3: The physiology of hemodynamics: Metabolism and circulation

In order to better understand the intricacies of cerebral hemodynamics and the mechanisms behind neurovascular coupling, one must begin by understanding the many functions of blood and the circulatory system in large multicellular organisms such as ourselves. Therefore Chapter 3 begins by discussing basic mammalian physiology, with an emphasis on the vascular system. Blood rheology, vascular hydrodynamics and gas exchange are discussed in Section 3.2. The concepts of global and local hemodynamics are introduced in Sections 3.3 and 3.4 respectively, and Section 3.4 discusses the control of local blood flow in great detail, with emphasis on our current understanding of the neural and humoral influences on vascular function. Sources of biogenic interference are then presented in Section 3.5, and the concept of a biogenic noise floor is introduced and explained. Section 3.6 concludes the chapter with a discussion on the effect of wavelength and optode spacing on image contrast, and how judicious selection of the optical source wavelengths might be useful in optimizing DOT signal quality.

3.1 Basic mammalian metabolic physiology

The energetics of metabolism: phosphorylation and glycolysis

In all mammalian cells, energy is transported through adenosine triphosphate (ATP). The hydrolytic removal of the terminal phosphate group releases about 7kcal per mole of ATP. This is completely reversible, and by adding the same amount of energy, the phosphate group can be reattached to a molecule of ADP through a condensation reaction to yield ATP. Unfortunately this catabolic reaction is only about 40% efficient in transferring energy to the cells, and the remaining 60% is dissipated as heat [50].

ATP can be formed through both aerobic and anaerobic pathways. The aerobic pathway is commonly referred to as **phosphorylation**: glucose + oxygen + ADP + P = CO_2 + ATP + H_2O. Phosphorylation can only function in the presence of oxygen. Normally this oxygen is taken from the air into the lungs. It is then transported by hemoglobin in the blood to the capillaries, where it diffuses from the blood, through cell membranes and cytosol, and into organelles known as mitochondria. Here special enzymes called cytochromes transfer energy from the oxidation of glucose into the phosphorylation of ADP to form ATP. Phosphorylation is efficient, but it requires a steady source of oxygen to function. The water is a metabolic bonus – it compensates for much of the water evaporatively lost through the lungs. [Some desert creatures, such as the kangaroo rat, are so efficient at water conservation that they can survive on metabolic water alone.] The CO_2 is very soluble in cytosol, blood plasma, and lipids, so it is easily excreted through the lungs [50].

The anaerobic pathway is referred to as **glycolysis**: glucose + ADP + P = lactic acid + ATP + H_2O. The advantage of glycolysis is that it can proceed in the absence of oxygen. However it is less efficient at producing ATP than phosphorylation, and it leads to the production of lactate, a large acidic molecule that cannot be excreted easily. If lactic acid accumulates, it can lead to metabolic acidosis [50].

Metabolic properties of the human brain

Table 3.1 describes the nominal metabolic requirements of the human brain. Typically about 95% of the CMRglu is oxidatively phosphorylated and only about 5% is consumed anaerobically and glycolyzed to lactate.

Table 3.1. Metabolic properties of the typical human brain [49].

PARAMETER	DEFINITION	VALUE	UNIT
$CMRO_2$	Cerebral metabolic rate of oxygen consumption	3.5 - 5	cc / 100g / minute
		50	cc / minute
CMRglu	Cerebral metabolic rate of glucose consumption	5 – 6.5	mg / 100g / minute
		70	mg / minute
CBV	Cerebral blood volume	3 - 5	percent by volume
CBF	Cerebral blood flow	50 - 60	cc / 100g / minute
		800	cc / minute
		14-20	% of cardiac output

Blood flow and gas exchange within the cardiovascular system

Blood is the main transport system within the body. It performs a number of vital functions: It supplies food, oxygen, and nutrients to tissues throughout the body and it carries away the metabolic waste products. It protects against pathogens, both foreign and domestic. It also transports hormones, enzymes, and ions throughout the body, and maintains a homeostatic environment for normal cell function.

Blood consists of a suspension of erythrocytes, leucocytes and platelets in plasma. Erythrocytes (RBCs) are about 7-9um in diameter. They are anuclear, and contain hemoglobin in solution within the cytosol. Since they contain no mitochondria, they obtain their metabolic energy through glycolysis. The outer membrane is very deformable, and healthy RBCs can deform enough to safely pass through 3um vessels. Their main function is oxygen transport, and the normal volume fraction of RBCs in blood (the hematocrit) is 40-45%. Below this value, the overall oxygen capacity of the blood is reduced, limiting oxidative metabolic processes. Above this value, the RBC fraction is too high for blood to pass smoothly through the microvasculature, and "sludging" of RBCs can occur, which compromises tissue perfusion [49, 50].

Leucocytes are between 10 and 20um diameter, and are stiffer than RBCs. Unlike RBCs, they are nuclear and capable of ameboid motion, which facilitates their movement through the vascular endothelium. Their main function is to provide a systemwide immunological defense. They are part of the humoral (i.e. circulatory) immune system, which defends the body against microbial pathogens and many toxins as well [50].

Platelets are disk shaped cells about 2-3um diameter. They are involved in thrombogenesis: the process by which leaks in the vascular system are both sealed and healed. The combined volume fraction of both leucocytes and platelets is less than 1%.

Plasma, the liquid portion of the blood, consists of a suspension of organic and inorganic substances. In addition to proteins, nutrients, and metabolic waste products, plasma also contains many inorganic electrolytes [50].

The cardiovascular circuit

The cardiovascular system is, from a hydraulic standpoint, a balanced system. This means that the blood volume must always equal the vascular volume. An imbalance in this relationship could lead to either hypertension or vascular collapse. A diagram depicting blood flow throughout the body is shown in Figure 3.1.

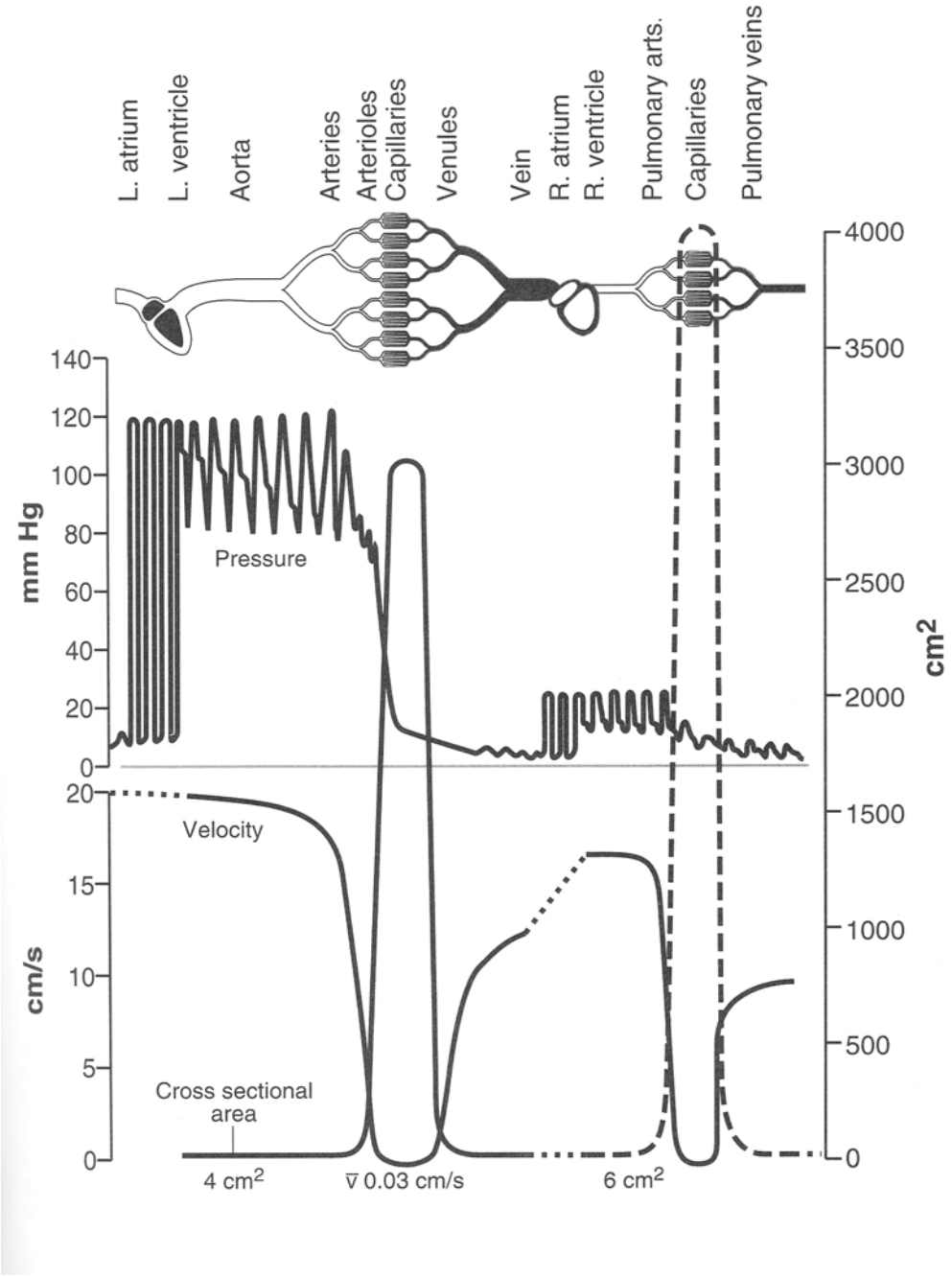


Figure 3.1. The relationship between pressure, flow, velocity, and cross-sectional area in the cardiovascular system. Most of the pressure drop occurs across the arterioles, which exert a significant influence over the blood flow through local regions of tissue within the parenchyma [50].

3.2 Blood rheology and vascular hydrodynamics

The typical viscosity of blood at normal shear rates is ~3-4 centipoise (cP). Although generally considered a Newtonian fluid, whole blood is actually pseudoplastic, and exhibits shear-thinning (a drop in viscosity with increasing shear rate), due mainly to RBC aggregation at low shear rates and

RBC deformation at higher shear rates. The hematocrit also has a strong effect on the viscosity of whole blood, as shown in Figure 3.2.

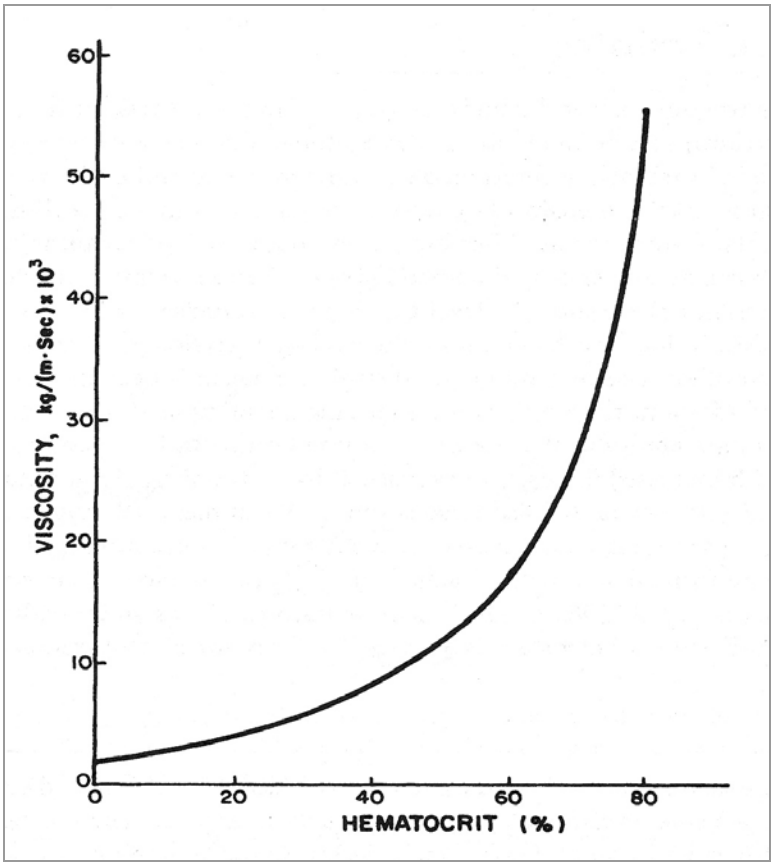


Figure 3.2. The effect of hematocrit on blood viscosity. As the hematocrit (the RBC volume fraction) increases above ~40%, the oxygen carrying capacity of the blood increases linearly, however viscosity increases supralinearly, impairing blood transport through the microvasculature and thus reducing the amount of oxygen actually delivered to the tissues [49].

