AUTOMATION IN CLINICAL ANESTHESIA

by

Stéphane Bibian

Master of Applied Science, *University of British Columbia*, Vancouver, Canada, 1999

Bachelor of Science, *E.S.I.E.E.*, Amiens, France, 1997

A thesis submitted in partial fulfillment of
the requirements for the degree of

DOCTORATE OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES

(Electrical and Computer Engineering)

THE UNIVERSITY OF BRITISH COLUMBIA

July 2006

Abstract

This thesis investigates the design and performance of a controller for the maintenance of anesthesia during surgery. The controller is designed to be robustly stable for a large population of patients.

Even though anesthetic drugs are amongst the most dangerous drugs used in today's clinical setting, anesthesia procedures are known to be very safe. Hence, the impact of automation in anesthesia in terms of patients' safety cannot be clearly established. However, there are a number of significant clinical advantages to be gained by closing the loop:

1. Recent evidences suggest that most patients undergoing anesthesia procedures are overdosed. This is one of the main reasons for patients' discomfort and slow recovery. Literature suggests that closed-loop systems can significantly reduce drug consumption and lessen recovery times, thus improving the patient outcome while reducing drug-associated costs and bed occupancy.

2. Anesthesiologists have access to intravenous agents with fast onset of action and fast metabolism. Using a closed-loop controller would allow for an infusion-type titration that provides smoother transitions, thus avoiding the respiratory and hemodynamic depression observed in a bolus-based manual regimen.

3. Closed-loop controllers are also particularly well-suited for solving complex optimization problems. The profound synergy that exists between intravenous anesthetics and opioids could then be fully exploited. This could be a significant factor contributing to a reduction in drug usage and the improvement of patients' comfort.

This project is particularly challenging. In particular:

1. There is no accepted measure of depth of anesthesia. Hence, it is necessary to work at the conceptual and sensor levels in order to define adequate feedback measures.

2. Drug effect modeling suffers from many shortcomings. In particular, published studies are often not in agreement regarding model parameters.

3. Uncertainty of dose/response models is daunting. Measuring this uncertainty is necessary in order to ensure stability of the control design.

While the anesthesia closed-loop concept has already been investigated in the past, no breakthrough has yet been achieved. We feel it is necessary to investigate the anesthesia system from a control engineering perspective.

This thesis is divided into two distinct parts. Part A contains the first 4 chapters and presents a thorough introduction to clinical anesthesia. The main concepts, terminology and issues are covered, including anesthesia monitors and basic pharmacology principles. A review of the prior closed-loop control attempts is presented in Chapter 4.

Part B contains the chapters 5 to 8. In these chapters, we investigate a new sensor technology to quantify both cortical and autonomic activity. This technology is used to derive drug effect models, from which uncertainty bounds are derived. Based on this uncertainty analysis, we derive robustly stable controllers achieving clinically adequate performances.

Finally, we invite the readers to refer to Chapter 9 for a complete synopsis and summary of this thesis.

Table of Contents

List of Tables

List of Figures

List of Acronyms

Acronym	Description
ACL	Anterior Cruciate Ligament
ASA	American Society of Anesthesiologists
ASSR	Auditory Steady State Response
BIS	Bispectral Index Scale
BSR	Burst Suppression Ratio
CNS	Central Nervous System
DOA	Depth of Anesthesia
DWT	Discrete Wavelet Transform
ECG	Electrocardiogram
ECT	Electro-Convulsive Therapy
EEG	Electroencephalogram
EMG	Electromyography
EOG	Electrooculogram
ESU	Electro-surgical unit
ERP	Evoked Response Potential
FDA	Food and Drug Administration
FIR	Finite Impulse Response
HRV	Heart Rate Variability
ICU	Intensive Care Unit
IIR	Infinite Impulse Response
LMA	Laryngeal Mask Airway
LOC	Loss Of Count
MAC	Minimum Alveolar Concentration

ACRONYM	DESCRIPTION
MEF	Median Frequency
MLAER	Mid-Latency Auditory Evoked Response
NMB	Neuromuscular Blocking Agent
OA	Ocular Artifact
OAs	Ocular Artifacts
OR	Operating Room
PD	Pharmacodynamics
PDF	Probability Density Function
PET	Positron Emission Tomography
PID	Proportional-Integrative-Derivative control law
PK	Pharmacokinetics
PKPD	Pharmacokinetic - Pharmacodynamic
PS	Patient Simulator
REM	Rapid Eye Movement
RMS	Root Mean Square
RSA	Respiratory Sinus Arrhythmia
RWT	Redundant Wavelet Transform
SCC	Semilinear Canonical Correlation
SEF	Spectral Edge Frequency
SEP	Somatosensory Evoked Potential
SP	Smith Predictor
STFT	Short Time Fourier Transform
SWT	Stationary Wavelet Transform
UBC	The University of British Columbia
UBCH	The University of British Columbia Hospital
VGH	Vancouver General Hospital
WAV	Wavelet-based Anesthesia Value
WAV_{CNS}	Wavelet-based Anesthesia Value for Central Nervous System monitoring
WAV_{CNS}	Wavelet-based Anesthesia Value for Autonomic Nervous System monitoring
WT	Wavelet Transform

Acknowledgements

As I am concluding this thesis, I would like to express my deepest gratitude to my supervisor, Professor Guy Dumont, for his valuable support, advice, and guidance throughout the course of this research. Being part of the multidisciplinary research team has given me the unique opportunity to familiarize myself to an exiting new field of research.

I am also particularly grateful to Dr. Mihai Huzmezan, my co-supervisor, for his constant encouragement and unfaltering support. His optimism and pragmatism have often smoothed over many difficulties and moments of doubt.

I am also deeply indebted to Dr. Craig Ries, my co-supervisor, for sharing with me his expertise in anesthesia, and for opening the many doors that have allowed this work to flourish. I am looking forward to our continuing collaboration.

I would also like to thank Ms. Tatjana Zikov, who shared with me her scientific expertise in signal processing, and her commitment to anesthesia research. I will always be grateful for the long hours we spent together reviewing data, and sharing in the excitement of new discoveries.

Thanks are extended to Dr. Kris Lundqvist, Dr. Henrik Huttunen, Dr. Jason Waechter, and Ms. Lou-Ann Mendoza from the UBC departments of Anesthesia and Electrical and Computer Engineering, for helping prepare and co-ordinate the clinical studies, which provided the basis for most of the results and contributions presented in this thesis.

I would also like to acknowledge Cleveland Medical Devices Inc. (Cleveland, OH, USA), and in particular Dr. Mohammad Modarreszadeh and Mr. Robert Schmidt, for their interest in carrying forward the development of the NeuroSENSE^{TM} anesthesia monitor, which integrates the sensing technology developed in part in this thesis.

Finally, I would like to thank all my friends here in Vancouver, for their support and friendship, and for having made my stay in Vancouver memorable.

Stéphane Bibian

July, 2006

This thesis is dedicated to the memory of my grandfather, Albert Vande Kerckhove, a truly righteous and altruistic man, who will always live in our memories.

PART A: The Anesthesia Framework

Chapter 1 introduces the readers to the terminology and concepts in use in the field of clinical anesthesia. A section dedicated to the most commonly administered drugs is presented, as well as information regarding drug delivery and standard of care in monitoring.

The second chapter addresses the problem of Depth of Anesthesia (DOA) measurement. This particular topic has received significant attention from both the anesthesia and engineering communities. Even as of today, there is still no definite consensus as to how this measurement should be performed. While there are commercially available monitors of the anesthetic state, current practice does not mandate these monitors as a standard of care. These monitors are often criticized by practitioners, and are the source of debates and controversies.

Knowledge of the process is vital in any control design venture. Since online identification of the process is not practical for this particular application, it is necessary to establish proper models relating infusion rates and observed effects, as well as the uncertainty range. Chapter 3 investigates the pharmacokinetic and pharmacodynamic concepts used in the clinical field to model the drug-response relationship.

Finally, Chapter 4 lays out the foundation for an anesthesia controller. Prior attempts are reviewed. Based on the experience gained from these attempts, as well as the concepts presented in the previous chapters, we propose a control framework suitable for this particular application. The main requirements and challenges are outlined.

Chapter 1

Clinical Anesthesia: Terminology, Concepts, and Issues

Industrialized societies have enjoyed the benefits brought by clinical surgery and modern medicine for over a century. Undergoing a surgical procedure has become nowadays a rather common event in one's life. However, this remarkable progress has only been made possible through the development of modern anesthesia practice since its discovery in the mid 19^{th} century.

While the aim of this chapter is not to present a thorough review of the practice of clinical anesthesia, we feel that it is necessary to introduce our readers with an overview of the concepts now in use in this medical field. The short history presented in the first section will emphasize the tremendous impact that the discovery of anesthesia had in medicine, as well as the numerous developments that have taken place ever since. The concepts and definition of anesthesia, as well as the different anesthetic and opioid agents used in today's practice, will then be discussed. Finally, we will conclude this chapter by briefly reviewing the different devices and sensors most commonly found in today's practice.

1.1 Clinical Anesthesia: A Key Specialty in Critical Care

Before the advent of anesthesia, surgical procedures demanded extremely fast execution. Early regional techniques such as nerve compression or the application of cold were used to provide a slight relief from the pain. Decreased cerebral perfusion obtained by compressing the carotid artery was also used to render patients unconscious. Quite clearly, the discovery of inhalation gases that could provoke a state of *anesthesia*, and thus make invasive surgeries possible, was a major event in the development of modern medicine.

Numerous books and textbooks give the historical perspective to clinical anesthesia. We invite interested readers to refer to [1], [2], and [3] for a more in depth look at this medical specialty.

1.1.1 The Early Days

While the alteration of the senses using drugs (opium, laudanum, mandragora, etc.) and alcohol was well known since antiquity, it's only in 1840 that the idea of using inhalation gases and vapors, otherwise reserved to entertain the public in village fairs, was put forward by Hickman. The first recorded anesthesia procedure was performed by Crawford Williamson Long using diethyl ether. Having often inhaled ether, he realized that this gas had the property of rendering him particularly insensitive to painful knocks and bruises. He concluded that this could be applicable to surgery, and tried out his theory in March 1842. While he continued using ether for surgeries, he did not publish his results until later. In 1844, Gardner Colton and Horace Wells performed the first dental surgery using nitrous oxide as an anesthetic. However, they failed in convincing the medical community, and it is only later, in October 1846, that William Morton succeeded in demonstrating to his colleagues that ether could be used to deprive patients of their sense during surgery. The success of his 'etherization' technique spread quickly throughout the civilized world. The term 'anesthesia' (lack of *esthesia*, *i.e.* sense) was later proposed to Morton by Oliver Wendell Holmes to describe this new phenomenon.

While ether has been advocated by many in the early days, other agents such as chloroform, nitrous oxide, and ethyl chloride were investigated. Nowadays, only nitrous oxide is still in use as a supplement to the modern agents developed in the 1960s and 1970s. The development of drugs used for anesthesia procedures will receive a more thorough review in Section 1.3.

Parallel to the development of anesthetic agents, the 19^{th} century witnessed the invention of many inhalation apparatuses and techniques for the administration of anesthetics. For instance, there was a need for an artificial ventilator, since a major side effect of many anesthetics and opioids is ventilatory depression. The first ventilator was developed in 1896 by Northrup, who had shown that patients who had opium poisoning could be maintained alive through artificial ventilation, using an endotracheal tube and bellows. But only in the 1950s, after the introduction of muscle relaxants and polio epidemics, did the use of ventilators became common in the operating room.

Due to the increasing complexity of the administration and management of anesthesia, it became clear that anesthesia and surgery were two complementary specialties that demanded their own practitioners. But it was only in 1935 that the first diploma of anesthesia was offered, and in 1936 that the American Society of Anesthesiologists (ASA) was founded.

1.1.2 Risks and Outcome in Anesthesia

1.1.2.1 Mortality rate

Development of clinical anesthesia since the 19^{th} century has resulted mostly from concerns about patients' safety.

The first recorded death attributed directly to anesthesia was that of Hannah Greener in 1848, after

inhaling chloroform for the removal of an ingrown toenail. The question whether the patient's death resulted from an inadequate anesthesia management or from undesirable side effects due to chloroform was hotly debated. Anesthesia and the use of chloroform being rather new, this case received considerable attention from both the medical community and the public, and was particularly well-publicized. Chloroform was by then a newly introduced drug. Less flammable than ether, chloroform offered an attractive alternative to surgeons. But after more deaths were reported throughout the world, the safety of chloroform as an anesthetic drug was questioned. Even then, it was only in the 1920s that chloroform was replaced by new, more potent, and safer gases. The cause of Hannah Greener's death remains unclear, and 150 years later this case still generates interest [4]. By the late 19^{th} century, the incidence of death due to anesthesia was less than 0.1% [5].

Nowadays, clinical anesthesia is probably one of the safest components of any surgical operation. A 1986 survey [6] revealed that the overall death rate attributable directly to anesthetic drugs was 1:185,056 (see Table 1.1). With approximately 28 million patients undergoing anesthesia and surgery in the United States [7], it is estimated that about 150 patients die each year in the U.S.A. from complications due to anesthesia. This very low mortality rate can be attributed mainly to the following three aspects of the clinical practice:

i. First, anesthesiologists select an appropriate combination of drugs and drug dosage according to the patient's age, weight, co-morbid disease, and the type and duration of the operation. In standard practice, anesthesiologists often use several drugs in order to reach a state of *balanced anesthesia* (refer to Section 1.2.3), thus limiting the potential lethal side effects of each drug.

ii. The second aspect concerns the equipment that monitors patients' vital signs and eventually warns the practitioner of possible complications. Modern equipment is fairly sophisticated and includes standard devices such as mass spectrometers, capnographs, pulse oximeters, heart rate and blood pressure monitors, etc.

iii. Finally, education has had a key role in making anesthesia a particularly safe and reliable procedure. Postgraduate training programs in the specialty of anesthesiology are offered by every major medical schools. The content of these programs are usually supervised by a centralized specialty college in most countries. Also, this medical specialty benefits from the publication of numerous clinical research journals (Anesthesia and Analgesia, Anesthesiology, British Journal of Anaesthesia, etc...).

STUDY	DATE	TOTAL CASES	MORTALITY RATE
Beecher	1948–52	599,548	1:1,560 [a]
Clifton	1952–62	205,640	1:3,955 [a]
Harrison	1967–76	240,483	1:4,537 [b]
Hatton	1977	190,389	1:2,885 [a]
Lunn	1979	1,147,362	1:6,789 [e]
Eichhorn	1976–85	757,000	1:151,400 [d]
Eichhorn	1985–88	244,000	0 [d]
CEPOD	1986	486,000	1:185,056 [e]
CEPOD	1986	486,000	1:1,185 [c]

a: All operative cases considered in calculation
b: Cases included if death occurred in less than 24 hours
c: Cases included if some contribution by anesthetic
d: Only ASA physical status I and II patients included
e: Only deaths directly attributable to anesthetic included
CEPOD: Confidential Enquiry into Postoperative Deaths

Table 1.1: Contemporary anesthetic mortality rates (adapted from Brown [5])

The method used for estimating the mortality rate depends greatly on the methodology used by each researcher. While deaths during surgery were easily attributed to the anesthetic drugs in the early days, such cases are now thoroughly investigated and documented. Also, the patient physical status is now accounted for. The outcome of any surgery depends indeed whether the patient is healthy (ASA Physical Status I) or in critical condition (ASA V). The mortality rate is obviously much higher in ASA V patients (about 10% [5]) who undergo emergency surgery following major trauma than for ASA I or II patients (0.08% and 0.27% respectively) who undergo elective surgery. The ASA (American Society of Anesthesiology) rating was introduced in 1941.

Human error is probably the most common cause of death (hypoxic gas mixture, airway obstruction, errors in drug administration, lapses in vigilance, etc...) [8]. According to a 1987 study [9], 75% of anesthetic related deaths are due to the anesthetists' failure to apply life saving knowledge, while only 1.7% of cases involve equipment failure. Also, where the most common reason cited leading to such events used to be overdosing (1960–1969), it is nowadays the inadequate preoperative preparation and patient assessment that leads to errors in anesthesia management. Finally the actual trend of increased efficiency in the operating room (so called *production pressures*) is an additional factor of anesthesiologists' fatigue that might provoke misjudgement, resulting in reduced patients' safety.

1.1.2.2 Intraoperative awareness

While patient's safety is still an important issue, new concerns arose from the use of *neuromuscular blocking agents* (NMBs) in the early 1940s. These drugs, used to block muscle movement and thus reduce muscle tone to facilitate surgery, have the unpleasant consequence of obtunding an important sign of light

anesthesia, *i.e.* movement. It is therefore possible for a patient to remain immobile while being aware of his/her surroundings, and experiencing the trauma caused by the surgery [10]. While not necessarily lethal, awareness with pain during surgery often results in severe psychological consequences. These cases are fortunately extremely rare (about 1 in 10,000 [11]) and result principally from faulty equipment or human error.

However, limited intraoperative awareness without the presence of pain is more common, mostly when patients are maintained in a shallower depth of anesthesia or during emergencies. A number of such cases have been reported in the literature. Very recently, two extensive studies have been conducted to determine the incidence of intra-operative awareness:

- **SAFE1 Trial**: Sandin *et al.* [12] have undertaken an extensive survey of intra-operative awareness cases in Sweden, spanning 2 major hospitals and 2 years of clinical activity (from 1997 to 1999). 11,785 patients were interviewed on three separate occasions (immediately after the operation, 1 to 3 days after surgery, and then once again 7 to 14 days after). A total of 18 cases (0.15%) of awareness were reported, and 14 of these cases occurred in surgeries involving NMBs. 11 of these 14 paralyzed patients experienced either pain, anxiety, or delayed neurotic symptoms, while the 4 non-paralyzed patients did not suffer during their period of wakefulness. This study also revealed that patients might recall intra-operative events only after some time after the surgery. For instance, this study revealed that only 11 cases would have been identified if the patients had only been interviewed once immediately following the operation.

- **AIM Trial**: A similar study was undertaken in the United States [13]. 19,576 patients over 7 different sites have been enrolled so far in this study (the study objective is 50,000). The incidence of intra-operative awareness has so far been found to be 1 to 2 cases per 1000 patients. Considering that there are 20 millions of anesthetic procedures carried out each year, it is estimated that about 100 patients suffer from intra-operative awareness every weekday worldwide.

It is estimated that up to 50% of the patients who experience intra-operative awareness will develop Post-Traumatic Stress Disorders (PTSD) in the following year.

1.2 Modern Concepts

In this section, we will discuss the practice of modern clinical anesthesia by introducing the concepts of the *anesthesia triad* and *balanced anesthesia.* These concepts play a key role in defining and understanding the proposed research project.

1.2.1 The Role of the Anesthesiologist

In lightly anesthetized patients, surgical stimulation leads to movement, cardiorespiratory changes (*i.e.*, rise in blood pressure, heart rate and respiration), and hormonal responses (release of adrenaline, cortisol,

etc.). During "normal" stress situations of course (*i.e.*, no surgery, no anesthesia), these cardiorespiratory and hormonal responses act as the body's self-defense system in that they assist the subject to fight or flee. During surgery, however, "stress responses" can be dangerous, and even life-threatening, as they can provoke strokes and heart attacks in susceptible individuals. Intra-operative awareness is another potential risk in under-anesthetized patients. While not fatal, it can provoke postoperative post-traumatic stress disorders [14, 15].

In order to blunt the effect of surgical stimulation, anesthesiologists use a combination of drugs to block sensation. However, the very mechanism of action of these drugs make them particularly dangerous, as they deprive the central nervous system from the information necessary to control normal body functions (*i.e.*, gag reflex, respiration, cardiac rhythm and blood pressure). An overdose may then stop a patient's breathing and might even provoke a cardiovascular collapse. Overdoses are usually associated with a lack of balance between the anesthetic regimen and the patient's pharmacological needs. When there is no surgical stimulation, the patients' needs are low and a small amount of drug may be sufficient to make them comfortable. However, during noxious stimuli (*i.e.*, stimuli associated with transmission of nerve pain signals), drug titration needs to be increased to limit the effect of surgery. As a result, a common side effect is the depression of the cardiorespiratory system when surgical stimulation suddenly disappears.

Therefore, anesthesiologists try to keep a balance between the toxicity of anesthetic drugs and the noxious stimulation of surgery.

1.2.2 Functional Components of Clinical Anesthesia

Although its scientific definition is still open to debate, anesthesia has been described as a state of "drug-induced unconsciousness, [where] the patient neither perceives nor recalls noxious stimuli" [16]. This functional definition proposed by Prys-Roberts in 1987 limits the term of anesthesia to an absence of both conscious awareness and memory formation (*i.e.*, *hypnosis* and *amnesia*).

However, the role of anesthesiologists goes beyond provoking a mere hypnotic state. They also ensure that autonomic reflexes involving the sympathetic and parasympathetic nervous system (to provide cardiorespiratory control) are not sensitive to surgical stress. This is achieved by inducing a state of *analgesia*. Opioid drugs are typically used to achieve this endpoint.

When both the hypnotic and analgesic states are adequately reached, patients do not usually exhibit purposeful movement to surgical stress. However, intra-abdominal surgeries require the blockade of reflex muscle activity in the abdominal wall in order to permit surgical exposure. To attain such a state of *paralysis*, it is necessary to use neuro-muscular blocking agents (NMBs). It is important to note that these drugs act peripherally at the level of the synaptic link between the nerve and the muscle, and not centrally in the brain or the spinal cord.

To summarize, it is common in the literature to consider that the state of general clinical anesthesia results from the combination of three functional components, that is, hypnosis, analgesia and immobility.

While this *anesthesia triad* [17] concept is somewhat simplistic, most engineering oriented authors contributing in this field are setting their work within this functional framework since each individual component can be accomplished by separate classes of intravenous drugs [18]. Note, however, that inhaled anesthetics have minor analgesic and relaxant properties and large doses of opioids can depress cognitive function. Not surprisingly then, anesthetics and opioids enhance or potentiate each other's action in a synergistic fashion. Conversely, interactions between NMBs and opioids, or NMBs and intravenous hypnotics are minimal.

1.2.3 Modern Balanced Anesthesia

As stated earlier, all anesthetics are by definition hypnotics. They first act at the level of cognition (cortex) by rendering patients unconscious. However, with increasing doses of anesthetic drugs, it is possible to go beyond hypnosis and blunt further the response to noxious stimuli. This is particularly true for inhalational anesthetics. Till the 1940s, it was common for anesthetists to use a single agent at high concentration to achieve adequate anesthesia. Unfortunately, using higher doses usually results in stronger side effects during surgery (ventilation depression, cardiac arrhythmia, etc.) and also during recovery (nausea, vomiting, etc.).

To alleviate undesirable side effects George Crile advocated in 1911 the use of local analgesics as a complement to light general anesthesia. In 1926 the term *balanced anesthesia* was introduced by John Lundy to describe a combination of agents that would achieve adequate anesthesia. This concept can be well-understood when considering that analgesia and areflexia can be achieved separately using drugs such as opioids and NMBs. These drugs are not hypnotics in the sense that they do not provoke unconsciousness. Nowadays, the focus in clinical anesthesia is to achieve an adequate balanced anesthetic state using a combination of hypnotics (inhalational/intravenous anesthetics), opioids and NMBs. This technique has the advantage that much lower concentrations of these drugs are then needed as they are used concurrently. Hence, side effects, recovery time, and post operative nausea and vomiting are considerably reduced. Balanced anesthesia is now the standard in the management and conduct of clinical anesthesia.

1.3 An Extensive Pharmacopoeia

A combination of anesthetics and opioids, with or without NMBs, are administered together to create the state of general anesthesia. These drugs, even when taken within the same family, have different properties. Since they provide the actuators through which the patient's state can be regulated (*i.e.*, allowing the control of the anesthetic state), it is necessary to provide control engineers with some knowledge of the mechanism of action of the most commonly used drugs.

1.3.1 Anesthetics (*i.e.* hypnotics)

Inhaled Anesthetics With the advent of fluorine technology in the 1940s, new inhaled anesthetics were developed. Compared to ether and chloroform, fluorine compounds have lower blood solubility (thus ensur-

ing rapid induction and recovery), lower toxicity, are less irritating to the airway, and are not flammable. Nowadays three agents are commonly used with or without nitrous oxide: isoflurane, desflurane and sevoflurane. All these agents provoke a decrease in the mean arterial blood pressure when administered to healthy subjects.

A major advantage of inhaled anesthetics is that the drug uptake in the arterial blood stream can be precisely titrated by measuring the difference between the inspired and expired concentrations. This measurement is done in real-time using a device such as a mass spectrometer. At steady state, the expired concentration correlates with the brain concentration. As a result, inhaled gases are used extensively to maintain a desired depth of anesthesia.

This discussion would not be complete without mentioning the concept of *Minimum Alveolar Concentration* (MAC). A value of 1 MAC is the minimum alveolar concentration of an inhaled anesthetic that prevents purposeful movement in response to a noxious stimuli in 50% of patients. The MAC value is thus used to compare dosage between inhaled anesthetics.

Intravenous Anesthetics Intravenous anesthetics can be classified into 5 families: Barbiturates (thiopental), Benzodiazepines (midazolam, diazepam, lorazepam), Phencyclidines (ketamine), Carboxylated imidazoles (etomidate), and Isopropylphenols (propofol). Compared with volatile agents, intravenous anesthetics (besides ketamine) do not provide analgesic effects – hence, they are defined as hypnotic rather than anesthetic drugs. However, opioids and intravenous anesthetics, when used in combination, are strongly synergistic, both in terms of hypnosis and analgesia.

Propofol was introduced in the early 1990s and has become the intravenous drug of choice in anesthesia. Two particular characteristic of propofol are its fast redistribution and its metabolism. As a result it can be easily used in infusion schemes as it provides very fast emergence, without cumulative effect.

Inhalational Anesthetics *vs.* Intravenous Agents Inhalational anesthetics are considered by many practitioners as near ideal anesthetics as they have both an hypnotic and an analgesic effect. This explains why past closed-loop anesthesia attempts used inhalational anesthetics as the sole actuator. Combined with the fact that the lung partial pressures of inhaled anesthetics is closely related to the vapour concentration in the brain, the control problem is significantly simplified since additional states are measurable.

As opposed to inhaled anesthetics, the brain concentration of an intravenous drug cannot be easily measured. As a result, the titration of these drugs is more difficult as the anesthesiologist does not have any feedback on how much of the administered drug has been metabolized or stored in inactive tissues. In the majority of cases, intravenous agents are given as large boluses (large doses given over a very short amount of time, *e.g.*, < 1 min) for the induction of anesthesia, while maintenance is accomplished with inhaled vapors (but more and more often with intravenous infusions). However, since intravenous agents are more specific than inhaled anesthetics, they give more flexibility in separately controlling the functional components of anesthesia.

1.3.2 Opioids

Opioids are unique in the sense that they provoke analgesia without loss of touch, temperature sensation and consciousness, when administered in small doses. They act as agonists at specific receptors within the Central Nervous System (CNS). Their principal effect may be the inhibition of neurotransmitter release, resulting in a significant analgesic effect.

Contrary to most anesthetics, opioids do not directly depress the heart, and are thus particularly suitable for cardiac anesthesia. Opioids can produce unconsciousness when used in very large doses. This observation has led some authors to believe that opioids should be considered to be anesthetics, as they fit Prys-Roberts' definition [19]. However, the state of unconsciousness brought by opioids is not reliable. It has been shown for instance that they cannot fully replace inhaled vapours to provoke an adequate state of hypnosis. However, their use can reduce the requirements of inhaled anesthetics by up to 50% [20]. Also, the sedative effect of opioids is opposed by the presence of acute pain. Hence, even though patients in severe pain receive very large amount of opioids, they can remain aware. In current practice, therefore, opioids are almost always supplemented by other anesthetics.

Five opioid compounds are used in clinical anesthesia: morphine, hydromorphone, fentanyl, sufentanil, and remifentanil. While they all have similar effects, their kinetic characteristics differ tremendously due to large differences in their lipid-solubility. Of particular interest is remifentanil, a new agent introduced in the mid 1990s. The potency of remifentanil is twice that of fentanyl and its effect-site equilibration time is slightly less than that of alfentanil (about 1.1 min). Remifentanil differs from all other opioids due to a unique structure where an ester linkage makes it susceptible to hydrolysis. This property results in a rapid degradation to inactive metabolites [21]. The main characteristics of remifentanil then are: rapid onset, brevity of action, noncumulative effects in inactive tissues and rapid recovery after termination of the infusion. Remifentanil is used mostly to provide the analgesic component of general anesthesia. Its brevity of action allows patients to recover rapidly from undesirable opioid-induced side-effects such as ventilatory depression.

1.3.3 Neuromuscular Blocking Agents (NMBs)

NMBs act locally at the level of the neuromuscular junction by interrupting the transmission of nerve impulses. Their principal use is to produce skeletal muscle relaxation to facilitate intubation of the trachea and to provide optimal surgical conditions. NMBs do not have any analgesics or hypnotic properties. They also do not interact in a clinically significant way with opioids and anesthetics. Succinylcholine is used whenever a rapid onset and short duration of action is needed. Cisatracurium, vecuronium, mivacurium and rocuronium are used when a longer effect is desired.

1.3.4 Drugs: Action, Effect and Interaction

The development of safer and more potent intravenous agents with faster onset of effect and, in certain cases, shorter duration of action, has greatly impacted the practice of anesthesia. Nowadays, small drug quantities used in combination can produce a balanced state of anesthesia while minimizing side effects.

In North America, inhalational gases are used to maintain the background of anesthesia. However, intravenous agents are increasingly employed for maintenance. The use of intravenous agents is geared towards facilitating intubation, compensate for undesirable changes in patients' state and also to anticipate painful surgical stimuli. In this realm, the short acting characteristic of intravenous agents such as remifentanil and propofol indicates that these drugs should best be used in infusion regimens, since their administration as boluses often result in too strong effects for too short periods of time.

The inability of measuring plasma drug concentration of intravenous agents raises questions. Currently this affects the ability of the anesthesiologist to set precise rates of infusion. Infusion regimens published in medical journals are prone to error due to model uncertainty. Hence, the resulting titration might not correspond to the patients' real needs.

In the context of closed-loop control, and when using intravenous drug as "actuators", it is necessary to account for both patient variability and drug synergism.

Patient variability results from differences in the way the drug distributes and is eliminated from the body which is further influenced by cardiac output, an individual's age, lean body mass, renal and liver functions, etc. Genetic differences and enzyme activity might also alter the sensitivity to the drug. For example, while some patients might be hyporeactive (*e.g.*, tolerance due to addiction), others may be hypersensitive.

When using different drugs in combination, significant interactions can be observed. An *additive* effect signifies that a second drug taken concurrently with a first will produce an effect equal to the superposition of their effects (e.g. the anesthetic effects of two inhaled anesthetics are additive [22]), whereas a *synergistic* effect means that the resulting effect is greater than what could be expected from superposition. Drug synergism often appears when using hypnotics in combination with opioids. In some particular cases, drugs can also be *antagonistic*, in which case they counter-act each other when administered concurrently (*e.g.*, an opioid antagonist such as naloxone provokes a complete reversal of the effect of opioids). From a control point of view, such drug interactions tend to generate significant cross-coupling in the multivariable system that models the patient. Only very few models of such couplings have been discussed in the literature; see [23, 24] and [25]. These models are mainly mathematical expressions that describe drug interaction in steady state. As such they do capture the static non-linearity of drug interactions, but fail to characterize the dynamics associated with state transitions during transitory events.

1.4 Conduct of Anesthesia

In most cases, we can divide the anesthesia procedure into 3 distinct phases: induction, maintenance and emergence.

Induction Changing a patients consciousness from an alert to an anesthetized state involves a transition phase called induction. An intravenous induction of anesthesia is typically used in adults since inhalational inductions are slower and can be associated with an intermediate "stage of excitation" with vomiting and/or spasm of the vocal cords. During intravenous induction with a hypnotic agent, the intermediate stage of excitation is generally not seen and airway, respiratory and cardiovascular reflexes are rapidly depressed with the sudden onset of unconsciousness. For example, the tongue relaxes to obstruct the airway and spontaneous breathing stops. As a result, the upper airway must be instrumented and ventilation must be, at least initially, manually or mechanically controlled. Airway instrumentation methods are invasive (*i.e.*, like surgery) and generally determine the initial need for opioids and NMBs, which in turn may further depress airway and ventilatory control. There is no clinically useful brain monitor as yet that can assist in the initial titration of intravenous anesthetic drug dosage, so cardiorespiratory stability is used to infer an absence of consciousness.

The induction of anesthesia is then a critical phase, as the first priority is to maintain the patient's airway. For example, major surgery generally requires the patient's airway to be secured by an endotracheal tube – a tube and air-filled cuff that is placed through the vocal cords – which then allows positive pressure and mechanical ventilation. An additional advantage of the endotracheal tube is that its intra-tracheal cuff prevents the pulmonary aspiration of gastric contents in the case of either gastric-esophagal reflux or gastric vomiting with food ingestion prior to emergency surgery. As discussed earlier, the insertion of the endotracheal tube is particularly painful and stimulating (*e.g.*, skin incision) as it passes through the vocal cords and thus can provoke strong gagging reflexes. To facilitate the insertion of the tube, it is common to first use an opioid such as fentanyl to blunt the pain, followed a few minutes later by a large bolus of propofol. The administration of such a bolus will rapidly provoke a deep level of hypnosis. To further blunt any motor reflex, a NMB is generally administered.

An alternative to the endotracheal tube is the *Laryngeal Mask Airway* – a tube and air-filled cuff that sits above the voice-box. Less invasive than the endotracheal tube, the laryngeal mask airway usually does not necessitate the use of NMBs for its insertion since it does not pass through the vocal cords. However, it is less secure in case of gastric reflux and does not therefore prevent pulmonary aspiration. As a result, laryngeal mask airways are used essentially on healthy patients undergoing minor elective surgery on empty stomachs.

