Vascular and neural basis of the BOLD signal
Patrick J Drew

Neural activity in the brain is usually coupled to increases in local cerebral blood flow, leading to the increase in oxygenation that generates the BOLD fMRI signal. Recent work has begun to elucidate the vascular and neural mechanisms underlying the BOLD signal. The dilatory response is distributed throughout the vascular network. Arteries actively dilate within a second following neural activity increases, while venous distensions are passive and have a time course that last tens of seconds. Vasodilation, and thus local blood flow, is controlled by the activity of both neurons and astrocytes via multiple different pathways. The relationship between sensory-driven neural activity and the vascular dynamics in sensory areas are well-captured with a linear convolution model. However, depending on the behavioral state or brain region, the coupling between neural activity and hemodynamic signals can be weak or even inverted.

Address
Departments of Engineering Science and Mechanics, Biomedical Engineering and Neurosurgery, Pennsylvania State University, University Park, PA 16802, United States

Corresponding author: Drew, Patrick J (pjd17@psu.edu)

Current Opinion in Neurobiology 2019, 58:61–69
This review comes from a themed issue on Computational neuroscience
Edited by Máté Lengyel and Brent Doiron

https://doi.org/10.1016/j.conb.2019.06.004
0959-4388/© 2019 Elsevier Ltd. All rights reserved.

“... Working by hindsight/Got the message from the oxygen...”
Talking Heads and Brian Eno, “Crosseyed and Painless”, 1980

Introduction
In the brain, local increases in neural activity are typically accompanied by a transient oversupply of blood flow, known as functional hyperemia, leading to an increase in blood and tissue oxygenation. This forms the basis of the blood-oxygen level dependent (BOLD) signal, which can be detected non-invasively in humans, and has provided an unparalleled window on human cognition. Despite the ubiquity of BOLD fMRI, we do not completely understand what BOLD signals are telling us about neural activity, or how these signals might change with behavioral state. Work done in animal models is beginning to elucidate the vascular changes accompanying functional hyperemia. How signals from astrocytes, excitatory neurons, and inhibitory neurons combine to control the BOLD signal is still an area of active exploration. There is evidence that spontaneous BOLD signals are less correlated with the underlying neural activity than sensory-evoked responses, and that the nature of neurovascular coupling can vary from brain region to brain region. Clearly, understanding the neural and vascular basis of the BOLD signal is essential for interpreting hemodynamic signals.

Vascular origins of BOLD signal-beyond the balloon model
The BOLD signal is generated in the vasculature, and for interpreting the BOLD signal the vascular mechanisms of functional hyperemia need to be understood. BOLD contrast is generated by changes in local deoxyhemoglobin, and depending on the field strength and scanning sequence used, the relative contributions of smaller venules and capillaries can be enhanced. For most field strengths used in humans, the signal primarily arises in the veins and capillaries. When vessels dilate, the increased flow results in an increase in oxygenation [1]. Arteries are surrounded by smooth muscle that relaxes following increases in neural activity, and this dilation can be a 20–30% increase in diameter in awake animals [2,3]. Arteries start dilating within a few hundred milliseconds of the onset of neural activity, reach peak dilation in response to sustained stimulation within 2–3 s of the stimulus onset [2,4,5,6,7], and return to their baseline diameter within a few seconds upon cessation of stimulation. Veins dilate much more slowly, their dilations are smaller (up to 10%), and take tens of seconds or more to reach maximal dilation [2,3,4]. The dilation of arteries starts in the deeper layers (4–6) and propagates up the arterial tree [7] (Figure 1). Increases in arterial volume are much larger than increases in venous volume [4,6,8]. These results are in sharp contrast to the older ‘balloon’ models of the hemodynamics response [9], which posited most of the volume change took place in the venous compartment, and which have been superseded by more realistic models [10,11]. Like the vasculature in other parts of the body [12], the endothelial cells of cerebral blood vessels can propagate electrical signals up the vascular tree into the large pial arteries [13]. This spreading of dilation can lead to a ‘blurring’ of the hemodynamic response relative to the underlying neural activity [14]...
(but see Ref. [15]), so the resolution of fMRI will be limited by the spatial spread of flow and dilation in the vasculature. The dilations propagate rapidly (~10 mm/s), and the increase in flow will happen instantaneously throughout the network, so the direction of these propagated dilations is too rapid to detect with conventional MRI.

