Progress in Neurobiology xxx (xxxx) xxx

Contents lists available at ScienceDirect

Progress in Neurobiology

journal homepage: www.elsevier.com/locate/pneurobio

Review Article

Imaging faster neural dynamics with fast fMRI: A need for updated models of the hemodynamic response

Jonathan R. Polimeni [a,b,c,*,1], Laura D. Lewis [a,d,**,1]

[a] Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Charlestown, MA, USA
[b] Department of Radiology, Harvard Medical School, Boston, MA, USA
[c] Division of Health Sciences and Technology, Massachusetts Institute of Technology, Cambridge, MA, USA
[d] Department of Biomedical Engineering, Boston University, Boston, MA, USA

ARTICLE INFO

Keywords:
High-resolution fMRI
Brain hemodynamics
Biophysical modeling
Blood-oxygenation-level-dependent contrast
Human neuroimaging
Fast fMRI
Hemodynamic response function

ABSTRACT

Fast fMRI enables the detection of neural dynamics over timescales of hundreds of milliseconds, suggesting it may provide a new avenue for studying subsecond neural processes in the human brain. The magnitudes of these fast fMRI dynamics are far greater than predicted by canonical models of the hemodynamic response. Several studies have established nonlinear properties of the hemodynamic response that have significant implications for fast fMRI. We first review nonlinear properties of the hemodynamic response function that may underlie fast fMRI signals. We then illustrate the breakdown of canonical hemodynamic response models in the context of fast neural dynamics. We will then argue that the canonical hemodynamic response function is not likely to reflect the BOLD response to neuronal activity driven by sparse or naturalistic stimuli or perhaps to spontaneous neuronal fluctuations in the resting state. These properties suggest that fast fMRI is capable of tracking surprisingly fast neuronal dynamics, and we discuss the neuroscientific questions that could be addressed using this approach.

1. Introduction

The most widespread tool for measuring whole-brain activity non-invasively in humans is functional magnetic resonance imaging (fMRI), which tracks changes in blood flow, volume, and oxygenation that occur alongside neuronal activity. Together these changes generate the blood oxygenation-level-dependent (BOLD) signal, which is the most robust and commonly used fMRI signal today. The fundamental limitation of hemodynamics-based techniques such as fMRI is that they reflect vascular responses to neuronal activity and not the neuronal activity itself, which creates challenges for the interpretation of the fMRI signals as measures of brain function (Logothetis, 2008). While fMRI has been shown to be capable of localizing brain activity at sufficient spatial and temporal scales to map large-scale functional organization of the human brain (Eickhoff et al., 2018; Glasser et al., 2016), today it is still unclear what are the spatial and temporal limits of the technique (Buxton et al., 2014). The spatial specificity of human fMRI has substantially advanced in recent years, and now enables imaging of fine-scale features of functional architecture such as cerebral cortical columns (De Martino et al., 2015; Nasr et al., 2016; Yacoub et al., 2008, 2007; Yang et al., 2019), cortical layers (Huber et al., 2017; Kok et al., 2016; Muckli et al., 2015; Norris and Polimeni, 2019), and small subcortical nuclei (Arcaro et al., 2015; De Martino et al., 2013; Denison et al., 2014; Faull et al., 2015; Satpute et al., 2013; Schneider et al., 2004; Sclocco et al., 2018), yet the temporal specificity of fMRI has received less attention. This may be, in part, because until recently fMRI acquisition technology was not capable of sampling the fMRI response faster than about a 2-second interval with a large enough brain coverage to be practically useful for

Abbreviation: ASL, arterial spin labeling; BOLD, blood-oxygenation-level-dependent; CBF, cerebral blood flow; CBV, cerebral blood volume; $CMRO_2$, cerebral metabolic rate of oxygen consumption; FWHM, full-width at half-maximum; GM, gray matter; HDR, hemodynamic response; HRF, hemodynamic response function; IRF, impulse response function; LGN, lateral geniculate nucleus; MEG, magnetoencephalography; MTT, mean transit time; SC, superior colliculus; SNR, signal to noise ratio; V1, primary visual cortex; VAN, vascular anatomical network; WM, white matter.

* Corresponding author at: Athinoula A. Martinos Center for Biomedical Imaging, 149 Thirteenth St. Suite 2301, Charlestown, MA, 02129, USA.
** Corresponding author at: Kilachand Center for Integrated Life Sciences and Engineering, 610 Commonwealth Ave, Boston, MA, 02115, USA.
E-mail addresses: jonp@nmr.mgh.harvard.edu (J.R. Polimeni), ldlewis@bu.edu (L.D. Lewis).
[1] Authors contributed equally.

https://doi.org/10.1016/j.pneurobio.2021.102174
Received 4 October 2020; Received in revised form 30 July 2021; Accepted 8 September 2021
Available online 12 September 2021

neuroscience studies. However, with the advent of rapid parallel imaging technologies (Barth et al., 2015; Chiew et al., 2018; Hennig et al., 2007; Lin et al., 2008; Moeller et al., 2010; Setsompop et al., 2016, 2012) it has become possible to rapidly (< 400 ms) sample the BOLD response with spatial coverage of the whole brain.

Now that image acquisition can be fast, the temporal resolution of the hemodynamic response itself is the critical limiting factor in the temporal resolution of fMRI. There is a long-held belief that the hemodynamic response (HDR) to neuronal activity is intrinsically slow or "sluggish", lagging behind by up to 6 s, and thus represents a delayed surrogate of neuronal activity. Indeed the full HDR timecourse is slow relative to neuronal activity, which is why many assume it does not contain neuronally relevant information at time scales faster than those of the HDR. However, several studies have demonstrated that some aspects of BOLD respond far more rapidly to neuronal activity than is commonly appreciated, and can contain information about neural activity with a higher temporal resolution than that of the full HDR—suggesting that fMRI is capable of reflecting what may be surprisingly fast neuronal dynamics.

While the HDR dynamics are fundamentally slower than neuronal dynamics, the intrinsic temporal resolution limit of fMRI is still unknown—despite keen interest in this question dating back to the introduction of fMRI (as evidenced by prior review articles; see Bandettini, 2002, 1999; Menon and Kim, 1999; and references therein). Early fMRI studies tested the ability of the BOLD response to reflect brief neural activity using the instrumentation available at the time (Savoy et al., 1995, 1994) and demonstrated that BOLD exhibited sufficient temporal *resolution* to detect neuronal activity driven by distinct stimuli separated in time by 8 s (i.e., 0.13 Hz) (Bandettini, 1999; Thomas and Menon, 1998) and potentially down to 4 s (i.e., 0.25 Hz) (Bandettini and Cox, 2000); see Fig. 1 for examples of the data used in these early studies. Because temporal resolution, typically defined as the minimal temporal separation required to discriminate or detect two stimuli as distinct BOLD responses, is fundamentally dependent on SNR (Green and Swets, 1966; Helstrom, 1964), estimates of temporal resolution also depend on the sensitivity of the fMRI measurement. Recently, using new fMRI technologies that provide higher sensitivity to detect small, rapidly varying fluctuations, it has been demonstrated that the BOLD signal can track neuronal oscillations up to 0.75 Hz (Lewis et al., 2016) and perhaps even 1 Hz (Lewis et al., 2017) in human visual cortex, with similar observations in auditory (Frühholz et al., 2020) and somatosensory (Cao et al., 2019) cortices. Other studies demonstrated that the HDR possesses sufficient temporal *precision* across trials to be able to detect the relative temporal order of BOLD activations in different brain regions down to hundreds of milliseconds (Lin et al., 2018, 2013; Menon et al., 1998; Miezin et al., 2000; Misaki et al., 2013; Savoy et al., 1994), and that BOLD response onset time differences on the order of 100 ms could be detected (Bellgowan et al., 2003). Thus there have been many successful efforts into finding meaningful sub-second differences in the BOLD response timing. The BOLD response has also been shown to be capable of reporting on differences in neural activity at the millisecond scale (Ogawa et al., 2000). Beyond this, there is mounting evidence that the earliest phases of the BOLD response may be more neuronally specific (Błażejewska et al., 2019; Goodyear and Menon, 2001; Shmuel et al., 2007). Also several recent demonstrations have shown that pattern classifiers across multiple voxels can decode responses as early as 2 s after stimulus presentation (Kohler et al., 2013; Miyawaki et al., 2020; Vizioli et al., 2018; Vizioli and Yacoub, 2018; Vu et al., 2016), indicating again that the fast components of the BOLD response carry neurally relevant information. Taken together, these studies suggest the fMRI signal responds rapidly to neuronal activity and reflects meaningful neuronal dynamics approaching the time scales of many cognitive processes.

This insight could be highly valuable for human neuroscience, as it could enable a new class of fMRI experiments that study the rapid dynamics of naturalistic high-level cognition. Attention, language, sleep, and perceptual awareness are all associated with 0.1–1 Hz dynamics that cannot currently be precisely localized, and these could be studied directly using fast fMRI approaches. In addition, while typical experimental designs are constrained to slow stimulus presentation, these fast responses could enable more naturalistic and freeform experimental designs, e.g., using movie stimuli or abstract reasoning tasks, tracking

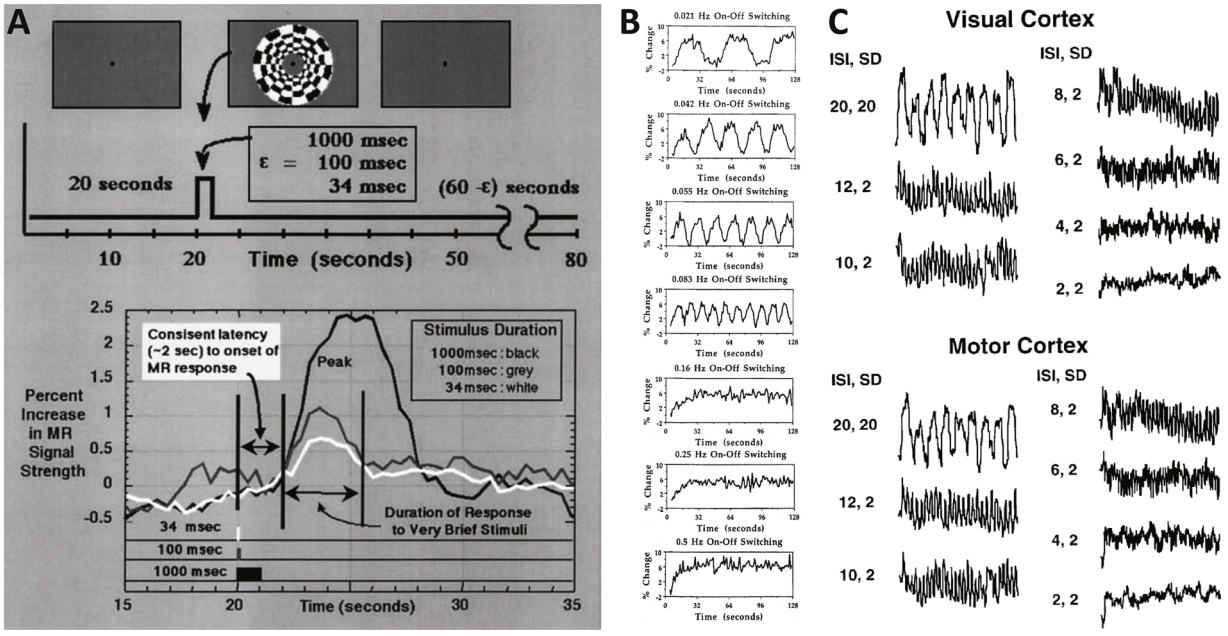

Fig. 1. Early investigations of the temporal limits of the BOLD response in humans. (**A**) BOLD responses measured in visual cortex to brief stimuli including 1-s, 100-ms, and 34-ms duration flickering checkerboards. The response to the shortest stimulus, 34 ms, was about 3.5 s in duration. (From Savoy et al., 1995, 1994; reproduced from Savoy, 2001, with permission.) (**B**) BOLD response measured in motor cortex during finger movement, with on-off frequency increasing from 0.024 Hz to 0.5 Hz. Distinct responses can be detected up to 0.083 Hz. (Reproduced from Bandettini, 1999, with permission.) (**C**) BOLD responses measured in visual and motor cortex to stimuli of varying inter-stimulus intervals (ISI) and stimulus duration (SD). Weak BOLD modulation can be detected in the 2-s on, 2-s off case (0.25 Hz). (Reproduced from Bandettini and Cox, 2000, with permission; see also Hathout et al., 1999a.)

cognition as it operates on its natural timescales. If a more complete understanding of the temporal resolution of fMRI can be established, together with tools for acquiring and appropriately analyzing fast fMRI data, it will likely enable new classes of experiments investigating brain computation.

Today there are several ongoing and emerging developments in instrumentation and pulse sequence technologies that will enable routine acquisition of fast fMRI with sufficient sampling rates and sensitivity (Bandettini and Wong, 2015; Budinger and Bird, 2018; Polimeni and Wald, 2018; Setsompop et al., 2016; Yacoub and Wald, 2018). These acquisition technology developments have been inspired in part by the demonstrated benefits of faster sampling seen in functional connectivity studies (Feinberg et al., 2010; Smith et al., 2012b; Uğurbil et al., 2013). Recent developments in time-series data preprocessing and analysis have addressed new considerations that arise with such rapidly-sampled data to perform proper statistical inference (Purdon and Weisskoff, 1998; Woolrich et al., 2001) by generating statistical models that account for the statistical properties of these data (Agrawal et al., 2020; Bollmann et al., 2018; Chen et al., 2019a; Corbin et al., 2018; Olszowy et al., 2019; Sahib et al., 2016). One of the most important remaining steps to fully realize the potential of fast fMRI is the development of appropriate biophysical models of the HDR that capture fMRI responses to rapidly-presented, brief stimuli/tasks—accurate detection and interpretation of the fast fMRI responses requires more accurate models of the BOLD signal for these unconventional paradigms. Although there have been many successful efforts at statistical modeling of fast fMRI data, biophysical models have received less attention. As we review below, it has long been known that the BOLD response for short-duration stimuli/tasks does not conform to the standard HDR models. Not only are improved biophysical models required to analyze fast fMRI data, they are also needed to guide future experimental design and push beyond existing paradigms to enable new discoveries. Motivated by this, the goal of this review article is to make the case that there is a need for the neuroimaging community to adopt an updated view of the hemodynamic response that is in line with these observations. Specifically, we discuss how using a single HDR can be problematic for fast fMRI, and discuss updated models that can help tap the potential of fast fMRI to image rapid neural dynamics.

2. Physiological basis of the fast hemodynamic response

Before considering the models of the HDR used in BOLD fMRI, we first review the physiological basis of the fast hemodynamic responses to neuronal activity. The speed of the HDR is determined by the metabolic and signaling pathways of neurovascular coupling and by the mechanical responses of both the vasculature and the blood. These mechanical responses are far slower and are thus thought to be the limiting factor on the delay between the neuronal activity and the corresponding HDR. The BOLD response to sustained neuronal activity peaks about 4–6 s after the start of the stimulus or task, then persists for the duration of the stimulus or task, and finally at the end of the stimulus or task returns slowly to baseline, often after a prolonged "post-stimulus undershoot" (see Fig. 2). For this reason, the BOLD response is often interpreted as a delayed and dispersed version of the neuronal activity, lagging behind by 4–6 s. However, the BOLD response onset begins almost immediately following neuronal activity, indicating that the vasculature in fact responds quickly to the increased metabolic demand of the neurons, with a build-up of the response over time and a peak occurring up to 4–6 s later.

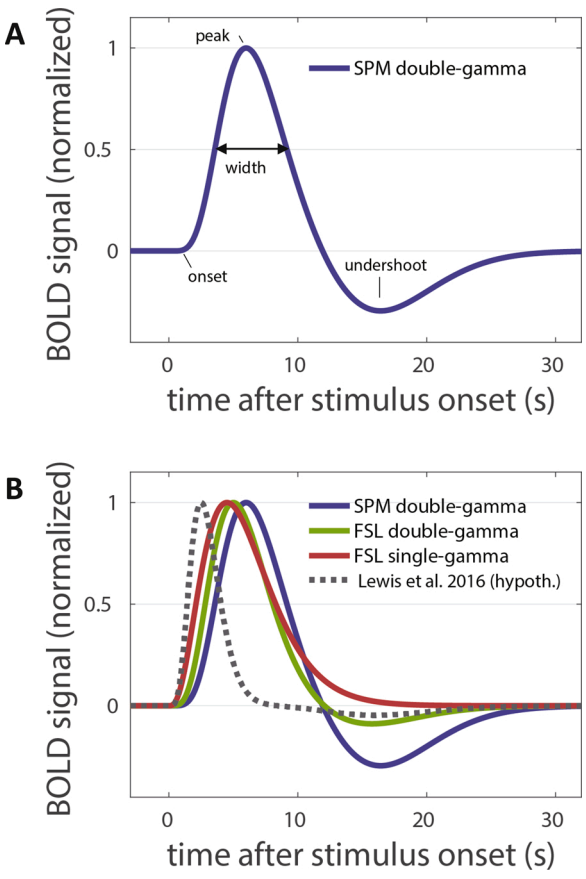

Fig. 2. Example HDR models that represent the BOLD impulse response. (**A**) The default HRF provided by the SPM package (Ashburner, 2012). The temporal features are characterized by the onset time, the peak time, the width (i.e., the FWHM) and the time and relative amplitude of the undershoot. This is often taken as the canonical HRF. (**B**) The default SPM HRF alongside two default HRFs provided by the FSL package (Jenkinson et al., 2012), including a double-gamma that includes a small undershoot and a single-gamma that does not; these HRFs have a slightly faster time to peak. Also included is a *hypothesized* HRF—consistent with data from a recent study (Lewis et al., 2016) reflecting measured responses to rapidly-oscillating stimuli—that is substantially faster than the canonical (peak time = 2.5 s, FWHM = 2.6 s). The HRF known to be faster for short-duration stimuli, thus the HDR is considered to be *nonlinear*, and thus the canonical HRF is not appropriate for rapidly-varying stimuli.

What may be underappreciated is that relevant neuronal information can be extracted from the early phases of the BOLD response.

Unlike heart and skeletal muscle, brain tissue lacks an oxygen store (Buxton, 2010; Buxton and Frank, 1997), and neurons therefore require a rapid supply of oxygen to support energy production through metabolism. Recent invasive optical microscopy studies in small animal models have demonstrated that individual intracortical blood vessels respond rapidly to local neuronal activity,[2] with the fastest responses occurring ~0.5 s or less after neuronal activity in animal models (Hirano et al., 2011; Kleinfeld et al., 2011; Nizar et al., 2013; Rungta et al., 2018; Tian et al., 2010; Wei et al., 2016). These rapid vascular dilations appear to be driven actively by smooth muscle cells on arterioles or pericyte

[2] These changes in vessel diameter appear to be driven by the same presynaptic signaling pathways that trigger postsynaptic neuronal activity, therefore vessels react *alongside* the postsynaptic neuronal activity (Attwell and Iadecola, 2002; Iadecola, 2004) (the "feedforward hypothesis" of neurovascular coupling (Iadecola, 2017; Uludağ et al., 2004)), not *in response to* the postsynaptic activity (the "feedback hypothesis" (Buxton, 2010)).

cells on capillaries or pre-capillary arterioles (Hall et al., 2014; Hamilton et al., 2010; Hill et al., 2015; Mazzoni et al., 2015; Peppiatt et al., 2006), although whether capillaries dilate *in vivo* remains controversial (Drew, 2019; Uludağ and Blinder, 2018). While the HDR is known to be much faster in rodents than in humans (de Zwart et al., 2005; Lambers et al., 2020; Silva et al., 2007), this suggests that there are mechanisms in place for rapid regulation of blood flow in brain microvasculature that are likely to be present in humans as well. New evidence points to additional mechanisms such as a rapid increase of flux through the capillaries due to changes in their mechanical compliance that accompany oxygen release, providing another potential mechanism for fast hemodynamic changes accompanying local increases in metabolism (Grutzendler and Nedergaard, 2019; Wei et al., 2016; Zhou et al., 2019).

Following the initial vascular response near to the site of neuronal activity, a cascade of vascular dilation occurs, including a *retrograde propagation of active vessel dilation* (i.e., in the opposite direction of blood flow) running upstream along the feeding arterioles (Chen et al., 2014; Figueroa et al., 2007; Segal et al., 1989; Segal and Duling, 1986), and in many cases an *anterograde propagation of passive vessel dilation* running downstream along the draining venules, indicating that there is a spread of the vascular response with time, away from the site of neuronal activity. After the neuronal activity ends, the vessels appear to return to their baseline diameter in a time-reversed manner (Chen et al., 2011; Devor et al., 2003). These observations of a rapid HDR based on optical imaging data have also been made in fast fMRI studies using small animal models that have observed an onset time of ~500 ms (Hirano et al., 2011; Silva and Koretsky, 2002; Tian et al., 2010; Yu et al., 2014). The spatial spread of the BOLD response outwards with time has been observed in both high-resolution animal (Hirano et al., 2011; Tian et al., 2010; Yu et al., 2012) and human studies (Błażejewska et al., 2019; Shmuel et al., 2007). The implication of these findings is that the HDR within the microvasculature appears to provide the highest spatial and temporal specificity, whereas the HDR within the macrovasculature (either on the arterial or venous sides) provides lower spatial and temporal specificity. It is thus reasonable to assume that the microvascular signal should provide the most specific surrogate for neuronal activity both in space and in time. While the potential of improved *spatial* specificity of the capillary response has been long appreciated (Frostig et al., 1990; Grinvald et al., 1986; Lai et al., 1993; Menon et al., 1995), the *temporal* specificity of the capillary response has received less attention.

This spreading of the HDR outwards from the capillary bed with time has other important implications for the temporal resolution of fMRI. The lag between the initial onset of the HDR at the capillary bed and the time that the HDR reaches the feeding arteries or draining veins at the pial surface immediately above the capillaries can approach 1 s (Chen et al., 2011; Gutiérrez-Jiménez et al., 2016; Hutchinson et al., 2006; Kriegeskorte et al., 2010; Rungta et al., 2018; Schmid et al., 2017; Yu et al., 2012), which induces a temporal dispersion of the HDR even within a relatively small volume. Also, because any given surface artery or vein will feed or drain multiple capillary beds within the cortex, such as the multiple beds present across cortical depths that span groups of cortical layers (Bell and Ball, 1985; Duvernoy et al., 1981; Pfeifer, 1928; Schmid et al., 2017), if multiple capillary beds at a given cortical location respond—which is often the case for stimuli/tasks that engage multiple neuronal populations within a local neuronal circuit—the responses seen at the large surface vessels will reflect an aggregation of the responses in these capillary beds. This effect will lead to temporal smearing or blurring seen in the responses measured from these large vessels. For fMRI signals that are sensitive to changes within the micro- and macrovasculature, such as gradient-echo BOLD, the different temporal phase lags between the micro- and macrovasculature will lead to further temporal smearing or blurring as well as *temporal phase cancellation effects* (see Fig. 8C below), which will act to reduce the amplitude of the measured response (Lewis et al., 2016). Therefore, acquisition techniques that can achieve microvascular specificity should provide higher temporal resolution as well as higher neuronal specificity.

The increased neuronal specificity of the early phase of the BOLD response has been explored in the context of the "initial dip", which has been exploited in fMRI studies to achieve improved neuronal specificity (Duong et al., 2000; Kim et al., 2000; Yacoub et al., 2001). Although it has been demonstrated in invasive optical imaging studies (Devor et al., 2003), its detection in fMRI studies remains elusive (Buxton, 2001). It is also conceivable that this initial dip may only be reliably detected with both fine spatial and temporal sampling (Tian et al., 2010). Recent evidence suggests that this initial dip may not only reflect a transient increase in the cerebral metabolic rate of oxygen (a *metabolic* signal) (Malonek et al., 1997) but may represent a rapid initial increase in cerebral blood volume (a *vascular* signal) (Hathout et al., 1999b; Havlíček and Uludağ, 2020; Hillman, 2014; Lindauer et al., 2001; Sirotin et al., 2009; Uludağ, 2010), yet it still appears to be more spatially specific than the peak positive response. This indicates that fMRI signals induced by *vascular* changes still precisely reflect neuronal activation: the former view, where metabolic signals are naturally more specific to neural activity than vascular signals, has been superseded by a view in which vascular signals themselves can be highly specific to neuronal activity. Whether or not a dip in the BOLD response can be detected, or whether it reflects a metabolic or vascular change, the earliest regime of the HDR appears to provide tighter spatial specificity (Menon and Goodyear, 1999; Shmuel et al., 2007; Yu et al., 2012). Thus, while the early phase of the HDR is increasingly thought to reflect a rapid vascular response, it still appears to provide increased specificity relative to the later positive peak (Sheth et al., 2004a).[3]

While there is strong evidence for fast hemodynamic response onset times that reflect neuronal activity, less attention has been paid to the temporal dispersion or width of the response (see Figure 2a). This temporal property also relates to the temporal precision of the fMRI signals as it will alter the low-pass filtering properties of the HDR. An examination of recently published data suggests that, for short-duration (2 s) neuronal activity, the width of the arteriolar dilation responses may be as narrow as 3 s (FWHM) in small animal models (Tian et al., 2010; Uhlirova et al., 2016), with a narrower response expected for shorter-duration stimuli. Again this feature is expected to vary across the various branches of the microvascular tree, and will most likely be longer in the human brain compared to the rodent brain. Still, this feature will strongly affect the frequency content of the HDR and therefore will also partly determine the temporal resolution of the fMRI signal.

Overall it appears that the vasculature does respond quickly to local neuronal activity, and while the HDR spreads out with time, if it can be sampled with sufficient spatial and temporal resolution to avoid smearing and cancellation effects fast fMRI should be capable of measuring much faster brain dynamics than the canonical view of the hemodynamic response might predict (Bandettini and Cox, 2000; Lewis et al., 2016).

[3] Because during neuronal activity the BOLD response begins by expanding outwards spatially, then when the neuronal activity ends the BOLD response returns to baseline by contracting inwards spatially, it suggests that either the early response before the peak or the late response after the peak should be more specific to the site of neuronal activity (Uludağ, 2010). In agreement with this view, the post-stimulus undershoot has been shown to provide improved neuronal specificity relative to the positive peak in some cases (Yacoub et al., 2006).

3. The canonical hemodynamic response function – "wrong but useful" for slow designs, yet problematic for fast fMRI

The standard model used for the analysis of fMRI time-series data assumes a standard or "canonical" hemodynamic response function (HRF)[4] . This default model is overly simplistic (Handwerker et al., 2012) but is useful for conventional fMRI experimental designs and is assumed in the majority of fMRI studies. While there are slight differences in the exact shape of this canonical HRF across software implementations, the basic features are the same (see Fig. 2). Typically it is modeled as a gamma function or a double-gamma function, which provide either a single positive peak or a positive peak followed by a post-stimulus undershoot, respectively, where the peak occurs at a delay of 4–6 s. Newer experimental evidence suggests that, for some stimulus conditions, the HRF may be substantially faster, as discussed further below.

The BOLD response to a task with known timing is often estimated as the stimulus/task timing convolved with the HRF, and in this way the BOLD response is modeled as a linear function of the neuronal activity timing (typically assumed to be equivalent to the stimulus/task timing, sometimes called the stimulus/task waveform), with the *convolution function* (i.e., a convolution kernel in time) acting as an Impulse Response Function (IRF).