The induction of anesthesia is usually a very short event lasting only a few minutes.

Maintenance The effect of the propofol induction dose being very short (due to re-distribution), it usually wears off once the tube or the laryngeal mask airway is inserted. Hence, it is necessary to administer more anesthetic to maintain an adequate depth of hypnosis. The maintenance of anesthesia is often obtained via an inhalational agent set around 1 MAC. Prior to surgical incision or any other stimulating surgical event, it is customary to reinforce the analgesic component of anesthesia by administering a small bolus of opioid. The inhalational gas being also an analgesic, another alternative is to increase the concentration of the inhaled agent to 1.5 MAC. If during the surgery there is a long period of low stimulation, the end tidal gas concentration may need to be reduced to lessen the level of cardiac depression. The widespread use of inhalational agents for the maintenance of anesthesia is mostly attributed to their ease of use. As mentioned earlier, the measurement of the end-expired vapour concentration provides the anesthesiologist with a reliable feedback quantity.

In some cases, an infusion of propofol is given concurrently with an inhaled agent. Usually, the infusion is set at a fixed rate and the concentration of the inhaled anesthetic is tuned according to the patient's need. This method makes it possible for the anesthesiologist to reduce the concentration of the inhaled anesthetic to about 0.5 MAC, as the intravenous anesthetic will deepen the level of hypnosis. This usually results in less side-effects, especially post-operative nausea.

In recent years, *Total Intravenous Anesthesia* (TIVA) has been strongly advocated by many anesthesiologists [26], and has gained popularity with the introduction of faster agents with shorter duration of effect [27]. In TIVA schemes, no inhalational agent is used. Hypnosis and analgesia are achieved using only 2 agents: an anesthetic drug (usually propofol) and an opioid (alfentanil or remifentanil in recent years). The major difficulty for TIVA schemes is the lack of arterial blood measurement of the concentration of each drug. Without this feedback information, precise titration of the quantity of drug necessary to achieve a particular endpoint is not possible. Hence, the development of TIVA has been made easier through the development of pharmacokinetic model-driven infusion devices. These devices reach a desired plasma concentration by using a computer-controlled infusion pump driven by the known pharmacokinetic parameters of the drug. The first *Target Controlled Infusion* (TCI) devices were made available in the late 1980s, but they have not been yet approved for use in North America. The main advantage of TIVA is that patients wake up faster and experience less postoperative nausea, thus shortening their stay in post operative care units.

However, TIVA is still a marginal technique, as the anesthesiologist remains blinded to the true plasma concentration. Also, the incidence of intraoperative awareness and movement is higher. Finally, recent economic studies have concluded that TIVA usually results in higher costs (up to three times the cost of standard practice) [28], [29]. In today's cost sensitive environment, this is a major deterrent to this technique.

Emergence The emergence from anesthesia is simply achieved by turning off the vaporizer as well as any infusion device used during the surgery. This is usually done during skin closure so that the patient wakes

up faster at the end of surgery. An additional bolus of a long acting opioid may be given for postoperative pain management. If the patient does not breathe spontaneously and still needs artificial ventilation, the anesthesiologist may use naloxone to antagonize the effect of opioids. However this practice is rare, as it may be associated with pain and nausea.

1.5 Drug Delivery and Monitoring in Anesthesia

This introduction of clinical anesthesia would not be complete without introducing unfamiliar readers to the different devices that are currently used. Most of the information presented here is from [30] and [2].

1.5.1 Drug Delivery Devices

With the exception of nitrous oxide, all anesthetic agents can be found in liquid form. Their administration is therefore ensured by either a vaporizer, an infusion pump, or directly as a bolus (manual injection).

The Anesthesia Machine The Anesthesia Machine delivers gaseous drugs and vapors (air, N_20) and has the ability to control ventilation. Its role is essential to ensure patient safety in terms of adequate respiration and blood oxygenation (even for procedures where no inhaled anesthetic is used, patients are intubated in order to support their breathing through mechanical ventilation).

The anesthesia machine comprises four distinct devices: an oxygen and N_2O flow meter, a vaporizer, a breathing circuit, and a mechanical ventilator. The vaporizer mixes the anesthetic vapor with fresh gas. A knob permits the anesthesiologist to set the anesthetic vapour concentration at the outlet of the vaporizer. The breathing circuit provides the patient's lungs with the gas mixture via a face mask, the endotracheal tube or LMA. Pressure, volume, gas and vapour composition can be monitored and manually controlled by the anesthesiologist. Sophisticated breathing circuits can maintain temperature, water content, and usually incorporate a CO_2 absorption cannister to remove the excess carbon dioxide in close circuit systems (systems in which the expired gas is recirculated for maximum efficiency in terms of anesthetic gas usage). The breathing cycle can be either controlled directly by the patient (spontaneous ventilation) or by a ventilator (mechanical ventilation) which maintains the respiration set by the anesthesiologist.

Patients undergoing general anesthesia and unconscious sedation in the ICU are often connected to a mechanical ventilator. However, mechanical ventilation has truly an active role in two cases:

- **when using opioids**: large doses of opioids can significantly depress the respiration cycle. The brainstem mechanism regulating the blood CO_2 concentration becomes less sensitive, thus inducing a slowing down of the respiration rate and a reduction of the inspired volume. In some cases, the administration of opioid boluses can be followed by a prolonged apneic state during which proper oxygen concentration in the brain could be seriously compromised. The mechanical ventilator then assumes the role of the autonomic function by ventilating the patient.

- **when using NMBs**: NMB drugs block the transmission of synaptic motor information. Hence, even though the autonomic respiratory function sends the appropriate activation signals to the lungs, their action is inhibited and no respiratory movement follows. As a result, mechanical ventilation is necessary.

Using the ventilator, the anesthesiologist can decrease the blood CO_2 concentration below the threshold where the autonomous respiratory function does not drive respiration. This can be helpful when the breathing pattern needs to be very regular and thereby avoid the use of NMBs or large opioid doses.

In terms of anesthesia delivery, the anesthesia machine is mainly used to maintain the anesthetic adequacy during the maintenance phase. It is generally not used to induce the patient.

Infusion Pumps Most intravenous drugs are administered in the form of boluses to correct for inadequacies in the anesthesia regimen (*e.g.*, during period of intense surgical stress). However, to limit hemodynamic depression, post-operative nausea and airway irritation, infusion pumps are used to supplement the anesthetic gas with a continuous infusion of intravenous anesthetic such as Propofol. While opioids can also be delivered via an infusion pump, this practice is rare in the OR, and is limited to specific cases (*e.g.*, neurosurgery).

The infusion pump usually comprises a syringe plunger actuated by a high precision stepping motor. In newer versions, the anesthesiologist can input the type of drug infused, the syringe caliber, the patient's weight, and the desired infusion rate, usually in $\mu g \cdot \text{min}^{-1} \cdot \text{kg}^1$ or in $\text{ng} \cdot \text{min}^{-1} \cdot \text{kg}^1$. Based on this information, the pump automatically controls the rate of descent of the plunger. An alarm warns the practitioner when the plunger reaches the end position. These pumps are also usually fitted with a RS232 serial port. Using the proper communication protocol, it is possible to remotely control the pump from an external device such as a laptop. In some countries, TCI devices can be used. These devices target the blood plasma concentration by calculating the infusion profile according to the pharmacokinetic model of the infused drug.

1.5.2 Anesthesia Monitors

Standard of care during anesthesia procedures includes the use of a number of sensors:

- Pulse Oximeter: a light transducer clamped conveniently on one of the patient's fingers permits the calculation of heart rate and oxygen saturation.

- Capnograph: a small tube connects the endotracheal tube or LMA to the capnograph. A sample of the gas expired by the patient is continuously analyzed. The capnograph permits the measurement of the end-tidal CO_2.

- Mass Spectrometer: this device is integrated to the breathing circuit. Samples of the inspired and expired gas mixture are analyzed by mass spectrometry to determine the inspired and expired con-

centrations of the anesthetic gas. This is particularly convenient to calculate the quantity of the anesthetic agent that enters the patient's arterial blood stream.

- Hemodynamics Monitors: ECG leads located on the patient's chest record the Electrocardiogram (ECG). The ECG signal is mainly used to determine the heart rate. However, the very shape of the ECG (referred to as QRS complex) can be indicative of cardiovascular difficulties. The ECG is displayed continuously. A pressure cuff also allows the measurement of the brachial blood pressure. A rise or fall in systolic or diastolic blood pressure also indicates cardiovascular difficulties. This measurement is intermittent in order not to block systemic circulation in the instrumented limb.

The information gathered by these sensors is displayed on a unique screen, usually referred to as the *Anesthesia Monitor*. The anesthesiologist's expertise in interpreting these readings permit the assessment of the patient's state and needs.

1.6 Summary

Paradoxically, surgeons achieve healing by first inflicting injury. Anesthesiologists use anesthetics and opioids to prevent the awareness of pain and attenuate the body's stress response to injury. The unconsciousness produced by general anesthesia is accompanied by a depression of the respiratory, cardiovascular and endocrine responses to surgery. Since the degree of surgical stimulation changes during surgery, anesthesiologists must constantly adjust the extent of anesthetic depression to avoid both under- and overdosing. Otherwise, excessive activation of the sympathetic nervous system or pharmacological depression could in turn lead to injury of critical organs, especially in patients with limited respiratory and cardiovascular reserves.

Until the mid-20^{th} century, only inhaled anesthetics were available to create the state of general anesthesia. However, their onset of action was slow and often accompanied by vomiting and signs of respiratory irritation. In the 1940s, the introduction of intravenous agents revolutionized the medical specialty of anesthesia, and over the past 50 years, further refinements in anesthetic pharmacology, equipment and technology have occurred. Nowadays, anesthesiologists have access to agents that can act within a minute of their administration and can block specific mechanisms such as cognition, awareness, memory, stress response and muscle movement. These agents are further characterized by their fast metabolism and elimination. Hence, constant monitoring of their titration is necessary to provide patients with an adequate drug regimen during surgery.

Chapter 2

Quantifying Depth of Anesthesia: a Review

In today's anesthesia practice, patient monitoring is ensured by a set of transducers to measure cardiovascular and respiratory parameters. These measurements enable experienced anesthesiologists to assess their patients' state and take appropriate actions. For instance, a rise in blood pressure and heart rate is usually associated with surgical stress. Conversely, a drop in blood pressure and respiratory rate is often associated with unnecessary overdosing. However, these physiological measurements must be interpreted only within the surgical context. Movements and other visual cues (lachrimation, grimacing, etc.) are often the ultimate warning signs of an inadequate anesthetic regimen. Immediate corrective actions must then be taken to avoid intra-operative awareness.

With the introduction of NMB drugs in the early 1940s, this safety net disappeared. A growing number of intra-operative cases associated with the use of NMBs were reported. The search for a Depth of Anesthesia (DOA) monitor to quantitatively assess the anesthetic state of the patient became an important focus in research.

In terms of close loop anesthesia, the availability of a feedback measure is a *sine que non* condition to the success of such an endeavor. While we have developed our own solutions (see Chapter 5), this chapter presents a review and gives a historical perspective on this very active research field.

Some notions about the use of surrogate measures are presented in Section 1. We will then take a more in-depth look into the techniques that have been developed to measure hypnosis. One of these technique, based on bispectral analysis, was first commercialized in 1996 and has received considerable attention from the clinical community. The experience gained by anesthesiologists in using this type of monitor sheds some light on the eventual advantages of automation in anesthesia. As such, the clinical relevance, interests and potentials of this type of monitoring is reviewed in Section 3. Finally, we will conclude this Chapter with a short section dedicated to the research in the field of analgesia measurement. This particular topic has not received much attention as of yet, which explains the rather limited amount of available information.

2.1 The Use of Surrogate Measures

Hypnosis and analgesia are the result of different pharmacological mechanisms within the CNS. It is not possible to measure them directly. However, there are physiological signs that can be sufficiently correlated with these two anesthesia endpoints. For instance, it is known that cortical activity mirrors the state of hypnosis for a patient. The fact that the brain can or cannot process sensory information can be observed in the EEG. Similarly, it is a well known fact that opioids depress the respiratory autonomic response of the body in a dose dependent fashion. Monitoring the changes in respiratory rate and blood CO_2 can thus provide an indirect measurement of the opioid effect.

The use of surrogate measurements to assess hypnosis and/or analgesia is subject to a limited therapeutic window. When using the EEG to determine the effect of anesthetics and opioids, great care has to be taken into defining a proper therapeutic window. For instance, in their study of opioids, Billard and Shafer [31] observed that the EEG could represent adequately the opioid effect, but only for a limited range of doses, see Fig. 2.1. Small doses do not have a marked influence on the EEG, while larger doses used for cardiac surgery reach a maximum EEG effect. Therefore, the use of the EEG seems to be irrelevant in preoperative and postoperative ICU (where often only a shallow analgesic state is provided) and in surgeries where cardiac stability is ensured by large opioid amounts.

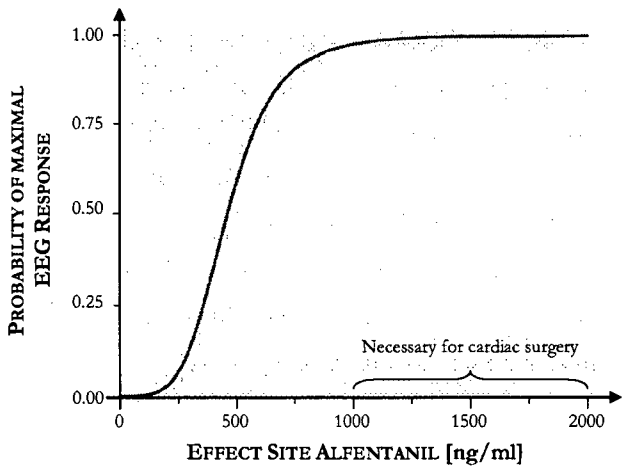


Figure 2.1: Therapeutic window of the EEG for measuring the opioid effect (from [31]).

Various authors ([32], [33], [34]) have been trying to define the characteristics of an ideal index of Depth Of Anesthesia (DOA). These authors were referring to either hypnosis or analgesia. They all agree on the following characteristics that an ideal index of hypnosis must possess:

i. vary in a predictable and consistent fashion as the drug concentration increases,

ii. show similar changes for different agents of the same family (anesthetics, opioids),

iii. indicate shallow levels of hypnosis or analgesia,

iv. have a stable baseline value that varies only minimally,

v. characterize the maximal change induced by the drug, and,

vi. recover to the baseline value on discontinuation of the drug, or upon termination of the effect.

2.2 Quantifying the Depth of Hypnosis: a Review

The Central Nervous System, and in particular the brain (cerebral cortex) are the target organs of anesthetic drugs. Since sensory perception and consciousness are processed at this level, it is reasonable to assume that electroencephalogram signals (EEG) - which reflect cortical activity - can be used to determine the depth of drug-induced unconsciousness.

The effects of anesthetic drugs on the EEG have been known since the early 40s [35] where neurophysiologists observed that the EEG of anesthetized patients contain slower waves with higher amplitudes, see Fig. 2.2. Later on, Faulconer and Bickford [36] published a review cataloguing the different patterns of the EEG associated with ranging depths of anesthesia for different anesthetic agents. Bickford also showed that the EEG measured from the scalp is similar and synchronous to the EEG that can be measured at depth in the human brain. An extensive review published in 1973 by Clark and Rosner [37] concluded that all anesthetics have an effect on the EEG.

But even though all studies have shown that there is a clear correlation between the EEG and the anesthetic depth, they have also shown that patterns generated through anesthesia were different according to the drug that was used. Some researchers concluded that neuroelectric recordings couldn't provide a simple and uniform measure of anesthetic depth. Stanski [33] concluded that in order to use parameters derived from the EEG as a measure of the anesthetic depth, one must first study, for each drug, the correlation between the plasma concentration, the EEG effect and the resulting clinical anesthetic state of the patient. He also stressed out that the pharmacokinetic and pharmacodynamic model of the drug must be associated with the parameter to accurately represent and predict the anesthetic depth.

2.2.1 Review of Tools and Techniques

There are a number of signal processing tools and techniques available to quantify the EEG in order to derive a surrogate measurement of hypnosis. Here is a brief summary of these tools:

2.2.1.1 Time Domain Methods

Bickford, Faulconer and Soltero ([38], [39], [40]) quantified the energy of the EEG by rectifying and integrating the signal. When the resulting integrated signal reached a pre-specified value, the system generated a pulse and reset the integrator. The frequency at which pulses were generated was an indication of the hypnotic depth.

Later on, Bellville [41] developed a method using cyclopropane anesthesia. He observed that increasing concentration of cyclopropane was also increasing the EEG amplitude. He thus used the raw amplified

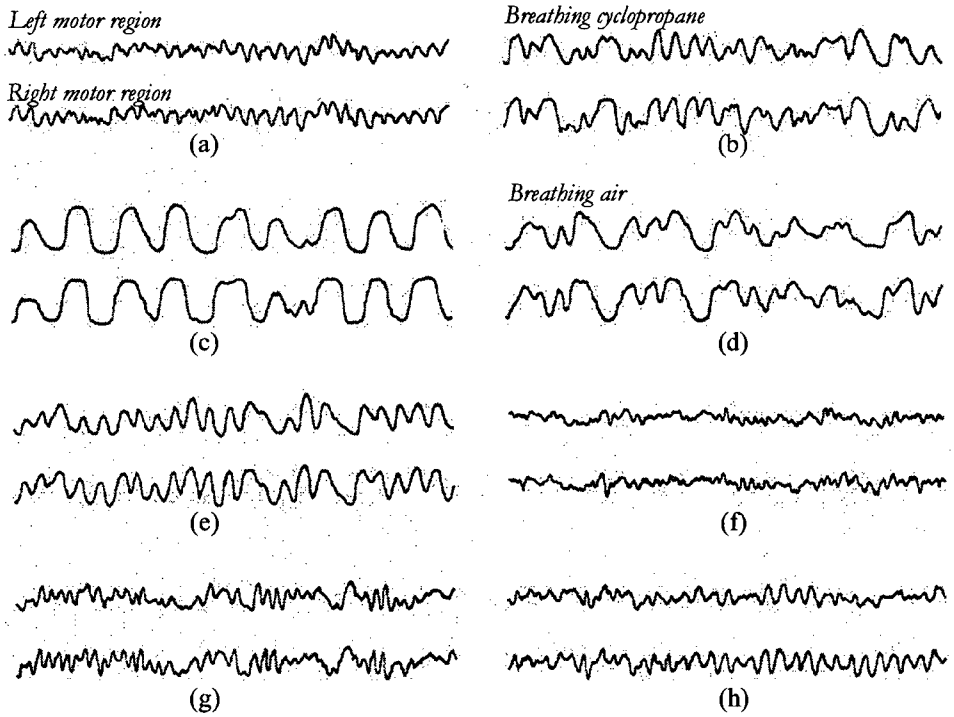


Figure 2.2: Simultaneous records from the left and right motor regions, illustrating the changes in frequency during cyclopropoane anesthesia and recovery. (a) awake (b) breathing cyclopropane for 2 min. (c) 1 min later (d) immediately on substitution of room air for cyclopropane (e) 3 seconds later (f) 2 min. later (g) 5 min. later (h) complete recovery (from [35]).

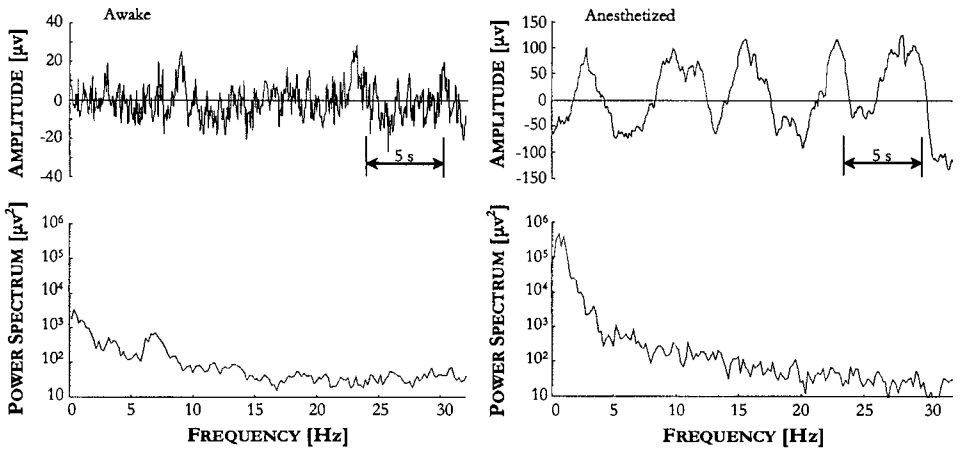


Figure 2.3: Awake and anesthetized EEG and their power spectrum.

signal directly as a feedback quantity.

2.2.1.2 Power Spectrum Analysis

With the development of microprocessors and signal processing tools, researchers have focused their attention on Fourier analysis of the EEG. Power spectrum analysis is used to obtain a frequency distribution of the EEG. Hence, any change in the frequency content of the signal can be visualized. Pichlmayr *et al.* [42] published a thorough review of the effect of the different anesthetic agents on the EEG spectral distribution. The EEG spectral distribution is characterized by different modes. It is common practice to distinguish between 5 frequency bands: δ band (0.25 Hz - 3.5 Hz), θ band (3.5 Hz - 7.5 Hz), α band (7.5 Hz - 12.5 Hz), β band (12.5 Hz - 32 Hz), and γ band (32 Hz - 70 Hz). For a normal awake patient, the EEG activity is principally concentrated in the delta and alpha bands. With increasing level of anesthetics, the activity of the alpha band tends to reduce, while the low frequency content of the delta band is increased, see Fig. 2.3.

To quantify the effect of anesthetics on the EEG, researchers have tried to derive univariate indexes based on Fourier analysis that characterizes the spectral distribution. Among the parameters that have been thoroughly investigated, we can mainly distinguish between the following two:

Median Frequency (MEF) The MEF is the frequency that splits the power spectrum distribution into two parts of equal power. This descriptor is advocated by Schwilden and co-workers ([43], [44], [45] and [46]) who used it for close loop anesthesia.

Spectral Edge Frequency (SEF) The SEF is the frequency at which 95% of the EEG power is present. This descriptor was proposed in 1980 by Rampil *et al.* [47]. According to its authors, the SEF is highly repeatable but presents large inter-subject variations. One advantage of the SEF over the MEF is its high sensitivity to deepening hypnotic levels. As a result, this descriptor is still being used today, in research.

Levy [48] argued that, since EEG activity exhibits multimodal patterns, no univariate descriptor based on the processed EEG can serve as a consistent and adequate representation of the power spectrum. Levy thus doubted the reliability of univariate descriptors of DOA based on the EEG power spectrum.

2.2.1.3 EEG modeling

The idea of modeling the EEG using auto-regressive techniques dates back to the early 1970s. But its application to anesthesia is more recent ([49]).

As compared to univariate descriptors, auto-regressive (AR) modeling generates a set of parameters that can further be correlated to anesthetic depth. In order to derive a single index from the AR parameters, a neural network is trained. Sharma *et al.* [50] showed that this technique can lead to accurate results for measuring analgesia in dogs using halothane. However, the size of the proposed network is rather large (43 inputs, 43 nodes in the primary hidden layer, 6 nodes in the secondary hidden layer and one output). One of these inputs is actually the MAC value. Results are seriously compromised when removing this information, thus indicating that the anesthetic depth measured by this technique depends heavily on the dose-response characteristic of the inhaled anesthetic.

2.2.1.4 Bispectral Analysis: The BIS™ Monitor

Rampil, a leading researcher in the field of neuro-anesthesia, recently argued that anesthetic agents tend to synchronize the generation of postsynaptic potentials [51], thus resulting in slower waves of higher amplitude in the EEG. With increasing levels of anesthesia, it is expected that some of the frequency components of the EEG will shift in time (change of phase). This phenomenon should be mostly predominant in the lighter anesthetic states. This change in latency is not observable by spectral analysis, as phase information is usually discarded.

To illustrate this point, Bowles *et al.* [52] proposed the following example: let us consider a signal composed of 3 sinusoid waveforms. The phases of two of the frequency components are fixed, while the phase of the third one is allowed to slowly drift (see Fig. 2.4). As a result, the shape of the signal itself changes according to the phase of the third component. Clearly, even though the shape of the signal is different, standard power spectral decomposition is not able to capture the phase drift. However, when calculating the bicoherence value, a large discrepancy can be noted when the phase of the third component is coupled with the phases of the first two components. Hence, bispectral analysis can differentiate between these two signals and capture phase drifts.

Ning *et al.* [53] applied bispectral analysis to the EEG in order to characterize sleep patterns in rats. They realized that there was a strong coupling between the frequencies of 6 and 8 Hz during Rapid-Eye Movement (REM) sleep. Sleep patterns being close to patterns obtained during anesthesia procedure, Ning assumed that this technique might lead to interesting results in monitoring the depth of anesthesia. Their

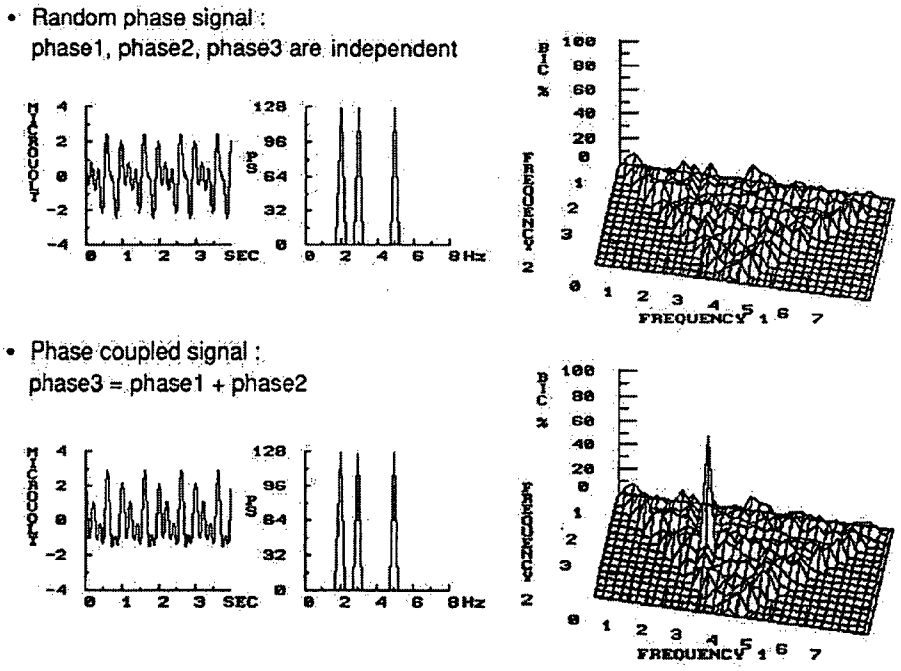


Figure 2.4: Changes in latencies in a signal can be tracked by the bicoherence index (from [52]).

assumption was validated in 1990, when Kearse *et al.* [54] reported that an index based on bispectral analysis was more accurate than the spectral edge frequency for opioid-induced anesthesia (alfentanil and sulfentanil). These findings were confirmed by Sebel *et al.* [55] with isoflurane and later on with propofol [56]. A study by Muthuswamy and Roy in 1993 on anesthetized dogs (halothane) also led to similar conclusions [57]. However, they also noticed that the accuracy of the analysis could probably be improved by using bicoherence indexes in association with other parameters such as the MEF, SEF and hemodynamic parameters.

Probably the most compelling result was obtained by a research team from Aspect Medical Systems Inc. who derived no less than 33 variables (11 bispectra, 11 bicoherence indexes, 11 power spectral values). These variables were combined into a dimensionless index in order to predict somatic responses to surgical incision [52] [1]. This index, referred to as BIS (Bispectral Index Scale), is a weighted sum of each spectral and bispectral variable. In the earliest versions, the weights were tuned using discriminant analysis and based on data collected from 170 standard surgical procedures.

In comparison with other methods (SEF, MEF, etc.), the accuracy of the BIS was significantly higher. A major advantage of bispectral analysis over more conventional techniques is its wider therapeutic window that allows the differentiation between different levels of sedation [58].

Based on these findings, Aspect Medical Systems has developed a monitor of consciousness, referred to as the BIS Monitor, which integrates this technology. A single EEG signal is recorded through contact

[1] Note once again the confusion between hypnosis and analgesia.

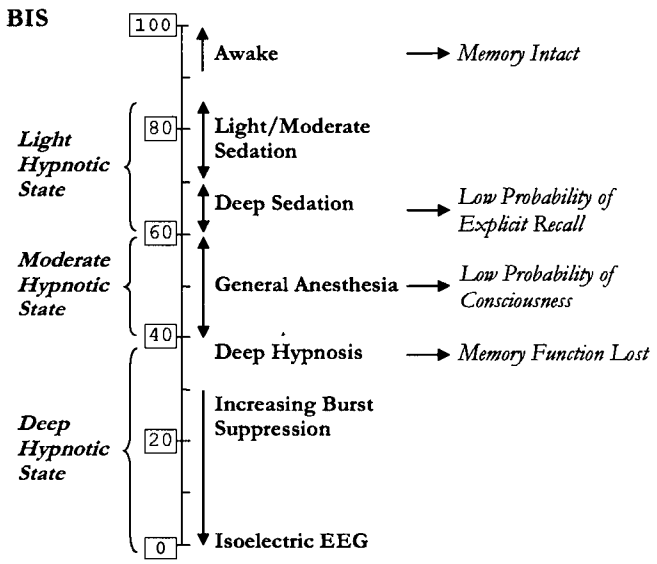


Figure 2.5: Bispectral Index Scale and its meaning (from [59]).

electrodes placed on the patient's forehead. The BIS value and other relevant information are then displayed on the monitor screen. To simplify the interpretation, the BIS is scaled between 0 and 100. A value of 100 represents the awake state. With increasing concentration of anesthetics, the index decreases. General anesthesia is obtained for an index between 60 and 40, see Fig. 2.5. Lower values represent deep hypnotic states, while values between 90 and 60 generally represent sedation levels.

2.2.1.5 Quantitative Evoked Potentials: the A-Line™ Monitor

Somatosensory information provoked by auditory, visual, or tactile stimulation generate transitory oscillatory signals within the EEG itself. It has been advocated that such transient signals, if properly analyzed, can reveal information concerning a patients' state of consciousness ([32, 60, 61, 62, 63]). For instance, *midlatency auditory evoked potentials* (MLAEP) have a very distinct shape whether the subject is awake or asleep, see Fig. 2.6. The most remarkable feature beside the change in amplitude of the signal is the change in latency of the P_a and P_b waves. A very interesting work by Huang *et al.* [64] has shown that it is possible to measure the hypnotic depth of a dog under anesthesia by using wavelet transform of the MLAEP signal and feeding the wavelet coefficients to a properly trained neural network.

A significant work has been carried out in that particular field. One major disadvantage of using evoked potentials is their very low signal to noise ratio, which makes them particularly difficult to acquire, as they are embedded inside the EEG signal. Averaging over a large number of samples is necessary in order to filter out the EEG noise from the signal. New techniques such as wavelet de-noising have been successfully applied to partially alleviate this problem [65].

The latest developments in AEP extraction using ARX modeling have yielded results that allow the

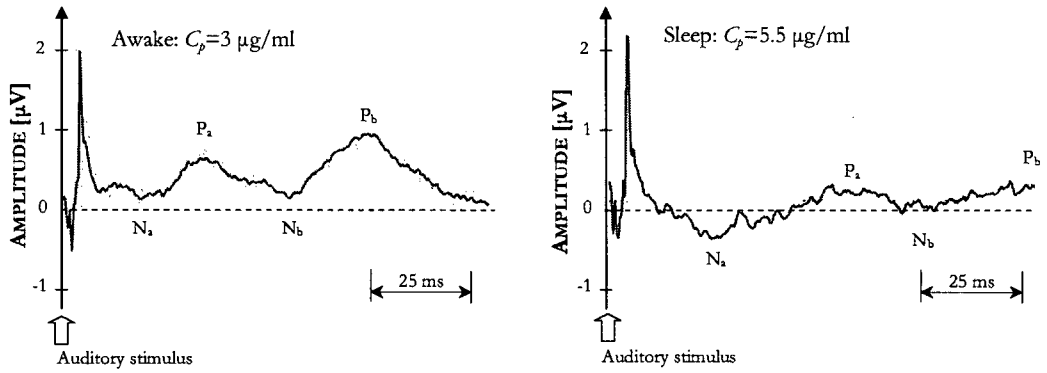


Figure 2.6: Midlatency Auditory Evoked Potentials obtained from a responsive (awake) and non-responsive (asleep) dog subject to tail clamping. The waveforms have been obtained after averaging of 1000 samples (adapted from [64]).

estimation of the anesthetic state based on 6 second epochs [66, 67], thus necessitating the averaging of only 15 sweeps. This new technology has been embedded in the A-Line Monitor (Alaris Medical Systems, California) and is now commercially available.

2.2.1.6 Other Monitors

Following the success of the BIS Monitor, medical companies have taken a keen interest in this new market segment. New monitors have recently become available:

- The PSA 4000™ (Physiometrix, Massachusetts): this monitor displays an index of hypnosis calculated on 1.25 s EEG epochs. Numerous quantitative measurements are derived (power gradients, absolute power in specific frequency bands, etc.). These measurements are further combined into a unique dimensionless index, referred to as the Patient State Index (PSI). The PSI is scaled between 100 (awake) and 0 (isoelectric EEG). Three extensive EEG databases were used for the tuning and calibration of the PSI. While the PSI is based on similar principles as the BIS Monitor (*i.e.* composite index tuned from EEG databases and clinical cases), it differs in that it does not use bispectral parameters. Instead, the PSI focuses on the power shift in specific frequency components between the frontal cortex and the posterior lobes. This power shift is directly related to the hypnotic depth. Paradoxically, the commercial version of the PSA 4000 uses only frontal electrodes.

