subjective user decisions such as selection of the number of clusters in ICASSO or the reproducibility cutoff in RAICAR or a cutoff on intra cluster distance in sogICA.

4.1 Results for publicly available rsfMRI data

We applied RAICAR-N to publicly available $N = 23$ subject rsfMRI data from <http://www.nitrc.org/>. We analyzed the data in 2 different ways:

1. $n_C = 40$ ICs were extracted for each of the $N = 23$ subjects. The $K = 23$ single subject ICA runs were subjected to a RAICAR-N analysis (after registration to standard space). In single subject ICA based RAICAR-N analysis (Figures 8, 9), we are able to declare 6 out of the 8 ICs reported in [8] (which used group ICA) as “reproducible” (4 ICs have p -values < 0.05 and 2 ICs have p -values < 0.06). This is consistent with the 5 reproducible RSNs reported in [32] using single subject ICA analysis.
2. $L = 5$ subjects were randomly drawn from $N = 23$ subjects to create one group of subjects which was subjected to a group ICA analysis in which $n_C = 40$ components were extracted. This process was repeated $K = 50$ or 100 times and the resulting group ICA runs were subjected to a RAICAR-N analysis. In group ICA based RAICAR-N analysis (Figures 10, 11, 12, and 13), we are able to declare all 8 components reported in [8] as “reproducible” (at $p < 0.05$). Some of the ICs detected as “reproducible” in the group ICA based RAICAR-N on human rsfMRI data are not shown in [8] but do appear in the more recent paper [31]. RAICAR-N results for $K = 50$ are almost identical to those for $K = 100$ suggesting that $K = 50$ runs of group ICA are sufficient for a RAICAR-N reproducibility analysis.

Figure 9. The top 8 “reproducible” ICs from a RAICAR-N analysis on single subject ICA runs compared with standard RSN maps reported in literature [8]. We are able to declare 4 “standard” RSNs as significantly reproducible at a p -value < 0.05 . There are 2 other “standard” RSNs that achieve a reproducibility p -value between 0.05 and 0.06 as well as 2 “non-standard” RSNs that achieve p -values of 0.0125 and 0.05699 respectively. We also could not find 2 of the published RSNs in [3] as reproducible in single subject ICA runs.
doi:10.1371/journal.pone.0027594.g009

4.2 Single subject ICA vs Group ICA

Based on our results, it appears that single subject ICA maps are less reproducible compared to group ICA maps as illustrated in Figures 8 and 10. A single subject ICA based analysis is more resistant to subject specific artifacts. On the other hand, a group ICA based analysis makes the strong assumption that ICs are spatially identical across subjects. If this assumption is true, group ICA takes advantage of temporal concatenation to constrain the ICs spatially across subjects thereby reducing their variance. Hence, when there are no gross artifacts in individual rsfMRI data sets, group ICA is expected to be more sensitive for reproducible IC detection. As seen in Figures 9 and 11, our results agree with this proposition. All ICs declared as “reproducible” in the single subject based RAICAR-N analysis continue to remain “reproducible” in the group ICA based RAICAR-N analysis.

4.3 How should subjects be grouped for group ICA?

This raises the question of how the subjects should be grouped together for individual group ICA runs in preparation for

RAICAR-N. If all N subjects are used in all group ICA analyses then there is no diversity in the individual group ICA runs. In this case, a RAICAR-N analysis will capture algorithmic variability due to non-convexity of ICA objective function but not dataset variability. Hence, our conclusions might not be generalizable to a different set of N subjects.

Another option is to randomly select L subjects out of N for each group ICA run and submit the resulting K group ICA runs to RAICAR-N. In this case, we will account for both algorithmic and data set variability via a RAICAR-N analysis. In other words, we will be able to determine those ICs that are “reproducible” across different sets of L subjects and across multiple ICA runs. A key question is: How should we choose L and K ? In section 2.5, we proposed a simple method to determine the number of subjects L to be used in a single group ICA run out of the N subjects - the key idea is to form groups with enough “diversity”. Multiple such group ICA runs can then be submitted to a RAICAR-N analysis for reproducibility assessment.