The cross-sectional distribution of RBCs within the vasculature is not uniform. RBCs tend to travel in the center of the vessel, where the flow is fastest, and so they are jacketed by a thin (2-5um) layer of plasma in contact with the vascular endothelium [49]. As a consequence of this plasma sheathing layer, the microvascular hematocrit is effectively lower than the feed or discharge hematocrit, since the RBCs collectively travel faster on average than the blood plasma. This reduction in microvascular hematocrit is referred to as the Fahraeus Effect. The observation that the apparent viscosity of flowing blood within the vasculature is lower than the bulk viscosity (as measured with a rotational viscometer) is referred to as the Fahraeus-Lindqvist Effect [49].

Poisseuile's Law

The apparent viscosity η of a Newtonian fluid flowing in a laminar fashion through a cylindrical vessel can be defined to the first order using Poisseuile's Law:

$$\eta = \frac{\Pi \Delta P R^4}{8QL} \qquad 3.1$$

Where: η is the apparent viscosity,
$\Pi = 3.14159...$,
ΔP is the pressure difference along the length of the vessel,
R and L are the radius and length of the vascular lumen,
and Q is the volumetric flow rate.

The actual viscosity of blood varies with hematocrit, vessel radius, blood flow rate, temperature, and other parameters, but Poisseuile's Law provides a reasonable estimate of blood flow and vascular resistance for understanding cerebral hemodynamics as it pertains to DOT [49].

3.2.1 Microvascular structure

The microvasculature consists of four distinct tissue types, as shown in Figure 3.3. The innermost tissue in direct contact with the blood is called the intima, and consists of a single layer of contiguous endothelial cells. The adventitia surrounds the intima, and consists primarily of collagen fibers. The media is the muscular tissue surrounding the intima, and consists of one or more layers of smooth muscle cells, and the elastica lamina consists of fibers situated between the smooth muscle cells and the endothelium.

Arterioles are typically 10-150um in diameter and venules are typically 10-200um in diameter. The larger vessels contain all three layers, which gradually thin as the vessel size decreases. The smooth muscle and collagen fiber layers terminate at the capillaries, which consist of a single layer of endothelial cells.

Capillaries range from 4-8um in diameter. Their walls must be thin in order to allow dissolved gases, solutes, and other substances to diffuse between the blood and the extravascular fluid. As a result, capillaries are fragile vessels that are unable to sustain large transmural pressures [49].

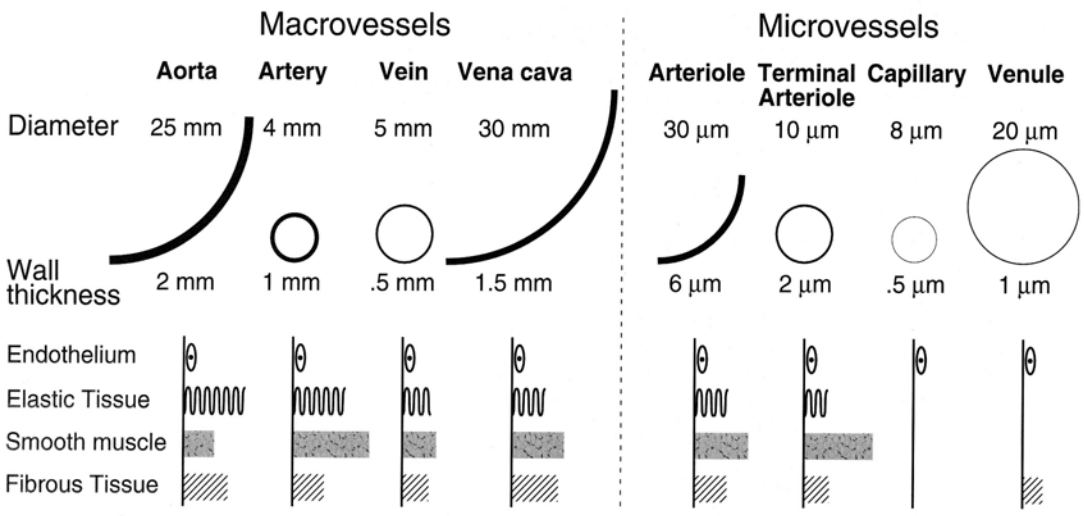


Figure 3.3. Tissue composition of the macro- and microvasculature. Arterioles are well endowed with smooth muscle and can significantly vary their luminal diameter to modulate blood flow. Venules are mostly devoid of smooth muscle fibers and respond passively to vascular pressure changes. Capillaries consist of only a single layer of endothelial cells, and are very permeable to gases, nonionic solutes and some ionic solutes as well [49].

3.2.2 Gas transport and diffusion

Oxygen transport

Oxygen is transported in two forms within the blood: chemically bound to hemoglobin molecules dissolved within the cytosol of RBCs, and physically dissolved in blood plasma. At sea level, whole blood can hold about 20cc of O_2 per deciliter, of which only 0.3cc is physically dissolved and the rest is chemically bound to hemoglobin. Although the amount of O_2 dissolved in plasma is a linear function of oxygen tension, the relationship between hemoglobin oxygen saturation and oxygen tension is nonlinear, and varies in a physiologically significant manner as a function of pH, CO_2 tension, and temperature as shown in Figure 3.4 [50].

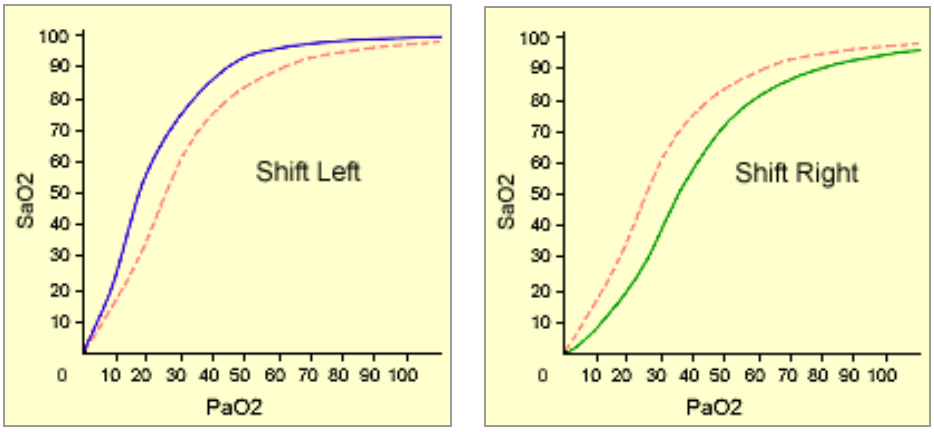


Figure 3.4. Hemoglobin oxygen saturation curves, showing hemoglobin saturation in percent vs. oxygen tension in Torr. Increases in pH and decreases in CO_2 tension and temperature shift the normal saturation curve (dashed red line) to the left, facilitating the uptake of oxygen. Decreases in pH and increases in CO_2 tension and temperature shift the curve to the right, enhancing oxygen unloading and delivery to the tissues.

Carbon dioxide transport

Unlike O_2, CO_2 is transported in a number of different forms within the blood. Most of the CO_2 (about 70%) is in the form of bicarbonate anions, mostly contained within the RBCs. A compensatory inward shift of chloride ions occurs to reestablish charge balance. The pH buffering ability of hemoglobin assists in this process. About 10% is transported in physical solution, dissolved in blood plasma. Amino acids and aliphatic (straight-chain) amines react with the CO_2 to form carbamino compounds, primarily with hemoglobin. About 20% of CO_2 is transported in this form. These mechanisms are depicted in Figure 3.5 [50].

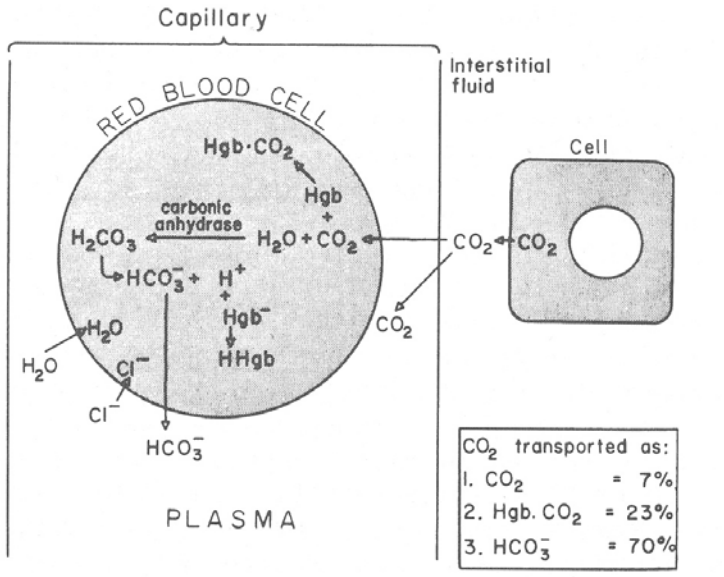


Figure 3.5. Carbon dioxide transport in the blood. About 70% is in the form of bicarbonate anions, ~23% is bound in carbamino form to hemoglobin, and ~7% is physically dissolved in blood plasma [50].

Gas diffusion

One of the most important functions of the blood is the delivery and removal of the two gases most critical to metabolic function: oxygen and carbon dioxide. From within the capillaries oxygen must diffuse from the RBCs through the intravascular plasma sheath through the endothelial cells of the intima and into the extravascular space, then through the parenchymal cell walls and cytosol to ultimately reach the mitochondria where oxidative metabolism occurs. CO_2 takes essentially the same route in reverse.

The first practical model of oxygen transport was proposed by August Krogh in 1918, and the Krogh-Erlang equation describes the radial variation of O_2 tension in tissue:

$$\mathrm{Pti}_{O_2} = \mathrm{Pc}_{O_2} + \left(\frac{M}{8D_{ti}\alpha_{ti}}\right) * \left[3R^2 - r^2 - \left(\frac{4R^4}{R^2 - r^2}\right) * \log_e R/r\right] \qquad 3.2$$

Where: Pti_{O_2} is the tissue P_{O_2},
Pc_{O_2} is the end-capillary P_{O_2},
M is the O_2 consumption per unit volume of Krogh cylinder (not including the intracapillary volume),
D_{ti} is the O_2 diffusion coefficient,
α_{ti} is the O_2 solubility in the tissue,
R is the radius of the Krogh's cylinder,
and r is the capillary radius.

The Krogh-Erlang equation is based on the assumption that O_2 diffusion is adequate to supply the metabolic need of O_2 throughout the cylinder. It appears that a major source of the resistance in the O_2 transport chain, as much as 50%, arises from the limited diffusivity of oxygen through the thin plasma layer between RBCs and the vascular wall [49].

The Bohr and Haldane effects.

The affinity of hemoglobin for CO_2 and O_2 are inversely related: changes in the CO_2 tension affect the strength of O_2 binding, and hence the O_2 tension in the blood – and vice versa. These are referred to as the Bohr effect (CO_2 affecting O_2 tension) and the Haldane effect (O_2 affecting CO_2 tension).

The **Bohr effect** refers to the reduction in binding affinity of hemoglobin for O_2 as the result of a drop in cytosolic pH (through an increase in carbonic acid, formed locally by bicarbonate ions).

The **Haldane effect** refers to a complementary process, in which the release of O_2 converts oxyhemoglobin (a stronger acid) to deoxyhemoglobin (a weaker acid) which favors uptake of CO_2 through carbamino formation.

Recently it was discovered that arterioles and, to a lesser extent, venules also contribute to gas exchange, but to a smaller degree than capillaries. Since CO_2 is far more soluble in tissue and plasma than is O_2, it is thought that countercurrent exchange of CO_2 between venules and arterioles may play a significant role in CO_2 transport, and that equilibration of CO_2 tension between the blood and the surrounding tissue may occur before the blood even reaches the capillaries [49].

