The neural basis of functional brain imaging signals

David Attwell and Costantino Iadecola

The haemodynamic responses to neural activity that underlie the blood-oxygen-level-dependent (BOLD) signal used in functional magnetic resonance imaging (fMRI) of the brain are often assumed to be driven by energy use, particularly in presynaptic terminals or glia. However, recent work has suggested that most brain energy is used to power postsynaptic currents and action potentials rather than presynaptic or glial activity and, furthermore, that haemodynamic responses are driven by neurotransmitter-related signalling and not directly by the local energy needs of the brain. A firm understanding of the BOLD response will require investigation to be focussed on the neural signalling mechanisms controlling blood flow rather than on the locus of energy use.

The blood-oxygen-level-dependent (BOLD) magnetic resonance signal used in functional imaging of the brain reflects the loss of oxygen from haemoglobin, causing its iron to become paramagnetic, which influences the magnetic field experienced by protons in surrounding water molecules [1]. During neuronal activity, an increase of oxygen usage is followed within a few seconds by a larger fractional increase in blood flow and an increase in blood volume, resulting in a net decrease of the amount of deoxygenated haemoglobin present [2, 3]. This is the signal that is detected in most BOLD imaging experiments. Positron-emission tomography (PET) imaging can also detect activity-induced increases in blood flow. Understanding the increased blood flow associated with neural activity, a relationship first proposed 100 years ago [4–8], is thus fundamental to understanding functional imaging. It is often assumed that this increase of blood flow is a direct consequence of the oxygen consumption or energy usage of the tissue [9–15] – an idea that we assess in this article. We initially consider which cellular processes consuming most energy. Brain energy usage has been attributed mainly to activity in presynaptic terminals [15] or to the energy needed to take up the neurotransmitter glutamate and convert it to glutamine in astrocytes [17]. These ideas have been investigated by constructing an energy budget based on the measured properties of individual ion channels and synapses [18]. This analysis concluded that most of the energy used by grey matter on signalling is, in fact, expended on reversing the ion fluxes underlying excitatory postsynaptic currents and action potentials (Fig. 1). In rodents, 47% of the energy used on signalling was predicted to support action potentials and 34% to support postsynaptic responses (>33% for postsynaptic currents and <1% for metabotropic responses). It was estimated that 10% supports the resting potential of neurons, with only 5% being used in glia (3% for the resting potential and 2% for glutamate recycling) and 3% used in synaptic terminals (<3% for Ca2+ fluxes and ~0.5% for vesicle filling and recycling). In primates, the greater number of synapses per neuron results in postsynaptic responses being predicted to be the dominant energy consumer (accounting for 74% of the energy used on signalling) [18], consistent with the high proportion (62%) of mitochondria found in dendrites [19]. Action potentials were predicted to consume 10% of the signalling energy used in primates, the neuronal resting potential to use 2%, glia to use 6% (5% for recycling glutamate and 1% for the resting potential) and presynaptic terminals to use 7%. Thus, assuming that experiments verify these predictions, if the BOLD signal from primates did directly reflect energy consumption it would be dominated by excitatory postsynaptic currents. In the following section, we argue that BOLD signals are, indeed, coupled to postsynaptic events - but not via energy use.

Controversies over the locus of brain energy use

Energy used on action potentials

The 10–47% of signalling energy predicted to be expended on action potentials [18] contrasts with a classic estimate by Creutzfeldt [20]. This estimate was based on heat generation by action potentials in peripheral nerves and suggested that only 0.3–3.0% of energy used in the brain was needed to support action potentials in human and cat cortex. However, Creutzfeldt’s calculations employed a value of nerve heat production that was 6.6-fold lower than later measurements [21], he ignored the glial contribution to the nerve mass and, for cat cortex, he omitted a factor of ten for the action potential frequency. Redoing his calculation for human cortex, by dividing the heat that action potentials produce at 10 Hz (93 µcal.g−1 per action potential [21], assuming the same neural mass fraction as for peripheral nerve) by the total energy usage of grey matter (6870 µcal.g−1s−1, from a primate cortical glucose usage of 57 µmol per 100 g.min−1 [10]), suggests that
The neurotransmitter glutamate (Glu) is recycled through uptake into glia and conversion to glutamine (Gln). This was observed that rodent cortical glucose consumption is equal to the rate of glutamine formation from glutamate taken up into astrocytes [17] (Fig. 1). In the context of a model in which all ATP production in glia is glycolytic and glia then release lactate to neurons so that these can generate ATP by oxidative phosphorylation [12,13], it was concluded that the power needed for glutamate neurotransmitter cycling…accounts for >80% of total glucose oxidation [17]. This is a discrepancy with the prediction that only 2% of brain energy supports glutamate recycling in glia [18]. Sibson et al. [17] apparently considered only glycolytically derived ATP, overlooking both the 94% of ATP produced by oxidative phosphorylation in mitochondria and the fact that glia need energy for tasks other than glutamate recycling, such as maintaining the resting potential [23,24]. This resulted in a 15-fold over-estimate of the fraction of ATP expended on glial glutamate recycling.