Figure 10. $L=5$ subjects were randomly drawn from the set of $N=23$ subjects and submitted to a temporal concatenation based group ICA. This process was repeated $K=50$ times and the resulting ICA maps were submitted to a RAICAR-N analysis. Figure (a) shows the observed values of normalized reproducibility (bottom) as well as the “null” distribution of normalized reproducibility across $R=100$ simulations (top). Figure (b) shows the p -values for each IC along with the 0.05 and 0.1 cutoff lines.
doi:10.1371/journal.pone.0027594.g010

Clearly, the larger the value of N , the larger the value of L . Hence, increasing the number of subjects N in a study will allow us to make conclusions that are generalizable to a larger set of L subjects. Also, conclusions generalizable to L_1 subjects are expected to hold for $L_2 > L_1$ subjects but not vice versa.

4.4 RAICAR-N for group comparisons of reproducible ICs

In the present work, our focus was on enabling the selection of reproducible ICs for a given single group of subjects. However, RAICAR-N can be extended for between group analysis of reproducible components as well. Before we describe how to do so, it is useful to discuss other approaches for group analysis of RSNs described in **Appendix S1**. Suppose we have two groups of subjects A and B .

4.4.1 Discussion of single group ICA based approaches. 1. Subject specific maps corresponding to group ICA maps derived using ICA back projection or dual regression are not true ICs, i.e., they are not solutions to an ICA problem.

2. These approaches do not account for either the algorithmic or the data set variability of an ICA decomposition. The single group ICA decomposition will contain both reproducible and non-reproducible ICs, but there is no systematic way to differentiate between the two.

3. Both dual regression and ICA back projection using data derived IC templates are circular analyses. First, group ICA using **all** data is used to derive template IC maps or template time courses. Next least-squares based ICA back projection or dual regression using a subset of the same data is used to derive subject specific maps and time courses corresponding to each IC. Thus model 1 (group ICA) on data $\mathcal{D}$ is used to learn an assumption $\mathcal{A}$ (template IC maps or template time courses) that is then used to fit model 2 (dual regression or ICA back projection) on a subset of the **same** data $\mathcal{D}$. This is circular analysis [33,34].

It is easy to avoid circular analysis in a dual regression approach via cross-validation. For example, one can split the groups A and B into two random parts, a “training” set and a “test” set. First,

the “training” set can be used to derive template IC maps using group ICA. Next, the “training” set based template IC maps can be used as spatial regressors for dual regression on the “test” set. Alternatively, the template ICs for dual regression can also come from a separate ICA decomposition on an independent data set unrelated to groups A and B such as human rsfMRI data. This train/test approach cleanly avoids the circular analysis problem. It is not clear how to use cross-validation for an ICA back projection approach since template time courses cannot be assumed to remain the same across ICA decompositions.

4. Subject specific structured noise is quite variable in terms of its spatial structure. Hence, a group ICA analysis cannot easily model or account for subject specific structured noise via group level ICs. Consequently, subject specific spatial maps in ICA back projection or dual regression will have a noise component that is purely driven by the amount of structured noise in individual subjects. On the other hand, a single subject ICA based analysis can accurately model subject specific structured noise via single subject ICs.

4.4.2 Discussion of multiple ICA run approaches. 1. [35] report that using different sets of template ICs in template based methods using spatial correlation such as [36] can result in the selection of different ICs in individual ICA runs. This is not surprising since IC correspondence derived from template based methods does depend on the particular template used. This is similar to a seed based correlation analysis being dependent on the particular seed ROI used. It is worth noting that template free approaches such as sogICA and RAICAR do not rely on any template.

2. [22] state that individual runs across subjects (or groups of subjects) can be quite variable in terms of the spatial structure of the estimated ICs. For example [22], point out that an IC might be apparently split into two sub-components in some subjects but not others. The real problem is that the same model order could lead to over-fitting in some subjects (or groups of subjects) but not in others. Hence, the observed differences in a group comparison might be biased by the unknown difference in the amount of over-fitting across groups A and B .

