The BOLD Response

Douglas C. Noll
Department of Biomedical Engineering
University of Michigan
Outline

• Mechanism of the BOLD response
  – Susceptibility of blood
  – Fick’s relationship/mass balance

• Spatial specificity of BOLD and alternatives

• Evidence for and mechanisms of non-linearity

• Implications for fMRI Studies

• Conclusions
Dynamic Vascular Changes

Optical imaging of intrinsic signal (OIS) from a rat

Anatomical Image

620 nm (BOLD-like)

Movie courtesy of Alberto Vazquez, University of Pittsburgh
Interesting Temporal Features

Early OIS Decrease Map

Well-localized temporally and spatially

Late OIS Increase Map

Less well-localized temporally and spatially, but larger response
The BOLD Effect

Brain Function

Metabolic Function

Physiological Effects

Physical Effects

MRI Properties

Neural Activity
(spiking, synaptic activity, …)

Glucose and Oxygen Metabolism

Neurovascular Effects

Cerebral Blood Flow (CBF)

Venous Blood Volume (vCBV)

Blood Oxygenation

Magnetic Field Uniformity
(microscopic)

Decay Time (T2*)

T2*-weighted (BOLD) Image Intensity

Other Factors
Hematocrit
Vessel Diameter
Vessel Orientation
Blood Volume Fraction
Field Strength

Noll
What Part of Neural Activity does BOLD Represent?

• Spiking or synaptic activity?
  – Post-synaptic recycling of neurotransmitters is the very metabolically intensive

Local field potentials (LPF) correlate more strongly to BOLD than multi-unit activity (MUA)

Synaptic activity > spiking

(Logothetis et al, Nature, 2001)
Consensus is that BOLD more strongly reflects the input to a neuronal population, rather than only its spiking output to other regions.

- A number of studies that have found BOLD changes independent of neural firing

In this study, spiking correlates with BOLD better than LPF

Spiking > synaptic activity

(Lima et al, J Neurosci, 2014)
What Drives the BOLD Response?

- Evolutionarily speaking:
  - Need for O$_2$ for increased metabolism
  - Need to take away metabolic products
- Delivery of O$_2$ is driven by diffusion
  - Rate of delivery depends on concentration gradients
  - Increase delivery by decreasing tissue O$_2$ (happens with metabolism) and increasing cap O$_2$ (happens with increased flow)
- Can you deliver more O$_2$ without flow increases?
  - Yes, but that’s not what happens in healthy tissues
Diffusion Delivery of O2

What are the mechanisms of increased flow (CBF)?

- Complicated!
- Blood flows increase through increases in arteriole diameter
- Smooth muscle surrounding arterioles is relaxed by variety of metabolic products

The BOLD Response

• Changes in cerebral blood flow (CBF) drive the BOLD changes
  – Temporal characteristics of the CBF changes dominate the BOLD temporal response

BOLD and CBF changes with graded hypercapnia and visual stimulation

BOLD and CBF changes have similar temporal characteristics


Magnetic Susceptibility of Blood

• The distortion of the magnetic field is:
\[ \Delta B = \chi_m B_0 \]

• Some relative susceptibility values

<table>
<thead>
<tr>
<th>Material (relative to water)</th>
<th>( \chi_m ) (x 10(^{-6}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water</td>
<td>0</td>
</tr>
<tr>
<td>Room Air</td>
<td>9.4</td>
</tr>
<tr>
<td>Pure oxygen gas</td>
<td>11</td>
</tr>
<tr>
<td>Deoxygenated Blood (Hct = 1)</td>
<td>2.3</td>
</tr>
</tbody>
</table>
Magnetic Susceptibility of Blood

• An expression for the susceptibility of blood is:

\[ \chi_{\text{blood}} = \text{Hct}(Y\chi_{\text{oxy}} + (1-Y)\chi_{\text{deoxy}}) + (1-\text{Hct})\chi_{\text{plasma}} \]

