### fMRI as a Biomarker Reliability, QA, Multisite

Gary H. Glover

Radiological Sciences Laboratory Center for Advanced MR Technology

Stanford University School of Medicine

**Department of Radiology** 



# The consistency problem

- Where are the "emotion regions?"
- Look at coordinates reported in published studies of emotion:



163 studies of emotion

Slide courtesy Tor Wager

### **Consistency: fMRI of WM**



60 PET/fMRI studies <u>TD Wager</u>, 2003 The Scientist, July 19, 2004 v18 i14 p64(1)

Fake Method for Research Impartiality (fMRI): behavioral sciences bid for enhanced status fails short. (Closing Bell) Sam Jaffe.

Full Text: COPYRIGHT 2004 Scientist Inc.

For decades, the behavioral sciences have been at a dramatic disadvantage to the hard sciences. When a biologist hypothesizes that the addition of a particular ligand to a cell will cause a certain gene to turn on and thus produce a certain protein, all she has to do is to introduce the enzyme and then test for the protein. If it's there, she publishes a paper; if it's not, she quietly discards the work.

The psychologist has a much steeper hill to climb. Let's say he's trying to prove his hypothesis that most people who hate their fathers also secretly desire their mothers. Relying on the subject to tell you how he feels has too many obvious landmines that can corrupt the data. How can the psychologist scientifically prove that the connection exists?

Well, now he can. Or at least he can claim that it's a provable hypothesis. Thanks to fMRI (functional magnetic resonance imaging), dozens of studies are pouring out of the humanities aisle of academia claiming that the yellow and red blotches on fMRI scans reveal scientific evidence that can be used to

### Outline

- Issues with fMRI reliability
- Reducing confounds
  - HRF
  - calibration of vasoreativity
  - latency
- Physiological noise
- Multicenter studies

### **fMRI** as a biomarker: Motivation

Examples where quantifying activation may be important in drawing inferences about cognition:

- Inter-group comparisons Age, health
- Longitudinal studies normal/abnormal development, therapy
- Multi-center studies fBIRN schizophrenia fMRI trial



### Variance in fMRI

- inter-subject variability (probably what you want)
- inter-trial variability (attention, behavior)
- inter-run/session variability (attention, behavior, scanner)
- hemodynamic confounds (calibration)
- physiologic noise (measure and remove)
- small effect size (average)
- task design (control for unrelated effects)
- a lot more...

### **fMRI** Calibration: Motivation

What's the problem?



### **BOLD Contrast**

... BOLD signal is an epiphenomenological indicator of neural processing: many confounds to quantification

- HRF characteristics
   Amplitude
   Latency
   Baseline rCBF
- Physiological noise cardiac pulsation respiration head motion

-> calibration

-> denoising

# Activation map derives from thresholding a statistical estimate of BOLD CNR:

$$y_{meas}(t) = \beta_{exp} d_{exp}(t) + \beta_{cntl} d_{cntl}(t) + \beta_0 + \varepsilon(t)$$

$$GLM \Rightarrow \beta' s$$

$$T = \frac{\beta_{exp} - \beta_{cntl}}{\sigma} = \frac{effect}{resid}$$

$$T_{crit} = tpdf(p,df)$$



#### What does it mean?

#### Does "activation" = metabolic up-regulation consequent to neural firing? $y_{meas}(t) = \beta d(t) + \varepsilon(t)$

#### No, not directly... HRF is in the way

$$y_{meas}(t) = \beta(d(t) * h(t)) + \varepsilon(t)$$

### **General Linear Model**

$$y = (\beta_1 d_1 + \beta_2 d_2 + \beta_3 d_3 \dots + \beta_n d_n) * h + \beta_0 + \tilde{n}$$



### **Hemodynamic Response Function**

• Definition: BOLD response to an impulsive stimulus

 may include neuronal and vascular responses
 -> use a cognitively simple task to reduce neuronal component

may be nonlinear
 -> superposition does not hold





**M.** Thomason

# **Timing error**





### **Individual differences: HRF**

 temporal differences in HRF important in event-related designs

-> measure individual HRFs



### Outline

- Issues with fMRI reliability
- Reducing confounds
  - HRF
  - calibration of vasoreativity
  - latency
- Physiological noise
- Multicenter studies

