XI

Brain Imaging

Steven Laureys, Serge Goldman and Philippe Peigneux

INTRODUCTION

While philosophers have, for centuries, pondered upon the relation between mind and brain, neuroscientists have only recently been able to explore the connection analytically — to peer inside the black box. This ability stems from recent advances in technology and emerging neuroimaging modalities. It is now possible not only to produce remarkably detailed images of the brain's structure (i.e. anatomical imaging) but also to capture images of the physiology associated with mental processes (i.e. functional imaging). We are able to see how specific regions of the brain 'light up' when activities such as reading this book are performed, and how our neurons and their elaborate cast of supporting cells organize and coordinate their tasks. As demonstrated in the other chapters of this book, the mapping of the human mind (mostly by measuring regional changes in blood flow, initially by positron emission tomography (PET) and recently by functional magnetic resonance imaging or (fMRI)) has provided insight into the functional neuroanatomy of neuropsychiatric diseases.

Amazingly, the idea that regional cerebral blood flow (rCBF) is related intimately to brain function goes back more than a century. As is often the case in science, this idea was initially the result of unexpected observations. The Italian physiologist Angelo Mosso first expressed the idea while studying pulsations of the living human brain that keep pace with the heartbeat (Mosso, 1881). These brain pulsations can be observed on the surface of the fontanelles in newborn children. Mosso believed that they reflected blood flow to the brain. He observed similar pulsations in an adult with a post-traumatic skull defect over the frontal lobes. While studying this subject, a peasant named Bertino, Mosso observed a sudden increase in the magnitude of the 'brain's heart-beats' when the church bells signalled 12 o'clock, the time for a required prayer. The changes in brain pulsations occurred independently of any change in pulsations in the forearm. Mosso understood that the bells had reminded Bertino of his obligation to say a silent Ave Maria. Intrigued by this observation, Mosso then asked Bertino to perform a mental calculation; again, he observed an increase in pulsations and, presumably, blood flow as the subject began the calculation and a second rise just as he answered. This was the first study ever to suggest that measurement of cerebral blood flow might be a way of assessing human cognition.

Charles Roy and Charles Sherrington characterized this relationship further at Cambridge University. Based on animal experiments, they suggested that 'the brain possesses an intrinsic mechanism by which its vascular supply can be varied locally in correspondence with local variations of functional activity'. Their observations were fundamental, and their brilliant deductions have dominated the brain-imaging field since. One of the most extraordinary examples of the relationship posited by Roy and Sherrington was observed in Walter K., a German-American sailor who consulted Dr John Fulton at Boston's neurosurgery clinic for headache and failing vision. Walter also reported to hear a humming noise in his head. Fulton, when listening with a stethoscope at the back of his patient's head, confirmed this bruit and organized an exploratory intervention. Dr Harvey Cushing performed the neurosurgery and found a large arteriovenous malformation overlying the visual cortex. An attempt to remove the vascular malformation failed and left Walter with a bony defect overlying his visual cortex. His physicians could now hear the bruit even more clearly. During the course of his stay, Walter mentioned that the noise in his head became louder when he was using his eyes. As Fulton (1928) published later in the journal *Brain*:

> It was not difficult to convince ourselves that when the patient suddenly began to use his eyes after a prolonged period of rest in a dark room, there was a prompt and noticeable increase in the intensity of his bruit ... Activity of his other sense organs, moreover, had no effect upon his bruit. Thus, smelling tobacco or vanilla did not influence it, straining to hear the ticking of a distant watch produced no effect, and ordinary quiet conversation was without demonstrable influence.

In order to document his remarkable observations for others to appreciate, Fulton recorded the sounds of the bruit while his patient 'was allowed to lie on his stomach in a comfortable position, with his chin resting on the edge of a chaise-longue in such a way that he could close his eyes and open them to read a newspaper lying on the floor'. Within 20–30 seconds after Walter began to use his eyes to read the newspaper, there was a noticeable increase in the intensity of the bruit. If the lights were put out, the bruit continued for nearly a minute afterward; then it gradually subsided, and at the end of two minutes it had returned almost to resting level. Recent research capable of recording changes in blood flow within milliseconds has provided remarkable confirmation of these pioneer observations made with only a stethoscope and a simple recording device. After many such studies, Fulton 'gained the impression' that it was the effort of trying to discern objects that were just at the limit of his patient's acuity that brought on the increases of the bruit. Merely shining light into his eyes when he was making no mental effort had no effect. This was, indeed, a remarkable observation, the significance of which would not be appreciated for many years. It was probably the first ever recorded result of top-down influences on sensory processing (Posner and Raichle, 1994).

In what follows, we will introduce the vast area of anatomical brain imaging (X-ray computed tomography (CT) and magnetic resonance imaging (MRI)) and functional brain imaging (PET, single photon emission tomography (SPECT), fMRI, electroencephalography (EEG), event-related potentials (ERPs), magnetoencephalography (MEG), magnetic resonance spectroscopy (MRS),

Biological Psychiatry: Edited by H. D'haenen, J. A. den Boer and P. Willner. ISBN 0 471 49198 5
© 2002 John Wiley & Sons, Ltd.

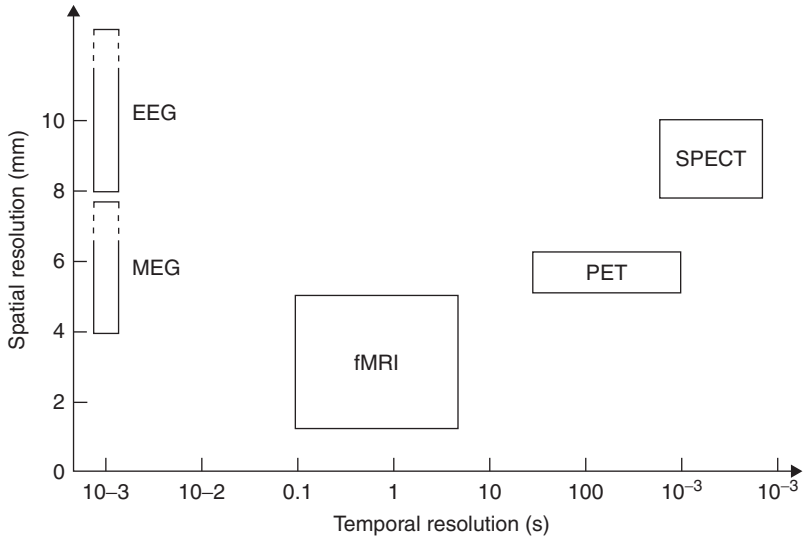

Figure XI.1 Approximation of the resolution in time and space of the most commonly employed functional neuroimaging techniques based on measurements of haemodynamic (fMRI, PET, SPECT) and electrical (EEG, MEG) activity of the brain

transcranial magnetic stimulation (TMS), and near-infrared spectroscopy (NIRS)). Each technique provides different information and has its own advantages and disadvantages in terms of cost, safety and temporal and spatial resolution (Figure XI.1). After briefly discussing their history and basic principles, we will present a short overview of study design and methods to process and analyse functional neuroimaging data.

COMPUTERIZED TOMOGRAPHY

History

The modern era of medical imaging began in the early 1970s, with the introduction of a remarkable technique called X-ray computed axial tomography, now known as CAT, X-ray CT or just CT. The theories that underlie tomographic imaging have existed since 1917, when the German mathematician Radon published his work dealing with reconstruction from image projections. Extension of this work by the South African physicist Allan Cormack in the early 1960s, and its practical applications by the British engineer Sir Godfrey Hounsfield, resulted in the construction of the first CT scanner. Both investigators received the Nobel Prize in 1979 for their independent contributions. Crucial in the development of X-ray CT was the emergence of clever computing and mathematical techniques to process the vast amounts of data necessary to reconstruct the images.

The development of X-ray CT had two consequences. First, it changed forever the practice of medicine because, for the first time, clinicians could non-invasively clearly view living human tissue such as the brain (standard X-rays reveal only bone and some surrounding tissues). Second, it immediately stimulated engineers and scientists to consider alternative ways of creating images of the body's interior using similar mathematical and computerized strategies for image reconstruction (e.g. SPECT and PET) (Posner and Raichle, 1994).

Despite its wide availability, CT has been replaced by the more sensitive MRI as the procedure of choice for cerebral imaging. CT is useful mainly when rapid information about the state of the brain is desired. In particular, it helps in making the choice between surgical and medical management of patients with sudden onset of neurological symptoms; such conditions include head trauma and stroke (where a differentiation between haemorrhage and infarction is important). CT is also used widely for the evaluation of lesions that involve bone (e.g. fractures or bone metastases) and calcifications within lesions of the brain.