3.3 Global hemodynamics: Regulating blood pressure

Blood pressure must be maintained over a wide range of both internal and external conditions in order to maintain the ability to perfuse all of the tissues throughout the body. Intuitively, this is no different than maintaining water pressure throughout a large town. Many of the same issues apply:

- Consumers' needs vary with time of day and with their activity level
- Water is consumed at different elevations and in different amounts
- Pressure must be maintained despite water main ruptures, droughts, and other unforeseen events
- Waste water must be returned to the treatment facility primarily through the force of gravity, since the use of distributed return pumps is difficult and impractical

In order to address similar physiologic needs, mammals have evolved a complex feedback system equipped with features like feedforward compensation in order to adapt to short-term demands while still maintaining long-term barostasis. Baroreceptors in the carotid sinus and the aortic arch detect the instantaneous pressure and send this information to the brain. The brain then sends control signals through autonomic nerves to control the heart and vasculature to maintain a normotensive state, as depicted in Figure 3.6.

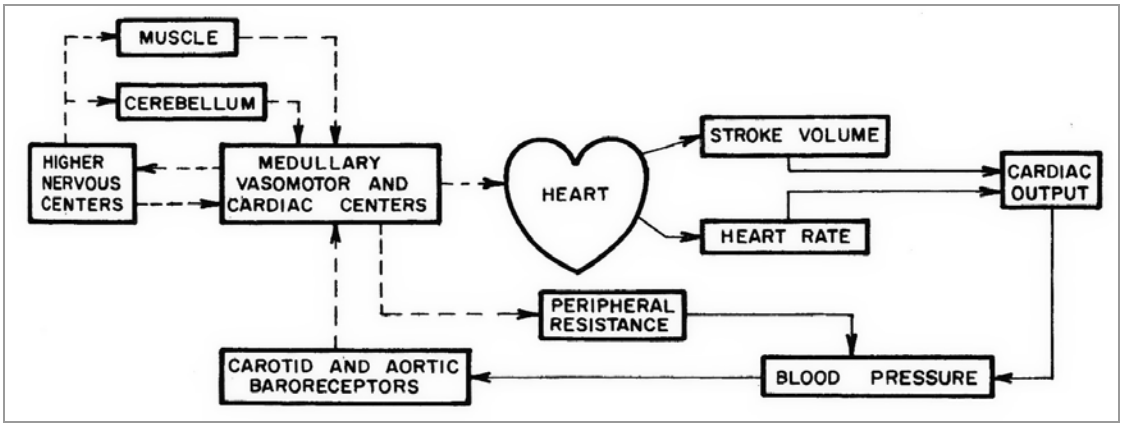


Figure 3.6. The general mechanisms responsible for systemic blood pressure regulation. Dashed lines indicate neural feedback and solid lines indicate direct mechanical effects [49].

Since moving fluids have inertia, they appear reactive from an impedance perspective. The systemic circulation can thus be modeled as a reactive load on the (current) output of the heart, and Ohm's Law can be used to model blood flow and organ perfusion: blood flow is directly proportional to blood pressure and inversely proportional to systemic vascular impedance. Under laminar flow conditions (which applies to the majority of the circulatory system), the vascular resistance of a large vessel segment can be modeled reasonably well by Poisseuile's Law (eq. 3.1, p.4), however the non-Newtonian rheology of blood becomes more important as the vessel size drops and the shear rate increases. According to Poisseuile's Law, vessel diameter exerts a fourth-order influence on vascular resistance. This suggests that a vessel diameter ratio of only 4:1 can modulate perfusion by a factor of a thousand or more - well in excess of the normal metabolic range for most tissues [50].

3.4 Local hemodynamics: Regulating local blood flow

Tissue perfusion is controlled locally throughout the body. Blood flow through most organs is lowest at rest and rises to meet the metabolic demands as activity rises. The brain and kidneys are exceptions to this rule. Cerebral blood flow remains relatively constant over a wide range of neuronal activity. Due to the unique role of the kidneys as a filtering organ, their metabolic rate is a direct function of the glomerular filtration rate. So for the kidneys, metabolism is a function of blood flow, unlike with most other organs [50].

Arteriolar vasomodulation and the "vascular filter" problem

Metabolic activity within tissue parenchyma generates waste products, most notably CO_2 and lactic acid, both of which are vasoactive. They interact with the vascular endothelium to generate nitric oxide (NO), a soluble and reactive gas. Nitric oxide diffuses through the endothelium to reach the vascular smooth muscle, which dilates in response. The subsequent increase in perfusion (due to the decrease in vascular resistance) sweeps away these metabolites, reducing NO production, and thus closing the feedback loop. An example of the vasodilatory temporal response to a change in cortical activation is shown in Figure 3.7 [55].

The problem, from a neuroimaging perspective, is that neuronal activation is convolved with the hemodynamic response in all vascular-based imaging modalities, including fMRI and DOT. Much effort is thus being spent in an effort to better understand the mechanisms behind neurovascular and neurometabolic coupling, with the hope that this will lead to better methods of deconvolving neural activity from the "vascular filter" created by this hemodynamic response [12, 55, 56].

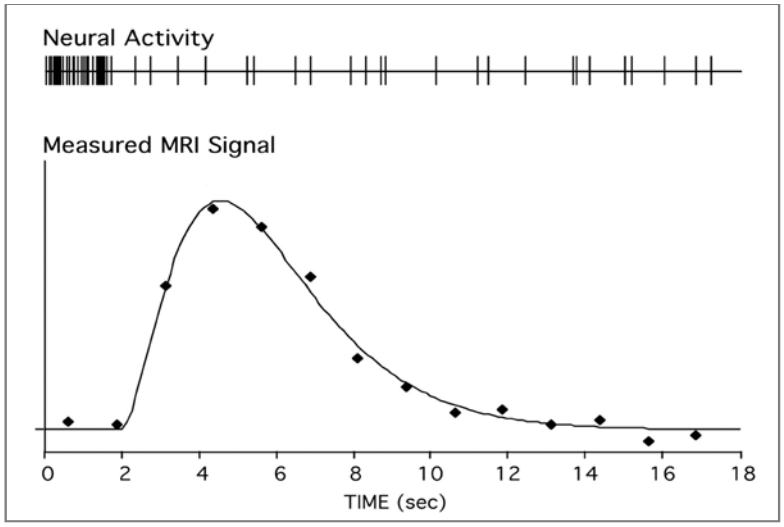


Figure 3.7. The vasodilatory temporal response to neuronal activation. There is a response latency of about 1-2 seconds and a FWHM of about 7 seconds, which varies slightly depending on the region of activation [55].

3.4.1 Vascular tone and the local control of blood flow

Arterioles and venules exhibit what is referred to as vascular tone – a basal level of smooth muscle activity that varies as a function of neural, thermal, hydrostatic, humoral, and likely many other stimuli. Vascular tone affects vessel diameter, which strongly affects blood flow [50].

Vascular compliance

Vascular compliance is defined as the change in vascular volume caused by an incremental change in transmural pressure. Elastance, a related term, is equal to the reciprocal of compliance. Vascular compliance varies significantly with transmural pressure, exhibiting three quasi-linear regimes, very much like a latex balloon. An example of a typical arterial pressure-volume curve is shown in Figure 3.8. Arteries operate near the region labeled "A," while veins can operate over the entire range below this point.

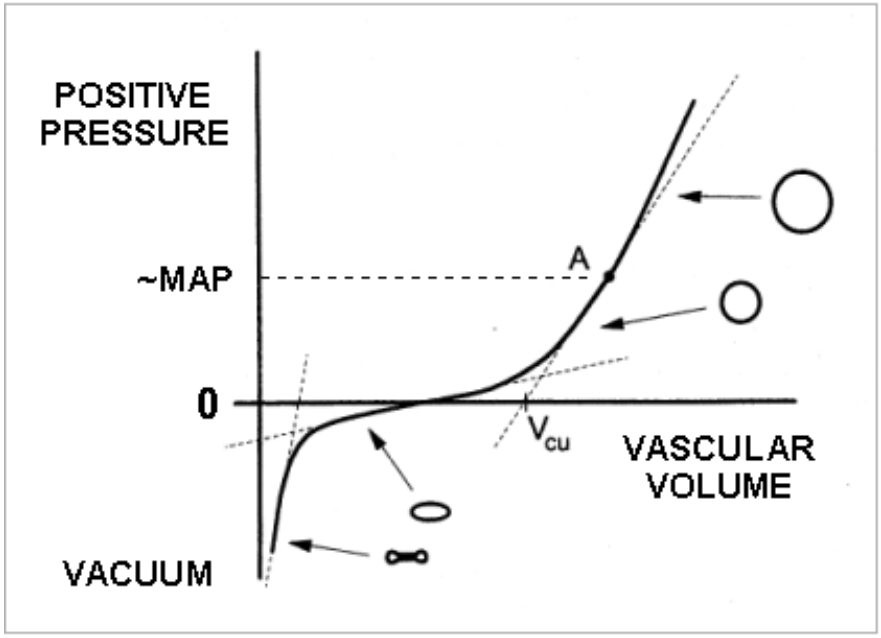


Figure 3.8. A typical static arterial pressure-volume curve showing the three compliance regimes. Below a certain pressure, the vessel is fully collapsed, and the compliance is very low. As the transmural pressure increases, the vessel begins to fill like a bladder. Since the overall circumference remains fixed, the compliance is very large. Once the vessel has filled fully, it begins to dilate. Further expansion forces the vessel to stretch, elongating both collagen and smooth muscle fibers, and resulting in a progressively decreasing compliance. The normal operating point for arteries is the mean arterial pressure (MAP), and is signified by the point "A" on the dilatory region of the P-V curve [49].

Viscoelasticity

Microvessels exhibit a time-dependent change in elastic properties, referred to as "viscoelastic" behavior. The initial reaction to a stepwise change in transluminal pressure is a passive elastic vasodilatory response followed by a slower myogenic compensatory response. This biphasic nonlinear response is seen following changes in transluminal pressure as well as changes in the metabolic activity of local parenchymal tissue [25, 57].

Myogenic response

Arterioles also exhibit a myogenic response: an increase in transmural pressure raises the vascular tone and thus reduces the lumen diameter. This provides a local autoregulatory function, actively stabilizing blood flow as the transluminal pressure varies [57].

Vasomotion

If viewed from the standpoint of feedback control, the phase shift introduced by the slow myogenic vascular response can erode the phase margin of the loop, leading to spontaneous oscillations in vessel diameter referred to as vasomotion or Mayer waves. These are discussed further below [57].

3.4.2 Neural and chemical mediation of vascular function

Autonomic neural influences

The autonomic nervous system has three primary functions. It regulates heart rate and cardiac output, it regulates most secretory (ducted, or exocrine) glands, and it also innervates smooth muscle (muscles of the bronchi, blood vessels, urogenital system, and gastrointestinal tract).

The autonomic nervous system consists of both sympathetic and parasympathetic pathways, each working in a complementary fashion to achieve balanced neural control. When an organ is innervated by both divisions of the autonomic nervous system, one pathway - either the sympathetic or the parasympathetic - provides most of the basal control. The pathway that controls basal function most of the time is said to provide the predominant tone to that organ. In most organs, the parasympathetic nervous system provides the predominant tone. The vascular system, which is regulated almost exclusively by the sympathetic nervous system, is an exception to this general rule [50].

Sympathetic pathways regulate the cardiovascular system. Stimulation of sympathetic nerves to the heart increases cardiac output. Sympathomimetics (drugs which increase the firing rate of sympathetic nerves) such as epinephrine and aminophylline are thus inotropic agents. Stimulation of sympathetic nerves to arterioles and veins causes vasoconstriction. Cutaneous vasoconstriction following strong sympathetic outflow hints at the origin of the term "white with fear." Regulation of body temperature is also sympathetically mediated [50].

The parasympathetic nervous system acts in an antagonistic manner at most sympathetically mediated sites. For example, increased parasympathetic activity, or the administration of parasympathomimetics such as physostigmine, result in a slowing of the heart rate and a reduction in ventricular myocardial contractility – a negative inotropic effect. Excess parasympathetic activity also constricts bronchial smooth muscle, a problem for asthmatics, hence the use of sympathomimetics such as aminophylline to counter this effect [50].