Figure 11. The top 15 “reproducible” ICs from $K=50$ runs of $L=5$ subject group ICA RAICAR-N analysis compared with standard RSN maps reported in literature [8]. We are able to declare 8 “standard” RSNs as significantly reproducible at a p -value of <0.05 . There are 6 other “non-standard” RSNs that can be declared as significantly reproducible at a p -value of <0.05 and 1 other “non-standard” RSN that achieves a p -value of 0.05299.

doi:10.1371/journal.pone.0027594.g011

Figure 12. $L=5$ subjects were randomly drawn from the set of $N=23$ subjects and submitted to a temporal concatenation based group ICA. This process was repeated $K=100$ times and the resulting ICA maps were submitted to a RAICAR-N analysis. Figure (a) shows the observed values of normalized reproducibility (bottom) as well as the “null” distribution of normalized reproducibility across $R=100$ simulations (top). Figure (b) shows the p -values for each IC along with the 0.05 and 0.1 cutoff lines. doi:10.1371/journal.pone.0027594.g012

As described in 2.1.3, over-fitting can lead to the phenomenon of component “splitting” in ICA. This is not limited to single subject ICA but can also occur in group ICA. For instance [35], report the “default mode” network as split into three sub networks using group ICA and note that component “splitting” can also reflect functional segregation or hierarchy within a particular IC and is not necessarily a consequence of model order overestimation in every case.

Over-fitting can be correctly accounted for by a reproducibility analysis. This is because we expect the real and stable non-Gaussian sources to be reproducible across multiple ICA runs (algorithmic variability) and across different subjects or groups of subjects (data set variability).

If we want the results of a between group ICA analysis to be generalizable to an independent group of subjects then we must account for both the algorithmic and data variability of ICA. We propose to modify RAICAR-N for enabling between group comparisons of “reproducible” ICs as follows:

1. Enter multiple within and across subject (or within and across sets of subjects) ICA runs for groups A and B into a RAICAR analysis. Perform the RAICAR component matching process across groups A and B .
2. Use RAICAR-N to compute reproducibility p -values separately for group A and B for each matched component across groups A and B .
3. Only ICs that are *separately* reproducible in both groups A and B and that are *maximally* similar to each other are used for between group comparisons.

4.5 Limitations of a RAICAR-N analysis

In this work, we focussed on developing an objective method for reproducible IC detection across multiple ICA runs. However, our approach has some limitations that are worth mentioning:

1. RAICAR-N gives no indication to the extent of biological relevance of a reproducible IC. For example, an artifactual component consistently appearing across fMRI runs could be highly reproducible and yet non-interesting.

2. The p -values in RAICAR-N can be conservative. Certain borderline reproducible ICs may not achieve statistical significance in RAICAR-N due to the differences in the amount of structured noise in individual ICA decompositions. One approach to increase the sensitivity of RAICAR-N is to denoise the individual fMRI runs via ICA and remove gross artifacts prior to a RAICAR-N analysis.
3. RAICAR-N does not allow us to *relate* the reproducible ICs to each other. These relationships might be important for the identification of functionally related brain networks. A recent paper [37] proposes to use the mutual information between spatial ICs as a similarity measure for agglomerative hierarchical clustering of ICs. The number of IC clusters is decided using the well-known ANOVA based approach developed in [38]. Each IC cluster can then be thought of as a functionally related brain network.

RAICAR-N is useful for objective and non-parametric reproducibility assessment but does not attempt to relate the ICs to each other. On the other hand, the work by [37] uses a subjective definition of “qualified clusters” and a hard cutoff on the “quality index” to identify the *single most stable run* of ICA for IC clustering. RAICAR-N and the work by [37] have different primary objectives (IC reproducibility vs. IC clustering) and it appears that the benefits of both can be realized by feeding the reproducible ICs from RAICAR-N into the algorithm of [37].

To summarize, a RAICAR-N analysis:

- can be applied for “reproducible” component detection either within or across subjects in any component based analysis - not necessarily ICA. For instance, a set of PCA maps across subjects can be submitted to a RAICAR-N analysis.
- is simple to implement and accounts for both algorithmic and data set variability of an ICA decomposition.
- avoids any user decisions except the final p -value cutoff which can be objectively pre-set at standard values such as 0.05.
- can be extended to enable comparisons of reproducible ICs between groups A and B .