Basic Principles

X-rays (i.e. a form of electromagnetic radiation travelling at the speed of light and carrying a large amount of energy) are capable of knocking an electron out of its orbit, therefore they are a form of ionizing radiation. A CT scanner delivers a narrow beam of X-rays that pass through the head. The exiting beam is then collected by a set of detectors, converted into digital data, and fed into a computer for image reconstruction. As the beam travels through the brain, it undergoes attenuation due to interaction with the various tissues it encounters. The degree of attenuation depends on the tissue density: very dense tissue, such as bone, attenuates lots of X-rays, cerebral grey matter attenuates some X-rays, and fluid attenuates even fewer. X-ray detectors positioned in a circle around the head collect attenuation readings as the beam is delivered from multiple angles. A computerized algorithm reconstructs a slice from these multiple readings. Thus, the contrast in a CT image is due to differences in X-ray attenuation among various tissues.

The use of exogenous contrast media in CT (and MRI) of the brain improves the sensitivity of detection and delineation of pathological structures, such as tumours, inflammation and ischaemia. CT contrast agents, like all commonly employed radiographic contrast agents, utilize substances with a high electron density (typically iodated agents) to absorb X-rays and thus produce a contrast effect. In both CT and MRI, a disruption of the blood–brain barrier results in accumulation of the intravenously administered contrast material in the extravascular space, leading to signal enhancement.

MAGNETIC RESONANCE IMAGING

History

MRI comprises a vast and varied array of techniques that use no ionizing radiation and provide an enormous range of information.

From an established ability to provide high-quality structural information, magnetic resonance techniques are advancing rapidly and providing other clinically relevant physiological information, such as spectroscopic studies illuminating the details of biochemical status (MRS; see below), blood oxygenation level allowing functional activation studies (fMRI; see below), cerebral blood compartment (magnetic resonance angiography or (MRA)); perfusion (perfusion-weighted imaging (PWI)), water molecular diffusion (diffusion-weighted imaging (DWI)), cerebral microstructure and fibre tracking (using diffusion anisotropy effects measured by diffusion tensor imaging (DTI)), magnetization transfer imaging, etc.

MRI derives from a potent laboratory technique, nuclear magnetic resonance (NMR), which was designed to explore detailed chemical features of molecules. In 1946, Felix Bloch of Stanford University and Edward Purcell of Harvard University discovered independently the phenomenon of NMR. They were awarded the Nobel Prize 6 years later. In the period between 1950 and 1970, the use of NMR was limited to chemical and physical molecular analysis. In 1971, Raymond Damadian showed that the nuclear magnetic relaxation times of tissues and tumours differed, thus motivating scientists to consider magnetic resonance for the detection of disease. NMR moved from the laboratory to the clinic when Paul Lauterbur found that it could form images by detecting protons. He used a back-projection technique similar to that used in X-ray CT to generate these images. Protons are interesting because they are abundant in the human body; by acting as little compass needles, they respond sensitively to magnetic fields. In 1975, Richard Ernst proposed MRI using phase and frequency encoding and the Fourier transform. This technique is the basis of current MRI techniques. It resulted in excellent images of the anatomy of the brain and other organs that far surpassed in detail those produced by CT. Six years later, Ernst was rewarded the Nobel Prize. In 1977, Peter Mansfield developed the echo-planar imaging (EPI) technique, which would permit the production of images at video rates (30 ms per image). By 1986, the NMR microscope was developed, allowing 10-mm resolution. In 1987, Charles Dumoulin perfected MRA, which could image flowing blood without the use of contrast agents. fMRI was developed in 1993 (see below). Clearly, the magnetic resonance technique is powerful and continuously expanding. Because of the negative connotations associated with the word 'nuclear' in the late 1970s, NMR is now commonly known as magnetic resonance imaging.

At present, MRI is the procedure of choice for the structural imaging of the brain. However, it is susceptible to movement artefacts, and patients who are on life-support systems, have gunshot wounds, or who have implanted MRI-incompatible material (pacemakers, prostheses, etc.) still represent problems. The main limit on the wealth of diagnostic information that can be obtained for each patient is in the duration of the procedure. Approximately 1 hour of magnetic resonance examination time is the practical limit of patient discomfort and cooperation for routine clinical studies. As scan times decrease through the use of new sequencing techniques, the available examination time can be used to improve the quality of the data or to obtain additional types of data. A range of magnetic resonance physiological measurement techniques is poised to fill any reclaimed imaging time. It is expected that further refinements of fMRI, MRA, MRS, PWI, DWI, DTI and other magnetic resonance techniques will allow them to fit into routine clinical practice.

Basic Principles

Some atoms, such as hydrogen (^{1}H), behave like spinning charges. Spinning nuclei will induce microscopic loops of electric current. As a result, they generate a very small magnetic field. MRI depends on the fact that these spinning atoms (mainly hydrogen nuclei, i.e. protons) act as little compass needles in the presence of a magnetic field. A strong magnetic field (approximately 10 000 times stronger than the earth's magnetic field) can align most of the atoms. By applying radiowave pulses to the tissue, the atoms can be perturbed in a precise manner. As a result, they emit detectable radio signals unique to the number and state of the particular atoms in the tissue. The strength of the MRI signal (i.e. the signal emitted when nuclei return to their equilibrium state) depends primarily on three parameters: the proton density in a tissue (the higher the density of protons, the larger the signal), the T_1 relaxation time (longitudinal magnetization), and the T_2 relaxation time (transverse magnetization). These properties are variable among tissues and are the predominant factors responsible for contrast among tissues. The signal intensity of one tissue compared with another (contrast) can be manipulated by varying the time elapsed between application of the radiofrequency pulse and sampling of the emitted signal (Pykett, 1982). The signal in magnetic resonance images is high or low (bright or dark) depending on the pulse sequence used and the type of tissue under study. On a T_1-weighted image, fat, subacute haemorrhage, melanin protein-rich fluid, slowly flowing blood, and paramagnetic substances (e.g. gadolinium) will appear bright. On a T_2-weighted image, increased water, as in oedema, tumour, infarction, inflammation, infection or subdural collection will appear bright (dark on T_1-weighted images). MRI contrast involves the use of relatively non-toxic paramagnetic agents, such as gadolinium diethylenetriamine pentaacetic acid (Gd-DTPA). Unlike CT contrast agents, MRI contrast enhancement is indirect. Indeed, paramagnetic substances will act as relaxation centres for other nuclei in the local microenvironment and shorten the magnetic relaxation times of the surrounding hydrogen nuclei. MRI is a more sensitive method of detection of contrast, requiring dosages in the range of one-twentieth of that for CT.

The techniques used in MRI of the brain depend on the primary goal of its procedure (e.g. anatomical detail is best obtained using T_1 sequences, while inflammation and oedema are better visualized using T_2 sequences). As mentioned previously, MRI can provide a wealth of information not only about intrinsic T_1 and T_2 relaxation properties and spin density but also regarding blood and cerebrospinal fluid flow, bulk motion, diffusion, diffusion anisotropy, perfusion, local oxygenation, local iron content, membrane permeability, temporal dynamics of contrast agent interaction, etc.

POSITRON EMISSION TOMOGRAPHY

History

PET has its roots in tissue autoradiography, a method used for many years in animal studies to investigate organ metabolism and blood flow. In tissue autoradiography, a radioactively labelled compound is injected into a vein. After the compound has accumulated in the organ under interest (such as the brain), the animal is sacrificed and the organ removed for study. The organ is sectioned carefully and the individual slices are laid on a piece of film sensitive to radioactivity. This X-ray film records the distribution of radioactively labelled compound on each slice of tissue. After film development, a picture of the distribution of radioactivity within the organ is obtained, hence regional information on the organ's specific functions can be deduced. The injected radioactive compound determines the type of information. Radioactively labelled water, for example, measures cerebral blood flow. In the late 1940s, Seymour Kety developed this technique for autoradiography in laboratory animals. Accumulation of a radioactively labelled form of glucose measures cerebral metabolism because glucose is the primary source of energy for neurons. In 1977, thanks to such a tracer, Louis Sokoloff introduced a now widely used autoradiographic method for the regional investigation of neuronal activity.

Researchers in the field of tissue autoradiography became fascinated when CT was introduced in the 1970s. They realized that if the anatomy of an organ could be reconstructed by passing an X-ray beam through it, then the distribution of a previously administered radioisotope could also be reconstructed *in vivo*. They simply had to measure the emission of radioactivity from the body section. With this insight was born the idea of autoradiography of living human subjects. A crucial element was the choice of the radioisotope. A class of radioisotopes was selected that emitted positrons (i.e. particles identical to electrons, except that they carry a positive charge). A positron will combine immediately with a nearby electron. They will annihilate each other, emitting two gamma rays in the process. Because the gamma rays travel in opposite directions, detectors around the sample can detect the gamma rays and locate their origin. The crucial role of positrons in human autoradiography gave rise to the name positron emission tomography (Ter-Pogossian *et al.*, 1980).