  – Hct is hematocrit and
  – Y is blood oxygenation

• But, \( \chi_{\text{oxy}} = \chi_{\text{plasma}} = \chi_{\text{tissue}} = \chi_{\text{water}} = 0 \), so:

\[ \chi_{\text{blood}} = \text{Hct}(1-Y)\chi_{\text{deoxy}} \]

• Example
  – If Hct = .4 and Y = .7, \( \chi_{\text{blood}} = 0.12 \chi_{\text{deoxy}} = 0.28 \times 10^{-6} \)
Extravascular Contribution

\[ \Delta \omega (r, \theta, \phi) = \Delta \omega_0 \sin^2(\theta) \left(\frac{a}{r}\right)^2 \cos(2\phi) \]

Orientation and radius dependent

Courtesy of P. Jezzard
Intravascular Contribution: heterogeneous orientation

\[ \Delta \omega (\theta) = \Delta \omega' \left[ 3 \cos^2(\theta) - 1 \right] / 3 \]

Courtesy of P. Jezzard
Tissue and Blood Contributions

Field strength dependence

Noll

Boxerman et al., MRM 34, 1995

Courtesy of P. Jezzard
What is T2*?

- T2* has two parts:
  - Inter-molecular interactions leading to dephasing, a.k.a. T2 decay
  - Macroscopic or mesoscopic static magnetic field inhomogeneity leading to dephasing, a.k.a. T2’

\[
\frac{1}{T2^*} = \frac{1}{T2'} + \frac{1}{T2}
\]

- Pulse sequence issues:
  - Spin echoes are sensitive to T2
  - Gradient echoes are sensitive to T2*
What is $R2^*$?

- We often talk about “rates” instead of time constants, e.g.:

\[ R2^* = \frac{1}{T2^*} \]

- So the decay rate is given by the sum of the rates:

\[ R2^* = R2 + R2' \]
Dephasing from Deoxyhemoglobin

60% vs. 90%

Different Fields Vs. Uniform Field
BOLD Signal Changes

T2* - T2 decay with Susceptibility Effects

- $R2'$ relates to susceptibility properties of blood:
  
  $$R2' = \kappa \gamma \Delta BV$$

  
  $$= \kappa \gamma B_0 \text{Hct}(1 - Y) \chi_{\text{deoxy}} V$$

  - $\kappa$ is a constant
  - $\Delta B$ is the distortion of the field induced by the blood
  - $V$ is blood volume fraction
  - Observe that the BOLD decay rate is proportional to $B_0$

- As oxygenation ($Y$) goes down, the decay rate $R2'$ goes up (faster decay)
BOLD Changes

• Consider that the total amount of deoxyhemoglobin is:

\[ Q = \nu \text{Hct} (1-Y) \]

• Thus

\[ R2' = \kappa \gamma B_0 \kappa_{\text{deoxy}} Q \]

• \( R2^* \) changes are directly proportional to changes in total deoxyhemoglobin (in a voxel):

\[ \Delta R2^* = \Delta R2' \propto \Delta Q \]
R2* and R2’

- Some have suggested, based on computer modeling a more nonlinear relationship with respect to blood O2:

\[ R2' = \kappa \gamma B_0 \text{Hct} \chi_{\text{deoxy}} V (1 - Y)^\beta \]

- $\beta$ is a constant that ranges from 1-2, typically assumed to be about 1.5
- We’ll use this relationship later…
MRI Signal Changes with R2’

- Let the signal at the echo time be:
  
  \[ m = m_0 e^{-\frac{TE}{T2^*}} = m_0 e^{-TE \cdot R2^*} \]

  \( R2^*b \) – the baseline rate

  \( R2^*a \) – the activation relaxation rate

- Fractional (%) signal change:
  
  \[ \Delta\% = \frac{m_a - m_b}{m_b} = \frac{m_0 e^{-TE \cdot R2^*_a} - m_0 e^{-TE \cdot R2^*_b}}{m_0 e^{-TE \cdot R2^*_b}} \]

  \[ \Delta\% = e^{-TE \cdot \Delta R2^*} - 1 \]

  \[ \approx -TE \cdot \Delta R2' \]
MRI Signal Changes with Deoxyhemoglobin

• Fractional (%) Signal Change:

\[ \Delta \% = \frac{m_a - m_b}{m_b} \approx -TE \cdot \Delta R2' \]