Although the original articles that established the standard analysis approaches for fMRI assumed linearity of the BOLD response (Bandettini et al., 1993; Boynton et al., 1996; Friston et al., 1995), the linearity of the BOLD response was directly established by Boynton et al., who codified the linear systems analysis approach for BOLD fMRI (Boynton et al., 1996). In this linear model, the HDR is viewed as a linear time-invariant system that is fully characterized by an IRF in the time domain (or, equivalently, a "transfer function" in the frequency domain). This linear, time-invariant system is presumed then to exhibit linearity properties of *additivity* and *scaling* (if input $x_1(t)$ produces an output $y_1(t)$ and an input $x_2(t)$ produces an output $y_2(t)$, then a sum of the scaled input $\beta_1 x_1(t) + \beta_2 x_2(t)$ produces an output $\beta_1 y_1(t) + \beta_2 y_2(t)$) and time-invariance (whether the input occurs at time t or at a later time $t + \delta$ the output will be identical except for a delay of δ seconds). The eigenfunctions of linear time-invariant systems are complex exponentials (Oppenheim et al., 1997), therefore if the input is a sinusoid, the output will be equal to a sinusoid of the same frequency only scaled in amplitude and delayed in time.

Perhaps the most relevant implication the linearity assumption is that it allows for the *prediction* of a measured BOLD response by convolving the presumed input neuronal activity by the system IRF regardless of the timing (via the additivity property) and amplitude (via the scaling property) of the neuronal activity. While the linear system model of the BOLD response is often an excellent approximation, and the assumption of linearity is implicit in nearly all fMRI analyses performed today (in the context of the General Linear Model framework), the linearity assumption does not hold for many stimulus types. It is well known that for sustained stimulation longer than about 4 s in duration the linearity assumption holds well, while for shorter duration stimuli/tasks the BOLD response exhibits clear nonlinearity (Glover, 1999; Miller et al., 2001; Vazquez and Noll, 1998; Yeşilyurt et al., 2008), as the additivity property does not hold. In fact, some evidence for nonlinearity can be seen in the BOLD response to a 3-second stimulation in the original study establishing the linear system model (Boynton et al., 2012, 1996). The experimental evidence for this nonlinearity is reviewed further below. The fact that the assumption of a temporally linear system[5] does not hold for all stimuli means that the convolution kernel relating the stimulus waveform to the BOLD response itself changes with the details of the stimulus or task. The key point is that there is no one convolution kernel that generalizes to all stimuli or tasks, therefore there is no true IRF that characterizes the system. Critically, these timescales in which linearity breaks down correspond to those on which many if not most naturalistic neural dynamics take place; few real-world cognitive tasks involve a constant, unchanging stimulus for >4 s durations.

Not only is the HDR dependent on stimulus/task intensity and duration, but it is well known that the HDR varies systematically with many factors (Handwerker et al., 2012). For example, the HDR varies: across individuals (Aguirre et al., 1998; Miezin et al., 2000); with healthy aging (D'Esposito et al., 2003, 1999; Huettel et al., 2001) and development (Kang et al., 2003; Moses et al., 2014; Richter and Richter, 2003; Wenger et al., 2004); with common stimulants such as caffeine (Liu et al., 2004); and with baseline physiological state, such as basal vasodilation manipulated through CO_2 inhalation (Behzadi and Liu, 2005; Cohen et al., 2002; Kemna and Posse, 2001). The HDR is well-known to vary across the brain—not only across distant cerebral cortical areas (Chang et al., 2008; Handwerker et al., 2004) but also at more local scales such as within individual areas (Saad et al., 2003). This inconsistency of the HDR further complicates efforts to define the temporal resolution of fMRI.

While conventional block-design experimental paradigms, with relatively long stimulus/task durations and inter-stimulus intervals, are somewhat forgiving with respect to the exact form of the HRF (Aguirre et al., 1998; Greve et al., 2013; Handwerker et al., 2012; Lindquist et al., 2009)—especially when combined with acquisitions using slow temporal sampling rates—for faster event-related designs (Bandettini and Cox, 2000; Buckner et al., 1996; Dale, 1999; Dale and Buckner, 1997) a mismatch between the HDR model and the measured BOLD response can be problematic, leading to both increased estimation bias and variance (Greve et al., 2013; Lindquist et al., 2009; Wager et al., 2005). Although these modeling challenges can be partly addressed by more flexible frameworks that, instead of assuming a single fixed HRF model, utilize a basis set of functions that span the space of possible expected HRFs (Friston et al., 2000; Woolrich et al., 2004), our inability to predict the HRF for a given stimulus or task does produce challenges in experimental design (Dale, 1999; Lindquist et al., 2009; Liu et al., 2001). As mentioned in the previous section, it has been demonstrated experimentally that in particular the BOLD responses to brief, rapidly presented stimuli are far larger in amplitude than what would be expected from the canonical HRF (Bandettini, 1999; Bandettini and Cox, 2000; Cao et al., 2019; Lewis et al., 2018b, 2016; Thomas and Menon, 1998). In some cases, these estimates are off by an order of magnitude, suggesting that a more complete understanding of the HDR to weak or brief neural activation is required to design and analyze fast fMRI studies of rapid neural dynamics.

4. Nonlinear hemodynamic responses to weak or brief neural activation

There are several forms of nonlinearity that have been considered in the context of fMRI, including neuronal and hemodynamic nonlinearities. Neuronal nonlinearities include short-term effects such as neuronal transients at stimulus onset (Boynton et al., 1996; Buxton et al., 2004; Logothetis et al., 2001; Miller et al., 2001; Yeşilyurt et al., 2008) likely reflecting short-term neural adaptation (Gruber and Müller, 2005; Sharpee et al., 2006; Wark et al., 2007), neuronal refractoriness (Ogawa

[4] Here we use the term *hemodynamic response* to refer to the biological response to neuronal activity including the changes in blood flow and volume mediated by neurovascular coupling and the term *hemodynamic response function* to refer to the mathematical function that is used to describe this response; we have adopted the standard abbreviations HDR and HRF, respectively.

[5] A less commonly considered form of linearity is *spatial* linearity between nearby cortical locations (Hansen et al., 2004; van Dijk et al., 2020), which can be assessed by utilizing orthogonal stimulus conditions (e.g., in retinotopic cortex) (Hansen et al., 2004).

et al., 2000), or changes in excitatory/inhibitory balance (Havlíček et al., 2017b), which lead to a nonlinear transformation between the input stimulus/task timing and the neuronal response, as well as effects such as long-term adaptation or fluctuating attentional states occurring on the order of tens of seconds to minutes (Bandettini et al., 1997; Grill-Spector and Malach, 2001; Heckman et al., 2007; Krekelberg et al., 2006; Macdonald et al., 2011; Moradi and Buxton, 2013; Ou et al., 2009; Sadaghiani et al., 2009). While it is challenging to predict the neuronal response to a particular stimulus/task in human volunteers, several studies have compared MEG/EEG recordings with fMRI measures to estimate the neuronal response and compare it directly with the BOLD response (Chen et al., 2020a; de Zwart et al., 2009; Janz et al., 2001; Lewis et al., 2016; Liu et al., 2010; Yeşilyurt et al., 2010). Although the nonlinear neuronal response contributes to the observed nonlinear BOLD response, purely *hemodynamic* nonlinearities appear to contribute as well. In the current BOLD modeling framework used today (Buxton, 2012; Havlíček and Uludağ, 2020), which is explained further below, the prevailing view is that hemodynamic nonlinearities exist in the cascade linking neuronal activity to the BOLD response, such that Cerebral Blood Flow (CBF) appears to be linearly related to the neuronal activity, and the BOLD response appears to be nonlinearly related to CBF (Fig. 4) (Miller et al., 2001). While more direct experimental evidence for the tight coupling of neuronal activity and CBF will help determine whether this linear relationship holds in general, in this prevailing view a large part of the observed nonlinearity in the BOLD signal may be attributed to the transformation of the CBF response into a BOLD response (Friston et al., 1998; Miller et al., 2001; Yeşilyurt et al., 2008), therefore hemodynamics contribute to the nonlinearities seen in the BOLD response. Invasive electrophysiological recordings have demonstrated that neural nonlinearity cannot fully explain the observed nonlinearity seen in BOLD fMRI data (Li and Freeman, 2007; Sheth et al., 2004b). In addition, substantial differences in the observed BOLD nonlinearities can be observed with a fixed stimulus across magnetic field strength (Pfeuffer et al., 2003), and—since the dominant components of the hemodynamic response change with field strength (Uludağ et al., 2009)—this finding presents further evidence that there is an appreciable hemodynamic nonlinearity that accompanies any nonlinear components that are strictly neuronal in origin.

As described above, the nonlinearity of the HDR is typically characterized in terms of the additivity and, to a lesser extent, scaling properties. However, another category of effects that violate the assumption of a linear time-invariant system, known as hemodynamic "refractoriness", has also been explored, especially in the context of rapid event-related designs (Friston et al., 2000, 1998; Huettel and McCarthy, 2001, 2000; Inan et al., 2004; Soon et al., 2003; Wager et al., 2005; Zhang et al., 2008). This effect can be seen when comparing HDRs from a single stimulus to those of paired stimuli: when the second stimulus of the pair is presented during the refractory period (or post-stimulus undershoot) of the first stimulus of the pair, the HDR for the second stimulus experiences a reduction in amplitude and therefore is not a time-shifted version of the HDR for the single stimulus. While this may be viewed as a nonlinearity in that the HDR for the single stimulus cannot predict the HDR for the second stimulus of the pair, it may also be viewed as a violation of time-invariance in that the HDR for a given stimulus depends on when the stimulus is presented, and therefore this refractory effect indicates that the HDR can change with time when using short inter-stimulus intervals.

The most commonly considered form of nonlinearity that is perhaps the most relevant for fast fMRI is the violation of the additivity property, which can be seen readily as a disproportionately large BOLD response to short-duration stimuli (Birn and Bandettini, 2005; de Zwart et al., 2009; Glover, 1999; Hirano et al., 2011; Janz et al., 2001; Liu and Gao, 2000; Miller et al., 2001; Nangini et al., 2005; Pfeuffer et al., 2003; Robson et al., 1998; Soltysik et al., 2004; Vazquez and Noll, 1998; Wager et al., 2005; Yeşilyurt et al., 2008), which we will refer to as *amplitude effects*. In addition, the temporal properties of the HDR are different for short-duration stimuli compared to long-duration stimuli: the HDR to brief stimuli is faster than predicted from the linear system (Lewis et al., 2018b, 2016; Liu and Gao, 2000; Pfeuffer et al., 2003; Soltysik et al., 2004; Vazquez and Noll, 1998; Yeşilyurt et al., 2008), which we will refer to as *temporal effects*. The implication of this faster-than-expected HDR for brief stimuli is that the temporal resolution of fMRI is not a fixed quantity (even for a given subject at a given brain location) but rather itself depends on the stimulus. This means that some stimuli may be capable of probing faster neural dynamics than others, complicating the definition of temporal resolution in the context of fMRI.

Our recent study demonstrated this phenomenon by employing an oscillating visual stimulus to induce oscillating neural activity at various known frequencies of interest (Fig. 3), enabling quantification of fMRI response amplitude in the context of known neural timing (Lewis et al., 2016). This experiment demonstrated that the BOLD response can track neuronal oscillations up to at least 0.75 Hz, which was faster than previous reports. However, these data also demonstrated that the magnitude of these BOLD response oscillations far exceeded the predictions from the canonical HRF (i.e., the standard linear model), and the measured BOLD amplitudes became increasingly larger than the predictions with faster oscillation frequencies—up to an order of magnitude larger than predicted by the canonical HDR model for the 0.75 Hz stimuli. Simultaneous EEG recordings confirmed that the amplitude of neuronal activity did not change appreciably across stimulus frequencies. This suggests again that the HDR is able to track rapidly-fluctuating neural activity and can be used to detect short-duration neural events spaced closely together in time in a way that is not captured by existing models.

One possible explanation that is consistent with the findings of a nonlinear BOLD response to short-duration stimuli is that CBF is tightly coupled to neuronal activity, and, for short-duration stimulation, the CBF, Cerebral Blood Volume (CBV) and BOLD signal time courses are approximately scaled versions of each other. In this short-duration stimulus regime, the nonlinearity observed between the stimulus waveform and the measured BOLD response is viewed to be dominated by neuronal nonlinearities (e.g., a large initial transient neural response followed by a reduced sustained response amplitude) that cause the BOLD HRF to vary with stimulus duration. The components of the HDR and their relationship for short-duration stimuli is depicted in Fig. 4A as predicted using the current balloon model framework. These neuronal nonlinearities may be able to account for some observed amplitude effects such as BOLD responses to short-duration stimuli overpredicting BOLD responses to long-duration stimuli (Miller et al., 2001); however, neuronal nonlinearities alone cannot fully explain these amplitude effects in all cases (Lewis et al., 2018b).

For long-duration stimulation, additional hemodynamic nonlinearities become relevant that contribute to the observed temporal effects. Specifically, for long-duration stimuli there is *dynamic decoupling* between the CBF, CBV and BOLD signal time courses such that each component exhibits distinct temporal features that are seen most prominently during the time-course transients—e.g., the venous CBV response, which is the CBV component most relevant to the BOLD signal, is slower and therefore lags the CBF response (Kim and Kim, 2011a), and consequently the BOLD response is also slower than the CBF response (Fig. 4B). Recent empirical human fMRI data provides additional evidence in support of this dynamic decoupling between CBF and venous CBV following long-duration stimulation (Havlíček et al., 2017a). This dynamic decoupling likely occurs because, for longer-duration stimuli with sustained active arterial vasodilation of the feeding arterioles, there is a build-up of pressure in the draining venules causing a slow passive venous vasodilation (due to venous compliance (Mandeville et al., 1999b)). After the stimulus ends, there is a rapid active arterial vasoconstriction during the return to baseline and a slow passive venous vasoconstriction. Therefore, the asymmetry between the active control on the arterial side, which is directly coupled to neuronal activity, and the passive compliance on the venous side, which is partly uncoupled

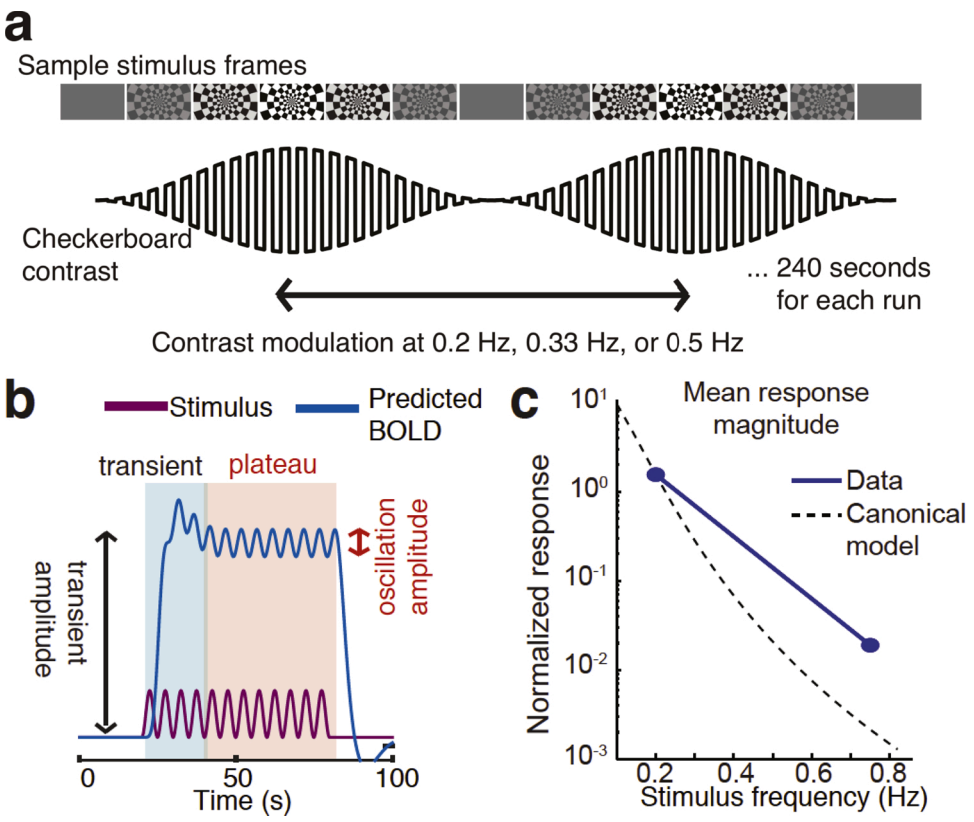

Fig. 3. Oscillatory stimuli elicit rapid BOLD oscillations an order of magnitude larger than canonical models predict. (**A**) Design of visual oscillatory visual stimulus used to probe fast fMRI responses in human visual cortex. Stimulus consisted of a standard radial checkboard flashing at 12 Hz, and the luminance contrast was modulated at several frequencies of interest (0.1 Hz, 0.2 Hz, 0.5 Hz, and 0.75 Hz). (**B**) Quantifying BOLD responses to oscillating stimuli. The contrast modulation is depicted by the purple trace and the expected BOLD responses depicted by the blue trace. After stimulus onset, the BOLD response exhibits a transient phase followed by a plateau (similar to the measured data presented in Figs. 1B & C); the BOLD oscillations and their amplitudes are measured around that plateau. (**C**) Mean BOLD response amplitude measured in visual cortex at 3 T (blue trace). The dashed line depicts the predicted amplitude of the BOLD response to this stimulus as a function of stimulus modulation frequency. The BOLD responses to the 0.75 Hz stimulus modulation were approximately 10× larger than those predicted from the canonical model. The data demonstrate the need for updated models of the hemodynamic response to predict rapidly varying BOLD signals. (Adapted from Lewis et al., 2016, with permission.)

from neuronal activity, can explain the observed nonlinearities that cause the temporal features of the HDR to change with stimulus duration. Note that this also implies that the BOLD response to short-duration stimuli may provide a much more faithful representation of neuronal activity (similar to CBF) compared to the BOLD response to long-duration stimuli, which can be seen by comparing the CBF and BOLD responses shown in Fig. 4A and B. This testable prediction is consistent with the current view of the BOLD nonlinearity (based largely on the balloon model), and may be plausible, however to our knowledge there is, to date, no empirical data to support this.

If a dynamic decoupling between the components of the HDR occurs for long-duration stimuli, why then is the observed BOLD response reported to be "nonlinear" only for short-duration stimuli? The current account is that for short-duration stimuli, the reported BOLD response nonlinearity is viewed to reflect mainly the nonlinearity in the transformation from the stimulus waveform to the neuronal activity, which changes with stimulus duration. Because short-duration stimuli generate short-duration BOLD responses, the entirety of the BOLD response occurs during this regime of tight CBF-BOLD coupling. The nonlinearity in the neural response, attributed to short-term adaptation, results in a substantial reduction in the amplitude of the neural activity after the first few seconds of the stimulus. Thus the current account implies that this "nonlinearity regime" is determined by the time scale of short-term neuronal adaptation.

As the stimulus continues, the prevailing view is that CBF and BOLD slowly become more and more decoupled, and as a consequence the BOLD response during the later phases of sustained stimulation is less closely reflecting neuronal activity. The gradual decoupling of the CBF and BOLD signals appears to take place over a longer time scale, perhaps on the order of 10–20 s. This dynamic decoupling will also result in a nonlinear HDR that changes with stimulus duration reflecting hemodynamic nonlinearities, however these temporal effects are less pronounced than the amplitude effects reflecting neuronal nonlinearities.

Another effect that must be considered is that, for longer-duration stimuli, the HDR reaches a steady-state during the period of sustained stimulation about ~6 s after the stimulus onset. During this sustained period, the neuronal activity, CBF, CBV and BOLD signals are all unchanging, such that any BOLD response to a sufficiently long-duration stimulus (i.e., beyond the "nonlinearity regime") can be reasonably predicted from a single convolution kernel, and the BOLD response is approximately "linear". In other words, the HDR for, e.g., an 8-second-duration stimulus and the HDR for a 32-second-duration stimulus should be well-predicted by the same convolution kernel. This specific convolution kernel generated from the 8-second-duration stimulus response will not be an IRF, since it will not be able to predict the BOLD response to a short stimulus "impulse" that is beyond the nonlinearity regime, but this convolution kernel can predict BOLD response generally for any stimuli with sufficiently long duration beyond the "linearity regime".

This current account does however leave several open questions. First, it implies that a "neural" nonlinearity contributes to amplitude effects manifesting in responses to short-duration stimuli, and a "hemodynamic" nonlinearity contributes to temporal effects that become pronounced with long-duration stimuli. However, even after factoring out neural nonlinearity effects, recent data suggest that the HDR to rapidly-varying stimuli exhibits a substantially larger BOLD response amplitude than predicted from standard models, suggesting some form of hemodynamic nonlinearity (Lewis et al., 2016). Today this observation is not fully explained. Similarly, recent work has explored whether a simple model of neuronal nonlinearity might explain observed temporal effects seen across a range of short-duration stimuli, and found that neuronal nonlinearity alone may not suffice to explain the measured BOLD fMRI data (Lewis et al., 2018b), suggesting again that hemodynamic nonlinearities may contribute to BOLD response nonlinearities associated with short-duration stimuli. Second, despite the many observations of this BOLD nonlinearity, existing quantitative modeling frameworks (reviewed below) do not seem capable of capturing the BOLD response behavior within both the nonlinearity regime and the linearity regime, and therefore do not generalize across a wide range of

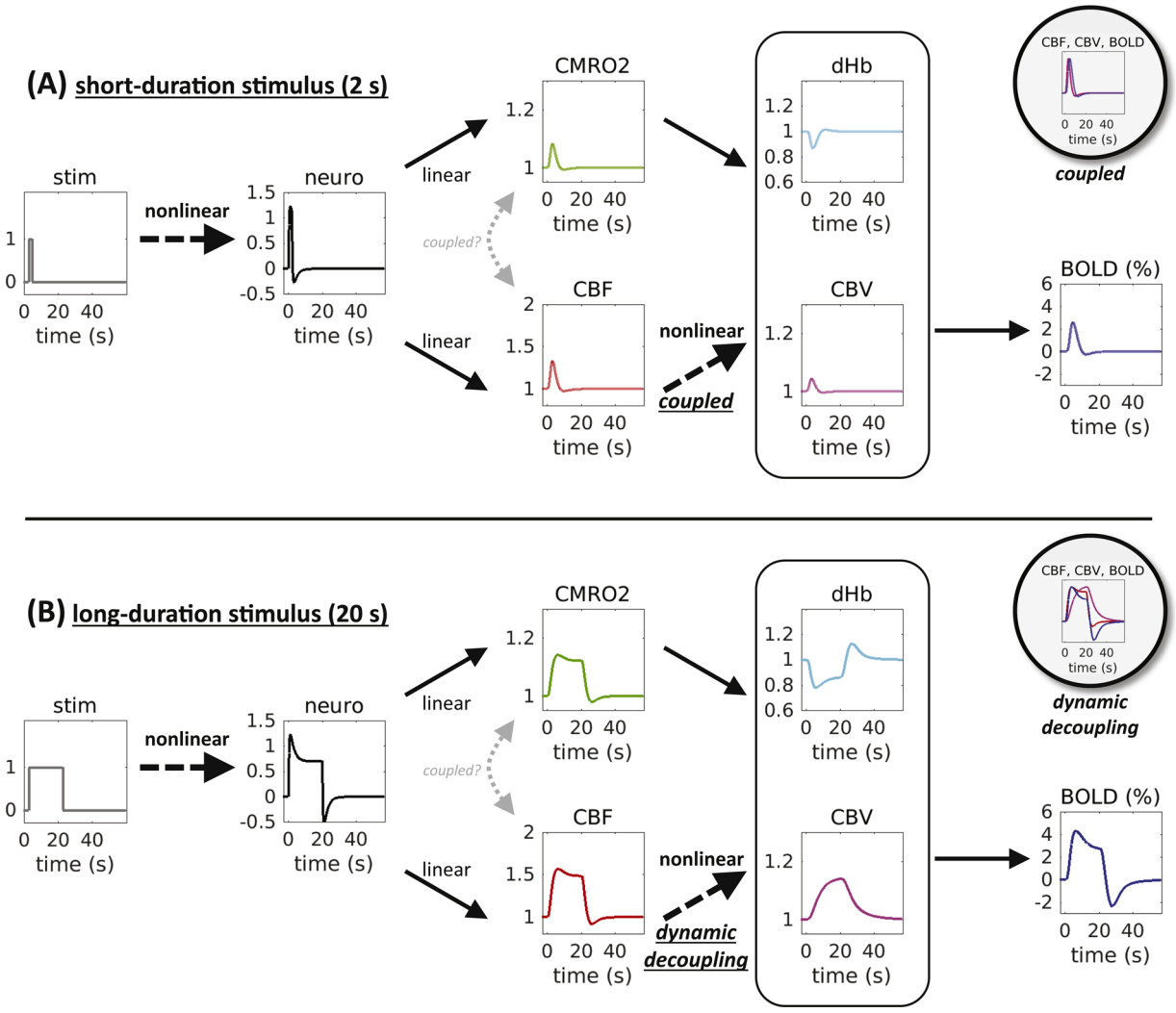

Fig. 4. Nonlinearities seen in the cascade linking the stimulus/task waveform to the BOLD response. (**A**) Current Balloon model framework applied to predicting the BOLD response to a short-duration (2 s) stimulus, using parameter values $\alpha = 0.15$ and $\tau_v = 0$ s (see text for description). (**B**) This same framework applied to predicting the BOLD response to a more conventional long-duration (20 s) stimulus, using parameter values $\alpha = 0.35$ and $\tau_v = 25$ s. The neuronal response begins with a positive transient reflecting excitation, followed by a reduced response during the sustained period reflecting short-term adaptation, and ends with a negative transient reflecting inhibition. *Dashed arrows* indicate stages of the cascade where stimulus duration is believed to affect the nonlinearity of the transformation, including the transformation between the stimulus waveform and the neuronal response, and the transformation between the CBF and CBV responses. For short-duration stimuli, a strong coupling is exhibited between CBV and CBF, however for long-duration stimuli a dynamic coupling occurs between CBV and CBV. The CBF-CBV decoupling translates into a consequent decoupling between CBF and BOLD, as shown in the *circle inset* in the top-right of each panel. (Also indicated in gray arrows is the potential for a dynamic decoupling between $CMRO_2$ and CBF, such as a $CMRO_2$ response that is faster than the hemodynamic response which would potentially give rise to an initial dip; here the CBF response is modeled as having no delay relative to the $CMRO_2$ response.) All quantities are plotted normalized to their baseline values, with the exception of the BOLD response which is in units of percent signal change. Parameter values were taken from the literature (Chen and Pike, 2009; Havlíček and Uludağ, 2020; Jones et al., 2001; Mandeville et al., 1999a; Uludağ and Blinder, 2018).

stimulus/task durations.

Beyond this nonlinear behavior that varies with stimulus duration, there are less-well-characterized sources of nonlinearity in the BOLD response. As noted above, the scaling property has also been shown to be violated, in that the HDR appears to be dependent on stimulus/task intensity. This manifests in two potentially distinct ways. First, several studies have shown that the BOLD response amplitude scales nonlinearly with stimulus/task amplitude, such that a large increase in stimulus intensity (e.g., luminance contrast in the case of visual stimuli) leads to only a modest increase in BOLD percent signal change, which then causes a modest increase in the spatial extent of activation for a fixed statistical threshold (Goodyear and Menon, 1998; Mohamed et al., 2002; Thompson et al., 2014). Some of this amplitude scaling can be attributed to purely neuronal effects—increasing stimulus luminance contrast by a factor of 2 is not expected to increase neuronal activity by a factor of 2, and so the main cause of this nonlinearity may not be hemodynamic in nature.[6] The HDR may also potentially saturate when local CBF and CBV reach their maximal values, causing a "ceiling" effect on the BOLD signal amplitude (Buxton et al., 2004), introducing a saturation nonlinearity that may also be reached more or less quickly as a function of stimulus intensity. This ceiling or plateau may represent a saturation of CBF response corresponding to a steady state in which sufficient numbers of

[6] By combining arterial spin labeling (ASL) and BOLD, and estimating $CMRO_2$ with the calibrated BOLD method (Davis et al., 1998; Hoge et al., 1999), it has also been shown that the coupling between CBF and $CMRO_2$ itself varies with many stimulus features including stimulus intensity (Buxton et al., 2014; Liang et al., 2013), which calls into question how to interpret the amplitude of the BOLD response in terms of neuronal activity.

feeding arteries are dilated to such that the energetic supply of oxygenated blood matches the metabolic demand of the neurons, or perhaps more likely a saturation of the BOLD response itself as further increases in CBF cause negligible decreases in the concentration of deoxyhemoglobin (Davis et al., 1998; Miller et al., 2001). It is unclear, however, under which physiological conditions and with which common stimuli/tasks this saturation in CBF or BOLD would be expected. It is also unclear how to properly model this potential effect. Simultaneous electrophysiological recordings may yield insight into the relative neuronal and hemodynamic contributions to the amplitude nonlinearity (Chen et al., 2020a; Liu et al., 2010; Yeşilyurt et al., 2010), although the balance of neuronal and hemodynamic effects are expected to vary with the details of the stimulus/task. While this nonlinear change in the BOLD response amplitude with stimulus/task duration provides insights into the physiology of the HDR, in practice it may not be relevant for fMRI studies because the amplitude of the BOLD response is typically not the quantity of interest (Buxton et al., 2014). However if this form of nonlinearity varies spatially then this phenomenon can cause problems (see below).

