What does brain damage tell us about the mechanisms of sleep?

B M Evans FRCP

Much is now known about the factors that influence the timing of sleep onset, such as the circadian rhythm and the length of time since a previous sleep[1]. There is also information about hormonal factors that influence sleeping and waking and the cerebral areas associated with them[2]. Less is known about the nerve fibres and cells involved in transferring the function of the brain and the body from a state of wakefulness to sleep and back again at least every twenty-four hours. This paper outlines what is known about these neural processes and suggests that information about them can be derived from the normal brain and also the damaged brain.

WAKEFULNESS, RAPID EYE MOVEMENT (REM) SLEEP AND NON-REM SLEEP

Wakefulness is characterized by a state of arousal with an activated cerebral cortex, high cerebral blood-flow and glucose metabolism, and fast activity in the electroence-phalogram (EEG); autonomic activity and muscular tone are also high. Non-REM sleep begins at sleep onset. The cortex is progressively deactivated as slow waves appear in the EEG, cerebral blood-flow and glucose metabolism decline and autonomic activity and muscular tone decrease. This process culminates in deep sleep, sometimes called stage 4 or slow-wave sleep from the EEG appearance. During deep sleep the heart rate and blood pressure are low and steady and the muscles are relaxed but not paralysed. Non-REM sleep comprises deep sleep and the intermediate states between it and wakefulness; it makes up about 80–85% of total sleep time, and is interrupted every 60–90 minutes by episodes of REM sleep. In these episodes the cortex is highly activated with fast EEG activity, enhanced cerebral metabolism and raised blood pressure and heart rate; the voluntary muscles are paralysed apart from the eye muscles, which cause the jerky eye movements that give this sleep its name. REM sleep makes up 15–20% of a total night's sleep and is associated with dreaming. REM and non-REM sleep are controlled by separate brainstem mechanisms. This paper discusses only the mechanisms of non-REM sleep, the predominant sleep type.

Department of Clinical Neurophysiology, Mapother House, King's College Hospital, Denmark Hill, London SE5 0RS, UK
E-mail: bidievans@theorchard.swinternet.co.uk

WHAT DO WE KNOW ABOUT THE MECHANISMS THAT CONTROL NON-REM SLEEP?

In 1949 a reticular activating mechanism (RAM) was described that switches the brain from deep sleep to wakefulness[3]. The RAM arises in the brainstem reticular grey matter, a collection of nuclei surrounding the central
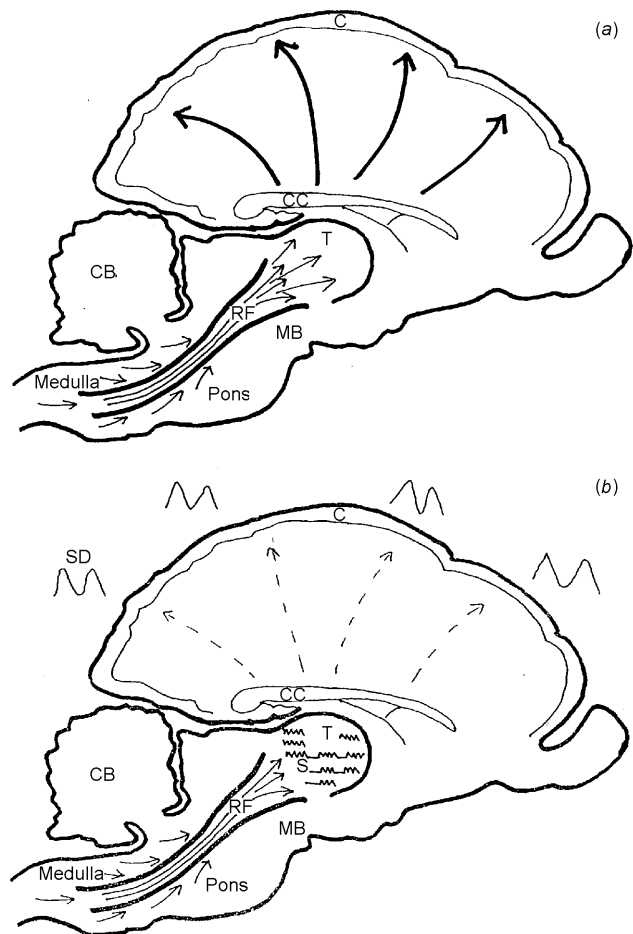