• But \( \Delta R2^* = \Delta R2' \) is roughly proportional to change in total deoxyhemoglobin, so:

\[ \% \text{BOLD} \propto -\Delta Q \]
Optimization of TE

Maximum signal difference occurs at $TE \approx T2^*$

Problem – $T2^*$ varies across the head, but you have to pick 1 TE
(Note that we usually pick a TE that is shorter than the optimal.)
Relationship to Metabolism

• Of interest is the relationship between the BOLD signal and oxygen metabolism
  \( \text{CMRO2} = \) cerebral metabolic rate of \( \text{O}_2 \)

\[
\text{CMRO2} = C \cdot f \cdot \text{Hct} \cdot \text{OEF}
\]

  – \( f \) = perfusion rate (CBF) in ml/min/(100 g tissue)
  – \( \text{OEF} \) = oxygen extraction fraction
  – \( C \) = constant related of \( \text{O}_2 \) capacity of hemoglobin

• This expression says “amount consumed is equal to amount delivered”
  – Known as Fick’s Principle
Assuming that most of the fMRI signal comes from the venous side, then:

\[ \text{OEF} = 1 - Y \]
Fick’s Principle

• We can now solve for some constants:

\[
\frac{\text{CMRO}_2}{f(1-Y)} = C \cdot \text{Hct}
\]

• And equate these constants for baseline and activation states:

\[
\frac{\text{CMRO}_2_a}{f_a(1-Y_a)} = \frac{\text{CMRO}_2_b}{f_b(1-Y_b)}
\]

or

\[
\frac{(1-Y_a)}{(1-Y_b)} = \frac{f_b}{f_a} \cdot \frac{\text{CMRO}_2_a}{\text{CMRO}_2_b}
\]
Connection to BOLD

• Recall that the fractional signal change is:

\[
\frac{m_a - m_b}{m_b} \propto \Delta \left[ V (1 - Y)^\beta \right] \propto V_b (1 - Y_b)^\beta - V_a (1 - Y_a)^\beta
\]

• Rearranging, we get:

\[
\frac{m_a - m_b}{m_b} = M \left[ 1 - \frac{V_a}{V_b} \left( \frac{1 - Y_a}{1 - Y_b} \right)^\beta \right]
\]

where \( M \) is a constant (includes baseline properties)
The BOLD/CMRO2/CBF Connection

• Combining the above expressions we get:

\[
\frac{m_a - m_b}{m_b} = M \left[ 1 - \frac{V_a}{V_b} \left( \frac{f_b}{f_a} \frac{\text{CMRO2}_a}{\text{CMRO2}_b} \right)^\beta \right]
\]
The CBF/CBV Connection

• We can also assume a volume flow relationship:

\[ \frac{V_a}{V_b} = \left( \frac{f_a}{f_b} \right)^\alpha \]

Grubb’s Relationship (\( \alpha = 0.38 \))

• Warning – this is problematic because most volume changes are on the arterial side and most deoxy-Hb changes are on the venous side. Nevertheless, we will use it ☺.
Volume Changes

- Mostly on the arterial side

Images courtesy of Alberto Vazquez, University of Pittsburgh
The BOLD/CMRO2/CBF Connection

Finally, we get:

\[
\frac{m_a - m_b}{m_b} = M \left[ 1 - \left( \frac{f_b}{f_a} \right)^{\beta - \alpha} \left( \frac{\text{CMRO2}_a}{\text{CMRO2}_b} \right)^\beta \right]
\]

Observations:
- As flow increases, BOLD increases
- As CMRO2 increases, BOLD decreases
The BOLD/CMRO2/CBF Connection

This difference due to CMRO2

GHC = Graded Hypercapnia
GVS = Graded Visual Stimulation

Lines of constant CMRO2

Can CMRO2 Ratios Be Determined?