Throughout the late 1970s and early 1980s, PET was developed rapidly to measure various activities in the brain, such as glucose metabolism, blood flow (see below), oxygen consumption, and uptake of drugs. Although PET is primarily a research tool for brain imaging, its increasing availability in medical centres for oncology and cardiac imaging makes likely its more widespread application to neuropsychiatric diseases. The most frequently performed PET studies measure resting regional cerebral metabolic rates for glucose (rCMRGlu) or changes in rCBF as indirect indices of neural synaptic activity (Magistretti and Pellerin, 1999). Recent developments are PET/CT combined imaging (offering improved attenuation correction and co-registration or fusion of the functional PET image with a high-anatomic-resolution CT image) and improved detector materials, such as lutetium oxyorthosilicate (LSO) scintillators (offering higher stopping power and counting rates).

Basic Principles

PET scanning involves the administration of positron-emitting radionuclides with short half-lives in which particle disintegration is captured by multiple sensors positioned around the head. The radiotracer is administered into a vein in the arm and is taken up by the brain through the bloodstream. After a course of a few millimetres, the positron will interact with an electron in the brain tissue and produce two high-energy photons at approximately 180 degrees apart from each other. In the PET scanner, a ring of detectors around the patient's head can detect these coincident photons. As the radioactive compound accumulates in different regions of the brain, and positron annihilations occur, the scanner detects the coincident rays produced at all positions outside the head and reconstructs an image that depicts the location and concentration of the radioisotope within a plane of the brain. This emission scan is then corrected by comparison with the attenuation image made from a transmission scan of the subject's head. PET studies involve the use of a cyclotron to produce the radioactive tracers. The type of information on the PET image is determined by the administered radiolabelled compound. Oxygen-15, fluorine-18, carbon-11 and nitrogen-13 are common radioisotopes that can combine with other elements to create organic molecules that can substitute for natural substances, such as water, glucose, levodopa(L-dopa), benzodiazepine-receptor ligands, etc. Using different compounds, PET can assess regional blood flow, oxygen and glucose metabolism, and neurotransmitter and drug uptake in the tissues of the working brain. PET can sample all parts of the brain with equal resolution and sensitivity. Typically, it can locate changes in activity with an accuracy of about 6 mm. PET measurements of radioligand binding are discussed elsewhere (see Chapter XIX.10).

Cerebral Metabolic Rate for Glucose

To study regional cerebral glucose utilization, a positron-labelled deoxyglucose tracer is used, [^{18}F]fluorodeoxyglucose (FDG) (Huang *et al.*, 1980). This tracer is taken up by active brain regions as if it was glucose. However, once inside the cell, FDG is phosphorylated by hexokinase to FDG-6-phosphate, which is not a substrate for glucose transport and cannot be metabolized by phosphohexoseisomerase, the next enzyme in the glucose metabolic pathway. Thus, labelled FDG-6-phosphate becomes metabolically trapped within the intracellular compartment. The amount of radioactive label that eventually remains in each discrete region of the brain is related to the glucose uptake and metabolism of that particular region. An FDG-PET scan summates approximately 30 minutes of cerebral glucose metabolism and allows assessment of regional variations. However, given the half-life of ^{18}F (2 h), it is less suited for brain-activation studies.

Cerebral Blood Flow

Most PET activation studies rely on the administration of radioactively labelled water, specifically hydrogen combined with oxygen-15, a radioactive isotope of oxygen ($H_2{}^{15}O$). The labelled water emits copious numbers of positrons as it decays (hydrogen isotopes cannot be used because they do not emit positrons). In just over a minute after intravenous injection, the radioactive water accumulates in the brain, forming an image of blood flow. The radioactivity of the water produces no deleterious effects. Oxygen-15 has a half-life of only 2 minutes, and an entire sample decays almost completely in about 10 minutes (five half-lives) into a nonradioactive form. The rapid decay substantially reduces the exposure of subjects to the potentially harmful effects of radiation. Moreover, only low doses of the radioactive label are necessary. The fast decay and small amounts permit many measures of blood flow to be made in a single experiment. In this way, $H_2{}^{15}O$-PET can take multiple pictures of the brain at work in different experimental conditions. Each picture represents the average neural activity of about 45 seconds. The total number of scans that can be made per subject (typically about 12 images) is limited by the exposure to radiation.

In the last decade, PET has been the technique used most widely to assess the neural substrates of cognitive processes at the macroscopic level. At present, it is superseded by fMRI for many applications (see below). It remains, however, a powerful tool in receptor imaging (e.g. assessment of neurotransmitter or drug uptake) and molecular imaging (e.g. assessment of gene expression or protein synthesis) in both normal and pathological states (Phelps, 2000).

SINGLE PHOTON EMISSION COMPUTED TOMOGRAPHY

History

The history of SPECT parallels that of PET. The development of nuclear tomographic techniques paralleled the advances in X-ray CT in the early 1970s (see above). In 1977, Keyes constructed the first rotating gamma camera SPECT system. In this technique, multiple angular projections are accumulated as the camera turns in orbit around the subject's head. (The mathematically equivalent technique of holding a gamma camera head steady while rotating the patient in a chair in front of the camera has been tried successfully but has not caught on commercially.) Hybrid SPECT systems, including coincidence cameras (which can perform both SPECT and PET imaging), and hybrids that incorporate a CT

scanner along with a SPECT system (which offer improved attenuation correction and perfect co-registration or fusion of the SPECT scan with a high-anatomical-resolution CT scan), are currently being developed (Groch and Erwin, 2001).

In general, SPECT tracers are more limited than PET tracers in the kinds of brain activity they can monitor, but they are longer lasting. Thus, SPECT does not require an on-site cyclotron. However, most SPECT technology is relatively nonquantitative, does not permit measured attenuation correction, and has a spatial resolution inferior to that of PET. On the other hand, SPECT is less expensive and more widely available.

Basic Principles

Similar to PET, SPECT also uses radioactive tracers, but it involves the detection of individual photons (low-energy gamma rays) rather than positrons emitted at random from the radionuclide to be imaged. A SPECT scanner uses two or three cameras that rotate around the patient to record data at different angles. Images of active brain regions are then reconstructed by combining a finite number of projections. Typical radionuclides include technetium-99m (^{99m}Tc) and iodine-123 (^{123}I) with half-lives of 6 and 13 hours, respectively. On average, SPECT acquisition times are 20–30 minutes.

Frequently used radiolabelled agents for brain perfusion SPECT are: (1) Tc-99m-hexamethyl propylamine oxime (Tc-99m-HMPAO), a lipid-soluble macrocyclic amine with a rapid brain uptake (reaches its maximum within 10 min of injection, and its distribution remains constant for many hours post-injection); (2) Tc-99m-bicisate ethyl cysteinate dimer (Tc-99m-ECD), with rapid uptake and very slow clearance from the brain (blood clearance is also rapid, resulting in high brain-to-soft-tissue activity ratios); (3) I-123-isopropyliodoamphetamine (I-123-IMP); and (4) the inert gas xenon-133, which permits the quantitative measurement of rCBF without the need for arterial sampling (however, this requires a technically difficult inhalation technique, and it has a poor spatial resolution). The long half-life, rapid brain uptake and slow clearance of most radiolabelled agents for brain perfusion SPECT offer the opportunity to inject the tracer at a time when scanning is impossible (e.g. during an epileptic crisis) and to scan (post-event) the associated distribution of activated brain regions.

In addition to their use in determining perfusion, radiotracers can also be used to determine biochemical interactions, such as receptor binding. Iodine-123-labelled ligands, such as iodo-hydroxy-methoxy-N-[(ethyl-pyrrolidinyl) methyl]-benzamide (IBZM), have been developed for imaging the dopamine receptor system and are used in studies of movement disorders and schizophrenia (IBZM is a D2 receptor agonist that shows high uptake in the striatum and can be displaced by haloperidol). Iodine-123-quinuclidinyl-iodobenzilate (I-123-QNB), an acetylcholine muscarinic antagonist, has been used to image these receptors in the brains of normal subjects and patients with Alzheimer's disease. Other iodine-123 labelled ligands have been used for imaging the benzodiazapine and serotonin receptors. These radiotracers, however, have not yet reached the stage of routine clinical practice, as a consensus on measuring receptor binding using single-photon tracers has not yet been reached, nor have diagnostic clinical strategies been identified.