Figure 1 **Arousal and arousal inhibitory mechanisms of brainstem (vertical section of monkey brain).** C=cortex; CC=corpus callosum; CB=cerebellum; MB=mid-brain; T=thalamus; RF=reticular formation. (*a*) Reticular activating mechanism. Arrows show sensory input from spinal cord and mid-brain into reticular substance of mid-brain nuclei, which then form relays in the thalamus. Fibres of the thalamo-cortical radiation reach all areas of the cortex. (*b*) Arousal inhibitory mechanism. Reticular formation input remains the same as in (*a*) but the upward flow into the thalamus is partly blocked in relation to the thalamic sleep spindles (S). The cortex shows slow and very slow activity (SD). The slowest waves are related to the sleep spindle
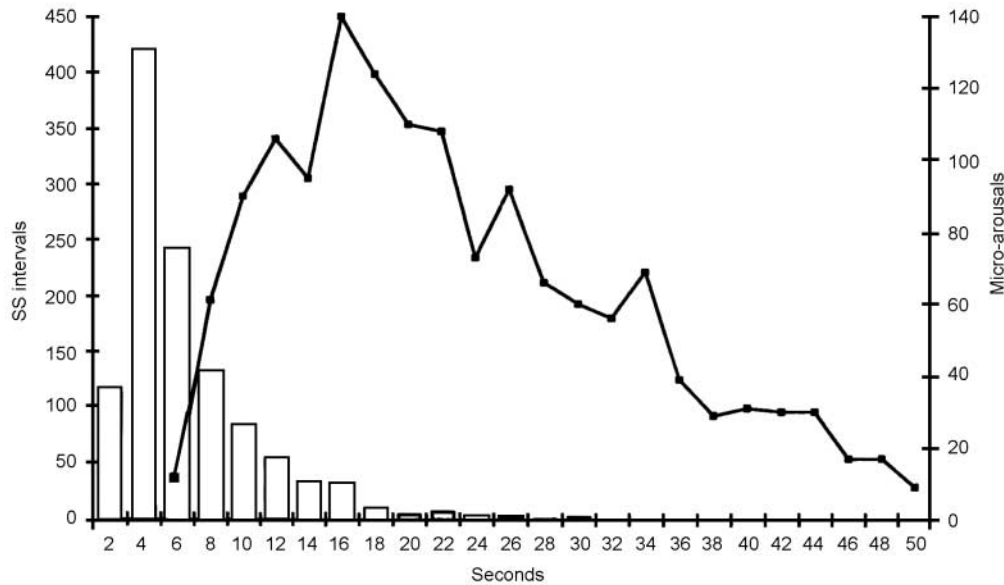
Figure 2 **Relation between sleep spindles (ss) and cyclical alternating pattern in early sleep.** The bar graph shows the intervals between the onsets of 1148 sleep spindles; pooled data from 32 individuals in stage 2–3 sleep. The preferred interval is between 3 and 5 s. The line graph shows the interval between the onsets of 1463 episodes of higher arousal activity in stage 1 sleep; pooled data from 52 individuals. The preferred interval is about 16 s with subpeaks at shorter and longer intervals (the subpeaks were consistent during collection of the data). The intervals between the subpeaks are at about 4 s—i.e. similar to the intervals between sleep spindles

canal. These nuclei receive inputs from all the ascending tracts of the spinal cord and brainstem and then form relays in the thalamus. From the thalamus the RAM connects with the cerebral cortex by means of a cortico-thalamic network of nerve cells and fibres (Figure 1). The RAM is the 'wake-up' mechanism of the brain and mediates the change from sleep to wakefulness. It is also responsible for periods of increased arousal during non-REM sleep.

The opposite process, from wakefulness to sleep, is not well understood. It is often thought to be passive, related to the withdrawal of sensory input, but recent work both in experimental electrophysiology and with positron emission tomography indicates an active process that can be thought of as an arousal inhibitory mechanism. First suggested in 1993[4], this seems to consist mainly of an intermittent gating mechanism in the thalamus related to the sleep spindle in the EEG. The sleep spindle, a brief burst of fast activity, is seen in the scalp EEG in the early stages of non-REM sleep and is present in the thalamus from the onset of drowsiness throughout the whole of non-REM sleep. The effect of the sleep spindle on the RAM has been demonstrated electrically by Steriade and his co-workers in naturally sleeping cats[5]. They showed that each thalamic sleep spindle was associated with an intermittent interruption in the up-flow of brainstem activity. The spindle occurs regularly at intervals of 3–5 s and whilst it is present in the thalamus the brainstem activity becomes intermittent at the same frequency. The same 3–5 s interval has been observed between the sleep spindles in man[6] (Figure 2). More recent research from Steriade's group has shown that the slowest

waves from the cerebral cortex in deep sleep are also related to the sleep spindle[7]. Thus two factors, cortical and thalamic, combine to produce a mechanism that prevents activation of the cortex by the RAM during slow wave sleep (Figure 1*b*). Because the thalamic gating is intermittent strong stimuli can still reach the cortex and restore wakefulness, which explains why the cry of a baby will always wake its mother from sleep.