Figure 13. The top 15 “reproducible” ICs from $K = 100$ runs of $L = 5$ subject group ICA RAICAR-N analysis compared with standard RSN maps reported in literature [8]. We are able to declare 8 “standard” RSNs as significantly reproducible at a p -value of < 0.05 . There are 6 other “non-standard” RSNs that can be declared as significantly reproducible at a p -value of < 0.05 and 1 other “non-standard” RSN that achieves a p -value of 0.05824.

doi:10.1371/journal.pone.0027594.g013

Multiple group ICA runs using groups of subjects with enough “diversity” can be used to account for the run-to-run variability in ICA algorithms both due to the non-convex ICA objective function as well as across subjects data variability. These group ICA runs can be subjected to a RAICAR-N “reproducibility” analysis. RAICAR-N enables the objective detection of “reproducible components” in any component based analysis of fMRI data such as ICA and can also be used for a between group comparison of “reproducible” ICs.

References

- Yang Z, LaConte S, Weng X, Hu X (2008) Ranking and Averaging Independent Component Analysis by Reproducibility (RAICAR). *Human Brain Mapping* 29: 711–725.
- Jutten C, Herault J (1991) Blind separation of sources, Part I: An adaptive algorithm based on neuromimetic architecture. *Signal Processing* 24: 1–10.
- Comon P (1994) Independent Component Analysis, a new concept? *Signal Processing* 36: 287–314.
- Bell AJ, Sejnowski TJ (1995) An information-maximization approach to blind separation and blind deconvolution. *Neural Computation* 7: 1129–1159.
- Attias H (1999) Independent factor analysis. *Neural Computation* 11: 803–851.
- McKeown MJ, Makeig S, Brown GG, Jung TP, Kindermann SS, et al. (1998) Analysis of fMRI data by blind separation into independent spatial components. *Human Brain Mapping* 6: 160–188.
- Tohka J, Foerke K, Aron AR, Tom SM, Toga AW, et al. (2008) Automatic independent component labeling for artifact removal in fMRI. *NeuroImage* 39: 1227–1245.
- Beckmann CF, DeLuca M, Devlin JT, Smith SM (2005) Investigations into resting-state connectivity using independent component analysis. *Philos Trans R Soc Lond B (Biol Sci)* 360: 1001–1013.
- Calhoun VD, Adali T, Pearson GD, Pekar JJ (2001) A method for making group inferences from functional MRI data using independent component analysis. *Human Brain Mapping* 14: 140–151.
- Beckmann CF, Smith SM (2004) Probabilistic independent component analysis for functional magnetic resonance imaging. *IEEE Transactions on Medical Imaging* 23: 137–152.
- Cardoso J (1998) Blind signal separation: statistical principles. *Proceedings of the IEEE* 9: 2009–2025.
- Rao CR (1969) A Decomposition Theorem for Vector Variables with a Linear Structure. *The Annals of Mathematical Statistics* 40: 1845–1849.
- Theis FJ (2004) A New Concept for Separability Problems in Blind Source Separation. *Neural Computation* 16: 1827–1850.
- Davies M (2004) Identifiability Issues in Noisy ICA. *IEEE Signal Processing Letters* 11: 470–473.
- Himberg A, Hyvarinen A, Esposito F (2004) Validating the independent components of neuroimaging time-series via clustering and visualization. *NeuroImage* 22: 1214–1222.
- Esposito F, Scarabino T, Hyvarinen A, Himberg J, Formisano E, et al. (2005) Independent component analysis of fMRI group studies by self-organizing clustering. *NeuroImage* 25: 193–205.
- Kass RE, Raftery AE (1993) Bayes factors and model uncertainty. *Technical Report Report 254*, March 1993. Department of Statistics, University of Washington, USA.
- Kritchman S, Nadler B (2009) Non-Parametric Detection of the Number of Signals: Hypothesis Testing and Random Matrix Theory. *IEEE Transactions on Signal Processing* 57: 3930–3941.