FUNCTIONAL MAGNETIC RESONANCE IMAGING

History

In 1935, Nobel Laureate Linus Pauling discovered that the amount of oxygen carried by haemoglobin affects its magnetic properties. In 1990, Sejji Ogawa and his colleagues at AT&T Bell Laboratories demonstrated that MRI could detect these small magnetic fluctuations. Several research groups realized immediately the importance of this observation. By the middle of 1991, it was shown that MRI could detect the functionally induced changes in blood oxygenation in the human brain. This ability has led to the term 'functional MRI'. The ability of fMRI to monitor the oxygen signal in real time is limited not by technique but by physiology. Indeed, the stumbling block is the speed of neural activity with the rate of change of oxygenation levels. Signals from one part of the brain can travel to another in 10 ms or less. Unfortunately, changes in blood flow and blood oxygenation are much slower, occurring hundreds of milliseconds to several seconds later. Hence, fMRI is not able to keep up with the 'conversations' between brain areas. For the time being, the only methods that respond quickly enough are electrical recording techniques such as EEG and MEG (see below).

fMRI is taking the place of ^{15}O-labelled-water-PET as the procedure of choice for haemodynamic functional activation measurements. It has several advantages over PET activation studies. First, it does not require the injection of radioactive tracers as the signal comes directly from functionally induced changes in brain tissue (i.e. changes in venous oxygen concentration). Second, the spatial resolution is better, distinguishing parts as small as 1 or 2 mm (better than the PET resolution of about 6 mm). Third, the temporal resolution is better, monitoring changes in blood-flow-induced oxygen signal in real time, using echoplanar imaging (better than $H_2{}^{15}O$-PET resolution of about 45 s). Fourth, MRI provides both anatomical and functional information in each subject, hence permitting a more accurate structural identification of the active regions.

Some concerns have been raised about the intensity of the magnetic field to which the tissues are exposed in MRI, but so far there are no known harmful biological effects. The largest limiting factor is the claustrophobia some subjects may suffer as in most instrument designs the entire body must be inserted into a relatively narrow tunnel. Other limiting drawbacks are its susceptibility to the subject's movement artefacts and artefacts related to the use of metal-containing devices in the magnet (e.g. EEG electrodes, electrical wires).

Basic Principles

fMRI can detect an increase in oxygen that occurs in an area of heightened neuronal activity. The basis for this capacity comes from the way neurons make use of oxygen. Functionally induced increases in blood flow are accompanied by alterations in the amount of glucose the brain consumes but not in the amount of oxygen it uses. Indeed, despite the presence of abundant oxygen, the normal brain resorts to anaerobic metabolism during spurts of neuronal activity. Apparently, this physiological behaviour relies on tactics similar to those present in a sprinter's muscles. It is not understood fully why the brain acts in this way. Additional blood flow to the brain without a concomitant increase in oxygen consumption leads to a heightened concentration of oxygen in the small veins, draining the active neural centres. The reason for this is that supply has increased but the demand has not. Therefore, the extra oxygen delivered to the active part of brain simply returns to the general circulation by way of the draining veins.

The commonest form of fMRI is blood-oxygenation-level-dependent (BOLD) imaging (Ogawa *et al.*, 1990). The BOLD signal depends on the ratio of oxygenated to deoxygenated haemoglobin. In regions of neuronal activity, this ratio changes as increased flow of oxygenated blood temporarily surpasses consumption, decreasing the level of paramagnetic deoxyhaemoglobin. These localized changes cause increases in the magnetic resonance signal, which are used as markers of functional activation. Ultrafast scanning can measure these changes in the signal, which are mapped directly on to a high-resolution scan of the subject's anatomy. fMRI studies require magnets with field strengths superior to 1 tesla.

ELECTROENCEPHALOGRAPHY

History

EEG detects spontaneous brain electrical activity from the scalp. In 1848, the German physiologist Dubois-Reymond demonstrated that an externally recordable electrical signal occurred concomitantly with passage of a nerve impulse along a peripheral nerve. This discovery led the English physiologist Richard Caton to explore the possibility that nerve impulses flowing within brain cells might also produce electrical signals. In 1875, Caton published in the *British Medical Journal*, 'feeble currents of varying direction pass through the multiplier when one electrode is placed on the gray matter and one on the surface of the skull'. Caton demonstrated that the cerebral cortex had a tonic level of oscillatory electrical activity, and that additional phasic electrical activity could be evoked in response to peripheral sensory stimulation. The diagnostic potential of EEG was first hinted at in 1912 by the discovery of Kaufman that abnormal neuroelectric discharges could be recorded in experimentally induced epilepsy in animals. Despite the exciting developments in animal EEG, it was not until 1929 that Hans Berger, a German neuropsychiatrist, recorded the first human EEG. Hereafter, there was an explosion of human EEG studies with (of particular interest) the discovery of 'the EEG in epilepsy and in conditions of impaired consciousness'(Gibbs *et al.*, 1935). Today, EEG is considered a routine clinical procedure of substantial diagnostic value in neurology, neurosurgery and psychiatry, and it is proving to be a useful measure of brain physiology in neuroscientific research.

EEG provides temporal resolution in the millisecond range. However, traditional EEG technology and practice provide insufficient spatial detail to identify relationships between brain electrical events and structures and functions visualized by fMRI or PET. Recent advances help to overcome this problem by recording EEGs from more electrodes (experimental laboratories use 32–64 or even 120 or more electrodes), by registering EEG data with anatomical images, and by correcting the distortion caused by volume conduction of EEG signals through the skull and scalp. In addition, statistical measurements of sub-second interdependences between EEG time series recorded from different locations can help to generate hypotheses about the instantaneous functional networks that form between different cortical regions during mental processing. Physiological and instrumental artefacts (e.g. subject's eye or head movements, heartbeats, poor electrode contacts) can contaminate the EEG (and MEG), so care must be taken to correct or eliminate such artefacts before further analyses are performed.

Basic Principles

Scalp-recorded EEGs in the waking state in healthy adults normally range from several to about 75 μV. The EEG signal is largely attributable to graded postsynaptic potentials of the cell body and large dendrites of vertically oriented pyramidal cells in cortical layers 3 to 5. These are synchronized by rhythmic discharges from thalamic nuclei, with the degree of synchronization of the underlying cortical activity reflected in the amplitude of the EEG. Most of the EEG signal originates in cortical regions near the recording electrode. The columnar structure of the cerebral cortex facilitates a large degree of their electrical summation rather than mutual cancellation. Thus, the EEG recorded at the scalp represents the passive conduction of currents produced by summating activity over large neuronal aggregates. Regional desynchronization of the EEG reflects increased mutual interaction of a subset of the population engaging in 'cooperative activity' and is associated with decreases in amplitude.

To measure the EEG, electrodes are attached to the scalp with a conducting paste. Each electrode is connected with an electrically neutral lead attached to the ear, nose, chin or chest (i.e. reference montage) or with an active lead located over a different scalp area (i.e. bipolar montage). Differential amplifiers are used to record voltage changes over time at each electrode. These signals are then digitized with 12 or more bits of precision, and are sampled at a rate high enough to prevent aliasing of the signals of interest. EEGs are conventionally described as patterns of activity in five frequency ranges: delta (less than 4 Hz), theta (4–7 Hz), alpha (8–12 Hz), beta (13–35 Hz; sometimes subdivided into beta1 at 13–20 Hz and beta2 at 21–35 Hz), and gamma (above about 35 Hz).

EVOKED POTENTIALS

History

An evoked potential (EP) or ERP is the time-locked average of the EEG in response to a specific sensory, motor or cognitive event. Because of their low amplitudes, especially in relation to the background EEG activity, a number of stimuli have to be recorded and averaged with a computer in order to permit their recognition and definition. The background EEG activity, which has no fixed temporal relationship to the stimulus, will be averaged out by this procedure. As mentioned above, Richard Caton reported a clear reactivity of rabbit EEG to sensory stimuli in the form of a cortical-event-related response in 1875.

Basic Principles

Sensory evoked or 'exogenous' potentials are recordings of cerebral or spinal potentials elicited by stimulation of specific sensory pathways, e.g. visual evoked potentials (VEPs) elicited by monocular stimulation with a reversing chequerboard pattern; brainstem auditory evoked potentials (BAEPs) elicited by monaural stimulation with repetitive clicks; and somatosensory evoked potentials (SEPs) elicited by electrical stimulation of a peripheral nerve. They are an important and routinely used means of monitoring the functional integrity of these pathways.

Certain EP components depend on the mental attention of the subject and the setting in which the stimulus occurs, rather than simply on the physical characteristics of the stimulus. Such 'event-related' or 'endogenous' potentials are related in some manner to the cognitive aspects of distinguishing an infrequently occurring target stimulus from other stimuli occurring more frequently. For clinical purposes, attention has been directed particularly at the so-called P_{300} or P_3 component of the ERP (named after its positive polarity and latency of approximately 300–400 ms after onset of an auditory target stimulus).

As a research tool, ERPs can provide valuable information about the precise timing and cortical distribution of the neuroelectrical activity generated during mental activity. An averaged EP waveform consists of a series of positive and negative waves; a significant difference in latency, amplitude, duration or topography of one or more of these waves between experimental conditions that differ in one specific cognitive factor is assumed to reflect the mass neural activity associated with that cognitive factor. Measurements of changes in the amplitudes and timings of peaks in the series of EP waves allows inferences to be made about the sequence and timing of task-associated processes, such as prestimulus preparation, encoding of stimulus features, operations such as matching or comparison of stimulus codes and memory codes, evaluation of the meaning of the stimulus, and response selection and execution. However, it is important to bear in mind that the averaged EP method assumes that the component subprocesses comprising a cognitive behaviour do not vary in time from trial to trial (Gevins, 1998).