While the dynamics of arterial and venous dilation are well understood, the dynamics of capillaries are controversial, as dilations of 5–10% are close to the limit of a spatially resolvable differences in two-photon microscopy. Understanding the degree to which capillaries diameter is under neural control is highly relevant to imaging, as very local control of capillary dilation would produce highly localized hemodynamic signals that will allow high-field MRI to measure laminar-specific signals [16,17]. Optical ablation of a single pericyte (mural cells, found in close apposition to capillaries) causes a transient, localized dilation of the adjacent capillary [18], showing that pericytes can regulate capillary diameter on the timescale of hours or more. In vitro pericytes are contractile and can dilate in response to neural activity [19,20**]. However, in vivo experiments have been split as whether or not capillaries dilate in response to sensory stimulation, with some groups finding no dilation [21–23] and other groups seeing dilation [19,20**]. Some of this controversy is due to what is considered to be a capillary, which varies from study to study [25]. The capillary bed is not composed of a single type of vessel [25], but rather a continuum of smooth muscle cells, pericytes and endothelial cells between the arteries and veins [26,27**]. There is a robust arterial–venous gradient in gene expression patterns [28] and in smooth muscle/pericyte coverage...
[26, 27**], which likely corresponds to variations in the dilatory response. Smooth muscle actin is found on the penetrating arterioles and is present several branches down from penetrating arterioles, while the smallest vessels lack this coverage [27**]. Recent work tracing sensory-evoked flow and diameter changes from the smallest capillaries up to the surface arteries has found that penetrating arteries and their precapillary branches dilate first [29**], indicating that they are the primary controller of flow (Figure 1). This drives flow increases in the capillaries (though there actually is a brief slowdown of blood flow in the arterioles). Any dilation in the smaller capillaries is delayed and passive [29**], consistent with the surrounding structure composed of sparse, smooth-muscle actin free pericytes around them [26, 27**]. Integrating the dynamics of vessels dilations into realistic computer models will be key to understanding the roles of different vessels types in flow control [30, 31] and the generation of the BOLD signal [10*].

**Neural and astrocytic control of the hemodynamic response**

Both neurons and astrocytes have been implicated in controlling cerebral blood flow (Figure 1). Excitatory neuron activity is correlated with blood flow and oxygenation increases [32, 33] and optogenetic stimulation of pyramidal neurons can drive increases in blood flow comparable to that seen in the sensory evoked condition [15, 34–37], potentially mediated by vasodilatory prostaglandins [38]. Optogenetic stimulation of interneurons drives robust vasodilation [39, 40*, 41**], even though it will (at least transiently) silence many nearby neurons. Different subtypes of interneurons produce different vasoactive substances, including nitric oxide (NO) [42], and neuropeptides [43*, 44], such as the vasoconstrictor NPY [40*]. While optogenetic stimulation can produce neural responses similar to sensory stimulation [15], the patterns of activity produced probably differ from normal neural activity. While activity of a neuronal subtype may cause vasodilation, its net contribution to the BOLD signal depends both on the evoked increase in oxygenated blood flow and on its metabolic demand [45]. Recent experiments utilizing optogenetic stimulation and pharmacology to independently activate excitatory and inhibitory neuron populations show that the ratio of evoked metabolic demand to blood supply increases is much larger in pyramidal neurons than interneurons [41**]. These results suggest that interneuron activation is the primary driver of the BOLD response.

Astrocytes play a role in regulating basal artery tone [46, 47, 48*] but whether they play a role in rapid stimulus-evoked responses has been controversial. The timing of astrocytic calcium signals relative to vasodilation has been hard to work out in vivo due to the contribution from the neuropil, with some groups finding rapidly evoked calcium changes in astrocytes [49], and some groups seeing astrocytic calcium signals lagging arterial dilations [50]. Genetically encoded calcium sensors in olfactory bulb astrocytes show rapid stimulus-evoked calcium signals [51], though both rapid [52] and delayed [53] astrocytic responses have been seen in the cortex of awake mice. Neither chemogenetic disruption of astrocytic signaling [54] or knockouts of IP3R2, which play a role in Ca2+ signaling [50] (but see Ref. [52]) alter the evoked hemodynamic response. In slice, electrical stimulation can induce vasodilation without inducing calcium signals in astrocytes [55]. Recent evidence has suggested that astrocyte-released PGE2 targets pericytes on capillaries, rather than smooth muscle on arteries [20**].