3.4.3 Humoral control of the vascular system

Cholinergic humoral influences

Cholinergic receptors mediate responses to acetylcholine. Drugs that activate cholinergic receptors on blood vessels (cholinomimetics) will cause vasodilation, which will in turn cause a drop in blood pressure. Muscarinic cholinergic receptors on blood vessels are not directly associated with the nervous system. No nerves terminate at vascular muscarinic receptors. It is not clear as to how, or even if, these receptors are ever activated physiologically. However, regardless of their physiologic relevance, muscarinic cholinergic drugs can induce vasodilation [50].

Adrenergic humoral influences

Adrenergic receptors mediate responses to epinephrine and norepinephrine. **Alpha1 receptors** are present in veins and on arterioles in many capillary beds. Activation of alpha 1 receptors in blood vessels produces vasoconstriction.

Beta1 receptors are located in the heart and kidney. Activation of these receptors increases heart rate, force of contraction (positive inotropism), and velocity of impulse conduction through the atrioventricular node. Activation of beta1 receptors in the kidney causes release of renin into the blood. Since renin promotes synthesis of angiotensin, a powerful vasoconstrictor, activation of renal beta1 receptors is a means by which the nervous system can elevate blood pressure. **Beta2 receptors** are located in arterioles of the heart, lungs and skeletal muscle. Beta2 receptor activation causes vasodilation (a complementary effect to that of alpha1 receptor activation).

Release of epinephrine from the adrenal medulla results in vasoconstriction in most vascular beds and vasodilation in certain others. Epinephrine then produces responses much like those elicited by stimulation of postganglionic sympathetic nerves [50].

Vasoactive metabolites

Most metabolites are either directly or indirectly vasoactive. pH/H_3O^+ and CO_2 can act directly on the vascular endothelium to elicit the release of endothelium-derived releasing factor (most likely the compound nitric oxide, NO), which then diffuses through the vessel wall and into the vascular smooth

muscle located nearby, resulting in profound vasodilation. The NO then either decomposes spontaneously or is scavenged by nitrolytic enzymes [57]. [This is why organic nitrates such as nitroglycerine and PETN ease the discomfort of angina pectoris – the nitrate moiety slowly hydrolyzes to form NO in the blood. The arteriolar vasodilation improves myocardial perfusion, thus relieving the symptoms. Viagra relieves certain forms of impotence in a complementary fashion. By blocking the nitrolytic enzymes, it prolongs the vasoactive effects of NO, thus facilitating the vasodilation required for normal sexual function.] The mechanism behind how various agonists affect vascular muscle tone is shown in Figure 3.9 [50].

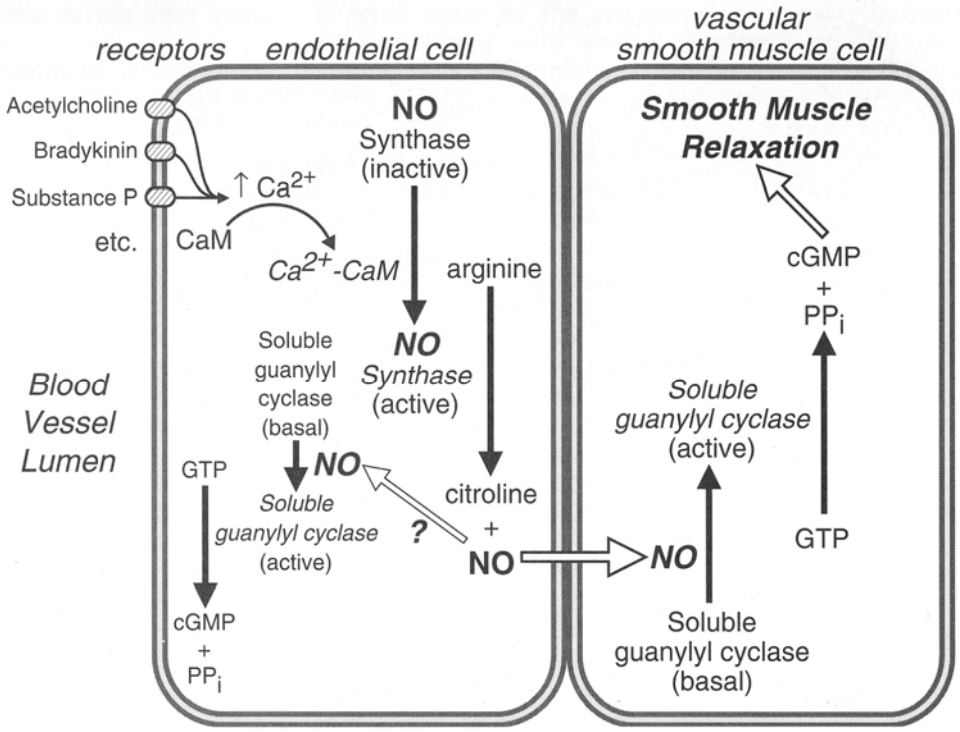


Figure 3.9. The effects of various agonists on vascular smooth muscle tone. Stimulation of receptors on the vascular endothelium leads to the formation of nitric oxide. Nitric oxide then diffuses through the endothelial wall and into the smooth muscle cells, where it modulates the formation of guanylyl cyclase, which facilitates the formation of cGMP and smooth muscle relaxation [50].

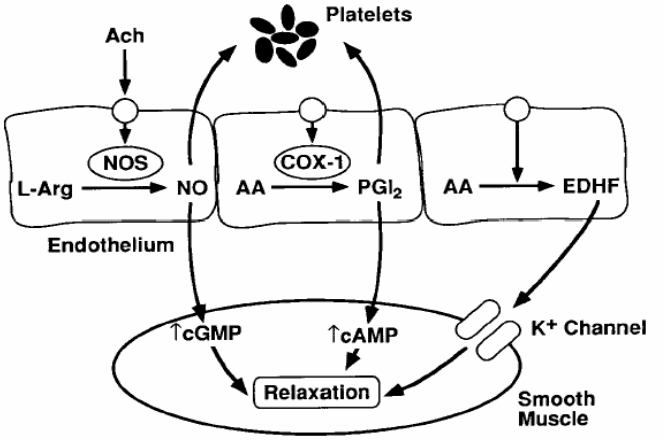


Figure 3.10. Some of the mechanisms of endothelium-dependent relaxation of cerebral vascular smooth muscle. NO is produced by NO synthase (NOS) from the amino acid L-arginine (L-Arg). NO diffuses into the vascular muscle, where it activates soluble guanylate cyclase, causing increased production of cyclic guanosine monophosphate (cGMP), which results in vascular relaxation. Prostacyclin (PGI_2) is normally produced by cyclooxygenase-1 (COX-1) from arachidonic acid (AA). PGI_2 diffuses to vascular muscle where it activates potassium ion (K^+) channels. Increased activity of potassium channels produces hyperpolarization that relaxes vascular muscle [57].

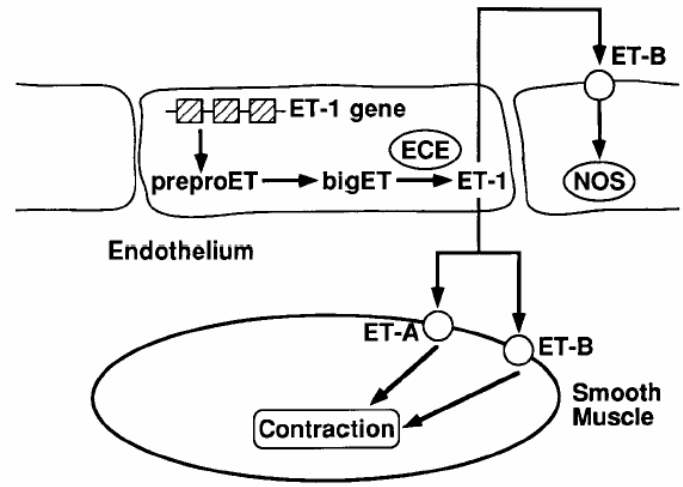


Figure 3.11. A diagram depicting the synthesis of endothelin-1 (ET-1) within the endothelium. ET-1 gene transcription results in the formation of preproET-1 mRNA. PreproET-1 is then converted to big ET-1 by an endopeptidase. Big ET-1 is then cleaved to ET-1 by endothelin-converting enzyme (ECE). ET-1 can activate both ET-A and ET-B receptors on the vascular muscle to cause contraction. ET-1 can also activate ET-B receptors on the endothelium, which causes activation of NO synthase and production of NO [57].

3.4.4 Neural, metabolic, and chemo-vascular interactions

Hb is constantly being generated from HbO_2 as a consequence of oxidative metabolism within the brain. A number of investigators have noted that increases in lactate production during neural stimulation were too small to account for the increase in metabolic activity. It was thus assumed that the dominant source of energy for neural activation was derived oxidatively, through phosphorylation.

Therefore neural activity and $CMRO_2$ should be closely linked. Ances observed that CBF and $CMRO_2$ both responded within a few seconds to the onset and termination of neural stimulation [27, 58]. Their time courses exhibited similar overshoots, however the steady state value for CBF was greater than expected for the needs of $CMRO_2$ alone. These observations suggest that CBF is coupled to $CMRO_2$, and thus CBF should be closely coupled to neural activity as well.

It turns out that changes in CBF do not scale with $CMRO_2$ changes in a linear stoichiometric fashion as one might expect [34, 42]. In fact CBF far exceeds $CMRO_2$, often by a factor of 3 or more, as depicted in Figure 3.12. It is this counterintuitive feature that actually generates the fMRI BOLD signal: increases in CBF act to dilute [Hb], despite the combined mitigating effects of increased $CMRO_2$ and rising intravascular volume [57, 59].

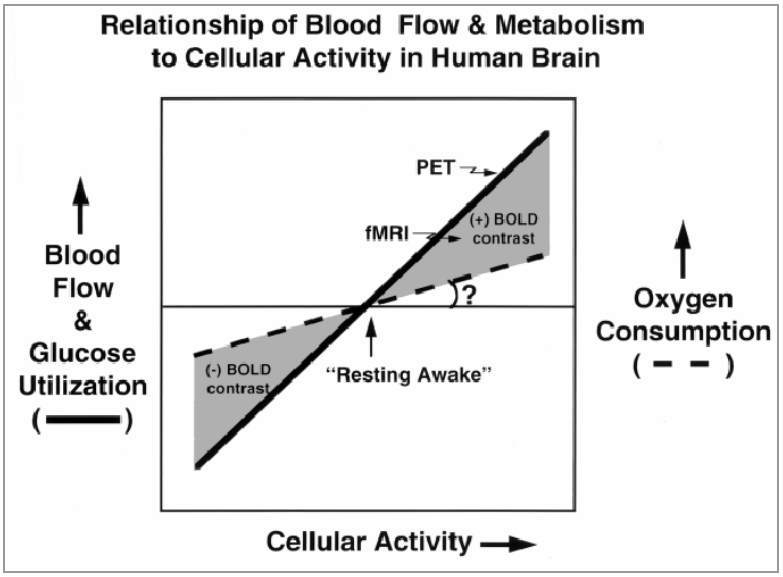


Figure 3.12. Local blood flow (CBF) exceeds local oxygen consumption ($CMRO_2$) in response to increased neural activity in the brain. This produces an increase in blood volume (CBV) and a *decrease* in oxyhemoglobin concentration ($[HbO_2]$), both of which are detectable with DOT and fMRI [34].

Generally accepted experimental observations

- Following cerebral activation, the increase in local CBF exceeds the increase in both CMR_{glu} and $CMRO_2$.
- The scaling between CBF and CMR_{glu} is nearly linear [34] while the scaling between CBF and $CMRO_2$ is supralinear [58].
- During continuous activation, the intracellular O_2 tension rises slowly to reach a plateau after about 20 seconds, and decays in similar fashion after cessation of stimulation [58].
- Regional CBV changes remain more localized than do regional [Hb] changes and BOLD signals [12].

These observations lead to a number of important assumptions:

- An increase in CBF creates an increase in both cerebral blood velocity and CBV.