- The Narcotrend® Monitor (MonitorTechnik, Bad Bramstedt, Germany): this monitor is based on a pattern recognition algorithm that classifies the EEG into 14 different classes using Kugler's classification and denomination [68] (A, B0, B1, B2, C0, C1, C2, D0, D1, D2, E0, E1, F0, F1). The range D0-E1 indicates an adequate depth of anesthesia for surgery. The range F0-F1 indicates burst suppression and isoelectric EEG. The Narcotrend index exhibits step-wise changes during increasing or decreasing hypnotic depth.

- The AS/5 M-Entropy Module (Datex Ohmeda, Helsinski, Finland): Datex Ohmeda is a leading

manufacturer of anesthesia machines and monitors. They recently developed an EEG measure based on entropy analysis that captures the complexity of the EEG signal [69]. During the awake state, the EEG is a highly complex signal, resembling that of noise (no apparent pattern). With increasing depth of hypnosis, patterns emerge, thus reducing the complexity of the signal.

2.2.2 Clinical Relevance, Interest and Potentials of the BIS Monitor

The BIS Monitor has received considerable attention from the research and clinical communities. Over 1,250 papers discussing its pros and cons and its clinical relevance have been published since its commercial introduction in 1996. The BIS Monitor was the only FDA approved monitor of drug-induced unconsciousness until 2002.

Intra-operative Awareness The most publicized role of the BIS Monitor is the reduction of the incidence of intra-operative awareness. However, validating such an assertion is a difficult task due to the rather limited occurrence of such cases. It is only very recently that two large studies were able to assess the impact of BIS monitoring to prevent intra-operative awareness:

- **SAFE2 Trial**: following the SAFE1 trial (see Section 1.1.2.2), Sandin *et al.* conducted a prospective study where 5,057 patients undergoing surgery requiring general anesthesia including the use of muscle relaxants or intubation were monitored with the BIS monitor. Only 2 patients experienced recall of intra-operative events, as compared to 14 in the control group (from the SAFE1 Trial). The authors concluded that the BIS monitor can reduce intra-operative cases by up to 80%. Even though the study population is still rather limited (only 5,000 cases so far), this result was deemed statistically relevant.

- **B-AWARE Trial**: the B-AWARE Trial targets a population of patients at risk undergoing relaxant general anaesthesia. A total of 2,503 patients were enrolled in this study. This population was divided into 2 groups: control, and BIS monitored. In the BIS group, the index was kept in the 40-60 range. The incidence of intra-operative recall was about 0.89% (11 cases) in the control group, and 0.16% (2 cases) in the BIS group.

Clearly, the BIS monitor succeeded in reducing intra-operative awareness in a significant manner.

Drug Usage, Post-operative Acute Care Unit Discharge Times and Outcome Another well-publicized advantage of a BIS guided titration (*i.e.*, maintaining the BIS value in-between a pre-determined range during the maintenance phase) is the reduction of drug usage and discharge time. For instance, several studies [70, 71, 72] concluded that an average reduction of 20% to 30% in the total amount of anesthetic drugs (inhaled and/or intravenous) could be achieved during the maintenance phase by keeping the BIS value in the 40-60 range, while still ensuring the adequacy of the anesthetic state. These studies

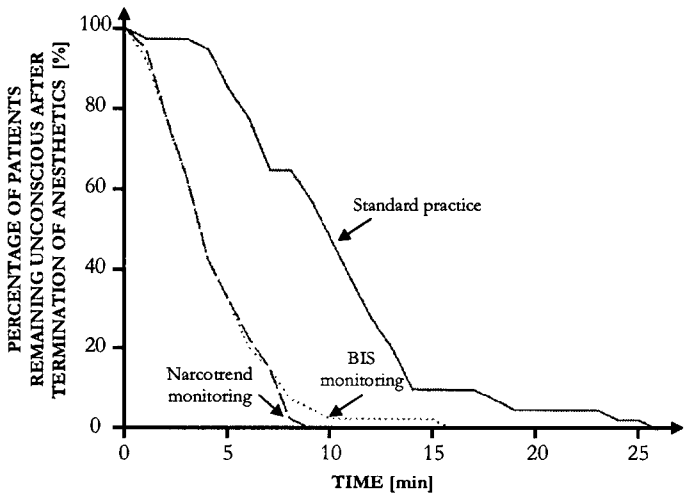


Figure 2.7: Time of arousal following discontinuation of propofol and remifentanil infusion (from [70]).

imply that, on average, patients without BIS are overdosed. Processed electroencephalography variables such as the BIS can thus offer practitioners a guideline to titrate each patient according to his/her specific needs.

With the reduction of drug usage, it has been shown that patients awake sooner from their surgery and are discharged faster than in standard practice [73, 70, 74]. Discharge times in the Post Anesthetic Care Unit (PACU) are reduced by up to 50%, see Fig. 2.7. The incidence of post operative nausea and vomiting, and patients' comfort are also significantly improved.

The reduction in drug usage, discharge times and adverse post-anesthetic events (nausea and vomiting) have the added advantage of reducing anesthesia-associated costs. In today's cost-sensitive society, this could be seen as a major advantage. However, some studies [75] have shown that the savings compensate for the cost of the BIS sensors and the maintenance of the monitor only for long surgeries (>4 hours).

2.2.2.1 Long Term Benefits of BIS Monitoring

In a recent abstract, Weldon *et al.* [76] showed an interesting correlation between one year post-operative death rates and the average BIS value. The study included 907 patients scheduled for major surgical procedures. Logisitc regression modeling showed that increasing age and lower BIS levels were both independently associated with higher one year post-operative mortality rates. In the elderly population, the mortality rate was 16.7% for an average BIS value <40, as compared to 4.2% for an average BIS >60. This study implies that maintaining middle-aged and older patients into very deep anesthetic levels might not be advisable. While this result was later confirmed by Lennmarken [77], details from these studies have still not been released. One could theorize that patients driven into deep hypnotic states received more inhaled anesthetic and perhaps less opioids, and as a result, suffered from cardiovascular instability.

2.2.2.2 Bad Press

There are still many controversies concerning the use and reliability of the BIS monitor. One major argument fuelling the debate is whether a single index derived from the EEG could reflect the anesthetic adequacy in all types of anesthetic procedures. Many argue that in order to validate such monitors one would need to study the effect of each anesthetic and opioid drugs (and combination thereof) on the index. A second argument is that the BIS is tuned to reflect the anesthesiologist's assessment. As such it is an empirical measurement which might not reflect the underlying neurophysiologic mechanism of anesthetics. Finally, while many studies have concluded that a BIS guided titration results in lower drug usage and improved outcome, some consider that such titration would increase the risk of intra-operative awareness [78]. However, this argument seems to have been refuted by the recent B-AWARE trial.

Particular cases where the BIS has shown paradoxical behavior were also reported:

- **Low BIS values during recovery:** Sleigh and Donovan [79] report that 14 out of 37 patients could respond and follow verbal commands during emergence while the BIS value was <75. They observed that the BIS value regains its baseline value during emergence only 30 s after patients start reacting to verbal commands. They attributed this delay to the update delay of the BIS and its 30 s averaging window. A similar observation was later on reported by Lehmann [80].

- **Effect of EEG amplitude on the BIS:** Schnider *et al.* [81] reported that a consistent BIS value of 40-50 can be obtained if the amplitude of the EEG is low enough. Muncaster *et al.* [82] in 2003 also reported having a patient with a very low BIS value during emergence due to a low EEG amplitude.

- **NMBs and BIS:** The effect of NMBs on the BIS value has also recently been investigated. In one study on volunteers, Messner *et al.* showed that a paralyzed fully conscious subject can have a BIS value as low as 9 [83]. If confirmed, this could be a very damaging result for a monitor that is supposed to detect episodes of consciousness in paralyzed patients. A similar finding was reported by Vivien *et al.* [84]. In their study of 45 patients in intensive car unit, they measured an average drop of 25% in the BIS value after the administration of NMBs. A similar decrease in electromyographic (EMG) activity was measured. A conclusion from these studies is that the BIS value is strongly correlated to the EMG and should not be used in sedated patients whenever NMBs are used concurrently. Note however that NMBs do not affect the BIS value in patients under general anesthesia [85]. This could be due to an already very low EMG activity, where the NMB impact remains very limited[2]. One can therefore hypothesize that the BIS value cannot reach levels above 60-70 without some EMG activity. As a result, paralyzed patients should not be allowed to be maintained at levels close to 60 since an episode of intra-operative awareness might not be picked up by the monitor[3].

[2]The contribution of the EMG signal in the EEG remains unaffected by the NMB (however the drug still effectively blocks muscle movements).

[3]Personal interpretation.

Finally, a cornerstone of the mistrust and recurring complaint by clinicians is that the BIS algorithm has never been made public, and hence cannot be rigourously evaluated by independent researchers [86].

2.2.2.3 Comparison with Other Monitors

Because of the relative novelty of the other monitors, only a few clinical studies have been published comparing the bispectral index with its competitors. So far, results from the PSA 4000 [87] and the Narcotrend Monitor [70] seem to show a good correlation between these indexes and the BIS. No clear advantages have been found in using one particular technology, and both the PSA 4000 and the NarcoTrend monitor reported similar findings in terms of drug usage and PACU discharge times.

2.3 Quantifying the Depth of Analgesia

The measurement of antinociception is part of the critical care given by anesthesiologists in their daily practice. Surgical trauma is usually accompanied with strong sympathetic and parasympathetic activity (*e.g.*, heart rate and blood pressure changes, sweating, lachrimation, somatic movements, etc.).

Deriving descriptors of antinociception that can be used as a feedback quantity is a challenge. For instance, blood pressure alone is not a reliable measure, as other parameters such as blood loss and the action of vasoactive drugs can affect the cardiovascular system. As a result, the analgesia functional component hasn't yet received enough attention. In recent years, only a few monitors have been introduced, without full disclosure of their scientific base.

2.3.1 Using Physiological Measures

2.3.1.1 The EEG as a Measure of Analgesia

The effect of opioids on the EEG has been thoroughly investigated. In 1984, Smith *et al.* [88] concluded that the EEG probably reflects the depth of anesthesia with high-dose narcotics. Later studies by Scott *et al.* ([89] and [90]) resulted in similar conclusions. The same techniques (MEF, SEF, BIS) described previously can be used for measuring the effect of opioids as well. However, the therapeutic window of these techniques is very limited. Also, many agree today that cortical activity, while appropriate for consciousness monitoring, cannot be used to assessed the effect of opioids since their site of action is principally localized at the level of the spinal cord [91].

2.3.1.2 End-tidal CO_2

All opioids depress ventilation in a dose-dependent fashion. As a result, the amount of carbon dioxide (CO_2) in the arterial blood during spontaneous ventilation is an indication of the opioid effect. The advantage of using the CO_2 arterial partial pressure is that it can be easily measured using the capnograph. However, the usefulness of this measurement is limited to cases where patients breathe spontaneously (a majority of

day care surgeries), thus limiting the therapeutic window of this parameter to doses of opioids that do not provoke apnea.

2.3.1.3 Heart Rate and Blood Pressure

A rise in heart rate and blood pressure is usually indicative of increased sympathetic activity (*i.e.* stress). Monitoring these parameters can be useful to determine the level of stress. However, heart rate and blood pressure baselines are patient-dependent. Also, numerous drugs other than opioids can depress or alter these signs. Finally, changes in heart rate and blood pressure are only indicative of strong stimulation and do not provide a graded response to increasing stimulation.

2.3.2 Using Heart Rate Variability

In healthy subjects, parasympathetic and sympathetic activity are balanced according to stress levels, circadian rhythms, position, etc. This balance is regulated by the CNS. In most subjects under stress, sympathetic activity dominates. Conversely, relaxation increases parasympathetic activity and depresses sympathetic tone.

While the field of Heart Rate Variability (HRV) is quite prolific, most of the research done till now was oriented towards the development of diagnosing tools for awake subjects. In standard analysis of the HRV, researchers differentiate between the HF band (0.2 - 0.5 Hz) which reflects parasympathetic activity, and the LF band (0.01 - 0.15 Hz) which is believed to reflect both sympathetic and parasympathetic activity (the exact mechanism remains unknown). High frequency activity is usually concentrated around the respiratory frequency. One explanation could be that the regulation of respiration is mediated by the same center that regulates the activation of the vagal nerve. Typically, the HF measurement usually quantifies the activity concentrated around ± 0.06 Hz around the respiratory frequency. In order to normalize this measurement, standard practice usually involves the ratio LF/HF, which is believed to yield a more consistent measurement of the sympathetic/parasympathetic balance.

The effect of anesthesia on the HRV signal has not yet been thoroughly investigated. A standard measurement practice for instance has not been adapted, since the respiratory rate in anesthetized patients can drop to a point where both the HF and LF activity share the same frequency band. Hence, it seems that normalizing one with respect to the other is not indicated.

2.3.2.1 Existing Technologies

Two monitors have been advocated as analgesia monitors: the Anemon I, Medical Control System SA, and the Fathom, Amtec. There is, however, no published literature reviewing the merits of these technologies.

The Fathom monitor is still at the prototype stage and is not yet commercially available. This particular device analyzes the respiratory pattern fluctuations. These fluctuations are difficult to measure. However, they strongly correlate with fluctuations of the cardiac rhythm called sinus arrythmia. Observing the

changes of the interval between R-waves of the electrocardiogram and the respiration cycle forms the basis for this measurement ([92] and [93]). Clinical trials are currently being conducted in England to validate the efficacy of this technique.

2.4 Summary

One hundred and fifty years after the discovery of anesthesia, the current standard of care in anesthesia management still does not include monitors to assess quantitatively the effect of anesthetic drugs on the target organ, *i.e.*, the brain. The current practice still relies uniquely on secondary signs to warn the practitioner of either pharmacological toxicity or anesthetic inadequacy.

It is only in the past 10 years that serious advances have been made. Since 1996, practitioners have access to monitors to measure the degree of drug-induced unconsciousness. As of today, monitors, such as the BIS, are still considered by many to be nothing more than gadgets. Also, the price of the BIS sensor deters many to use the device in their everyday practice. Further, the lack of predictability in the BIS behavior during transients (*e.g.*, unconscious patients after induction can have a stable BIS >80 while a reacting patient after emergence can have a BIS <60) has brought many practitioners to question its reliability. However, with the new awareness studies, the use and interest of such monitors might be re-evaluated once the economic and social impact of intra-operative awareness are properly studied.

In terms of analgesia, no real breakthrough has been achieved yet. This can be explained by the fact that anesthesiologists can assess this state better than hypnosis, since they monitor cardiorespiratory functions.

Chapter 3

Modeling Anesthetic Drugs: the Traditional Approach

In this chapter, we present the drug modeling approach favored in traditional pharmacological studies. In this approach, it is usual to first consider how the administered drug distributes within the body. This analysis leads to a *pharmacokinetic model* (PK) which can be used to predict the blood plasma concentration of the drug.

The second step is to relate this concentration to the drug effect itself. This yields a second mathematical expression referred to as the *pharmacodynamic model* (PD).

The organization of this chapter mirrors this approach. The pharmacokinetic concepts are covered in Section 3.1, while pharmacodynamic modeling is discussed in Section 3.2.

Propofol and remifentanil being the fastest anesthetic and opioid agents available in the current practice, we believe that a close loop anesthesia delivery system will benefit the most from these agents. Hence, most of the discussion presented here is focused on these two drugs.

3.1 Pharmacokinetics

Anesthesiologists act through the administration of drugs. A pharmacokinetic model of a drug is a mathematical expression relating the drug blood plasma concentration $C_p(t)$ to the administered dose $I(t)$. The aim of this section is thus to define the transfer function[1] $PK(s)$:

$$PK(s) = \frac{C_p(s)}{I(s)} \tag{3.1}$$

While special emphasis will be given to propofol and remifentanil, most of the concepts presented here can be broadened to a wider variety of drugs. For a more in-depth look into these concepts, we invite

[1] In contrast with the usual approach in pharmacology, $PK(s)$ is expressed here in the Laplace domain (or frequency domain). In linear system theory, the Laplace transform is used extensively to express time-domain differential equations into algebraic expressions, which are thus easier to solve. Continuous time domain system analysis and control design are often conducted in the Laplace domain.

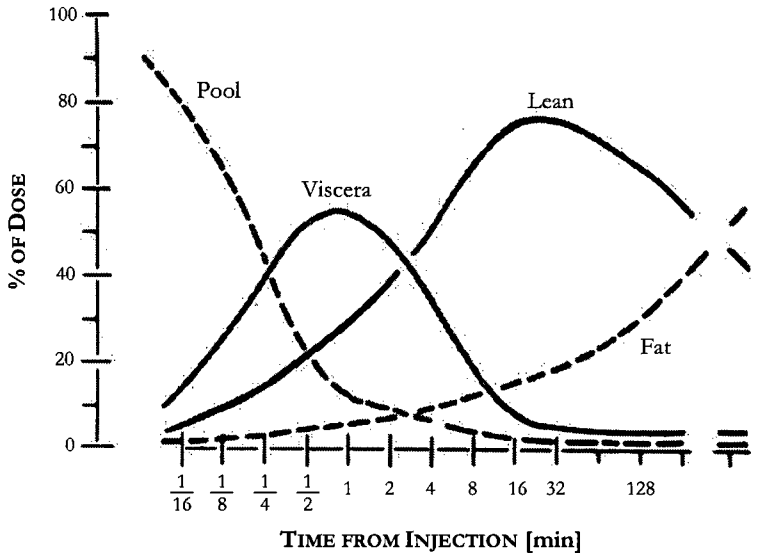


Figure 3.1: Simulation of the time course of the percentage of a thiopental bolus in the central blood pool, viscera, lean tissue and fat as a function of time, assuming no elimination clearance (adapted from [97]).

interested readers to refer to the textbooks from Stanski and Watkins [94], Prys-Roberts and Hug [95], and Fragen [96].

3.1.1 Principles and Concepts

Pharmacokinetics can best be understood when considering the time course of the concentration of any given drug within the plasma and other tissues of the human body (see Fig. 3.1). During the absorption phase following an intravenous bolus administration, a drug mixes rapidly within the central blood pool, resulting in a plasma peak concentration. While at first glance the peak plasma concentration seems to occur instantaneously, a small delay elapses between the actual injection of the drug and its mixing within the blood pool. Systemic circulation then distributes the drug to a variety of tissues within the body. A part of the drug enters these tissues through molecular diffusion, according to the affinity of the tissues for the drug, the rate of perfusion and the relative concentration of the drug in the blood and in the tissues. Highly perfused, relatively low volume tissues such as the brain equilibrate rapidly with the concentration of the central blood pool. Since the site of action of anesthetics and opioids is the *central nervous system* (CNS), this explains the rapid onset of effect of such drugs. Concurrently to this equilibrium, a quantity of the drug will also pass from the blood through other tissues such as muscles, viscera, fat and bones, hence slowly reducing the drug plasma concentration. As the concentration decreases, a part of the drug accumulated inside the highly perfused tissues will then be re-distributed inside the blood, hence decreasing the concentration of the drug in these tissues.

The reversible transfer of the drug from one location to another is referred to as *distribution*, and involves molecular movement across lipid membranes and capillary walls. The distribution kinetics depend

then mostly on the lipid solubility characteristic of the drug. Another important aspect that affects the distribution phase of the drug is protein binding. While the bound fraction of a drug does not contribute to the pharmacologic action of the drug, it acts as a depot: as the free fraction of the drug is removed from the body due to metabolism and excretion, protein bindings are broken to reestablish the initial ratio between bound and free drug.

The elimination of a drug is usually achieved through metabolism and excretion. Metabolism transforms lipophilic substances (such as most anesthetics and opioids) into hydrophilic substances, hence facilitating their excretion through renal mediation. This process, done at the hepatic level, usually leads to the inactivation of the drug. There can be a large individual variability in the activity of the enzymes responsible for metabolism. This variability can be as great as 40-times which might be the result of genetic differences between individuals. Renal and hepatic excretion are responsible for most of the drug elimination within the body.

3.1.2 Modeling

The phenomenon of drug uptake, distribution and elimination can be expressed mathematically. To that end, a number of model architectures have been proposed.

3.1.2.1 Exponential Models

For most drugs, the time course of their concentration within the blood plasma after rapid intravenous administration and uptake can be fitted to resemble a decaying function, with two distinct modes corresponding to the distribution and elimination phase respectively, see Fig. 3.2. This behavior can be mathematically expressed as:

$$C_p(t) = A \cdot \mathrm{e}^{-\alpha \cdot t} + B \cdot \mathrm{e}^{-\beta \cdot t} \quad [\mathrm{ng/ml}] \text{ or } [\mu\mathrm{g/ml}] \tag{3.2}$$

where $C_p(t)$ is the time course of the drug concentration expressed in nanogram per milliliter (opioids) or microgram per milliliter (barbiturates, propofol), α is the rate constant of the distribution phase (slope of the tangent at the origin), and β is the rate constant of the elimination phase (slope of the curve once the distribution phase is completed). In many cases, a tri-exponential model will capture significantly better the kinetic of the drug[2]:

$$C_p(t) = P \cdot \mathrm{e}^{-\pi \cdot t} + A \cdot \mathrm{e}^{-\alpha \cdot t} + B \cdot \mathrm{e}^{-\beta \cdot t} \quad [\mathrm{ng/ml}] \text{ or } [\mu\mathrm{g/ml}] \tag{3.3}$$

A major advantage of exponential models is that they can be easily derived using graphical means. The identification can be carried out directly by using either bolus data and analyzing the decaying blood plasma characteristic (such as shown in Fig. 3.2), or by using infusion data and analyzing how the plasma concentration increases over time.

[2]The use of β being usually reserved to model and represent the slow elimination phase, researchers have introduced the notation P and π to describe the fast dynamics corresponding to the distribution phase.

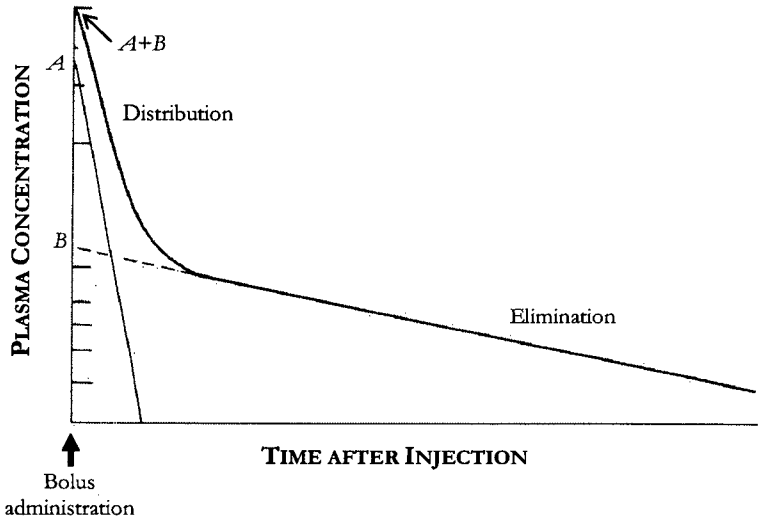


Figure 3.2: Time course of plasma concentration with two distinct phases: distribution and elimination (adapted from [98]).

Note that in pharmacology it is usual to replace the rate constants by *half-lives* which represent the amount of time that is necessary to observe a reduction of 50% of the initial dose during a given phase. Half-lives are calculated as:

$$T_{\frac{1}{2}}^{\pi} = \frac{\ln 2}{\pi} \quad \text{and} \quad T_{\frac{1}{2}}^{\alpha} = \frac{\ln 2}{\alpha} \quad \text{and} \quad T_{\frac{1}{2}}^{\beta} = \frac{\ln 2}{\beta} \quad [\text{min}] \tag{3.4}$$

In terms of control and system engineering, the exponential model (3.3) can be directly expressed in the Laplace domain:

$$PK(s) = \frac{C_p(s)}{I(s)} = P \cdot \frac{1}{s+\pi} + A \cdot \frac{1}{s+\alpha} + B \cdot \frac{1}{s+\beta} \tag{3.5}$$

3.1.2.2 Compartmental Models

Pharmacologists usually consider exponential models to be counter-intuitive and prefer the use of mamillary compartmental models. Since the literature abounds in pharmacokinetic parameters given in the compartmental framework, it is necessary to introduce these models here as well. Note that mamillary compartmental models are used extensively in pharmacology, and are not restricted to modeling only anesthetic drugs.

Pharmacokinetically, a compartment represents a tissue group that has similar kinetic characteristic. For instance, it is possible to consider that the body can be divided into three compartments: a small central compartment that contains the arterial blood and highly perfused tissues such as the brain and liver, a larger compartment which contains muscles and viscera, and a third compartment consisting mostly of fat and bones, see Fig. 3.3. Following the administration of a bolus, the amount of drug delivered into the central compartment is eliminated according to the rate constant k_{10} (usually expressed in min^{-1}). This elimination corresponds to metabolism and hepatic and/or renal excretion. In parallel to the elimination

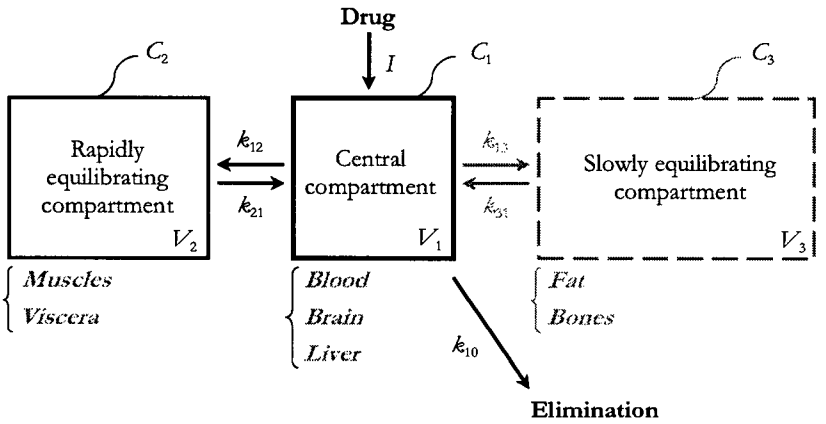


Figure 3.3: 2-compartment pharmacokinetic model. A third compartment is often added for improved accuracy (adapted from [99]).

process, the drug is distributed in the two peripheral compartments at a rate of k_{12} and k_{13}. Following the administration of a drug, the concentration C_1 of the first compartment decreases rapidly while the concentrations C_2 and C_3 of the second and third compartments rise. Once the concentrations in the central compartment and any of the peripheral compartments reach equilibrium, the distributive process reverses and the drug stored in the peripheral compartment returns to the central compartment at the rate of k_{21} or k_{31}. Since the blood of the central compartment acts as a carrier for the drug, we can assume that there is no direct drug exchange between the two peripheral compartments. For a similar reason, only the drug presents in the central compartment can be eliminated.

The mathematical expressions governing this model can be obtained by writing the mass balance equations (expressed here in a state space representation):

$$\begin{cases} \begin{bmatrix} \dot{C}_1(t) \\ \dot{C}_2(t) \\ \dot{C}_3(t) \end{bmatrix} = \begin{bmatrix} -(k_{10}+k_{12}+k_{13}) & k_{21} & k_{31} \\ k_{12} & -k_{21} & 0 \\ k_{13} & 0 & -k_{31} \end{bmatrix} \cdot \begin{bmatrix} C_1(t) \\ C_2(t) \\ C_3(t) \end{bmatrix} + \begin{bmatrix} \frac{1}{V_1} \\ 0 \\ 0 \end{bmatrix} \cdot I(t) \\ C_p(t) = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} C_1(t) \\ C_2(t) \\ C_3(t) \end{bmatrix} \end{cases}, \tag{3.6}$$

where V_1 is the volume of the central compartment. Note that, by definition, the plasma blood concentration equals the drug concentration of the central compartment, *i.e.*, $C_p(t) = C_1(t)$.

Volume of Distribution The *volume of distribution* V_d is a measurement of the extent of the distributive process. It is defined as the ratio between the total amount of the drug in the body versus the drug plasma concentration prior to the beginning of the elimination phase. The volume of distribution does not have any physiological significance, and it is also referred to as the *apparent volume of distribution.*

The volume of distribution is dependent on the size of the tissues into which the drug distributes, the partition coefficient of the drug, and the extent of protein binding. This volume can be extremely large, even exceeding the anatomic size of the body, meaning that most of the drug is distributed inside peripheral tissues.

For a 3-compartment model, the volume of distribution is simply the sum of all three compartment volumes:

$$V_d = V_1 + V_2 + V_3 \quad [\text{l}] \tag{3.7}$$

where V_2 and V_3 are the volumes of the rapidly equilibrating and slowly equilibrating compartments.

Clearance Another important concept used by pharmacologists is the intercompartmental and total drug *clearance* concept. It has been mentioned earlier that the elimination of a drug from the body is achieved through a variety of mechanisms, involving metabolism and excretion. Total drug clearance can thus be thought of as the ability of the body to remove a given drug from the blood or plasma. Clearance is usually expressed in [ml·h^{-1}]. Using a 2- or 3-compartment model, the total body clearance Cl_1 can be calculated as:

$$Cl_1 = V_1 \cdot k_{10} \quad [\text{ml}\cdot\text{h}^{-1}] \tag{3.8}$$

The intercompartmental clearances Cl_{12}, Cl_{21}, and Cl_{13}, Cl_{31} represent a similar concept. Instead of the elimination process, they quantify the ability of each compartment to exchange the drug with one another. Intercompartmental clearances are not very useful. However, considering that $Cl_{12} = Cl_{21} = Cl_2$ and $Cl_{13} = Cl_{31} = Cl_3$, and that $Cl_{ij} = V_i \cdot k_{ij}$, we reach the following equalities:

$$\frac{V_2}{V_1} = \frac{k_{12}}{k_{21}} \qquad \frac{V_3}{V_1} = \frac{k_{13}}{k_{31}} \tag{3.9}$$

3.1.2.3 Equivalence Between Exponential and Compartmental Models

The models expressed in (3.3) and (3.6) are essentially equivalent as they describe the same input/output relationship. Because of this equivalence, a PK parameter set can take different forms:

- $\{P, A, B, T_{\frac{1}{2}}^{\pi}, T_{\frac{1}{2}}^{\alpha}, T_{\frac{1}{2}}^{\beta}\}$: standard exponential parameter set,
- $\{V_1, P^*, A^*, B^*, T_{\frac{1}{2}}^{\pi}, T_{\frac{1}{2}}^{\alpha}, T_{\frac{1}{2}}^{\beta}\}$: exponential parameter set with normalized partial coefficients (*i.e.*, $P^* + A^* + B^* = 1$),
- $\{V_1, k_{10}, k_{12}, k_{13}, k_{21}, k_{31}\}$: standard compartmental parameter set,
- $\{V_1, V_2, V_3, Cl_1, Cl_2, Cl_3\}$: compartmental parameter set expressed as volumes and clearances.

All of these sets are equivalent as they define the same PK relationship. Most of the PK parameter sets found in the literature are expressed in either of these forms.

In order to simplify the notations, we propose to express the PK model as a SISO transfer function $PK(s)$ using both the exponential and compartmental parameters:

$$PK(s) = \frac{C_p(s)}{I(s)} = \frac{1}{V_1} \cdot \frac{(s+k_{21}) \cdot (s+k_{31})}{(s+\pi) \cdot (s+\alpha) \cdot (s+\beta)} \tag{3.10}$$

This hybrid notation has the advantage of clearly expressing each one of the system modes, hence allowing a trained observer to readily identify the frequency response of the system.

3.1.3 Literature Review

Medical literature is quite prolific in reviewing drugs and their PK properties. In this section, we present a summary of a survey of published PK parameter sets for propofol and remifentanil. The complete survey is presented in the Annex B. Each model is summarized under the hybrid form of (3.10) in Table B.1 and B.3.

Propofol Propofol has received considerable attention over the past 20 years. A large number of pharmacokinetic models have been published by various authors, and under various conditions and methods of identification.

In the mid-1990s, researchers started to use the NONMEM[3] software (NONMEM Project, University of California, San Francisco, CA) to derive the PK models from data sampled across a large population of patients. This software package allows the user to identify the most significant covariates and include them in the final model. In terms of propofol, the age, weight and Lean Body Mass (LBM) of patients were found to alter the PK parameters. The sampling site (venous *vs.* arterial) and the method of administration (bolus *vs.* infusion) were also significant covariates. Schnider *et al.* [100], and later on Schüttler and Ihmsen [101], used this technique in their analysis.

As of today, Schüttler and Ihmsen PK sets are the most up-to-date. In their study (spanning hundreds of patients and thousands of blood samples) the intra-patient variability was found to be less than 20%. The mean absolute weighted residual[4] (MAWR) was about 25% and the mean weighted residual (MWR) was -3.4%. The PK parameter set is presented in the Table 3.1 and Table 3.2.

Remifentanil Remifentanil being a relatively new drug, the derivation of its PK model has directly benefitted from the NONMEM analysis software. As such, the models published by Egan *et al.* in 1996 [102], and later on by Minto *et al.* [103], are consistent and accurate (the MAWR was less than 25% in both studies). As a result, no further attempts at modeling this drug has been carried out since then.

Minto's PK parameter sets (see Table 3.3) have the advantage of accounting for the age and lean body mass of the patients. However, they can only be used to predict arterial blood concentration during

[3]NONMEM: NON-linear Mixed Effect Models

[4]The MAWR is a standard PK accuracy parameter. It is defined as $\text{MAWR} = \sum_{i=1\ldots N} \sum_{j=1\ldots M} \frac{|c_{ij} - cp_{ij}|}{cp_{ij}}$, where c_{ij} is the j^{th} measured plasma concentration of the i^{th} individual, and cp_{ij} denotes the corresponding predicted value. N is the total number of individuals and M is the number of samples. The MWR is a similar accuracy parameter without the absolute value.