### **Measurement of HRF**

#### **Use short stimulus, long ISI:**



## HRF: Measure h(t) with 1s task

#### motor

#### auditory



Finger tapping & tones at 3Hz, N=5

Glover, NI 1999

### **Measurement of HRF**

### • Event related designs are inefficient (T. Liu)



### **Detection or Estimation?**



Jittered (random) designs → maximum estimation efficiency

Block designs → maximum detection power

Liu et al. NI (2001)

### **Spectral content of h(t)**



### **Fourier Measurement of HRF: (FHRF)**

Design has on/off blocks of duration 4s, 6s, 8s, 10s, 12s, 16s, 20s, 30s, 40s, ...4s





### **Measurement Efficiency**



### Effect of HRF on Activation

#### Canonical Gamma variate

Measured Linear HRF







### **Measurement of HRF**

- Can provide characteristic info for each subject
  - requires a task
  - may be difficult to obtain in relevant regions
- Key features are amplitude & latency

   may be obtained without invoking
   a task

### Outline

- Issues with fMRI reliability
- Reducing confounds
  - HRF
  - calibration of vasoreativity
  - latency
- Physiological noise
- Multicenter studies

### Vasoreactivity



$$\begin{split} &\Delta R2^* \propto rCBV_a [Hb]_a^\beta - rCBV_0 [Hb]_0^\beta \\ &BOLD \propto -TE \cdot \Delta R2 * \\ &rCBV \propto rCBF^\alpha \\ &BOLD_a = M[f^\alpha (\frac{m}{f})^\beta - 1] \\ &m \equiv CMRO2_a / CMRO2_0 \\ &f = rCBF_a / rCBF_0 \end{split}$$

*M* represents a 'gain' factor related to vascular reactivity

### Measuring vasoreactivity

Use a task that does not involve change in metabolism:

- Hypercapnia
   O2 or CO2
  - Breath holding

 $\Delta R2^* \propto rCBV_a [Hb]_a^\beta - rCBV_0 [Hb]_0^\beta$   $BOLD \propto -TE \cdot \Delta R2^*$   $rCBV \propto rCBF^\alpha$   $BOLD_a = M[f^\alpha (\frac{m}{f})^\beta - 1]$   $m \equiv CMRO2_a / CMRO2_0$   $f = rCBF_a / rCBF_0$ 

### **BH Task**

### Block trial: 15s off/on 8 cycles, 4 min, 15 s





### **BH-induced BOLD signal**



# Vascular Responsivity: BH



### **BOLD Signal**



 $\Delta R2^* \propto rCBV_a[Hb]_a^\beta - rCBV_0[Hb]_0^\beta$  $BOLD \propto -TE \cdot \Delta R2 *$  $rCBV \propto rCBF^{\alpha}$  $BOLD_a = M[f^{\alpha}(\frac{m}{f})^{\beta} - 1]$  $m \equiv CMRO2_a / CMRO2_0$  $f = rCBF_a / rCBF_0$  $BOLD_{BH} = M[f_{BH}^{\alpha-\beta} - 1]$  $BOLD_{a} = BOLD_{BH} \frac{[f_{a}^{\alpha}(\frac{m}{f_{a}})^{\beta} - 1]}{[f_{BH}^{\alpha-\beta} - 1]}$ 



Working memory BOLD signal

### **BH Calibration**


## **BH Calibration: Individual Subs**

#### No calib

#### Calib



M. Thomason et al., 2007

## **BH Calibration: Group Activation**



 $3.5 \le T \le 10$ 

M. Thomason et al., 2007

## **Activation Response**



## **Calibration: SWM**



## **Correlation: WM & BH**

### SWM

### BH



**Top: adults Bottom: children** 

Thomason, et. al, 2005

# Outline

- Issues with fMRI reliability
- Reducing confounds
  - HRF
  - calibration of vasoreativity
  - latency
- Physiological noise
- Multicenter studies

# **Timing error**



Auditory WM (N. Gaab)



## **BH** to measure vascular latency

Can a BH task be used to quantify relative differences in vascular latency across the brain?