Second, there is mounting evidence that changes in stimulus intensity also modulate the *temporal* features of the BOLD response (Chen et al., 2020a, 2019b; Thompson et al., 2014; Vazquez and Noll, 1998; Yeşilyurt et al., 2008). This form of nonlinearity represents both a violation of additivity and of scaling, or an interaction between the two. Stronger stimuli appear to result in a BOLD response that begins earlier in time (Chen et al., 2020a; Thompson et al., 2014); however the earlier onset time (i.e., the time at which the BOLD response amplitude is detectable by being above the noise level) may be caused in part by the overall increased amplitude of the BOLD response to the more intense stimulus. What may be more surprising is that the temporal properties such as the time to peak and temporal dispersion (quantified as the temporal width of the main positive response) also increase with stimulus intensity (Chen et al., 2020a; Vazquez and Noll, 1998; Yeşilyurt et al., 2008). Although, as stated above, some amount of the observed nonlinearity in the measured fMRI signal can and should be attributed to neuronal effects, the differences in HDR timing in the range of 1 s have been observed across stimulus/task intensities (Chen et al., 2020a; Thompson et al., 2014; Vazquez and Noll, 1998; Yeşilyurt et al., 2008), which may be too long to be completely attributable to differences in neuronal activity. Given the scale of these timing differences it is likely that at least part of this dependence on stimulus/task intensity is due to hemodynamic effects. This form of nonlinearity may be interpreted in terms of the dynamic decoupling hypothesis stated above: strong stimuli may induce a larger CBF response and more downstream pressure in the venous compartment causing a venous CBV effect that induces an increased uncoupling between CBF and BOLD, whereas weaker stimuli may induce less downstream pressure engaging the venous compartment less leading to a decreased decoupling between CBF and BOLD. This speculative explanation would have to be validated empirically.

A potentially related finding is that other major temporal features of the HDR appear to be dependent on details of the stimulus presentation; for example, studies have shown that the post-stimulus undershoot is present after flickering visual stimuli but absent after static but otherwise identical visual stimuli (Havlíček et al., 2017a; Mullinger et al., 2017; Sadaghiani et al., 2009), or that the BOLD response is selective to other aspects of the stimulus timing (Gonzalez-Castillo et al., 2015, 2012; Harms and Melcher, 2003; Uludağ, 2008), giving rise to unexplainable features of the observed BOLD response. These changes in the temporal features may be more likely due to neuronal effects than purely hemodynamic effects that are the focus of this review, however these examples do highlight the complexities in relating stimulus/task dynamics to the dynamics of the measured fMRI signal.

In summary, the HDR violates the assumptions of a linear time-invariant system, and thus the temporal properties of the BOLD response, including its speed and amplitude, vary systematically with the stimulus configuration. In particular, both brief stimuli and weak stimuli lead to rapid HDRs that are faster than the canonical HRF. Both of the phenomena have been interpreted physiologically in terms of coupling/decoupling between the CBF and BOLD responses, as described above: for either brief or weak stimuli, there is less decoupling between CBV and BOLD. In other words, the vascular system is less perturbed away from its baseline steady-state for either brief or weak stimuli, under which circumstances the BOLD response is relatively rapid. While some evidence exists for differing ratios of CBF and BOLD response amplitudes with stimulus intensity (Buxton et al., 2014; Liang et al., 2013) as well as a dependence of BOLD temporal dynamics on baseline CBF (Behzadi and Liu, 2005; Cohen et al., 2002; Kemna and Posse, 2001), the gradual transition from a tight CBF-BOLD coupling for short-duration or low-intensity stimuli to decoupling for long-duration or high-intensity stimuli should be tested empirically in humans. A recently introduced experimental paradigm succeeded in demonstrating dynamic decoupling between CBF and BOLD for long-duration stimuli (Simon and Buxton, 2015) and could perhaps be extended to test whether short-duration or low-intensity stimuli exhibit stronger coupling between CBF and BOLD.

5. Spatial heterogeneity of fast BOLD responses

Because the temporal properties of the HDR dictate the temporal resolution of fMRI, and, as mentioned above, the HDR varies substantially across the brain (Handwerker et al., 2012), it follows that the detectability of BOLD responses to brief or weak stimuli will also vary across the brain. The question is whether the HDR varies systematically with any aspect of brain anatomy, and to what extent the degree of HDR *nonlinearity* (i.e., the change in HRF shape with changing stimulus duration or intensity) and the departure of the local HRF from the canonical HRF also varies systematically—which might allow investigators to know *a priori* how to interpret responses to brief or weak stimuli within individual brain regions.

As noted above, the HRF shape changes across brain regions, across cortical areas, and within cortical areas. The HRF shape even varies systematically across cortical depths, which may largely be attributable to proximity to large pial vessels, as demonstrated by directly measuring vascular anatomy using angiographic techniques (Bause et al., 2020; Chen et al., 2019b; Moerel et al., 2018). The effect in voxels adjacent to these large pial vessels is dramatic because large vessels on the pial surface are quite sparse in the human cerebral cortex (Bollmann et al., 2020; Duvernoy et al., 1981). Just as the HDR exhibits spatial heterogeneity, the severity of the observed nonlinearity in the BOLD response varies across cortical gray matter voxels (Birn et al., 2001; Huettel and McCarthy, 2001; Pfeuffer et al., 2003), across cortical areas (Glover, 1999; Miller et al., 2001; Soltysik et al., 2004), across cortical and subcortical responses (Lau et al., 2011; van Raaij et al., 2012) and even across individuals (Handwerker et al., 2004). Recent evidence also points to an increased linearity of the BOLD response within the cerebral cortical parenchyma (Gomez et al., 2020; Lewis et al., 2018b; Zhang et al., 2009), where the response is less influenced by large draining vessels on the surface, suggesting that BOLD responses from the microvasculature may exhibit more linearity and perhaps a stronger CBF-BOLD coupling compared to BOLD responses dominated by the microvasculature (Tak et al., 2015, 2014). This finding may also be interpreted in terms of the dynamic decoupling hypothesis stated above: regions with either higher venous density or higher baseline venous CBV may exhibit stronger venous CBV changes which induce decoupling between CBF and BOLD and influence the degree of nonlinearity observed. Veins farther downstream from the capillary beds may also experience a larger build-up of pressure during the HDR (Krieger et al., 2012), which may similarly induce larger delayed venous CBV changes that may also impact the nonlinear BOLD effect.

Beyond these observations of spatially varying nonlinearity, indicating a dependence of the HDR on the stimulus/task, even for a given stimulus/task the speed of the HDR varies across brain regions,

suggesting that some brain regions may be more suitable for fast fMRI experimental paradigms. Several studies have compared the HDR between cerebral cortical and subcortical gray matter and have shown that for a specific stimulus/task that elicits concurrent activation in multiple regions that the HDR is faster in subcortical nuclei (Lau et al., 2011; Lewis et al., 2018b; Yen et al., 2011) as shown in Fig. 5A. Within the visual system, the differences in HDR speed between the thalamus and visual cortex are greater than the differences across cortical depths within the visual cortex, as shown in Fig. 5B. As mentioned above, while some differences across brain regions may be attributable to differences in the timing of neuronal activity across these regions (since neuronal dynamics are likely to vary between e.g., thalamus and cortex), the observed differences in HDR timing are on the order of seconds, which is far too long to be explainable by neuronal differences alone. The observed differences may be due to regional differences in neurovascular coupling (Devonshire et al., 2012) or in vascular anatomy such as capillary density and path length between the capillary bed and the largest proximal draining veins (Duvernoy, 1999a, 1999b), or simply in the speed or reactivity of the local vasculature.

Comparing the HDR speed across brain regions highlights a potential trend in the speed of the HDR. Although data comparing the speed of the HDR across the brain are sparse, there is some preliminary evidence for a slowest HDR seen in the cerebral white matter (Li et al., 2019), then progressively faster responses seen in the cerebellar gray matter (Boillat and van der Zwaag, 2019; Chen and Desmond, 2005; Marvel and Desmond, 2012), followed by superficial then parenchymal cerebral gray matter (Kriegeskorte et al., 2010; Lewis et al., 2018b; Tian et al., 2010; Yu et al., 2012), subcortical gray matter such as the thalamus (Lau et al., 2011; Lewis et al., 2018b; Yen et al., 2011), then deeper subcortical gray matter in the brainstem (Faull and Pattinson, 2017; Lewis et al., 2018b), and potentially the fastest responses seen within the central nervous system just outside of the brain in the cervical spinal cord (Marcus et al., 1977; Nix et al., 1976; Piché et al., 2017; Sasaki et al., 2002). Although again some of these differences in the HDR across regions is likely attributable to differences in neuronal timing, the hypothesized progression of BOLD response delays across brain regions is summarized in Fig. 6. This hypothesis requires rigorous testing, and the biological significance of such gradient in HDR latency or speed across brain regions is unclear—and, naturally, some exceptions are expected given the heterogeneity of the metabolic demand and vascular density and morphology across the brain (Lauwers et al., 2008; Shannon et al., 2013; Shokri-Kojori et al., 2019; Tsai et al., 2009; Villien et al., 2014)—systematic comparison of regional differences in the HDR would be practically useful for fMRI studies. For example, an atlas of region-dependent HDR responses, combined with an atlas of regional vascular density (Bernier et al., 2018), may provide some insight into the anatomical/physiological determinants of the HDR timing (Chen et al., 2020b), which could potentially yield some ability to predict or model the HDR timing on the basis of additional anatomical information. Until there is a better understanding of the HDR in the form of regional models of the expected fMRI response to rapidly-varying neuronal activity, for now these regional differences in the speed of the HDR do indicate that, in practice, some brain regions lend themselves more to fast fMRI paradigms than others.

6. Updated models of the HDR

Given this nonlinearity in the BOLD response that is seen for brief or weak stimuli, there is a need for updated models of the HDR for fast fMRI that can account for these observations. Ideally an updated model would not only generalize sufficiently so that it could be used to design novel experimental paradigms, but it would also provide biophysical interpretability that could guide new acquisition and analysis approaches as well.

Early attempts to model the HDR in both linear and nonlinear response regimes proposed mathematical frameworks based on systems identification theory to generate parameterized convolution kernels, or expansions of convolution kernels into basis sets, to account for the observed deviations from the canonical HDR (Friston et al., 1998; Vazquez and Noll, 1998; Woolrich et al., 2004) but despite their success for conventional experimental paradigms these *convolution models* did not have a biophysical interpretation and therefore are not easily extended to account for new physiological insights. The first models inspired by vascular physiology were the venous "balloon" model (Buxton et al., 2004, 1998; Mildner et al., 2001; Obata et al., 2004) and the post-arteriole "windkessel" model (Mandeville et al., 1999b), which both viewed each voxel as containing a full vascular tree including feeding arterioles, draining venules, and the capillary bed. These *single--compartment lumped models* accounted for the steady-state and dynamic relationships between CBF as the input component, CBV and the cerebral metabolic rate of oxygen consumption ($CMRO_2$) as intermediate components, and the BOLD signal as the output component. Later versions of the models using this framework also accounted for the nonlinear transformation between the stimulus timing and the neuronal response (Buxton et al., 2004; Miller et al., 2001). These models fit a wide variety of data with a small number of parameters, although with

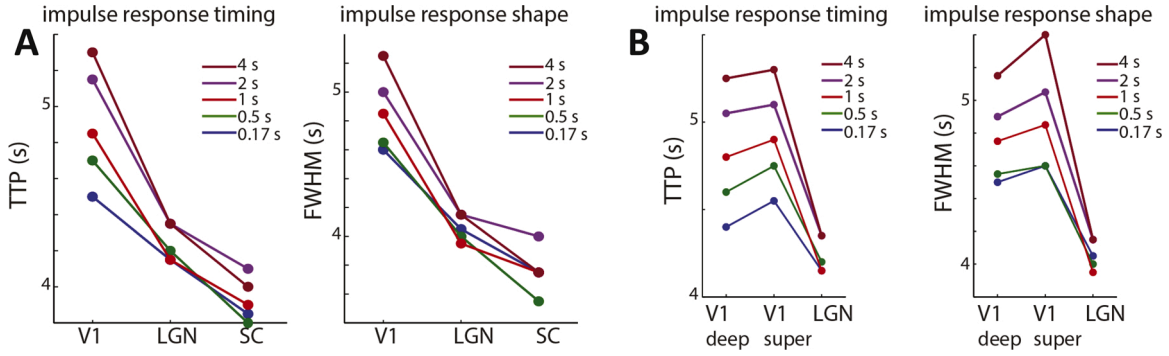

Fig. 5. The temporal properties of the BOLD "impulse response function" vary with stimulus duration and with brain region. Impulse response functions (i.e., convolution kernels) were estimated from BOLD response to visual stimuli with varying durations presented in a conventional block design; on-block durations were 4 s, 2 s, 1 s, 0.5 s, and 0.17 s. (**A**) The stimulus-specific impulse response function timing (measured as the time to peak or TTP of the impulse response) and shape (measured as the full-width at half-maximum or FWHM of the impulse response) were estimated along the visual pathway at the primary visual cortex (V1), the lateral geniculate nucleus (LGN) of the thalamus, and the superior colliculus (SC) of the brainstem. Both TTP and FWHM became systematically shorter with shorter stimulus durations in all three brain regions. (**B**) A similar analysis applied to the superficial and deep cortical depths of V1, with the timing and shape parameters plotted from LGN as a reference. Both response timing and response shape were systematically faster in the deeper cortical depths compared to the superficial depths, but both were consistently slower than the responses measured within the LGN, suggesting that the fastest responses in V1 are still slower than the responses in LGN. (Reproduced from Lewis et al., 2018b, with permission.)

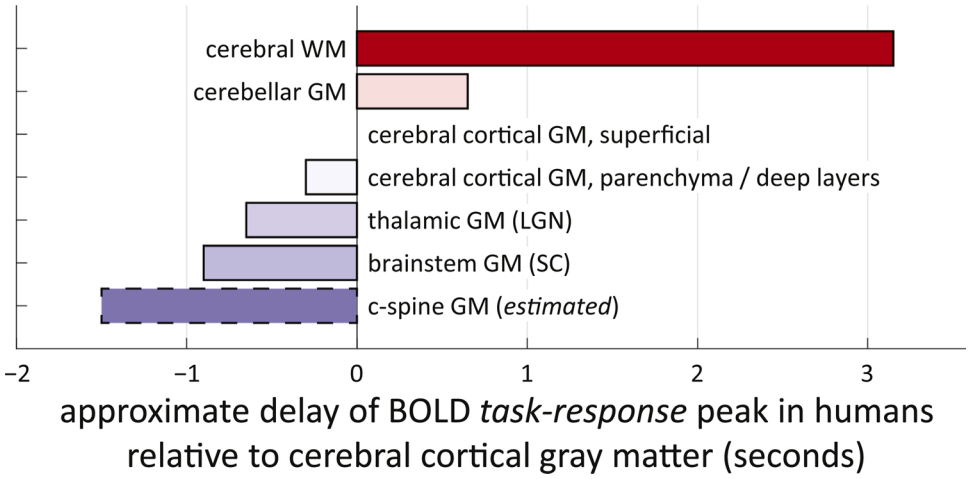

Fig. 6. Tentative hypothesis of the progression of BOLD response delays across brain regions, relative to the cerebral cortical GM as a reference. Each bar represents the approximate delay of the block response relative to cerebral cortical GM as reported in the original studies, with the exception of the c-spine GM which did not record cerebral cortical GM responses; this value is therefore roughly estimated. While these measurements across brain regions were each obtained using different stimuli and different experimental conditions, thus should be interpreted with caution, plus there are likely differences in the timing of neuronal response, given that the time scales are longer than the expected neuronal timing differences this suggest that at least some of this trend may be explained by hemodynamic difference. Values were obtained from the following studies: cerebellar GM (Boillat and van der Zwaag, 2019; Chen and Desmond, 2005); cerebral WM (Li et al., 2019); cerebral cortical GM, deep layers (Lewis et al., 2018b); thalamic GM (lateral geniculate nucleus, LGN) (Lewis et al., 2018b; Yen et al., 2011); brainstem GM (superior colliculus, SC) (Lewis et al., 2018b), c-spine GM (Sasaki et al., 2002).

the accumulation of more specific experimental data the exact physiological interpretation of the model parameters came into question (Buxton, 2012; Gagnon et al., 2016a; Griffeth and Buxton, 2011).

One of the first updates to the single-compartment lumped model was to include the possibility of arterial dilation. With the availability of invasive optical imaging recordings of blood flow and volume at the individual capillary scale (Devor et al., 2007; Hillman et al., 2007; Kleinfeld et al., 1998) it became clear that a key assumption of the original lumped models—that there should be negligible arterial response to neuronal activity—did not match the experimental data, and arteries did exhibit larger-than-expected dilations (Buxton, 2009). Similar effects had been seen in MR-based measures that could track CBV changes specific to the arterial compartment (Kim et al., 2007; Lee et al., 2001). In fact, in many cases venous CBV was instead found to be negligible: intracortical veins did not appear to dilate in response to short-duration stimuli (Berwick et al., 2005; Fernández-Klett et al., 2010; Hillman et al., 2007), but only dilated in response to longer, sustained stimulation (Drew et al., 2011) that presumably caused a build-up in pressure on the venous side of the vascular tree. Note that these experimental observations of negligible venous dilation following short-duration stimuli support the dynamic decoupling model described above (Havlíček et al., 2017a; Uludağ and Blinder, 2018). The findings of an arterial change in CBV and a stimulus-duration-dependent venous change in CBV subsequently inspired *multi-compartment lumped models* such as the "bagpipe" model that allowed for independent arterial and venous dilation (Drew et al., 2011), and other models that included arterial, capillary, and venous sections (Griffeth and Buxton, 2011; Kim et al., 2013; Kim and Ress, 2016; Zheng et al., 2005).

Still, most conventional lumped models consider only large fMRI voxels that each contain the entire vascular tree. Motivated by the increasing availability of fMRI data with high spatial resolution, recent extensions of the multi-compartment lumped hemodynamic models explicitly account for the hierarchy of vascular compartments within the cerebral cortical gray matter. With conventional voxel sizes, each voxel contains a similar heterogeneous mix of vasculature spanning the large supplying pial surface arteries, the intracortical diving arterioles, the small pre-capillary arterioles, the capillary bed, the small post-capillary venules, the intracortical ascending venules, and the large draining pial surface veins. With newly-available smaller voxel sizes far smaller than the cortical thickness, there is less heterogeneity *within* any given voxel, however there is more heterogeneity *across* voxels, as neighboring voxels will sample from different levels of the vascular hierarchy. To address this, newer *hierarchical lumped models* account for the HDR within different cortical depths, which corresponds roughly to different levels of the vascular hierarchy, and also account for the coupling of the HDRs across neighboring voxels (Havlíček and Uludağ, 2020; Heinzle et al., 2015; Markuerkiaga et al., 2016). In the context of fast fMRI, allowing for the increased degrees of freedom in these extended, spatially-coupled models (and, with it, more free parameters to assign) can better account for the vascular anatomy and physiology, yet thus far these more flexible hierarchical lumped models have only just begun to account for the delays of the various venous and arterial components of the HDR across the vascular hierarchy from parenchyma to pial surface (Havlíček and Uludağ, 2020). Therefore for fast fMRI there is a need for models that can account for these timing differences across the vascular hierarchy.

While these lumped models are often derived directly from experimental data, there is another class of models, *vascular anatomical models*, that attempt to generate biophysical simulations of the fMRI signals from first principles using explicit representations of vascular anatomy. These biophysical models originally represented the vasculature as randomly spaced and oriented infinite cylinders and the blood as being static (Bandettini and Wong, 1995; Boxerman et al., 1995; Fisel et al., 1991; Ogawa et al., 1993; Uludağ et al., 2009) and were used to characterize the physics of the fMRI signal formation. These *random cylinder models* could be as the basis for numerical simulations (Bandettini and Wong, 1995; Boxerman et al., 1995; Fisel et al., 1991; Ogawa et al., 1993; Uludağ et al., 2009) and based on realistic distributions of vessel size and type (Berman et al., 2018), or as the basis for analytic expressions of simplified systems (He and Yablonskiy, 2007; Kiselev et al., 2005; Kiselev and Posse, 1999; Troprès et al., 2001; Yablonskiy and Haacke, 1994), but because they were not connected they could not be used to model blood flow through the vasculature or fMRI signal dynamics. Vascular anatomical models based on random cylinders were later extended into *vascular anatomical network (VAN) models* that similarly accounted for the coupling of the HDR between adjacent vascular compartments. Although the HDR and BOLD signal components in hierarchical lumped models are typically phrased in terms of *mesoscopic quantities* that are measurable non-invasively with MRI such as CBF (i.e., perfusion) and CBV (Havlíček and Uludağ, 2020; Heinzle et al., 2015), the hemodynamics represented these in VAN models are expressed in terms of *microscopic quantities* that are measurable with optical microscopy such as blood velocity, vessel diameter, blood oxygen saturation, vessel compliance, and blood pressure, making it

somewhat challenging to directly relate these two frameworks. Early simplified VAN models (Boas et al., 2008) therefore were used to model blood flow and volume dynamics that, when coupled with an MR physics simulation engine, could simulate BOLD dynamics.

Perhaps the most significant recent step for fMRI modeling was the extension of the simplified VAN model into the *realistic VAN model* in which a realistic reconstruction of vascular anatomy based on optical microscopy served as the basis of the modeling (Gagnon et al., 2016a, 2015). These realistic VAN models combine full reconstructions of the microvascular network (Blinder et al., 2013) with *in vivo* measures of blood flow and vessel diameters both in the baseline and active states recorded across the vascular hierarchy (Tian et al., 2010; Uhlirova et al., 2016), then use numerical simulations to estimate blood flow and oxygenation throughout the vascular tree (Gould and Linninger, 2015; Lorthois et al., 2011a, 2011b; Peyrounette et al., 2018; Reichold et al., 2009; Secomb et al., 2004), from which the BOLD signal can be computed numerically using conventional techniques (Bandettini and Wong, 1995; Boxerman et al., 1995; Fisel et al., 1991; Ogawa et al., 1993). These realistic VAN models have been applied to simulate baseline physiology (Gould et al., 2017; Schmid et al., 2017) and MR signal characteristics such as model parameters used in lumped models (Báez-Yánez et al., 2017; Cheng et al., 2019; Gagnon et al., 2016a; Pfannmoeller et al., 2019; Pouliot et al., 2017), but they have also been applied to modeling BOLD dynamics (Gagnon et al., 2016a, 2016b, 2015). Remarkably, these models, based on microvascular anatomy and physiology, were able to predict previously undiscovered BOLD signal characteristics that were subsequently confirmed experimentally in conventional human fMRI data (Gagnon et al., 2015; Viessmann et al., 2019b, 2019a). However, perhaps due to their complexity, these realistic VAN models have not yet been applied to modeling temporal features such as nonlinearities in response to brief or weak stimuli, and more experimental data from invasive optical imaging may be needed to enable these models to generalize to new stimulus/task configurations.

How then might an existing modeling framework be used to explain the nonlinearities that are central to predicting fast responses to brief or weak stimuli? As articulated in the dynamic decoupling hypothesis, stated above, the observed nonlinearity of the BOLD signal can be understood in terms of delayed venous CBV relative to CBF, and during the early phases of the HDR CBF and BOLD are tightly coupled while in the later phases the increase in venous CBV induces a decoupling between CBF and BOLD. This could be captured in the framework of a realistic VAN model through a venous compliance parameter, however additional experimental data may be required to constrain this parameter properly. Instead we can consider the simple single-compartment lumped model (a.k.a., the updated balloon model (Buxton et al., 2004; Griffeth and Buxton, 2011)). In this model, the *steady-state* relationship between CBF and CBV is summarized by the empirical Grubb equation $\mathrm{CBF} = \mathrm{CBV}^{1/\alpha}$ (Davis et al., 1998; Grubb et al., 1974) parameterized by exponent α, while during transient states the relationship (ignoring normalization and scaling factors) is typically represented roughly as $\mathrm{CBF}(t) = \mathrm{CBV}(t)^{1/\alpha} + \tau_v\, \mathrm{dCBV}(t)/dt$, where the parameter τ_v, known as the venous CBV viscoelastic constant, controls the extent to which venous CBV lags CBF (Buxton et al., 2004; Havlíček and Uludağ, 2020; Mildner et al., 2001). The venous compartment "initially resists a change in volume, but eventually settles into a new steady-state" (Buxton et al., 2004). In this form, if the value of the viscoelastic parameter τ_v were set to zero, the changes in CBV and this in the BOLD signal would closely track changes in CBF (see Fig. 4). It has been suggested that this value of τ_v is not a fixed quantity, and may be different during venous inflation and deflation (Buxton et al., 2004, 1998), and the value of τ_v may also change with stimulus duration such that for short-duration stimuli $\tau_v \approx 0$ whereas for long-duration stimuli τ_v takes on a larger value (Havlíček et al., 2017b; Uludağ and Blinder, 2018). If in fact these parameter values themselves evolve during the HDR and can be viewed as functions of stimulus duration this introduces yet another potential nonlinearity into the modeling framework. It may also be the case that different vascular compartments may exhibit different levels of the venous ballooning effect and may contribute differently to the BOLD nonlinearity—for example, it may be that the pial veins and parenchymal veins contribute in different ways to the ballooning effect (Gao et al., 2015; Uludağ and Blinder, 2018), which would further contribute to the spatial heterogeneity of the observed nonlinearities. Other parameters of the established lumped model also control the nonlinearity of the BOLD response, including the Grubb parameter α, which controls the steady-state relationship between CBF and CBV, and controls the influence of CBV on the BOLD signal and thus affects the temporal properties of the BOLD signal including the transients (Havlíček et al., 2015; Uludağ and Blinder, 2018); whether the value of α is stimulus dependent (Chen and Pike, 2009; Jones et al., 2001; Mandeville et al., 1999a; Zheng and Mayhew, 2009) warrants further investigation. Additionally, the baseline mean transit time (MTT) through the venous compartment is an additional parameter that represents the relationship between baseline CBF and CBV (Boas et al., 2003; Buxton et al., 1998) that has also been shown in simulations to affect BOLD response transients and nonlinearity (Lewis et al., 2016); given that baseline physiology has been shown experimentally to affect the BOLD response dynamics (Behzadi and Liu, 2005; Cohen et al., 2002; Kemna and Posse, 2001), it is likely that this parameter may also play an important role in modeling BOLD nonlinearity.