Considerable support for the thalamic inhibitory mechanism has come from positron emission tomography. Maquet and several other workers[8–11] have studied the differences in blood-flow in the sleeping and waking brain in REM and non-REM sleep. These show that the blood-flows in the thalamus, basal ganglia and upper mid brain during non-REM sleep are very low indeed, indicating deactivation at this level. Maquet considers that his findings provide support for those of Steriade.

CLINICAL EVIDENCE OF INTERMITTENT CHANGES IN AROUSAL AT SLEEP ONSET

Regular changes are seen at sleep onset in the normal human brain with alterations in the level of cerebral arousal[4,12–18] (Figure 3)—a process often referred to as the cyclical alternating pattern. The changes involve the level of arousal in the EEG, muscular tone, autonomic activity and also cerebral blood-flow[14] and cerebrospinal fluid pressure[14,19,20]. The changes are rhythmic, with varying frequencies between about 8 and 60s. They respond to stimulation, so that states of low arousal can be switched to high by simple stimuli such as touching or calling. By

contrast, periods of lower arousal develop spontaneously— an indication that the motive force is in this direction (i.e. inhibitory rather than activatory). After interruption by a stimulus, the inherent rhythm resumes[4,12]. As sleep progresses from stage 1 to stage 4 the changes become slower and slower, and in deep sleep no intermittent arousal activity is seen[4,12]. No intervals as short as the 3–5 s interval between the sleep spindles have been reported in any of the studies of cyclical alternating pattern; the shortest intervals described are in the early part of sleep[12]. When all the intervals between arousals in early sleep are pooled, a preferred interval length of 16 s is revealed (Figures 2 and 3) but there are other faster and slower preferred intervals that produce the subpeaks shown on the line graph in Figure 2. The intervals between these subpeaks are about 4 s, suggesting that there is a link between the sleep spindle in the thalamus and the cyclical alternating pattern in early sleep. It is probable that all the varying intervals between arousals in the cyclical alternating pattern are subharmonics of the interval between the sleep spindles.

FINDINGS IN THE DAMAGED BRAIN THAT RESEMBLE THE CYCLICAL ALTERNATING PATTERN

The first evidence of intermittent episodes of high and low arousal came from studies of the damaged brain[21] and such episodes have since been reported in coma of many different causes[12,13,22,23]. These arousal changes are accompanied by simultaneous changes in the EEG, the autonomic system[4,12] (Figures 4 and 5), cerebrospinal fluid pressure[19,20], and somatic motor activity[12]. Compared with the intervals between the sleep spindles and even the arousal alternations of the cyclical alternating pattern the changes in activation in the damaged brain are very slow, lasting many seconds or minutes, and are not always rhythmic (Figure 5). They are also often more extreme than in the normal situation and may even be life-threatening, especially those involving the cerebrospinal fluid pressure[19]. The contrast between the excessive changes seen in the damaged brain and the modest ones in the normal brain suggests a process in which control mechanisms have been lost or damaged.
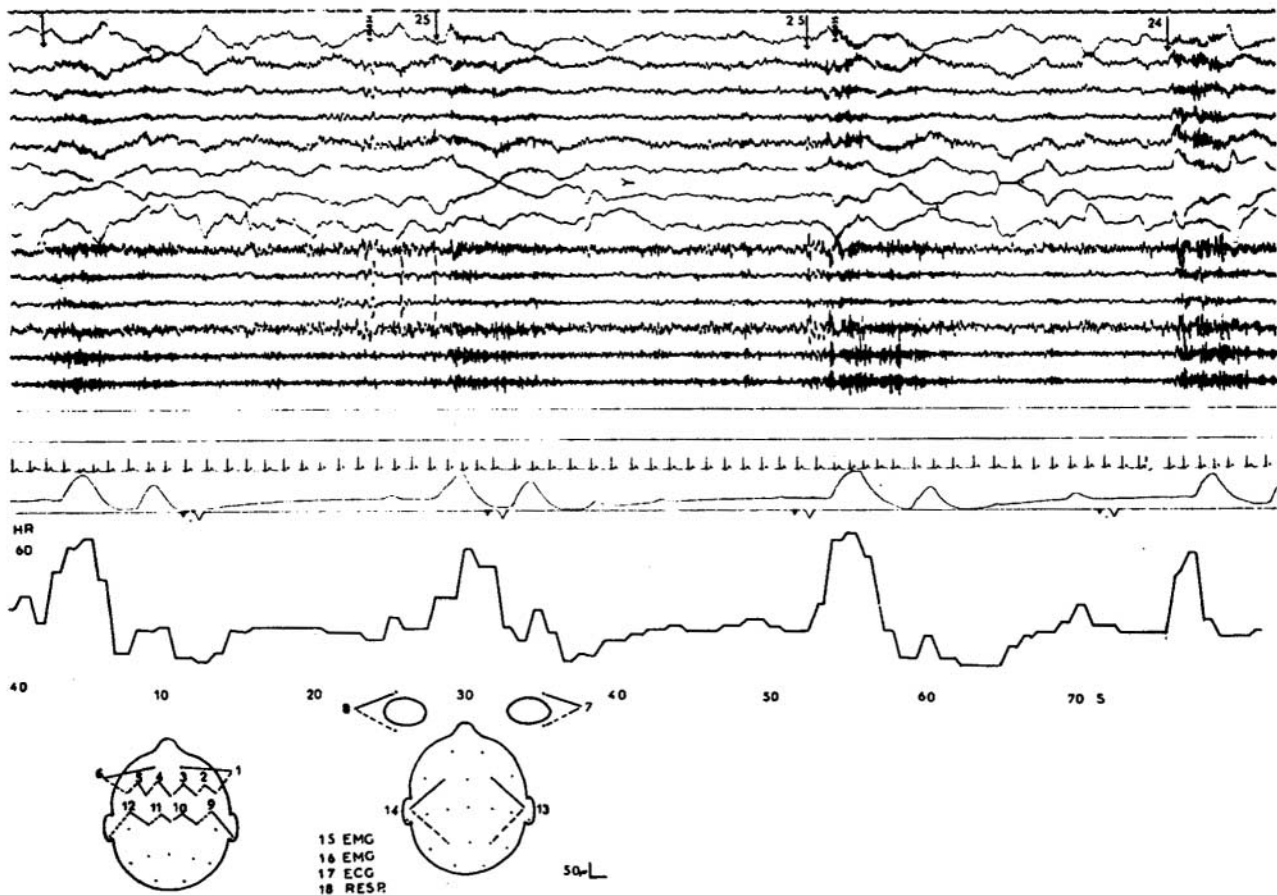