MAGNETOENCEPHALOGRAPHY

History

MEG and magnetic source imaging (MSI) measure the magnetic fields generated by electrical activity within the brain. In the early nineteenth century, Hans Christian Oersted, a Danish physicist, discovered that electrical currents generate magnetic fields. The first measure of magnetic fields generated by human bioelectric currents was a magnetocardiogram recorded by a 2-million-turn hand-wound induction coil magnetometer at room temperature (Baule and McFee, 1963). Five years later, using a similar conduction coil and signal-averaging techniques, the brain's alpha rhythm (generating a magnetic field 100 times weaker than that of the heart) was recorded in a specially designed magnetically shielded room constructed at the Massachusetts Institute of Technology (MIT) (Cohen, 1968). The subsequent development of point-contact superconducting quantum interference devices (SQUIDs), operating at liquid helium temperatures of -269°C, improved the sensitivity of the magnetometers, and the first stimulus-elicited cerebral magnetic fields could be recorded (Brenner *et al.*, 1975). The early technical difficulties were overcome in the late 1980s, such that instruments covering the whole scalp and containing more than 100 channels are now available commercially. Nowadays, magnetic field tomography (MFT), a technique based on distributed source analysis of MEG data, makes possible the three-dimensional reconstruction of dynamic brain activity in humans with a temporal resolution better than 1 ms and a spatial accuracy of 2–5 mm at the cortical level (which deteriorates to 1–3 cm at depths of 6 cm or more) (Ribary *et al.*, 1991).

The major advantage of techniques based on the measurements of cerebral electrical activity (i.e. EEG and MEG) is their uncompromized time resolution. Their major drawback, however, is their limited spatial resolution. Indeed, accurate localization of the source of brain activity remains difficult. Furthermore, the resolution becomes poorer the deeper into the brain we attempt to image. The main advantages of MEG over EEG are its superior spatial accuracy and ease of use, particularly when a large number of channels are involved. On the other hand, EEG complements MEG in detecting source components not detected by MEG (i.e. radially oriented sources) (Naatanen *et al.*, 1994). For the time being, MEG, MFT and MSI remain expensive and largely experimental research tools unavailable to most clinical settings.

Basic Principles

Electric currents in the brain produce a change in the magnetic field that can be detected outside the head by SQUID magnetometers. The signals measured can be used to compute the distribution of cerebral activity as a function of time. This method is related closely to EEG, in which the electric field pattern on the scalp is measured.

Electrical currents generate magnetic fields. The direction of the magnetic field is described by a simple right-hand rule. When the thumb of the right hand is pointed in the direction of current flow, the fingers curl in the direction of the surrounding magnetic field. This is true not only for currents flowing within power lines but also for all bioelectric currents, such as those flowing within neurons. Biomagnetic fields directly reflect electrophysiological events of the brain and pass through the skull without distortion. Hence, currents initiated at the synapses, and guided postsynaptically by cell structure, produce the magnetic field detectable outside the head. Magnetic field lines encircle the flow path of this primary current, extending outside the skull. Because pyramidal cells are predominantly oriented perpendicular to the cortex, the direction of the primary current is also perpendicular to the cortex. MEG is therefore most sensitive to activity in the fissural cortex, where the current is oriented parallel to the skull, whereas it does not detect sources that are oriented exactly radially to the skull. Because MEG detects only the tangential component of the primary current, amplitude comparison between differently oriented sources is possible only if source orientations can be estimated, e.g. on the basis of MRI.

The average electromagnetoencephalogram is about 10 picotesla (10^{-12}T) in amplitude, nine orders of magnitude smaller than the earth's steady magnetic field and seven orders of magnitude smaller than the magnetic fields generated by power lines, cars and elevators. The magnetic field produced by a single postsynaptic potential is too weak to be detected outside the head. Instead, what is detected is the macroscopic coherent activity of thousands of neurons. Still, cerebral magnetic fields are so weak that measurements are preferably performed inside magnetically shielded rooms. Sensitivity to such weak signals requires the use of cryogenic technologies. Current MEG instruments consist of an induction loop of niobium wire, which becomes superconducting (i.e. loses its resistivity to the flow of electrical currents) at approximately -258°C. Recording neuromagnetic signals has been compared to listening for the footsteps of an ant in the middle of a rock concert. Noise cancellation is improved by measuring gradients of the magnetic field instead of the field itself. In quiet environments, disturbances can be compensated for sufficiently by using elaborated compensation methods to enable MEG studies to be carried out even without shielded rooms.

Inverse Problem

As for EEG, MEG data have to be subjected to an inverse problem algorithm to obtain an estimate for the distribution of the activity in the brain. Similar to PET, fMRI and EEG, these can then be displayed on cross-sectional anatomical images (obtained by MRI) of the same subject. The inverse problem relates to the difficulty of determining internal sources on the basis of measurements performed outside the head. The most common way to tackle this problem is to determine the single-source current element (dipole) that explains most completely the EEG or MEG pattern. This can be done with a computer algorithm that starts from a random dipole position and orientation and keeps changing these parameters as long as the field pattern computed from the dipole keeps approaching the observed EEG or MEG pattern. When no further improvement is obtained, a minimum in the cost function has been reached; a source corresponding to this solution is called the equivalent current dipole (ECD). In most cases, however, the EEG or MEG data pattern cannot be explained accurately by a single source. In these cases, two or more dipoles could be used to explain the data, but this leads to computational difficulties in trying to determine the best multisource solution. Alternatively, continuous solutions, such as the minimum norm estimate, might also be constructed (Nenonen *et al.*, 1994). When interpreting EEG or MEG results, it should be borne in mind that the inverse problem is fundamentally non-unique. This means that even if the complete electric and magnetic field around the head could be measured precisely, an infinite number of current distributions in the brain could still be constructed that would explain the measured fields. It is always possible that some sources are missed, whatever the measurement set-up. For example, MEG alone is insensitive to radially oriented sources, but even when combined with EEG, silent sources are possible. Full use of available techniques requires the use of estimation theory to derive optimal solutions based on all available information, including MRI, PET and fMRI.

TRANSCRANIAL MAGNETIC STIMULATION

History

TMS is a tool for the non-invasive stimulation of the superficial cortex. In 1980, Merton and Morton surprised neuroscientists by

showing that it was possible to stimulate the motor areas of the human brain electrically through the intact scalp (transcranial electrical stimulation (TES)). They used a brief, high-voltage electric shock to activate the motor cortex and produce a relatively synchronous muscle response, the motor evoked potential (MEP). The problem was that TES was painful because of activation of pain fibres in the scalp. In 1985, Barker and colleagues showed that it was possible to stimulate both nerve and brain using external magnetic stimulation (TMS) with little or no pain. TMS is now used commonly in clinical neurology to study central motor conduction time. Depending on stimulation parameters, TMS can excite or inhibit the arbitrary sites of the superficial cortex, allowing functional mapping and creation of transient functional lesions (Hallett, 2000).

In neuropsychology, the classical paradigm is that of studying the effects of brain lesions on behaviour. With TMS, this paradigm can be applied in spatially and temporally restricted fashion to healthy volunteers. It is now used widely as a research tool to study aspects of human brain physiology, including motor function, vision, language, and the pathophysiology of brain disorders. Combined with other brain-imaging techniques, such as PET, EEG and fMRI, it can be used to evaluate cortical excitability and connectivity (Paus, 1999). It may also be useful in treating various neuropsychiatric disorders, most notably depression.

TMS is still a relatively young technique, and questions remain unanswered regarding its impact on brain function. Nonetheless, recent work with TMS has demonstrated that it allows the investigation of the relationship between focal cortical activity and behaviour to trace the timing at which activity in a particular cortical region contributes to a given task, and to map the functional connectivity between brain regions (Pascual-Leone *et al.*, 2000).

Basic Principles

A brief, high-current pulse is produced in a coil of wire, called the magnetic coil, placed above the scalp. A magnetic field is produced, with lines of flux passing perpendicular to the plane of the coil. An electric field is induced perpendicular to the magnetic field. Magnetic coils may have different shapes: round coils are relatively powerful; figure-of-eight-shaped coils are more focal, producing maximal current at the intersection of the two round components. The precise extent of neuronal activation is not known, but it varies with the intensity of stimulation. Ordinarily, TMS does not activate corticospinal neurons directly; rather, it activates them indirectly through synaptic inputs. Single-pulse TMS, which is very safe, has been used most commonly. Devices are now available that can deliver high-frequency (1–30 Hz), repetitive TMS (rTMS). This has greater effects than single-pulse TMS, but it also has the potential to cause seizures, even in normal individuals. Intracortical inhibition (ICI) and intracortical facilitation (ICF) are obtained using paired-pulse studies, and reflect the activity of interneurons in the cortex. Safety guidelines have been published that should prevent problems (Wassermann, 1998).

rTMS can produce effects that last after the stimulation period. The mechanisms of these changes are not clear, but the analogies to long-term potentiation (LTP) and long-term depression (LTD) of individual synapses in the central nervous system are apparent. Therapeutic effectiveness depends on the exact site of stimulation, intensity, and the precise pattern of pulses, including rate, train length, intertrain interval, and number of trains. This is clearly difficult to clarify and is currently an active area of psychiatric research (George *et al.*, 1999).