Finally, the question arises of how to best utilize these models in practice. While, as stated above, updated models of the HDR can be useful in planning new experiments, and determining, for example, what might be the fastest-detectable BOLD response for a given stimulus paradigm, the complexity of the modern modeling frameworks presented above suggests that they are usually not suitable for direct use in data analysis. Several groups have proposed methodologies for fitting nonlinear models with several free parameters directly to BOLD time-series data (Deneux and Faugeras, 2006; Havlíček et al., 2015; Vakorin et al., 2007; Vazquez et al., 2006). Recent work has shown how these nonlinear physiological models can be used to simultaneously explain multiple physiological components underlying the observed BOLD response, including experimentally measured neural activity derived from electrophysiological recordings (Havlíček et al., 2017b). However, since fast fMRI signals are very low-amplitude, it will typically not be practical to apply these complex models to fast fMRI studies of neural processes themselves. Instead, we suggest that these modeling studies are highly valuable for developing an understanding of how hemodynamics are expected to appear in the context of more naturalistic neural activity. These models can then be translated into more simple analysis approaches to analyze fast fMRI data. While many studies have used stimulus-specific, and even voxel-specific, HRFs derived directly from the measured fMRI data in order to reduce bias (Kay et al., 2008a; Pedregosa et al., 2015), these estimated HRFs themselves will be noisy and will likely increase the variance of the measured fMRI activation. Also, because the HDR changes with stimulus configuration, caution should be exercised when attempting to use an HRF derived from auxiliary fMRI data (akin to a functional localizer) to analyze the fMRI data of interest. Nevertheless, although there is ongoing work in these new modelling directions, and continued progress is expected, the existing models adapted to brief stimuli have already provided some insight into the possible origins of several features of the HDR. A promising future direction is to use the valuable insights from these modeling and experimental studies of context-dependent HDR dynamics to establish principles for how the HDR varies across tasks and regions (for example, how the values of lumped model parameters α and τ_v vary with stimulus/task configuration or with brain region) (Pfannmoeller et al., 2021), and then in turn use this to implement simple HDR models that can be applied to fast fMRI data.

Updating our models of the fast fMRI response, and deriving appropriate parameter settings, will likely require additional data on vascular anatomy and physiology, which often comes from invasive recordings in animal models. However, these parameters cannot be

directly applied to humans. The hemodynamics are likely matched to the spacing and dimensions of vessels (Turner, 2002), and it is well known that there are substantial species differences in vascular anatomy, including pial vessel topology (Mchedlishvili and Kuridze, 1984), the ratios of intracortical arterioles and venules (Schmid et al., 2019) and even the capillary network (Smith et al., 2019). Furthermore, there have been reports of species differences in the temporal properties of the BOLD response (de Zwart et al., 2005; Lambers et al., 2020; Silva et al., 2007); a recent example of the differences in the HRF between rats and humans is presented in Fig. 7, showing that the rat HRF is substantially faster than the human HRF. Another major consideration is the well-known effects of anesthetics on hemodynamics (Gao et al., 2017; Martin et al., 2006; Pisauro et al., 2013; Schlegel et al., 2015; Xie et al., 2020) and some evidence suggests that hemodynamics are faster in awake animals (Ma et al., 2016) and the nonlinearity of the BOLD response may differ between awake and anesthetized states (Martin et al., 2006). Therefore, any updated models will require careful validation to confirm applicability in humans.

7. Extensions of fast fMRI to resting-state designs, non-BOLD fMRI, and laminar fMRI

The canonical HRF was originally derived from long-duration stimuli commonly used for block-design experimental paradigms, but does not generalize to all stimuli, in particular brief or weak stimuli. This recognition is important, not only because it indicates that caution must be exerted when analyzing responses to brief or weak stimuli, but also because any estimation of the temporal resolution of fMRI based on considering the temporal properties of canonical HRF will be inaccurate. It also highlights that improved temporal resolution may be achieved with rapidly changing, subtle stimuli or cognitive tasks. Similarly, fast fMRI responses may be prevalent in the resting state, if these resting-state fluctuations can be viewed as responses to unknown yet brief or weak neuronal events (Chen and Glover, 2015), although other studies have reported slower HRFs in the resting state (de Munck et al., 2007; Logothetis et al., 2009; Yeşilyurt et al., 2010). Future work may further test what is the temporal (and spatial) specificity of these resting-state fluctuations.

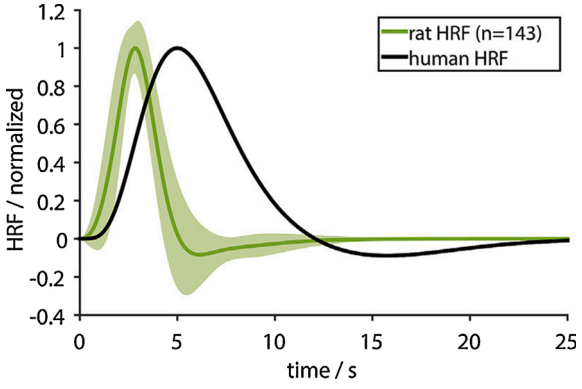

Fig. 7. Differences in HDR timing across species. Estimated BOLD impulse response from the rat cerebral cortical gray matter obtained from averaging 143 BOLD measurements in the primary somatosensory cortex of 76 animals under anesthesia; the green trace represents the average impulse response function and the shading represents its standard deviation. The black trace represents the canonical impulse response function for humans based on the SPM HRF. The maximum occurs after 2.8 ± 0.8 s and the undershoot occurs at 6.1 ± 3.7 s, compared to 5 s and 15 s for the human impulse response The HRF is expected to vary across brain regions, and may even become faster for under awake/behaving preparations, this example demonstrates that hemodynamics may occur at different time scales in the experimental animal model most often used for investigating neuro-vascular coupling. (Reproduced from Lambers et al., 2020, with permission.)

Although in this review we focus primarily on the fast fMRI responses of the BOLD signal, there is increasing availability and interest in non-BOLD fMRI signals (Huber et al., 2019). What is less clear is how the temporal resolution varies between the fMRI contrasts. Several studies have reported distinct dynamics for various fMRI contrasts (de Zwart et al., 2018; Hirano et al., 2011; Hulvershorn et al., 2005a, 2005b; Mandeville et al., 1998; Silva et al., 2007), including differences between gradient-echo versus spin-echo BOLD (Hulvershorn et al., 2005a, 2005b), and differences in the measured non-BOLD responses as a function of stimulus duration for CBF (Gu et al., 2005; Kim et al., 2020; Leite and Mandeville, 2006; Su et al., 2004; Yang et al., 2000) and venous CBV (Kim and Kim, 2011a). Arterial CBV, which can be measured separately from venous CBV (Hua et al., 2018; Kim et al., 2008; Kim and Kim, 2011b, 2006), is expected to exhibit a different time course (Kim and Kim, 2011a). While either arterial CBV or CBF may be anticipated to have improved temporal specificity compared with BOLD, still the ultimate resolution of these techniques will ultimately depend both on the speed with which the vasculature responds to neuronal activity, as well as on the voxel size—since temporal resolution will be decreased due to the abovementioned temporal dispersion or "smearing" effects if multiple levels of the vascular hierarchy contribute to the fMRI signal.

Another extrinsic factor that may influence achievable temporal resolution of fMRI is the magnetic field strength. Not only does imaging performance and sensitivity increase with field strength, allowing increased imaging resolution, but fMRI signal transients have been shown to exhibit a weak dependency on field strength for short-duration stimuli and a stronger dependency of field strength for long-duration stimuli duration (Pfeuffer et al., 2003; Uludağ and Blinder, 2018). This suggests that, while most high-resolution fMRI studies are conducted at ultra-high field strengths (Polimeni and Uludağ, 2018), fast neural dynamics may be observable with fast fMRI techniques at conventional field strengths as well—provided sufficient imaging resolution and sensitivity can be achieved.

Finally, fast fMRI techniques may also be used to better localize brain activity in time, to investigate the temporal sequence of neural processing across brain regions and map large-scale brain circuitry. While prior studies have utilized the temporal precision of the BOLD response to attempt to map out the temporal progression of activation within specific brain networks (Bartels and Zeki, 2004; Katwal et al., 2013; Menon et al., 1998) and even infer the causality of these connections on the basis of fMRI response timing (Stephan and Roebroeck, 2012; Valdes-Sosa et al., 2011), the variability of the temporal properties of the HDR including its delay across brain regions reviewed above must be considered when interpreting measures based on observing the relative lag of fMRI response across brain regions (Miezin et al., 2000; Smith et al., 2012a). Analogous questions arise in the context of laminar fMRI, where, because certain cortical layers are well-known inputs or outputs along feedforward, feedback, or horizontal pathways in multi-area brain circuits, localizing fMRI activation within a specific cortical layer can provide information about which brain circuit is engaged (Norris and Polimeni, 2019). However, because within a given cortical area, neuronal populations across multiple cortical layers may be engaged during local neuronal processing, multiple layers within a cortical area will be activated by a given stimulus/task, therefore identifying the input cortical layer will not be possible by simply localizing activity alone. Several laminar fMRI studies therefore have attempted to identify the input layer on the basis of fast fMRI acquisitions (Hirano et al., 2011; Petridou and Siero, 2019; Silva and Koretsky, 2002; Yu et al., 2014) under the assumption that the input cortical layer within a cortical area should activate before the other cortical layers. This is a promising direction, although proper accounting of the systematic differences in the HDR across cortical layers will be required in this scenario as well.

8. Limitations in modeling the HDR and the importance of experimental design

While improvements in HDR modeling can help directly relate fMRI signals to the underlying neural activity, it is important to acknowledge the challenges in generating a fully accurate model given the many complexities underpinning the HDR. First, deriving a HDR model that is based on experimental observations requires some ability to separate out temporal features of the measured fMRI signal that are attributable to neural versus hemodynamic effects. Indeed the brain exhibits a broad repertoire of neural responses to the stimuli and tasks used in fMRI, and abstracting the neural response as simply following the stimulus or task timing is clearly an oversimplification (Gonzalez-Castillo et al., 2012; Havlíček et al., 2017b; Uludağ, 2008). Second, neurovascular coupling and the many pathways through which neurons communicate with the vasculature and trigger the initial hemodynamic changes are notoriously complex (Iadecola, 2017, 2004; Uhlirova et al., 2016), and it is possible that no one mechanism can describe this relationship for all brain regions, brain states, or neuronal sub-populations that are engaged by during any given stimulus or task. Finally, the hemodynamics themselves are complex—there are many open questions regarding the mechanisms of functional hyperemia (Attwell et al., 2016; Cai et al., 2018; Drew, 2019; Hartmann et al., 2021), plus the local hemodynamics are likely shaped by several factors including the local microvascular architecture suggesting that they will vary systematically across the brain (Schmid et al., 2019). Altogether, while impressive progress has been made in modeling the HDR, relying entirely on accurate HDR models to infer patterns of neural activity from fast fMRI data may not be viable. What is clear, however, is that the standard linear model based on the canonical HRF is not appropriate for fast fMRI data. Fast task designs in particular will not be well modeled by canonical models (Fig. 8A & B), leading to low sensitivity and missed activation when studying many common timescales of neural activity.

Given this limited understanding, there are other strategies for extracting meaningful information about neural dynamics from fast fMRI data. Approaches that focus on extracting mean responses that are consistently time-locked across trials have uncovered meaningful neural activity that did not conform to standard expectations (Gonzalez-Castillo et al., 2012). Similarly, periodic designs are amenable to Fourier-based analyses akin to the classic "phase-encoded" stimulus design (Bandettini et al., 1993; Engel et al., 1994; Sereno et al., 1995) and do not require explicit HDR models; this approach can be used to examine the frequency response of the measured fast fMRI data during steady state (Lewis et al., 2016) as well as the spectral content of fast fMRI signals across the course of an experiment (Lewis et al., 2018a). Rapid periodic stimulus designs also have the potential advantage saturating some components of the HDR and may yield improved neural specificity (Fukuda et al., 2021; Moon et al., 2007). Lastly one can consider designs with carefully crafted controls that account for hemodynamic effects. For example, by comparing temporal features of the observed fMRI response to multiple stimuli with constant stimulus intensity and duration while varying other features hypothesized to generate different neural dynamics, it is possible to compare the changes in the HDR across stimuli (Bellgowan et al., 2003). Thus while continued development of HDR models improve our ability to estimate of neural activity from fMRI measurements, today, through careful experiment design, it is still possible to exploit the intrinsically high temporal specificity of the HDR to extract rich temporal dynamics of the underlying neural activity.

9. Neuroscience advances within reach of fast fMRI

Since many cognitive processes involve neural sequences unfolding over hundreds of milliseconds, corresponding to the temporal resolution now enabled with fast fMRI, fast fMRI studies that take into account these faster dynamics of the HDR could be used to test neuroscience questions in this temporal range. A large number of studies have reported fast BOLD dynamics in task-driven and in resting-state studies (Cao et al., 2019; Chen and Glover, 2015; Lee et al., 2013; Lewis et al., 2018b, 2016; Lin et al., 2015, 2013; Sasai et al., 2021), suggesting that these updated HDR models could both provide more accurate analysis of signals in naturalistic or resting-state paradigms, as well as enabling a new class of neuroscientific studies aiming to image subsecond neural dynamics directly with fMRI.

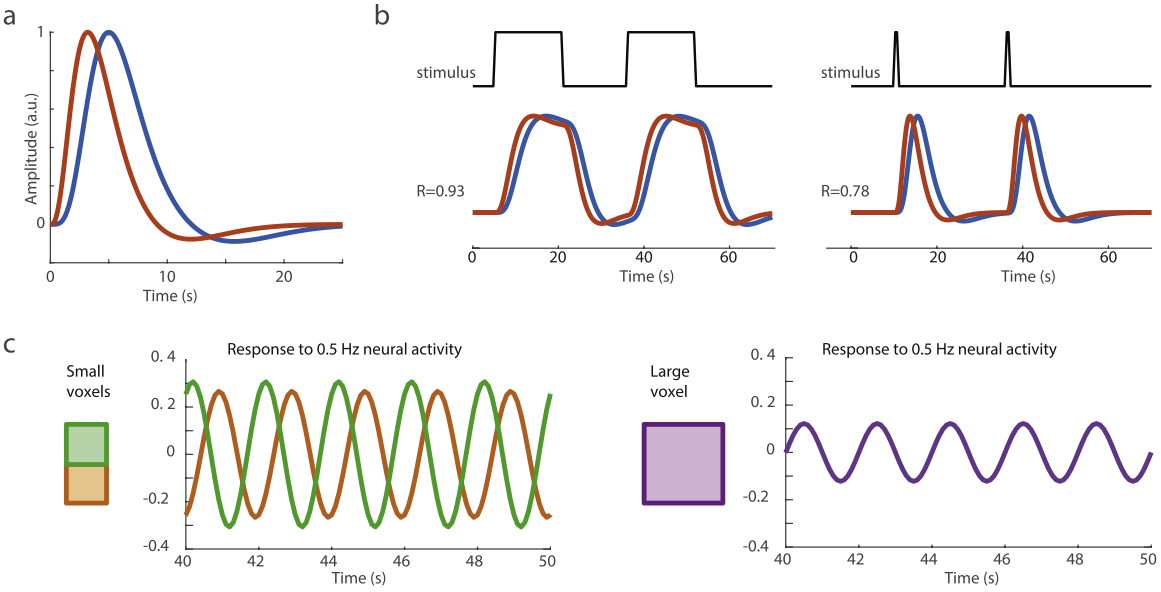

Fig. 8. Accurate hemodynamic response models become increasingly important when studying fast neural activity. (**A**) Example of a fast HRF vs. the canonical HRF. (**B**) Schematic of how a fast voxel would respond to a block vs. a rapid event. When the stimulus design is slow, a mismatched HRF has little effect; however, when the stimulus is fast, a mismatch between the HRF and canonical model leads to lower accuracy. This can cause lower sensitivity to activation when using conventional models to analyze rapid task designs, or in regions with fast hemodynamics such as subcortical structures or deep cortical gray matter. (**C**) Schematic of temporal phase cancellation during a 0.5-Hz stimulus. When analyzing rapidly fluctuating activity, the typical biological heterogeneity of HDRs (e.g., 1-s phase delay) across voxels can average out the responses, leading to smaller signals.

An immediate application area is cognitive neuroscience studies seeking to disambiguate rapid sequences of neural events. Lin et al. (2013) demonstrated that events as little as 100 ms apart can be distinguished using fast fMRI, by presenting motor and visual stimuli with controlled onset timing. Many cognitive tasks involve timing in this range—notably, language, abstract reasoning, attentional selection, and many other tasks we aim to study in the human brain involve sequences on this timescale. Previous work at slower timescales has already used fMRI to investigate neural sequences and transient events engaged in cognitive processing (Dux et al., 2006; Stigliani et al., 2017). Recent work has demonstrated how the BOLD dynamics of rapidly-presented short-duration stimuli could be used to test models of temporal summation and adaptation in the context of the visual system (Zhou et al., 2018); extending this model to other areas, and including modeling of how HRF timing changes with stimulus duration, could further enhance our understanding of temporal patterning of neural dynamics. A recent study with faster sampling has now demonstrated the use of fast fMRI to disentangle responses to individual words within a narrative (Rocca et al., 2020), and other recent studies have applied other rapid investigations to visual sequence detection (Wittkuhn and Schuck, 2020) and auditory dynamics (Frühholz et al., 2020). An important consideration in these types of rapid neural sequence detection studies will be consideration of how HDR models may differ across regions, complicating interpretation of different response time dynamics in different brain areas (Miezin et al., 2000). While keeping this issue in mind, approaches using within-region analyses and faster models of the HDR should enable a wide range of studies using subsecond task dynamics.

Spontaneous aspects of neural activity may also induce fast hemodynamic responses and contain higher frequency dynamics than are analyzed in conventional resting-state studies, which primarily focus on < 0.1 Hz bands. The low-frequency BOLD dynamics observed at conventional temporal sampling rates have been associated with infraslow spontaneous fluctuations in cortical potential (He et al., 2008; He and Raichle, 2009; Hiltunen et al., 2014; Pan et al., 2013), as well as low-frequency modulations of higher-frequency neural dynamics (Mateo et al., 2017; Schölvinck et al., 2010). As the attainable frequency range of fMRI increases, qualitatively distinct aspects of spontaneous neural activity should be in range of fast fMRI. A range of particular interest is the human slow oscillation band, peaking around ~0.7 Hz (Achermann and Borbély, 1997; Steriade et al., 1993). Slow oscillations reflect coherent dynamics across much of the cortex (Destexhe and Contreras, 2011; Massimini et al., 2004), and are implicated in sleep, memory, and awareness (Batterink et al., 2016; Mölle et al., 2004; Ngo et al., 2013). The rapid HDR observed in response to oscillatory neural activity suggests that this frequency range is now in reach of fast fMRI (Lewis et al., 2016). Another detectable aspect will be alterations in brain state occurring in the 0–1 Hz frequency band, such as amplitude envelope modulations of higher-frequency rhythms (Mateo et al., 2017; Schölvinck et al., 2010) or sequences of neural microstates (Van De Ville et al., 2010; Vidaurre et al., 2016; Yuan et al., 2012). Finally, endogenous fluctuations in physiology and autonomic state, while posing challenges for fMRI due to their associated effects on the BOLD signal (Birn et al., 2006; Chang et al., 2009; Murphy et al., 2013), may also be detectable with fast fMRI (Manuel et al., 2020). Each of these spontaneous neural dynamics falls into the subsecond range in which the HDR is expected to be faster, and faster HDR models may be needed in order to capture the > 0.1 Hz dynamics produced by these signals.

In addition, the advent of naturalistic tasks now poses both an opportunity for and an argument in favor of fast fMRI. Naturalistic stimuli fall squarely into the temporal regime where the conventional, slow and linear HDR is no longer expected to hold, as they typically do not involve slowly alternating high-intensity blocks. fMRI during naturalistic movies or auditory stimuli is increasingly being used to study high-level visual and semantic structures (Finn et al., 2020; Hasson et al., 2004; Huth et al., 2016; Kay et al., 2008b; Nishimoto et al., 2011; Simony et al., 2016) and early investigations suggest that high-frequency content may also be present in these responses (Gao, 2015). Accurate modeling of the HDR will be needed to capture how fMRI signals fluctuate in concert with these relatively broadband neural signals.

Fast fMRI also offers the potential of studying dynamics in cortical mesoscale structures at high spatial resolution. Laminar fMRI is being increasingly used to assess cognitive processes across deep and superficial cortical layers (Finn et al., 2019; Huber et al., 2017; Lawrence et al., 2019). The faster hemodynamic responses of deep cortical layers make them an ideal region for imaging fast fMRI signals, suggesting that subsecond dynamics could potentially be best studied within the cortical gray matter parenchyma, as these signals are attenuated in the larger surface veins.

Finally, fast fMRI with appropriate HDR models will be particularly well suited to studies seeking to image the subcortex, as it is the only noninvasive tool with millimeter resolution in deep brain. Diverse nuclei in structures such as the brainstem, thalamus, and basal ganglia play critical roles in many fundamental aspects of brain function, and their dynamics can be simultaneously imaged with fMRI. Furthermore, while sensitivity can be lower due to the distance from the RF receive coils, the observation that hemodynamics are even faster in subcortex will enhance the ability of fast fMRI to detect rapid signals in these regions. For example, a recent study demonstrated >0.1 Hz signals in human brainstem, which could not have been identified through other techniques (Manuel et al., 2020). Fast fMRI in brainstem, thalamus, and other subcortical structures may thus be able to uncover rapid activity sequences and higher-frequency oscillations due to the temporal characteristics of responses in these regions, providing a window into how subcortical systems regulate sleep, arousal, pain, mood, autonomic function (Sclocco et al., 2018), as well as the complex cognitive processes implemented in thalamic circuits (Saalmann and Kastner, 2015), in the human brain. Furthermore, using conventional HDR models in these regions could lead to missing their activity patterns even when imaging at conventional temporal resolution (Fig. 8b). The substantial decrease in time-to-peak of the brainstem signal (Lewis et al., 2018b) is expected to lead to poor model fits with a canonical HRF, which could prevent detection of the rapid responses in these structures.

As this example illustrates, a critical limitation for all fast fMRI studies—and indeed, for conventional fMRI studies as well, in which the problem may simply be harder to detect—is the lack of established HDR models that can generalize across brain regions and tasks. In many cases, inaccurate models could simply lead to missing detection of activity that is in fact present, as the incorrect HDR leads to poor model fit. While fitting complex VAN models to fast fMRI data is unlikely to be practical, using HRFs derived from the extensive work that has been done on nonlinear and heterogeneous HRF properties may enhance the sensitivity of these studies, particularly in subcortex and in deep cortical gray matter, where hemodynamics are fastest. Alternatively, in the absence of pre-identified HRFs, the use of FIR models (Gonzalez-Castillo et al., 2012) can allow discovery of local activity timecourses with fewer assumptions about the HDR shape, although these approaches do not resolve whether temporal differences are expected to be due to hemodynamic differences. Interpreting the neural dynamics underlying fast fMRI signals will thus benefit from incorporating information on how hemodynamics vary with stimulus type and brain region, informing conclusions about which temporal differences are neural, vs. hemodynamic, in origin.

In addition, a challenge for all these applications will be the low amplitude of fast fMRI signals, which are extremely small at high frequencies. With available technology, sensitivity likely poses the greatest limit to the use of fast fMRI. While high-frequency signals are clearly detected in structures that can be strongly driven by sensory signals, such as the visual pathway (Lewis et al., 2018b), more complex, high-level cognitive tasks already frequently elicit relatively low-amplitude fMRI responses that will be further attenuated at faster timescales. Fast fMRI may thus be best applied initially to cognitive questions that are expected to elicit large effect sizes, such as sensory

tasks, attentional dynamics, and arousal state modulations.

Two approaches may help increase the efficacy of fast fMRI analyses, expanding its reach beyond these initial studies. First, increasing spatial resolution will also greatly enhance detection of fast fMRI signals, as the spatial heterogeneity of HDR timing leads to signal cancellation when using large voxels (Fig. 8C) that can be mitigated by decreasing voxel size (Chen et al., 2019b), avoiding spatial smoothing, and restricting analyses to the cortical gray matter parenchyma (Polimeni et al., 2018). fMRI at 7 Tesla can help with this, as it increases CNR as well as providing additional spatial resolution with which to reduce phase cancellation effects, allowing enhancement of the contributions from the gray matter parenchyma. Second, analysis methods and statistical models that account for the distinct statistical properties of fast fMRI data (Agrawal et al., 2020; Bollmann et al., 2018; Chen et al., 2019a; Corbin et al., 2018) and carefully account for physiological noise, will enhance detection and estimation of these signals. Both of these improvements could be used in conjunction with updated HDR models that take into account the faster hemodynamics in response to transient or fluctuating neural activity.

Together, these studies suggest that fast fMRI will be well suited to examining fast HDRs in studies with brief, weak, or oscillatory patterns of neural activity, which correspond to the naturalistic dynamics of many aspects of both task-driven and spontaneous human cognition. Furthermore, it provides a unique tool with which to image subsecond dynamics in mesoscale structures such as brainstem nuclei, opening the door to new questions in the neuroscience of human subcortical circuits.

10. Conclusions and outlook

In this review we have presented evidence for the ability of modern fast fMRI methods to probe surprisingly fast neuronal dynamics. We have surveyed recent experimental data on the microvascular responses to neural activity which demonstrate that the changes in both the vasculature and the blood occurring alongside neuronal activity are localized in both space and time; fundamentally, the “biological” spatial and temporal resolution of hemodynamics-based measures of brain function such as fMRI appear to be finer than our imaging resolution today, motivating new advances in fMRI technology to better exploit this untapped potential. We summarized recent experimental data suggesting that the observed temporal resolution of the BOLD response is far beyond what was predicted by the canonical HRF, and posited that this is in part because the linear system assumption of the BOLD response does not hold for brief or weak stimuli. We surveyed the vast literature on the nonlinearity of the BOLD response, from which a picture of an HDR that is dependent on many features of the stimulus/task emerges. We then summarized a wide range of BOLD models and noted that while there has been recent progress there is still an unmet need for models that can generalize across stimulus and task configurations to account for these nonlinearities and thereby be capable of predicting and interpreting the observed fast fMRI signal. We then presented several examples where the temporal resolution of fast fMRI is approaching the temporal scales needed to track cognitive processes that may be inaccessible to conventional fMRI techniques. Measuring these dynamics will require updated analysis approaches tailored to these newly available paradigms.

In conclusion, fast fMRI offers the potential to study dynamics at timescales and contexts that are important for many aspects of human cognition: sequences over hundreds of milliseconds, involving brief, low-intensity, or rapidly fluctuating neural activity. While these signals are small, the increasing sensitivity of fMRI suggests that they are nevertheless now within reach. Using appropriate, updated models of the HDR will be essential for enhancing the detectability of these signals, ultimately enabling imaging of neural dynamics on timescales of hundreds of milliseconds.

Acknowledgements

We gratefully acknowledge useful discussions with Peter Bandettini, Avery Berman, Saskia Bollmann, Jingyuan Chen, Julien Cohen-Adad, Daniel Gomez, Alex Huth, Yoichi Miyawaki, Robert Savoy, Luca Vizioli, and Robert Weisskoff. We also in particular thank Kâmil Uludağ for helpful explanations of the nonlinearities in the hemodynamic response and their interpretation. We also thank Martin Havlíček for making his *Laminar BOLD Response* software publicly available. We finally thank the anonymous reviewers for their insightful comments and helpful suggestions. This work was supported in part by the NIH NIBIB (grants P41-EB015896, P41-EB030006 and R01-EB019437) NINDS (grant R21-NS106706), NIMH (grant R00-MH111748), NIA (grant R01-AG070135), by the *BRAIN Initiative* (NIH NIMH grants R01-MH111438 and R01-MH111419 and NIBIB grant U01-EB025162), the Searle Scholars Program, and by the MGH/HST Athinoula A. Martinos Center for Biomedical Imaging. Material from Moonen, C.T.W. and Bandettini, P.A., Functional MRI, published 1999 by Springer-Verlag, Berlin, reproduced with permission of SNCSC.