Figure 3 **Polygraphic record at 1.5 cm/s showing intermittent arousal activity of early sleep (stage 1).** The upper 14 traces show the EEG with alternating periods of faster activity (dark areas associated with higher arousal). The intervals between the onsets of the higher arousal periods are shown by the figures at the top. (These are the intervals used in the line graph of Figure 3.) Channel 17 is an electrocardiogram; channel 18 is respiration showing periodic apnoea, each group of breaths associated with higher arousal. Below is a heart rate graph showing intermittent heart rate increases with each arousal
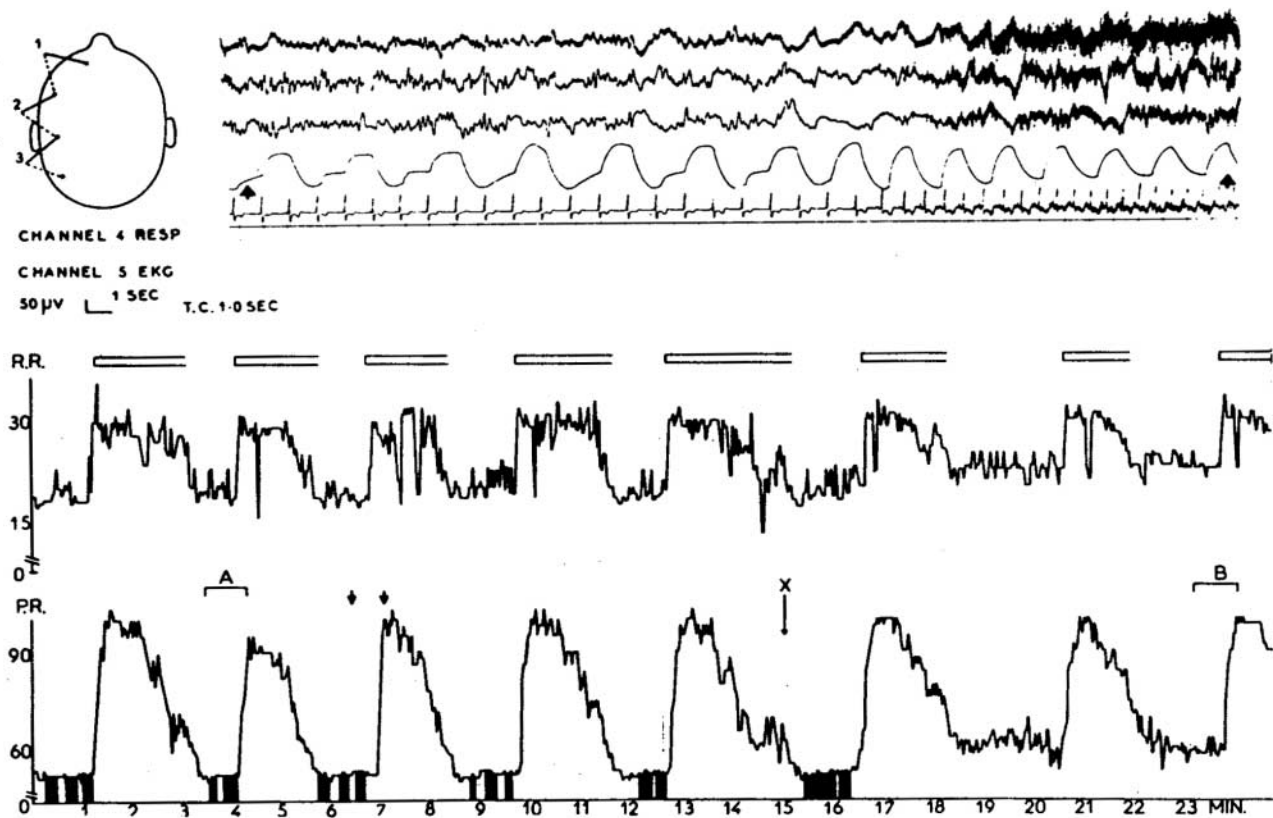
Figure 4 **Recordings from a patient aged 66 with subarachnoid haemorrhage from anterior communicating aneurysm**. *Above*: channels 1–3 show EEG at 1.5 cm/s. Low voltage slower activity is followed, at the time of a spontaneous arousal, by flattening of the EEG and then by slow waves with muscle artefact. Channel 4, respiration from a thermistor in the nose, channel 5, electrocardiogram. Note marked increase in respiratory and heart rates at the moment of arousal and a wandering pacemaker during the period of lower arousal. *Below*: graphs of respiration rate (R.R.) and heart rate plotted at 1 s intervals. This shows that arousal events are occurring about every 4 min. Open blocks above the respiration graph illustrate periods of higher arousal in the EEG. Filled areas below the heart rate graph illustrate periods when wandering pacemaker was evident. Short arrows show the section of the cardiorespiratory graphs associated with the EEG polygraph above. X marks a stimulus