PK PARAMETER	VALUE	UNITS
Cl_1	$\theta_1 \cdot (\text{BW}/70)^{\theta_7}$ if age ≤ 60	$[\text{l}\cdot\text{min}^{-1}]$
	$\theta_1 \cdot (\text{BW}/70)^{\theta_7} - (\text{age} - 60) \cdot \theta_{10}$ if age > 60	$[\text{l}\cdot\text{min}^{-1}]$
Cl_2	$\theta_3 \cdot (\text{BW}/70)^{\theta_8} \cdot (1 + \text{ven} \cdot \theta_{14}) \cdot (1 + \text{bol} \cdot \theta_{16})$	$[\text{l}\cdot\text{min}^{-1}]$
Cl_3	$\theta_5 \cdot (\text{BW}/70)^{\theta_{11}} \cdot (1 + \text{bol} \cdot \theta_{18})$	$[\text{l}\cdot\text{min}^{-1}]$
V_1	$\theta_2 \cdot (\text{BW}/70)^{\theta_{12}} \cdot (\text{age}/30)^{\theta_{13}} \cdot (1 + \text{bol} \cdot \theta_{15})$	[l]
V_2	$\theta_4 \cdot (\text{BW}/70)^{\theta_9} \cdot (1 + \text{bol} \cdot \theta_{17})$	[l]
V_3	θ_6	[l]

Table 3.1: Propofol PK parameter sets from [101] where BW stands for body weight, ven= 1 for venous sampling, ven= 0 for arterial sampling, bol= 1 for bolus administration, and bol= 0 for infusion administration.

PARAMETER ESTIMATES	VALUES	UNITS	SE
θ_1	1.44	[l/min]	0.09
θ_2	9.3	[l]	0.9
θ_3	2.25	[l/min]	0.31
θ_4	44.2	[l]	6.1
θ_5	0.92	[l/min]	0.15
θ_6	266	[l]	43
θ_7	0.75		0.06
θ_8	0.62		0.09
θ_9	0.61		0.11
θ_{10}	0.045		0.012
θ_{11}	0.55		0.13
θ_{12}	0.71		0.26
θ_{13}	-0.39		0.15
θ_{14}	-0.40		0.10
θ_{15}	1.61		0.36
θ_{16}	2.02		0.41
θ_{17}	0.73		0.23
θ_{18}	-0.48		0.12

Table 3.2: Parameter estimates from the NONMEM analysis ([101]).

PK PARAMETER	VALUE	UNITS
Cl_1	$2.2 - 0.0162 \cdot \text{age} + 0.0191 \cdot \text{LBM}$	$[\text{l}\cdot\text{min}^{-1}]$
Cl_2	$3.25 - 0.0301 \cdot \text{age}$	$[\text{l}\cdot\text{min}^{-1}]$
Cl_3	$0.121 - 0.00113 \cdot \text{age}$	$[\text{l}\cdot\text{min}^{-1}]$
V_1	$1.94 - 0.0201 \cdot \text{age} + 0.072 \cdot \text{LBM}$	[l]
V_2	$7.12 - 0.0811 \cdot \text{age} + 0.108 \cdot \text{LBM}$	[l]
V_3	5.42	[l]

Table 3.3: Pharmacokinetic parameters of remifentanil as a function of the age and lean body mass (from [103]).

remifentanil infusion. This is hardly a limitation since infusion is the method of administration of choice for this drug. Also, due to the drug susceptibility to blood esterases, venous blood concentration significantly lags the arterial concentration [104]. As such, PK models based on venous blood sampling should not be used as a basis to predict drug effect.

3.1.4 Concluding Remarks

3.1.4.1 Drug Uptake

In PK analysis and identification, it is usually assumed that the drug mixes instantaneously within the central compartment upon administration. However, this assumption is invalid when considering the kinetics of the drug during the uptake phase, *i.e.*, right after administration. Since the uptake phase is extremely short, this phenomenon is difficult to measure properly as it would require numerous blood samples taken in the first minute after administration. However, some authors [105] consider that a time delay of a few seconds should be accounted for in the model to describe the initial drug uptake phase.

3.1.4.2 Linearity

Drugs pharmacokinetics are not linear in the sense that the rate of injection of the drug affects the overall drug distribution. This aspect of pharmacokinetics is well known by clinicians and practitioners who do not titrate anesthetic drugs the same way whether they use boluses or infusions.

In that respect, most of the recent PK studies make a clear distinction between PK models obtained from bolus data and PK models obtained from infusions. Note for example in Fig. 3.4 the two-times difference between the propofol impulse response of the infusion *vs.* bolus models (from Schüttler [101]). As expected, this difference occurs mostly within the first few minutes following the administration.

One of the only study addressing this issue was the one published in 1998 by Schnider *et al.* [100] for propofol. The authors observed that propofol kinetics are linear for an infusion range of 25 to 200 $\mu\text{g}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$.

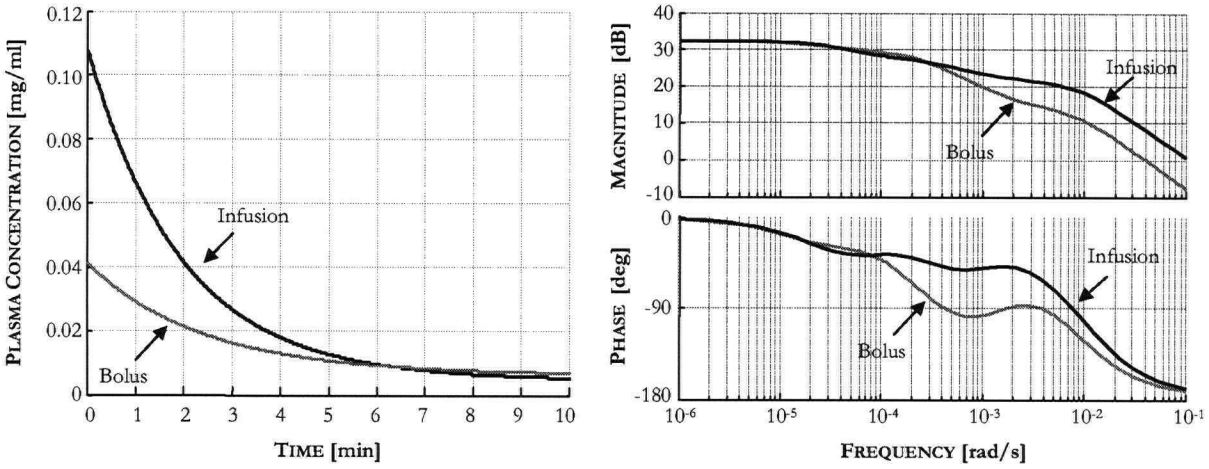


Figure 3.4: Bolus *vs.* infusion pharmacokinetic of propofol (using [101]). (a) Impulse response (b) Frequency response.

3.1.4.3 Time Variance and Other Factors Affecting PK Models

There are a number of physiological differences between individuals that can affect pharmacokinetic parameters. Understanding and quantifying these differences can reduce the model uncertainty.

For instance, most pharmacokinetic parameters are derived for a given age population. Pediatric, adult or elderly patients can have very different hepatic and renal extraction ratio, thus affecting the elimination of the drug. Also, since the distribution of a lipophilic drug is affected by the presence of adipose tissues, obesity in patients should be accounted for, at it increases the elimination half-life of the drug. Hence, weight, fat and lean body mass are often cited as covariates of pharmacokinetic parameters.

Finally, in recent years, it has been shown that cardiac output also affects drugs kinetics ([106], [107]). For instance, it has been shown that the initial arterial concentration of propofol, following its rapid infusion in sheep, is inversely related to the cardiac output. This issue hasn't yet been widely documented. The question whether or not the inclusion of the cardiac output in the pharmacokinetic parameters would reduce the uncertainty in the model remains opened. A preliminary study by Adachi *et al.* [108] tends to indicate that pharmacokinetic variability can be reduced when accounting for the cardiac output. However, since cardiac output can only be measured by invasive means, this issue hasn't yet received much attention.

While age, lean body mass, and eventually cardiac output are covariates of pharmacokinetic parameters in normal healthy subjects, factors such as renal dysfunction and hepatic diseases will also bear a tremendous impact on drugs kinetics, mostly during the elimination phase. We invite interested readers to refer to [98] for a more in-depth analysis on how these diseases can affect total drug clearance.

3.1.4.4 Pharmacokinetic Interactions

Effect of opioids administration on propofol A limited number of studies have documented the pharmacokinetic interactions between propofol and opioids. A first study by Briggs *et al.* in 1985 [109]

concluded that fentanyl significantly increases the level of the propofol blood concentration profile. To reach this conclusion, they compared the parameters of a 3-compartment model of propofol pharmacokinetics obtained from two different groups of patients. In the first group, patients received a single bolus of propofol. In the second group, a bolus of fentanyl was administered as a premedicant before the injection of propofol. They observed that the apparent volume of distribution, as well as the central compartment volume, were reduced by 45% in the group that received fentanyl. The total drug clearance was also significantly reduced.

A similar study by Gill *et al.* in 1990 [110] yielded a very different conclusion. Their experimental protocol also involved a patient population divided into two groups (one receiving only propofol, and the other receiving both propofol and fentanyl). The authors reported no significant differences in the pharmacokinetic profile of propofol between the two groups.

As of today, the effect of the administration of an opioid on propofol pharmacokinetics remain unclear and a subject of debate.

Effect of propofol administration on opioids Conversely, studies by Gepts *et al.* [111] and, more recently, Mertens *et al.* [112] in 2001 reported that the administration of propofol significantly reduces the elimination and inter-compartmental clearances of alfentanil. The authors also mentioned that a reason for this interaction could be linked to the fact that propofol tends to reduce the mean arterial blood pressure, thus affecting opioid distribution and elimination. Further studies are necessary to evaluate the precise mechanisms that can cause these interactions [24].

3.2 Pharmacodynamics

The role of pharmacodynamic modeling is to mathematically express the observed effect of a drug as a function of its plasma concentration:

$$PD(s) = \frac{E(s)}{C_p(s)} \tag{3.11}$$

where $PD(s)$ is the pharmacodynamic model and $E(s)$ is the drug effect.

Note that for a given drug there can be a multiplicity of effects. Hence, pharmacodynamic models are not unique as they describe only one over the many different possible endpoints of a drug.

3.2.1 The Dose/Response Relationship: a Steady-State Model

The characteristic of any given drug can be expressed by *dose-response curves*, see Fig. 3.5. As expected, at low doses, drugs have only a minimal effect. As the dose increases, the effect increases as well. Often, dose-response curves exhibit a linear characteristic (constant slope) for a given range of doses. At higher doses, a saturation phenomenon appears, indicating that the full effect of the drug has been reached. The

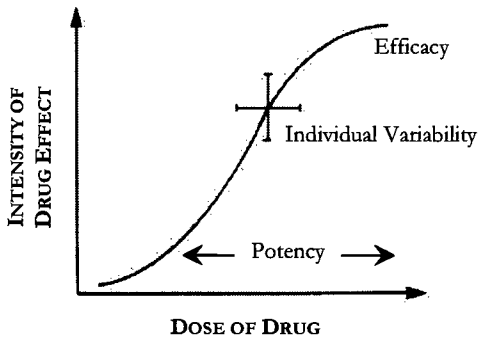


Figure 3.5: Dose-response relationships differ according to the drug potency, efficacy, slope and subjects' variability.

sigmoid characteristic of the dose/response is usually captured using the Hill equation:

$$\overline{E} = E_0 + E_{max} \cdot \frac{\overline{C}_p^{\gamma}}{EC_{50}^{\gamma} + \overline{C}_p^{\gamma}} \tag{3.12}$$

where $\overline{E}$ is the steady state effect, $\overline{C}_p$ is the steady state plasma concentration, and EC_{50} is the concentration which yields 50% of the maximal effect, or which yields the maximal effect in 50% of a given population of patients. The parameter γ is a measure of the steepness of the dose-response curve. Finally, E_0 and E_{max} are the minimum and maximum effects.

When considering anesthetics and opioids, it can be difficult to quantify drug effects. Very often, only quantal (*e.g.*, response *vs.* no response) observations can be made. For example, in the case of opioids, the absence of reaction (*i.e.*, movement) to intubation, skin incision, and skin closure is often taken as a relevant endpoint (see Fig. 3.6). The results obtained from different patients are then pooled in order to obtain the sigmoid-like shape which represents the probability of response versus the administered dose or the drug plasma concentration. In this case, the dose/response curve can be used to determine the dose or plasma concentration which would yield the desired effect (*e.g.*, no response to intubation or skin incision) in a given percentage of a population. This type of dose/response curve are the most commonly found in the literature preceding the mid-1990s.

Conversely, it is sometime possible to target an effect which can be quantifiable. The resulting dose/response curve would then indicate the dose or plasma concentration which are necessary to obtain a partial effect (*e.g.*, the plasma concentration of propofol to reach a certain BIS value). Most of the recent research done in pharmacodynamic modeling seems to follow this new trend.

When a number of effects or observational endpoints can be selected to model a drug, it is important to carefully select the effect or endpoint which yields a useable characteristic with respect to the therapeutic window of the drug. For instance, effects leading to dose/response characteristics such as in Fig. 3.7.a-b are not suitable since they are either on/off or they fail to characterize the whole therapeutic range of the drug. An ideal endpoint for modeling the dose/response characteristic would be an effect leading to the linear characteristic of Fig. 3.7.c. This characteristic simply implies that a variation of the effect can be

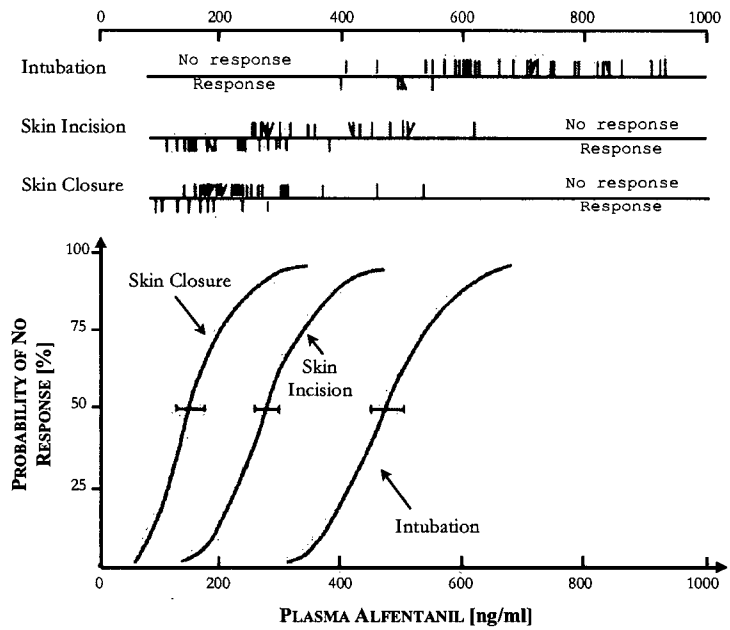


Figure 3.6: Drug dose-response curves can be obtained by observing whether a patient responds to a particular stimulus at a given concentration (from [113]).

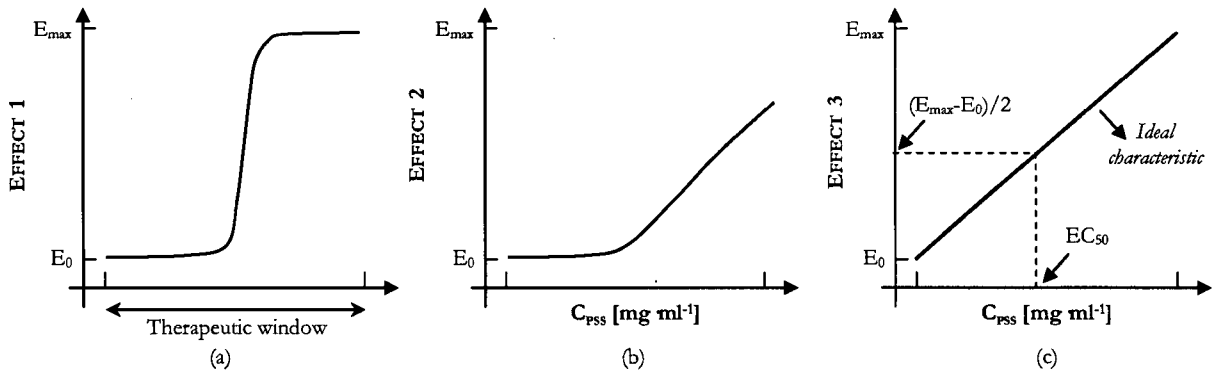


Figure 3.7: Different dose/response characteristics (a) On/Off effect (b) The observed endpoint cannot characterize the effect of small doses (c) Ideal characteristic.

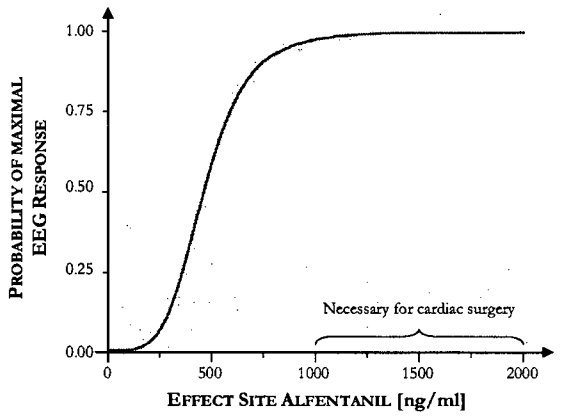


Figure 3.8: Dose/response model for Alfentanil and using the probability of maximal EEG response (from [31]).

achieved by applying a proportional variation in the drug dosing regimen, independently of the operating point of the system. To illustrate this discussion, let us consider the use of the EEG in measuring the effect of opioid agents, see Fig. 3.8. In this example, the smallest doses used in cardiac surgeries already induce a maximal EEG response (pronounced delta waves). Hence, this particular endpoint is useless in cardiac surgeries. However, it might be appropriate in non-cardiac cases where the opioid doses are generally much lower.

3.2.2 The k_{e0} Parameter: a Measure of the Effect Dynamics

Drug-response relationships are obtained based on steady state observations. As such, they do not capture the dynamic behavior that exists between the drug plasma concentration and the observed effect[5].

The effect of rapid intravenous administration is well known and well documented. When plotting the observed effect versus the plasma concentration, a hysteresis-like cycle can be observed (see for instance Fig. 3.9) indicating a disequilibrium between the plasma concentration and the effect site concentration of the drug. A direct consequence of this disequilibrium is that the effect lags the plasma concentration. An explanation for this disequilibrium is that the drug concentration within the biophase[6] lags the drug concentration in the blood. Hence, predicting the effect site concentration C_e is relevant in order to devise a proper infusion profile. In 1979, Sheiner *et al.* [114] proposed an addition to the conventional pharmacokinetic models in the form of an effect compartment connected to the central compartment. This hypothetical effect compartment models the temporal aspect of pharmacodynamics where a rate constant k_{e0} expresses the dynamics of transfer of the plasma concentration to the biophase. Therefore, it is possible to mathematically express the effect site drug concentration $C_e(s)$ as a function of the plasma concentration $C_p(s)$ as:

$$C_e(s) = \frac{k_{e0}}{s + k_{e0}} \cdot C_p(s) \tag{3.13}$$

[5]Often in the clinical literature, pharmacodynamics models are limited to only the static Hill dose *vs.* response relationship.
[6]The biophase is defined as the physiological milieu targeted by the drug. It usually cannot be accessed for sampling.

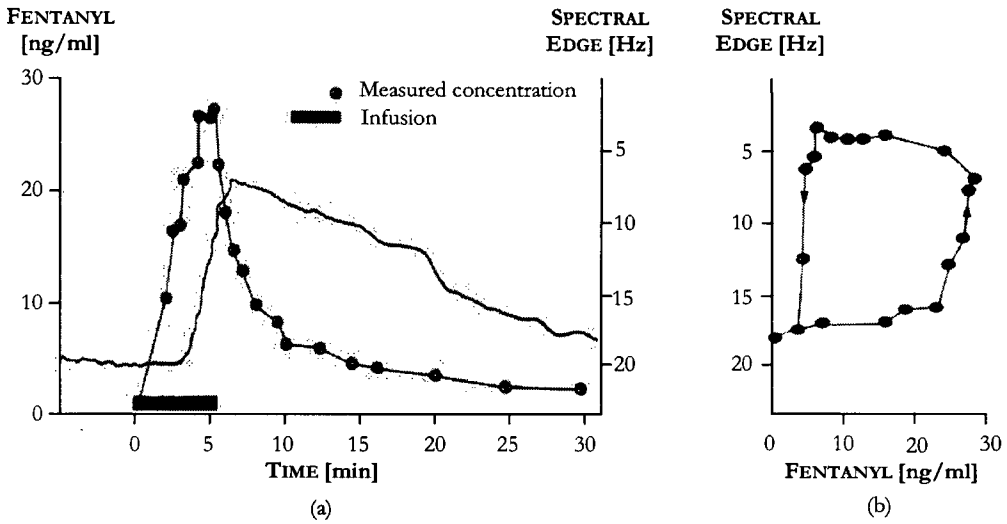


Figure 3.9: Effect and plasma concentration of fentanyl and alfentanil following rapid intravenous administration (adapted from Stanski [89]).

This equation only captures the effect dynamic. It has to be used in conjunction with the non-linear gain of the steady-state model (3.12):

$$E(s) = E_0 + E_{max} \cdot \frac{C_e(s)^\gamma}{EC_{50}^\gamma + C_e(s)^\gamma} \tag{3.14}$$

This modeling approach to pharmacodynamics has become the mainstream approach followed by pharmacologists since the early 1980s. It will be clear by anyone well-versed in System Engineering that this approach suffers from a number of limitations. In particular, the model (3.13) was never rigorously validated and tested using proper identification data. For instance, the absence of a time delay in (3.13) may hide the inherent latency of the drug effect. In addition, the dynamics of the sensor used to measure the effect is not included in the model. Finally, the use of the Hill saturation characteristics as the primary modeling tool cannot be recommended, as it can falsely characterize linear dynamics as non-linear. These limitations are further addressed in Chapter 6, where a new modeling approach based on System Identification know-how is proposed.

3.2.3 Literature Review

3.2.3.1 Propofol

Only 6 studies involving quantitative EEG have been published since the mid-1990s. These studies involve a limited number of volunteers and/or patients.

In terms of the equilibration time constant k_{e0}, results vary from 0.1 to >0.6 min^{-1}. This large variation can be attributed to differences in the type of identification methodology used in the studies. Also, the fact

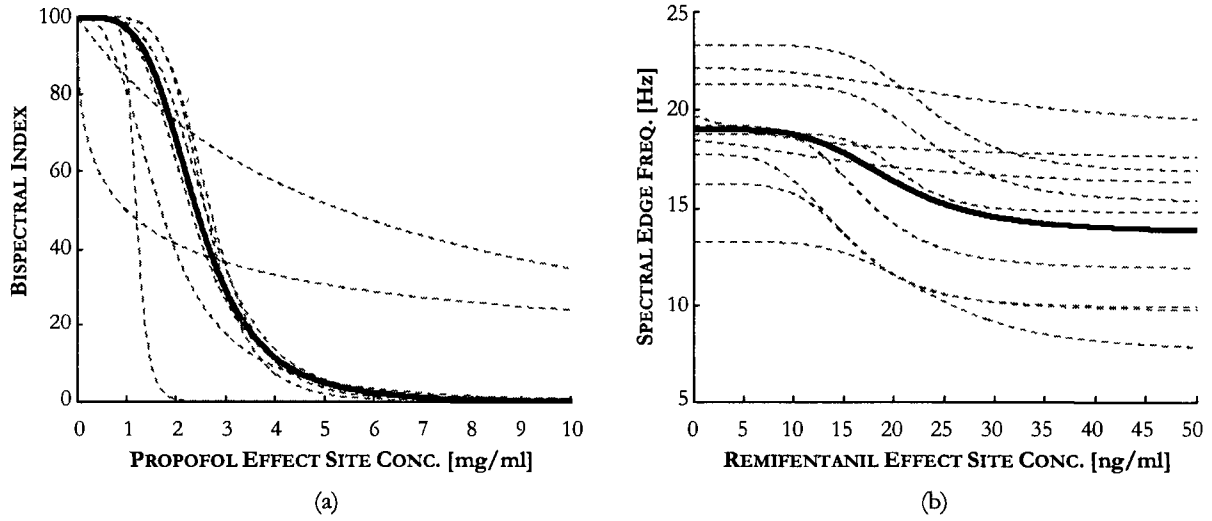


Figure 3.10: Individual (dashed lines) and mean (solid lines) Hill non-linear characteristics of published pharmacodynamic models. (a) Propofol study from Kuizenga *et al.* [118] (b) Remifentanil study from Egan *et al.* [102].

that each study was using a different PK set to calculate the blood plasma concentration, or used their own blood measurements, probably resulted in additional variability.

The EC_{50} parameter is also significantly different between studies. Using auditory evoked potentials, White *et al.* [115] found that a concentration of 2.17 mg/ml was sufficient to provoke unconsciousness. A similar result (about 2 mg/ml) was reported by Schnider at al. [100] for their semilinear canonical correlation index. Much higher values (about 6 mg/ml) were reported by Kazama *et al.* [116] for the bispectral index[7]. Conversely, Billard *et al.* [117] and Kuizenga *et al.* [118] reported intermediate values (3 to 4 mg/ml) for BIS, which are closer to the usual recommended concentration suitable for surgery.

Probably the most remarkable limitation of the PD models presented in some of these studies is the rather large inter-individual variability which can be observed from the Hill non-linear characteristic. For instance, Kuizenga *et al.* published the Hill parameters for each one of their subjects, see Fig 3.10.a. According to the Hill relationship, some of their subjects experienced a complete depression of their cortical activity at concentrations that barely had any effects in other subjects. The averaged Hill relationship might thus be useless in the face of so much variability.

The only study which reported a reasonable PD variability is the one by Kazama *et al.* [116]. One of the reason for the consistency of their results was that they compensated for the dynamics of the BIS sensor and removed the initial time delay between the drug administration and the start of the BIS descent. While this technique had the advantage of leading towards a more consistent PD model, the fact that part of the system dynamics was removed from the identification data limits the usability of the proposed models. However, one very interesting conclusion of the study was that age is a PD covariate. However, conversely to many other drugs, they found that the EC_{50} of propofol tends to increase with age, indicating that older

[7]Accordingly, a concentration of 6 mg/ml is necessary to reach a BIS of 50.

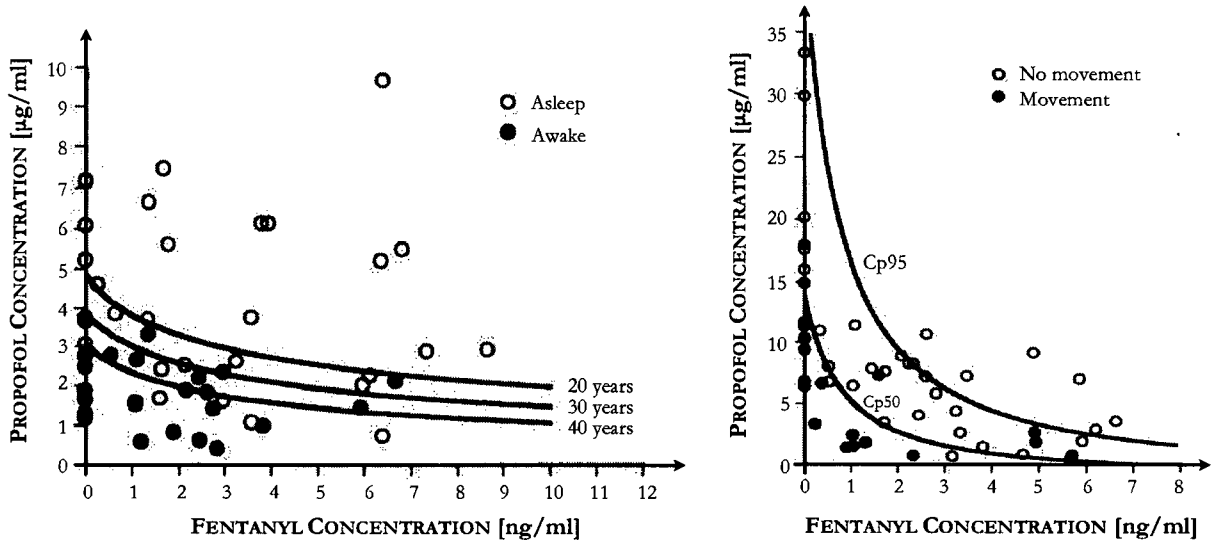


Figure 3.11: Drug interaction between propofol and fentanyl (adapted from [120]).

patients are *less* sensitive to propofol than younger patients. This result, while surprising, was corroborated by another study on rats [119]. They also found that the average EC_{50} in the adult population was about 6 mg/ml, which is much higher than most of the other studies.

Remifentanil The effect of remifentanil on the spectral edge frequency has been studied by both Egan *et al.* [102] and Minto *et al.* [103] in the late 1990s. Both studies also reported a considerable variability in the Hill parameters (see for instance Fig. 3.10.b). This variability could be due in part to the fact that the PD modeling was carried out using spectral edge frequency which is less sensitive to opioids than to anesthetics. In some cases, only a marginal decrease in the spectral edge could be observed.

3.2.4 Concluding Remarks

3.2.4.1 Pharmacodynamic Interactions

Propofol and opioids are known to interact with each other in a synergistic fashion in their anesthetic/analgesic effect, but not in their toxicity. This observation constitutes the basic assumption on which the concept of balanced anesthesia is based.

In an interesting set of studies, Smith *et al.* [120] have measured the reduction of propofol and fentanyl effective dose when both drugs were used concomitantly. Their results are displayed in Fig. 3.11. Quite clearly, there is an optimal titration that ensures adequate hypnosis and analgesia. For instance, the effect of the combined use of propofol and, in this case, fentanyl shows that the effective concentration EC_{50} needed to suppress movement in response to a noxious stimulus in 50% of patients can be drastically reduced. Where a concentration of 12 $\mu g \cdot ml^{-1}$ of propofol (or 6 $ng \cdot ml^{-1}$ of fentanyl) was necessary, using both drugs simultaneously reduces this requirement to 4 $\mu g \cdot ml^{-1}$ of propofol combined with 1.5 $ng \cdot ml^{-1}$ of fentanyl.

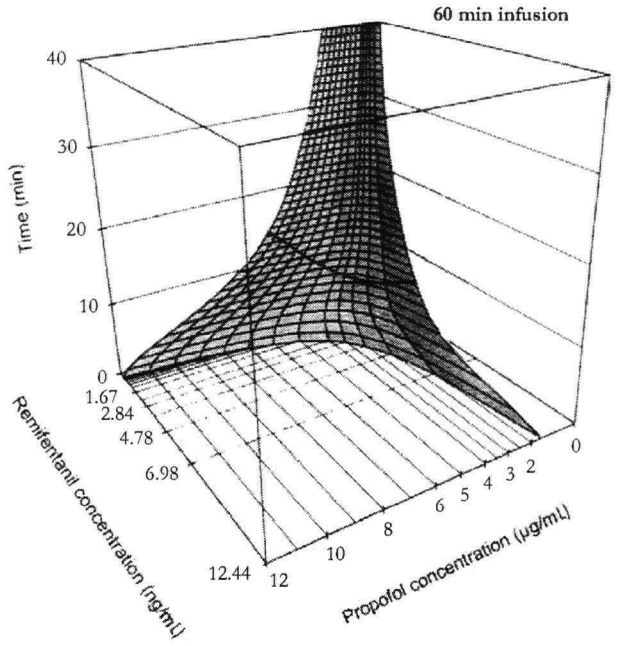


Figure 3.12: Simulation of propofol and remifentanil pharmacodynamic interactions (adapted from [23]).

Synergism is also apparent in hypnosis, but to a lesser extent.

Pharmacodynamic interactions between propofol and remifentanil have also been studied [23]. Surface plots were generated to estimate the optimal titration that minimizes the recovery time upon termination of the infusion, see Fig. 3.12. The curve in the lower plane represents the EC_{50} of both propofol and remifentanil. At the extremes, a concentration of either 12 μg·ml^{-1} of propofol or 12.44 ng·ml^{-1} of remifentanil is required to provide adequate anesthesia (in this case: no movement to painful stimuli). In this simulation the infusion rate of both drugs is maintained constant. After 1 hour, the infusion is discontinued, resulting in a decrease in both drugs plasma concentrations. The thick line is the isobole that predicts emergence in 50% of the patients. A minima at $EC_{50}^{p} = 2.5$ μg·ml^{-1} (propofol) and $EC_{50}^{r} = 5$ ng·ml^{-1} (remifentanil) exits, indicating that this combination of drugs would optimize the emergence time, while providing adequate titration for the maintenance of anesthesia.

Of interest, a 2000 study by Minto *et al.* [25] has extended this concept to 3 drugs (propofol, midazolam and alfentanil).

While surface plots clearly stress out the synergism between drugs, their interest is limited to steady state. No *dynamic* model of pharmacokinetic/pharmacodynamic drug interactions have been published so far. This represents a very fertile area for medical research.