BH causes activation "everywhere" BH causes a BOLD signal response that is uncoupled from neural activation (CMRO2)



C. Chang (2008)

## Latency Measurement







Chang et al., 2008

# Latency Map



## **Impact on default-mode network**



subject 1

subject 2

## **Impact on Granger causality**

#### **Before correction**

#### After correction



Chang et al., 2008

# Outline

- Issues with fMRI reliability
- Reducing confounds
  - HRF
  - calibration of vasoreativity
  - latency
- Physiological noise
- Multicenter studies

# BOLD/Physio noise: sources

- Neuronal activation
- Respiration cycle
- Cardiac cycle
- Respiration volume
  (≈CO<sub>2</sub>)
- Heart rate



# Physio noise: reduction

- Neuronal activation
- Respiration cycle
- Cardiac cycle
- Respiration volume ( $\approx CO_2$ )
- Heart rate

RETROICOR

RVHRCOR

## **Cardiac/Respiratory Noise**



## **Retrospective sorting by cardiac phase**



## **Cardiac/Respiratory Motion Correction**

### **Timeseries in a voxel**



### after correction





**TR 1000ms** 

# **Retrospective Corrections**

### Before K-space I-space

ECG



Resp

# **Resting State**

### No retroicor

### With retroicor



# RETROICOR

#### No retroicor

#### With retroicor



# Physio noise: reduction

- Neuronal activation
- Respiration cycle
- Cardiac cycle
- Respiration volume ( $\approx CO_2$ )
- Heart rate

### RETROICOR

RVHRCOR

# Physio noise: sources

### • Variation in respiration volume (per time)



- Air intake is inversely related to the amount of  $CO_2$  in your blood.
  - CO<sub>2</sub> is a vasodilator (causes blood vessels to expand); this decreases vascular resistance, causing blood flow to increase
  - Known to affect BOLD (Wise, 2004)
- Heart rate
  - Affect cerebral blood flow (CBF)/volume; coupled to respiration
  - Not well known!

# Physio noise: RVHRCOR

- Method to remove artifacts due to low-frequency respiration (RV) and heart rate (HR) (Chang et al, 2009, Birn et al., 2008)
- Model: RV-related HR-related Voxel time series =  $RV \otimes RRF + HR \otimes CRF + (brain signal, etc.)$

Remove these

- Ä denotes convolution.
- RRF and CRF are impulse responses that describe the mapping between RV ←→ BOLD signal, and HR ←→BOLD signal, respectively (just like the hemodynamic response function (HRF) maps between stimuli ←→ BOLD signal)

# RVHRCOR: RV

1. Compute RVT from the raw respiration trace



# RVHRCOR: RV

#### 2. Convolve RVT with the RRF



correlation = 0.71



# **RVHRCOR: RV**



• Correlation between RVx and each voxel

# $RVx \approx CO2$

(after shifting CO<sub>2</sub> forward by about 10s)



- $RV(T) \sim ventilation \sim 1/PaCO_2$
- So, model could be:  $RV \rightarrow CO_2$  changes  $\rightarrow$  BOLD changes

Chang, 2009

# RVHRCOR: HR

- RV:
  - 1. Compute RV from the raw respiration trace
  - 2. Convolve RV with the RRF
- HR:
  - 1. Compute HR from the cardiac/PPG triggers
  - 2. Convolve HR with the CRF

## Cardiac response function (CRF)



# Variance explained: RVx & HRx



Chang et al., 2009

# Impact on activation: SM task

Retroicor only



Retroicor & rvhrcor



T maps: after - before

## **Impact on WM task**



WM Activation increases (corrected - uncorrected) *T=0.1-1.0* 

> Chang et. Al (2009)

# Impact on resting-state networks

• Decreases "false" positive correlations w/ default-mode



Impact on resting-state networks

### Correction can

- reduce spurious
  "connectivity" in DFM
- increase anticorrelated network connectivity



Chang et al., 2009





Chang, 2009
# Outline

- Issues with fMRI reliability
- Reducing confounds
  - HRF
  - calibration of vasoreativity
  - latency
- Physiological noise
- Multicenter studies