EFFECTS OF BRAIN DAMAGE ON MECHANISMS OF SLEEP

Deep non-REM sleep can be regarded as a state of physiological reversible unconsciousness. Many forms of brain damage result in disturbance of consciousness. In this paper my working hypothesis is that coma or stupor is the result of interference with the non-REM sleep system by some types of brain damage. The non-REM sleep system can be considered at three levels—an upper level, consisting of the cerebral cortex and hemisphere white matter; a middle level, involving the thalamus and upper brainstem; and a lower level, comprising the lower mid-brain and pons. Brain damage at any or all these levels can produce coma or in less serious cases stupor. Most coma-producing conditions are capable of damaging any level individually or two or three levels together but some conditions commonly damage one level rather than the others.

Level 1: the cerebral hemispheres

Level 1 damage can affect the cerebral cortex or hemisphere white matter[24] or both together. Malfunction here can make it difficult or impossible for the RAM to activate the cortex, even though the lower levels of the sleep mechanism in the thalamus and brainstem remain intact. Clinically, comas of this type range from stupor to deep unconsciousness. In milder cases some ability to sleep and wake may be retained. There may be additional evidence of focal cortical damage, producing localized motor or sensory changes.

Cerebral cortex

Damage to cortex is most commonly produced by metabolic imbalance, anoxia, encephalitis or meningitis, and epilepsy (especially non-convulsive status epilepticus).

Hemisphere white matter

Damage here is most commonly due to severe head injury. Violent shaking of the brain within the skull tears the deep hemisphere white matter[24].