- Since increases in CBF exceed increases in $CMRO_2$, the blood flowing through the activated tissue not only flows faster, but it enters the venules and draining veins with more oxygen and thus a higher oxygen saturation than before, since less oxygen (per unit volume of blood) was consumed by the activated tissue.
- Increases in neural activity above baseline levels raises both [HbT] and $[HbO_2]$ and decreases [Hb].
- CBV changes remain better localized than BOLD signal changes because Hb is mobile and blood with decreased [Hb] eventually flows into nearby draining veins and reenters the circulation, thus shifting the locus of the BOLD and Δ[Hb] signals.

The observation that increases in CMR_{glu} exceed that of $CMRO_2$ leads to the following assumptions:

- Some of the metabolic energy must be derived through glycolysis.

The CBF measured at a distance from the activated regions sometimes decreases below baseline levels.

- Total CBF may be conserved throughout the brain, with local increases being compensated by regional decreases to hold the overall CBF constant.

Other observations

- Dopamine may directly affect local blood flow [60].
- Neuron-to-astrocyte signaling is a key mechanism in functional hyperemia [61].
- Receptor activation by neurohumoral substances referred to as "EDHF" opens potassium ion channels in endothelial cells [62].
- Local field potentials, not spiking activity, correlate better to hemodynamic changes [63].

Current hypotheses concerning neurovascular and activation-flow coupling

Magistretti's hypothesis: Glutaminergically mediated increases in neuronal activity cause large increases in glycolytic metabolism within astrocytes. This energy is used to convert glutamate back to glutamine, which is then recycled by neurons. Thus the primary metabolic change associated with increases in neuronal activity is glycolytic and occurs within astrocytes, not in the neurons themselves [64-66].

Frahm [67] and Hyder [68] have seen evidence that during extended stimulation, $CMRO_2$ may increase to match CMR_{glu}, since measurements of cortical lactate concentration peaked and then returned to baseline after ~6 minutes of continuous stimulation. Bandettini [69] claims that CMR_{glu} still exceeds $CMRO_2$ during prolonged periods of activation, so long as habituation does not occur. This is consistent with Fox [42] and Raichle's [34] initial PET observations.

Malonek and Grinvald [14] propose a different hypothesis: Increases in neuronal activity are powered almost exclusively by phosphorylation. The temporal phase shift introduced by the slow and delayed hemodynamic response drains the local tissue oxygen reserves. The need to repay this oxygen debt results in a proportionally greater hemodynamic response, well in excess of the real-time metabolic demand.

This may be supported by Buxton and Frank [70] , who calculated that CBF would have to exceed the stoichiometric changes in $CMRO_2$ in a nonlinear fashion, simply as a consequence of both the poor diffusivity and low solubility of oxygen in brain tissue, along with a drop in the oxygen extraction fraction due to the shorter capillary dwell time of the faster flowing blood. So in theory, CBF is directly coupled to $CMRO_2$, but the relation is nonlinear in order to compensate for the nonlinear impediments to increased oxygen delivery.

A metabolism-flow model that explains how ΔCBF could greatly exceed $\Delta CMRO_2$ while preserving the presumption that they are tightly coupled is the "diffusion-limited transport" model for oxygen delivery, proposed by Buxton [70]. He suggests that CBF and $CMRO_2$ may be tightly coupled if the imbalance between them results from a limitation in the available O_2 actually reaching the tissue. This is because the oxygen extraction decreases with increasing flow, due to a shorter capillary dwell time, so a large change in blood flow would be needed to support even a small increase in $CMRO_2$. This model relies on three basic assumptions: 1) All capillaries are always perfused, even at rest – i.e. there is no capillary "recruitment" occurring, only an increase in CBF. 2) All of the oxygen extracted from the capillaries is metabolized. 3) The dominant influence on blood flow regulation is that required increase in the rate of oxygen delivery to the activated tissue.

The intravascular Hb concentration is a simple function of CBF and $CMRO_2$ and can be estimated using Fick's Law. The intravoxel Hb concentration, however, is more complex. The capillaries and, to a greater extent, the veins, are highly distensible structures. Capillaries are somewhat elastic, and will swell slightly as blood flow, and hence intravascular pressure, increase. They appear to relax quickly when the pressure is reduced. Venules and veins show a much greater elastic swelling that actually grows with time. Their relaxation is both rapid and delayed. The "delayed compliance" probably results from the viscoelastic nature of the perivascular smooth muscle.

This time-variant vascular rheology imparts a temporal nonlinearity to the BOLD signal, since CBV, which is dominated by venous vascular volume, is now a nonlinear function of CBF in both magnitude and time. Although tissue is elastic, imaging voxels are not. As the blood vessels swell, parenchymal tissue is "squeezed" into adjacent voxels, generating spatial and temporal artifacts in the BOLD image. One model which helps explain the temporal features of the BOLD response is the "vascular balloon" model initially proposed by Buxton [71] and later analyzed using windkessel theory by Mandeville [25, 71]. It appears that the temporal features of the BOLD signal may result from a combination of fast (elastic) and slower (viscoelastic) vascular responses to changes in intraluminal pressure.

A non-Newtonian version of the balloon model can explain the bimodal nature sometimes seen in the CBV time course: a rapid initial rise in CBV from elastic compliance of the capillaries, with a slower exponentially-limited increase due to the viscoelastic properties of the perivascular smooth muscle surrounding the venules. It also suggests that the BOLD post-stimulus undershoot may result from delayed viscoelastic recovery of the vascular wall. Delayed venous compliance offers an alternative to the hypothesis that the BOLD undershoot was due to a mismatch between CBF and $CMRO_2$, thus the balloon model is consistent with Buxton's diffusion-limited transport model.

How DOT and fMRI detect cerebral activation

The blood flow and oxygenation changes following cerebral activation are what make both DOT and fMRI measurements possible.

A rise in blood volume within the local vasculature increases the optical absorption in that region of tissue. This increase in absorption is easily detected with DOT. Changes in oxygen saturation of the capillary and venous blood leaving the activated region alter the spectral absorption of the blood itself, which is also detectable with DOT.

fMRI operates by a completely different mechanism. It detects changes in blood oxygenation indirectly, through the effects of weakly paramagnetic deoxyhemoglobin (or an intravascular superparamagnetic contrast agent, for cerebral blood volume measurement), which alters the local magnetic field in the surrounding tissue. Changes in the magnetic field alter the nuclear precession rate of nearby protons (the nuclei of "light" hydrogen atoms bound to water molecules), which is then detected with sensitive RF antennas and receivers [3].

CBF changes cannot be measured with DOT, but they can still be detected optically. Blood velocity changes can be detected with optical Doppler techniques. These are similar to acoustic Doppler, but employ photons instead of acoustic compression waves [49]. The main advantage of optical Doppler is better spatial localization (a consequence of the far shorter wavelength of light vs. sound waves in tissue). The main disadvantage is poor depth penetration, since optical scattering from within the tissue itself destroys the spatial coherence needed to generate the Doppler signal. Acoustic Doppler signals are more weakly scattered in tissue, so spatial coherence of the acoustic signal is preserved at tissue depths well below those possible for optical Doppler measurements.

3.5 Biogenic interference

A number of naturally occurring (i.e. biogenic) events affect hemodynamics. Systemic volume and flow changes in the arterial compartment are caused by pulsatile contractions of the chambered heart. Spontaneous ventilation produces periodic thoracic pressure changes, while also influencing vagal tone [50]. These act to modulate both venous return pressure and, to a lesser extent, arterial pressure. An example of the hemomodulating effects of cardiac and pulmonary function on the magnitude of the cortical optical signal in a rat is shown in Figure 3.13. An example of human biogenic signals displayed in the frequency-domain is shown in Figure 3.14.

The modulation depth of these signals typically varies with source wavelength. For example, a 690nm source will likely experience a larger respiratory component (due to the greater concentration of Hb in the draining veins) and a weaker cardiac component (since the arterial blood is mostly HbO_2, and absorbs little at 690nm) than an 805nm source.

Pulmonary modulation is larger in amplitude but lower in frequency than cardiac modulation. As a result, it is potentially more corrupting to hemodynamic measurements, since the fundamental respiratory frequency in humans falls within the same range as most neurogenic and somatosensory-induced hemodynamic changes. Other poorly understood cyclic arterial pressure fluctuations, such as Mayer waves and vasomotion, appear at around 0.1Hz under certain conditions [72, 73].

Unfortunately, due to the biological nature of these signals, they are neither stable in frequency nor in amplitude [74]. Their natural chaotic instability, along with a phenomenon called spontaneous time-locking (in which the respiration and/or heartrate spontaneously couples to any repetitive stimulus; an autonomic version of toe tapping to a tune) means that the removal of biogenic confounds is not an easy task. Simple averaging through coaddition of the data from multiple trials works only up to a point, because the interfering signals (heartrate, respiration) will often time-lock to the evoked response, confounding the data. Frequency-domain methods (i.e. FFT techniques) are hampered by these rate instabilities, which broaden the bandwidth of respiration and heartrate components enough to overlap with the DOT signal components.

If sufficiently robust correlates to these biogenic noise sources can be found, then these limitations may be overcome. For example, the inter-beat interval, which varies for a variety of reasons, can be measured independently through ECG or pulse oximetry. DOT data can then be combined with the ECG data using decorrelation techniques to significantly improve the signal quality.

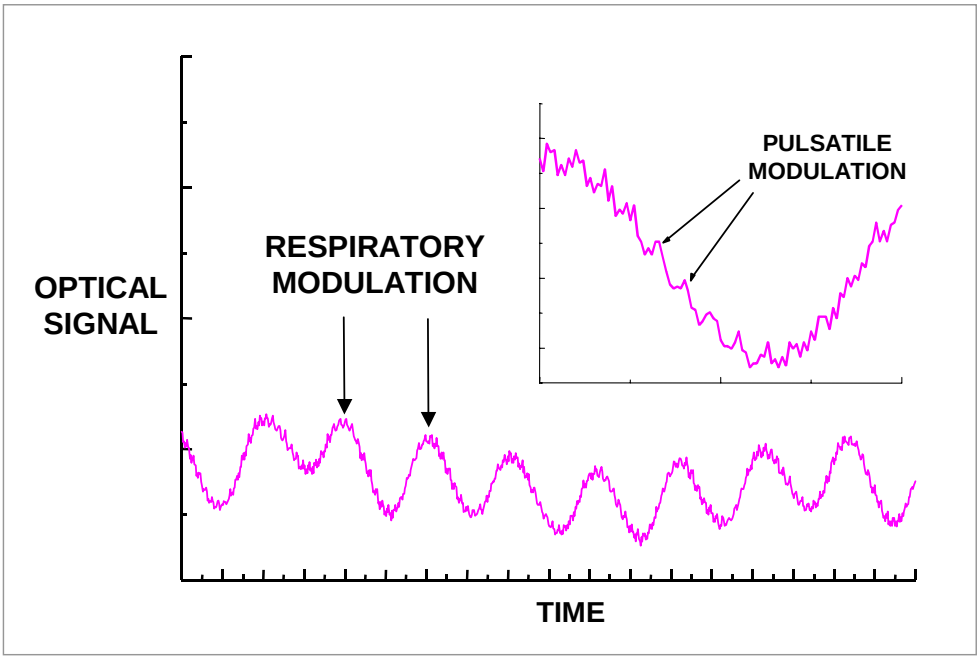


Figure 3.13. An example of how cardiac and pulmonary activity can modulate the intensity of a diffuse optical signal. Most of the static (DC) signal component has been subtracted for clarity. The actual modulation depth of biogenic signals typically ranges between 1% and 5%, and varies with source wavelength, probe position, and patient physiology [46].

Exploiting biogenic information

Although usually viewed as a nuisance, biogenic signals contain a significant amount of physiologic information. For example, the cardiac signal alone can provide information on heart rate, interbeat interval, cardiac output, and pulse pressure, as well as information on breathing rate, posture, and general neural activity (through respiratory sinus arrhythmia) [74]. The pulmonary signal contains information on ventilation rate and depth. It also provides some information on acid-base balance and general state of arousal [50]. There is doubtless far more physiologic information available that has yet to be uncovered.