• Maybe...

\[ \frac{m_a - m_b}{m_b} = M \left[ 1 - \frac{V_a}{V_b} \left( \frac{f_b}{f_a} \frac{\text{CMRO2}_a}{\text{CMRO2}_b} \right)^\beta \right] \]

- Need BOLD fractional signal change
- Need CBF measurements \((f_a/f_b)\)
- Need volumes or assume Grubb’s relationship
- Need constant M
  (This can be done making the above measurements in a hypercapnia challenge – no CMRO2 changes!)
Complicating Factors…

O2 Saturation by Vessel Size

Graph courtesy of Alberto Vazquez, University of Pittsburgh
One interesting feature of the fMRI response is the “post-stimulus undershoot”.

Post-Stimulus Undershoot

- Most likely sources
  - Persistent elevation of (venous) volume
  - Persistent elevation of CMRO2

- These data suggest continued CMRO2 is the dominant source

Outline

• Mechanism of the BOLD response
• Spatial specificity of BOLD and alternatives
• Evidence for and mechanisms of non-linearity
• Implications for fMRI Studies
• Conclusions
Spatial Specificity

• BOLD vs. other mechanisms
  – Since BOLD is venous weighted, it has substantial signals along the cortical surface

• Perfusion and Volume changes are mostly deeper
  – Perfusion – capillaries
  – CBV – arterioles

CBF and CBV Methods

• Cerebral Blood Flow (CBF) with MRI
  – Most comment in Arterial Spin Labeling (ASL) with the most common version of that being:
    – pCASL (pseudo-continuous ASL)
    – Works well, but low SNR and slower than BOLD

• Cerebral Blood Volume (CBV) with MRI
  – Contrast agent methods
  – Increasingly popular – Vascular Space Occupancy (VASO) methods
  – Also works well, but hard to get large volumes and lower SNR than BOLD
Continuous ASL (CBF)


ASL Acquisition. Hernandez-Garcia, 2019
CBV Methods (e.g. VASO)

- VASO inverts the entire volume and images where the blood is nulled out
- Signal decreases as blood volume increases


VASO/ALS vs. BOLD

- ASL – spin exchange focused on capillaries
- VASO – signals focused where volume changes most
- BOLD – signals where blood oxygenation changes

Vascular Specificity vs. SNR

Note that these images are 7T MRI – not everything is the same at 3T
VASO/ALS vs. BOLD

- VASO is often pitched for high-resolution, layer- or column-specific fMRI
  - Requires small voxels and is more appropriate for 7T MRI (though there is a little bit of work at 3T)
- SNR (sensitivity) and speed issues have limited the interest in VASO and ALS at 3T
- Still, it may be of interest if subtle localization changes (e.g. topographic mapping) is important
Outline

• Mechanism of the BOLD response
• Spatial specificity of BOLD and alternatives
• Evidence for and mechanisms of non-linearity
• Implications for fMRI Studies
• Conclusions
Linearity

A “linear system” satisfies the following:

• **Scaling**
  – Increasing stimulation by some ratio will increase the output by the same ratio

• **Superposition (additivity)**
  – Combining (adding) any two stimuli will lead to an output that is the sum of the two responses

• **Time-invariance**
  – A response is the same irrespective of when it comes or what precedes or follows it
Superposition and Predicted Responses

- Time-Invariance
- Scaling
- Superposition of responses

Impulse response or HRF Responses to individual stimuli
Summed response
Why Is Linearity Good?

• Extensive theory of linear systems

• Allows definition of an “impulse response function”
  – Commonly known as the hemodynamic response function (HRF)

• Predicted responses are easily determined
  – Can also allow “deconvolution” of response to get estimates of the system input
Is the BOLD response linear?

- In many cases, it is almost linear.
- In other cases, it is not very linear, notably:
  - Manipulations of task (block) duration for short (less than 4 s) blocks
  - Repeated stimuli after long (e.g. 30 s) rest periods
  - Other effects
Manipulations in Task Duration

Linearity implies that responses from short stimuli should predict responses to longer stimuli.

Response from 3 s stimulus poorly predicted responses to 6, 12 and 24 s stimuli.

But, response from 6 s stimulus was a good predictor.

→ The “system” behaves nonlinearly for 3 s stimuli.