Appendix A. The Peer Review Overview and Supplementary data

The Peer Review Overview and Supplementary data associated with this article can be found in the online version, at doi:https://doi.org/10.1016/j.pneurobio.2021.102174.

References

Achermann, P., Borbély, A.A., 1997. Low-frequency (< 1 Hz) oscillations in the human sleep electroencephalogram. Neuroscience 81, 213–222. https://doi.org/10.1016/S0306-4522(97)00186-3.

Agrawal, U., Brown, E.N., Lewis, L.D., 2020. Model-based physiological noise removal in fast fMRI. Neuroimage 205. https://doi.org/10.1016/j.neuroimage.2019.116231.

Aguirre, G.K., Zarahn, E., D'Esposito, M., 1998. The variability of human, BOLD hemodynamic responses. Neuroimage 8, 360–369. https://doi.org/10.1006/nimg.1998.0369.

Arcaro, M.J., Pinsk, M.A., Kastner, S., 2015. The anatomical and functional organization of the human visual pulvinar. J. Neurosci. 35, 9848–9871. https://doi.org/10.1523/JNEUROSCI.1575-14.2015.

Ashburner, J., 2012. SPM: a history. Neuroimage 62, 791–800. https://doi.org/10.1016/j.neuroimage.2011.10.025.

Attwell, D., Iadecola, C., 2002. The neural basis of functional brain imaging signals. Trends Neurosci. 25, 621–625.

Attwell, D., Mishra, A., Hall, C.N., O'Farrell, F.M., Dalkara, T., 2016. What is a pericyte? J. Cereb. Blood Flow Metab. 36, 451–455. https://doi.org/10.1177/0271678X15610340.

Báez-Yánez, M.G., Ehses, P., Mirkes, C., Tsai, P.S., Kleinfeld, D., Scheffler, K., 2017. The impact of vessel size, orientation and intravascular contribution on the neurovascular fingerprint of BOLD bSSFP fMRI. Neuroimage 163, 13–23. https://doi.org/10.1016/j.neuroimage.2017.09.015.

Bandettini, P.A., 1999. The temporal resolution of functional MRI. In: Moonen, C.T.W., Bandettini, P.A. (Eds.), Functional MRI. Springer-Verlag, Berlin, pp. 205–220.

Bandettini, P.A., 2002. The spatial, temporal, and interpretive limits of functional MRI. Neuropsychopharmacology: The Fifth Generation of Progress. Lippincott Williams & Wilkins, pp. 343–356.

Bandettini, P.A., Cox, R.W., 2000. Event-related fMRI contrast when using constant interstimulus interval: theory and experiment. Magn. Reson. Med. 43, 540–548.

Bandettini, P.A., Wong, E.C., 1995. Effects of biophysical and physiologic parameters on brain activation-induced R2* and R2 changes: simulations using a deterministic diffusion model. Int. J. Imaging Syst. Technol. 6, 133–152. https://doi.org/10.1002/ima.1850060203.

Bandettini, P., Wong, E.C., 2015. The future of functional MRI. FMRI: From Nuclear Spins to Brain Functions. Springer US, pp. 895–929. https://doi.org/10.1007/978-1-4899-7591-1_30.

Bandettini, P.A., Jesmanowicz, A., Wong, E.C., Hyde, J.S., 1993. Processing strategies for time-course data sets in functional MRI of the human brain. Magn. Reson. Med. 30, 161–173.

Bandettini, P.A., Kwong, K., Davis, T., Tootell, R.B.H., Wong, E.C., Fox, P., Belliveau, J. W., Weisskoff, R., Rosen, B.R., 1997. Characterization of cerebral blood oxygenation and flow changes during prolonged brain activation. Hum. Brain Mapp. 5, 93–109.

Bartels, A., Zeki, S., 2004. The chronoarchitecture of the human brain–natural viewing conditions reveal a time-based anatomy of the brain. Neuroimage 22, 419–433. https://doi.org/10.1016/j.neuroimage.2004.01.007.

Barth, M., Breuer, F.A., Koopmans, P.J., Norris, D.G., Poser, B.A., 2015. Simultaneous multislice (SMS) imaging techniques. Magn. Reson. Med. 75, 63–81. https://doi.org/10.1002/mrm.25897.

Batterink, L.J., Creery, J.D., Paller, K.A., 2016. Phase of spontaneous slow oscillations during sleep influences memory-related processing of auditory cues. J. Neurosci. 36, 1401–1409. https://doi.org/10.1523/JNEUROSCI.3175-15.2016.

Bause, J., Polimeni, J.R., Stelzer, J., In, M.-H., Ehses, P., Kraemer-Fernandez, P., Aghaeifar, A., Lacosse, E., Pohmann, R., Scheffler, K., 2020. Impact of prospective motion correction, distortion correction methods and large vein bias on the spatial accuracy of cortical laminar fMRI at 9.4 Tesla. Neuroimage 208, 116434. https://doi.org/10.1016/j.neuroimage.2019.116434.

Behzadi, Y., Liu, T.T., 2005. An arteriolar compliance model of the cerebral blood flow response to neural stimulus. Neuroimage 25, 1100–1111. https://doi.org/10.1016/j.neuroimage.2004.12.057.

Bell, M.A., Ball, M.J., 1985. Laminar variation in the microvascular architecture of normal human visual cortex (area 17). Brain Res. 335, 139–143.

Bellgowan, P.S.F., Saad, Z.S., Bandettini, P.A., 2003. Understanding neural system dynamics through task modulation and measurement of functional MRI amplitude, latency, and width. Proc. Natl. Acad. Sci. U. S. A. 100, 1415–1419. https://doi.org/10.1073/pnas.0337747100.

Berman, A.J.L., Mazerolle, E.L., MacDonald, M.E., Blockley, N.P., Luh, W.M., Pike, G.B., 2018. Gas-free calibrated fMRI with a correction for vessel-size sensitivity. Neuroimage 169, 176–188. https://doi.org/10.1016/j.neuroimage.2017.12.047.

Bernier, M., Cunnane, S.C., Whittingstall, K., 2018. The morphology of the human cerebrovascular system. Hum. Brain Mapp. 39, 4962–4975. https://doi.org/10.1002/hbm.24337.

Berwick, J., Johnston, D., Jones, M., Martindale, J., Redgrave, P., McLoughlin, N., Schiessl, I., Mayhew, J.E.W., 2005. Neurovascular coupling investigated with two-dimensional optical imaging spectroscopy in rat whisker barrel cortex. Eur. J. Neurosci. 22, 1655–1666. https://doi.org/10.1111/j.1460-9568.2005.04347.x.

Birn, R.M., Bandettini, P.A., 2005. The effect of stimulus duty cycle and "off" duration on BOLD response linearity. Neuroimage 27, 70–82. https://doi.org/10.1016/j.neuroimage.2005.03.040.

Birn, R.M., Saad, Z.S., Bandettini, P.A., 2001. Spatial heterogeneity of the nonlinear dynamics in the FMRI BOLD response. Neuroimage 14, 817–826. https://doi.org/10.1006/nimg.2001.0873.

Birn, R.M., Diamond, J.B., Smith, M.A., Bandettini, P.A., 2006. Separating respiratory-variation-related fluctuations from neuronal-activity-related fluctuations in fMRI. Neuroimage 31, 1536–1548. https://doi.org/10.1016/j.neuroimage.2006.02.048.

Błażejewska, A.I., Bernier, M., Nasr, S., Polimeni, J.R., 2019. Testing temporal dependence of spatial specificity in BOLD fMRI at 7T: comparing short versus long stimulus duration. Proc. Intl. Soc. Mag. Reson. Med. 27, 0536.

Blinder, P., Tsai, P.S., Kaufhold, J.P., Knutsen, P.M., Suhl, H., Kleinfeld, D., 2013. The cortical angiome: an interconnected vascular network with noncolumnar patterns of blood flow. Nat. Neurosci. 16, 889–897. https://doi.org/10.1038/nn.3426.

Boas, D.A., Strangman, G., Culver, J.P., Hoge, R.D., Jasdzewski, G., Poldrack, R.A., Rosen, B.R., Mandeville, J.B., 2003. Can the cerebral metabolic rate of oxygen be estimated with near-infrared spectroscopy? Phys. Med. Biol. 48, 2405–2418. https://doi.org/10.1088/0031-9155/48/15/311.

Boas, D.A., Jones, S.R., Devor, A., Huppert, T.J., Dale, A.M., 2008. A vascular anatomical network model of the spatio-temporal response to brain activation. Neuroimage 40, 1116–1129. https://doi.org/10.1016/j.neuroimage.2007.12.061.

Boillat, Y., van der Zwaag, W., 2019. Whole brain measurements of the positive BOLD response variability during a finger tapping task at 7 T show regional differences in its profiles. Magn. Reson. Med. 81, 2720–2727. https://doi.org/10.1002/mrm.27566.

Bollmann, S., Puckett, A.M., Cunnington, R., Barth, M., 2018. Serial correlations in single-subject fMRI with sub-second TR. Neuroimage 166, 152–166. https://doi.org/10.1016/j.neuroimage.2017.10.043.

Bollmann, S., Bernier, M., Robinson, S., Polimeni, J.R., 2020. Geometrically accurate imaging of the pial arterial vasculature of the human brain in vivo using high-resolution non-contrast angiography at 7T. Proc. Intl. Soc. Mag. Reson. Med. 28, 1105.

Boxerman, J.L., Bandettini, P.A., Kwong, K.K., Baker, J.R., Davis, T.L., Rosen, B.R., Weisskoff, R.M., 1995. The intravascular contribution to fMRI signal change: Monte Carlo modeling and diffusion-weighted studies in vivo. Magn. Reson. Med. 34, 4–10.

Boynton, G.M., Engel, S.A., Glover, G.H., Heeger, D.J., 1996. Linear systems analysis of functional magnetic resonance imaging in human V1. J. Neurosci. 16, 4207–4221. https://doi.org/10.1523/jneurosci.16-13-04207.1996.

Boynton, G.M., Engel, S.A., Heeger, D.J., 2012. Linear systems analysis of the fMRI signal. Neuroimage. https://doi.org/10.1016/j.neuroimage.2012.01.082.

Buckner, R.L., Bandettini, P.A., O'Craven, K.M., Savoy, R.L., Petersen, S.E., Raichle, M.E., Rosen, B.R., 1996. Detection of cortical activation during averaged single trials of a cognitive task using functional magnetic resonance imaging. Proc. Natl. Acad. Sci. U. S. A. 93, 14878–14883. https://doi.org/10.1073/pnas.93.25.14878.

Budinger, T.F., Bird, M.D., 2018. MRI and MRS of the human brain at magnetic fields of 14 T to 20 T: technical feasibility, safety, and neuroscience horizons. Neuroimage 168, 509–531. https://doi.org/10.1016/j.neuroimage.2017.01.067.

Buxton, R.B., 2001. The elusive initial dip. Neuroimage 13, 953–958. https://doi.org/10.1006/nimg.2001.0814.

Buxton, R., 2009. Introduction to Functional Magnetic Resonance Imaging. Principles and Techniques, 2nd ed. Cambridge University Press, Cambridge, UK. https://doi.org/10.1034/j.1600-0455.2002.430519_5.x.

Buxton, R.B., 2010. Interpreting oxygenation-based neuroimaging signals: the importance and the challenge of understanding brain oxygen metabolism. Front. Neuroenergetics 2, 8. https://doi.org/10.3389/fnene.2010.00008.

Buxton, R.B., 2012. Dynamic models of BOLD contrast. Neuroimage 62, 953–961. https://doi.org/10.1016/j.neuroimage.2012.01.012.

Buxton, R.B., Frank, L.R., 1997. A model for the coupling between cerebral blood flow and oxygen metabolism during neural stimulation. J. Cereb. Blood Flow Metab. 17, 64–72. https://doi.org/10.1097/00004647-199701000-00009.

Buxton, R.B., Wong, E.C., Frank, L.R., 1998. Dynamics of blood flow and oxygenation changes during brain activation: the balloon model. Magn. Reson. Med. 39, 855–864.

Buxton, R.B., Uludağ, K., Dubowitz, D.J., Liu, T.T., 2004. Modeling the hemodynamic response to brain activation. Neuroimage 23 (Suppl 1), S220–33. https://doi.org/10.1016/j.neuroimage.2004.07.013.

Buxton, R.B., Griffeth, V.E.M., Simon, A.B., Moradi, F., Shmuel, A., 2014. Variability of the coupling of blood flow and oxygen metabolism responses in the brain: a problem for interpreting BOLD studies but potentially a new window on the underlying neural activity. Front. Neurosci. 8, 139. https://doi.org/10.3389/fnins.2014.00139.

Cai, C., Fordsmann, J.C., Jensen, S.H., Gesslein, B., Lønstrup, M., Hald, B.O., Zambach, S. A., Brodin, B., Lauritzen, M.J., 2018. Stimulation-induced increases in cerebral blood flow and local capillary vasoconstriction depend on conducted vascular responses. Proc. Natl. Acad. Sci. U. S. A. 115, E5796–E5804. https://doi.org/10.1073/pnas.1707702115.

Cao, J., Lu, K.H., Oleson, S.T., Phillips, R.J., Jaffey, D., Hendren, C.L., Powley, T.L., Liu, Z., 2019. Gastric stimulation drives fast BOLD responses of neural origin. Neuroimage 197, 200–211. https://doi.org/10.1016/j.neuroimage.2019.04.064.

Chang, C., Thomason, M.E., Glover, G.H., 2008. Mapping and correction of vascular hemodynamic latency in the BOLD signal. Neuroimage 43, 90–102. https://doi.org/10.1016/j.neuroimage.2008.06.030.

Chang, C., Cunningham, J.P., Glover, G.H., 2009. Influence of heart rate on the BOLD signal: the cardiac response function. Neuroimage 44, 857–869. https://doi.org/10.1016/j.neuroimage.2008.09.029.

Chen, S.H.A., Desmond, J.E., 2005. Temporal dynamics of cerebro-cerebellar network recruitment during a cognitive task. Neuropsychologia 43, 1227–1237. https://doi.org/10.1016/j.neuropsychologia.2004.12.015.

Chen, J.E., Glover, G.H., 2015. BOLD fractional contribution to resting-state functional connectivity above 0.1 Hz. Neuroimage 107, 207–218. https://doi.org/10.1016/j.neuroimage.2014.12.012.

Chen, J.J., Pike, G.B., 2009. BOLD-specific cerebral blood volume and blood flow changes during neuronal activation in humans. NMR Biomed. 22, 1054–1062. https://doi.org/10.1002/nbm.1411.

Chen, B.R., Bouchard, M.B., McCaslin, A.F.H., Burgess, S.A., Hillman, E.M.C., 2011. High-speed vascular dynamics of the hemodynamic response. Neuroimage 54, 1021–1030. https://doi.org/10.1016/j.neuroimage.2010.09.036.

Chen, B.R., Kozberg, M.G., Bouchard, M.B., Shaik, M.A., Hillman, E.M.C., 2014. A critical role for the vascular endothelium in functional neurovascular coupling in the brain. J. Am. Heart Assoc. 3, e000787 https://doi.org/10.1161/JAHA.114.000787.

Chen, J.E., Polimeni, J.R., Bollmann, S., Glover, G.H., 2019a. On the analysis of rapidly sampled fMRI data. Neuroimage 188, 807–820. https://doi.org/10.1016/j.neuroimage.2019.02.008.

Chen, J.E., Polimeni, J.R., Fultz, N.E., Glover, G., Lewis, L.D., 2019b. Fast fMRI responses supported by inter-voxel variability of hemodynamic response functions. Proc. Intl. Soc. Mag. Reson. Med. 27, 3726.

Chen, J.E., Fultz, N.E., Glover, G., Rosen, B.R., Polimeni, J.R., Lewis, L.D., 2020a. Probing the neuronal and vascular origins of task contrast-dependent hemodynamic response functions. Proc. Intl. Soc. Mag. Reson. Med. 28, 1107.

Chen, J.E., Lewis, L.D., Chang, C., Tian, Q., Fultz, N.E., Ohringer, N.A., Rosen, B.R., Polimeni, J.R., 2020b. Resting-state "physiological networks". Neuroimage 213, 116707. https://doi.org/10.1016/j.neuroimage.2020.116707.

Cheng, X., Berman, A.J.L., Polimeni, J.R., Buxton, R.B., Gagnon, L., Devor, A., Sakadžić, S., Boas, D.A., 2019. Dependence of the MR signal on the magnetic susceptibility of blood studied with models based on real microvascular networks. Magn. Reson. Med. 81, 3865–3874. https://doi.org/10.1002/mrm.27660.

Chiew, M., Graedel, N.N., Miller, K.L., 2018. Recovering task fMRI signals from highly under-sampled data with low-rank and temporal subspace constraints. Neuroimage 174, 97–110. https://doi.org/10.1016/j.neuroimage.2018.02.062.

Cohen, E.R., Uğurbil, K., Kim, S.-G., 2002. Effect of basal conditions on the magnitude and dynamics of the blood oxygenation level-dependent fMRI response. J. Cereb. Blood Flow Metab. 22, 1042–1053. https://doi.org/10.1097/00004647-200209000-00002.

Corbin, N., Todd, N., Friston, K.J., Callaghan, M.F., 2018. Accurate modeling of temporal correlations in rapidly sampled fMRI time series. Hum. Brain Mapp. 39, 3884–3897. https://doi.org/10.1002/hbm.24218.

D'Esposito, M., Zarahn, E., Aguirre, G.K., Rypma, B., 1999. The effect of normal aging on the coupling of neural activity to the bold hemodynamic response. Neuroimage 10, 6–14. https://doi.org/10.1006/nimg.1999.0444.

D'Esposito, M., Deouell, L.Y., Gazzaley, A., 2003. Alterations in the BOLD fMRI signal with ageing and disease: a challenge for neuroimaging. Nat. Rev. Neurosci. 4, 863–872. https://doi.org/10.1038/nrn1246.

Dale, A.M., 1999. Optimal experimental design for event-related fMRI. Hum. Brain Mapp. 8, 109–114.

Dale, A.M., Buckner, R.L., 1997. Selective averaging of rapidly presented individual trials using fMRI. Hum. Brain Mapp. 5, 329–340. https://doi.org/10.1002/(SICI)1097-0193(1997)5:5<329::AID-HBM1>3.0.CO;2-5.

Davis, T.L., Kwong, K.K., Weisskoff, R.M., Rosen, B.R., 1998. Calibrated functional MRI: mapping the dynamics of oxidative metabolism. Proc. Natl. Acad. Sci. U. S. A. 95, 1834–1839.

De Martino, F., Moerel, M., Van de Moortele, P.-F., Uğurbil, K., Goebel, R., Yacoub, E., Formisano, E., 2013. Spatial organization of frequency preference and selectivity in the human inferior colliculus. Nat. Commun. 4, 1386. https://doi.org/10.1038/ncomms2379.

De Martino, F., Moerel, M., Uğurbil, K., Goebel, R., Yacoub, E., Formisano, E., 2015. Frequency preference and attention effects across cortical depths in the human primary auditory cortex. Proc. Natl. Acad. Sci. U. S. A. 112, 16036–16041. https://doi.org/10.1073/pnas.1507552112.

de Munck, J.C., Gonçalves, S.I., Huijboom, L., Kuijer, J.P.A., Pouwels, P.J.W., Heethaar, R.M., Lopes da Silva, F.H., 2007. The hemodynamic response of the alpha rhythm: an EEG/fMRI study. Neuroimage 35, 1142–1151. https://doi.org/10.1016/j.neuroimage.2007.01.022.

de Zwart, J.A., Silva, A.C., van Gelderen, P., Kellman, P., Fukunaga, M., Chu, R., Koretsky, A.P., Frank, J.A., Duyn, J.H., 2005. Temporal dynamics of the BOLD fMRI impulse response. Neuroimage 24, 667–677. https://doi.org/10.1016/j.neuroimage.2004.09.013.

de Zwart, J.A., van Gelderen, P., Jansma, J.M., Fukunaga, M., Bianciardi, M., Duyn, J.H., 2009. Hemodynamic nonlinearities affect BOLD fMRI response timing and amplitude. Neuroimage 47, 1649–1658. https://doi.org/10.1016/j.neuroimage.2009.06.001.

de Zwart, J.A., van Gelderen, P., Schindler, M.K., Sati, P., Liu, J., Reich, D.S., Duyn, J.H., 2018. Impulse response timing differences in BOLD and CBV weighted fMRI. Neuroimage 181, 292–300. https://doi.org/10.1016/J.NEUROIMAGE.2018.07.011.

Deneux, T., Faugeras, O., 2006. Using nonlinear models in fMRI data analysis: model selection and activation detection. Neuroimage 32, 1669–1689. https://doi.org/10.1016/j.neuroimage.2006.03.006.

Denison, R.N., Vu, A.T., Yacoub, E., Feinberg, D.A., Silver, M.A., 2014. Functional mapping of the magnocellular and parvocellular subdivisions of human LGN. Neuroimage 102 (Pt 2), 358–369. https://doi.org/10.1016/j.neuroimage.2014.07.019.

Destexhe, A., Contreras, D., 2011. The fine structure of slow-wave sleep oscillations: from single neurons to large networks. Sleep and Anesthesia. Springer, New York, pp. 69–105. https://doi.org/10.1007/978-1-4614-0173-5_4.

Devonshire, I.M., Papadakis, N.G., Port, M., Berwick, J., Kennerley, A.J., Mayhew, J.E. W., Overton, P.G., 2012. Neurovascular coupling is brain region-dependent. Neuroimage 59, 1997–2006. https://doi.org/10.1016/j.neuroimage.2011.09.050.

Devor, A., Dunn, A.K., Andermann, M.L., Ulbert, I., Boas, D.A., Dale, A.M., 2003. Coupling of total hemoglobin concentration, oxygenation, and neural activity in rat somatosensory cortex. Neuron 39, 353–359.

Devor, A., Tian, P., Nishimura, N., Teng, I.C., Hillman, E.M.C., Narayanan, S.N., Ulbert, I., Boas, D.A., Kleinfeld, D., Dale, A.M., 2007. Suppressed neuronal activity and concurrent arteriolar vasoconstriction may explain negative blood oxygenation level-dependent signal. J. Neurosci. 27, 4452–4459. https://doi.org/10.1523/JNEUROSCI.0134-07.2007.

Drew, P.J., 2019. Vascular and neural basis of the BOLD signal. Curr. Opin. Neurobiol. 58, 61–69. https://doi.org/10.1016/j.conb.2019.06.004.

Drew, P.J., Shih, A.Y., Kleinfeld, D., 2011. Fluctuating and sensory-induced vasodynamics in rodent cortex extend arteriole capacity. Proc. Natl. Acad. Sci. U. S. A. 108, 8473–8478. https://doi.org/10.1073/pnas.1100428108.

Duong, T.Q., Kim, D.S., Uğurbil, K., Kim, S., 2000. Spatiotemporal dynamics of the BOLD fMRI signals: toward mapping submillimeter cortical columns using the early negative response. Magn. Reson. Med. 44, 231–242.

Duvernoy, H.M., 1999a. Human Brain Stem Vessels. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-07813-6.

Duvernoy, H.M., 1999b. The Human Brain: Surface, Three-Dimensional Sectional Anatomy with MRI, and Blood Supply, Second. ed. SpringerWien, New York.

Duvernoy, H.M., Delon, S., Vannson, J.L., 1981. Cortical blood vessels of the human brain. Brain Res. Bull. 7, 519–579.

Dux, P.E., Ivanoff, J., Asplund, C.L., Marois, R., 2006. Isolation of a central bottleneck of information processing with time-resolved fMRI. Neuron 52, 1109–1120. https://doi.org/10.1016/j.neuron.2006.11.009.

Eickhoff, S.B., Yeo, B.T.T., Genon, S., 2018. Imaging-based parcellations of the human brain. Nat. Rev. Neurosci. https://doi.org/10.1038/s41583-018-0071-7.

Engel, S.A., Rumelhart, D.E., Wandell, B.A., Lee, A.T., Glover, G.H., Chichilnisky, E.J., Shadlen, M.N., 1994. fMRI of human visual cortex. Nature 369, 525. https://doi.org/10.1038/369525a0.

Faull, O.K., Pattinson, K.T.S., 2017. The cortical connectivity of the periaqueductal gray and the conditioned response to the threat of breathlessness. Elife 6. https://doi.org/10.7554/eLife.21749.

Faull, O.K., Jenkinson, M., Clare, S., Pattinson, K.T.S., 2015. Functional subdivision of the human periaqueductal grey in respiratory control using 7 Tesla fMRI. Neuroimage 113, 356–364. https://doi.org/10.1016/j.neuroimage.2015.02.026.

Feinberg, D.A., Moeller, S., Smith, S.M., Auerbach, E., Ramanna, S., Glasser, M.F., Miller, K.L., Uğurbil, K., Yacoub, E., 2010. Multiplexed echo planar imaging for sub-second whole brain fMRI and fast diffusion imaging. PLoS One 5, e15710. https://doi.org/10.1371/journal.pone.0015710.

Fernández-Klett, F., Offenhauser, N., Dirnagl, U., Priller, J., Lindauer, U., 2010. Pericytes in capillaries are contractile in vivo, but arterioles mediate functional hyperemia in the mouse brain. Proc. Natl. Acad. Sci. U. S. A. 107, 22290–22295. https://doi.org/10.1073/pnas.1011321108.

Figueroa, X.F., Chen, C.-C., Campbell, K.P., Damon, D.N., Day, K.H., Ramos, S., Duling, B. R., 2007. Are voltage-dependent ion channels involved in the endothelial cell control of vasomotor tone? Am. J. Physiol. Heart Circ. Physiol. 293, H1371–83. https://doi.org/10.1152/ajpheart.01368.2006.

Finn, E.S., Huber, L., Jangraw, D.C., Molfese, P.J., Bandettini, P.A., 2019. Layer-dependent activity in human prefrontal cortex during working memory. Nat. Neurosci. 22, 1687–1695. https://doi.org/10.1038/s41593-019-0487-z.

Finn, E.S., Glerean, E., Khojandi, A.Y., Nielson, D., Molfese, P.J., Handwerker, D.A., Bandettini, P.A., 2020. Idiosynchrony: from shared responses to individual differences during naturalistic neuroimaging. Neuroimage 215. https://doi.org/10.1016/j.neuroimage.2020.116828.

Fisel, C.R., Ackerman, J.L., Buxton, R.B., Garrido, L., Belliveau, J.W., Rosen, B.R., Brady, T.J., 1991. MR contrast due to microscopically heterogeneous magnetic susceptibility: numerical simulations and applications to cerebral physiology. Magn. Reson. Med. 17, 336–347.

Friston, K.J., Frith, C.D., Turner, R., Frackowiak, R.S.J., 1995. Characterizing evoked hemodynamics with fMRI. Neuroimage 2, 157–165. https://doi.org/10.1006/nimg.1995.1018.

Friston, K.J., Josephs, O., Rees, G., Turner, R., 1998. Nonlinear event-related responses in fMRI. Magn. Reson. Med. 39, 41–52. https://doi.org/10.1002/mrm.1910390109.

Friston, K.J., Mechelli, A., Turner, R., Price, C.J., 2000. Nonlinear responses in fMRI: the Balloon model, Volterra kernels, and other hemodynamics. Neuroimage 12, 466–477. https://doi.org/10.1006/nimg.2000.0630.

Frostig, R.D., Lieke, E.E., Ts'o, D.Y., Grinvald, A., 1990. Cortical functional architecture and local coupling between neuronal activity and the microcirculation revealed by in vivo high-resolution optical imaging of intrinsic signals. Proc. Natl. Acad. Sci. U. S. A. 87, 6082–6086.

Frühholz, S., Trost, W., Grandjean, D., Belin, P., 2020. Neural oscillations in human auditory cortex revealed by fast fMRI during auditory perception. Neuroimage 207. https://doi.org/10.1016/j.neuroimage.2019.116401.