3.2.4.2 Time Variance

Propofol In a recent study from Kuizenga [118], the time invariance of propofol pharmacodynamic models has been seriously questioned. Three propofol infusions were given to 10 patients. Each successive infusion

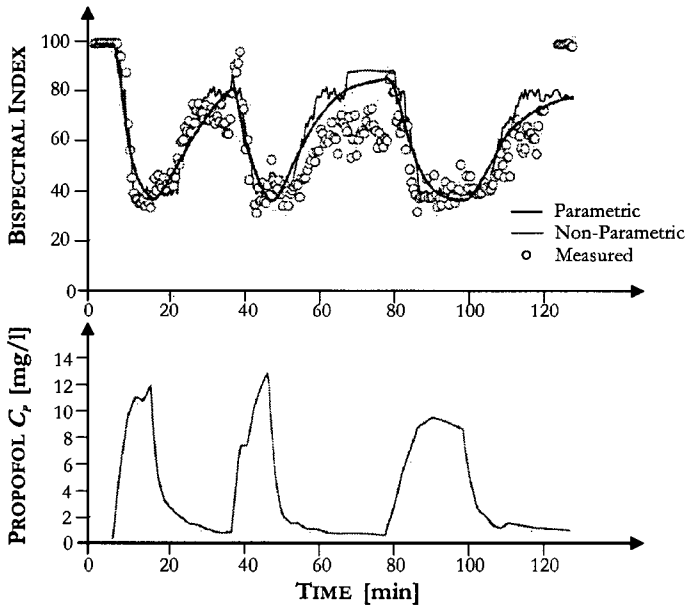


Figure 3.13: Measured and predicted time courses of the BIS following a series of induction sequence administered to a volunteer (from [118]). A pharmacodynamic model obtained based on the first repeat was used to predict the BIS time course. Note the good agreement between measurement and prediction for the first repeat. Results worsen considerably for the two other repeats.

was started immediately after the patient regained consciousness. A set of pharmacodynamic parameters were obtained from the first infusion data. This set was used to predict the effect of the subsequent infusions based on the plasma concentration. The results have shown an important discrepancy between the predicted effect and the observed effect, see Fig. 3.13. The authors concluded that pharmacodynamic models cannot be used accurately over time to predict the effect of propofol. In their study, they analyzed the effect of propofol on both the bispectral index and the amplitude of the patients' electroencephalogram in the 11 to 15 Hz band. A second set of pharmacodynamic parameters were derived based on the observation on the two successive infusions. Both sets are presented in the Table B.2. Note the large discrepancy between the first and the second set.

Remifentanil The long term administration of opioids usually results in the development of tolerance to analgesia followed by a strong physical dependence to the drug. In the case of remifentanil, a study by Vinik and Kissin [121] involving 13 volunteers have shown that tolerance to analgesia is profound and develops very rapidly. A constant-rate infusion of 0.1 $\mu g \cdot kg^{-1} \cdot min^{-1}$ was administered over a period of 4 hours. The analgesic effect was evaluated by measuring pain tolerance to thermal (cold water) and mechanical (pressure) noxious stimuli. Tolerance to pain was measured as the length of time that the subject was able to tolerate the stimulation. A baseline value was recorded before the start of the infusion. Painful stimulation was applied every 30 minutes. The study has shown that tolerance to pain is maximum after a 60 to 90 min infusion. Once the peak is reached, analgesia begins to decline steadily, see Fig. 3.14. After

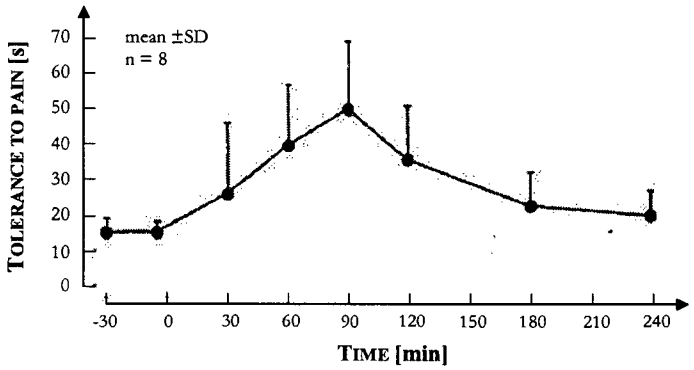


Figure 3.14: Development of acute tolerance to analgesia during short-term constant-rate infusion of remifentanil (adapted from [121]).

4 hours, the effect of remifentanil was no longer observable. The authors of the study concluded that calculations for target-controlled infusion of remifentanil must include corrections for tolerance. However, this would require models which have not yet been developed.

3.3 Summary

Since its inception in the early 1950s by Dost, the study of pharmacokinetics has received considerable attention from the research community. Recent advances in non-linear regression analysis have allowed pharmacologists to define PK parameter sets which account for the effect of covariates such as age, weight, gender and ethnicity. It is now possible to predict with a fair accuracy the blood plasma concentration of most of the anesthetic drugs in use today. In terms of control engineering, there are three important aspects in pharmacokinetics, which require some attention:

- PK parameters are dependent on the infusion rate of the drug. For instance, slow infusions (less than 200 $\mu g \cdot kg^{-1} \cdot min^{-1}$) are linear with respect to the input amplitude, while bolus-type infusions (above 2500 $\mu g \cdot kg^{-1} \cdot min^{-1}$) have a different initial drug distribution, which results in a different PK parameters set.

- A second aspect is the interaction between drugs, and in particular propofol and remifentanil. While the literature is very confused about this issue, most researchers tend to agree that, if any interactions occur at the level of pharmacokinetics, these interaction mostly affect the re-distribution and elimination of the drugs.

- Finally, the blood plasma concentration of intravenous drugs cannot be observed nor controlled directly since there is no available real-time measure. The pharmacokinetics of intravenous drugs can thus only be used to *predict* the plasma concentration. Accounting for patient variability appears then necessary.

Patients with hepatic or renal diseases may have a widely different PK characteristic. The PK models presented in this section should not be used in these instances.

Since Schüttler and Ihmsen in 2000, no further studies were conducted to derive the PK model of propofol. While each one of the reported studies (see Section 3.1.3) have contributed in improving the goodness of the fit (mostly by adding the age, weight and lean body mass as covariates), it was argued that no further improvements can be made. The error between the measured plasma concentration and the predicted plasma concentration is now about the same when considering population-normed PK models or individualized PK models (*i.e.*, PK models derived specifically for one single individual). This indicates that further improvements can only be achieved by considering more complex physiological models instead of a mamillary compartmental model.

While dose-response curves have been used intensively in the past to derive dosing guidelines (one of the brightest example is the MAC concept for gaseous anesthetics), it is only recently that PD models based on quantifiable endpoints such as EEG-based parameters have been researched and published in the literature. Unfortunately, these PD models do not seem to be appropriate for use in control-oriented frameworks. For instance, and conversely to PK models now used extensively in Target Controlled Infusion (TCI) pumps in Europe, PD models have not yet been used successfully in any clinical application. It is likely that the large variability observed between patients and the discrepancies between the model parameters published in the literature explain the lack of enthusiasm in the anesthesia community for PD models. However, we believe that the poor performance of these models stems mostly from a poor choice of the model structure rather than an inherent limitation of the system. In Chapter 6, we revisit in details the problem of PD modeling and identification. In particular, we propose a new system-oriented approach that yields better prediction performance and a more accurate description of the frequency response of the system dynamics.

Chapter 4

Closing the Loop in Anesthesia: a Review

In terms of control engineering, the idea of the anesthesiologists striving to keep a balance between the pharmacological toxicity of anesthetic drugs and the effect of surgical noxious stimuli is particularly appealing. The anesthetic and opioid titration needs to be constantly adjusted in order to avoid both under- and overdosing. The idea of an automated system that would regulate drug titration in order to maintain the adequacy of the anesthetic regimen is immediate.

The concept of a closed-loop anesthesia system is thus very close to that of automated flight control. The role of the anesthesiologist is similar to that of a pilot: after take-off (induction), the pilot usually maintains an adequate flight trajectory (hypnosis, analgesia, paralysis). Nowadays such tasks are performed by flight controllers able to plan ahead, optimize fuel consumption and minimize the duration of the flight. Closed-loop anesthesia is somehow similar in the sense that, by changing the titration of intravenous drugs, the anesthesiologist can drive the patient into a deeper or lighter hypnotic and/or analgesic state, according to the requirements of the surgical procedure.

Using proper feedback quantities and drug models the possibility to automate the drug titration and to allow the practitioner to concentrate on higher level tasks seems viable. In keeping the comparison, closed-loop anesthesia would not replace the anesthesiologist. On the contrary, the workload of the anesthesiologist will be reduced during the maintenance stage, leaving room for full attention in case of an emergency, or when the override of the controller is required.

4.1 Review

The concept and implementation of closed-loop anesthesia has been investigated for the past half century via numerous attempts at controlling anesthetic drugs titration through feedback control. A selected chronological survey of this prior art is summarized in Table 4.1 and 4.2. For each mentioned attempt, we listed the feedback quantity, the drug used, and the control technique employed. In spite of a good number

of such attempts, no clinically satisfying results have been obtained so far. This survey is also limited to attempts which focused more precisely on the control of hypnosis. Other work focusing on the control of heart rate and blood pressure, as well as the control of muscle relaxation in the context of anesthesia has not been reported here.

In recent years, researchers have been using either a simple PID or a lookup table of the drug pharmacodynamic model to set the target plasma concentration of a target controlled infusion devices in order to reach and maintain a given hypnotic reference. The successes reported by Struys et al. [122] and Absalom et al. [123] can be easily attributed to the fact that their system titrated the drug according to the index of consciousness provided by the bispectral monitor, rather than to the performance of the closed-loop controller itself.

The results produced by controllers embedding advanced techniques, as shown in the work by Frei et al. [124] and Gentilini et al. [125], emphasize that the problem is far from being solved due to the aforementioned challenges posed by the intra- and inter-patient variability. There have also been attempts at closing the loop by Linkens et al. ([126], [127], [128], [129]) using a variety of intelligent control techniques such as expert systems and fuzzy logic. Linkens et al. were probably among the first to attempt the control of distinct anesthesia components simultaneously (analgesia and areflexia) using different agents (atracurium and isoflurane). An in-depth analysis of such cases reveals the need for strong knowledge of the patient model. The intra- and inter-patient variability makes the establishment of *a priori* rules very difficult. Stability of fuzzy-based controllers may be impossible to prove.

From the perspective of interaction between drugs, and of particular interest, is the attempt by Zhang et al. in 1998 [130] at controlling an intravenous anesthetic (propofol) together with an opioid (fentanyl). This approach was limited to the control of the plasma concentration of propofol and fentanyl in dogs, where the setpoints were chosen to minimize the wake up time.

Probably the most interesting results obtained in recent years were those of Locher *et al.* and Liu *et al.*, see Table 4.2. Not only have they shown that closed-loop control is particularly effective in maintaining a proper BIS setpoint during anesthesia, they have also proven that closed-loop control brings further clinical advantages in terms of drug consumption and wake up time, as compared to manual titration [142]. Closed-loop controllers also react faster to the abrupt changes brought by surgical stimulation [141].

4.2 Summary

We have seen in Chapter 2 that the use of the new monitors of hypnosis, such as the BIS monitor, results in faster wake up, fewer side effects, and a reduction in drug usage, while ensuring that patients remain unaware during their surgery. For many years, however, the advantage of closed-loop system based on these monitors remained unclear. In particular, while closed-loop systems inherit *de facto* the advantages brought by the feedback sensors, doubt still persists whether there is any further clinical advantage to having a fully closed-loop system as compared to manually adjusting the titration based on the sensor information [143].

Study	Feedback quantity(ies)	Controlled Agent(s)	Control Technique	Population	Comments and Limitations
Bickford *et al.*, 1950–1960 [38] [39] [40]	EEG energy in the 4 to 12 Hz band	Pentothal or Ether	On/Off control	Rabbits, cats, monkeys. Clinical trial on 50 patients undergoing various abdominal surgeries	Oscillations due to the control technique. Method sensitive to extraneous interferences. No opioids have been administered concurrently, thus seriously limiting this technique in today's practice.
Bellville *et al.*, 1955–1960 [41]	EEG amplitude in a defined frequency band	Cyclopropane	Analog control (P or PI?)	Not Disclosed	No results have been presented. The proposed technique is merely an improvement of Bickford's servo anesthesia. The authors mention speed (infusion pump) and position (vaporizer) control, however no specific information are disclosed. Method limited to the control of a unique anesthetic agent.
Schwilden *et al.*, 1985–1995 [44] [45] [131] [132]	Median Edge Frequency	Methohexital Propofol Alfentanil Isoflurane	Model-Based Adaptive Control Adaptation was done if the system output was diverging too far from its reference	13 volunteers (22-29 yr ; 44-85 kg) 11 volunteers (24-31 yr ; 54-87 kg) 11 patients 25 female patients (31-47 yr), ASA I or II	Results on volunteers have shown that a constant excitation is necessary the guarantee the reliability of the feedback quantity (otherwise the volunteers were drifting from a drug-induced unconsciousness into a natural sleep). This technique works also for opioids. The controlled drug was used as the only anesthetic agent during the maintenance phase.
Kenny *et al.*, 1990–1995 [133] [134] [135]	Auditory Evoked Potentials	Propofol	Outer PI controller setting the reference to an inner TCI device.	27 patients	The authors recommended feedback control of anesthesia as a research tool to better identify pharmacodynamic models and the interaction between drugs.
Roy *et al.*, 1995–2000 [136] [64]	Auditory Evoked Potentials	Halothane Propofol	Fuzzy rule-based control system controlling either the vaporizer or giving a reference to a TCI device	10 sessions conducted on 6 mongrel dogs with tail clamping stimulation 9 sessions conducted on 5 mongrel dogs	These papers emphasize mostly the hypnosis index derived from midlatency auditory potentials using wavelet analysis. Due to the extensive averaging needed, a value quantifying the level of hypnosis was calculated every 3 minutes.
Gentilini *et al.*, 1995–2000 [124] [125] [18] [137] [138]	Mean Arterial Pressure Bispectral Index	Isoflurane	Model Predictive Control Cascade Internal Model Control. An inner loop controls the end-tidal concentration. The reference of the inner loop is set by an outer loop that regulates BIS variations	20 patients 40 patients (20-65 yr)	The control of inhalational gas such as isoflurane has the advantage that the drug plasma concentration is closely related to the end-tidal expired gas concentration which is readily available. This is a clear advantage over intravenous agents for which measurement of drug concentration is impractical. The authors have shown that isoflurane can both control the mean arterial blood pressure (MAP - sensitive to noxious stimuli) and the BIS (sensitive to the hypnotic level). The authors proposed to investigate a SIMO controller where isoflurane would be used to control both the MAP and the BIS

Table 4.1: Prior Art: a Literature Survey (Part I)

Study	Feedback quantity(ies)	Controlled Agent(s)	Control Technique	Population	Comments and Limitations
Struys *et al.*, 2001 [122]	Bispectral Index	Propofol	A lookup table (Hill model) acquired during induction serves to calculate the required effect site concentration changes. A TCI device tracks these changes.	10 female patients (12-60 yr)	A continuous infusion of remifentanil was started 2 min before induction, thus the necessity to acquire a Hill curve that models propofol effect on the BIS with remifentanil acting as a bias. The results clearly indicate that the closed-loop control of propofol significantly reduces recovery time as compared to the standard anesthesia practice. However, this benefit could only be the result of titrating the drug according to the BIS.
Absalom *et al.*, 2002-2003 [123] [139] [140]	Bispectral Index	Propofol	Outer PID controller setting the reference to a TCI device. There is an additional constraint limiting the maximum change of the infusion rate	10 patients (67 yr $\pm$11 ; 79 kg $\pm$11)	The authors present a comprehensive control algorithm based on a PID control structure. The way they calculated the gains and the time constant of the controller is not clear. 3 patients out of 10 presented severe oscillations where the BIS value was clearly leading the target plasma concentration, which is a clear sign of the instability of the outer PID controller. The authors also showed in a following publication that the same system can be effectively applied to sedation.
Locher *et al.*, 2004 [141]	Bispectral Index	Isoflurane	Modified version of Gentilini's cascade controller. The outer controller is a standard PID controller. The inner controller uses the end-tidal expired gaz concentration as feedback.	10 patients (+13 control) (29-59 yrs old)	This cascade control system makes use of a feedback measure of the end-tidal concentration in its inner loop. The authors report significant improvements as compared to the manual control of BIS. They also exercised caution when designing the outer PID controller to ensure robustness. However, no mathematical proof of stability was given. The authors reported that the controller reacted promptly during surgical disturbances. They also commented on the fact that the wake up time in the manually controlled group presented much more disparity as compared to the controlled group.
Liu *et al.*, 2005 [142]	Bispectral Index	Propofol	Outer PID controller setting the reference to a TCI device. No further information were given as to the controller parameters. It is likely that the TCI pump was age and weight adjusted.	83 patients (closed-loop) ; 81 patients (manual TCI)	This study is the largest study involving closed-loop control of propofol using BIS. The authors successfully demonstrated that the closed-loop system was able to maintain the patients within $\pm$10 of the BIS setpoint for 90% of the time, vs. 70% of the time during manual control of the TCI pump. In the closed-loop group, patients received in average 23% less drug and woke up in 30% less time. Note that the controller was also used to induce patients.

Table 4.2: Prior Art: a Literature Survey (Part II)

In other words, is the rigorous maintenance of a BIS setpoint brought by a closed-loop system worth the risk of having a machine deciding on the titration profile?

Only recent evidence in the literature suggest that there are significant clinical advantages to be gained when closing the loop. Liu *et al.*, in particular, reported a 30% decrease in wake up time, and a 23% decrease in drug usage. Locher *et al.* have shown that the closed-loop system responds faster to changes in the patient's state and results in a more predictable recovery time.

Closed-loop control in anesthesia remains essentially the domain of clinical researchers from European countries. Only the work of Gentilini *et al.* has been driven by a control engineering perspective. However, their work is mainly based on existing anesthesia sensors and drug effect models.

To ensure the success of closed-loop control for anesthesia, it appears essential to re-investigate the anesthesia system from a control engineering oriented point of view.

Before designing and testing controllers, it is first important to assess the performance and adequacy of today's anesthesia monitors for use as feedback sensors. For instance, it is interesting to note that none of the cortical monitors currently on the market has been properly identified (*i.e.*, there is no published model of their dynamic response to changes in the patient's cortical state). This stems from the fact that their behavior is highly non-linear[1]. The lack of a sensor model is a limiting issue from a control engineering perspective.

Also, the review of the propofol pharmacodynamic model presented in the Annex B reveals that there is a large discrepancy in terms of model parameters between the different published models. While this is, in part, due to the large variability that exists between patients, this is also due to a poor model identification methodology. This issue must be revisited.

Finally, the great majority of closed-loop attempts did not account for the uncertainty that stems from patient variability, even though some others do mention that their controllers were sufficiently derated to ensure robustness. As a result, some of these attempts resulted in instability characterized by an oscillatory behavior [123]. Analyzing patient variability in order to quantify the uncertainty that exists in anesthesia drug delivery systems is therefore a mandatory aspect of the control design.

The goal of this thesis is therefore to assess the feasibility of obtaining a control design that can maintain the adequacy of anesthesia in a large population of patients, while meeting the minimum requirements set forth by anesthesiologists in terms of closed-loop performance during setpoint changes and disturbance rejection. One additional condition to a successful design is that the stability of the controller with regard to patient variability to drug effect must be *mathematically* proven.

[1]This issue has recently been brought up by Dr. G. Schneider during the 2005 Advanced Modeling and Control in Anesthesia conference (AMCA 2005), where he has shown that the inherent time delay of most cortical monitors depends on the direction of change of the patient's state (*e.g.*, when patients loose consciousness, the BIS Monitor delay is in average 15 seconds, while the delay is about 30 seconds when patients regain consciousness.

PART B: Sensing, Modeling and Control

In the following 4 chapters, we review in details every aspect of the anesthesia system from a control engineering point of view. We bring new insights in terms of anesthesia sensors and models, and propose a control system design that is robustly stable and achieves the required clinical performance.

In Chapter 5, the inadequacy of existing monitoring technologies for use in close loop systems has prompted us to develop our own solutions, both from the point of view of Hypnosis and Analgesia. Significant time and efforts were invested into developing new tools and validating them in clinical studies. We give a detailed presentation of the derivation of the WAV_{CNS} (hypnosis). We also show how the same technique can be applied to the analysis of the Heart Rate Variability (HRV) signal for the quantification of analgesia.

In Chapter 6, we propose a new approach to pharmacodynamic modeling. Conversely to traditional PD modeling, this new approach accounts directly for the sensor dynamics. We also thoroughly validate the model structure using residual analysis. Forty-four (44) propofol PD models were derived based on induction data obtained during the LMA study (see Annex D). This makes of this study the second largest published propofol PD study.

The 44 PKPD patient models established in the previous chapter are then analyzed in order to quantify the expected system uncertainty originating from inter- and intra-patient variability. This analysis is carried out in Chapter 7. We found that uncertainty is cause for concern. We also investigate different method to reduce the uncertainty to a more manageable level. This analysis results in a number of nominal PKPD models and uncertainty weights that depend on the operating mode of the system, and the age of the patient. These models are fully disclosed in the Annex C.

In Chapter 8, we design a number of controllers using either a classical PID loop shaping design, or the more sophisticated H_∞ weighted mixed-sensitivity design procedure. We show that controllers achieving relevant clinical performance can be derived. We further show that the stability of these controllers is guaranteed, even under the very conservative assumptions made in Chapter 7. We conclude that patient variability can be effectively dealt with when proper modeling approach and uncertainty analysis are implemented.

Chapter 5

Quantifying Cortical and Autonomic Activity Using Wavelets

Closed-loop control performances rely directly on the availability and reliability of process output measurements. The selection of appropriate feedback sensors to perform these measurements is a critical aspect of any close loop design.

In terms of anesthesia, this issue is particularly challenging. As mentioned in Chapter 4, the control of anesthesia can only be achieved through the control of both its hypnotic and analgesic components. Hence, developing appropriate feedback sensing strategies for *both* endpoints is a priority.

Hypnosis and Cortical Activity In terms of hypnosis, most of the efforts pursued in recent years were based on the fact that hypnotic drugs exert their effect at the level of the Central Nervous System (CNS). Most researchers now agree that quantifying cortical activity provides a surrogate measure of hypnosis. One such measure is the BIS monitor (Aspect Medical Systems Inc., MA) developed in the mid-1990s (see Section 2.2.1.4). This device has become the reference in consciousness monitoring, and, as a result, most research groups working in control of anesthesia are now using the BIS as a feedback sensor.

However, its inadequacy for use in a closed-loop framework became apparent in the early stage of our work. The BIS algorithms were designed for stand alone monitoring. As a result, they suffer from many shortcomings that limit their applicability for use within a closed loop control framework. In particular, the BIS inherent non-linearity and time delay, as well as the lack of a closed form transfer function describing its dynamic behavior, imply serious closed loop performance limitations.

Analgesia and Autonomic Activity There is currently no technology that can express the patient's analgesic state as a quantified metric. In recent years, however, a renewal of interest in the anesthesia research community for developing such technologies has been observed.

One approach, investigated by a research group in Finland [144, 145], is to quantify the magnitude of the noxious stimulation perceived by the patient's body. When considering that surgical stimulation acts

as an output disturbance, this technology would then provide a measure of such disturbance. However, since it cannot quantify analgesic-based pharmacological depression, this technology cannot be used as an analgesia feedback sensor for controlling analgesic drug titration.

Another approach, which we favor, stems from the fact that the Autonomic Nervous System (ANS) is ultimately responsible for managing the response to noxious stimulation. It is also the target of opioids and other analgesic drugs. Furthermore, considering that the ability of anesthesiologists to maintain the autonomic balance dictates the success of the procedure, we believe that the development of a quantified metric of the autonomic activity can provide a surrogate feedback measure of the analgesic state. In 1985, it was proposed that Heart Rate Variability (HRV), *i.e.*, the variability between consecutive R-peaks of the ECG, be used as a source signal to measure the autonomic balance. A number of patents have been issued, but only one intra-operative monitor using that signal (Anemon-I) was commercialized for that particular purpose. No clinical studies validating this technology were published. The Anemon-I has been withdrawn from the market since the early 2000's. However, recent evidences in the literature suggest that the HRV does carry all of the necessary information to derive a meaningful metric [146, 147, 148].

Quantifying both the CNS and ANS may therefore provide the necessary feedback sensors to close the loop and achieve meaningful clinical performance. In terms of the CNS quantification, the EEG has already been identified as a non-invasive signal containing information relative to the patient's hypnotic state. Some success has been achieved by commercial stand alone monitors. However, their utility as feedback sensors is questionable. In terms of ANS quantification, the HRV signal has emerged as a good candidate to quantify autonomic activity. However, it is a difficult signal to analyze and is sensitive to many different confounding factors.

In this chapter, we present a signal processing technique based on wavelet analysis. This technology, referred to as Wavelet-based Anesthesia Value (WAV), quantifies the state of a system with respect to its two most extreme states (*e.g.*, awake (performing a mental task), and comatose (EEG electrical quiescence) for quantifying cortical activity). As the signal evolves from one state to the next, the changes are quantified and expressed into a bounded scale. In this Chapter, we show that this technology can be used to analyze and characterize both the CNS and ANS activity. We therefore propose two new anesthesia indices: the WAV_{CNS}, which quantifies cortical activity based on EEG analysis, and the WAV_{ANS}, which quantifies autonomic activity based on HRV analysis.

In Section 5.1, we present the wavelet transform, and, in particular, its ability for detecting rapid temporal changes in non-stationary signals. We present in details the WAV technology in Section 5.2 and its application to the quantification of cortical activity based on a single frontal EEG signal. The resulting index, referred to as WAV_{CNS}, was implemented in real-time and coupled to an EEG signal data acquisition system for use in the operating room. The BIS monitor being a clinical reference for consciousness monitoring, this setup allowed us to carry out a clinical study aimed at comparing both

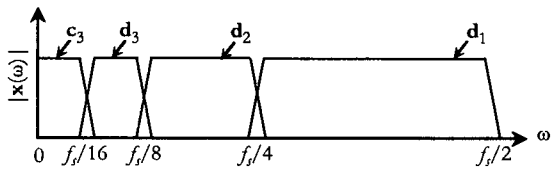


Figure 5.1: Dyadic frequency tiling with $L = 3$.

sensors. The results of this study are summarized in this section. We then show in Section 3 how the WAV technology can be applied to the analysis of the HRV signal. For illustration purposes, we present results from 3 clinical cases, where we compared the WAV_{ANS} to the WAV_{CNS} and the anesthesiologist's assessment of the patient's analgesic state. We also compared the WAV_{ANS} to the heart rate, blood pressure and drug titration recorded during the cases. These preliminary results confirmed the strong potential of the WAV_{ANS} for assessing the autonomic balance. Finally, Linear Time Invariant (LTI) transfer functions modeling the dynamic behavior of both WAV sensors are presented in Section 4.

5.1 The Discrete Wavelet Transform

This section is mainly an overview of wavelets and Wavelet Transform (WT). It is an opportunity to introduce the notations that will be used throughout this Chapter. We also present the rationale behind the use of WT to analyze the EEG in order to estimate the hypnotic depth. We invite interested readers to refer in particular to [149] and [150] for a thorough review about WT and its applications.

5.1.1 Standard Wavelet Dyadic Decomposition

A dyadic wavelet transform w is a transform that decomposes a given signal $\mathbf{x}$ into a number of coefficient sets. Each of these sets represents the activity of the signal in a particular frequency band:

$$w : \mathbf{x} \longrightarrow \left\{ \begin{array}{ll} \mathbf{c}_L^w & \text{band: } [0, fs/2^L] \\ \mathbf{d}_L^w & \text{band: } [fs/2^L, fs/2^{L-1}] \\ \vdots & \\ \mathbf{d}_l^w & \text{band: } [fs/2^{l-1}, fs/2^l] \\ \vdots & \\ \mathbf{d}_1^w & \text{band: } [fs/4, fs/2] \end{array} \right\} \tag{5.1}$$

For instance, a L-level dyadic wavelet decomposition yields $L + 1$ sets of coefficients corresponding to a dyadic tiling of the frequency domain, see Fig. 5.1.

Furthermore, each set $\{\mathbf{c}_L\}$ and $\{\mathbf{d}_l\}_{l=1,\ldots,L}$ is a time series of coefficients describing the time evolution of the signal in the corresponding frequency band. Hence, two signals having the same frequency content will yield different sets of coefficients if the signal patterns are different (*i.e.*, if the frequency components have different phases).

Similarly to Fourier analysis, which was made practical by the discovery of the Fast Fourier Transform (FFT) algorithm, wavelet analysis remained the apanage of mathematicians untill the mid-1990s when the use of filter banks appeared to yield much faster and efficient ways of performing the necessary calculations. Before then, wavelet analysis existed only in its continuous form, involving a very large computational complexity ill-suited for real-time applications. In a Discrete Wavelet Transform (DWT) framework, the wavelet decomposition is obtained through the use of a filter bank and 2 FIR filters, H_0 (low pass) and H_1 (high pass). The characteristic of the coefficient sets $\{\mathbf{c}_L\}$ and $\{\mathbf{d}_l\}_{l=1,\ldots,L}$ therefore depends on the wavelet filters H_0 and H_1 associated to the transform w.

5.1.2 Change in Phase and Wavelets

The coefficients obtained through the Wavelet Transform (WT) capture both the frequency and time information of a given signal. Wavelet analysis is therefore particularly well suited to track both time and frequency changes in a signal. The analysis of many biological signals have benefited from wavelet analysis.

To illustrate this, let us consider a signal composed of 3 frequency components. The phase of one of the component is allowed to drift slowly in time from 0 to π following the time course shown in Figure 5.2.d. Even though the change in the signal pattern is obvious (Figure 5.2.a), the power spectra (Figure 5.2.c) was unsuccessful in tracking the changes in the signal latency, while the wavelet decomposition (Figure 5.2.b) clearly shows an evolution of the signal over time.

It is interesting here to make a parallel with the arguments presented by Bowles *et al.* [52] in favor of bispectral analysis (see Figure 2.4). Following Rampil's observation [51] that anesthetic drugs tend to synchronize the generation of postsynaptic potentials and thus affect the signal latency, it appears that the time-frequency localization property of the WT makes it well suited for capturing the evolution of the EEG with increasing anesthetic depth. This is illustrated in Figure 5.2.

5.2 Estimating the Anesthetic Drug Effect: the WAV_{CNS}

Note that contributions in the development of the WAV_{CNS} are shared with Ms. T. Zikov [151].

5.2.1 Concepts and Derivation

The brain is the target organ of anesthetic drugs.. The electroencephalogram (EEG) signal being a non-invasive measurement of cortical activity, the effects of anesthetic drugs onto this signal has been thoroughly studied. Because of differences in the mechanism of action, the effect of different anesthetics usually results in different EEG patterns. However, for most drugs and with increasing blood concentration, the EEG signal evolves from a low-amplitude, large-bandwidth, noise-like signal, to a high-amplitude slow-waves signal. If large amounts of anesthetic drugs are used, the EEG activity eventually disappears, resulting in

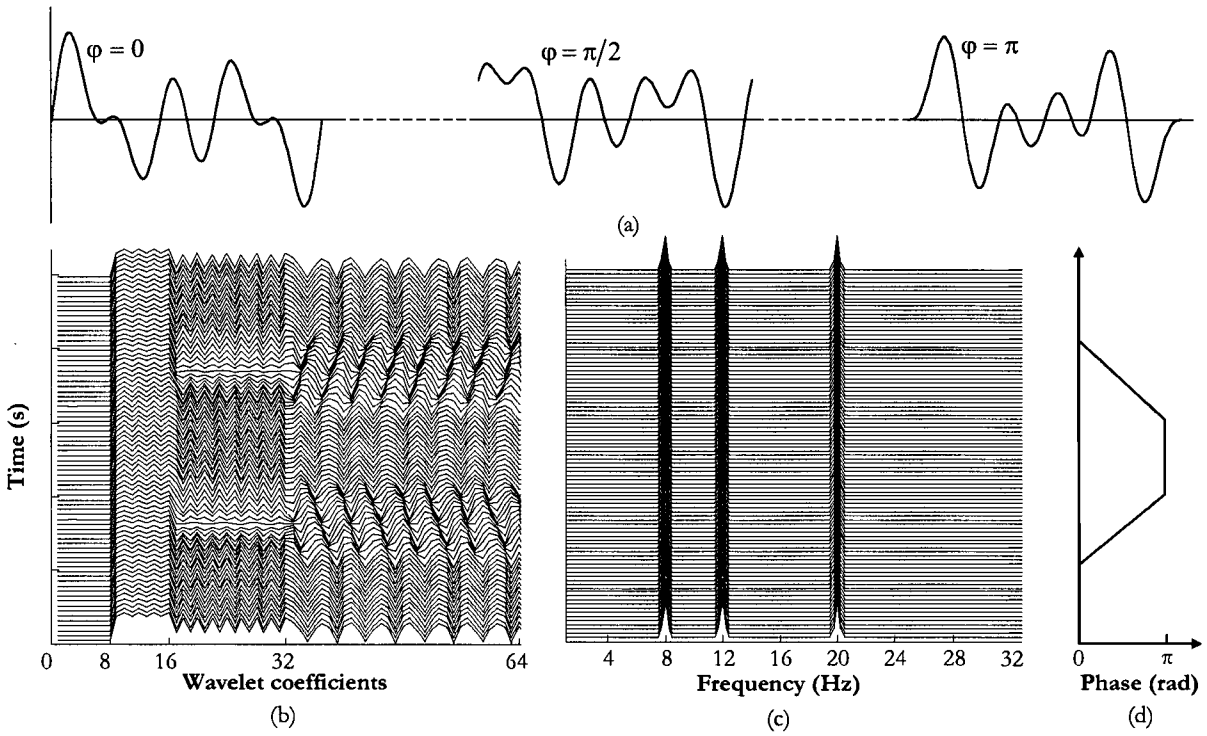


Figure 5.2: Changes in latencies in a signal can also be tracked by the wavelet coefficients (from [151]). (a) Time series signal. Note how the signal pattern changes significantly when changing the phase of one of the frequency component. (b) 'Periodogram' of the wavelet coefficients (note how the change in phase is clearly noticeable in the wavelet domain). (c) Classical Fourier periodogram. (d) Time profile of the phase change.