MAGNETIC RESONANCE SPECTROSCOPY

History

The fundamentals of NMR spectroscopy were studied and developed at least 30 years before being extended to clinical imaging. During this time, it was a major tool in physical and organic chemistry for determining molecular structure. As MRS technology evolved to allow the study of larger samples at higher field strengths, a logical step was to compare the spectral characteristics of normal and pathological tissue specimens. Following the discovery that normal and cancerous tissue samples had different NMR signals (Damadian, 1971), the first clinical NMR scanning machine was patented in the early 1970s. Hence, MRS has been in use much longer than clinical proton MRI. MRI is generally associated with the signals from hydrogen nuclei (i.e. protons) because of the large amounts of hydrogen atoms in human tissue and brain and the strong signals they provide. However, MRS makes measurements not only of protons but also of nuclei, such as phosphorus (^{31}P), carbon (^{13}C) and fluorine (^{19}F) (Dacey *et al.*, 1991). MRS offers the potential of assessing brain function at metabolic and molecular levels. At present, much of the work in this area is experimental.

Basic Principles

This technique uses natural emissions from atomic nuclei activated by magnetic fields to measure the concentration of endogenous molecules. Potential nuclei include ^{31}P, ^{13}C, ^{23}Na and ^{7}Li, in addition to ^{1}H. The ^{31}P magnetic resonance spectrum can detect tissue concentrations of the phosphomonoesters phosphocholine and inorganic orthophosphate, the phosphodiesters glycerol-3-phosphoethanolamine and glycerol-3-phosphocholine, the triphosphate adenosine triphosphate (ATP), and other phosphorus-containing molecules, including phosphocreatinine. ^{1}H spectroscopy offers the ability to measure lactate concentrations and neuronal markers such as *N*-acetyl aspartate. MRS permits quantitative analysis of these compounds *in vivo*, with the potential of three-dimensional resolution within the brain.

NEAR-INFRARED SPECTROSCOPY

History

NIRS and event-related optical signals (EROS) are relatively new methods to measure *in-vivo*• changes in cerebral haemodynamics and oxygenation. Changes in the intrinsic optical properties of the tissue are associated with changes in the level of physiological activity in neuronal tissue. As a consequence, it is possible to optically monitor neuronal activity without the use of dyes or other contrast-enhancing agents. Such optical techniques have been applied in the laboratory for more than 50 years. In previous studies of exposed brain tissue, optical imaging of brain activity has been achieved at high temporal and microscopical spatial resolution. Now, using near-infrared light, which can penetrate biological tissue reasonably well, it has become possible to assess brain activity in human subjects through the intact skull non-invasively. Recent developments in NIRS and intraoperative optical imaging have suggested a number of clinically important applications of this technology. After early studies employing single-site NIRS, first near-infrared imaging devices are being applied successfully for low-resolution functional brain imaging (Villringer and Chance, 1997). Q1

Basic Principles

Brain activity is associated with changes in optical properties of brain tissue. Optical measurements during brain activation can assess haemoglobin oxygenation, cytochrome-c-oxidase redox state, and two types of changes in light scattering, reflecting either membrane potential (fast signal) or cell swelling (slow signal). The physiological basis of the most common form of NIRS is wavelength-specific absorption of photons by oxygenated and deoxygenated haemoglobin. Thus, the contrast mechanism for NIRS signals is related closely to that of intrinsic optical imaging of exposed cortex using visible light. The much lower baseline absorption levels at the longer wavelengths used in NIRS allow the light to travel further through skin, skull and brain tissue, thus allowing non-invasive imaging of haemodynamics, albeit with lower spatial resolution. Although it is possible to sample optical signals quite rapidly (>1 kHz), the effective temporal resolution is limited by the indirect nature of the coupling of the haemodynamic processes affecting the optical signals and the underlying neuronal electrical activity. However, there is some evidence that it may be possible to detect optical signals related more directly to neuronal activation (i.e. EROS). The physiological basis of EROS is not well understood, but it may include cell swelling or membrane polarization associated with neuronal activity, resulting in local light-scattering changes. Thus, optical imaging may provide insights into both the electrophysiological (fast) and haemodynamic (slow) processes underlying other brain-imaging signals. However, the spatial resolution afforded by optical methods alone is limited by the diffuse nature of photon transport through tissues.

Although promising, the application of NIRS in functional brain imaging is in an early experimental state. Advantages of the optical methods include biochemical specificity, a temporal resolution in the millisecond range, the potential of measuring intracellular and intravascular events simultaneously, and the portability of the devices enabling bedside examinations. Caveats of cerebral NIRS include insufficient light shielding, optode displacement, and a sample volume including muscle or the frontal sinus mucous membrane.

FUNCTIONAL NEUROIMAGING STUDY DESIGN

History

Mapping the human brain is distinct from the assumptions held by phrenologists of the nineteenth century. According to the German physician Franz Josef Gall, thought processes are localized in single brain areas identified by bumps on the skull. Gall posited that complex behavioural traits (e.g. ideality, cautiousness, imitation, self-esteem, calculation) could be related to the size of these bumps. Although the 'bumps theory' was fanciful, the idea of a functional segregation of the brain was not. In 1861, by carefully studying the brain of a man who had lost the faculty of speech after a left inferior frontal lesion, Paul Broca became convinced that different functions could be localized in different parts of the cerebrum. Now, more than a century of neuropsychological investigations in brain-damaged patients has confirmed that a cortical area can be specialized for some aspects of perceptual or sensorimotor processing, and that this specialization is segregated anatomically in the cortex. In our current vision on brain function, however, functional segregation holds for simple processes rather than for complex behaviours or traits, such as those described by phrenologists. Now, the view is that the cortical infrastructure supporting a single function (and a fortiori a complex behaviour) may involve many specialized areas that combine resources by functional integration between them. Hence, functional integration is mediated by the interactions between functionally segregated areas, and functional segregation is meaningful only in the context of functional integration, and vice versa.

In this framework, the foundation for most functional neuroimaging studies is that complex behaviors can be broken down into a set of constituent mental operations. In order to read this book, for example, you must recognize that a string of letters is a word; then recognize the meanings of words, phrases and sentences; and finally create mental images. The methodological challenge is first to separate each of these tasks from a cognitive perspective, and second to determine those parts of the brain that are active and those that are dormant during their performance. In the past, cognitive neuroscientists have relied on studies of laboratory animals and patients with localized brain lesions to gain insight into the brain's functions. Imaging techniques, however, permit us to visualize, safely, the anatomy and the function of the human brain in both normal and pathological conditions.

It is amazing that the strategy used most widely for functional neuroimaging over the past 15 years is based on an idea first introduced to psychology in 1868. Indeed, Franciscus C. Donders, a Dutch ophthalmologist and physiologist, then proposed a general method to estimate cognitive processes based on a simple logic. He subtracted the time needed to respond to a light (with, say, a press of a key) from the time needed to respond to a particular colour of light. He found that discriminating colour required about 50 ms more than simply responding to the light. In this way, Donders (1969) was the first to isolate a basic mental process and to obtain a measure of the time needed by the brain to perform this specific process.

Basic Approaches

The current strategy in functional neuroimaging is designed to accomplish a similar subtraction but in terms of the brain areas implementing the mental process. In particular, images of neural activity (be it blood flow measured by PET or fMRI or electrical activity measured by EEG or MEG) taken before a task is begun can be compared with those obtained when the brain is engaged in that task. The two periods are referred to as the control state and the task state. It is important to choose carefully each state so as to isolate as best as possible a limited number of operations. Subtracting neural activity measurements made in the control state from each task indicates those parts of the brain active during a particular task. To achieve reliable data, averages are made of many experimental trials in the same person (e.g. fMRI) or of responses across many individual subjects (e.g. PET). Averaging enables the detection of changes in neural activity associated with mental activity that would otherwise be confused easily with spurious shifts resulting from noise.

It is important to stress that this methodological approach, known as the cognitive subtraction paradigm, has an important drawback. Indeed, in order to isolate the neural substrate of a given cognitive component of interest, it must be assumed that the only difference between the control state and the task state is the component of interest to the exception of any other stimulus- or task-related processes. Unfortunately, this cannot always be guaranteed easily and fully. Analytic strategies, however, have been devised to circumvent this problem (see below), and cognitive subtraction designs remain the foundation of most functional neuroimaging experiments.