- Tipping ME (1999) Probabilistic Principal Component Analysis. *Journal of the Royal Statistical Society Series B (Statistical Methodology)* 61: 611–622.
- Minka TP (2000) Automatic choice of dimensionality for PCA. *Technical Report 514*. MIT Media Laboratory Vision and Modeling Group, MIT.
- Hastie T, Tibshirani R, Friedman JH (2009) *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York: Springer.
- Cole DM, Smith SM, Beckmann CF (2010) Advances and Pitfalls in the Analysis and Interpretation of Resting-State fMRI Data. *Frontiers in Systems Neuroscience* 4.
- Feller W (1966) *An Introduction to Probability Theory*, Vol. II. New York: Wiley.
- Włodzimierz B (1995) *The normal distribution: Characterizations with applications*. New York: Springer-Verlag.
- Eriksson J, Koivunen V (2004) Identifiability, separability, and uniqueness of linear ICA models. *Signal Processing Letters, IEEE* 11: 601–604.
- Eriksson J, Koivunen V (2006) Complex random vectors and ICA models: identifiability, uniqueness, and separability. *IEEE Transactions on Information Theory* 52: 1017–1029.
- Hyvarinen A (1998) New approximations of differential entropy for independent component analysis and projection pursuit. *Advances in Neural Information Processing Systems* 10: 273–279.
- Hyvarinen A (1999) Fast and Robust Fixed-Point Algorithms for Independent Component Analysis. *IEEE Transactions on Neural Networks* 10(3): 626–634.
- Hyvarinen A, Karhunen J, Oja E (2001) *Independent Component Analysis* John Wiley and Sons.
- van Albada SJ, Robinson PA (2007) Transformation of arbitrary distributions to the normal distribution with application to EEG test-retest reliability. *Journal of Neuroscience Methods* 161: 205–211.
- Smith SM, Fox PT, Miller KL, Glahn DC, Fox PM, et al. (2009) Correspondence of the brain's functional architecture during activation and rest. *Proc Natl Acad Sci USA* 106: 13040–13045.
- DeLuca M, Beckmann CF, DeStefano N, Matthews PM, Smith SM (2005) fMRI resting state networks define distinct modes of long-distance interactions in the human brain. *NeuroImage* 29: 1359–1367.
- Kriegeskorte N, Simmons WK, Bellgowan PSF, Baker CI (2009) Circular analysis in systems neuroscience: the dangers of double dipping. *Nature Neuroscience* 12: 535–540.
- Vul E, Kanwisher N, Hanson S, Bunzl M (2010) *Foundational issues in human brain mapping* MIT Press, chapter 6. pp 71–92.
- Zuo X, Kelly C, Adelstein JS, Klein DF, Castellanos FX, et al. (2010) Reliable Intrinsic Connectivity Networks: Test-Retest Evaluation Using ICA and Dual Regression Approach. *NeuroImage* 49: 2163–2177.
- Harrison BJ, Pujol J, Lopez-Sola M, Hernandez-Ribas R, Deus J, et al. (2008) Consistency and functional specialization in the default mode brain network. *Proc Natl Acad Sci U S A* 105: 9781–9786.
- Ma S, Correa N, Li X, Eichele T, Calhoun V, et al. (2011) Automatic Identification of Functional Clusters in fMRI Data using Spatial Dependence. *IEEE Transactions on Biomedical Engineering* 2011 Sep 6 [Epub ahead of print].
- Ratcliff D, Pieper RD (1981) Deciding final clusters: An approach using intra and intercluster distances. *Plant Ecology* 48: 83–86.

Supporting Information

Appendix S1 Background on group comparison of ICA results.

(TEX)

Acknowledgments

We would also like to thank Dr. Christian Beckmann for making the IC image files from his 2005 paper [8] available to us.

Author Contributions

Conceived and designed the experiments: GVP DB LB. Performed the experiments: GVP. Analyzed the data: GVP. Contributed reagents/materials/analysis tools: GVP. Wrote the paper: GVP. Revised the article critically for important intellectual content: DB LB.