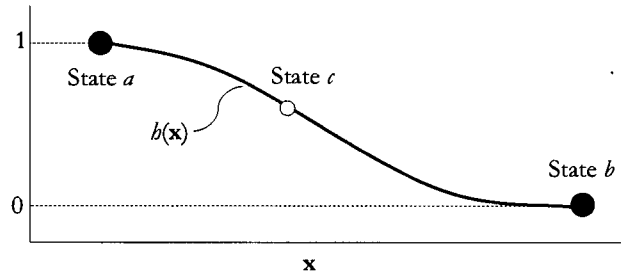


Figure 5.3: The function h quantifies the state of a system by attributing a unique value to each of the operating modes of the system. If the system evolves in a monotonous fashion from state a to state b, it is required that h is also monotonous.

an isoelectric signal (*i.e.*, complete absence of cortical activity).

Since this trend can be observed for most anesthetic agents, it is commonly assumed that the EEG can be used to estimate the hypnotic state of the patient, and thus provide a measurement of the drug effect. In order to avoid complex and time-consuming interpretations of raw EEG signals, a common approach is to extract a single univariate value that represents the patient's state. Since there is no gold standard for measuring depth of anesthesia, the various analysis techniques used to calculate this value are usually derived such that the resulting quantitative value is sufficiently correlated to the anesthesiologist's qualitative assessment of the patient's state.

In order to illustrate our methodology for deriving an index of hypnosis, let us consider the patient to be a system which can evolve between two states a and b. Our goal is to derive an index that characterizes the operating mode of this system, *i.e.*, finding a function h, which, given an observation $\mathbf{x}$ acquired while the system operates between a and b, yields a single value i comprised between two finite bounds, *e.g.*, 1 and 0 (see Fig. 5.3):

$$h : \mathbf{x} \in [a, b] \longrightarrow i \in [0, 1]. \tag{5.2}$$

In terms of hypnotic depth, the observation $\mathbf{x}$ can be a short EEG epoch, long enough to contain the necessary information, and short enough to allow frequent updates of i. The states a and b correspond to the two extremes of the signal observational window, *i.e.*, the awake and isoelectric EEG. As far as anesthetic drugs are concerned, this window is sufficient since it is usually not desired to titrate patients beyond the point where cortical activity is totally suppressed.

In this section, we will first give some very general guidelines on how to define the function h. We will then apply these guidelines to the estimation of the anesthetic depth, using data obtained from healthy subjects and anesthetized patients.

5.2.1.1 Main Concept and Approach

In order to define the function h, we first need to assume that both states a and b correspond to well established operating modes from which observation data are available. In the case of EEG analysis, this

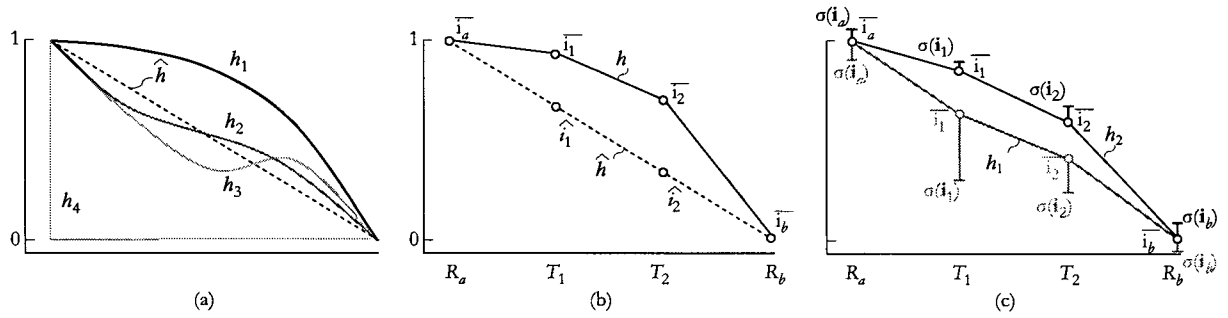


Figure 5.4: The characteristic of the function h depends on the selection of the feature function f. (a) Both h_1 and h_2 are valid functions, whereas h_3 and h_4 cannot be used to represent adequately intermediate states. (b) The linearity of a given function h can be assessed by introducing new data sets. (c) Variability is another important aspect. For instance, even though h_1 is nearly linear, it fails in characterizing properly the intermediate states T_1 and T_2. The function h_2 is preferred here since it allows for the proper discrimination between consecutive states.

assumption is verified since a corresponds to the awake state (*e.g.*, when able to perform a mental task) and b corresponds to an isoelectric signal. Since these states are well defined, we can then obtain two reference data sets corresponding to observations of the system in the states a and b:

$$\begin{cases} R_a = \{\mathbf{x}_{a,k}, \quad k = 1, 2, \ldots M\} & \text{(state } a\text{)}, \\ R_b = \{\mathbf{x}_{b,k}, \quad k = 1, 2, \ldots M\} & \text{(state } b\text{)}, \end{cases} \tag{5.3}$$

where each observation vector $\mathbf{x}_{a,k}$ and $\mathbf{x}_{b,k}$ contains a finite number of samples and represent the k^{th} epoch of the data sets R_a and R_b.

To characterize the data sets, a chosen feature $\mathbf{f}$ is calculated from each epoch. We denote the feature function f as:

$$f : \mathbf{x} \longrightarrow f(\mathbf{x}) = \mathbf{f}. \tag{5.4}$$

Each epoch of the reference data sets is then associated with a feature $\mathbf{f}_{a,k}$ or $\mathbf{f}_{b,k}$. This feature can be either a scalar (such as average value, Root Mean Square (RMS) amplitude, maximum or minimum value, standard deviation, *etc.*) or a vector (such as histogram, probability density function, *etc.*) derived in the original signal domain or in a transformed signal domain. A particular state can then be characterized by averaging the feature sets $\{\mathbf{f}_{a,k}\}_{k=1,\ldots,M}$ and $\{\mathbf{f}_{b,k}\}_{k=1,\ldots,M}$. This results in two references which are the averaged features $\overline{\mathbf{f}_a}$ and $\overline{\mathbf{f}_b}$:

$$\begin{cases} \overline{\mathbf{f}_a} = \dfrac{1}{M} \cdot \sum_{k=1}^{M} \mathbf{f}_{a,k}, \quad \text{and} \\ \\ \overline{\mathbf{f}_b} = \dfrac{1}{M} \cdot \sum_{k=1}^{M} \mathbf{f}_{b,k}. \end{cases} \tag{5.5}$$

Let us now assume that the system operates in the intermediate state c and that an observation $\mathbf{x}_c$ is provided. Using the corresponding feature $\mathbf{f}_c = f(\mathbf{x}_c)$, it is then possible to estimate how far the system has evolved from the state a towards the state b by comparing the value $\mathbf{f}_c$ to the references $\overline{\mathbf{f}_a}$ and $\overline{\mathbf{f}_b}$. Two

indices j_α and j_β are defined such that:

$$\begin{cases} j_\alpha = ||\mathbf{f}_c - \overline{\mathbf{f}_a}||_1, & \text{and} \\ j_\beta = ||\mathbf{f}_c - \overline{\mathbf{f}_b}||_1, \end{cases} \tag{5.6}$$

where the norm $||\cdot||_1$ for a vector $\mathbf{v}$ is defined as:

$$||\mathbf{v}||_1 = \sum_{n=1}^{N} |v(n)|. \tag{5.7}$$

The norm $||\cdot||_1$ quantifies the difference between $\mathbf{f}_c$ and $\overline{\mathbf{f}_a}$ (or $\overline{\mathbf{f}_b}$) by integrating the distance between these two vectors. The indexes j_α and j_β thus measure a distance of the system from either state a or b. Note that higher degree norms can be used for this analysis, however, they would emphasize larger differences which might be due to outliers, hence leading to noisier indexes.

In order to maximize the signal-to-noise ratio, the two indexes j_α and j_β can be combined to yield a single descriptor $j = j_\beta - j_\alpha$. It is convenient here to define a function g that, given an observation $\mathbf{x}$, associates the resulting index j:

$$g : \mathbf{x} \longrightarrow g(\mathbf{x}) = j = ||f(\mathbf{x}) - \overline{\mathbf{f}_b}||_1 - ||f(\mathbf{x}) - \overline{\mathbf{f}_a}||_1. \tag{5.8}$$

In order to scale the final index i such that $i \to 1$ (state a) and $i \to 0$ (state b), the function g is applied to the entire reference data sets. This yields two sets $\mathbf{j}_a$ and $\mathbf{j}_b$:

$$\begin{cases} g : R_a \longrightarrow \mathbf{j}_a = \{j_{a,k}, \quad k = 1, 2, \ldots M\} \\ g : R_b \longrightarrow \mathbf{j}_b = \{j_{b,k}, \quad k = 1, 2, \ldots M\} \end{cases} \tag{5.9}$$

The function h is a scaled version of g, and is thus defined as:

$$h : \mathbf{x} \longrightarrow h(\mathbf{x}) = i = g(\mathbf{x}) \cdot \frac{1}{\overline{j_a} - \overline{j_b}} - \frac{\overline{j_b}}{\overline{j_a} - \overline{j_b}}, \tag{5.10}$$

where:

$$\begin{cases} \overline{j_a} = \frac{1}{M} \cdot \sum_{k=1}^{M} j_{a,k}, & \text{and} \\ \\ \overline{j_b} = \frac{1}{M} \cdot \sum_{k=1}^{M} j_{b,k}. \end{cases} \tag{5.11}$$

5.2.1.2 Feature Function Selection

The performance of the index i as a substitute to the qualitative assessment of the patient's state clearly depends on the characteristic of the function h over the state continuum $[a, b]$. This is illustrated in Fig. 5.4.a where h_1 and h_2 are adequate as they uniquely define each intermediate state. Conversely, a function such as h_3 should be avoided as there exist different states which yield similar index values, thus leading to confusion in the interpretation of the index. As a rule of thumb, any function whose characteristic exhibits

a change of sign in its first derivative should be avoided. Similarly, a function such as h_4 is not very useful since it does not discriminate properly between different states.

The selection of f is critical in shaping the characteristic of h. For instance, a good discrimination between consecutive intermediate states is achieved if a feature function f is determined such that the evolution between the states a and b is linear. We define $\hat{f}$ (resp. $\hat{h}$) such a function.

While linearity is an important aspect, it is not sufficient to guarantee proper discrimination between states. Another important aspect is the variability of the index i while the system operates in a stationary mode. A large variability may cause a significant overlap between consecutive states, indicating that either the observation signal or the feature function are poorly suited for this analysis.

In order to facilitate the selection of the feature function f, we then introduce two design parameters: L (linearity parameter) and V (variability parameter). Both parameters rely on the use of new data sets T_1, T_2, ..., T_N corresponding to intermediate states of operation (*e.g.*, sedation, general anesthesia, deep hypnotic state, *etc.*). For each data set, R or T, we can associate a desired index value $\hat{i}$:

$$\hat{h} : \{R_a, T_1, \ldots, T_N, R_b\} \longrightarrow \{1, \hat{i_1}, \ldots, \hat{i}_N, 0\}. \tag{5.12}$$

Note that, by definition of h, $\hat{i_a} = 1$, $\overline{\hat{i_a}} = 1$, $\hat{i_b} = 0$, and $\overline{\hat{i_b}} = 0$.

Linearity Linearity can be easily assessed by calculating the corresponding index for each epoch of each data set. We thus obtain a set of vectors $\{\mathbf{i}_n\}$ of length equal to the number of epochs in the corresponding data set:

$$h : \{R_a, T_1, \ldots, T_N, R_b\} \longrightarrow \{\mathbf{i}_a, \mathbf{i}_1, \ldots, \mathbf{i}_N, \mathbf{i}_b\}. \tag{5.13}$$

We can assess the characteristic of h by simply comparing the average vector values $\overline{i_n} = \text{mean}(\mathbf{i}_n)$ with the set of desired values $\hat{i}_n$ in (5.12), see Fig. 5.4.b. The parameter L is the normalized root mean square error between the characteristics of h and $\hat{h}$:

$$L = \sqrt{\frac{\sum_{n=1}^{N} \left(\overline{i_n} - \hat{i}_n\right)^2}{\sum_{n=1}^{N} \hat{i}_n{}^2}}. \tag{5.14}$$

L is bounded between 1 and 0. Note that to normalize L, we divided the root mean square error of $(h - \hat{h})$ by the maximum possible root mean square error (*e.g.*, the function h_4 in Fig. 5.4.a. would typically have $L = 1$). To optimize the linearity of h, it is therefore necessary to minimize L.

Variability The variability of h can be assessed by calculating for each vector $\mathbf{i}_n$ of (5.13) the standard deviation $\sigma(\mathbf{i}_n)$. If a statistically significant discrimination of each intermediate state T_1, T_2, ..., T_N is a critical feature of h, it is necessary to ensure that:

$$\sigma(\mathbf{i}_n) + \sigma(\mathbf{i}_{n+1}) < |\overline{i_n} - \overline{i_{n+1}}|, \tag{5.15}$$

i.e., that the standard deviation intervals of two consecutive states does not overlap, see Fig. 5.4.c. It is, however, important to keep in mind that the standard deviation $\sigma(\mathbf{i})$ is strongly influenced by the eventual

non-stationarity of the data set T. This is mostly true when T is composed by data acquired from different systems or under different recording conditions. This is also true when the data sets are collected based on the evaluation of a human operator.

We define the design parameter V as:

$$V = \sqrt{\frac{1}{N+1} \cdot \sum_{n=0}^{N} \left[\frac{\sigma(\mathbf{i}_n) + \sigma(\mathbf{i}_{n+1})}{\overline{i_n} - \overline{i_{n+1}}} \right]^2} \tag{5.16}$$

where $\sigma(\mathbf{i}_0) = \sigma(\mathbf{i}_a)$ and $\sigma(\mathbf{i}_{N+1}) = \sigma(\mathbf{i}_b)$. For the same reasons outlined above, it is advised to keep $V < 1$, or as small as possible if the stationarity of each set T cannot be properly guaranteed.

5.2.1.3 Data Sets

In the search for a function h that would adequately estimate the anesthetic depth, 5 different observation sets corresponding to 5 distinct hypnotic states were recorded, see Fig. 5.5 for representative samples:

- R_a (awake state): 15 minutes recorded from 5 healthy adult subjects (3 minutes each). Subjects were asked to keep their eyes shut and minimize muscle activity while concentrating on a mental task.

- T_1 (light REM sleep): 15 minutes recorded from 3 subjects (5 minutes each). Only epochs without ocular artifacts were included in this set.

- T_2 (general anesthesia): 18 minutes recorded from 6 patients (3 minutes each) undergoing minimally invasive arthroscopy surgery.

- T_3 (deep hypnotic state): 9 minutes recorded from 5 patients. Delta activity is usually transitory, which explains the limited amount of data available.

- R_b (isoelectric EEG): $\approx$5 minutes recorded from 2 patients exhibiting electrical quiescence.

The data sets T_2, T_3 and R_b were collected from surgical audit cases. While a variety of anesthetic regimens were used during these cases, only drugs provoking a concentration dependent depression of the cortical activity were used (vapour anesthetics (isoflurane, desflurane, sevoflurane, NO_2), propofol and midazolam).

While a simple observation of the EEG was sufficient to extract periods of isoelectric activity (R_b), the classification of EEG epochs between T_2 and T_3 was done following the anesthesiologist's assessment for each patient. The anesthesiologist was first asked to identify periods of adequate anesthetic depth, *i.e.*, when patients were unconscious in the context of surgical anesthesia. The EEG corresponding to these periods were then reviewed post hoc to identify periods of deeper hypnosis (apparent delta waves) which were then classified as the T_3 set. The remaining EEG data were then included in the T_2 set.

Each EEG signal was recorded from a frontal differential channel (F_{p1}-AT_1 (the left outer malar bone) and F_{pz} as ground), using the Crystal Monitor Model 16^{TM} (Cleveland Medical Devices Inc., OH), a

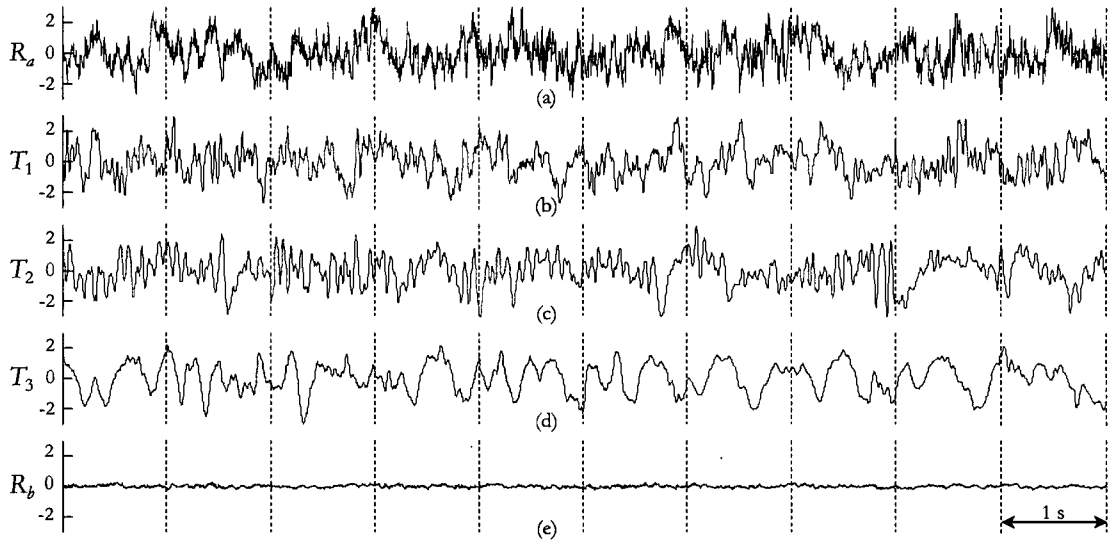


Figure 5.5: Normalized EEG signals at different anesthetic depths. Each 1-second epoch is first detrended and then normalized using the RMS amplitude. (a) Awake healthy subject (b) Light REM sleep (the spikes due to REM activity were manually removed) (c) General anesthesia (d) Deep anesthetic state (δ-waves) (e) Isoelectric state (normalized by the RMS amplitude of the last non-isoelectric epoch).

sampling frequency of 480 S/s and a resolution of 16 bits. The signals were then resampled at 256 S/s and notch-filtered at 50/60 Hz prior to analysis.

The EEG signals were further divided into 1-second epochs (256 samples per epoch). Because the signals were measured through contact electrodes with a high input impedance amplifier, there are a number of factors (*e.g.*, skin conductivity and thickness, electrode impedance and placement) which can affect the amplitude of the recorded signal. To minimize the influence of these factors, each epoch is first detrended and then normalized by its RMS value prior to analysis:

$$\tilde{\mathbf{x}} = \frac{\mathbf{x} - \overline{\mathbf{x}}}{\mathrm{RMS}(\mathbf{x})}. \tag{5.17}$$

This also compensates for the bias and calibration errors that are expected between different acquisition devices. Note that for the isoelectric epochs, the normalization was performed using the RMS value of the last EEG epoch preceding the occurrence of isoelectricity.

Finally, the desired values in terms of the mean index values were defined for each data set to reflect the anesthesiologist's assessment:

$$\hat{h}_{\mathrm{WAV_{CNS}}} : \{R_a, T_1, T_2, T_3, R_b\} \longrightarrow \{1, 0.75, 0.5, 0.25, 0\} \tag{5.18}$$

5.2.1.4 Feature Function $f_{\mathbf{WAV}}$

In order to extract information from the EEG signal pertinent to depth of anesthesia, the Probability Density Function (PDF) of the wavelet coefficient set $\mathbf{d}_l^w$ was chosen as the feature function, and we denote

it as $f_{\text{WAV}_{\text{CNS}}}$:

$$f_{\text{WAV}_{\text{CNS}}} : \tilde{\mathbf{x}} \longrightarrow f_{\text{WAV}_{\text{CNS}}}(\tilde{\mathbf{x}}) = \text{PDF}(\mathbf{d}_l^w). \tag{5.19}$$

This choice is motivated by a number of facts.

First, the information regarding the anesthetic depth of a patient lies in both the spectral distribution of the EEG and the pattern of the signal itself. The wavelet transform is therefore an appropriate tool to extract this time-frequency information. However, the resulting coefficient sets cannot be used right away. It is necessary to format the information they contain using statistical tools. The interest in using the PDF in this case lies in the fact that the area of a PDF curve is, by definition, always equal to 1. Therefore, the indexes j_α and j_β are naturally bounded between 0 and 2.

Secondly, even though we have already chosen the particular structure of f, this choice offers design flexibility through the selection of the frequency band l and wavelet transform w. Thus, by iterating between different wavelet filters and frequency bands, a search algorithm can easily be written to determine the best l and w.

5.2.1.5 Selection of the Frequency Band and Wavelet Filter

Let us denote with WAV_{CNS} the index i defined in (5.10), and corresponding to the feature function $f_{\text{WAV}_{\text{CNS}}}$. To select the appropriate frequency band l and wavelet transform w, we assess the linearity and variability of the WAV_{CNS} for different frequency bands and wavelet filters using the design parameters L and V defined in Section II (5.14) and (5.16).

The linearity and variability of the WAV_{CNS} were assessed for the three frequency bands $\mathbf{d}_1^w$, $\mathbf{d}_2^w$ and $\mathbf{d}_3^w$, and for the first 16 Daubechies wavelet filters, see Fig. 5.6.

In addition to the standard definition of L given in the previous section, we added a sign information to this otherwise positive value. A change in the sign of the linearity parameter displayed in Fig. 5.6 indicates a change in the dominant curvature of h (a positive value indicates concavity, while a negative value indicates convexity). This observation is of interest when considering that, depending on the concavity or convexity of the function, the resulting index will either be more sensitive to changes in the lighter hypnotic states and less sensitive to changes in the deeper states (convexity), or conversely, more sensitive to changes in the deeper hypnotic states and less sensitive to changes in the lighter states (concavity). The selection of the frequency band and filter allows for the modification of this characteristic.

The linearity analysis indicates that, either the γ-band (32-64 Hz) in conjunction with any wavelet filter of order >4, or the EMG band (64-128 Hz) and the Daubechies filter #2, yield a nearly linear characteristic. These choices also agree with the variability criterion according to which the parameter V should remain well below 1.

Note that these results are in agreement with the recent findings suggesting a fundamental relationship between consciousness and γ activity. It has also been reported by John *et al.* [152] that loss of consciousness after anesthetic administration showed a marked drop in the γ-band activity, along with an increase in slower

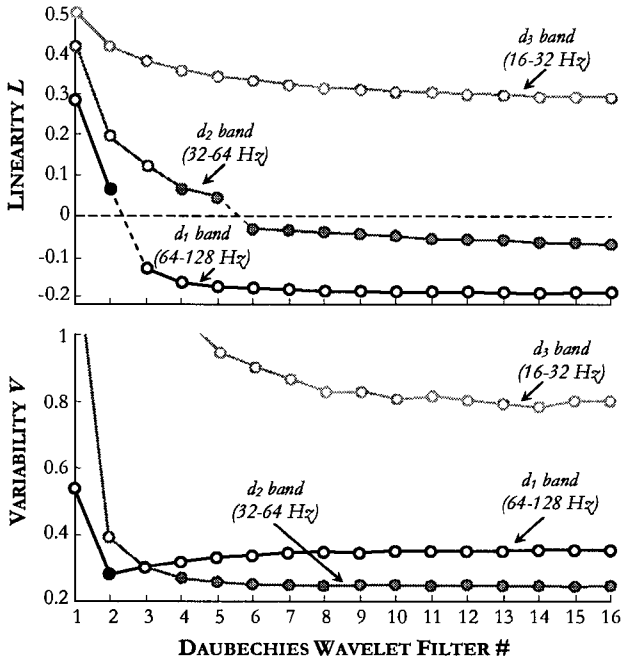


Figure 5.6: Selection of the frequency band and wavelet filter for the sampling frequency of 256 S/s. (a) Linearity parameter L. A positive value indicates a concave characteristic, while a negative value indicates convexity. By appropriately choosing the wavelet filter, the characteristic of h can change from concave to convex. (b) Variability parameter.

waveforms, which is in line with our findings.

It is also interesting to note that the EMG band yields a similar result for the particular wavelet filter Db_2. However, this is explained by the fact that this low order filter still preserves a significant amount of γ-band information. Higher order filters, which extract mostly EMG information, lead to a convex characteristic of the index which does not give a good discrimination between deeper states.

Given these choices, it is advisable to use the γ-band instead of the EMG band. First, the use of the γ-band decreases computational complexity by allowing to lower the sampling rate to 128 S/s, and thus, use the $\mathbf{d}_1^w$ coefficient set instead of $\mathbf{d}_2^w$ via a single convolution operation. Also, the signal corresponding to the EMG band is of lower power than that of the γ-band, which makes it more susceptible to environmental

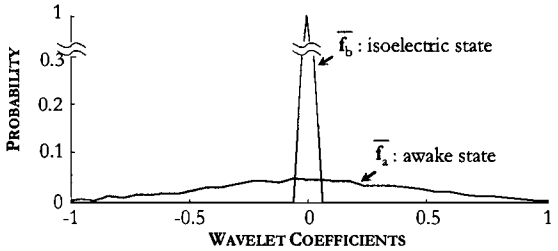


Figure 5.7: Awake and isoelectric reference PDFs based on $f^{\text{opt}}_{\text{WAV}_{\text{CNS}}}$.

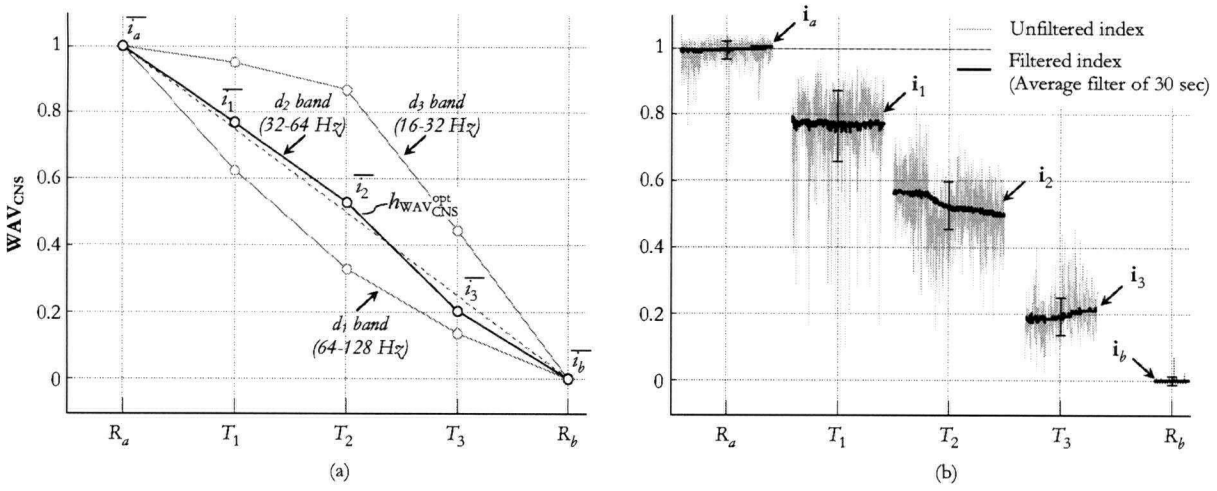


Figure 5.8: Characteristic of the optimal function $h^{\text{opt}}_{\text{WAV}_{\text{CNS}}}$ for the sampling frequency of 256 S/s. (a) Linearity characteristic. (b) WAV_{CNS} values for the data sets R_a, T_1, T_2, T_3 and R_b. Note the larger variability in the intermediate data sets.

noise. Furthermore, the use of the EMG band would also require to notch-filter both at 50/60 Hz and 100/120 Hz to reduce the electrical noise from the mains. Finally, the EMG band is more susceptible to neuromuscular blocking agents, which aim at suppressing muscle activity. An index based on this band would therefore be particularly sensitive to this type of drugs.

For the γ-band, an optimal result is obtained when using the Daubechies wavelet filter #6 ($w[\text{Db}_6]$). The optimal feature function $f^{\text{opt}}_{\text{WAV}_{\text{CNS}}}$ is thus defined as:

$$f^{\text{opt}}_{\text{WAV}_{\text{CNS}}} : \tilde{\mathbf{x}} \longrightarrow f^{\text{opt}}_{\text{WAV}_{\text{CNS}}}(\tilde{\mathbf{x}}) = \text{PDF}(\mathbf{d}_1^{w[\text{Db}_6]}), \tag{5.20}$$

where the signal $\tilde{\mathbf{x}}$ is sampled at 128 S/s. Figure 5.7 illustrates the reference awake and isoelectric PDFs based on $f^{\text{opt}}_{\text{WAV}_{\text{CNS}}}$ averaged over the sets R_a and R_b. For more details on the derivation of the PDFs, see Section 5.2.2.2.

The corresponding characteristic of $h^{\text{opt}}_{\text{WAV}_{\text{CNS}}}$ is plotted in Fig. 5.8.a. Results show a nearly linear characteristic where:

$$h^{\text{opt}}_{\text{WAV}_{\text{CNS}}} : \{R_a, T_1, T_2, T_3, R_b\} \longrightarrow \{100, 77, 54, 20, 0\} \tag{5.21}$$

Note that the values in (5.21) are average values obtained from the training data sets.

For each data set, the corresponding WAV_{CNS} values are plotted in Fig. 5.8.b. The variability in the awake and isoelectric states is small as compared to the variability observed from the intermediate states. This is due to the fact that the awake and isoelectric states are, by definition, very well identified. Conversely, the intermediate states are subject to interpretation. This results in a higher non-stationarity of the T data sets.

5.2.2 Practical Issues and Implementation

This section addresses practical issues concerning the implementation of the WAV_{CNS} algorithm.

5.2.2.1 Pre-Processing

Notch Filtering The EEG signal is a low power signal particularly susceptible to environmental noise. For instance, electrical heaters used as warming units in the operating rooms draw large amount of current from the main electrical outlets. When these units are in proximity of the EEG acquisition device, strong 50/60 Hz and 100/120 Hz perturbations may be superimposed to the EEG signal. While appropriate shielding can alleviate this problem, it is necessary to use a 50/60 Hz notch filter to remove this noise.

Resampling The EEG signal needs to be resampled to a rate suitable for the analysis. Since we have determined that the γ-band (32-64 Hz) is of interest, a sampling rate of 128 S/s is optimal. Note again that an advantage of using this sampling rate is that only one convolution of the filter $w[\text{Db}_6]$ with the signal is needed to obtain the $\mathbf{d}_1^{w[\text{Db}_6]}$ coefficients. When higher sampling rates are used, the wavelet coefficients in the band of interest are obtained through a series of filter banks, which adds to the computational complexity.

Artifact Detection and De-noising Ocular Artifacts (OAs) and head movements can make the WAV_{CNS} particularly unstable in lighter hypnotic levels (awake and light sedation), when the patient is still conscious. This difficulty arises mostly before and during induction.

One strategy to deal with these artifacts is to systematically reject corrupted epochs from the analysis and simply set the index to its previous value. However, this can result in substantial data loss.

A better approach is to remove the artifact from the raw EEG signal, without perturbing the high frequency content of the EEG. This can be easily done for two reasons. First, OAs and head movements typically perturb the EEG from 0 Hz up to 16 Hz. Thus, the high frequency information is not affected. However, they substantially affect the RMS value of the signal, and thus the entire analysis. Second, this type of artifact signal, which is superimposed with the true cortical EEG signal, is of high amplitude compared to the relatively small EEG amplitude. Using a redundant wavelet decomposition, which provides a good temporal resolution, this signal induces large coefficients that are easily distinguishable from the smaller background coefficients of the true EEG. A simple technique such as thresholding of wavelet coefficients in the lower frequency bands (up to 16 Hz) is sufficient to remove the perturbing artifacts from the source EEG, see Fig. 5.9.a [59]. This technique is particularly effective in removing most of the artifacts while keeping the exact amplitude and phase of the high frequency components, see Fig. 5.9.b.

A difficulty arises during the state transition from awake to anesthetized. This transition is fairly rapid. Further, the initial induction bolus usually drives the patient into a deep hypnotic state. As a result, δ-waves are likely to occur within a few seconds from the loss of consciousness. The δ-waves being particularly

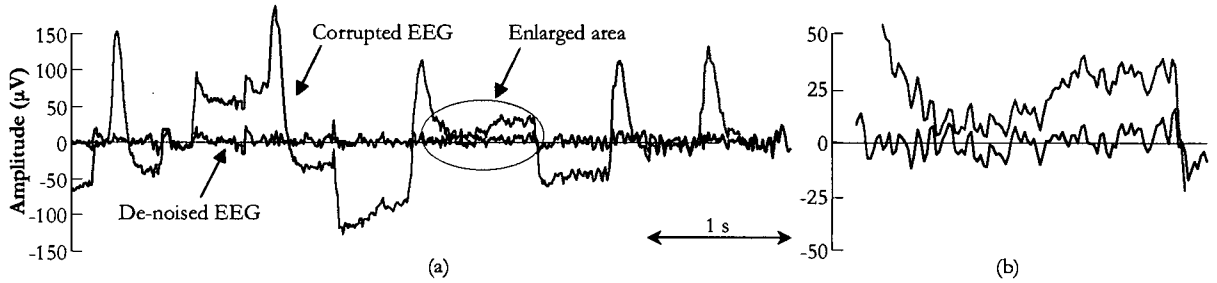


Figure 5.9: Effect of ocular artifact de-noising during an awake period. (a) Raw EEG and de-noised EEG. (b) Note how the high frequency information remains unaltered by the de-noising technique, both in terms of amplitude and phase.

similar to OAs (*i.e.*, low frequency, high amplitude), it is necessary to switch off the de-noising technique as soon as the patient does not exhibit ocular activity.