Manipulations in Task Duration

A system model should remain the same regardless of stimulus duration

- The “system” produced shorter and more intense responses for stimuli less than 4 s in duration.

- Estimated model parameters varied for short stimuli.

- The same study also showed some nonlinearities with manipulations in stimulus intensity.

Nonlinearity in predicting responses also appears in cerebral blood flow (CBF) as well as BOLD.

Response from 2 s stimulus poorly predicted responses to 6 s stimulus for.

- Nonlinearities occur broadly in the vascular response.

Explicit manipulation of inter-stimulus interval demonstrates that subsequent responses are delayed and less intense.

→ The “system” response is time-variant and the degree depends on time to preceding stimulus.

→ Long latencies (as much as 1 s) imply that some effects are not neural in origin.
Mechanisms for Nonlinear Behavior

• Vascular
  – Non-linear (e.g. visco-elastic) behavior of vessels
  – Non-linear coupling/control of vascular response
  – Non-linear extraction of oxygen from blood pool

• Neural
  – Non-linear neural response to stimulus or task (e.g. novelty, threshold effects)
  – Synaptic activity (including inhibitory) vs. firing rate

• Metabolic
  – Non-linear relationship between metabolism and neural activity

(Not meant to be an exhaustive list.)
Vascular Mechanisms for Nonlinearity

• Manipulations in basal flow conditions through hypercapnia and hypocapnia

• Increased basal CBF leads to
  – Decreased intensity
  – Increased response width and latency
  – And vice-versa.

• Basal flow conditions affect BOLD response

(Cohen, et al., JCBFM 2002; 22:1042-1053.)
The “balloon” model (Buxton RB, et al. Magn Reson Med 1998; 39:855-864) can predict some of the observed nonlinearities. This model can explain the persistent elevation of cerebral blood volume (CBV), perhaps through viscoelastic behavior.
Vascular Mechanisms for Nonlinearity

- Cerebral blood volume (CBV) response returns to baseline more slowly than the CBV onset

- Elevated CBV effectively changes the basal physiological conditions

- Could cause time-variant behavior

Neural Mechanisms for Nonlinearity

• Neural networks are nonlinear.

• Some observed nonlinearities could be caused by neural mechanisms, for example:
  – Changes in apparent BOLD magnitude response with stimulus duration are similar to measured in local field potentials (LPF’s).
  – BOLD response found to be closely correlated with LPF’s.

Metabolic Mechanisms for Nonlinearity

- There have been suggestions that oxygen metabolism may persist in a nonlinear fashion following activation
  - Alternate explanation of the post-stimulus undershoot of the BOLD response
Mechanisms for Nonlinear Behavior

• BOLD nonlinearity is a likely result of:
  – A combination of vascular and neural effects
  – Other mechanisms including metabolic rates and the coupling between neural activity and physiological responses.

• The good news is that we can deal many of these effects.
Outline

- Mechanism of the BOLD response
- Spatial specificity of BOLD and alternatives
- Evidence for and mechanisms of non-linearity
- Implications for fMRI Studies
- Conclusions
Avoiding Nonlinearities

• Blocked task designs
  – Long blocks do not exhibit much nonlinearity and are relatively immune to changes in shape of HRF

• Event-related task designs
  – Avoid combinations of long (> 30 s) and short (< 10 s) intertrial intervals
  – Rapid, random stimulus presentations give mostly linear response
    • No opportunity to return to basal neural or vascular conditions
Explicit Modeling of Nonlinearities

• One can also explicitly model the nonlinear effects and incorporate that into the development of reference functions for GLM analysis.

• One can also explicitly model the nonlinearity through:
  – Many other approaches.

Conclusions

• The temporal character of the BOLD response:
  – Results from complex chain of physiological events
  – Is fairly slow (time constants >= 5 s)
  – Exhibits some nonlinear characteristics

• BOLD is biased towards draining veins
  – Alternatives (CBF, CBV) are costly (SNR, speed)

• Linearity is important for accuracy in analysis and temporal feature extraction
  – Nonlinearity is likely caused by several mechanisms and is currently topic of ongoing research