It has been suggested that deoxygenation may play a role in reducing blood flow resistance [56]. However in awake animals there is no deoxygenation preceding the hemodynamic response [57], the local oxygen concentration around RBCs is very high [58], the speedup in the flow in the smallest capillaries follows the dilation of the arterioles and intermediate capillary dilations [29**], and changing systemic oxygenation in vivo does not change vasodilatory responses [59, 60], so it is unlikely that this mechanism plays a role in the normoxic brain. In slice, elevated oxygen concentrations can invert the sign of astrocytes neurovascular coupling [61]. However, the oxygen concentrations at which these inversions are seen in the slice (~500 mmHg) are more than 10-fold higher than those typically seen in vivo [60, 62, 63]. Another metabolic pathway involves voltage-gated inwardly rectifying K+ channels which are present on capillary endothelial cells, and are activated by extracellular increases in potassium accompanying neural activity [64**]. The hyperpolarization of endothelial cells propagates upstream to arteries [64**], and contributes to the upstream spread of vasodilation [13*]. Thus, there is no single signaling mechanism from neurons or astrocytes that causes vasodilation, and it is likely that both capillaries and arterioles receive and transduce these signals.

**Electrophysiological correlates of BOLD signals**

The vast majority of studies investigating the electrophysiological basis of the BOLD signal have found that the gamma band power of the local field potential (nominally 40–100 Hz) is the band most correlated with the subsequent vasodilation and/or increased oxygenation [65–67, 68**, 69] (Figure 2). Multiunit spiking activity also tends to be strongly correlated with the hemodynamic response [53, 65–67, 68**, 70]. One should not think of these electrical signals themselves as driving vasodilation and the BOLD response, but rather that vasodilator release is elevated during the patterns of neural activity that generate gamma-band oscillations. These electrophysiological correlates of the hemodynamic response are stable across behavioral state (passive stimulation, active sensation and ‘rest’ [68**]), though the
signal to noise may vary. Given the link between gamma band power increases and the BOLD signal, and the critical role interneurons play in generating gamma band oscillations, the most parsimonious explanation is that the activity level of interneurons is the primary driver of the BOLD signal [39,40*,41**,42,43*,44]. The electrophysiological correlates of the negative BOLD [71] signal are murkier. These signals could be produced by high levels of activity outstripping oxygen demand, or decreased neural activity accompanying vasoconstriction [72,73]. Although evoked potentials have been used to assay neural activity in many studies, evoked potentials are volume-conducted responses and do not reflect local neural activity [74].

There has been speculation that slow (<1 Hz) cortical potentials are the neural correlate of the hemodynamic response [75]. However, these potentials are not generated by neural activity, but rather by the dilations of blood vessels themselves, as there is electrical potential difference between the blood and brain [76–78]. Consistent with vascular dilation driving slow potentials independent of neural activity, inhalation of CO₂ drives large increases in arterial diameter and corresponding slow cortical potentials without corresponding changes in neuronal membrane potential [76–78].

The hemodynamic response function—relating neural activity to hemodynamic signals

Changes in cerebral blood volume (CBV) and BOLD are typically taken to be convolutions of a linear kernel (known as the hemodynamic response function (HRF)) with underlying neural activity changes [79]. The CBV and cerebral blood flow (CBF) HRFS will have faster time courses than the BOLD HRF [6**,11,80], and are usually modeled as being a gamma-distribution function, or the difference between two gamma distribution functions (in order to account for any post-stimulus undershoot) (Figure 2c). The CBV HRF can be decomposed into a rapid, spatially specific arterial component and slower, diffuse venous component [81,82]. Anesthesia slows [83,84] and decreases the amplitude [2*,83–85] of the hemodynamic response. Additionally, oxygen levels are also lower in awake animals than anesthetized ones [63], so HRFS obtained in anesthetized animals will be substantially different from those in awake brains.