The moment of anesthesia induction, which corresponds to the bolused administration of an intravenous anesthetic, is suitable for switching off the de-noising technique as the patient rapidly looses consciousness. From the perspective of automation in clinical anesthesia, this solution is viable since the computer is allowed to directly command the infusion device.

5.2.2.2 The WAV_{CNS} Algorithm

Normalization For the reasons outlined in Section 5.2.1.3, the EEG epochs are first detrended and normalized prior to analysis. This normalization is carried out on EEG epochs with RMS amplitudes above 4 μV_{RMS}.

During electrical quiescence (*i.e.*, when the RMS amplitude of the EEG is less than 4 μV_{RMS}), the normalization using RMS values close to zero would lead to the amplification of measurement noise. Therefore, we directly assign the feature $\overline{\mathbf{f_b}}$ (see Fig. 5.7) to any isoelectric EEG epoch:

$$f^{\text{opt}}_{\text{WAV}_{\text{CNS}}} : \mathbf{EEG}_{isoelectric} \rightarrow \overline{\mathbf{f_b}}. \tag{5.22}$$

The Redundant Wavelet Transform For the derivation and implementation of the WAV_{CNS} algorithm, we use a redundant DWT, so-called Stationary Wavelet Transform (SWT), which provides the same time resolution in all frequency bands of decomposition.

There is a number of reasons for using a redundant transform. First, when choosing an optimal frequency band for the analysis, the SWT provides the same number of coefficients in each frequency band (*i.e.*, it is equal to the sampling rate for a 1-second long epoch), and therefore enables easier comparison between them. (For a non-redundant transform, the number of coefficients decreases by a factor 2 for each level of decomposition.) In addition, a greater number of coefficients results in better temporal resolution in the decomposition bands, and, therefore, smother reference PDFs.

Secondly, at the pre-processing stage, the better time resolution of the SWT provides an improved

characterization of artifacts, as well as a smoother estimation of the signal of interest (*i.e.*, the true cortical EEG) after thresholding in the wavelet domain [153, 154].

However, the utilization of the SWT somewhat increases the computational demand, since the complexity of the SWT is $O(N \cdot log_2 N)$ for a signal of N samples, in comparison to $O(N)$ for the DWT.

Finally, note that caution has to be taken when performing the wavelet decomposition on epochs of short length, since boundary effects can significantly perturb the result. Therefore, prior to the decomposition, each epoch is extended using a standard procedure such as symmetrization. After applying the SWT, only the coefficients that are not affected by boundary effects are kept for further analysis.

Probability Density Function When calculating the PDF of the wavelet coefficients, special care has to be taken when selecting an appropriate number of bins. It has been shown [155] that the optimal histogram bin size, which provides the most efficient, unbiased estimation of the probability density function, is achieved when:

$$B_w = 3.49 \cdot \sigma \cdot N^{-1/3}, \tag{5.23}$$

where B_w is the width of the histogram bin, σ is the standard deviation of the distribution and N is the number of available samples.

However, when it comes to estimating the depth of consciousness, the statistics of the EEG signal can significantly vary between the different states. For example, the variance of the wavelet coefficients in the γ-band for the awake state is much greater than for the anesthetized state. Furthermore, for the isoelectric EEG, the variance is almost equal to zero (see Fig. 5.7).

Hence, an optimal number of bins defined by (5.23) for the proper representation of the awake state might actually be too small to adequately represent deeper states, and might cause them to be mistaken for the isoelectric state. Similarly, if the number of bins is too large, the PDF waveform might exhibit a 'comb' effect in the awake and light sedation states, while precisely defining the deeper states. Therefore, a compromise needs to be reached when choosing the number of bins for the analysis so that all states are represented in a satisfactory manner. Based on the length of an epoch, a number of bins equal to the sampling frequency of the signal appears to provide an adequate compromise. In fact, the isoelectric reference PDF becomes zero everywhere except at the origin, where it is equal to one.

5.2.2.3 Post-Processing

Trending Using an analysis epoch as short as 1 second can lead to a particularly noisy WAV_{CNS} index. To overcome this problem, a filtering stage has to be added to the WAV_{CNS} algorithm to extract the most significant information.

There are a number of techniques which can be used. The simplest one is to use an averaging window, where the displayed index becomes the mean value of the previous index values over a fixed horizon. For example, the BIS monitor uses a window of 30 or 15 seconds, since it yields good results in terms of

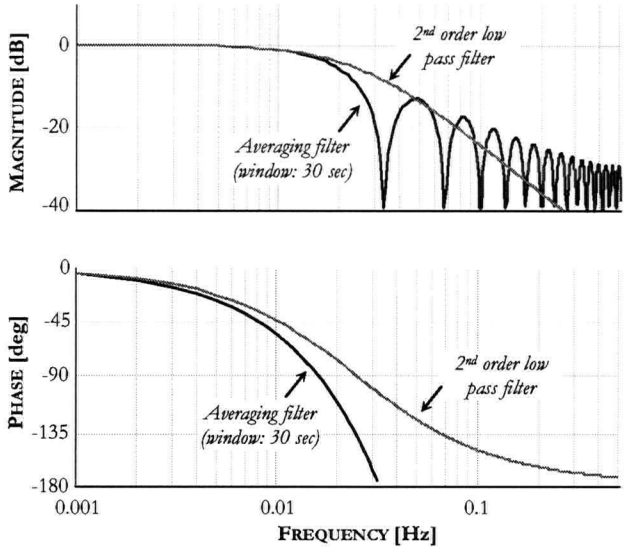


Figure 5.10: Characteristic of a 30-second averaging filter and a 2^{nd} order low pass filter ($\omega_0 = 0.02$ Hz and ζ=1).

noise reduction. During the acquisition of our clinical data, the BIS monitor used a 30-second averaging period. Therefore, we have chosen the same averaging period for our analysis, in order to perform the direct comparison between the WAV_{CNS} and the BIS. A property of this filter is that the filtered index only depends on the past 30 seconds of data, which makes its interpretation straightforward.

However, while FIR filters are simple to implement, they introduce a non-minimum phase element which is detrimental in applications involving control and identification. In that respect, the use of an IIR filter is more suitable. For instance, a simple second order low pass filter characterized by its parameters ω_0 and ζ, such as described in Fig. 5.10, gives a better high frequency noise rejection profile, while being minimum phase. We found that an IIR filter with a cutoff frequency of 0.02 Hz and damping factor of 1 is an adequate alternative to the 30-second averaging window filter.

Scaling To allow for a better comparison with the BIS Monitor (see Section VI), the WAV_{CNS} was scaled between 100 (awake state) and 0 (isoelectric state) by multiplying the right hand side of (5.10) by 100.

5.2.3 Clinical Results

To validate the WAV_{CNS}, a clinical study was conducted in 2002 at the University of British Columbia Hospital (UBCH). The aim of the study was to compare the real-time WAV_{CNS} to the BIS v.3.4 (Aspect Medical Systems, MA) within a clinical setting. Note that, to allow for a better comparison, the WAV_{CNS} IIR trending filter was replaced by a 30 seconds averaging filter similar to that of BIS. The study protocol and demographics are summarized in the Annex D.1.

We present in Fig. 5.11 two cases representative of the time course of both indices. Probably one of the most surprising results from this study is the striking overall similarity between the BIS and the WAV_{CNS}.

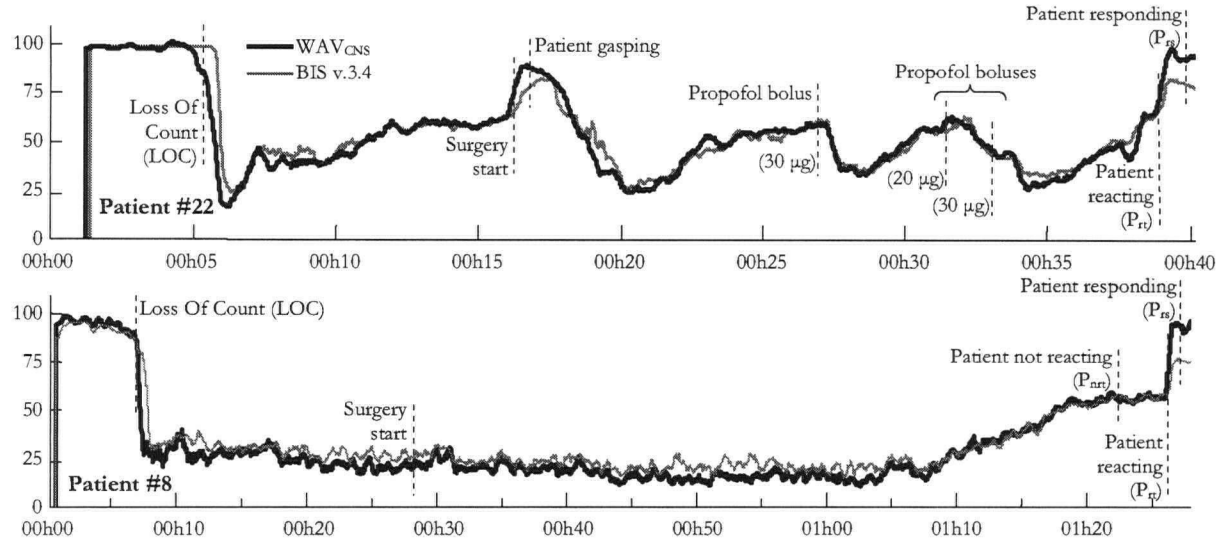


Figure 5.11: Time course of the BIS and WAV$_{CNS}$ for Patient #22 and Patient #8.

Yet, after closer inspection, differences during periods of large transients become apparent. This observation led us to separate steady state behavior from transitory behavior in the comparison of the two indices. Our results and conclusions are the subject of the two following subsections. We also observed that periods of burst suppression patterns are yielding different time courses. We felt that this result could carry strong clinical interest. We therefore present a short discussion concerning the behavior of both BIS and WAV$_{CNS}$ at the end of this section.

5.2.3.1 Steady State Behavior

The BIS index is derived and tuned specifically to reflect the anesthesiologist assessment of the patients' anesthetic depth. The BIS value is nowadays a reference against which other indexes are compared [156, 157, 158]. However, this statement holds true only in steady state, *i.e.*, when the patient's state of consciousness remains constant, since transitory states are affected by the time delay induced by the bispectral analysis [159].

In order to evaluate the accuracy of the WAV$_{CNS}$, epochs of steady state behavior were extracted from each case using a decision algorithm based on the BIS value. A 1-minute sliding window was used to assess the stationarity of the BIS. If the BIS value remained constant within this window (with bounds set at ± 5 of the average BIS value) the epoch was classified as steady-state. Consecutive steady-state epochs were pooled together. The first (or last) 30 seconds corresponding to the beginning (or the end) of each steady-state period were removed under the assumption that they are transitory. A total of 13 hours of recorded data was identified as stationary (73% of the total recording time), representing more than 45,000 EEG epochs.

The WAV$_{CNS}$ *vs.* BIS correlation density plot is presented in Fig. 5.12. There is a strong correlation

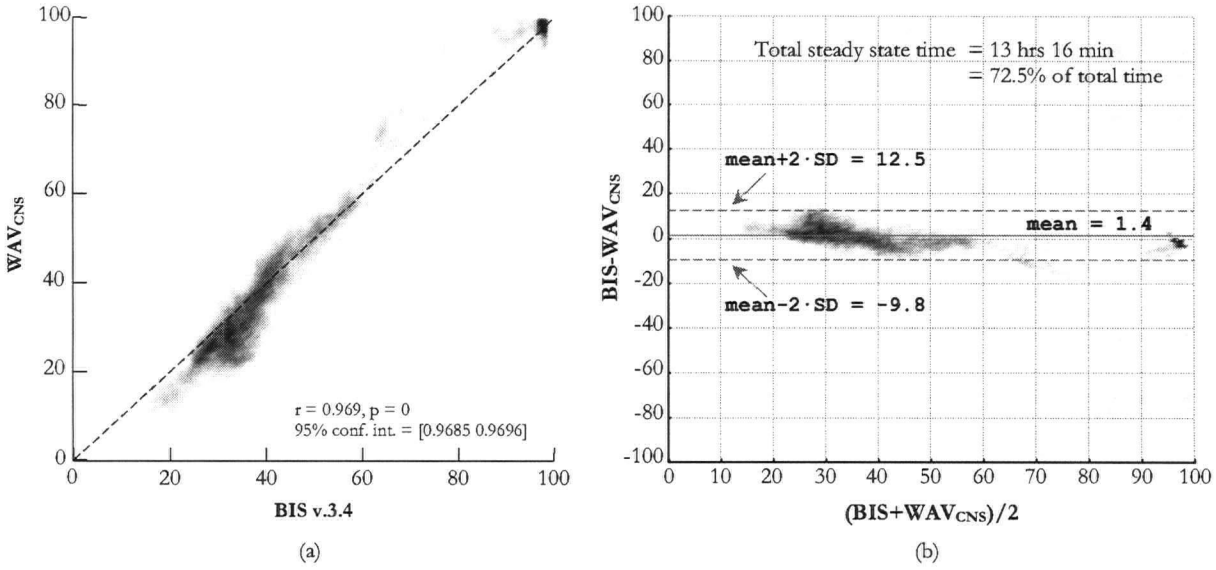


Figure 5.12: Correlation during steady-state operation between the BIS (v.3.4) and WAV_{CNS}. A total of 13 hours of steady-state data were collected from the arthroscopy study. (a) Correlation density plot (b) Bland-Altman test.

between the two values in the 30-60 range. The WAV_{CNS} was slightly higher in the light anesthesia/awake range (60-100). However, most of the data points from this region were obtained from the emergence phase where the BIS value is known to be underestimated [79].

5.2.3.2 Transitory Behavior

The WAV_{CNS} algorithm was designed to exhibit a smooth transitory behavior. In order to evaluate the performance of the WAV_{CNS} during large transients, we observed the time courses of both indexes during induction and emergence.

Induction For this analysis, the study population was separated into two groups depending on the reaction to airway manipulation and LMA insertion. Note that airway management and the insertion of the LMA typically occurred 40 to 80 seconds after LOC. Patients exhibiting no reaction or mild extremity movements were classified in Group 1 ($n = 10$). Patients exhibiting airway reaction (biting, coughing, swallowing, *etc.*) or purposeful motor movements were classified in Group 2 ($n = 9$). One case (patient #15) was excluded due to the use of a neuro-muscular blocking agent during induction (the presence or absence of reaction could not be assessed).

In Fig. 5.13, the time courses of the BIS and WAV_{CNS} are plotted for both groups. Each individual time course is synchronized using the LOC event. In Group 1, the induction is fast and profound. The WAV_{CNS} typically starts decreasing with the LOC event. As compared to the WAV_{CNS}, the BIS value exhibits a more erratic behavior during this transitory phase.

Conversely to Group 1, the induction titration in Group 2 was not adequate to warrant the absence of

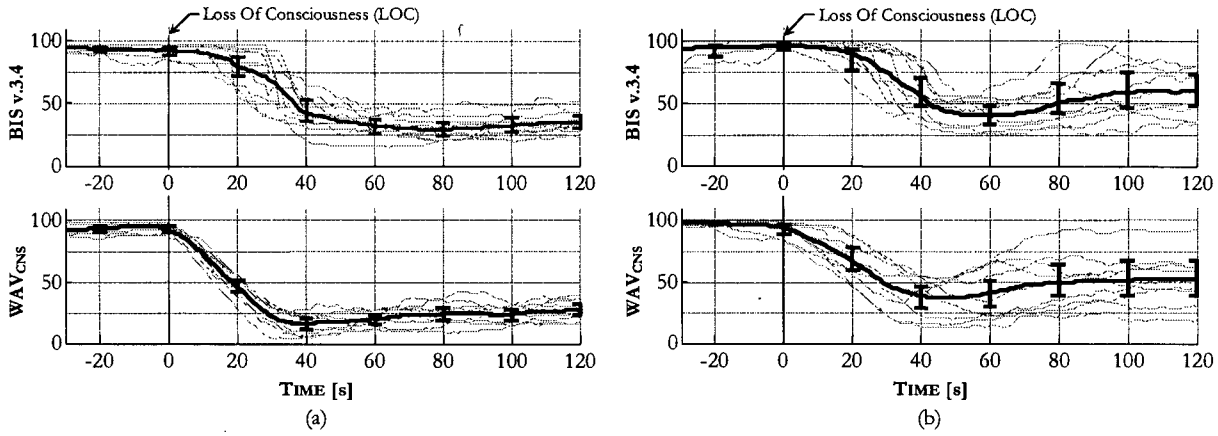


Figure 5.13: Time courses of the BIS (v.3.4) and WAV_{CNS} during induction for both patient groups (each case is synchronized at the LOC). (a) Population showing no reaction to LMA insertion (n=10). (b) Population showing some reaction to the LMA insertion (n=9).

reaction to the insertion of the laryngeal mask. The post-insertion BIS and WAV_{CNS} time courses exhibit more variability than those of Group 1. It is interesting to notice that there is no significant difference in the BIS time course prior to LMA insertion between the two groups. Conversely, the rate of descent and nadir value observed in the WAV_{CNS} shows significant difference between the groups. Thus, it can be hypothesized that the WAV_{CNS} time course during the awake-to-anesthetized transition might be capable of predicting the reaction to LMA insertion. This finding was further corroborated by Lundqvist *et al.* [160] in a study involving 50 patients.

In Fig. 5.14.a, the average time courses of the BIS and WAV_{CNS} are plotted together. We observe a marked delay between the LOC event and the BIS values indicating loss of consciousness (*i.e.*, < 80). The WAV_{CNS} precedes the BIS by an average of 15 seconds in both groups. Note again how the WAV_{CNS} drop is well synchronized with the LOC event.

Emergence Following the clinical protocol, 3 events were recorded during the anesthesia emergence:

- P_{nrt} (Patient Not Reacting): the patient was not reacting to the anesthesiologist's touch and/or voice,
- P_{rt} (Patient Reacting): the patient was reacting (movement, vocalization) to the anesthesiologist's touch and/or voice, or to the surgical environment,
- P_{rs} (Patient Responding): the patient responded to a verbal command.

The P_{rs} event only indicates a time when the patient was able to follow verbal command. However, it is not the *earliest* time when the patient was able to process cognitive thoughts. This is due to the fact that the clinical protocol of this observational study did not make provisions for how frequently the anesthesiologist was to assess the patient's cognitive state during emergence. Therefore, the P_{rs} event cannot be used as a marker of Return Of Consciousness (ROC).

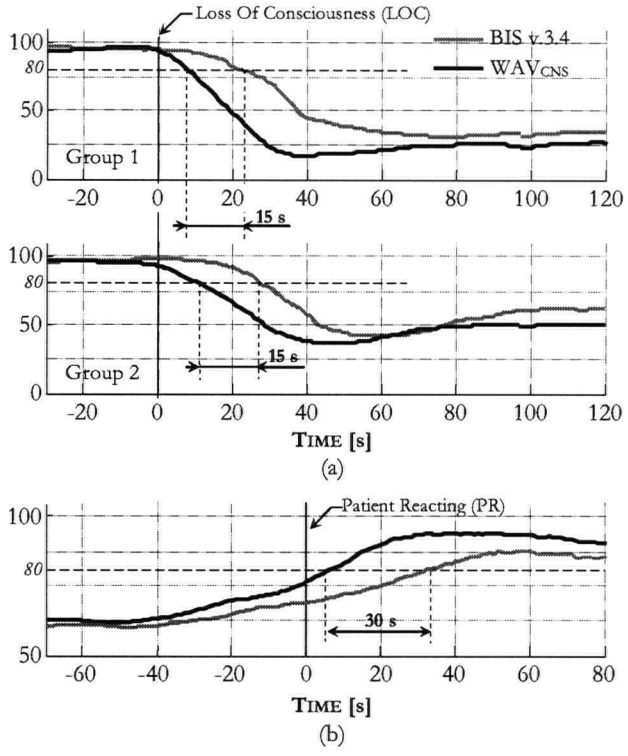


Figure 5.14: Time delay between the BIS and WAV_{CNS}. (a) Induction (both study groups). (b) Emergence.

Conversely, the P_{rt} event indicates the earliest time when a patient reaction was observed. This reaction was triggered either by the anesthesiologist's probing, or by environmental factors. Since the P_{rt} corresponds to a well defined point in time, which is easily observable, it can be used to synchronize the WAV_{CNS} and BIS time courses during emergence, see Fig. 5.14.b. Similarly to induction, we observed a delay in the BIS as compared to the WAV_{CNS} (about 30 seconds in average).

Another interesting difference between the BIS and WAV_{CNS} is the difference of the index value during emergence. In particular, the WAV_{CNS} was 90.6 ($\pm$10.2), while the BIS value remained in the range below 80 (79.0 $\pm$18.8) at the time of the P_{rs} event, *i.e.*, when the patients were responsive. This result is comparable to the results reported by Sleigh *et al.* [79] who found BIS values under 80 while patients were responding to verbal command.

5.2.3.3 Behavior During Burst Suppression

The WAV_{CNS} was derived based on the methodology and data sets presented in Section 5.2.1. The frequency band and wavelet filter were selected to optimize the linearity of the index characteristic, while limiting its variability during periods of steady state. When considering the intermediate states T_1, T_2 and T_3, the results presented in Fig. 5.8 show that the WAV_{CNS} index has a nearly linear characteristic. However, this result does not necessarily hold true for other intermediate states.

In particular, during burst suppression patterns, the EEG alternates between periods of fast, large-

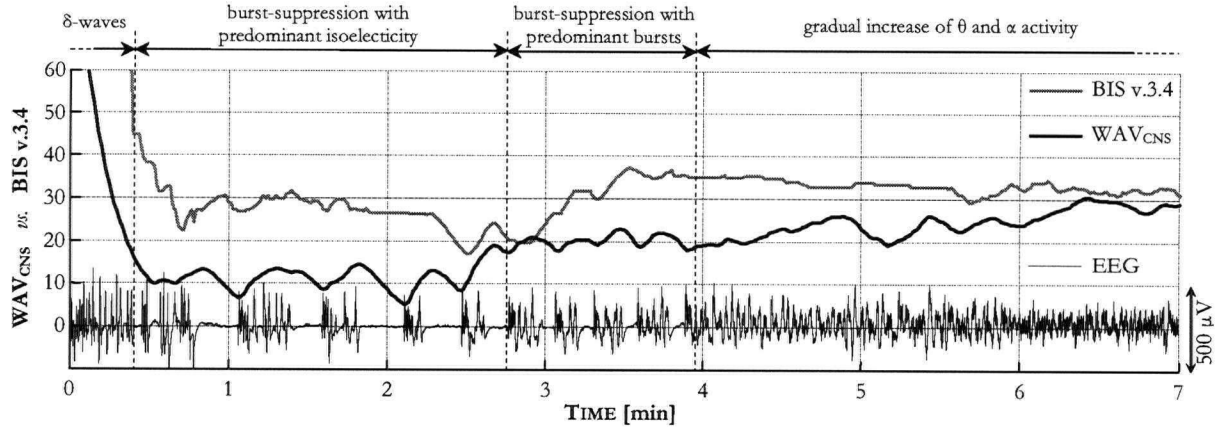


Figure 5.15: Time course of the BIS (v.3.4) and WAV_{CNS} during an episode of burst suppression (patient #4). The corresponding EEG is also plotted for comparison purposes.

amplitude waves, and periods of electrical quiescence. This non-stationarity of the EEG is a well known phenomenon, and has posed a challenge for assessing the patient's cortical state. Questions may therefore arise as to how the WAV_{CNS} behaves during such episodes.

Among the 20 patients of our study, 3 patients exhibited a short period of burst suppression. We present in Fig. 5.15 the time courses of the BIS and WAV_{CNS} during such an episode, as well as the corresponding EEG. In this particular case, burst suppression with prominent isoelectric patterns occurred following a high-dose propofol induction (>5 mg/kg). This EEG pattern lasted 3 to 4 minutes during which the WAV_{CNS} dropped to values of about 10. This level is consistent with the WAV_{CNS} scale since this EEG signal corresponds to an intermediate state between T_3 and R_b.

As the periods of isoelectric EEG got shorter (possibly due to the elimination of the propofol from the blood plasma), the WAV_{CNS} increased up to values of about 20. A gradual increase in θ and α activity 5 minutes into this case provoked a further increase in the WAV_{CNS}. Once steady state was reached ($t > 6min$), both the WAV_{CNS} and BIS time courses converged towards similar levels.

During periods of burst suppression patterns, the WAV_{CNS} may exhibit an oscillatory behavior corresponding to intermittent periods of bursts and electrical quiescence. In particular, we can observe an increase in WAV_{CNS} during EEG bursts, and a decrease during periods of isoelectricity. This is consistent with the implementation of the index described in Section 5.2.2.

The BIS time course does not seem to follow the changes in the EEG signal during periods of burst suppression in the similar way as the WAV_{CNS}. This difference stems from the fact that the BIS monitor uses a different signal processing algorithm to deal with the non-stationarity of the EEG in this state. Conversely to the BIS, the WAV_{CNS} algorithm is unique across the whole 100-0 scale, *i.e.*, EEG epochs are processed similarly whether the patient is awake or experiences periods of bursts. We therefore expect the WAV_{CNS} to differ from the BIS during burst suppression. However, note that further studies are needed to provide more clinical insight into the behavior of the WAV_{CNS} during this particular state.

5.3 Measuring the Autonomic Nervous System (ANS) State: the WAV_{ANS} Index

Anesthesiology is commonly regarded as "the practice of autonomic medicine" [Lawson:2001], where the greater part of the anesthesiologist's expertise relies on his or her ability to control autonomic functions through the administration of drugs. Yet, there is currently no available monitor of the Autonomic Nervous System (ANS). Appropriate therapeutic actions are devised through the observation and interpretation of trends in patients' vital signs. The goal of this section is to close this gap by developing a feedback sensor for the direct assessment of the ANS state. This measure can then be used as a surrogate measure of the analgesia component of anesthesia.

In particular, it is desired to obtain a bounded dimensionless index which would increase during autonomic activation (leading to heart rate and blood pressure increase) and decrease during autonomic depression (leading to heart rate and blood pressure decrease). In anesthetized patients, autonomic activation can be obtained through noxious surgical stimulation and the use of anti-cholinergic drugs such as atropine. Conversely, the depression of the autonomic system can be achieved through the use of analgesic drugs, vapour anesthetics and adrenergic drugs such as β-blockers.

To derive this measure, we propose to use the HRV signal, and use the same signal processing technique as used previously for the derivation of the WAV_{CNS}.

We present results in the form of 3 case reports obtained from data gathered from clinical audits at the University of British Columbia Hospital.

5.3.1 The Use of Heart Rate Variability

The HRV signal is derived directly from the ECG signal. The HRV is based on the R-R interval between two consecutive beats, see Figure 5.16.a. The last two decades have witnessed the recognition of the HRV as a promising quantitative markers of autonomic activity. The apparently easy derivation of this measure has popularized its use; from being an independent predictor of mortality following myocardial infarction, to detecting autonomic neuropathy in diabetic patients [161].

In terms of the quantification of the autonomic state, one particularly interesting aspect of the HRV signal is that its high frequency (HF) content is associated with vagal tone. When patients are under acute stress, such as surgical noxious stimulation, the vagal tone tends to be depressed, which is characterized by a decrease in HF activity. As a patient's body experiences increasing level of pain, beat-to-beat intervals tend to become more regular (see for instance Figure 5.16.b for illustration). Conversely, during periods of relaxation, consecutive beat-to-beat intervals are more irregular.

In addition to the effect of pain and noxious stimulation, there are a number of factors also known to affect the HRV high frequency content:

- Respiration (or Respiratory Sinus Arrhythmia - RSA): inspiration tends to increase the heart rate

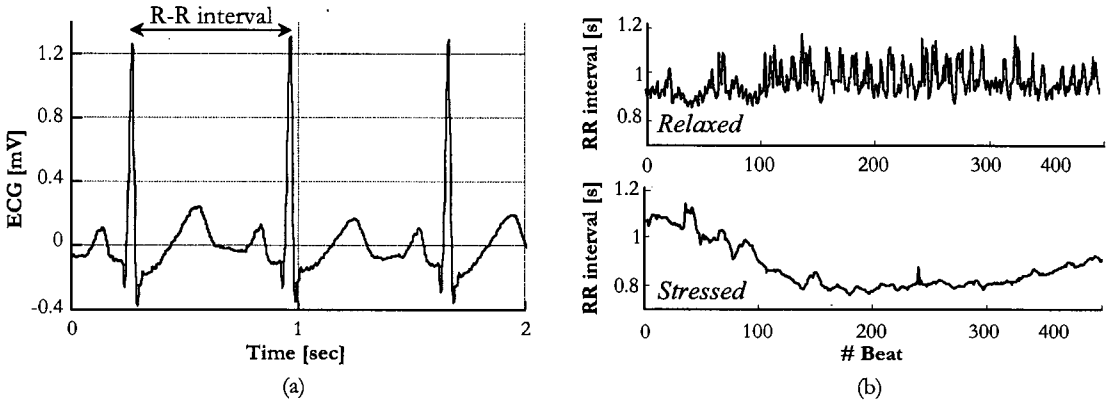


Figure 5.16: The HRV signal as a measure of autonomic activity. (a) 3 consecutive QRS complex (the HRV is obtained based on the detection of the R-wave and the measurement of the R-R interval between two consecutive R-waves). (b) Tachogram of a subject performing a relaxation exercise, and a patient reacting to surgical stimulation.

while expiration tends to slow it down. As a result, the HRV exhibits an oscillatory behavior following the respiration pattern. Hence, changes in the respiratory pattern strongly influence the HF content of HRV and must be accounted for.

- Opioids and inhalation agents: these agents are known analgesics. As such, they reduce the generation and transmission of pain signals and increase the HF of HRV.

- β-blocking drugs: these drugs are essentially used to depress the sympathetic system in order to prevent hypertension and tachycardia. However, evidence from the literature also support that the administration of β-blocking drugs is associated with an in-crease in the HF content of the HRV [162].

- Anti-cholinergic drugs: drugs such as atropine antagonize the action of acetylcholine. This results in vagal blockade and the disappearing of HF in the HRV.

The effect of pain, analgesic drugs, β-adrenergic and anti-cholinergic agents all follow the same trend: pain and stress decreases the HF (associated with a rise in blood pressure and heart rate). Conversely, analgesics and β-blockers increase the HF (associated with a decrease in blood pressure and heart rate). Therefore, the quantification of the HF activity of the HRV should provide a reliable metric of the effect of pain and analgesic/adrenergic drugs on the autonomic system.

The fact that HRV analysis can yield a usable 'depth of anesthesia' index was already known in 1985. Unfortunately, the very nature of the HRV signal makes it particularly difficult to obtain a usable index while achieving real-time operation. There are a number of difficulties related to the acquisition and processing of the HRV signal:

- ECG temporal resolution: the variability in terms of the R-R interval during periods of surgical stress can be as low as few milliseconds. As a result, the HRV should be derived based on an ECG signal

sampled at a high sampling rate (>500 samples per second) in order to gain good temporal resolution of the R-peak [161].

- Detection of the R-peak: the task of detecting the R-peak of the QRS complex is especially difficult during periods of artifactual activity. Electrosurgical Units (ESU) are the most significant source of interference in the operating room and completely obliterate the ECG [163]. This poses a severe obstacle for real-time intra-operative monitoring, since there is usually significant data loss during long periods of time.

- Slow signal: probably the most damaging aspect of the HRV signal is the slow update rate linked to the heart beat. Any real-time analysis must be optimized in order to provide measurements which minimally lag relevant physiological events. In today's current state of the art, existing methods of 'short-time' HRV analysis usually require at least 5 minutes recordings to provide reliable information [161]. This type of performance is not suitable for a real-time application such as the one envisaged here. A part of the problem lies in the fact that neither time nor spectral analyses are suitable for non-stationary signals such as HRV. We believe that this difficulty can be overcome by using Wavelet Transform.

5.3.2 The WAV_{ANS}

Similarly to the WAV_{CNS} index, the goal is to analyze the HRV signal in order to derive a metric, referred to as WAV_{ANS}, representative of the patient autonomic state. This metric should represent the autonomic state of the patient, and be bounded between 2 extreme states: relaxed (*i.e.*, no pain and comfortable) and extreme stress (*i.e.*, high level of pain triggering a significant autonomous reaction).

The WAV_{ANS} index is obtained by applying to the HRV signal a methodology similar to that of the WAV_{CNS}. A feature extracted from the HRV signal is compared to 2 reference features corresponding to the 2 extreme autonomic states. The WAV_{ANS} is then the combination of the likelihood of the patient being relaxed, and the likelihood of the patient being under large surgical stress.

To find the wavelet filter and frequency band which would carry information relative to the patient's autonomic state, we obtained 5 HRV data sets correspond-ing to different levels of stress, see Figure 5. The sets T_2 and T_3 were obtained from 2 clinical patients during periods of surgical stress (T_2) and intense surgical stress leading to a significant autonomic response (T_3). The sets R_a and T_1 were obtained from a volunteer performing a breathing relaxation exercise (R_a, see Figure 5) and during a cold pressure test (*i.e.*, moderate pain, T_2). Finally, the set R_b is a simple constant value representing the absence of vagal activity. In all cases, patients and subjects were breathing at 8 breaths per minute (a metronome was used to guide the breathing pattern of the volunteer).