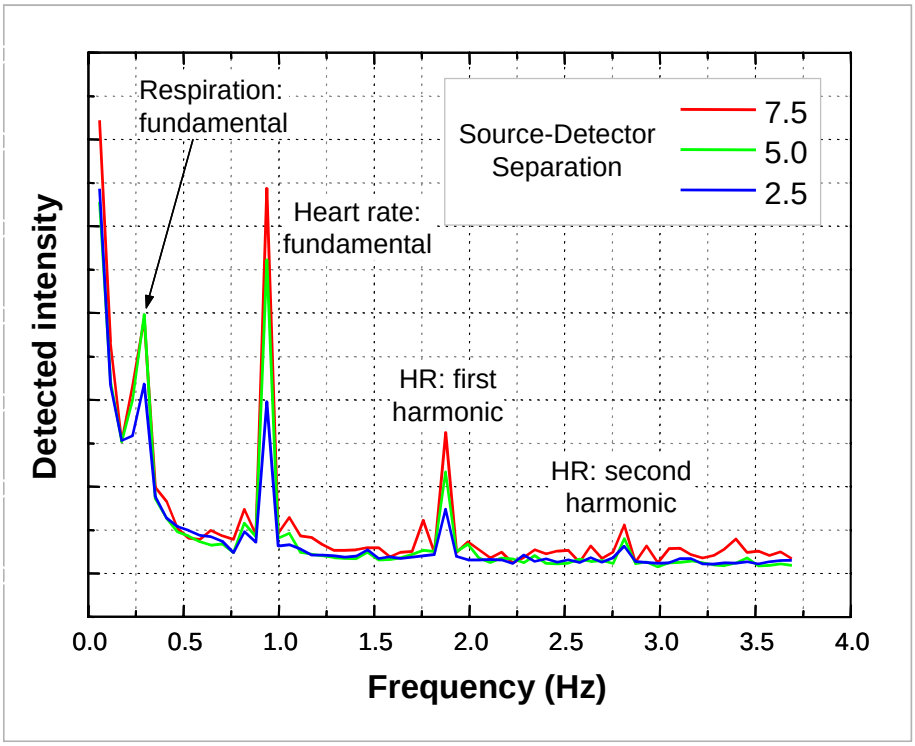


Figure 3.14. Cardiac and pulmonary modulation of the diffuse optical signal as seen in the frequency domain. The two largest frequency components are the pulmonary and heart rate fundamentals. It is interesting to note that the largest source-detector separation provided the strongest modulation. This suggests that the shallower tissue, transited by optical signals at the two shorter S-D separations, was more weakly perfused than the deeper tissue [46].

Mayer waves and vasomotion

Spontaneous periodic arterial pressure, blood flow, and heart rate fluctuations have been noted for more than a century. These periodic fluctuations are variously referred to as vasomotion or Mayer waves (after an article published in 1876 by Mayer), and typically occur within the 0.08Hz to 0.12Hz spectral band. One likely assumption is that Mayer waves represent transient ringing or low-level oscillation within the feedback loop of the vascular autoregulation system – a rather complex feedback control system. If so, then it is likely caused by erosion of the phase margin by the slow vascular response to changes in metabolic demand [49, 72, 73].

The origin and function of these low frequency vascular fluctuations are still a mystery, however a number of observations have been made. These include:

- Mayer waves persist even after destruction or inhibition of central nervous activity, thus it is likely that their origin is at the microvascular level [72].
- They are highly correlated with periodic changes in peripheral resistance.
- They are exaggerated under certain pathologic and disease states.
- A simulation of the vascular system using a simplified autoregulating Windkessel model predicts a damped resonant response at ~0.1Hz, as shown in Figure 3.15 [49].

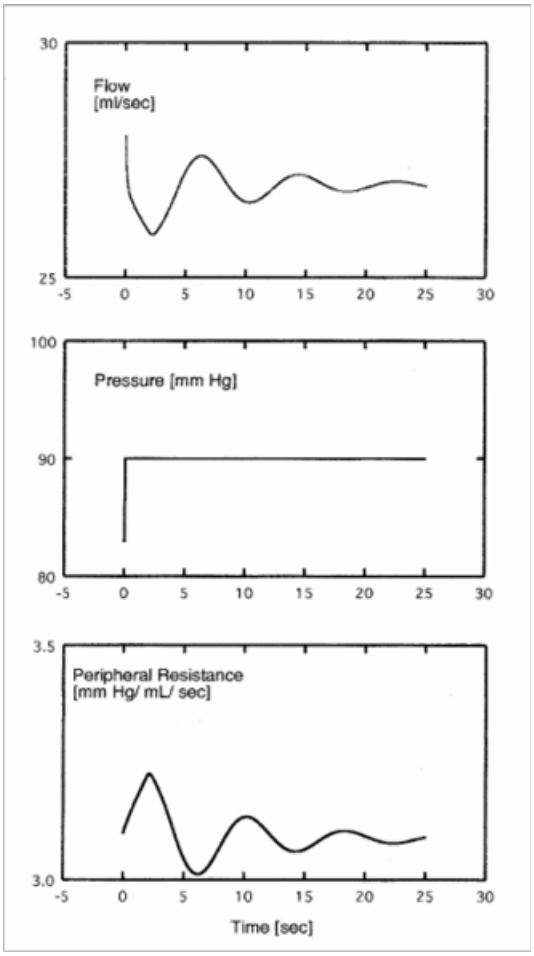


Figure 3.15. The pressure step response computed from an autoregulating Windkessel vascular model. The flow and peripheral resistance ring in direct opposition (i.e. 180' out of phase WRT each other) with a damped sinusoidal response at a resonance frequency of ~0.1Hz [49].

Mayer waves have been observed to time-lock to various stimulus paradigms, and allowances must be made for this effect during DOT measurements [75].

Compression ischemia

The blood flow to any region of the body can be slowed or stopped if the local extravascular pressure approaches or exceeds the maximum perfusion pressure. Since blood vessels are flexible and collapse readily, they cannot sustain a vacuum. Thus, blood cannot receive a vacuum assist from the vascular system: i.e. blood cannot be drawn "uphill". This is why blood pressure can be measured using a pressure cuff, and why your arm muscles fatigue so quickly when painting a ceiling.

The perfusion pressure is greatest in the arterial compartment and is relatively low (~10 to 50mmHg) in the capillary and venous compartments, as shown in Figure 3.1. Modest mechanical force can create hydrostatic pressures that can expel blood out of local veins and capillaries, thus blanching the skin. If the force is increased beyond the systolic pressure (~120mmHg, or about 2 psi), then regional perfusion will cease completely. As a result, cutaneous or shallow venous blood flow measurements must be performed with well-engineered optode assemblies, or else compression ischemia can introduce significant errors.

Compression ischemia can also be an asset. By applying sufficient pressure normal to the optode/skin interface, the perfusion through portions of the scalp can be reduced, decreasing blood

volume and increasing the optical throughput of the scalp. Compression ischemia can be used to monitor tissue metabolism. Since tumors are often hypermetabolic as compared to the surrounding tissue, it may be possible to detect breast tumors through their increased rates of phosphorylation, as depicted in Figure 3.16. By imaging the rate of blood deoxygenation throughout the compressed, and thus ischemic breast, the hypermetabolic regions should become visible [35].

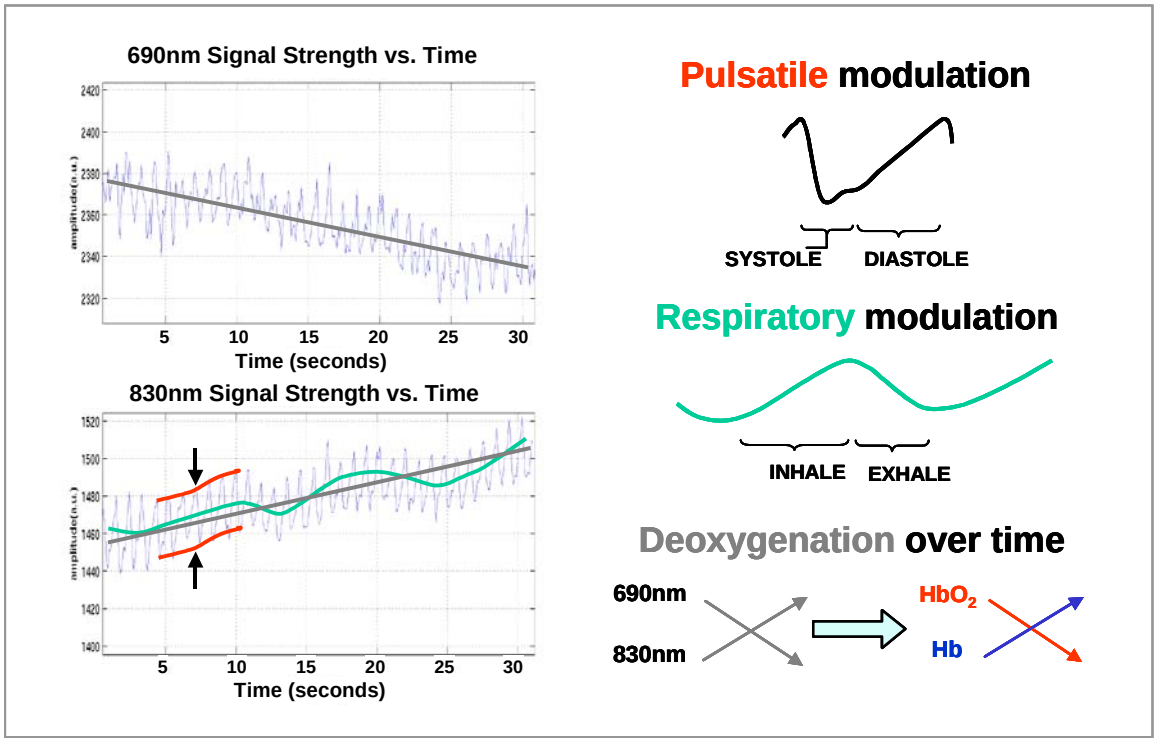


Figure 3.16. A diffuse optical measurement of breast tissue under compression. When combined, the 690nm and 830nm waveforms provide information on ventilation, heart rate, and blood deoxygenation vs. time. It may be possible to detect certain breast tumors by monitoring the deoxygenation rate within the tissue.

Other noise sources

Other biogenic noise sources can include:

- Involuntary motion artifacts (flinching, muscle spasms, restlessness, anticipatory tension, peristaltic activity)
- Voluntary motion artifacts (spontaneous and elicited muscular movement, swallowing, scratching)
- Fixturing issues (slippage or motion of the optode assembly during a measurement)
- The effects of stray light and source-to-detector light leakage from poorly shielded optodes

3.6 The Effect of Wavelength and Optode Spacing on Image Contrast

This analysis followed from an attempt to understand how absolute tissue absorption affects the quality of a hemodynamic optical signal. The premise of this analysis is the following:

1) For a given tissue absorber geometry, there should be an optimal level of overall (i.e. static and dynamic) optical attenuation which provides the maximum sensitivity to small absorption changes.

- If there is insufficient tissue absorption, the optode-to-optode throughput will be high, but the modulation depth will be small, resulting in poor SNR
- If there is too much tissue absorption, then the modulation depth will be large, but the flux reaching the detector will be very low, thus providing poor SNR

2) Control of the optical attenuation can best be achieved by controlling both the source wavelength and probe geometry.

- Since the extinction coefficients of Hb and HbO_2 are both strong functions of wavelength, it should be possible to select source wavelengths which optimize the tissue absorbance for each measurement and probe geometry

This section investigates how the selection of source wavelengths affects the image contrast for a given optode geometry and tissue properties. Based on this information, an investigator should be able to optimize image contrast by selecting the wavelengths used to illuminate the tissue. Although a combination of both engineering limitations and economics may determine the actual wavelengths selected, this knowledge could still be of great value to the investigator when designing future equipment, or for exploring the boundaries of the design space for other diffuse optical modalities.

Introduction

The majority of diffuse imaging has been performed within the 700nm to 1300nm near-infrared spectral band. There are a number of good reasons for this:

- Tissue absorption is very low within this band. Since the average optical pathlength through diffuse media can exceed the equivalent ballistic pathlength by as much as a factor of ten, minimizing the optical absorption has always been an important consideration.
- There are many optical sources available within this band. A variety of both coherent (laser diode, $Ti:Al_2O_3$, etc.) and incoherent (LED, SLD, tungsten lamps) sources are available.
- The broad spectral bandwidth of this region, in combination with the optical absorption characteristics of certain chromophores (Hb, HbO_2, NADH, and cytochrome oxidase) permits near-infrared spectroscopy (NIRS) using a number of different source wavelengths.
- For most investigators, the ultimate applications of diffuse imaging involve human subjects. These will require large optode separations, high source power levels, and very sensitive detectors in order to cover sizeable (>100cm^2) tissue regions and achieve deep (>3cm) penetration depths.