Conditions affecting both white and grey matter

Degenerative conditions such as Alzheimer's disease or Creutzfeldt–Jakob disease are associated terminally with a
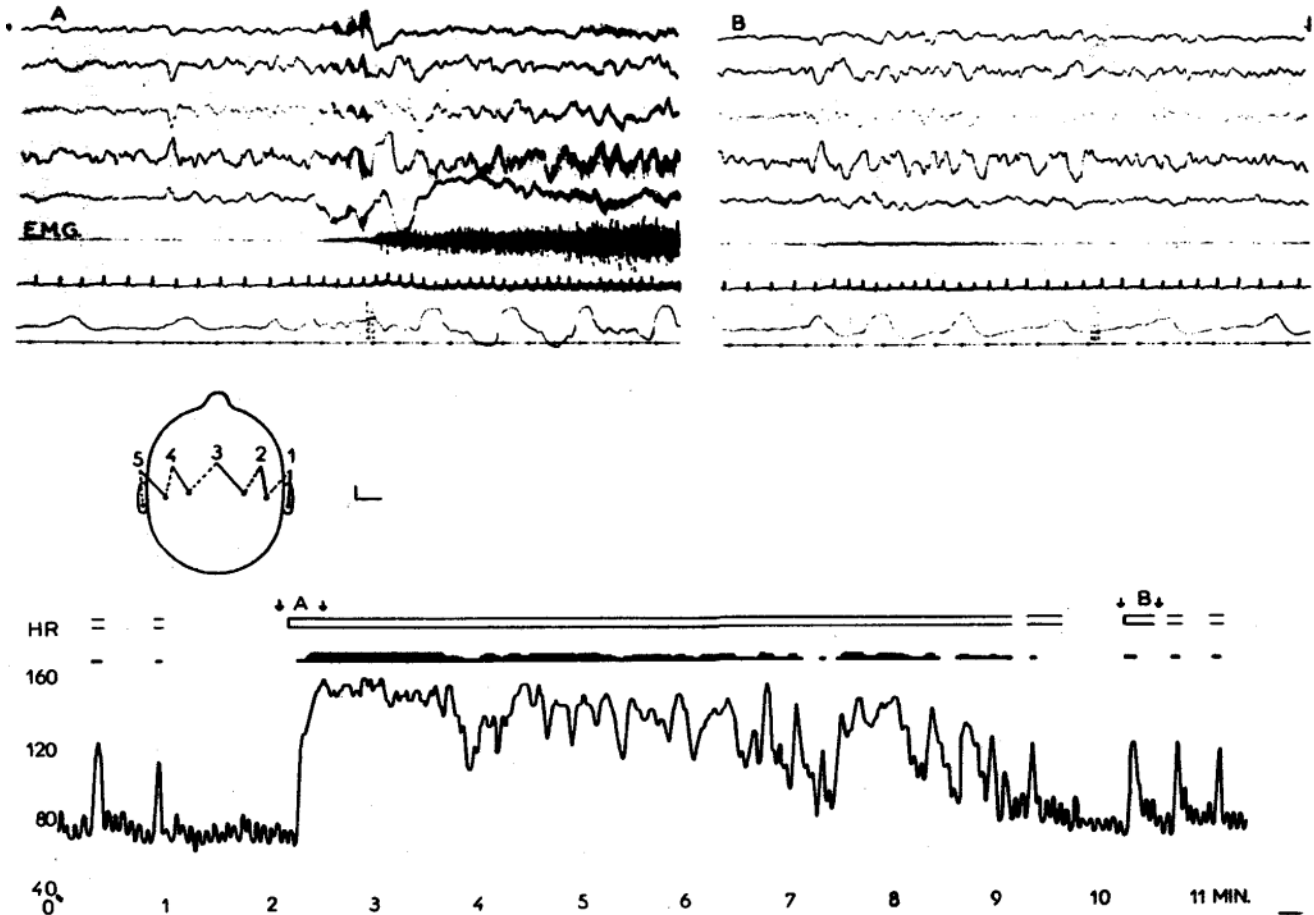
Figure 5 **Recording from boy aged 16 with closed head injury**. *Above*: two pieces of polygraph, A and B. Channels 1–5 EEG; channel 6, surface electromyogram (E.M.G.) from right leg; channel 7, electrocardiogram; channel 8, respiration from thermistor in nose. Section A is taken from a spontaneous arousal accompanied by a decerebrate spasm and shows change in the EEG from low voltage to higher voltage slow waves and marked increase in muscle activity and heart and respiration rate. Section B is taken from a brief arousal after the spasm shown in section A had subsided. EEG shows a brief burst of slow waves with an associated muscle spasm with heart and respiratory rate increases. *Below*: graph of the heart rate. Short arrows, A and B, mark the periods of the graph related to the polygraphic traces shown above. Open blocks illustrate occurrence of EEG change; closed blocks illustrate muscle spasms

stuporose or comatose state related to widespread grey and white matter damage.

Patients with level 1 damage have continuous, severe abnormal EEG activity consisting of high amplitude delta and subdelta with no response to stimulation (non-convulsive status epilepticus will show continuous spike-and-wave activity). This is indicative of widespread cortical and/or white matter damage or malfunction. If the damage is at level 1, with partial recovery the final result may be a state of vacant wakefulness alternating with sleep. The undamaged thalamic sleep mechanisms allow sleeping and waking periods to return but the cortex remains too abnormal to permit awareness. This is one form of the persistent vegetative state.