Fukuda, M., Poplawsky, A.J., Kim, S.-G., 2021. Time-dependent spatial specificity of high-resolution fMRI: insights into mesoscopic neurovascular coupling. Philos. Trans. R. Soc. B Biol. Sci. 376, 20190623 https://doi.org/10.1098/rstb.2019.0623.

Gagnon, L., Sakadžić, S., Lesage, F., Musacchia, J.J., Lefebvre, X., Fang, Q., Yücel, M.A., Evans, K.C., Mandeville, E.T., Cohen-Adad, J., Polimeni, J.R., Yaseen, M.A., Lo, E.H., Greve, D.N., Buxton, R.B., Dale, A.M., Devor, A., Boas, D.A., 2015. Quantifying the microvascular origin of bold-fMRI from first principles with two-photon microscopy and an oxygen-sensitive nanoprobe. J. Neurosci. 35, 3663–3675. https://doi.org/10.1523/JNEUROSCI.3555-14.2015.

Gagnon, L., Sakadžić, S., Lesage, F., Pouliot, P., Dale, A.M., Devor, A., Buxton, R.B., Boas, D.A., 2016a. Validation and optimization of hypercapnic-calibrated fMRI from oxygen-sensitive two-photon microscopy. Philos. Trans. R. Soc. B Biol. Sci. 371 https://doi.org/10.1098/rstb.2015.0359.

Gagnon, L., Smith, A.F., Boas, D.A., Devor, A., Secomb, T.W., Sakadžić, S., 2016b. Modeling of cerebral oxygen transport based on in vivo microscopic imaging of microvascular network structure, blood flow, and oxygenation. Front. Comput. Neurosci. https://doi.org/10.3389/fncom.2016.00082.

Gao, J.S., 2015. fMRI Visualization and Methods. University of California, Berkeley.

Gao, Y.R., Greene, S.E., Drew, P.J., 2015. Mechanical restriction of intracortical vessel dilation by brain tissue sculpts the hemodynamic response. Neuroimage 115, 162–176. https://doi.org/10.1016/j.neuroimage.2015.04.054.

Gao, Y.R., Ma, Y., Zhang, Q., Winder, A.T., Liang, Z., Antinori, L., Drew, P.J., Zhang, N., 2017. Time to wake up: studying neurovascular coupling and brain-wide circuit function in the un-anesthetized animal. Neuroimage. https://doi.org/10.1016/j.neuroimage.2016.11.069.

Glasser, M.F., Coalson, T.S., Robinson, E.C., Hacker, C.D., Harwell, J., Yacoub, E., Uğurbil, K., Andersson, J., Beckmann, C.F., Jenkinson, M., Smith, S.M., Van Essen, D. C., 2016. A multi-modal parcellation of human cerebral cortex. Nature 536, 171–178. https://doi.org/10.1038/nature18933.

Glover, G.H., 1999. Deconvolution of impulse response in event-related BOLD fMRI. Neuroimage 9, 416–429.

Gomez, D.E.P., Fultz, N.E., Nasr, S., Polimeni, J.R., Lewis, L.D., 2020. Apparent attenuation of BOLD macro-vascular contributions with high-frequency stimuli. Proc. Intl. Soc. Mag. Reson. Med. 28, 3826.

Gonzalez-Castillo, J., Saad, Z.S., Handwerker, D.A., Inati, S.J., Brenowitz, N., Bandettini, P.A., 2012. Whole-brain, time-locked activation with simple tasks revealed using massive averaging and model-free analysis. Proc. Natl. Acad. Sci. U. S. A. 109, 5487–5492. https://doi.org/10.1073/pnas.1121049109.

Gonzalez-Castillo, J., Hoy, C.W., Handwerker, D.A., Roopchansingh, V., Inati, S.J., Saad, Z.S., Cox, R.W., Bandettini, P.A., 2015. Task dependence, tissue specificity, and spatial distribution of widespread activations in large single-subject functional MRI datasets at 7T. Cereb. Cortex 25, 4667–4677. https://doi.org/10.1093/cercor/bhu148.

Goodyear, B.G., Menon, R.S., 1998. Effect of luminance contrast on BOLD fMRI response in human primary visual areas. J. Neurophysiol. 79, 2204–2207. https://doi.org/10.1152/jn.1998.79.4.2204.

Goodyear, B.G., Menon, R.S., 2001. Brief visual stimulation allows mapping of ocular dominance in visual cortex using fMRI. Hum. Brain Mapp. 14, 210–217. https://doi.org/10.1002/hbm.1053.

Gould, I.G., Linninger, A.A., 2015. Hematocrit distribution and tissue oxygenation in large microcirculatory networks. Microcirculation 22, 1–18. https://doi.org/10.1111/micc.12156.

Gould, I.G., Tsai, P., Kleinfeld, D., Linninger, A., 2017. The capillary bed offers the largest hemodynamic resistance to the cortical blood supply. J. Cereb. Blood Flow Metab. 37, 52–68. https://doi.org/10.1177/0271678X16671146.

Green, D.M., Swets, J.A., 1966. Signal Detection Theory and Psychophysics. John Wiley, Oxford, England.

Greve, D.N., Brown, G.G., Mueller, B.A., Glover, G., Liu, T.T., 2013. A survey of the sources of noise in fMRI. Psychometrika 78, 396–416. https://doi.org/10.1007/s11336-012-9294-0.

Griffeth, V.E.M., Buxton, R.B., 2011. A theoretical framework for estimating cerebral oxygen metabolism changes using the calibrated-BOLD method: modeling the effects of blood volume distribution, hematocrit, oxygen extraction fraction, and tissue signal properties on the BOLD signal. Neuroimage 58, 198–212. https://doi.org/10.1016/j.neuroimage.2011.05.077.

Grill-Spector, K., Malach, R., 2001. fMR-adaptation: a tool for studying the functional properties of human cortical neurons. Acta Psychol. (Amst) 107, 293–321.

Grinvald, A., Lieke, E., Frostig, R.D., Gilbert, C.D., Wiesel, T.N., 1986. Functional architecture of cortex revealed by optical imaging of intrinsic signals. Nature 324, 361–364. https://doi.org/10.1038/324361a0.

Grubb, R.L., Raichle, M.E., Eichling, J.O., Ter-Pogossian, M.M., 1974. The effects of changes in PaCO2 on cerebral blood volume, blood flow, and vascular mean transit time. Stroke 5, 630–639. https://doi.org/10.1161/01.STR.5.5.630.

Gruber, T., Müller, M.M., 2005. Oscillatory brain activity dissociates between associative stimulus content in a repetition priming task in the human EEG. Cereb. Cortex 15, 109–116. https://doi.org/10.1093/cercor/bhh113.

Grutzendler, J., Nedergaard, M., 2019. Cellular control of brain capillary blood flow: in vivo imaging veritas. Trends Neurosci. 42, 528–536. https://doi.org/10.1016/j.tins.2019.05.009.

Gu, H., Stein, E.A., Yang, Y., 2005. Nonlinear responses of cerebral blood volume, blood flow and blood oxygenation signals during visual stimulation. Magn. Reson. Imaging 23, 921–928. https://doi.org/10.1016/j.mri.2005.09.007.

Gutiérrez-Jiménez, E., Cai, C., Mikkelsen, I.K., Rasmussen, P.M., Angleys, H., Merrild, M., Mouridsen, K., Jespersen, S.N., Lee, J., Iversen, N.K., Sakadzic, S., Østergaard, L., 2016. Effect of electrical forepaw stimulation on capillary transit-time heterogeneity (CTH). J. Cereb. Blood Flow Metab. 36, 2072–2086. https://doi.org/10.1177/0271678X16631560.

Hall, C.N., Reynell, C., Gesslein, B., Hamilton, N.B., Mishra, A., Sutherland, B.A., O'Farrell, F.M., Buchan, A.M., Lauritzen, M., Attwell, D., 2014. Capillary pericytes regulate cerebral blood flow in health and disease. Nature 508, 55–60. https://doi.org/10.1038/nature13165.

Hamilton, N.B., Attwell, D., Hall, C.N., 2010. Pericyte-mediated regulation of capillary diameter: a component of neurovascular coupling in health and disease. Front. Neuroenergetics 2. https://doi.org/10.3389/fnene.2010.00005.

Handwerker, D.A., Ollinger, J.M., D'Esposito, M., 2004. Variation of BOLD hemodynamic responses across subjects and brain regions and their effects on statistical analyses. Neuroimage 21, 1639–1651. https://doi.org/10.1016/j.neuroimage.2003.11.029.

Handwerker, D.A., Gonzalez-Castillo, J., D'Esposito, M., Bandettini, P.A., 2012. The continuing challenge of understanding and modeling hemodynamic variation in fMRI. Neuroimage 62, 1017–1023. https://doi.org/10.1016/j.neuroimage.2012.02.015.

Hansen, K.A., David, S.V., Gallant, J.L., 2004. Parametric reverse correlation reveals spatial linearity of retinotopic human V1 BOLD response. Neuroimage 23, 233–241. https://doi.org/10.1016/j.neuroimage.2004.05.012.

Harms, M.P., Melcher, J.R., 2003. Detection and quantification of a wide range of fMRI temporal responses using a physiologically-motivated basis set. Hum. Brain Mapp. 20, 168–183. https://doi.org/10.1002/hbm.10136.

Hartmann, D.A., Berthiaume, A.A., Grant, R.I., Harrill, S.A., Koski, T., Tieu, T., McDowell, K.P., Faino, A.V., Kelly, A.L., Shih, A.Y., 2021. Brain capillary pericytes exert a substantial but slow influence on blood flow. Nat. Neurosci. 24, 633–645. https://doi.org/10.1038/s41593-020-00793-2.

Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., Malach, R., 2004. Intersubject synchronization of cortical activity during natural vision. Science (80-.) 303, 1634–1640. https://doi.org/10.1126/science.1089506.

Hathout, G.M., Gopi, R.K., Bandettini, P., Gambhir, S.S., 1999a. The lag of cerebral hemodynamics with rapidly alternating periodic stimulation: modeling for functional MRI. Magn. Reson. Imaging 17, 9–20. https://doi.org/10.1016/S0730-725X(98)00150-7.

Hathout, G.M., Varjavand, B., Gopi, R.K., 1999b. The early response in fMRI: a modeling approach. Magn. Reson. Med. 41, 550–554. https://doi.org/10.1002/(SICI)1522-2594(199903)41:3<550::AID-MRM18>3.0.CO;2-Q.

Havlíček, M., Uludağ, K., 2020. A dynamical model of the laminar BOLD response. Neuroimage 204, 116209. https://doi.org/10.1016/j.neuroimage.2019.116209.

Havlíček, M., Roebroeck, A., Friston, K., Gardumi, A., Ivanov, D., Uludağ, K., 2015. Physiologically informed dynamic causal modeling of fMRI data. Neuroimage 122, 355–372. https://doi.org/10.1016/j.neuroimage.2015.07.078.

Havlíček, M., Ivanov, D., Poser, B.A., Uludag, K., 2017a. Echo-time dependence of the BOLD response transients – a window into brain functional physiology. Neuroimage 159, 355–370. https://doi.org/10.1016/j.neuroimage.2017.07.034.

Havlíček, M., Ivanov, D., Roebroeck, A., Uludağ, K., 2017b. Determining excitatory and inhibitory neuronal activity from multimodal fMRI data using a generative hemodynamic model. Front. Neurosci. 11 https://doi.org/10.3389/fnins.2017.00616.

He, B.J., Raichle, M.E., 2009. The fMRI signal, slow cortical potential and consciousness. Trends Cogn. Sci. 13, 302–309. https://doi.org/10.1016/j.tics.2009.04.004.

He, X., Yablonskiy, D.A., 2007. Quantitative BOLD: mapping of human cerebral deoxygenated blood volume and oxygen extraction fraction: default state. Magn. Reson. Med. 57, 115–126. https://doi.org/10.1002/mrm.21108.

He, B.J., Snyder, A.Z., Zempel, J.M., Smyth, M.D., Raichle, M.E., 2008. Electrophysiological correlates of the brain's intrinsic large-scale functional architecture. Proc. Natl. Acad. Sci. U. S. A. 105, 16039–16044. https://doi.org/10.1073/pnas.0807010105.

Heckman, G.M., Bouvier, S.E., Carr, V.A., Harley, E.M., Cardinal, K.S., Engel, S.A., 2007. Nonlinearities in rapid event-related fMRI explained by stimulus scaling. Neuroimage 34, 651–660. https://doi.org/10.1016/j.neuroimage.2006.09.038.

Heinzle, J., Koopmans, P.J., den Ouden, H.E.M., Raman, S.S., Stephan, K.E., 2015. A hemodynamic model for layered BOLD signals. Neuroimage 125, 556–570. https://doi.org/10.1016/j.neuroimage.2015.10.025.

Helstrom, C.W., 1964. The detection and resolution of optical signals. IEEE Trans. Inf. Theory 10, 275–287. https://doi.org/10.1109/TIT.1964.1053702.

Hennig, J., Zhong, K., Speck, O., 2007. MR-Encephalography: fast multi-channel monitoring of brain physiology with magnetic resonance. Neuroimage 34, 212–219. https://doi.org/10.1016/j.neuroimage.2006.08.036.

Hill, R.A., Tong, L., Yuan, P., Murikinati, S., Gupta, S., Grutzendler, J., 2015. Regional blood flow in the normal and ischemic brain is controlled by arteriolar smooth muscle cell contractility and not by capillary pericytes. Neuron 87, 95–110. https://doi.org/10.1016/j.neuron.2015.06.001.

Hillman, E.M.C., 2014. Coupling mechanism and significance of the BOLD signal: a status report. Annu. Rev. Neurosci. 37, 161–181. https://doi.org/10.1146/annurev-neuro-071013-014111.

Hillman, E.M.C., Devor, A., Bouchard, M.B., Dunn, A.K., Krauss, G.W., Skoch, J., Bacskai, B.J., Dale, A.M., Boas, D.A., 2007. Depth-resolved optical imaging and microscopy of vascular compartment dynamics during somatosensory stimulation. Neuroimage 35, 89–104. https://doi.org/10.1016/j.neuroimage.2006.11.032.

Hiltunen, T., Kantola, J., Elseoud, A.A., Lepola, P., Suominen, K., Starck, T., Nikkinen, J., Remes, J., Tervonen, O., Palva, S., Kiviniemi, V., Matias Palva, J., 2014. Infra-slow EEG fluctuations are correlated with resting-state network dynamics in fMRI. J. Neurosci. 34, 356–362. https://doi.org/10.1523/JNEUROSCI.0276-13.2014.

Hirano, Y., Stefanovic, B., Silva, A.C., 2011. Spatiotemporal evolution of the functional magnetic resonance imaging response to ultrashort stimuli. J. Neurosci. 31, 1440–1447. https://doi.org/10.1523/JNEUROSCI.3986-10.2011.

Hoge, R.D., Atkinson, J., Gill, B., Crelier, G.R., Marrett, S., Pike, G.B., 1999. Linear coupling between cerebral blood flow and oxygen consumption in activated human cortex. Proc. Natl. Acad. Sci. U. S. A. 96, 9403–9408.

Hua, J., Liu, P., Kim, T., Donahue, M., Rane, S., Chen, J.J., Qin, Q., Kim, S.-G., 2018. MRI techniques to measure arterial and venous cerebral blood volume. Neuroimage. https://doi.org/10.1016/j.neuroimage.2018.02.027.

Huber, L., Handwerker, D.A., Jangraw, D.C., Chen, G., Hall, A., Stüber, C., Gonzalez-Castillo, J., Ivanov, D., Marrett, S., Guidi, M., Goense, J., Poser, B.A., Bandettini, P.A., 2017. High-resolution CBV-fMRI allows mapping of laminar activity and connectivity of cortical input and output in human M1. Neuron 96. https://doi.org/10.1016/j.neuron.2017.11.005, 1253-1263.e7.

Huber, L., Uludağ, K., Möller, H.E., 2019. Non-BOLD contrast for laminar fMRI in humans: CBF, CBV, and CMRO2. Neuroimage 197, 742–760. https://doi.org/10.1016/j.neuroimage.2017.07.041.

Huettel, S.A., McCarthy, G., 2000. Evidence for a refractory period in the hemodynamic response to visual stimuli as measured by MRI. Neuroimage 11, 547–553. https://doi.org/10.1006/nimg.2000.0553.

Huettel, S.A., McCarthy, G., 2001. Regional differences in the refractory period of the hemodynamic response: an event-related fMRI study. Neuroimage 14, 967–976. https://doi.org/10.1006/nimg.2001.0900.

Huettel, S.A., Singerman, J.D., McCarthy, G., 2001. The effects of aging upon the hemodynamic response measured by functional MRI. Neuroimage 13, 161–175. https://doi.org/10.1006/nimg.2000.0675.

Hulvershorn, J., Bloy, L., Gualtieri, E.E., Leigh, J.S., Elliott, M.A., 2005a. Spatial sensitivity and temporal response of spin echo and gradient echo bold contrast at 3 T using peak hemodynamic activation time. Neuroimage 24, 216–223. https://doi.org/10.1016/j.neuroimage.2004.09.033.

Hulvershorn, J., Bloy, L., Gualtieri, E.E., Redmann, C.P., Leigh, J.S., Elliott, M.A., 2005b. Temporal resolving power of spin echo and gradient echo fMRI at 3T with apparent diffusion coefficient compartmentalization. Hum. Brain Mapp. 25, 247–258. https://doi.org/10.1002/hbm.20094.

Hutchinson, E.B., Stefanovic, B., Koretsky, A.P., Silva, A.C., 2006. Spatial flow-volume dissociation of the cerebral microcirculatory response to mild hypercapnia. Neuroimage 32, 520–530. https://doi.org/10.1016/j.neuroimage.2006.03.033.

Huth, A.G., De Heer, W.A., Griffiths, T.L., Theunissen, F.E., Gallant, J.L., 2016. Natural speech reveals the semantic maps that tile human cerebral cortex. Nature 532, 453–458. https://doi.org/10.1038/nature17637.

Iadecola, C., 2004. Neurovascular regulation in the normal brain and in Alzheimer's disease. Nat. Rev. Neurosci. 5, 347–360. https://doi.org/10.1038/nrn1387.

Iadecola, C., 2017. The neurovascular unit coming of age: a journey through neurovascular coupling in health and disease. Neuron 96, 17–42. https://doi.org/10.1016/j.neuron.2017.07.030.

Inan, S., Mitchell, T., Song, A., Bizzell, J., Belger, A., 2004. Hemodynamic correlates of stimulus repetition in the visual and auditory cortices: an fMRI study. Neuroimage 21, 886–893. https://doi.org/10.1016/j.neuroimage.2003.10.029.

Janz, C., Heinrich, S.P., Kornmayer, J., Bach, M., Hennig, J., 2001. Coupling of neural activity and BOLD fMRI response: New insights by combination of fMRI and VEP experiments in transition from single events to continuous stimulation. Magn. Reson. Med. 46, 482–486. https://doi.org/10.1002/mrm.1217.

Jenkinson, M., Beckmann, C.F., Behrens, T.E.J., Woolrich, M.W., Smith, S.M., 2012. FSL. Neuroimage 62, 782–790. https://doi.org/10.1016/j.neuroimage.2011.09.015.

Jones, M., Berwick, J., Johnston, D., Mayhew, J., 2001. Concurrent optical imaging spectroscopy and laser-Doppler flowmetry: the relationship between blood flow, oxygenation, and volume in rodent barrel cortex. Neuroimage 13, 1002–1015. https://doi.org/10.1006/nimg.2001.0808.

Kang, H.C., Burgund, E.D., Lugar, H.M., Petersen, S.E., Schlaggar, B.L., 2003. Comparison of functional activation foci in children and adults using a common stereotactic space. Neuroimage 19, 16–28. https://doi.org/10.1016/S1053-8119(03)00038-7.

Katwal, S.B., Gore, J.C., Gatenby, J.C., Rogers, B.P., 2013. Measuring relative timings of brain activities using fMRI. Neuroimage 66, 436–448. https://doi.org/10.1016/j.neuroimage.2012.10.052.

Kay, K.N., David, S.V., Prenger, R.J., Hansen, K.A., Gallant, J.L., 2008a. Modeling low-frequency fluctuation and hemodynamic response timecourse in event-related fMRI. Hum. Brain Mapp. 29, 142–156. https://doi.org/10.1002/hbm.20379.

Kay, K.N., Naselaris, T., Prenger, R.J., Gallant, J.L., 2008b. Identifying natural images from human brain activity. Nature 452, 352–355. https://doi.org/10.1038/nature06713.

Kemna, L.J., Posse, S., 2001. Effect of respiratory CO2 changes on the temporal dynamics of the hemodynamic response in functional MR imaging. Neuroimage 14, 642–649. https://doi.org/10.1006/nimg.2001.0859.

Kim, T., Kim, S.-G., 2006. Quantification of cerebral arterial blood volume using arterial spin labeling with intravoxel incoherent motion-sensitive gradients. Magn. Reson. Med. 55, 1047–1057. https://doi.org/10.1002/mrm.20867.

Kim, T., Kim, S.-G., 2011a. Temporal dynamics and spatial specificity of arterial and venous blood volume changes during visual stimulation: implication for BOLD quantification. J. Cereb. Blood Flow Metab. 31, 1211–1222. https://doi.org/10.1038/jcbfm.2010.226.

Kim, T., Kim, S.-G., 2011b. Quantitative MRI of cerebral arterial blood volume. Open Neuroimag. J. 5, 136–145. https://doi.org/10.2174/1874440001105010136.

Kim, J.H., Ress, D., 2016. Arterial impulse model for the BOLD response to brief neural activation. Neuroimage 124, 394–408. https://doi.org/10.1016/j.neuroimage.2015.08.068.

Kim, D.S., Duong, T.Q., Kim, S., 2000. High-resolution mapping of iso-orientation columns by fMRI. Nat. Neurosci. 3, 164–169. https://doi.org/10.1038/72109.

Kim, T., Hendrich, K.S., Masamoto, K., Kim, S.-G., 2007. Arterial versus total blood volume changes during neural activity-induced cerebral blood flow change: implication for BOLD fMRI. J. Cereb. Blood Flow Metab. 27, 1235–1247. https://doi.org/10.1038/sj.jcbfm.9600429.

Kim, T., Hendrich, K., Kim, S.-G., 2008. Functional MRI with magnetization transfer effects: determination of BOLD and arterial blood volume changes. Magn. Reson. Med. 60, 1518–1523. https://doi.org/10.1002/mrm.21766.

Kim, J.H., Khan, R., Thompson, J.K., Ress, D., 2013. Model of the transient neurovascular response based on prompt arterial dilation. J. Cereb. Blood Flow Metab. 33, 1429–1439. https://doi.org/10.1038/jcbfm.2013.90.

Kim, J.H., Taylor, A.J., Wang, D.J.J., Zou, X., Ress, D., 2020. Dynamics of the cerebral blood flow response to brief neural activity in human visual cortex. J. Cereb. Blood Flow Metab. 40, 1823–1837. https://doi.org/10.1177/0271678X19869034.

Kiselev, V.G., Posse, S., 1999. Analytical model of susceptibility-induced MR signal dephasing: effect of diffusion in a microvascular network. Magn. Reson. Med. 41, 499–509.

Kiselev, V.G., Strecker, R., Ziyeh, S., Speck, O., Hennig, J., 2005. Vessel size imaging in humans. Magn. Reson. Med. 53, 553–563. https://doi.org/10.1002/mrm.20383.

Kleinfeld, D., Mitra, P.P., Helmchen, F., Denk, W., 1998. Fluctuations and stimulus-induced changes in blood flow observed in individual capillaries in layers 2 through 4 of rat neocortex. Proc. Natl. Acad. Sci. U. S. A. 95, 15741–15746.

Kleinfeld, D., Blinder, P., Drew, P.J., Driscoll, J.D., Muller, A., Tsai, P.S., Shih, A.Y., 2011. A guide to delineate the logic of neurovascular signaling in the brain. Front. Neuroenergetics. https://doi.org/10.3389/fnene.2011.00001.

Kohler, P.J., Fogelson, S.V., Reavis, E.A., Meng, M., Guntupalli, J.S., Hanke, M., Halchenko, Y.O., Connolly, A.C., Haxby, J.V., Tse, P.U., 2013. Pattern classification precedes region-average hemodynamic response in early visual cortex. Neuroimage 78, 249–260. https://doi.org/10.1016/j.neuroimage.2013.04.019.

Kok, P., Bains, L.J., van Mourik, T., Norris, D.G., de Lange, F.P., 2016. Selective activation of the deep layers of the human primary visual cortex by top-down feedback. Curr. Biol. 26, 371–376. https://doi.org/10.1016/j.cub.2015.12.038.

Krekelberg, B., Boynton, G.M., van Wezel, R.J.A., 2006. Adaptation: from single cells to BOLD signals. Trends Neurosci. https://doi.org/10.1016/j.tins.2006.02.008.

Krieger, S.N., Streicher, M.N., Trampel, R., Turner, R., 2012. Cerebral blood volume changes during brain activation. J. Cereb. Blood Flow Metab. 32, 1618–1631. https://doi.org/10.1038/jcbfm.2012.63.

Kriegeskorte, N., Cusack, R., Bandettini, P.A., 2010. How does an fMRI voxel sample the neuronal activity pattern: compact-kernel or complex spatiotemporal filter? Neuroimage 49, 1965–1976. https://doi.org/10.1016/j.neuroimage.2009.09.059.

Lai, S., Hopkins, A.L., Haacke, E.M., Li, D., Wasserman, B.A., Buckley, P., Friedman, L., Meltzer, H., Hedera, P., Friedland, R., 1993. Identification of vascular structures as a major source of signal contrast in high resolution 2D and 3D functional activation imaging of the motor cortex at l.5T preliminary results. Magn. Reson. Med. 30, 387–392. https://doi.org/10.1002/mrm.1910300318.

Lambers, H., Segeroth, M., Albers, F., Wachsmuth, L., van Alst, T.M., Faber, C., 2020. A cortical rat hemodynamic response function for improved detection of BOLD activation under common experimental conditions. Neuroimage 208, 116446. https://doi.org/10.1016/j.neuroimage.2019.116446.

Lau, C., Zhou, I.Y., Cheung, M.M., Chan, K.C., Wu, E.X., 2011. BOLD temporal dynamics of rat superior colliculus and lateral geniculate nucleus following short duration visual stimulation. PLoS One 6, e18914. https://doi.org/10.1371/journal.pone.0018914.

Lauwers, F., Cassot, F., Lauwers-Cances, V., Puwanarajah, P., Duvernoy, H.M., 2008. Morphometry of the human cerebral cortex microcirculation: general characteristics and space-related profiles. Neuroimage 39, 936–948. https://doi.org/10.1016/j.neuroimage.2007.09.024.

Lawrence, S.J.D., Formisano, E., Muckli, L., de Lange, F.P., 2019. Laminar fMRI: applications for cognitive neuroscience. Neuroimage 197, 785–791. https://doi.org/10.1016/j.neuroimage.2017.07.004.

Lee, S.P., Duong, T.Q., Yang, G., Iadecola, C., Kim, S., 2001. Relative changes of cerebral arterial and venous blood volumes during increased cerebral blood flow: implications for BOLD fMRI. Magn. Reson. Med. 45, 791–800.

Lee, H.-L., Zahneisen, B., Hugger, T., LeVan, P., Hennig, J., 2013. Tracking dynamic resting-state networks at higher frequencies using MR-encephalography. Neuroimage 65, 216–222. https://doi.org/10.1016/j.neuroimage.2012.10.015.

Leite, F.P., Mandeville, J.B., 2006. Characterization of event-related designs using BOLD and IRON fMRI. Neuroimage 29, 901–909. https://doi.org/10.1016/j.neuroimage.2005.08.022.