Although much thought has been spent on source wavelength selection for the purpose of minimizing differential pathlength errors, to the best of my knowledge no one to date has attempted to quantify the impact of wavelength selection on image contrast and sensitivity. Why is this even important? Because the modulating effects of a small dynamic absorber embedded in tissue are related to the attenuation experienced by the light as it passes through that absorber.

The absorption coefficient (μ_a) is defined as the optical pathlength required to attenuate the incoming light by a factor of 1/e – a reduction to 37% of its initial value. It is therefore inversely proportional to the absorptivity of the tissue (i.e. a large extinction means a small μ_a, and vice-versa). If the μ_a within a dynamic absorber that fills a large fraction of the banana pattern is very small, but the average optical pathlength through that region (as a result of the absorber's physical size) is large, then the ***regional absorbance*** will be high.

Now, if the μ_a within that region were to change by a small amount (say, due to hemodynamic changes following a neural stimulus), the fractional change in light intensity ***per unit length*** (i.e. the absorbance change) would be small as well, but the ***regional intensity change*** (what I call the

modulation depth, M) created by this dynamic absorber would be large, since attenuation increases multiplicatively, not linearly, with pathlength (assuming the regional attenuation exceeded about 3dB or so). This is when Beers Law applies. As the optode separation increases, the absorbance of the dynamic region might grow as well, increasing the modulation depth (a good thing – to a point). If only the absorbance of the inactive region within the banana pattern were to grow instead, M would drop accordingly, but it could still remain measurable.

At some point, however, the total (optode-to-optode) attenuation grows too large and the measurement then becomes optical throughput-limited. So, for situations requiring large optode separations along with a minimum temporal bandwidth (such as cognitive or sensorimotor studies in humans), the penetration depth will be restricted by the finite (FDA-mandated) optical power limits and the noise floor of the detectors. In this case, selecting source wavelengths for minimal absorption appears to be a wise choice. Thus, the design requirements for a system used for adult human cortical studies would be as follows:

- A very powerful source should be used to maximize the detected power.
- The detector should be very sensitive with lots of front-end gain. Since the optical throughput is very low, the dynamic range of the measurement will be low as well, and sensitivity will be important. [Note that this applies to a single optode pair measurement – source multiplexing strategies can greatly affect dynamic range requirements.]
- The source wavelengths should be chosen for minimal absorbance. Even a small spectral absorption difference will provide good Hb/dHb discrimination after many extinction lengths.
- The post-detection bandwidth should be set as low as practical to reduce noise.

Now let's consider the case where the objective is to monitor hemodynamic changes within the whisker barrel cortex of a rat. The activated region is on the order of 500μm in size, located at an average depth of around 700μm. With an optode separation of 1mm, tissue absorption at 800nm would be minimal, and the power should decay as $1/d^2$. Even if the geometry scaled identically from human cortical measurements, the μ_a of cortical tissue is substantially the same.

So, with a similar μ_a, the active region in the rat would have significantly less regional absorbance than in a human. What this means from a diffuse optical standpoint is that the optical throughput will be quite high yet the modulation depth will be very low, since the regional absorbance is so small. Detecting this modulation is not trivial, and the design requirements should therefore change to the following:

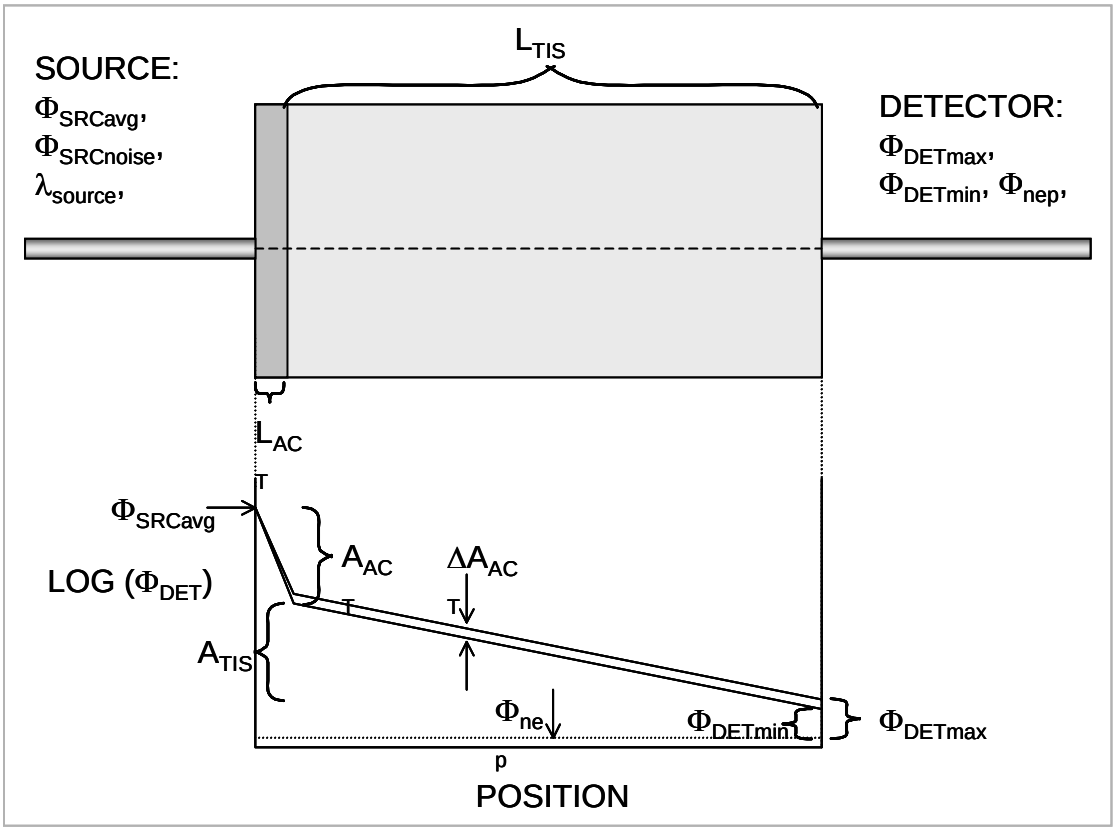
- Only moderate source power is required, since the overall absorbance will be relatively low.
- Source stability is now very important, since the modulation depth will be small. Ideally the temporal noise of the source (at the modulation frequency) should be below the shot noise limit at the receiver.
- A detector with lower gain in exchange for higher dynamic range is preferred. The detector noise floor is less critical in this case, since the large amount of detected flux will generate a shot noise level well above the noise floor of most detectors. Dynamic range is very important here, since a small AC change must be measured atop a very large DC signal. The ADC resolution must be commensurate with this range as well, to keep the quantization noise below the shot level. Solutions include sampling the average DC levels and then digitizing only the AC component of the signal, or using a technique such as individualized gain control, which is discussed in detail in Chapter 4. These allows for further amplification, and thus better matching to the ADC.

- The source wavelengths should be chosen to provide higher absorbance than for a human cortical measurement in order to improve the image contrast.

The balance of this section will consist of a theoretical discussion whose goal will be to prove these assertions.

Discussion

The optical path traveled by light through diffuse media is best represented statistically. In a planar tissue geometry, light emanating from one point at the surface and detected at another point nearby travels through a curved probability distribution function whose region of highest average flux density approximates the shape of a banana. Rather than resort to this level of detail, a simpler linear geometry will be used. Once this geometry has been fully explored, the model can be upgraded to include more complex tissue geometries. The diagram shown in Figure 3.17 illustrates the simple linear transmission geometry.



Φ_{SRCavg} = The average optical source flux entering the tissue (W)
$\Phi_{SRCnoise}$ = The temporal noise component of the optical source flux entering the tissue(W)
Φ_{nep} = The noise equivalent power of the detector at a specified bandwidth (W)
Φ_{DETmin} = The minimum flux reaching the detector (max. active area absorption) (W)
Φ_{DETmax} = The maximum flux reaching the detector (min. active area absorption) (W)
λ_{source} = The wavelength of the optical source (μm)
L_{ACT} = The physical length of the optically variable "active" region (m)
L_{TIS} = The physical length of the optically invariant "tissue" region (m)
A_{TIS} = The optical attenuation ratio for light transiting the tissue region (unitless)
A_{ACT} = The optical attenuation ratio for light transiting the active region (unitless)
ΔA_{ACT} = The change in attenuation through the active region due to hemodynamics (unitless)

Figure 3.17. A pictorial depiction of the linear approximation to a diffuse imaging scenario, showing a small active attenuating region in series with a much larger passive attenuating region. The active attenuator represents the small region of cortex which experiences hemodynamic changes following neural activation, while the passive attenuator represents the scalp, skull, and non-activated regions of the cortex. A more complex model would also include a "shunt" light term, to account for light that never reaches the activated region.

In order to minimize confusion between variable names and product terms, I have chosen to represent the product of two variables by placing a single space between them as follows: $\mathbf{\Phi_{SRC}\ A_{ACT}}$ represents the product of $\mathbf{\Phi_{SRC}}$ and $\mathbf{A_{ACT}}$. Most variable names also contain descriptive subscripts. Exponential terms are written in the form: exp(var), which represents "**e**" raised to the power of the variable expression "var".

Since my objective is to present a general argument to support my hypothesis rather than a rigorous analytical proof, I have chosen to make the following simplifying assumptions:

- The attenuation due to the inactive tissue, $\mathbf{A_{TIS}}$, is constant and uniform (i.e. it acts as an ideal linear attenuator), and thus its effect on the detected flux will be ignored for now. [It will become quite important when the detector noise floor is taken into account, though.]
- The effects of scattering will be lumped in as a simple (i.e. constant with respect to T λ) gain adjustment to the optical pathlength term $\mathbf{L_{ACT}}$. This will allow me to use Beer's Law for my initial calculations.

The optical power passing through the active region shown in Figure 3.17 is subject to attenuation. The simplest formula to describe this attenuation is Beer's Law:

$$I_{out} = I_o \exp(a\ c\ l)$$

Where: a = the absorptivity of the medium (1/(length x conc.))
c = the concentration of the absorber in the tissue (concentration)
l = the length of the absorbing medium (length)

So for our geometry, this becomes:

$$\Phi_{DET} = \Phi_{SRC} \exp(\mu_{ACT} L_{ACT}) = \Phi_{SRC} A_{ACT}$$

Where: μ_{ACT} = the absorption coefficient of the active tissue (1/length) and
L_{ACT} = the length of the active tissue region (length)

So: $A_{ACT} = \exp(\mu_{ACT} L_{ACT})$

For the sake of simplicity, let's assume that our source power $\mathbf{\Phi_{SRC}}$ remains constant during the measurement. What then matters, from a detection standpoint, is maximizing the change in $\mathbf{\Phi_{DET}}$ for a given change in $\mathbf{A_{ACT}}$.