Level 2: the thalamus and upper mid-brain

This level contains the intermittent gating mechanism of the arousal inhibitory mechanism, related to the thalamic sleep spindle. It is damage at this level that produces the exaggerated intermittent spontaneous changes in arousal described above, involving autonomic, muscular, cerebrospinal fluid pressure and EEG changes. The nature of the muscle activity depends on the level of the injury. Damage at the thalamic/diencephalic level is associated with semi-purposeful movements, but when the upper brainstem is involved the increase in muscle tone associated with each arousal may be decerebrate in character, producing recurrent decerebrate spasms[13] that have been mistaken for seizures. Intermittent arousal activity has been described in many types of brain damage including cerebral tumours, stroke, anoxia, encephalitis and Creutzfeldt–Jakob disease[12,13,20,22,23]. However, it is most commonly seen in subarachnoid haemorrhage and head injury.

Subarachnoid haemorrhage

This condition often causes spasm of the arteries of the circle of Willis, with consequent ischaemia at the thalamic/

diencephalic level. In addition, deeply seated hemisphere clots are common in relation to a bleeding aneurysm. Intermittent arousals occur with changes in the EEG— usually bursts of higher voltage slow waves[12,13] accompanied by periods of rapid respiration and heart rate. These changes may be hard to identify clinically, as the patient remains comatose throughout; often the only sign is a slight movement with rapid breathing. Figure 4 shows recordings from a patient with subarachnoid haemorrhage. Rhythmic arousals are seen at about 4 min intervals. Each arousal is associated with pronounced changes in heart rate and respiration rate. The electrocardiogram shows a wandering pacemaker during the low-arousal periods.

Head injury

Here the brain is damaged at a slightly lower level than in subarachnoid haemorrhage. Closed head injuries can result in a 'knock-out effect' when the hemispheres move in relation to the brainstem at the upper mid-brain level. This impacts on the arousal mechanisms in the upper mid-brain, causing instant unconsciousness. If the dysfunction is longlasting the patient often develops spontaneous changes in the level of arousal, with associated respiratory, heart rate and muscular changes. The motor changes may take the form of decerebrate spasms or uncoordinated movements, often violent. The changes are usually very irregular with long periods of high or low arousal lasting many minutes at a time. There are also brief arousal 'spikes', usually preceding or following a long high-arousal episode. Figure 5 shows recordings from a deeply unconscious 16-year-old boy with a closed head injury: frequent prolonged episodes of high arousal were associated with decerebrate movements, mainly involving the left arm. These episodes sometimes lasted up to 10 min and, when the arousal subsided, there were frequent brief decerebrate motor spasms followed by a long period of continuous low arousal. Each upward arousal change was associated with slow activity in the EEG, a precipitate increase in heart and respiration rate and decerebrate muscle activity. In head injury the EEG arousal activity is prognostically useful. The presence of any sort of EEG change to arousal, either spontaneous or in response to stimulation, predicts a favourable outcome[25]. The young man in Figure 5 did make a good recovery. In modern practice it is usual to sedate patients with severe head injury and ventilate them artificially, so that the clinical features described here may not be immediately apparent. Recognition of the underlying situation may make it easier to interpret sudden swings in blood pressure, cerebrospinal fluid pressure or cerebral blood-flow that can occur even in the sedated state, especially if the patient is stimulated.

Level 3: the lower mid-brain and upper pons

Damage at this level involves the lowest part of the RAM where most of the sensory input is received. This prevents even strong arousal stimuli from reaching the higher cerebral levels. There is always evidence of other brainstem malfunction in the pupils, eye movements and motor and reflex responses. Most fatal comas finish in this state[24]. The final result is brainstem death, with failure of spontaneous respiration, absent responses to stimulation and an iso-electric EEG. In this state there is no sleeping or waking[24].

CONCLUSION

Comatose states are universally assessed by the Glasgow coma scale[26]—simple clinical tests that yield the most useful description of coma and the most accurate prognosis of its outcome at present available. The tests deal with many of the phenomena described in this paper and will continue to offer a practical clinical measurement of comatose states. Nevertheless, the scale is empirical, rather than based on understanding of the underlying neural mechanisms.

This paper presents evidence from many sources to suggest that a dedicated neural mechanism, the non-REM sleep system, underlies the changes between sleep and wakefulness. The system consists of two opposing subsystems—the reticular activating mechanism, which wakes the brain up, and the arousal inhibitory mechanism, which sends it to sleep. When the non-REM system is damaged, coma or stupor results. An analogy can be drawn with the pyramidal system in voluntary movement, or the visual system in sight. The clinical phenomena accompanying any particular coma are related to the site and nature of the damage and provide information that helps in understanding the way the system works. Sleep, as a study, should not be confined to the disciplines of psychiatry, respiratory medicine and psychology but should enter the mainstream of neurological thinking as an important function whose action is still poorly understood.