Lewis, L.D., Setsompop, K., Rosen, B.R., Polimeni, J.R., 2016. Fast fMRI can detect oscillatory neural activity in humans. Proc. Natl. Acad. Sci. U. S. A. 113, E6679–E6685. https://doi.org/10.1073/pnas.1608117113.

Lewis, L.D., Setsompop, K., Stelzer, J., Bause, J., Ehses, P., Scheffler, K., Rosen, B.R., Polimeni, J.R., 2017. Identifying neural contributions to high frequency dynamics in the fMRI signal at 9.4 Tesla. In: 23rd Annu. Meet. Organ. Hum. Brain Mapping. June 25–29, 2017, Vancouver, BC, Canada, p. 23.

Lewis, L.D., Bonmassar, G., Gupta, K., Setsompop, K., Stickgold, R., Rosen, B.R., Polimeni, J.R., 2018a. Focal thalamic activity at the moment of awakening identified through simultaneous EEG and fast fMRI. Soc. Neurosci. Abstr. 192.08.

Lewis, L.D., Setsompop, K., Rosen, B.R., Polimeni, J.R., 2018b. Stimulus-dependent hemodynamic response timing across the human subcortical-cortical visual pathway identified through high spatiotemporal resolution 7T fMRI. Neuroimage 181, 279–291. https://doi.org/10.1016/j.neuroimage.2018.06.056.

Li, B., Freeman, R.D., 2007. High-resolution neurometabolic coupling in the lateral geniculate nucleus. J. Neurosci. 27, 10223–10229. https://doi.org/10.1523/JNEUROSCI.1505-07.2007.

Li, M., Newton, A.T., Anderson, A.W., Ding, Z., Gore, J.C., 2019. Characterization of the hemodynamic response function in white matter tracts for event-related fMRI. Nat. Commun. 10 https://doi.org/10.1038/s41467-019-09076-2.

Liang, C.L., Ances, B.M., Perthen, J.E., Moradi, F., Liau, J., Buracas, G.T., Hopkins, S.R., Buxton, R.B., 2013. Luminance contrast of a visual stimulus modulates the BOLD response more than the cerebral blood flow response in the human brain. Neuroimage 64, 104–111. https://doi.org/10.1016/j.neuroimage.2012.08.077.

Lin, F.-H., Witzel, T., Mandeville, J.B., Polimeni, J.R., Zeffiro, T.A., Greve, D.N., Wiggins, G.C., Wald, L.L., Belliveau, J.W., 2008. Event-related single-shot volumetric functional magnetic resonance inverse imaging of visual processing. Neuroimage 42, 230–247. https://doi.org/10.1016/j.neuroimage.2008.04.179.

Lin, F.H., Witzel, T., Raij, T., Ahveninen, J., Wen-Kai Tsai, K., Chu, Y.H., Chang, W.T., Nummenmaa, A., Polimeni, J.R., Kuo, W.J., Hsieh, J.C., Rosen, B.R., Belliveau, J.W., 2013. fMRI hemodynamics accurately reflects neuronal timing in the human brain measured by MEG. Neuroimage 78C, 372–384. https://doi.org/10.1016/j.neuroimage.2013.04.017.

Lin, F.H., Chu, Y.H., Hsu, Y.C., Lin, J.F.L., Tsai, K.W.K., Tsai, S.Y., Kuo, W.J., 2015. Significant feed-forward connectivity revealed by high frequency components of BOLD fMRI signals. Neuroimage 121, 69–77. https://doi.org/10.1016/j.neuroimage.2015.07.036.

Lin, F.H., Polimeni, J.R., Lin, J.F.L., Tsai, K.W.K., Chu, Y.H., Wu, P.Y., Li, Y.T., Hsu, Y.C., Tsai, S.Y., Kuo, W.J., 2018. Relative latency and temporal variability of hemodynamic responses at the human primary visual cortex. Neuroimage 164, 194–201. https://doi.org/10.1016/j.neuroimage.2017.01.041.

Lindauer, U., Royl, G., Leithner, C., Kühl, M., Gold, L., Gethmann, J., Kohl-Bareis, M., Villringer, A., Dirnagl, U., 2001. No evidence for early decrease in blood oxygenation in rat whisker cortex in response to functional activation. Neuroimage 13, 988–1001. https://doi.org/10.1006/nimg.2000.0709.

Lindquist, M.A., Meng Loh, J., Atlas, L.Y., Wager, T.D., 2009. Modeling the hemodynamic response function in fMRI: efficiency, bias and mis-modeling. Neuroimage 45, S187–98. https://doi.org/10.1016/j.neuroimage.2008.10.065.

Liu, H.L., Gao, J.H., 2000. An investigation of the impulse functions for the nonlinear BOLD response in functional MRI. Magn. Reson. Imaging 18, 931–938. https://doi.org/10.1016/S0730-725X(00)00214-9.

Liu, T.T., Frank, L.R., Wong, E.C., Buxton, R.B., 2001. Detection power, estimation efficiency, and predictability in event-related fMRI. Neuroimage 13, 759–773. https://doi.org/10.1006/nimg.2000.0728.

Liu, T.T., Behzadi, Y., Restom, K., Uludağ, K., Lu, K., Buracas, G.T., Dubowitz, D.J., Buxton, R.B., 2004. Caffeine alters the temporal dynamics of the visual BOLD response. Neuroimage 23, 1402–1413. https://doi.org/10.1016/j.neuroimage.2004.07.061.

Liu, Z., Rios, C., Zhang, N., Yang, L., Chen, W., He, B., 2010. Linear and nonlinear relationships between visual stimuli, EEG and BOLD fMRI signals. Neuroimage 50, 1054–1066. https://doi.org/10.1016/j.neuroimage.2010.01.017.

Logothetis, N.K., 2008. What we can do and what we cannot do with fMRI. Nature 453, 869–878. https://doi.org/10.1038/nature06976.

Logothetis, N.K., Pauls, J., Augath, M., Trinath, T., Oeltermann, A., 2001. Neurophysiological investigation of the basis of the fMRI signal. Nature 412, 150–157. https://doi.org/10.1038/35084005.

Logothetis, N.K., Murayama, Y., Augath, M., Steffen, T., Werner, J., Oeltermann, A., 2009. How not to study spontaneous activity. Neuroimage 45, 1080–1089. https://doi.org/10.1016/j.neuroimage.2009.01.010.

Lorthois, S., Cassot, F., Lauwers, F., 2011a. Simulation study of brain blood flow regulation by intra-cortical arterioles in an anatomically accurate large human vascular network: part I: methodology and baseline flow. Neuroimage 54, 1031–1042. https://doi.org/10.1016/j.neuroimage.2010.09.032.

Lorthois, S., Cassot, F., Lauwers, F., 2011b. Simulation study of brain blood flow regulation by intra-cortical arterioles in an anatomically accurate large human vascular network. Part II: flow variations induced by global or localized modifications of arteriolar diameters. Neuroimage 54, 2840–2853. https://doi.org/10.1016/j.neuroimage.2010.10.040.

Ma, Y., Shaik, M.A., Kozberg, M.G., Kim, S.H., Portes, J.P., Timerman, D., Hillman, E.M.C., 2016. Resting-state hemodynamics are spatiotemporally coupled to synchronized and symmetric neural activity in excitatory neurons. Proc. Natl. Acad. Sci. U. S. A. 113, E8463–E8471. https://doi.org/10.1073/pnas.1525369113.

Macdonald, J.S.P., Mathan, S., Yeung, N., 2011. Trial-by-trial variations in subjective attentional state are reflected in ongoing prestimulus EEG alpha oscillations. Front. Psychol. 2 https://doi.org/10.3389/fpsyg.2011.00082.

Malonek, D., Dirnagl, U., Lindauer, U., Yamada, K., Kanno, I., Grinvald, A., 1997. Vascular imprints of neuronal activity: relationships between the dynamics of cortical blood flow, oxygenation, and volume changes following sensory stimulation. Proc. Natl. Acad. Sci. U. S. A. 94, 14826–14831.

Mandeville, J.B., Marota, J.J., Kosofsky, B.E., Keltner, J.R., Weissleder, R., Rosen, B.R., Weisskoff, R.M., 1998. Dynamic functional imaging of relative cerebral blood volume during rat forepaw stimulation. Magn. Reson. Med. 39, 615–624.

Mandeville, J.B., Marota, J.J., Ayata, C., Moskowitz, M.A., Weisskoff, R.M., Rosen, B.R., 1999a. MRI measurement of the temporal evolution of relative CMRO(2) during rat forepaw stimulation. Magn. Reson. Med. 42, 944–951.

Mandeville, J.B., Marota, J.J., Ayata, C., Zaharchuk, G., Moskowitz, M.A., Rosen, B.R., Weisskoff, R.M., 1999b. Evidence of a cerebrovascular postarteriole windkessel with delayed compliance. J. Cereb. Blood Flow Metab. 19, 679–689. https://doi.org/10.1097/00004647-199906000-00012.

Manuel, J., Färber, N., Gerlach, D.A., Heusser, K., Jordan, J., Tank, J., Beissner, F., 2020. Deciphering the neural signature of human cardiovascular regulation. Elife 9, 1–14. https://doi.org/10.7554/eLife.55316.

Marcus, M.L., Heistad, D.D., Ehrhardt, J.C., Abboud, F.M., 1977. Regulation of total and regional spinal cord blood flow. Circ. Res. 41, 128–134. https://doi.org/10.1161/01.RES.41.1.128.

Markuerkiaga, I., Barth, M., Norris, D.G., 2016. A cortical vascular model for examining the specificity of the laminar BOLD signal. Neuroimage 132, 491–498. https://doi.org/10.1016/j.neuroimage.2016.02.073.

Martin, C., Martindale, J., Berwick, J., Mayhew, J., 2006. Investigating neural-hemodynamic coupling and the hemodynamic response function in the awake rat. Neuroimage 32, 33–48. https://doi.org/10.1016/j.neuroimage.2006.02.021.

Marvel, C.L., Desmond, J.E., 2012. From storage to manipulation: how the neural correlates of verbal working memory reflect varying demands on inner speech. Brain Lang. 120, 42–51. https://doi.org/10.1016/j.bandl.2011.08.005.

Massimini, M., Huber, R., Ferrarelli, F., Hill, S., Tononi, G., 2004. The sleep slow oscillation as a traveling wave. J. Neurosci. 24, 6862–6870. https://doi.org/10.1523/JNEUROSCI.1318-04.2004.

Mateo, C., Knutsen, P.M., Tsai, P.S., Shih, A.Y., Kleinfeld, D., 2017. Entrainment of arteriole vasomotor fluctuations by neural activity is a basis of blood-oxygenation-level-dependent "resting-state" connectivity. Neuron 96. https://doi.org/10.1016/j.neuron.2017.10.012, 936-948.e3.

Mazzoni, J., Cutforth, T., Agalliu, D., 2015. Dissecting the role of smooth muscle cells versus pericytes in regulating cerebral blood flow using in vivo optical imaging. Neuron 87, 4–6. https://doi.org/10.1016/j.neuron.2015.06.024.

Mchedlishvili, G., Kuridze, N., 1984. The modular organization of the pial arterial system in phylogeny. J. Cereb. Blood Flow Metab. 4, 391–396. https://doi.org/10.1038/jcbfm.1984.57.

Menon, R.S., Goodyear, B.G., 1999. Submillimeter functional localization in human striate cortex using BOLD contrast at 4 Tesla: implications for the vascular point-spread function. Magn. Reson. Med. 41, 230–235.

Menon, R.S., Kim, S.G., 1999. Spatial and temporal limits in cognitive neuroimaging with fMRI. Trends Cogn. Sci. https://doi.org/10.1016/S1364-6613(99)01329-7.

Menon, R.S., Ogawa, S., Hu, X., Strupp, J.P., Anderson, P., Uğurbil, K., 1995. BOLD based functional MRI at 4 Tesla includes a capillary bed contribution: echo-planar imaging correlates with previous optical imaging using intrinsic signals. Magn. Reson. Med. 33, 453–459. https://doi.org/10.1002/mrm.1910330323.

Menon, R.S., Luknowsky, D.C., Gati, J.S., 1998. Mental chronometry using latency-resolved functional MRI. Proc. Natl. Acad. Sci. U. S. A. 95, 10902–10907.

Miezin, F.M., Maccotta, L., Ollinger, J.M., Petersen, S.E., Buckner, R.L., 2000. Characterizing the hemodynamic response: effects of presentation rate, sampling procedure, and the possibility of ordering brain activity based on relative timing. Neuroimage 11, 735–759. https://doi.org/10.1006/nimg.2000.0568.

Mildner, T., Norris, D.G., Schwarzbauer, C., Wiggins, C.J., 2001. A qualitative test of the balloon model for BOLD-based MR signal changes at 3T. Magn. Reson. Med. 46, 891–899. https://doi.org/10.1002/mrm.1274.

Miller, K.L., Luh, W.M., Liu, T.T., Martinez, A., Obata, T., Wong, E.C., Frank, L.R., Buxton, R.B., 2001. Nonlinear temporal dynamics of the cerebral blood flow response. Hum. Brain Mapp. 13, 1–12.

Misaki, M., Luh, W.-M., Bandettini, P.A., 2013. Accurate decoding of sub-TR timing differences in stimulations of sub-voxel regions from multi-voxel response patterns. Neuroimage 66, 623–633. https://doi.org/10.1016/j.neuroimage.2012.10.069.

Miyawaki, Y., Handwerker, D., Gonzalez-Castillo, J., Huber, L., Khojandi, A., Chai, Y., Bandettini, P.A., 2020. Event-related decoding of visual stimulus information using short-TR BOLD fMRI at 7T. Proc. Intl. Soc. Mag. Reson. Med. 28, 4015.

Moeller, S., Yacoub, E., Olman, C.A., Auerbach, E., Strupp, J., Harel, N., Uğurbil, K., 2010. Multiband multislice GE-EPI at 7 Tesla, with 16-fold acceleration using partial parallel imaging with application to high spatial and temporal whole-brain fMRI. Magn. Reson. Med. 63, 1144–1153. https://doi.org/10.1002/mrm.22361.

Moerel, M., De Martino, F., Kemper, V.G., Schmitter, S., Vu, A.T., Uğurbil, K., Formisano, E., Yacoub, E., 2018. Sensitivity and specificity considerations for fMRI encoding, decoding, and mapping of auditory cortex at ultra-high field. Neuroimage 164, 18–31. https://doi.org/10.1016/j.neuroimage.2017.03.063.

Mohamed, F.B., Pinus, A.B., Faro, S.H., Patel, D., Tracy, J.I., 2002. BOLD fMRI of the visual cortex: quantitative responses measured with a graded stimulus at 1.5 Tesla. J. Magn. Reson. Imaging 16, 128–136. https://doi.org/10.1002/jmri.10155.

Mölle, M., Marshall, L., Gais, S., Born, J., 2004. Learning increases human electroencephalographic coherence during subsequent slow sleep oscillations. Proc. Natl. Acad. Sci. U. S. A. 101, 13963–13968. https://doi.org/10.1073/pnas.0402820101.

Moon, C.-H., Fukuda, M., Park, S.-H., Kim, S.-G., 2007. Neural interpretation of blood oxygenation level-dependent fMRI maps at submillimeter columnar resolution. J. Neurosci. 27, 6892–6902. https://doi.org/10.1523/JNEUROSCI.0445-07.2007.

Moradi, F., Buxton, R.B., 2013. Adaptation of cerebral oxygen metabolism and blood flow and modulation of neurovascular coupling with prolonged stimulation in human visual cortex. Neuroimage 82, 182–189. https://doi.org/10.1016/j.neuroimage.2013.05.110.

Moses, P., Dinino, M., Hernandez, L., Liu, T.T., 2014. Developmental changes in resting and functional cerebral blood flow and their relationship to the BOLD response. Hum. Brain Mapp. 35, 3188–3198. https://doi.org/10.1002/hbm.22394.

Muckli, L., De Martino, F., Vizioli, L., Petro, L.S., Smith, F.W., Uğurbil, K., Goebel, R., Yacoub, E., 2015. Contextual feedback to superficial layers of V1. Curr. Biol. 25, 2690–2695. https://doi.org/10.1016/j.cub.2015.08.057.

Mullinger, K.J., Cherukara, M.T., Buxton, R.B., Francis, S.T., Mayhew, S.D., 2017. Post-stimulus fMRI and EEG responses: evidence for a neuronal origin hypothesised to be inhibitory. Neuroimage 157, 388–399. https://doi.org/10.1016/j.neuroimage.2017.06.020.

Murphy, K., Birn, R.M., Bandettini, P.A., 2013. Resting-state fMRI confounds and cleanup. Neuroimage 80, 349–359. https://doi.org/10.1016/j.neuroimage.2013.04.001.

Nangini, C., Macintosh, B.J., Tam, F., Staines, W.R., Graham, S.J., 2005. Assessing linear time-invariance in human primary somatosensory cortex with BOLD fMRI using vibrotactile stimuli. Magn. Reson. Med. 53, 304–311. https://doi.org/10.1002/mrm.20363.

Nasr, S., Polimeni, J.R., Tootell, R.B.H., 2016. Interdigitated color- and disparity-selective columns within human visual cortical areas V2 and V3. J. Neurosci. 36, 1841–1857. https://doi.org/10.1523/JNEUROSCI.3518-15.2016.

Ngo, H.V.V., Martinetz, T., Born, J., Mölle, M., 2013. Auditory closed-loop stimulation of the sleep slow oscillation enhances memory. Neuron 78, 545–553. https://doi.org/10.1016/j.neuron.2013.03.006.

Nishimoto, S., Vu, A.T., Naselaris, T., Benjamini, Y., Yu, B., Gallant, J.L., 2011. Reconstructing visual experiences from brain activity evoked by natural movies. Curr. Biol. 21, 1641–1646. https://doi.org/10.1016/j.cub.2011.08.031.

Nix, W., Capra, N.F., Erdmann, W., Halsey, J.H., 1976. Comparison of vascular reactivity in spinal cord and brain. Stroke 7, 560–563. https://doi.org/10.1161/01.STR.7.6.560.

Nizar, K., Uhlirova, H., Tian, P., Saisan, P.A., Cheng, Q., Reznichenko, L., Weldy, K.L., Steed, T.C., Sridhar, V.B., Macdonald, C.L., Cui, J., Gratiy, S.L., Sakadzic, S., Boas, D. A., Beka, T.I., Einevoll, G.T., Chen, J., Masliah, E., Dale, A.M., Silva, G.A., Devor, A., 2013. In vivo stimulus-induced vasodilation occurs without IP3 receptor activation and may precede astrocytic calcium increase. J. Neurosci. 33, 8411–8422. https://doi.org/10.1523/JNEUROSCI.3285-12.2013.

Norris, D.G., Polimeni, J.R., 2019. Laminar (f)MRI: a short history and future prospects. Neuroimage 197, 643–649. https://doi.org/10.1016/j.neuroimage.2019.04.082.

Obata, T., Liu, T.T., Miller, K.L., Luh, W.M., Wong, E.C., Frank, L.R., Buxton, R.B., 2004. Discrepancies between BOLD and flow dynamics in primary and supplementary motor areas: application of the balloon model to the interpretation of BOLD transients. Neuroimage 21, 144–153. https://doi.org/10.1016/j.neuroimage.2003.08.040.

Ogawa, S., Menon, R.S., Tank, D.W., Kim, S., Merkle, H., Ellermann, J.M., Uğurbil, K., 1993. Functional brain mapping by blood oxygenation level-dependent contrast magnetic resonance imaging. A comparison of signal characteristics with a biophysical model. Biophys. J. 64, 803–812. https://doi.org/10.1016/S0006-3495(93)81441-3.

Ogawa, S., Lee, T.M., Stepnoski, R., Chen, W., Zhu, X.H., Uğurbil, K., 2000. An approach to probe some neural systems interaction by functional MRI at neural time scale down to milliseconds. Proc. Natl. Acad. Sci. U. S. A. 97, 11026–11031.

Olszowy, W., Aston, J., Rua, C., Williams, G.B., 2019. Accurate autocorrelation modeling substantially improves fMRI reliability. Nat. Commun. 10, 1220. https://doi.org/10.1038/s41467-019-09230-w.

Oppenheim, A.V., Willsky, A.S., Nawab, S.H., 1997. Signals and Systems. Prentice Hall.

Ou, W., Raij, T., Lin, F.H., Golland, P., Hämäläinen, M., 2009. Modeling adaptation effects in fMRI analysis. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). NIH Public Access, pp. 1009–1017. https://doi.org/10.1007/978-3-642-04268-3_124.

Pan, W.J., Thompson, G.J., Magnuson, M.E., Jaeger, D., Keilholz, S., 2013. Infraslow LFP correlates to resting-state fMRI BOLD signals. Neuroimage 74, 288–297. https://doi.org/10.1016/j.neuroimage.2013.02.035.

Pedregosa, F., Eickenberg, M., Ciuciu, P., Thirion, B., Gramfort, A., 2015. Data-driven HRF estimation for encoding and decoding models. Neuroimage 104, 209–220. https://doi.org/10.1016/j.neuroimage.2014.09.060.

Peppiatt, C.M., Howarth, C., Mobbs, P., Attwell, D., 2006. Bidirectional control of CNS capillary diameter by pericytes. Nature 443, 700–704. https://doi.org/10.1038/nature05193.

Petridou, N., Siero, J.C.W., 2019. Laminar fMRI: what can the time domain tell us? Neuroimage 197, 761–771. https://doi.org/10.1016/j.neuroimage.2017.07.040.

Peyrounette, M., Davit, Y., Quintard, M., Lorthois, S., 2018. Multiscale modelling of blood flow in cerebral microcirculation: details at capillary scale control accuracy at the level of the cortex. PLoS One 13, e0189474. https://doi.org/10.1371/journal.pone.0189474.

Pfannmoeller, J.P., Berman, A.J.L., Kura, S., Cheng, X., Boas, D.A., Polimeni, J.R., 2019. Quantification of draining vein dominance across cortical depths in BOLD fMRI from first principles using realistic Vascular Anatomical Networks. Proc. Intl. Soc. Mag. Reson. Med. 27, 3715.

Pfannmoeller, J.P., Hartung, G.A., Cheng, X., Berman, A., Boas, D., Polimeni, J.R., 2021. Simulations of the BOLD non-linearity based on a viscoelastic model for capillary and vein compliance. Proc. Intl. Soc. Mag. Reson. Med. 29, 2856.

Pfeifer, R., 1928. Die Angioarchitektonik der Grosshirnrinde. J. Springer, Berlin.

Pfeuffer, J., McCullough, J.C., Van de Moortele, P.-F., Uğurbil, K., Hu, X., 2003. Spatial dependence of the nonlinear BOLD response at short stimulus duration. Neuroimage 18, 990–1000. https://doi.org/10.1016/S1053-8119(03)00035-1.

Piché, M., Paquette, T., Leblond, H., 2017. Tight neurovascular coupling in the spinal cord during nociceptive stimulation in intact and spinal rats. Neuroscience 355, 1–8. https://doi.org/10.1016/j.neuroscience.2017.04.038.

Pisauro, M.A., Dhruv, N.T., Carandini, M., Benucci, A., 2013. Fast hemodynamic responses in the visual cortex of the awake mouse. J. Neurosci. 33, 18343–18351. https://doi.org/10.1523/JNEUROSCI.2130-13.2013.

Polimeni, J.R., Uludağ, K., 2018. Neuroimaging with ultra-high field MRI: present and future. Neuroimage 168, 1–6. https://doi.org/10.1016/j.neuroimage.2018.01.072.

Polimeni, J.R., Wald, L.L., 2018. Magnetic Resonance Imaging technology — bridging the gap between noninvasive human imaging and optical microscopy. Curr. Opin. Neurobiol. 50, 250–260. https://doi.org/10.1016/j.conb.2018.04.026.

Polimeni, J.R., Renvall, V., Zaretskaya, N., Fischl, B., 2018. Analysis strategies for high-resolution UHF-fMRI data. Neuroimage 168, 296–320. https://doi.org/10.1016/j.neuroimage.2017.04.053.

Pouliot, P., Gagnon, L., Lam, T., Avti, P.K., Bowen, C., Desjardins, M., Kakkar, A.K., Thorin, E., Sakadzic, S., Boas, D.A., Lesage, F., 2017. Magnetic resonance fingerprinting based on realistic vasculature in mice. Neuroimage 149, 436–445. https://doi.org/10.1016/j.neuroimage.2016.12.060.

Purdon, P.L., Weisskoff, R.M., 1998. Effect of temporal autocorrelation due to physiological noise and stimulus paradigm on voxel-level false-positive rates in fMRI. Hum. Brain Mapp. 6, 239–249.

Reichold, J., Stampanoni, M., Lena Keller, A., Buck, A., Jenny, P., Weber, B., 2009. Vascular graph model to simulate the cerebral blood flow in realistic vascular networks. J. Cereb. Blood Flow Metab. 29, 1429–1443. https://doi.org/10.1038/jcbfm.2009.58.

Richter, W., Richter, M., 2003. The shape of the fMRI BOLD response in children and adults changes systematically with age. Neuroimage 20, 1122–1131. https://doi.org/10.1016/S1053-8119(03)00347-1.

Robson, M.D., Dorosz, J.L., Gore, J.C., 1998. Measurements of the temporal fMRI response of the human auditory cortex to trains of tones. Neuroimage 7, 185–198. https://doi.org/10.1006/nimg.1998.0322.

Rocca, R., Coventry, K.R., Tylén, K., Staib, M., Lund, T.E., Wallentin, M., 2020. Language beyond the language system: dorsal visuospatial pathways support processing of demonstratives and spatial language during naturalistic fast fMRI. Neuroimage 216. https://doi.org/10.1016/j.neuroimage.2019.116128.

Rungta, R.L., Chaigneau, E., Osmanski, B.-F., Charpak, S., 2018. Vascular compartmentalization of functional hyperemia from the synapse to the pia. Neuron 99. https://doi.org/10.1016/j.neuron.2018.06.012, 362-375.e4.

Saad, Z.S., DeYoe, E.A., Ropella, K.M., 2003. Estimation of FMRI response delays. Neuroimage 18, 494–504. https://doi.org/10.1016/S1053-8119(02)00024-1.

Saalmann, Y.B., Kastner, S., 2015. The cognitive thalamus. Front. Syst. Neurosci. 9, 39. https://doi.org/10.3389/fnsys.2015.00039.

Sadaghiani, S., Uğurbil, K., Uludağ, K., 2009. Neural activity-induced modulation of BOLD poststimulus undershoot independent of the positive signal. Magn. Reson. Imaging 27, 1030–1038. https://doi.org/10.1016/j.mri.2009.04.003.

Sahib, A.K., Mathiak, K., Erb, M., Elshahabi, A., Klamer, S., Scheffler, K., Focke, N.K., Ethofer, T., 2016. Effect of temporal resolution and serial autocorrelations in event-related functional MRI. Magn. Reson. Med. 76, 1805–1813. https://doi.org/10.1002/mrm.26073.

Sasai, S., Koike, T., Sugawara, S.K., Hamano, Y.H., Sumiya, M., Okazaki, S., Takahashi, H.K., Taga, G., Sadato, N., 2021. Frequency-specific task modulation of human brain functional networks: a fast fMRI study. Neuroimage 224, 117375. https://doi.org/10.1016/j.neuroimage.2020.117375.

Sasaki, S., Yazawa, I., Miyakawa, N., Mochida, H., Shinomiya, K., Kamino, K., Momose-Sato, Y., Sato, K., 2002. Optical imaging of intrinsic signals induced by peripheral nerve stimulation in the in vivo rat spinal cord. Neuroimage 17, 1240–1255. https://doi.org/10.1006/nimg.2002.1286.

Satpute, A.B., Wager, T.D., Cohen-Adad, J., Bianciardi, M., Choi, J.-K., Buhle, J.T., Wald, L.L., Barrett, L.F., 2013. Identification of discrete functional subregions of the human periaqueductal gray. Proc. Natl. Acad. Sci. U. S. A. 110, 17101–17106. https://doi.org/10.1073/pnas.1306095110.