The BOLD HRFS of different brain regions vary in their temporal dynamics, as regions with larger draining veins will have more delayed HRFS [86*,87]. These spatially varying delays, coupled with systemic oscillations in blood flow, volume, or oxygenation [88*], may contribute to the generation of resting-state connectivity [89]. Lastly, the HRF may not be stable, but could change its dynamics in response to hormonal conditions [90]. In accordance with the regional and state-dependence of the HRF, one is wise to determine the parameters for relevant HRF empirically rather than relying on estimates of HRFS made under disparate conditions.

Resting-state neurovascular coupling

There has been extensive use of BOLD fMRI to examine the spontaneous, or resting-state dynamics of the brain. Some studies have found relatively weak correlations (Pearson’s correlation coefficient, R, in the range of 0.0–0.3) between spontaneous neural activity and spontaneous hemodynamic signals [68**,69,91,92**], others found stronger (R up to 0.7) correlations [33,67]. One should keep in mind that if the correlation coefficient (R) between neural activity and the hemodynamic/BOLD signal is 0.7, the signal to noise ratio is 1:1 (as the variance in the BOLD signal explained by neural activity is
proportional to $R^2$). A correlation coefficient of 0.3 implies a signal-to-noise ratio of $\sim 1:10$, where neural activity drives $<10\%$ of the hemodynamic signals. The correlation between neural activity and hemodynamic signals can be highly statistically significant, but still be low enough to only explain a fraction of the variance. While the correlations between spontaneous neural activity and hemodynamic signals are low, evoked neural activity and hemodynamic signals have found much higher correlations ($R \geq 0.9$) [65,68**,70]. Thus, the observed low correlations between neural activity and hemodynamic signals present during spontaneous activity are not due to methodological issues, but rather ongoing dynamics in the vasculature that are independent of neural activity. One of these processes is vessel-autonomous oscillations in the arterial diameter, which have been found in arteries throughout the body [93,94]. In the mouse cortex, spontaneous oscillations in arterial diameter in the 0.1–0.3 Hz band persist when local cortical spiking and glutamatergic transmission is pharmacologically blocked [68**]. The spontaneous oscillations in arterial diameters are independent of local neural activity and will add ‘noise’ to the hemodynamic signal, particularly when the brain is at rest and neural activity is low (Figure 3).

Additionally, awake animals (and humans) engage in a great deal of spontaneous movements (blinking, whisking, fidgeting movements) that drive spontaneous hemodynamic signals (discussed in Ref. [95]). These behavioral and arousal changes can cause large changes in the apparent strength of neurovascular coupling [96,97**,98]. Often, these state changes are not monitored, and may contribute to the differences in neurovascular correlations observed between studies. Awake mouse studies where behavior was not monitored have found higher correlations between ‘spontaneous’ neural and hemodynamic signals [33,67]. When these periods of movement or fidgeting are excluded, the correlations between spontaneous neural activity and hemodynamic signals are much lower [68**].

**Neurovascular decoupling and systemic confounds on BOLD imaging**

Not all BOLD increases are associated with detectable increases in neural activity, and conversely, not all neural activities seem to drive BOLD signals. Spontaneous BOLD events have been found to be independent from neural activity in anesthetized rodents [99**], and optogenetic stimulation of astrocytes can generate BOLD signals independent of neural activation [100]. Studies have shown no correlation or anti-correlations between hemodynamic and neural activity in the cortex [101] and caudate–putamen [91,102] under some conditions. Neural activity can rise in frontal cortex during locomotion with no detectable vasodilation or blood flow changes [82,103]. Perhaps most troublingly, both BOLD and CBV decrease in the caudate–putamen when local neural activity increases [104**]. These results show the coupling of the BOLD signal to local neural activity can vary greatly from region to region and condition to condition.