But if: $\mathbf{\Phi_{DET} = \Phi_{SRC} A_{ACT}}$ and $\mathbf{\Phi_{SRC}}$ is held constant, then:

$$\Delta\Phi_{DET} = \Phi_{SRC} \Delta A_{ACT} = \Phi_{SRC} \exp(\Delta\mu_{ACT} \Delta L_{ACT})$$

Now let's hold $\mathbf{L_{ACT}}$ constant during a measurement too, because we'll assume (also for the sake of simplicity) that the length of the active tissue region remains constant during activation. So the only variable left to vary is $\mathbf{\mu_{ACT}}$, which makes sense from our knowledge of hemodynamics. Thus we have:

$$\Delta\Phi_{DET} = \Phi_{SRC} \exp(\Delta\mu_{ACT} L_{ACT})$$

So at this point, $\mathbf{\Delta\Phi_{DET}}$ is a function of $\mathbf{\Phi_{SRC}}$, $\mathbf{L_{ACT}}$, and $\mathbf{\Delta\mu_{ACT}}$. The linear proportionality to $\mathbf{\Phi_{SRC}}$ is logical and expected, however the $\mathbf{\exp(\Delta\mu_{ACT} L_{ACT})}$ term is more complex, and bears further examination. Since $\mathbf{L_{ACT}}$ is held constant, the relevant variable is $\mathbf{\Delta\mu_{ACT}}$.
So what determines $\mathbf{\Delta\mu_{ACT}}$ anyway? Well, for a given region of tissue, the change in the optical absorption coefficient $\mathbf{\Delta\mu_{ACT}}$ will appear as some small percentage of the nominal $\mathbf{\mu_{ACT}}$ value. Thus:

$$\Delta\mu_{ACT} = C_{STIM}\ \mu_{ACT}$$

Where: $\mathbf{C_{STIM}}$ is a fractional constant that represents the level of hemodynamic activation in the tissue

Thus, $\Delta\mu_{ACT}$ is directly proportional to both $\mathbf{C_{STIM}}$ (which is the value we want to monitor during the experiment) and the nominal value of μ_{ACT}, so the only thing under our control (i.e. that can be altered to suit our needs) is μ_{ACT}. Now using the Beers Law model:

$\mu_{ACT} = (a\ c) = a_{ACT}\ c_{ACT}$

Where: $\mathbf{a_{ACT}}$ is the tissue absorptivity (1/(length x conc.)) and
$\mathbf{c_{ACT}}$ is the absorber concentration (conc.)

So: $\Delta\mu_{ACT} = C_{STIM}\ \Delta a_{ACT}\ \Delta c_{ACT}$

Now $\mathbf{a_{ACT}}$ is a function of the type of chromophores present in the tissue and the optical wavelength. Since this will remain fixed during any single (i.e. one optode, one wavelength) measurement, we'll hold it constant as well, but let's express it as a function of wavelength: $\mathbf{a_{ACT}}(\lambda)$. (Yes, it's true that both Hb and HbO_2 change with time, but for the purpose of this discussion, this detail is relatively minor and will be neglected for now). So what we now have is:

$\Delta\mu_{ACT} = C_{STIM}\ a_{ACT}(\lambda)\ \Delta c_{ACT}$

So: $\Delta\Phi_{DET} = \Phi_{SRC} \exp(\Delta\mu_{ACT}\ L_{ACT})$, or

. . . $\Delta\Phi_{DET} = \Phi_{SRC} \exp(C_{STIM}\ a_{ACT}(\lambda)\ \Delta c_{ACT}\ L_{ACT})$.

But, to keep the math correct, it should be rewritten as:

$$\Delta\Phi_{DET} = \Phi_{SRC} \exp(C_{STIM}\ a_{ACT}(\lambda)\ c_{ACT\text{-}post}\ L_{ACT}) - \Phi_{SRC} \exp(C_{STIM}\ a_{ACT}(\lambda)\ c_{ACT\text{-}pre}\ L_{ACT})$$

Where: $\mathbf{C_{STIM}}$ is a fractional constant that represents the (paradigm-dependent) degree of hemodynamic activation (unitless)

$\mathbf{a_{ACT}}(\lambda)$ is the tissue absorptivity vs. wavelength (1/(length x conc.))

$\mathbf{c_{ACT\text{-}pre}}$ is the pre-stimulus absorber concentration (concentration)

$\mathbf{c_{ACT\text{-}post}}$ is the post-stimulus absorber concentration (concentration)

$\mathbf{L_{ACT}}$ is the length of the active tissue region (length)

Now, $\mathbf{C_{STIM}}$ is a function of the specific stimulus protocol, so we don't want to alter it (or else it might invalidate the entire experiment) and $\mathbf{L_{ACT}}$ is a constant. But what about $\mathbf{a_{ACT}}$ and $\mathbf{\Delta c_{ACT}}$? Well, $\mathbf{\Delta c_{ACT}}$ is the hemodynamic concentration change that results from the stimulation, which is part of what we're trying to measure, so we can't do much about that. But $\mathbf{a_{ACT}}$ is not a measured variable. It's not a function of the stimulus paradigm. But it directly affects $\Delta\mu_{ACT}$, and thus $\Delta\Phi_{DET}$, which is the actual flux change we detect with our instrument. If we could somehow control the value of $\mathbf{a_{ACT}}$, then we could maximize the value of $\Delta\Phi_{DET}$, which would improve the sensitivity of our measurement.

Why? What does the term "sensitivity" really mean here anyway? Since the optical detector in our instrument has a finite electrical noise floor and it converts optical flux to an electrical signal, the noise floor can also be expressed in terms of flux units, i.e. a "flux noise floor". The common term for this is the Noise-Equivalent Delta Flux, or **NE$\Delta\Phi$**, for short. In our case, since the concentration change is small, **$\Delta\Phi$** can be assumed to be directly proportional to **Δc_{ACT}**. So a more relevant noise measure for our instrument is the **NEΔC**, the noise-equivalent concentration change.

Note that this concept can be applied to just about any type of sensor: For magnetic sensors, **NEΔB**. For thermal sensors, **NEΔT**. For charge sensors, **NEΔQ**, etc. For our purpose I'll define sensitivity to be inversely proportional to the "noise-equivalent **Δc_{ACT}**", or **NEΔC** of the instrument. Thus, anything which reduces the **NEΔC** will improve sensitivity. The two factors which most directly affect the **NEΔC** are:

* The magnitude of $\Delta\Phi_{DET}$ for a given concentration change, i.e. the $\Delta\Phi_{DET}/\Delta c_{ACT}$ ratio.
* The **NE$\Delta\Phi$** of the equipment.

So an analytical expression for **NEΔC** is: $(\Delta c_{ACT}/\Delta\Phi_{DET})$ **NE$\Delta\Phi$**. What this says is that there are two ways to improve the sensitivity of our system: To increase the $\Delta\Phi_{DET}/\Delta c_{ACT}$ ratio (by optimizing the a_{ACT} level), and to decrease the **NE$\Delta\Phi$** of the electronics. Since the electronics noise is often dominated by the irreducible uncertainties of photon statistics (in well-designed equipment, at least), the emphasis should also be placed on optimizing a_{ACT}.

I have already stated that a_{ACT} is a function of both the chromophores of interest and the optical wavelength. Well, we can't easily alter the optical properties of the blood without affecting the measurement, so if the chosen chromophores are Hb and HbO_2, then this term must be considered constant. [Note that this may not always be the case. If all that matters is the degree or rate of perfusion, then an "extrinsic" chromophore (a chemical dye such as ICG, for example), can be injected into the circulation upstream of the region of interest. In this way, a_{ACT} can be varied.] What we **can** change, however, is the optical wavelength!

Historically, achieving large optode separations has been an important goal – especially for diffuse imaging of deeper objects, where a greater range of optode spacing can lead to more accurate image reconstruction. As a consequence of this, most researchers have selected their optical wavelengths to permit the maximum optode separation. This means minimal tissue absorption. The prevailing assumption was that, since the detected flux levels for small optode spacings was so large, there was little concern over obtaining sufficient hemodynamic sensitivity.

For some animal models, however, optode separations can be as small as 1mm, which is a fifty-fold reduction in optical geometry from the human model. What this means is that if the same source wavelengths were used to monitor the cerebral hemodynamics in rodents, the **NEΔC** of the measurement could be up to fifty times worse. In order to prevent this loss in sensitivity from occurring, we must optimize our source wavelengths. So how can we determine what our source wavelengths should be? Well, this leads us to a more fundamental question:

- Is there an optimum value for a_{ACT}, and if so, how can it be determined?
- How big (or small) should a_{ACT} be for each subject type and optode spacing?

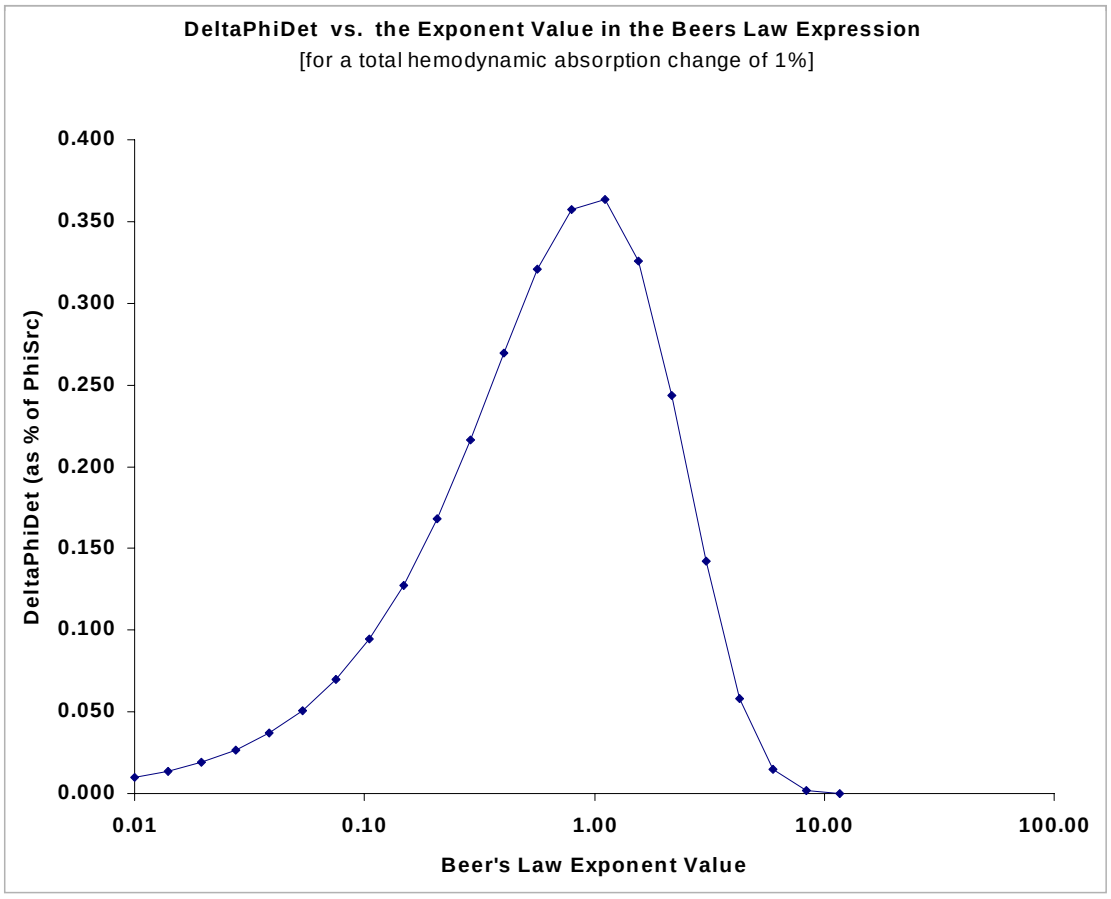


Figure 3.18. A plot of $\Delta\Phi_{DET}$ vs. the exponent value (C_{STIM} a_{ACT} c_{ACT} L_{ACT}). Note that the exponent value which generated the largest $\Delta\Phi_{DET}$ was 1.

The answer to both of these questions can be found in Figure 3.18. The Y-axis represents $\Delta\mathbf{\Phi_{DET}}$ and the X-axis represents the exponent value in the Beer's Law equation. Since this plot was made for a fixed $\mathbf{\Delta c_{ACT}}$ value of 1%, and the result scales linearly with $\mathbf{\Delta c_{ACT}}$ (trust me on this), we can assume that the plot also represents the $\Delta\mathbf{\Phi_{DET}}/\mathbf{\Delta c_{ACT}}$ ratio as well.

Conclusion

The rather intriguing result of this analysis is that the best hemodynamic contrast is achieved when the total value of the exponent is equal to 1. This means that the combined effects of the absorber concentration, the chromophore absorptivity, and the mean optical pathlength should all generate a static absorbance value within the active tissue region equal to "e" (2.7182).

So in order to achieve the best hemodynamic sensitivity, the optode spacing and the mean optical pathlength through the active tissue should be estimated, and the $\mathbf{a_{ACT}}$ should be adjusted (through selection of the appropriate optical source wavelengths, based upon the known optical properties of the chromophores of interest) to yield a total exponent value equal to 1.

Note that this analysis neglected real-world effects of the electronics noise floor on the measurement system. This can come from two sources: from fluctuations in the intensity of the (ideally) static optical sources, and from electrical and thermal noise in the optical detectors.