Savoy, R.L., 2001. History and future directions of human brain mapping and functional neuroimaging. Acta Psychol. (Amst) 107, 9–42. https://doi.org/10.1016/S0001-6918(01)00018-X.

Savoy, R.L., O'Craven, K.M., Weisskoff, R.M., Davis, T.L., Baker, J.R., Rosen, B.R., 1994. Exploring the temporal resolution boundaries of fMRI: measuring responses to very brief visual stimuli. In: FLAnnu. Meet. Soc. Neurosci. Miami Beach, 20, p. 1264.

Savoy, R.L., Bandettini, P.A., O'Craven, K.M., Kwong, K.K., Davis, T.L., Baker, J.R., Weiskoff, R.M., Rosen, B.R., 1995. Pushing the temporal resolution of fMRI: studies of very brief visual stimuli, onset variability and asynchrony, and stimulus-correlated changes in noise. Proc. Soc. Magn. Reson. Third Sci. Meet. Exhib. VIII 450.

Schlegel, F., Schroeter, A., Rudin, M., 2015. The hemodynamic response to somatosensory stimulation in mice depends on the anesthetic used: implications on analysis of mouse fMRI data. Neuroimage 116, 40–49. https://doi.org/10.1016/j.neuroimage.2015.05.013.

Schmid, F., Tsai, P.S., Kleinfeld, D., Jenny, P., Weber, B., 2017. Depth-dependent flow and pressure characteristics in cortical microvascular networks. PLOS Comput. Biol. 13, e1005392 https://doi.org/10.1371/journal.pcbi.1005392.

Schmid, F., Barrett, M.J.P., Jenny, P., Weber, B., 2019. Vascular density and distribution in neocortex. Neuroimage 197, 792–805. https://doi.org/10.1016/j.neuroimage.2017.06.046.

Schneider, K.A., Richter, M.C., Kastner, S., 2004. Retinotopic organization and functional subdivisions of the human lateral geniculate nucleus: a high-resolution functional magnetic resonance imaging study. J. Neurosci. 24, 8975–8985. https://doi.org/10.1523/JNEUROSCI.2413-04.2004.

Schölvinck, M.L., Maier, A., Ye, F.Q., Duyn, J.H., Leopold, D.A., 2010. Neural basis of global resting-state fMRI activity. Proc. Natl. Acad. Sci. U. S. A. 107, 10238–10243. https://doi.org/10.1073/pnas.0913110107.

Sclocco, R., Beissner, F., Bianciardi, M., Polimeni, J.R., Napadow, V., 2018. Challenges and opportunities for brainstem neuroimaging with ultrahigh field MRI. Neuroimage 168, 412–426. https://doi.org/10.1016/j.neuroimage.2017.02.052.

Secomb, T.W., Hsu, R., Park, E.Y.H., Dewhirst, M.W., 2004. Green's function methods for analysis of oxygen delivery to tissue by microvascular networks. Ann. Biomed. Eng. 32, 1519–1529. https://doi.org/10.1114/B:ABME.0000049036.08817.44.

Segal, S.S., Duling, B.R., 1986. Flow control among microvessels coordinated by intercellular conduction. Science (80-.) 234, 868–870. https://doi.org/10.1126/science.3775368.

Segal, S.S., Damon, D.N., Duling, B.R., 1989. Propagation of vasomotor responses coordinates arteriolar resistances. Am. J. Physiol. 256, H832–7.

Sereno, M., Dale, A., Reppas, J., Kwong, K., Belliveau, J.W., Brady, T., Rosen, B.R., Tootell, R.B.H., 1995. Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. Science (80-.) 268, 889.

Setsompop, K., Gagoski, B.A., Polimeni, J.R., Witzel, T., Wedeen, V.J., Wald, L.L., 2012. Blipped-controlled aliasing in parallel imaging for simultaneous multislice echo planar imaging with reduced g-factor penalty. Magn. Reson. Med. 67, 1210–1224. https://doi.org/10.1002/mrm.23097.

Setsompop, K., Feinberg, D.A., Polimeni, J.R., 2016. Rapid brain MRI acquisition techniques at ultra-high fields. NMR Biomed. 29, 1198–1221. https://doi.org/10.1002/nbm.3478.

Shannon, B.J., Dosenbach, R.A., Su, Y., Vlassenko, A.G., Larson-Prior, L.J., Nolan, T.S., Snyder, A.Z., Raichle, M.E., 2013. Morning-evening variation in human brain metabolism and memory circuits. J. Neurophysiol. 109, 1444–1456. https://doi.org/10.1152/jn.00651.2012.

Sharpee, T.O., Sugihara, H., Kurgansky, A.V., Rebrik, S.P., Stryker, M.P., Miller, K.D., 2006. Adaptive filtering enhances information transmission in visual cortex. Nature 439, 936–942. https://doi.org/10.1038/nature04519.

Sheth, S.A., Nemoto, M., Guiou, M., Walker, M., Pouratian, N., Hageman, N., Toga, A.W., 2004a. Columnar specificity of microvascular oxygenation and volume responses: implications for functional brain mapping. J. Neurosci. 24, 634–641. https://doi.org/10.1523/JNEUROSCI.4526-03.2004.

Sheth, S.A., Nemoto, M., Guiou, M., Walker, M., Pouratian, N., Toga, A.W., 2004b. Linear and nonlinear relationships between neuronal activity, oxygen metabolism, and hemodynamic responses. Neuron 42, 347–355.

Shmuel, A., Yacoub, E., Chaimow, D., Logothetis, N.K., Uğurbil, K., 2007. Spatio-temporal point-spread function of fMRI signal in human gray matter at 7 Tesla. Neuroimage 35, 539–552. https://doi.org/10.1016/j.neuroimage.2006.12.030.

Shokri-Kojori, E., Tomasi, D., Alipanahi, B., Wiers, C.E., Wang, G.J., Volkow, N.D., 2019. Correspondence between cerebral glucose metabolism and BOLD reveals relative power and cost in human brain. Nat. Commun. 10, 690. https://doi.org/10.1038/s41467-019-08546-x.

Silva, A.C., Koretsky, A.P., 2002. Laminar specificity of functional MRI onset times during somatosensory stimulation in rat. Proc. Natl. Acad. Sci. U. S. A. 99, 15182–15187. https://doi.org/10.1073/pnas.222561899.

Silva, A.C., Koretsky, A.P., Duyn, J.H., 2007. Functional MRI impulse response for BOLD and CBV contrast in rat somatosensory cortex. Magn. Reson. Med. 57, 1110–1118. https://doi.org/10.1002/mrm.21246.

Simon, A.B., Buxton, R.B., 2015. Understanding the dynamic relationship between cerebral blood flow and the BOLD signal: implications for quantitative functional MRI. Neuroimage 116, 158–167. https://doi.org/10.1016/j.neuroimage.2015.03.080.

Simony, E., Honey, C.J., Chen, J., Lositsky, O., Yeshurun, Y., Wiesel, A., Hasson, U., 2016. Dynamic reconfiguration of the default mode network during narrative comprehension. Nat. Commun. 7 https://doi.org/10.1038/ncomms12141.

Sirotin, Y.B., Hillman, E.M.C., Bordier, C., Das, A., 2009. Spatiotemporal precision and hemodynamic mechanism of optical point spreads in alert primates. Proc. Natl. Acad. Sci. U. S. A. 106, 18390–18395. https://doi.org/10.1073/pnas.0905509106.

Smith, S.M., Bandettini, P.A., Miller, K.L., Behrens, T.E.J., Friston, K.J., David, O., Liu, T., Woolrich, M.W., Nichols, T.E., 2012a. The danger of systematic bias in group-level FMRI-lag-based causality estimation. Neuroimage 59, 1228–1229. https://doi.org/10.1016/j.neuroimage.2011.08.015.

Smith, S.M., Miller, K.L., Moeller, S., Xu, J., Auerbach, E.J., Woolrich, M.W., Beckmann, C.F., Jenkinson, M., Andersson, J., Glasser, M.F., Van Essen, D.C., Feinberg, D.A., Yacoub, E.S., Uğurbil, K., 2012b. Temporally-independent functional modes of spontaneous brain activity. Proc. Natl. Acad. Sci. U. S. A. 109, 3131–3136. https://doi.org/10.1073/pnas.1121329109.

Smith, A.F., Doyeux, V., Berg, M., Peyrounette, M., Haft-Javaherian, M., Larue, A.E., Slater, J.H., Lauwers, F., Blinder, P., Tsai, P., Kleinfeld, D., Schaffer, C.B., Nishimura, N., Davit, Y., Lorthois, S., 2019. Brain capillary networks across species: a few simple organizational requirements are sufficient to reproduce both structure and function. Front. Physiol. 10, 233. https://doi.org/10.3389/fphys.2019.00233.

Soltysik, D.A., Peck, K.K., White, K.D., Crosson, B., Briggs, R.W., 2004. Comparison of hemodynamic response nonlinearity across primary cortical areas. Neuroimage 22, 1117–1127. https://doi.org/10.1016/j.neuroimage.2004.03.024.

Soon, C.S., Venkatraman, V., Chee, M.W.L., 2003. Stimulus repetition and hemodynamic response refractoriness in event-related fMRI. Hum. Brain Mapp. 20, 1–12. https://doi.org/10.1002/hbm.10122.

Stephan, K.E., Roebroeck, A., 2012. A short history of causal modeling of fMRI data. Neuroimage. https://doi.org/10.1016/j.neuroimage.2012.01.034.

Steriade, M., Nunez, A., Amzica, F., 1993. A novel slow (< 1 Hz) oscillation of neocortical neurons in vivo: depolarizing and hyperpolarizing components. J. Neurosci. 13, 3252–3265. https://doi.org/10.1523/jneurosci.13-08-03252.1993.

Stigliani, A., Jeska, B., Grill-Spector, K., 2017. Encoding model of temporal processing in human visual cortex. Proc. Natl. Acad. Sci. U. S. A. 114, E11047–E11056. https://doi.org/10.1073/pnas.1704877114.

Su, F.C., Chu, T.C., Wai, Y.Y., Wan, Y.L., Liu, H.L., 2004. Temporal resolving power of perfusion- and BOLD-based event-related functional MRI. Med. Phys. 31, 154–160. https://doi.org/10.1118/1.1634480.

Tak, S., Wang, D.J.J., Polimeni, J.R., Yan, L., Chen, J.J., 2014. Dynamic and static contributions of the cerebrovasculature to the resting-state BOLD signal. Neuroimage 84, 672–680. https://doi.org/10.1016/j.neuroimage.2013.09.057.

Tak, S., Polimeni, J.R., Wang, D.J.J., Yan, L., Chen, J., 2015. Associations of resting-state fMRI functional connectivity with flow-BOLD coupling and regional vasculature. Brain Connect. 5, 137–146. https://doi.org/10.1089/brain.2014.0299.

Thomas, C.G., Menon, R.S., 1998. Amplitude response and stimulus presentation frequency response of human primary visual cortex using BOLD EPI at 4 T. Magn. Reson. Med. 40, 203–209. https://doi.org/10.1002/mrm.1910400206.

Thompson, S.K., Engel, S.A., Olman, C.A., 2014. Larger neural responses produce BOLD signals that begin earlier in time. Front. Neurosci. 8 https://doi.org/10.3389/fnins.2014.00159.

Tian, P., Teng, I.C., May, L.D., Kurz, R., Lu, K., Scadeng, M., Hillman, E.M.C., De Crespigny, A.J., D'Arceuil, H.E., Mandeville, J.B., Marota, J.J.A., Rosen, B.R., Liu, T. T., Boas, D.A., Buxton, R.B., Dale, A.M., Devor, A., 2010. Cortical depth-specific microvascular dilation underlies laminar differences in blood oxygenation level-dependent functional MRI signal. Proc. Natl. Acad. Sci. U. S. A. 107, 15246–15251. https://doi.org/10.1073/pnas.1006735107.

Troprès, I., Grimault, S., Vaeth, A., Grillon, E., Julien, C., Payen, J.F., Lamalle, L., Décorps, M., 2001. Vessel size imaging. Magn. Reson. Med. 45, 397–408.

Tsai, P.S., Kaufhold, J.P., Blinder, P., Friedman, B., Drew, P.J., Karten, H.J., Lyden, P.D., Kleinfeld, D., 2009. Correlations of neuronal and microvascular densities in murine cortex revealed by direct counting and colocalization of nuclei and vessels. J. Neurosci. 29, 14553–14570. https://doi.org/10.1523/JNEUROSCI.3287-09.2009.

Turner, R., 2002. How much cortex can a vein drain? Downstream dilution of activation-related cerebral blood oxygenation changes. Neuroimage 16, 1062–1067. https://doi.org/10.1006/nimg.2002.1082.

Uğurbil, K., Xu, J., Auerbach, E.J., Moeller, S., Vu, A.T., Duarte-Carvajalino, J.M., Lenglet, C., Wu, X., Schmitter, S., Van de Moortele, P.-F., Strupp, J., Sapiro, G., De Martino, F., Wang, D., Harel, N., Garwood, M., Chen, L., Feinberg, D.A., Smith, S.M., Miller, K.L., Sotiropoulos, S.N., Jbabdi, S., Andersson, J.L.R., Behrens, T.E.J., Glasser, M.F., Van Essen, D.C., Yacoub, E., 2013. Pushing spatial and temporal resolution for functional and diffusion MRI in the Human Connectome Project. Neuroimage 80, 80–104. https://doi.org/10.1016/j.neuroimage.2013.05.012.

Uhlirova, H., Kılıç, K., Tian, P., Thunemann, M., Desjardins, M., Saisan, P.A., Sakadžić, S., Ness, T.V., Mateo, C., Cheng, Q., Weldy, K.L., Razoux, F., Vandenberghe, M., Cremonesi, J.A., Ferri, C.G., Nizar, K., Sridhar, V.B., Steed, T.C., Abashin, M., Fainman, Y., Masliah, E., Djurovic, S., Andreassen, O.A., Silva, G.A., Boas, D.A., Kleinfeld, D., Buxton, R.B., Einevoll, G.T., Dale, A.M., Devor, A., 2016. Cell type specificity of neurovascular coupling in cerebral cortex. Elife 5. https://doi.org/10.7554/eLife.14315.

Uludağ, K., 2008. Transient and sustained BOLD responses to sustained visual stimulation. Magn. Reson. Imaging 26, 863–869. https://doi.org/10.1016/j.mri.2008.01.049.

Uludağ, K., 2010. To dip or not to dip: reconciling optical imaging and fMRI data. Proc. Natl. Acad. Sci. U. S. A. 107 https://doi.org/10.1073/pnas.0914194107. E23; author reply E24.

Uludağ, K., Blinder, P., 2018. Linking brain vascular physiology to hemodynamic response in ultra-high field MRI. Neuroimage 168, 279–295. https://doi.org/10.1016/j.neuroimage.2017.02.063.

Uludağ, K., Dubowitz, D.J., Yoder, E.J., Restom, K., Liu, T.T., Buxton, R.B., 2004. Coupling of cerebral blood flow and oxygen consumption during physiological activation and deactivation measured with fMRI. Neuroimage 23, 148–155. https://doi.org/10.1016/j.neuroimage.2004.05.013.

Uludağ, K., Müller-Bierl, B., Uğurbil, K., 2009. An integrative model for neuronal activity-induced signal changes for gradient and spin echo functional imaging. Neuroimage 48, 150–165. https://doi.org/10.1016/j.neuroimage.2009.05.051.

Vakorin, V.A., Krakovska, O.O., Borowsky, R., Sarty, G.E., 2007. Inferring neural activity from BOLD signals through nonlinear optimization. Neuroimage 38, 248–260. https://doi.org/10.1016/j.neuroimage.2007.06.033.

Valdes-Sosa, P.A., Roebroeck, A., Daunizeau, J., Friston, K., 2011. Effective connectivity: influence, causality and biophysical modeling. Neuroimage. https://doi.org/10.1016/j.neuroimage.2011.03.058.

Van De Ville, D., Britz, J., Michel, C.M., 2010. EEG microstate sequences in healthy humans at rest reveal scale-free dynamics. Proc. Natl. Acad. Sci. U. S. A. 107, 18179–18184. https://doi.org/10.1073/pnas.1007841107.

van Dijk, J.A., Fracasso, A., Petridou, N., Dumoulin, S.O., 2020. Linear systems analysis for laminar fMRI: evaluating BOLD amplitude scaling for luminance contrast manipulations. Sci. Rep. 10 https://doi.org/10.1038/s41598-020-62165-x.

van Raaij, M.E., Lindvere, L., Dorr, A., He, J., Sahota, B., Foster, F.S., Stefanovic, B., 2012. Quantification of blood flow and volume in arterioles and venules of the rat cerebral cortex using functional micro-ultrasound. Neuroimage 63, 1030–1037. https://doi.org/10.1016/j.neuroimage.2012.07.054.

Vazquez, A.L., Noll, D.C., 1998. Nonlinear aspects of the BOLD response in functional MRI. Neuroimage 7, 108–118. https://doi.org/10.1006/nimg.1997.0316.

Vazquez, A.L., Cohen, E.R., Gulani, V., Hernandez-Garcia, L., Zheng, Y., Lee, G.R., Kim, S.-G., Grotberg, J.B., Noll, D.C., 2006. Vascular dynamics and BOLD fMRI: CBF level effects and analysis considerations. Neuroimage 32, 1642–1655. https://doi.org/10.1016/j.neuroimage.2006.04.195.

Vidaurre, D., Quinn, A.J., Baker, A.P., Dupret, D., Tejero-Cantero, A., Woolrich, M.W., 2016. Spectrally resolved fast transient brain states in electrophysiological data. Neuroimage 126, 81–95. https://doi.org/10.1016/j.neuroimage.2015.11.047.

Viessmann, O., Chen, J.E., Setsompop, K., Wald, L.L., Polimeni, J.R., 2019a. Variability, BOLD temporal SNR bias and variance across the HCP population as a function of cortical B0-orientation and orientation. Proc. Intl. Soc. Mag. Reson. Med. 27, 0369.

Viessmann, O., Scheffler, K., Bianciardi, M., Wald, L.L., Polimeni, J.R., 2019b. Dependence of resting-state fMRI fluctuation amplitudes on cerebral cortical orientation relative to the direction of B0 and anatomical axes. Neuroimage 196, 337–350. https://doi.org/10.1016/j.neuroimage.2019.04.036.

Villien, M., Wey, H.-Y., Mandeville, J.B., Catana, C., Polimeni, J.R., Sander, C.Y., Zürcher, N.R., Chonde, D.B., Fowler, J.S., Rosen, B.R., Hooker, J.M., 2014. Dynamic functional imaging of brain glucose utilization using fPET-FDG. Neuroimage 100. https://doi.org/10.1016/j.neuroimage.2014.06.025.

Vizioli, L., Yacoub, E., 2018. Probing temporal information in fast-TR fMRI data during attention modulations. Proc. Intl. Soc. Mag. Reson. Med. 26, 0153.

Vizioli, L., Bratch, A., Lao, J., Uğurbil, K., Muckli, L., Yacoub, E., 2018. Temporal multivariate pattern analysis (tMVPA): a single trial approach exploring the temporal dynamics of the BOLD signal. J. Neurosci. Methods 308, 74–87. https://doi.org/10.1016/j.jneumeth.2018.06.029.

Vu, A.T., Phillips, J.S., Kay, K., Phillips, M.E., Johnson, M.R., Shinkareva, S.V., Tubridy, S., Millin, R., Grossman, M., Gureckis, T., Bhattacharyya, R., Yacoub, E., 2016. Using precise word timing information improves decoding accuracy in a multiband-accelerated multimodal reading experiment. Cogn. Neuropsychol. 33, 265–275. https://doi.org/10.1080/02643294.2016.1195343.

Wager, T.D., Vazquez, A., Hernandez, L., Noll, D.C., 2005. Accounting for nonlinear BOLD effects in fMRI: parameter estimates and a model for prediction in rapid event-related studies. Neuroimage 25, 206–218. https://doi.org/10.1016/j.neuroimage.2004.11.008.

Wark, B., Lundstrom, B.N., Fairhall, A., 2007. Sensory adaptation. Curr. Opin. Neurobiol. https://doi.org/10.1016/j.conb.2007.07.001.

Wei, H.S., Kang, H., Rasheed, I.-Y.D., Zhou, S., Lou, N., Gershteyn, A., McConnell, E.D., Wang, Y., Richardson, K.E., Palmer, A.F., Xu, C., Wan, J., Nedergaard, M., 2016. Erythrocytes are oxygen-sensing regulators of the cerebral microcirculation. Neuron 91, 851–862. https://doi.org/10.1016/j.neuron.2016.07.016.

Wenger, K.K., Visscher, K.M., Miezin, F.M., Petersen, S.E., Schlaggar, B.L., 2004. Comparison of sustained and transient activity in children and adults using a mixed blocked/event-related fMRI design. Neuroimage 22, 975–985. https://doi.org/10.1016/j.neuroimage.2004.02.028.

Wittkuhn, L., Schuck, N.W., 2020. Faster than thought: detecting sub-second activation sequences with sequential fMRI pattern analysis. bioRxiv. https://doi.org/10.1101/2020.02.15.950667, 2020.02.15.950667.

Woolrich, M.W., Ripley, B.D., Brady, M., Smith, S.M., 2001. Temporal autocorrelation in univariate linear modeling of FMRI data. Neuroimage 14, 1370–1386. https://doi.org/10.1006/nimg.2001.0931.

Woolrich, M.W., Behrens, T.E.J., Smith, S.M., 2004. Constrained linear basis sets for HRF modelling using Variational Bayes. Neuroimage 21, 1748–1761. https://doi.org/10.1016/j.neuroimage.2003.12.024.

Xie, H., Chung, D.Y., Kura, S., Sugimoto, K., Aykan, S.A., Wu, Y., Sakadžić, S., Yaseen, M. A., Boas, D.A., Ayata, C., 2020. Differential effects of anesthetics on resting state functional connectivity in the mouse. J. Cereb. Blood Flow Metab. 40, 875–884. https://doi.org/10.1177/0271678X19847123.

Yablonskiy, D.A., Haacke, E.M., 1994. Theory of NMR signal behavior in magnetically inhomogeneous tissues: the static dephasing regime. Magn. Reson. Med. 32, 749–763. https://doi.org/10.1002/mrm.1910320610.

Yacoub, E., Wald, L.L., 2018. Pushing the spatio-temporal limits of MRI and fMRI. Neuroimage 164, 1–3. https://doi.org/10.1016/j.neuroimage.2017.11.034.

Yacoub, E., Shmuel, A., Pfeuffer, J., Van de Moortele, P.-F., Adriany, G., Uğurbil, K., Hu, X., 2001. Investigation of the initial dip in fMRI at 7 Tesla. NMR Biomed. 14, 408–412. https://doi.org/10.1002/nbm.715.

Yacoub, E., Uğurbil, K., Harel, N., 2006. The spatial dependence of the poststimulus undershoot as revealed by high-resolution BOLD- and CBV-weighted fMRI. J. Cereb. Blood Flow Metab. 26, 634–644. https://doi.org/10.1038/sj.jcbfm.9600239.

Yacoub, E., Shmuel, A., Logothetis, N.K., Uğurbil, K., 2007. Robust detection of ocular dominance columns in humans using Hahn Spin Echo BOLD functional MRI at 7 Tesla. Neuroimage 37, 1161–1177. https://doi.org/10.1016/j.neuroimage.2007.05.020.

Yacoub, E., Harel, N., Uğurbil, K., 2008. High-field fMRI unveils orientation columns in humans. Proc. Natl. Acad. Sci. U. S. A. 105, 10607–10612. https://doi.org/10.1073/pnas.0804110105.

Yang, Y., Engelien, W., Pan, H., Xu, S., Silbersweig, D.A., Stern, E., 2000. A CBF-based event-related brain activation paradigm: characterization of impulse-response function and comparison to BOLD. Neuroimage 12, 287–297. https://doi.org/10.1006/nimg.2000.0625.

Yang, J., Huber, L., Yu, Y., Chai, Y., Khojandi, A., Bandettini, P.A., 2019. High-resolution fMRI maps of columnar organization in human primary somatosensory cortex. Proc. Intl. Soc. Mag. Reson. Med. 27, 0617.

Yen, C.C.-C., Fukuda, M., Kim, S.-G., 2011. BOLD responses to different temporal frequency stimuli in the lateral geniculate nucleus and visual cortex: insights into the neural basis of fMRI. Neuroimage 58, 82–90. https://doi.org/10.1016/j.neuroimage.2011.06.022.

Yeşilyurt, B., Uğurbil, K., Uludağ, K., 2008. Dynamics and nonlinearities of the BOLD response at very short stimulus durations. Magn. Reson. Imaging 26, 853–862. https://doi.org/10.1016/j.mri.2008.01.008.

Yeşilyurt, B., Whittingstall, K., Uğurbil, K., Logothetis, N.K., Uludağ, K., 2010. Relationship of the BOLD signal with VEP for ultrashort duration visual stimuli (0.1 to 5 ms) in humans. J. Cereb. Blood Flow Metab. 30, 449–458. https://doi.org/10.1038/jcbfm.2009.224.

Yu, X., Glen, D., Wang, S., Dodd, S., Hirano, Y., Saad, Z.S., Reynolds, R., Silva, A.C., Koretsky, A.P., 2012. Direct imaging of macrovascular and microvascular contributions to BOLD fMRI in layers IV-V of the rat whisker-barrel cortex. Neuroimage 59, 1451–1460. https://doi.org/10.1016/j.neuroimage.2011.08.001.

Yu, X., Qian, C., Chen, D., Dodd, S.J., Koretsky, A.P., 2014. Deciphering laminar-specific neural inputs with line-scanning fMRI. Nat. Methods 11, 55–58. https://doi.org/10.1038/nmeth.2730.

Yuan, H., Zotev, V., Phillips, R., Drevets, W.C., Bodurka, J., 2012. Spatiotemporal dynamics of the brain at rest - exploring EEG microstates as electrophysiological signatures of BOLD resting state networks. Neuroimage 60, 2062–2072. https://doi.org/10.1016/j.neuroimage.2012.02.031.

Zhang, N., Zhu, X.-H., Chen, W., 2008. Investigating the source of BOLD nonlinearity in human visual cortex in response to paired visual stimuli. Neuroimage 43, 204–212. https://doi.org/10.1016/j.neuroimage.2008.06.033.

Zhang, N., Yacoub, E., Zhu, X.-H., Uğurbil, K., Chen, W., 2009. Linearity of blood-oxygenation-level dependent signal at microvasculature. Neuroimage 48, 313–318. https://doi.org/10.1016/j.neuroimage.2009.06.071.

Zheng, Y., Mayhew, J., 2009. A time-invariant visco-elastic windkessel model relating blood flow and blood volume. Neuroimage 47, 1371–1380. https://doi.org/10.1016/j.neuroimage.2009.04.022.

Zheng, Y., Johnston, D., Berwick, J., Chen, D., Billings, S., Mayhew, J., 2005. A three-compartment model of the hemodynamic response and oxygen delivery to brain. Neuroimage 28, 925–939. https://doi.org/10.1016/j.neuroimage.2005.06.042.

Zhou, J., Benson, N.C., Kay, K.N., Winawer, J., 2018. Compressive temporal summation in human visual cortex. J. Neurosci. 38, 691–709. https://doi.org/10.1523/JNEUROSCI.1724-17.2017.

Zhou, S., Giannetto, M., DeCourcey, J., Kang, H., Kang, N., Li, Y., Zheng, S., Zhao, H., Simmons, W.R., Wei, H.S., Bodine, D.M., Low, P.S., Nedergaard, M., Wan, J., 2019. Oxygen tension-mediated erythrocyte membrane interactions regulate cerebral capillary hyperemia. Sci. Adv. 5 https://doi.org/10.1126/sciadv.aaw4466.