ANALYSING BRAIN IMAGING DATA

History

Regional differences among brain scans have long been characterized thanks to hand-drawn regions of interest (ROIs). This approach reduced the information from hundreds of thousands of voxels (volume elements that in three dimensions correspond to a pixel with a given slice thickness) to a handful of ROI measurements, with a somewhat imprecise anatomical validity. The development of more powerful voxel-based statistical methods has made these ROI analyses obsolete. Although several solutions are in use in neuroscience laboratories, one of the most popular methods for the analysis of neuroimaging data is statistical parametric mapping (SPM). This is a standardized method that refers to the construction and assessment of spatially extended statistical processes used to test hypotheses about neuroimaging data (mainly PET, SPECT and fMRI). Statistical parametric maps can be thought of as 'X-rays' of the significance of an effect, which can be projected on a three-dimensional representation of the brain. These ideas have been instantiated in software (version SPM99 at time of going to press) by Karl Friston and coworkers at the Wellcome Department of Cognitive Neurology in London (http//www.fil.ion.ucl.ac.uk/spm). SPM has become the most widely used and validated method for analysing functional neuroimaging data. Since its first description in 1990, over 1500 citations now refer to its use.

Basic Approaches

As described above, there are two basic approaches when analysing and interpreting functional neuroimaging data. They are based on the distinction between functional segregation and integration.

Functional Segregation

Using a functional specialization concept of the brain, the following sets of approaches are based on detecting focal differences. They generally fall into one of three broad categories: (1) The subtractive or categorical designs are predicted on the assumption that the difference between two tasks can be formulated as a separable cognitive or sensorimotor component, and that the regionally specific differences in brain activity identify the corresponding functional area (i.e. the cognitive subtraction paradigm). Its utilization ranges from the functional anatomy of word processing to the functional specialization in visual cortex, an application that has been validated by electrophysiological studies in monkeys (Zeki, 1993). (2) The parametric or dimensional design assumes that regional physiology will vary systematically with the degree of cognitive or sensorimotor processing. Parametric designs may avoid many of the shortcomings of 'cognitive subtraction'. A fundamental difference between subtractive and parametric designs lies in treating a cognitive process not as a categorical invariant but as a dimension that can be expressed to a greater or lesser extent in relation to the brain's regional activity. (3) Factorial or interaction designs are also well suited to avoiding the drawbacks of simple subtraction paradigms. Two or more factors can be combined in the same experiment, and the interaction term will assess the effect of one factor while excluding the effect of the other.

Functional Integration

The functional role played by any component (e.g. a neuron or a specific brain area) of a connected system (e.g. the brain) is defined largely by its connections. Connectionist approaches to understanding the integration of brain functions are well established (Hebb, 1964). The nature and organizational principles of intracortical (Goldman-Rakic, 1988) and subcortical (Mesulam, 1990) connections have provided a basis for mechanistic descriptions of brain function, referring to parallel, massively distributed and interconnected (sub)cortical areas. Anatomical connectivity, determined mainly by neuroanatomic tracer experiments in animals, is a necessary underpinning for these models. The concepts of functional and effective connectivity were developed in the analysis of separable spike trains obtained from multi-unit electrode recordings. However, the neurophysiological measurements obtained from functional neuroimaging have a very different timescale (seconds versus milliseconds) and nature (metabolic or haemodynamic versus spike trains) than those obtained from electrophysiological studies.

Only recently have analytical tools become available to assess the functional or effective connectivity between distant cerebral areas (Friston *et al.*, 1997). Functional connectivity is defined as the temporal correlation of a neurophysiological index (e.g. blood flow) measured in different remote brain areas, whereas effective connectivity is defined as the influence that one neural system exerts over another (Buchel and Friston, 1997). In this context, a psychophysiological interaction can be assessed in the framework of the general linear model, as employed by SPM (Friston *et al.*, 1997), to explain the activity in one cortical area in terms of an interaction between the influences of another area in a given experimental context. Put simply, the statistical analysis will identify brain regions that show condition-dependent differences in the way their activity relates to the activity in another (chosen) area.

Preprocessing the Data

Voxel-based analyses require the data to be in the same anatomical space. This is obtained by realigning the data. Indeed, in functional neuroimaging experiments, movement-related variance components represent one of the most serious confounds. Therefore, scans from each subject are realigned using an optimization procedure minimizing the residual sum of squares (Friston *et al.*, 1995).

In a second step, the realigned images are normalized. They are subject to nonlinear warping so that they match a template that already conforms to a standard anatomical space approximating those described by Talairach and Tournoux (1988). Pooling neuroimaging data from grossly different individual brains requires a procedure to spatially normalize the individual brains to an idealized or standard brain for the purpose of achieving overlap between corresponding anatomical and functional areas in different subjects. The Talairach and Tournoux atlas was initially developed — and has proven very useful — for anatomical normalization required for neurosurgical procedures, particularly those at brain sites close to the origin of the reference system (e.g. the anterior and posterior commisures). Each point within Talairach space into which brains are transformed is defined using three coordinates (expressed in millimetres). The first coordinate defines the position in x, i.e. from left (negative) to right (positive), with 0 mm corresponding to the interhemispheric line. The second coordinate defines the position in y, i.e. from posterior (negative) to anterior (positive), with 0 mm corresponding to the anterior commisure. The third coordinate defines the position in z, i.e. from bottom (negative) to top (positive), with 0 mm corresponding to the plane through the anterior and posterior commisures. This standard coordinate system facilitates the reporting of results in a conventional way, and facilitates comparisons between peak voxels obtained in experiments from different laboratories.

After spatial normalization, images need to be smoothed (i.e. convolved with an isotropic Gaussian kernel). Smoothing individual

images before a statistical analysis offers (1) an improved signal-to-noise ratio; (2) conditioning of the data, so that they conform more closely to the Gaussian field model, which lies at the basis of the correction procedure for multiple statistical comparisons; and (3) a better overlap between the localization of anatomical and functional brain areas from different subjects, which permits intersubject averaging.

Statistical Analysis

The data obtained after preprocessing consist of a matrix of many hundred-thousandths of voxels for each subject and for each condition. Each of these voxels is characterized by the x, y and z spatial coordinates in the standard space and a value representing the functional activity in that voxel (e.g. blood flow, glucose metabolism, BOLD signal). The statistical analysis corresponds to modelling the data in order to partition observed neurophysiological states or responses into components of interest, confounds of no interest, and an error term. This partitioning is effected using the framework of the general linear model to estimate the components in terms of parameters associated with the design matrix. The analysis of regionally specific effects uses the general linear model to assess differences among parameter estimates (specified by a contrast) in a univariate sense, by referring to the error variance. The significance of each contrast is assessed with a statistic with a Student's t distribution under the null hypothesis for each and every voxel (i.e. SPM{t}). The SPM{t}is transformed to the unit normal distribution to give a Gaussian field or SPM{Z}.

Statistical Inference

The final stage is to make statistical inferences on the basis of the SPM and characterize the responses observed using the fitted responses or parameter estimates. On one hand, with an a priori anatomically constrained hypothesis about effects in a particular brain location, the Z value in that region in the SPM{Z} can be used to test the hypothesis (i.e. uncorrected P value). On the other hand, if an anatomical site cannot be predicted a priori, a correction for multiple non-independent comparisons is required. Therefore, the theory of Gaussian fields (Friston, 1997) provides a way for correcting the P value for the multiple non-independent comparisons implicit in the analysis. This correction depends on the search volume, the residual degrees of freedom due to error, and the final image smoothness estimate. The obtained corrected and uncorrected P values pertain to different levels of inference in terms of (1) the significance of the effect in a particular voxel, (2) the significance of the coactivation of a cluster of voxels in a specific region, and (3) the significance of the coactivation of several clusters in the whole brain. Importantly, only in cases of well-documented prior neuroanatomical knowledge about the expected result can uncorrected P values be accepted. By specifying different contrasts, one can test for the variety of effects described above, and the significance values above a chosen threshold are represented comprehensively in an SPM map, where each voxel is represented at its proper location on the brain template, and where the T value in this voxel for a given contrast is represented by use of a colour-intensity code.

FUTURE DIRECTIONS: MULTIMODALITY INTEGRATION

As discussed above, fMRI and $H_2{}^{15}O$-PET measure local changes in brain haemodynamics induced by cognitive or perceptual tasks. These measures have a uniformly high spatial resolution of millimetres or less but poor temporal resolution (about 1 s at best). Conversely, EEG and MEG measure instantaneously the current flows induced by synaptic activity, but the accurate localization of these current flows remains an unsolved problem. Recently, techniques have been developed that, in the context of brain anatomy visualized with structural MRI, use both haemodynamic and electromagnetic measures to get estimates of brain activation with higher spatial and temporal resolution. These methods range from simple juxtaposition to simultaneous integrated techniques. However, further advances in multimodality integration will require an improved understanding of the coupling between the physiological phenomena underlying the different signal modalities (Dale and Halgren, 2001).