REFERENCES

1 Borbely AA. Sleep circadian rhythm versus recovery process. In: Koukkou M, Lehmann D, Angst J, eds. *Functional States of the Brain, Their Determinants*. Amsterdam: Elsevier, 1980:51–61

2 Ebrahim IO, Howard RS, Kopelan MD, *et al*. The hypocretin/orexin system. *J R Soc Med* 2002;**95**:227–34

3 Moruzzi G, Magoun HW. Brain stem reticular formation and activation of the EEG. *Electroenceph Clin Neurophysiol* 1949;**1**:455–73

4 Evans BM. Cyclical activity in non-rapid eye movement sleep: a proposed arousal inhibitory mechanism. *Electroenceph Clin Neurophysiol* 1993;**86**:123–31

5   Steriade M. Brain electrical activity and sensory processing during waking and sleeping states. In: Kryger MH, Roth T, Dement WC, eds. *The Principles and Practice of Sleep Medicine*. New York: Saunders, 1989:86–103

6   Evans BM, Richardson N. Demonstration of 3–5s periodicity between spindle bursts in NREM sleep in man. *J Sleep Res* 1994;**4**:196–7

7   Amizica F, Steriade M. Coalescence of sleep rhythms and their chronology in corticothalamic networks. *Sleep Res Online* 1998;**1**:1–10

8   Maquet P. Functional neuroimaging of normal human sleep by positron emission tomography. *J Sleep Res* 2000;**9**:207–31

9   Braun AR, Balkin TJ, Wesensten NJ, *et al.* Regional cerebral blood flow throughout the sleep wake cycle—an (H2O)-O-15 PET study. *Brain* 1997;**120**:1173–97

10 Andersson JL, Onoe H, Hetta J, *et al.* Brain networks affected by synchronised sleep visualised by positron emission tomography. *J Cereb Blood Flow Metab* 1998;**18**:701–15

11 Kajimura N, Uchiyama M, Takayama Y, *et al.* Activity of midbrain reticular formation and neocortex during the progression of human non rapid eye movement sleep. *J Neurosci* 1999;**19**:10065–73

12 Evans BM. Periodic activity in cerebral arousal mechanisms—the relationship to sleep and brain damage. *Electro Enceph Clin Neurophysiol* 1992;**83**:130–7

13 Evans BM. Patterns of arousal in comatose patients. *J Neurol Neurosurg Psychiatry* 1976;**39**:392–402

14 Cooper R, Hulme A. Changes in the EEG, intracranial pressure and other variables during sleep in patients with intracranial lesions. *Electroenceph Clin Neurophysiol* 1969;**27**:564–70

15 Lugaresi E, Coccagna G, Mantovani M, Lebrun R. Some periodic phenomena arising during drowsiness and sleep in man. *Electroenceph Clin Neurophysiol* 1971;**32**:701–5

16 Evans BM. The arousal cycle—a physiological phenomenon. *Electroenceph Clin Neurophysiol* 1981;**52**:25–26P

17 Terzano MG, Parrino I, Spaggiavi MC. The cyclic alternating pattern in the dynamic organisation of sleep. *Electroenceph Clin Neurophysiol* 1988;**69**:437–47

18 Coccagna G, Mantovani M, Brignani F, *et al.* Arterial pressure changes during spontaneous sleep in man. *Electroenceph Clin Neurophysiol* 1971;**31**:277–81

19 Lundberg N. Continuous recording and control of venticular pressure in neurosurgical practice. *Acta Psychiatr Scand* 1960;**36**(suppl 149)

20 Munari C, Calbucci F. Correlations between intra cranial pressure and EEG changes during coma and sleep. *Electroenceph Clin Neurophysiol* 1981;**51**:170–6

21 Fischgold F, Matthis P. Obnubilations comas et stupeurs. *Electroenceph Clin Neurophysiol* 1959;**11**(suppl 11)

22 Evans BM. Cyclic EEG changes in subacute spongiform and anoxic encephalopathy. *Electroenceph Clin Neurophysiol* 1975;**39**:587–98

23 Terzano MG, Mancia D, Zaccetti O, Manzoni GC. The significance of cyclic EEG changes in Creutzfeldt–Jakob disease: prognostic value of their course in 9 patients. *Ital J Neurol Sci* 1981;**3**:243–54

24 Plum F, Posner JB. *The Diagnosis of Stupor and Coma*. Philadelphia: Davis, 1980

25 Evans BM. Prediction of outcome in severe head injury based on recognition of sleep related activity in the polygraphic electroencephalogram. *J Neurol Neurosurg Psychiatry* 1995;**59**:17–25

26 Teasdale G, Jennett B. Assessment of coma and impaired consciousness, a practical scale. *Lancet* 1974;**ii**:81–3