Finally, the BOLD signal can be influenced by systemic cardiovascular state and fluctuations, such as breathing rate [105]. The BOLD signal measured in the carotid arteries leads the global BOLD signal in humans [88] and is correlated with peripherally measured oscillations in blood volume [106], suggesting that systemic cardiovascular fluctuations account for some of the observed spontaneous observed BOLD signals. Interestingly, the absolute size of the neurally evoked blood flow increase is unaffected by baseline elevation of flow [107]. In subjects with elevated CBF, the amplitude of the evoked BOLD signal will be decreased without a corresponding change of neural activity.
Summary
BOLD imaging is the only non-invasive, whole-brain imaging modality available to study humans, and has been used to ‘decode’ visual perception and dreams in humans [108,109]. Activity in interneurons and other cells in the cortex releases vasoactive molecules. Dilations in the vasculature initiate in the arteries and their initial branches and passively spread to capillaries and veins. Sensory-evoked BOLD responses in primary sensory areas have been shown to be robustly correlated with neural activity. However, the strong sensory-evoked coupling between neural activity and hemodynamic signals seen in primary sensory cortex does not seem to be the default condition in the rest of cortex, or for spontaneous activity. Thus, inferring neural activity from hemodynamic signals (such as the BOLD signal) should be done with caution.

The weak coupling of hemodynamic signals to resting neural activity and the stronger coupling during sensory-evoked stimulation has important implications for computational models that use BOLD signals as surrogates measures of neural activity. In models of brain-wide resting activity, there is a mismatch between the observed BOLD dynamics and the dynamics predicted from structural connections in a computational model [109]. This mismatch has been interpreted as an emergent phenomenon of brain networks, but also rather could be a reflection the weak or non-canonical neurovascular coupling during rest in many brain regions. In contrast, computational models of visual object recognition and perception that are based on visually evoked BOLD responses in the visual cortices produce much better correspondence between the predicted and actual activity and/or perception [110–113]. This better correspondence of these models is likely because of the strong neurovascular coupling present in visual areas during visual stimulation. Understanding the limitations of hemodynamic signals in reporting neural activity is critical for using them to constrain and test computational models of brain function.

Funding
This work was supported by the National Institutes of Health grants R01EB021703, RF1MH114224, R01NS101353, and R01NS078168.

Conflict of interest statement
Nothing declared.

Acknowledgements
Thanks to C. Echagarrua and Q. Zhang for comments on the manuscript.

References and recommended reading
Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
- - of outstanding interest


The first two-photon measurements of arterial and venous dilations in the awake animal.


Single-vessel MRI images showing the contribution of individual arterioles and venules to the CBV and BOLD responses.


A biophysical model of the BOLD response using vascular reconstructions and oxygen measurements.


Arterial dilations in the brain propagate upstream rapidly through the endothelial cell layer.


This paper shows that neurovascular coupling mechanisms depend on the vessel type. Capillaries (pre-capillary arterioles) are dilated by astrocyte PGE2, while arteries are dilated by neurally released NO.


This paper systematically characterizes the anatomy of the mural cells present on the cerebral vasculature.


Sensory-evoked functional hyperemia is shown to initiate in the pene-trating arteries and their initial branches. Calcium signals in pericytes are driven by glutamate release.


This paper shows that optogenetic stimulation of interneurons produces vasodilation, followed by an NPY-dependent constriction.


Isolated optogenetic stimulation of pyramidal neurons causes a small increase in flow and a large consumption of oxygen, while isolated optogenetic stimulation of interneurons causes a small oxygen composition increase and large increases in flow.


First paper showing the activation of individual interneurons can cause vasodilations and vasoconstrictions in vitro.


First paper characterizing the differences in HRF dynamics across the brain.


BOLD signals in the internal carotid arteries are highly correlated with global BOLD signals, suggesting a systemic contribution.


This paper simultaneously measures spontaneous neural activity and BOLD signals simultaneously in anesthetized rats.


An extensive review of how ‘fidgeting’ behaviors (such as twitches, blinks and fidgets) can drive spontaneous hemodynamic and neural signals, and how these fidget-driven changes can be a potential confound for resting-state studies.


Arousal state transitions are tightly associated with large BOLD signals.


Spontaneous large amplitude BOLD signals in anesthetized rodents are driven by adrenergic activations of astrocytes, and are not coupled to local neural activity.


Noxious stimulation causes negative striatal BOLD and CBV signals, despite elevations in neural firing and metabolism.


Baseline level of blood flow affects the BOLD response.