The combination of TMS with PET or EEG permits the assessment of connectivity and excitability of the human cerebral cortex. PET and fMRI, working in a combination yet to be determined, can define the anatomy of the circuits underlying a behaviour of interest; electrical recording techniques can reveal the course of temporal events in these spatially defined circuits. Parallel information from different imaging modalities is beginning to be used to constrain the EEG or MEG inverse solutions discussed above to limited regions of the cerebrum. This approach provides optimal combined spatial and temporal resolution by exploiting the best aspects of each technology. Combining various techniques offers a more complete characterization of the different aspects of brain activity during cognitive processing. This is even more so regarding our understanding of transitory neuropsychiatric phenomena (e.g. single hallucinations).

Regardless of the particular mix of technologies that will ultimately be used to image human brain function, the field demands extraordinary resources. MRI, PET and MEG equipment cost from €1–4 million, require large teams to be run, and are expensive to maintain. Furthermore, success requires close collaboration within multidisciplinary teams of scientists gathering knowledge and expertise in basic sciences (physics, informatics, chemistry, mathematics, etc.) and neuroscience (neurology, psychiatry, psychology, neuropharmacology, neurobiology, etc.). Few institutions are fortunate enough to have the technical and human resources necessary to master these technologies. Such institutions should optimally make them available to all expert scientists with valid questions in order to increase our knowledge of all aspects of human mental activity. In contrast, EEG data can be collected with a modest amount of compact, lightweight and easy-to-use equipment. Clinical EEG recordings from ambulatory subjects have been made routinely for many years. In the future, then, it appears likely that EEG will become an important tool for studying the neurophysiology of cognition outside the laboratory in naturalistic settings.

Functional neuroimaging experiments provide a vast amount of information. Recent efforts to create neuroscience databases could organize and disseminate quickly such a repository of data. As demonstrated in many chapters of this book, wise use of these powerful tools and the information they produce can aid our understanding and management of many neuropsychiatric diseases. Clearly, neuroimaging is heading us towards a much richer grasp of the relation between the human mind and the brain.

ACKNOWLEDGEMENTS

This work was supported by grants from the Fonds National de la Recherche Scientifique de Belgique (FNRS), the Special Funds for Scientific Research of the University of Liège and by the Queen Elizabeth Medical Foundation. SL is postdoctoral researcher at the FNRS.

REFERENCES

Baule, G.M. and McFee, R., 1963. Detection of the magnetic field of the heart. *Am Heart J* 66, 95–96.

Brenner, D., Williamson, S.J. and Kaufman, L., 1975. Visually evoked magnetic fields of the human brain. *Science* 190, 480–482.

Buchel, C. and Friston, K.J., 1997. Modulation of connectivity in visual pathways by attention: cortical interactions evaluated with structural equation modelling and fMRI. *Cereb Cortex* 7, 768–778.

Cohen, D., 1968. Magnetoencephalography: evidence of magnetic fields produced by alpha-rhythm currents. *Science* 161, 784–786.

Dacey, R., Dikmen, S., Temkin, N., *et al.*, 1991. Relative effects of brain and non-brain injuries on neuropsychological and psychosocial outcome. *J Trauma* 31, 217–222.

Dale, A.M. and Halgren, E., 2001. Spatiotemporal mapping of brain activity by integration of multiple imaging modalities. *Curr Opin Neurobiol* 11, 202–208.

Damadian, R., 1971. Tumor detection by nuclear magnetic resonance. *Science* 171, 1151–1153.

Donders, F.C., 1969. On the speed of mental processes [translation]. *Acta Psychol* 30, 412–431.

Friston, K.J., 1997. Analysing brain images: principles and overview. In: Frackowiak, R.S.J., Friston, K.J., Frith, C.D., Dolan, R.J. and Mazziotta, J.C. (eds), *Human Brain Function*, pp. 25–41. Academic Press, San Diego.

Friston, K., Ashburner, J., Frith, C., *et al.*, 1995. Spatial realignment and normalization of images. *Hum Brain Mapp* 2, 165–189.

Friston, K.J., Buechel, C., Fink, G.R., *et al.*, 1997. Psychophysiological and modulatory interactions in neuroimaging. *Neuroimage* 6, 218–229.

Fulton, J.F., 1928. Observations upon the vascularity of the human occipital lobe during visual activity. *Brain* 51, 310–320.

George, M.S., Lisanby, S.H. and Sackeim, H.A., 1999. Transcranial magnetic stimulation: applications in neuropsychiatry. *Arch Gen Psychiatry* 56, 300–311.

Gevins, A., 1998. The future of electroencephalography in assessing neurocognitive functioning. *Electroencephalogr Clin Neurophysiol* 106, 165–172.

Gibbs, F.A., Davis, H. and Lennox, W.G., 1935. The electroencephalogram in epilepsy and in conditions of impaired consciousness. *Arch Neurol Psychiatry* 35, 1133–1148.

Goldman-Rakic, P.S., 1988. Topography of cognition: parallel distributed networks in primate association cortex. *Annu Rev Neurosci* 11, 137–156.

Groch, M.W. and Erwin, W.D., 2001. Single-photon emission computed tomography in the year 2001: instrumentation and quality control. *J Nucl Med Technol* 29, 12–18.

Hallett, M., 2000. Transcranial magnetic stimulation and the human brain. *Nature* 406, 147–150.

Hebb, D.O., 1964. *Organisation of Behavior*. Wiley, New York.

Huang, S.C., Phelps, M.E., Hoffman, E.J., Sideris, K. and Kuhl, D.E., 1980. Noninvasive determination of local cerebral metabolic rate of glucose in man. *Am J Physiol* 238, 69–82.

Magistretti, P.J. and Pellerin, L., 1999. Cellular mechanisms of brain energy metabolism and their relevance to functional brain imaging. *Philos Trans R Soc Lond B Biol Sci* 354, 1155–1163.

Mesulam, M.M., 1990. Large-scale neurocognitive networks and distributed processing for attention, language, and memory. *Ann Neurol* 28, 597–613.

Mosso, A., 1881. *Ueber den Kreislauf des Blutes in Menschlichen Gehirn*, pp. 66–67. Verlag von Viet and Company, Leipzig.

Naatanen, R., Ilmoniemi, R.J. and Alho, K., 1994. Magnetoencephalography in studies of human cognitive brain function. *Trends Neurosci* 17, 389–395.

Nenonen, J.T., Hamalainen, M.S. and Ilmoniemi, R.J., 1994. Minimum-norm estimation in a boundary-element torso model. *Med Biol Eng Comput* 32, 43–48.

Ogawa, S., Lee, T.M., Kay, A.R. and Tank, D.W., 1990. Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proc Natl Acad Sci USA* 87, 9868–9872.

Pascual-Leone, A., Walsh, V. and Rothwell, J., 2000. Transcranial magnetic stimulation in cognitive neuroscience–virtual lesion, chronometry, and functional connectivity. *Curr Opin Neurobiol* 10, 232–237.

Paus, T., 1999. Imaging the brain before, during, and after transcranial magnetic stimulation. *Neuropsychologia* 37, 219–224.

Phelps, M.E., 2000. Inaugural article: positron emission tomography provides molecular imaging of biological processes. *Proc Natl Acad Sci USA* 97, 9226–9233.

Posner, M.I. and Raichle, M.E., 1994. Images of the brain. In: Posner, M.I. and Raichle, M.E., *Images of Mind*, pp. 53–81. Scientific American Library, New York.

Pykett, I.L., 1982. NMR imaging in medicine. *Sci Am* 246, 78–88.

Ribary, U., Ioannides, A.A., *et al.*, 1991. Magnetic field tomography of coherent thalamocortical 40-Hz oscillations in humans. *Proc Natl Acad Sci USA* 88, 11 037–11 041.

Talairach, J. and Tournoux, P., 1988. *Co-Planar Stereotaxis Atlas of the Human Brain*. Georges Thieme Verlag, Stuttgart.

Ter-Pogossian, M.M., Raichle, M.E. and Sobel, B.E., 1980. Positron-emission tomography. *Sci Am* 243, 170–181.

Villringer, A. and Chance, B., 1997. Non-invasive optical spectroscopy and imaging of human brain function. *Trends Neurosci* 20, 435–442.

Wassermann, E.M., 1998. Risk and safety of repetitive transcranial magnetic stimulation: report and suggested guidelines from the International Workshop on the Safety of Repetitive Transcranial Magnetic Stimulation, June 5–7, 1996. *Electroencephalogr Clin Neurophysiol* 108, 1–16.

Zeki, S., 1993. *A vision of the Brain*. Blackwell Scientific Publications, Oxford.