### **Multicenter Neuroimaging Studies**

- Potential use of MRI/fMRI as a biomarker
  - structural/functional differences may predict disease; large study numbers are necessary for biodiversity
    monitoring drug efficacy or other therapy
- Generate large data sets rapidly
- Access wide or targeted demographic characteristics
- Provide image databases for other analyses

### **Multicenter MRI**

- Desire to pool results across sites equally requires standardization
- Different venders may have incompatible characteristics/ definitions- e.g.
  - pulse sequence contrast in FSPGR vs MPRAGE
  - meaning of BW/echo spacing in EPI imaging -> artifacts/SNR
  - k-space apodization filters -> smoothness/CNR
  - grad distortion correction
  - geometric calibration precision
  - temporal stability

### **Multicenter MRI**

- Need to qualify sites for entry into study
  - characteristics for acceptance
    - geometric accuracy
    - contrast/resolution
    - SNR, CNR, tSNR (SFNR)
    - temporal stability
    - reliability/reproducibility
    - artifacts (ghosts/distortion/eddy current-related, ...)
  - understand sensitivity of scanner characteristics relative to desired measurements
- Set criteria for acceptance
- Need to maintain minimum performance standards
  - develop a QA program

### **Multicenter MRI**

- Decide policy for upgrades (chances virtually 100% for at least one site to upgrade )
  - minor: software only
  - major: hardware & software
- Develop procedures to control for/reduce site effects
- Develop procedures to reduce data acquisition confounds, e.g. hemodynamics in BOLD fMRI- test scientific question



## **fBIRN**

- Goal: Develop methods for pooling fMRI data on schizophrenics at each of 11 centers
  - load manipulation in emotional WM
- Approaches: Reduce intersite/intersubject variability
  - scanner QA
  - measure/calibrate/normalize
    - **BOLD** sensitivity
    - **HRF**/vasoreactivity

# **Issues in multicenter studies**

- Standardization of protocols
- QA
- Site Equalization

### **Standardization**

- Study design
- Acquisition parameters
- Scanner characteristics
- Study procedures
- Analysis pipeline
- Database structures

### **fMRI Imaging Characteristics**

#### **Modest importance:**

• Geometric accuracy (since fMRI is low

resolution- e.g. 3.4x3.4x4 mm<sup>3</sup>)

#### **Highest importance:**

- Stability (short/long term)
- BOLD CNR- B1 uniformity (coil choice)
- Susceptibility-induced distortion/dropout (seq. params)
- Ghost/spike noise/other artifacts
- Standardization across vendors

### **Intersite smoothness differences**



Friedman et al. NI (2006)

### Why smoothness differences?

### k-space reconstruction kernel



• Apodization-lower resolution, higher SNR



### **Important fMRI Characteristics**

fMRI acquisition contrast/smoothness

control parameters: resolution/smoothness (resel != FOV/matrix\_size) BW: keep ESP constant across venders slice spacing/skip/orientation fat saturation vs. water excitation readout trajectory (EPI vs. spiral), affects smoothness, artifacts
field strength (affects SNR, CNR, vessels vs. tissue, artifacts)



 affects resolution, motion sensitivity, SNR, num slices/TR

### **RF** Excitation/Slice Select

Spectral-Spatial (default for GE EPI)

> FATSAT (default for Siemens EPI)



Bryon Mueller (UMinn)

### **Measured Slice Profile**



#### spectral-spatial

fatsat, 256x64, 4 shot, fov 10cm







fat sat

Kun Lu, Tom Liu (UCSD)

### **Slice select**





### **Important MRI Characteristics**

• Dynamic image stability

fMRI & ASL depend on subtraction to compare conditions scanner stability must be << brain noise



### Brain noise relative to thermal noise: to set acceptance criteria

$$\sigma^2 = \sigma_0^2 + \sigma_s^2 + \sigma_p^2$$
$$= \sigma_0^2 + (\lambda_s + \lambda_p)S^2(\alpha)$$

G. Krueger (MRM 2000)

Acquire data at 10°, 77°
Calc fraction of scanner/ brain noise vs. thermal noise, using human & phantom scans

> D. Greve (MGH, ISMRM 2008)