Nevertheless, as most brain signalling energy is used by neurons, for postsynaptic currents and the action potentials that evoke transmitter release, one might expect the signalling energy used in a particular area to correlate with the (much smaller amount of) energy expended in the same area on recycling transmitter [18,25].

**Energy used on presynaptic terminals**

The prediction that only 3–7% of energy is used in presynaptic terminals [18] contrasts with the observation that 23% of mitochondria are in synaptic terminals [19], which might suggest that presynaptic processes consume significant amounts of energy [15]. However, synaptic terminals are only 5 μm apart along cortical axons [22], implying a diffusion time for small molecules of only 25 ms. Thus, mitochondria in synaptic terminals might support energy use by axons, by providing the ATP needed to reverse the ion influx that drives axonal action potentials, as well as supporting presynaptic terminal activity (Fig. 1).

**What controls cerebral blood flow?**

*Energy use does not directly increase blood flow*

The preceding discussion suggests that, in primate grey matter, energy use by neurons and synapses is not as steep as that down inhibitory synapses is not as steep as that down which Na⁺ moves postsynaptically at inhibitory synapses, implying that less energy expenditure is required to pump the ions back. Waldvogel et al. [27] reported that inhibition evokes no BOLD signal, and concluded that inhibitory synapses use less energy than excitatory synapses. However, in their experiments, inhibition was associated with the generation of fewer action potentials, so the change of energy use evoked by their stimuli does not reflect just inhibitory synaptic activity. In fact, increased glucose usage is associated with inhibition of hippocampal pyramidal and auditory cells [28,29], although it is not clear whether the energy is needed to drive inhibitory synaptic currents or to power the action potentials that generate inhibitory transmitter release.

---

**Fig. 1.** Schematic diagram of an excitatory neuron and a glial cell. The percentage of grey matter signalling energy predicted to be expended on different subcellular processes is shown for neurons (yellow) and glia (mauve) [18]. Predictions are given for rodents (first value) and primates (second value, calculated from rodent figures [18] by assuming a tenfold increase in synapse density per neuron with no other parameter changes). Na⁺/K⁺ pump symbols illustrate the fact that almost all signalling energy is expended on reversing Na⁺ and K⁺ fluxes across the cell membrane (or reversing Ca²⁺ fluxes with Na⁺/Ca²⁺ exchanger, which is indirectly powered by the Na⁺/K⁺ pump). Presynaptic boutons represent release sites both on axon collaterals near the soma and on parts of the axon far from the soma in the grey matter. Mitochondrion icons illustrate sites of ATP production by oxidative phosphorylation: mitochondria in presynaptic terminals can provide ATP to Na⁺/K⁺ pumps in the axon. The neurotransmitter glutamate (Glu) is recycled through uptake into glia and conversion to glutamine (Gln).
brain [9], this need not imply that oxygen or energy usage directly regulate the blood flow. Indeed, the available evidence suggests that oxygen and glucose delivery to the tissue is not maintained by oxygen or glucose lack increasing the CBF [30,31]. Neither can signals related to energy production, such as CO₂ and H⁺, account for the changes in CBF evoked by neural activity [32,33]. How, then, can CNS tissue maintain an energy supply adequate to meet its changing requirements? Interestingly, although the oxygen usage associated with neuronal activity must colocalize with the activity, the subsequent increase of blood flow occurs in a larger area [34], suggesting that blood flow is controlled by factors other than a lack of energy.

Control of local blood flow by fast neurotransmitters

In the cerebellar cortex, hippocampus and neocortex, there is increasing evidence that blood flow is controlled locally by glutamate and, perhaps, by GABA. In the cerebellar cortex, activating the excitatory parallel or climbing fibre inputs leads to an increased blood flow that is blocked by inhibitors of non-NMDA glutamate receptors, nitric oxide synthase (NOS) and adenosine receptors [35–37]. In addition, exogenous glutamate mimics the vascular effects of parallel fibre stimulation [38]. Similarly, in neocortex and hippocampus, exogenous glutamate or NMDA dilates pial arterioles and/or precapillary microvessels—an effect attenuated by NOS inhibitors [39–41]. The increase in CBF produced by activation of the somatosensory cortex is also attenuated by inhibition or genetic deletion of cyclooxygenase-2, an enzyme that synthesizes prostanooids from arachidonic acid and is associated with glutamatergic synapses [42,43]. These data suggest that a glutamate-evoked Ca²⁺ influx in postsynaptic neurons activates the production of NO, adenosine and arachidonic acid metabolites. These agents, in turn, produce a vasodilation reflecting both the activity of neurons presynaptic to the cells releasing NO, adenosine and prostanoids, and the level of depolarization of the postsynaptic cell (which will alter Mg²⁺-block of NMDA receptors and the resulting Ca²⁺ influx). Interestingly, in the cerebellum the blood vessel dilation produced locally by parallel fibre activation somehow propagates to more distant blood vessels [44], providing a possible explanation for the observation that blood flow increases over a larger area than that in which there is neuronal activity [34].

Activity-evoked rises in extracellular 
K⁺ concentration ([K]o) have also been shown to increase blood flow [45,46], but they have a less of an effect than neural activity. In addition, much of the effect of the [K]o increase could be K⁺-depoloarizing neurons and, thus, releasing glutamate or evoking a Ca²⁺ influx via voltage-gated channels, which leads to release of nitric oxide (NO) [47].

In the light of the controversy over whether inhibition contributes to brain energy use or to functional imaging signals, it is interesting that exogenous GABA, acting via GABA_A receptors, dilates precapillary microvessels in hippocampus and neocortex, and that blockade of GABA_A receptors produces a constriction [48]. However, in the cerebellar cortex, GABA is not involved in activation-induced increases in CBF [36,49].

Local blood flow and the BOLD signal do not correlate with principal neuron spike rate

In the cerebellar cortex, it is not possible to relate neurally evoked blood flow changes (and, therefore, BOLD signals) to an increase in firing of the principal output neurons, the Purkinje cells. Indeed, parallel fibre stimulation increases blood flow but inhibits Purkinje cell spiking [49,50]. Similarly, in visual cortex the BOLD signal correlates somewhat better with synaptically evoked field potentials than it does with the spiking output of the area [51]. It is essential to determine whether these results, which could reflect NOS being located both in excitatory and inhibitory neurons [52], extend to other brain areas. In the cerebellar cortex, mice with cyclin D2 knocked out, which decreases the number of inhibitory NOS-containing stellate cells, there is a reduction in the blood flow response to neural activity, suggesting that stellate cells play a significant role in controlling cerebellar blood flow [53]. Although knockout of cyclin D2 has effects
Global control of blood flow by dedicated intrinsic neural networks

In the cerebral cortex, dopamine, noradrenaline and serotonin are present in neurons innervating microvessels and produce vasoconstriction [54–57]. Furthermore, cholinergic axons from the basal forebrain have close associations with neocortical arterioles and capillaries, and can produce vasodilation [58,59]. Thus, in addition to the spatially restricted control of blood flow provided by fast transmitters, there might be a more widespread regulation by dopaminergic fibres from the mesencephalic ventral tegmental area, noradrenergic fibres from the locus coeruleus, serotonergic fibres from the raphe nucleus, and cholinergic fibres from the basal forebrain. These projections are diffuse, and could increase the blood flow to a wide cortical area when a particular type of task is being carried out. Furthermore, activation of some neural pathways, such as those passing through the cerebellar fastigial nucleus or originating in area C1 of the rostral ventrolateral medulla, leads to a general rise in CBF without a change of metabolism [60]. This is mediated, in part, by muscarinic ACh receptors and release of NO [60,61], and raises the possibility of a global change of CBF independent of local energy needs.

Functional imaging of neuropsychiatric conditions

The findings that amines could be involved in CBF control and that CBF can be dissociated from energy utilization have important implications for the interpretation of functional magnetic resonance imaging (fMRI) data in disease states. Functional imaging is increasingly being used to investigate brain function in conditions with altered amine function, such as schizophrenia, Parkinson’s disease, attention-deficit hyperactivity disorder or infusion of drugs affecting amine receptors or transporters [62–65]. However if dopamine, noradrenaline and serotonin modulate the blood flow changes that generate the BOLD signal, then in such conditions it might be hard to determine whether any observed alteration of the BOLD signal is due to an effect of altered amine function on cortical processing, or simply an alteration of the neural control of blood flow. In such work, control experiments will need to be devised to test how the amine disorder affects neural control of CBF, independently of affecting neural processing.

Conclusions

The available evidence indicates that the haemodynamic response to neural activity is not initiated by signals arising from the energy deficit of the tissue but, rather, is driven locally by fast glutamate-mediated signalling processes, and more globally by amine- and ACh-mediated neural systems. Accordingly, the BOLD effect used in functional brain imaging should be interpreted as a reflection of neuronal signalling and not as a locus of increased energy utilization. BOLD signals could, as in cerebellum, reflect the neural processing occurring within a brain area, rather than the output from or input to that area. Consequently, a change of processing with no net change of energy usage could lead to a BOLD signal – and a change of spiking output with no change to the signalling systems controlling blood flow could fail to generate a BOLD signal. Although sites of increased neural activity can often colocalize with areas of increased metabolism, CBF and energy utilization can be dissociated and should be considered as the results of processes operating in parallel. The view that the haemodynamic response is coupled to signalling processes represents a conceptual shift from the traditional idea that the energy demands of the tissue directly determine the flow increase associated with neural activation. In summary, we suggest that understanding the BOLD response is a signalling problem, not an energy problem.