## **Issues in multicenter studies**

- Standardization of protocols
- QA
- Site Equalization

### **QA: What to measure?**

- Time series image stability
- Signal to noise ratio
- Signal intensity
- Xmtr/Rcvr Gains
- MRS characteristics
- Eddy currents
- Geometric accuracy

### SNR, SFNR





# Stability



# Site differences

#### RMS instability, %





# QA helped to bring scanners into spec



## **Issues in multicenter studies**

- Standardization of protocols
- QA
- Site Equalization
  - smoothness compensation
  - SFNR compensation

# **Smoothness/BOLD Differences**



Friedman, et al. 2006

## **Sensitivity vs. Smoothness**

5 traveling subjects at 10 sites performing sensorimotor task



Friedman, et al. 2006

### **Smoothness equalization**

- In first-level analysis: Use "smooth to" instead of "smooth by"
- Smooth each site to largest FWHM using Gaussian filter

$$FWHM_{out}^{-2} = FWHM_{meas}^{-2} + FWHM_{filter}^{-2}$$

e.g., AFNI program (thanks to R. Cox)

Friedman, et al. 2006

# **BOLD Sensitivity: Oddball Task**



Novel Tones - Effect Size

Friedman, et al.



# **fMRI equalization across sites**

- Compensation for smoothness
- Compensation by SFNR

### **fMRI equalization by SFNR**

### • Measure SFNR using

### **Original SFNR**

- GM Rest
- WM Rest
- GM SMresid
- WM SMresid
- Covary for SFNR



Friedman, et al. 2006



**1.5**T





Friedman, et al. 2006

# fBIRN Emotional Working Memory Task: Emotional Distraction



Stimulus Item

Eight pictures presented during the encode period.

Eight picture pairs presented during the forced choice period.

### Mean of Contrast 1 at Four Study Sites



# ICC Unadjusted Minus ICC w/BH Calib: Site effects



Functional Contrast 1

G. Brown, et al.

# Multisite Studies- items not discussed

- Task design
- Ancillary hardware: projection, sound, button box, physiological, bite bar/stabilization
- Acquisition script
- Data integrity:
  - QA of all scan data
  - automated upload of scan and meta data
  - automated analysis pipeline
- Employ a traveling site coordinator/scannee

# Summary

**BOLD** contrast confounded by

inter-subject, inter-regional variations in hemodynamic response amplitude/latency
use hypercaphic calibration (e.g. BH) or ASL to reduce vasoreactivity 'gain factor' variance

respiratory- and cardiovascular-induced BOLD signal changes

- use RETROICOR and RVHRCOR

(must measure card. and resp. functions)

• site differences in stability, SNR, pulse sequence, parameters, study administration

# Summary

- Reduction of these confounds can improve confidence in activation maps
- Calibration important in group comparisons, longitudinal studies
- QA, standardization, calibration crucial in multicenter studies

# **fMRI:** Many biomarker applications



Acknowledgements: fBIRN Calibration Working Group Greg Brown @ucsd.edu Lee Friedman @gmail.com Doug Greve @nmr.mgh.harvard.edu Tom Liu @ucsd.edu Bryon Mueller @cmrr.umn.edu Jessie Turner @uci.edu Jim Vovoyodic @duke.edu

My kids

Catie Chang Lara Foland Jason Hsu Christine Law David Ress Moriah Thomason

Caires

NIH: National Center for Research Resources P41-RR009784 U24-RR021992



### **Nonlinearities-** Motor

#### **Measured- average of 5 subjects** Calculated using h(t)\*rect(T) averaged motor cortex response, 5 subjects alculated average motor response 60 50 16s \8s 41 normalized response 1 c c c signal 2s Normalized 8 30 10 2 5 10 15 20 frame num (1/s) 0 25 5 10 15 time frame 20 25 Ω 2/3s 2s1/3s 15 **8**s **Finger tapping at 4**s 16s3Hz: 1/3s, 2/3s, 1s, 2s, 4s, 8s, 16s Glover, NeuroImage 9:416 (1999)

# **Effect of HRF on Activation**

Linear HRF





### **Nonlinear HRF**

# **BH Calibration Method**

- Use BH task (non-neuronal, no change in CMRO2) to normalize cognitive task
- Reduces signal change related to vasoreactivity
- Should reduce inter-subject variance