A preliminary analysis based on these data revealed that the wavelet filter Daubechies #2 and the frequency band 0.2-0.4 Hz are optimal to discriminate between these 5 HRV signals (this correspond the $\mathbf{d}_1^{\psi}$

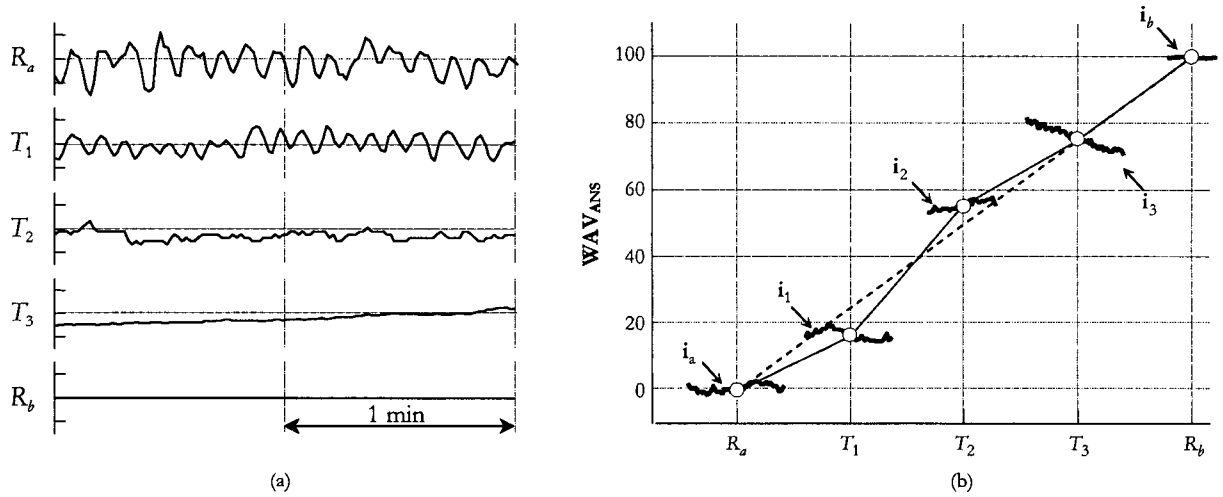


Figure 5.17: Application of the WAV technology to the quantification of autonomic activity using the HRV signal. (a) Representative data sets of HRV signals for different stress levels: R_a: relaxed subject (y-scale: 0.2 s/div), T_1: cold pressure test on volunteer (y-scale: 0.2 s/div), T_2: clinical patient undergoing surgical stimulation (y-scale: 0.1 s/div), T_3: clinical patient reacting to surgical stress (y-scale: 0.1 s/div). R_b: no vagal tone (i.e., no heart rate variability - theoretical state). (b) Result of the best wavelet and frequency band selection. Note that no post-processing filtering is required due to a long analysis epoch.

high frequency band). The resulting WAV_{ANS} metrics is expressed in the 0-100% scale, where 0% represents deep relaxation, and 100% represents the total lack of parasympathetic activity.

5.3.3 Implementation Issues

Even though the WAV_{ANS} is based on similar principles than the WAV_{CNS}, the nature of the HRV differs considerably to that of the EEG. As a result, the implementation of the WAV_{ANS} needs to account for the various specificities of the HRV signal.

HRV Signal Acquisition New HRV samples are only available at each heart beat. The R-peak of the ECG is used as a triggering event to measure the beat-to-beat interval. The temporal resolution of the R-peak is therefore particularly important to guarantee an adequate HRV resolution. In the current implementation, we use an ECG sampling rate of 960 S/s, which results in a temporal resolution close to 1 ms.

A spike detection algorithm detects the peak of the R-wave. Artifacts are detected by analyzing the ECG signal between two consecutive R-peaks. If the beat-to-beat waveform differs considerably from a reference waveform, or if the beat-to-beat interval is much smaller or larger than previous measurements, an artifact flag indicating that the current R-R interval may be corrupted is activated. This flag is used in the processing algorithm to limit the effect of corrupted HRV samples on the WAV_{ANS}.

Data flow In the current implementation, the HRV is sampled at the same sampling frequency than the ECG signal. A zero-order-hold is used to hold the R-R interval value between each heart beat. The signal is then resampled at a rate of 0.8 samples per second using a resampling filter bank. This choice of sampling frequency is motivated by the results found in Section 5.3.2.

As compared to the WAV_{CNS} implementation, the WAV_{ANS} is updated at every new HRV sample. This results in an update rate of 1.25 second. Also, the SWT decomposition is calculated based on the last 2-minute of HRV data (96 samples) in order to have enough wavelet coefficients to derive a meaningful PDF.

Note that, due to the large size of the analysis window (2 minutes), there is no need for further post-processing filtering.

Artifacts Dealing with corrupting artifacts is also challenging.

For EEG analysis, epochs containing samples corrupted by artifacts can be rejected without impairing significantly the availability and reliability of the WAV_{CNS}. A similar strategy is not applicable for HRV analysis. It would indeed result in a large amount of data loss during, *e.g.*, periods of electrocautery activity[1], as it would be necessary to wait at least a complete HRV period (in this case 2 minutes) to obtain the next artifact-free HRV epoch. As a result, even HRV epochs containing corrupted samples are used in the WAV_{ANS} calculation.

It is therefore necessary to limit the effect of artifactual HRV samples on the WAV_{ANS} calculation. In order to do so, wavelet coefficients obtained from corrupted samples must be removed from the PDF calculation. Because of the convolution operation, each corrupted HRV sample results in the removal of l wavelet coefficients, where l is the length of the high-pass wavelet filter. It is therefore particularly interesting to use low order wavelet filters to limit the effect of corrupted samples due to the convolution operation. In the current implementation, we use a 4^{th} order FIR wavelet filter (Daubechies #2). Note that, if there are less than 32 valid wavelet coefficients left to carry out the PDF calculation, the HRV epoch is discarded. The WAV_{ANS} value is then hold till the next cycle. If more than 12 cycles (corresponding to 15 seconds) have elapsed since the last WAV_{ANS} update, the monitor sends a warning message to the user and stop displaying the index.

5.3.4 Case Reports

Three female patients undergoing hysterectomy procedures were induced with a combination of propofol and remifentanil (Case #1), or propofol and fentanyl (Case #2 and #3). After securing the airway, maintenance was insured by a background inhalational anesthetic supplemented by opioid boluses. In all cases, the patients were maintained in a hypocapneic state by mechanical ventilation (8 breaths/minute). Aside from the WAV_{ANS}, all the data presented here were obtained on-line. The WAV_{CNS} was obtained

[1]Interferences due to electrosurgical units have been identified as the most perturbing operating room noise for ECG acquisition devices [163].

directly from a 2-channel frontal electroencephalographic montage. The WAV_{ANS} was calculated off-line based on a 3-lead ECG signal. Intra-operative events and the anesthesiologist's assessment of the patients' state were logged in by an independent observer. In Case #2 and #3, the observer also recorded manually the heart rate, blood pressure (each time this measurement became available), the MAC value and the anesthetic titration.

Case #1 (see Figure 5.18) Following the initial propofol/remifentanil induction bolus, the patient's ANS state reached a low 'relaxed' level. Upon the start of surgery (multiple skin incisions), the patient's heart rate and blood pressure in-creased significantly. This increase was mirrored by a significant increase in the WAV_{ANS} index. At 0h35, the patient was deemed 'light' by the anesthesiologist. Similarly, the WAV_{ANS} reached a high value (>60). At 0h45, a large bolus of fentanyl was administered, provoking a marked drop in the WAV_{ANS} index. In 2 other instances, the anesthesiologist estimated that the patient was light (increase in heart rate and blood pressure). In these 2 instances, the WAV_{ANS} also reached higher levels. From this case, it appears that a WAV_{ANS} level of 70 (and above) is associated with a higher incidence of significant autonomic reaction to surgical stimulation. The large drop during emergence (2h35) can be due to a change in respiration pattern, as neuro-muscular blockade was reversed to allow extubation (the patient was spontaneously breathing at 2h30). Note that the WAV_{CNS} reached a very low level (<30) during most of the procedure, thus indicating a rather deep hypnotic state. Following the indication of both CNS and ANS monitors, it appears that the sedation *vs.* analgesia balance was tilted towards the sedation endpoint. *Ad hoc* analysis of this case indicates that increasing the analgesia titration while decreasing hypnotic dosage would have been beneficial to this patient. The benefit of dual CNS and ANS monitoring is evident in this case.

Case #2 (see Figure 5.19) The WAV_{ANS} peaked 5 times during this case. The first 2 peaks (0h30 and 0h35) correspond to the most stimulating part of the surgery involving large skin incisions. Blood pressure and heart rate increased in response to the stimulation which prompted the anesthesiologist to give a fentanyl top-up dose. The next 3 WAV_{ANS} peaks (0h55, 1h10, 1h20) were also characterized by an increase in heart rate and were correlated to short episodes of electrocautery activity. In all instances, the duration of autonomic activation was very short (2-3 minutes) and may not have been captured by the blood pressure measurement (updated only every 5 minutes). Finally, at the time of 'adequate anesthesia' (as per the anesthesiologist's assessment), the WAV_{ANS} was less than 50.

Case #3 (see Figure 5.20) As compared to the previous 2 cases, the WAV_{ANS} was, in average, less than 50 during most of the surgery. Similarly, there was no indication of a 'light' analgesic state during the whole procedure. Both the heart rate and blood pressure were stable. The absence of peak in the WAV_{ANS}, or sudden increase in heart rate and blood pressure can be explained by the fact that this surgery is not as invasive as in the previous cases (limited skin incisions). Interestingly, the increase in WAV_{ANS} towards

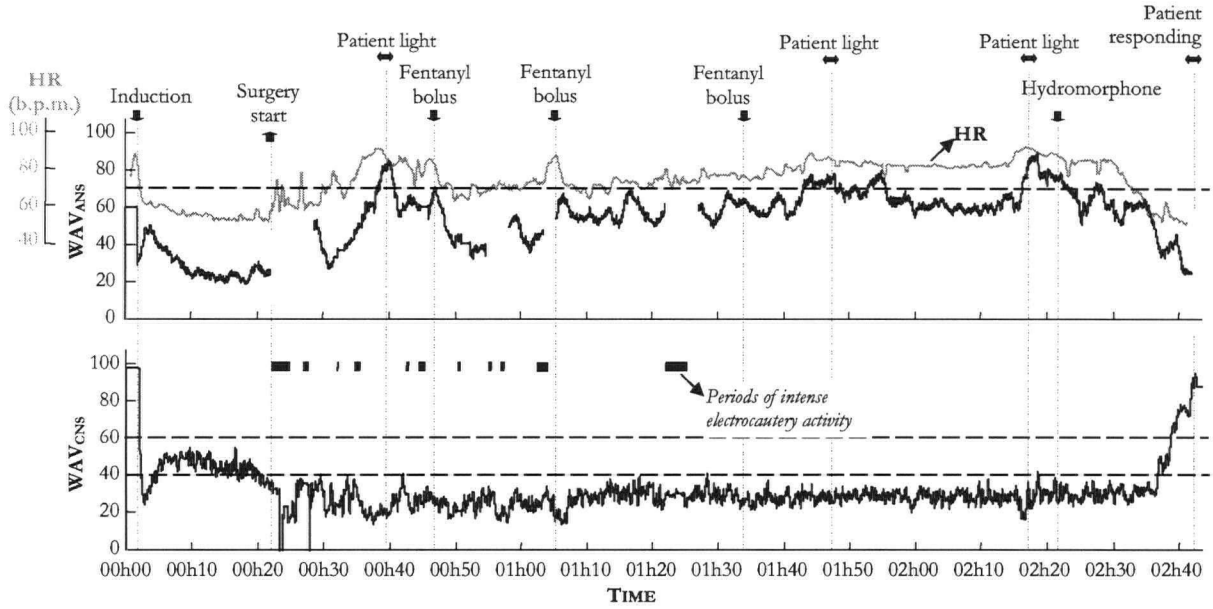


Figure 5.18: Case report #1 - Female patient, 55 years old, undergoing an hysterectomy procedure. The large gaps in WAV_{ANS} index correspond to periods of time heavily corrupted by electrocautery artifacts (the HRV signal could not be obtained from the ECG signal). Electrocautery also affected the EEG signal and created data loss in the WAV_{CNS} (indicated by vertical bars in the hypnosis index). The 'patient light' events correspond to the anesthesiologist's assessment of the patient during the surgery. Periods of intense electrocautery activity are indicated (it is assumed that they coincide with surgical noxious stimulation).

the end of the case mirrors the decrease in MAC value. This is consistent with the fact that MAC is a measure of the analgesic potency of the administered vapour. Upon emergence, the patient reacted to the anesthesiologist's voice, but was not able to respond to verbal commands.

These 3 cases illustrate how surgical stimulation triggers abrupt changes in heart rate and blood pressure. These abrupt changes are used by anesthesiologists to identify the presence of excessive autonomic activation, and then to proceed to take appropriate pharmacological action. Conversely, the WAV_{ANS} is a bounded and population-normed metric. As such, it can be used as a predictor of autonomic activation. For instance, in Case #1, the trend of the index shows that the patient's state is getting increasingly lighter (from 0h50 to 1h40), up to the point where the patient did actually react to the surgical act. Also, the index was quite high during the second half of the surgery, indicating the need to deepen the analgesic state. Conversely, in Case #3, the index was quite low and steady throughout most of the surgery. This is consistent with the rather aggressive sevoflurane titration (MAC>1.0) in a surgery involving only minor skin incisions and moderate stimulation. This may also explain the prolonged emergence. Finally in Case #2, the number of WAV_{ANS} 'spikes' and their rather short duration may indicate an overall shallow analgesic state that resulted in strong and acute autonomic activation whenever surgical stimulation was present.

These preliminary results are very encouraging. It appears that the WAV_{ANS} has the potential to

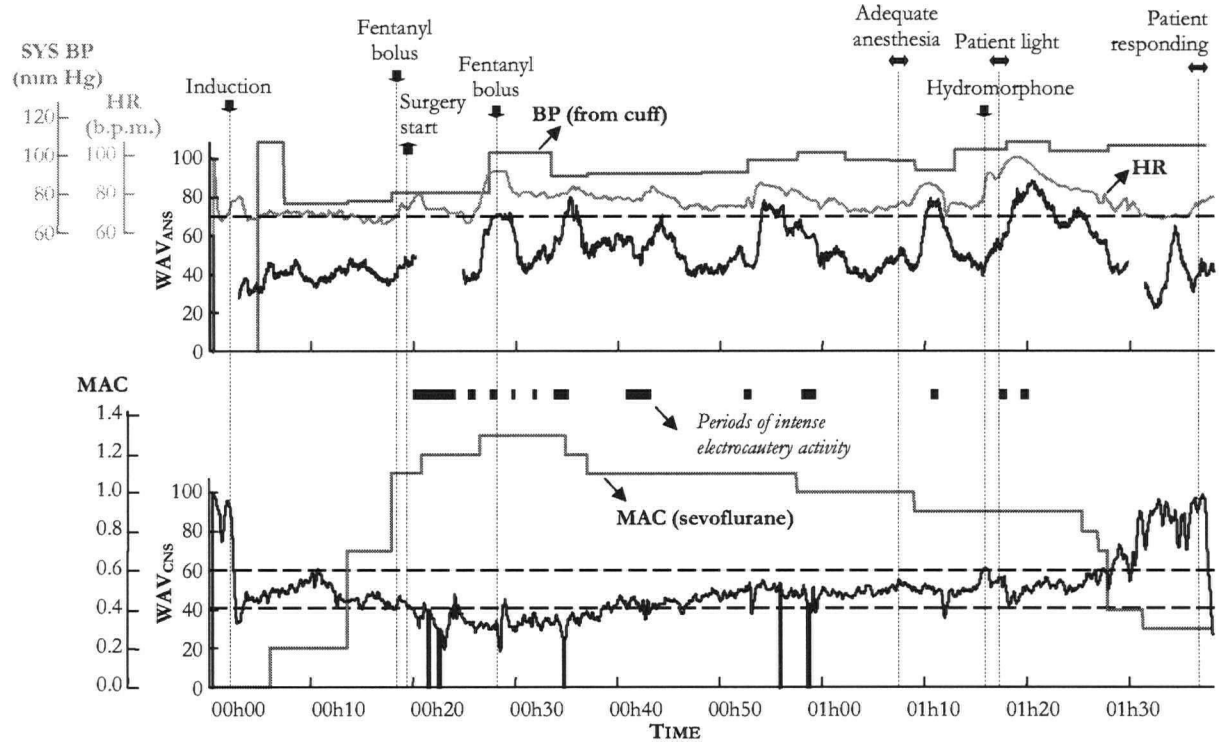


Figure 5.19: Case report #2 - Female patient, 39 years old, undergoing a hysterectomy procedure. The heart rate (top graph) was calculated ad hoc based on the EKG. The blood pressure from the cuff sensor was recorded manually each time this measurement was updated by the anesthesia monitor.

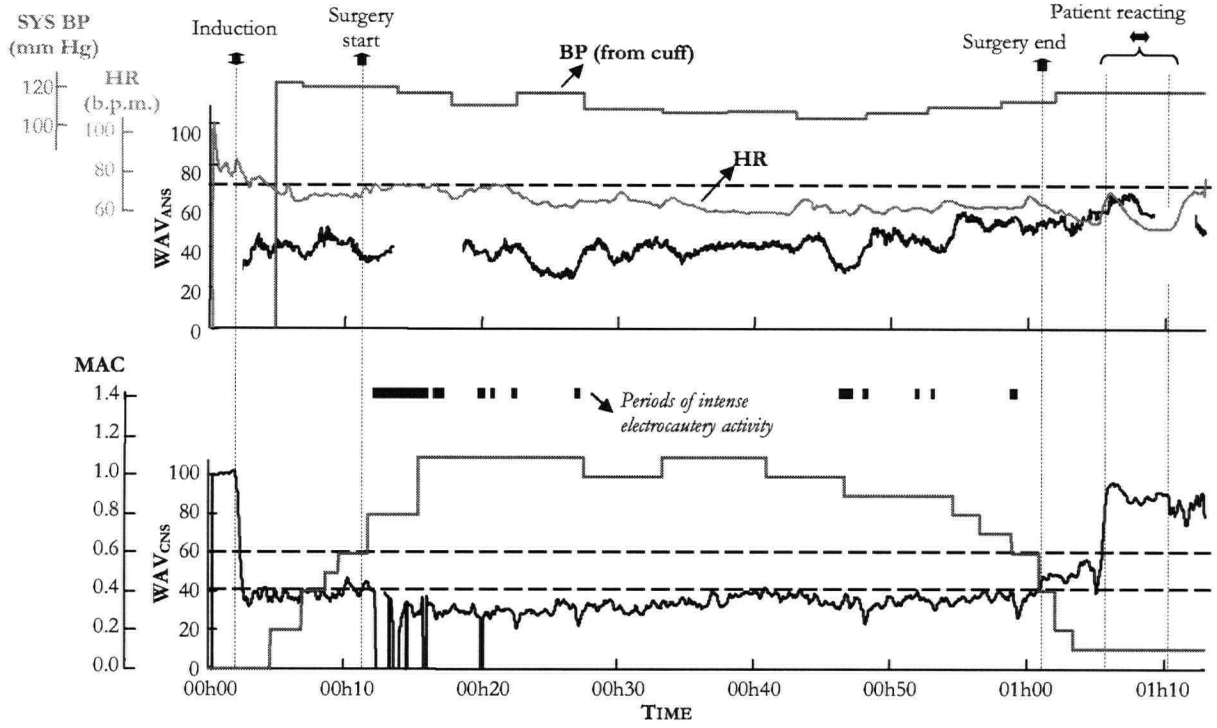


Figure 5.20: Case report #3 - Female patient, 38 years old, undergoing a laparotomy procedure. Due to the nature of the surgery, skin incision was kept to a minimum.

provide a guide for analgesic titration and guarantee patients' safety from a hemodynamic standpoint. It is important to point out, however, that this study suffers from a number of limitations:

- First, the WAV_{ANS} was calculated off-line. Artifactual periods were visually reviewed in order to facilitate the detection of the R-peak and limit the amount of data loss. Unfortunately, the limitation of the front end amplifier in attenuating the ESU interference resulted in large gaps in the HRV signal, which precluded the calculation of the WAV_{ANS} for long periods of time during the most important part of the surgery.

- Also, the 3 patients were mechanically ventilated, *i.e.*, the effect of a change in the respiration pattern cannot be inferred from this data. The analysis algorithm was specifically tuned for 8 breaths a minute, so a different respiratory rate may have resulted in different WAV_{ANS} levels (*e.g.*, the abrupt WAV_{ANS} drop at the end of Case #1 can be attributed to an erratic and rapid breathing pattern).

- Finally, this audit was simply an observational study, *i.e.*, the administration of opioids and change in gas concentration often coincided with surgical stimulation. It is therefore difficult to identify the contribution of noxious stimuli from the effect of opioids from the analysis of the WAV_{ANS}.

It is evident that more clinical data obtained under precisely defined clinical protocols are needed in order to further develop the proposed WAV_{ANS} index and assess its robustness in a clinical setting. In particular, it

will be important to clearly demonstrate the existence of a dose/response relationship between the WAV_{ANS} and drug dosage.

5.4 Dynamic Behavior

We conclude this Chapter by deriving deriving transfer functions for both indices in order to model their dynamic behavior in response to changes in the patient's state.

WAV_{CNS} Transfer Function The elements dictating the WAV_{CNS} dynamic behavior are the second order IIR low pass filter used to extract the main index trend (characterized by a cutoff frequency ω_0 and a damping factor ζ), and the 1-second update rate (zero-order-hold element). As a result, the sensor dynamics can be represented by the following LTI transfer function:

$$H_{\text{CNS}}(s) = \frac{\omega_0^2}{s^2 + 2 \cdot \zeta \cdot \omega_0 \cdot s + \omega_0^2} \cdot \frac{1 - \mathrm{e}^{-s}}{s} \tag{5.24}$$

With the current WAV_{CNS} implementation, $\zeta = 1$ and $\omega_0 = 0.02$ Hz:

$$H_{\text{CNS}}(s) = \frac{1}{(8 \cdot s + 1)^2} \cdot \frac{1 - \mathrm{e}^{-s}}{s} \tag{5.25}$$

If we assume no *a priori* knowledge about the WAV_{CNS} sensor dynamics, an identification procedure can be carried out using a database of various EEG signals corresponding to known sensor values. To illustrate this procedure, we consider the 5 data sets used for the derivation of the WAV_{CNS} in Section 5.2.1.3. Each of these data sets corresponds to a stationary hypnotic depth:

$$\overline{\text{WAV}_{\text{CNS}}} : \{R_a, T_1, T_2, T_3, R_b\} \rightarrow \{98, 78, 56, 19, 0\} \quad [\%] \tag{5.26}$$

A composite EEG signal is first obtained by arbitrarily selecting EEG epochs from each of the 5 data sets, see Fig. 5.21. An identification input signal is then obtained by attributing to each of the EEG epoch the corresponding WAV_{CNS} value according to (5.26). Finally, the composite EEG signal is processed by the sensor algorithm to yield an output identification signal. The sensor dynamic $H_{\text{CNS}}(s)$ is further identified by a least square estimation algorithm. In this case, the identification procedure resulted in a second order transfer function plus delay:

$$H_{\text{CNS}}(s) = \frac{0.36 \cdot s + 1}{(7.2 \cdot s + 1) \cdot (8.7 \cdot s + 1)} \cdot \mathrm{e}^{-s} \tag{5.27}$$

The difference between the identified (5.27) and the analytic (5.25) functions is negligible in the low frequency range, see Fig. 5.22. Differences are more apparent for higher frequencies. This discrepancy results from the limited input identification signal bandwidth (<1 rad$\cdot s^{-1}$ in this case) and the identification method whose accuracy can not be guaranteed near the 1 Hz sampling frequency. Since it is unlikely that the controller cross-over frequency be placed in this frequency range, no further refinements are necessary.

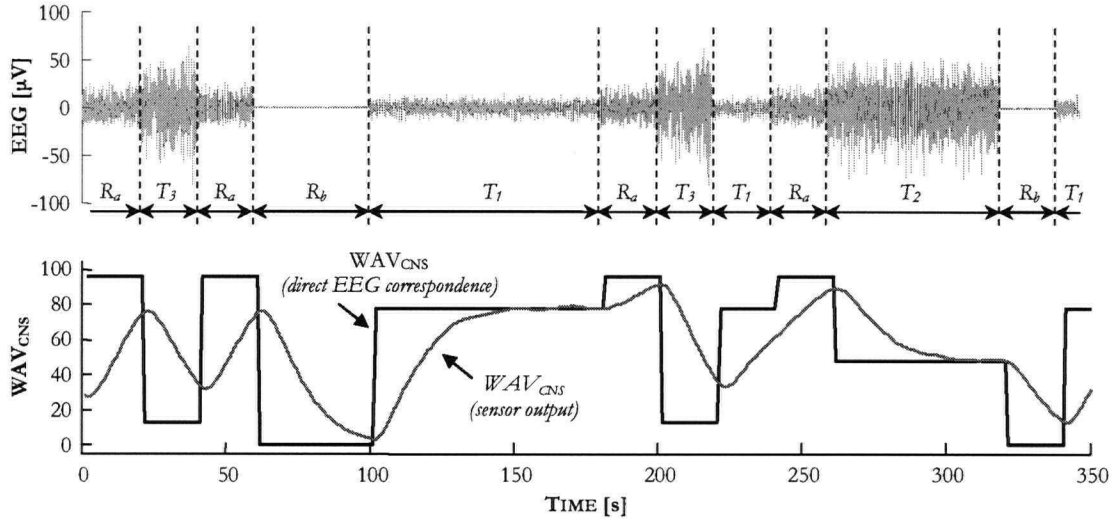


Figure 5.21: Experiment design for identifying the sensor dynamics. The EEG signal is composed based on EEG epochs arbitrarily chosen from a database of signals. The input identification signal corresponds to the simulated patient's instantaneous state (direct correspondence with the source EEG signal), while the output identification signal is the WAV_{CNS}.

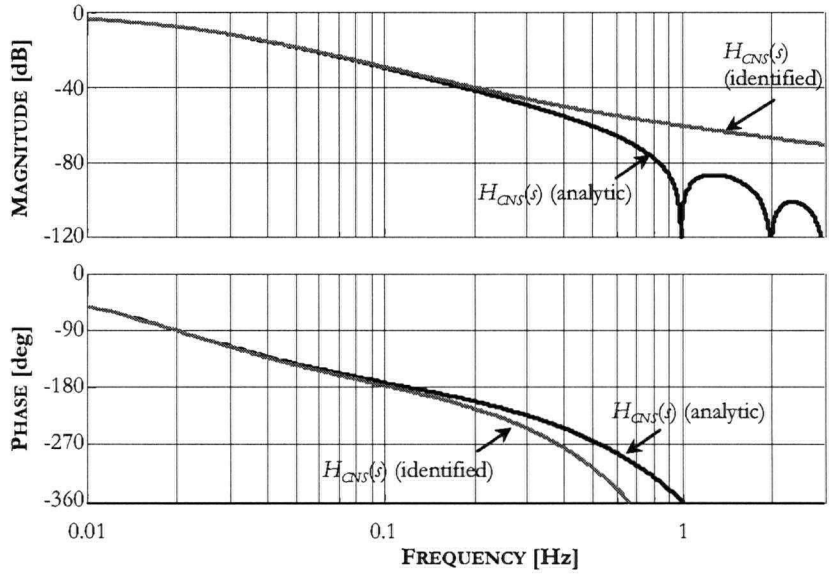


Figure 5.22: Bode plots of the sensor dynamic (analytic and identified models).

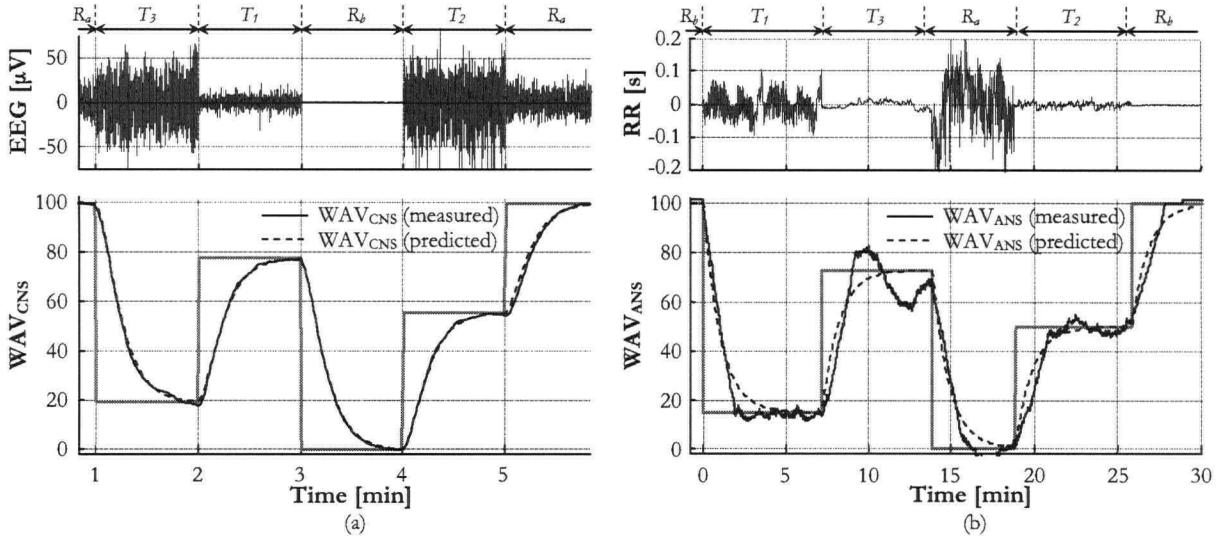


Figure 5.23: Accuracy of H_{CNS} and H_{ANS} for predicting the $\mathrm{WAV_{CNS}}$ and $\mathrm{WAV_{ANS}}$ time courses calculated for various step changes. (a) The H_{CNS} is able to predict accurately the $\mathrm{WAV_{CNS}}$ time course. (b) The H_{ANS} is an approximation of the index dynamics. As such, some discrepancies between measured and predicted outputs exist. Also, there is a noticeable index variation in T_3. This can be the result of the fact that the patient state was not as stable as in the other states.)

$\mathrm{WAV_{ANS}}$ Transfer Function Following the proposed implementation, the $\mathrm{WAV_{ANS}}$ dynamic is equivalent to a 2 minutes moving average due to the selected length of the HRV period upon which the analysis is carried out. Considering that the index is updated every 1.25 seconds, the $\mathrm{WAV_{ANS}}$ transfer function can be expressed as:

$$H_{\mathrm{ANS}}(s) = \frac{1}{96} \cdot \sum_{k=0}^{95} \mathrm{e}^{-1.25 \cdot k \cdot s} \frac{1 - \mathrm{e}^{-1.25 \cdot s}}{s} \tag{5.28}$$

The moving average filter can further be approximated as a first order transfer function whose time constant is one half the moving average window:

$$H_{\mathrm{ANS}}(s) = \frac{1}{(60 \cdot s + 1)} \cdot \frac{1 - \mathrm{e}^{-1.25 \cdot s}}{s} \tag{5.29}$$

The performance of H_{CNS} and H_{ANS} was assessed by comparing the time course of the indexes *vs.* their predicted time courses, see Figure 5.23. This test was done based on source signals (EEG or RR) obtained from a combination of the initial data sets.

The H_{CNS} was found to be remarkably accurate in predicting the $\mathrm{WAV_{CNS}}$ time course. Conversely, the transfer function H_{ANS} is only a coarse approximation of the $\mathrm{WAV_{ANS}}$ dynamics. This result can be improved by reducing the $\mathrm{WAV_{ANS}}$ analysis window and adding a post processing IIR trending filter. However, this would reduce the number of wavelet coefficients, which might in turn decrease the resolution of the PDF waveform and increase the measurement noise, while making the index more sensitive to artifacts.

5.5 Summary

The search for anesthesia feedback sensors has motivated the development of a methodology to analyze and characterize non-stationary physiological signals. This methodology, referred to as WAV, uses wavelet decomposition to capture rapid signal changes in both time and frequency. The statistical representation of the wavelet coefficients using probability density functions allowed us to derive bounded metrics, thus quantifying the signal into a pre-defined scale. The flexibility of wavelets can be used as an advantage to tune the static characteristic of the index in order to reflect the anesthesiologist's assessment.

We applied this methodology to the analysis of EEG signals in an attempt to quantify patients' cortical activity. The resulting index, the WAV_{CNS}, has demonstrated superior dynamic performance as compared to today's leading monitoring technology. Also, to our knowledge, the WAV_{CNS} is the only index of cortical activity whose dynamics are linear and can be expressed as an LTI transfer function. As such, it is an attractive solution for use as a feedback sensor for advisory and close loop systems, and to identify the pharmacodynamic models of anesthetic drugs.

The WAV methodology has also demonstrated strong potential for use in the analysis and characterization of the HRV signal. The case reports presented here show that the WAV_{ANS} was representative of the anesthesiologist's assessment of the patient's analgesic state. In particular, all intra-operative events leading to a significant change in heart rate and blood pressure were well correlated with an elevated WAV_{ANS} value (for instance, a WAV_{ANS} level of 70 was associated with a higher incidence of significant blood pressure and heart rate change). In addition, the administration of opioids (fentanyl, hydromorphone) was followed by a decrease in the WAV_{ANS} index. Similarly, changes in MAC value seemed to be well correlated with changes in WAV_{ANS}.

It is likely that other applications can benefit from the WAV technology, *e.g.* automatic sleep scoring, depression and Alzheimer's decease diagnosing, etc... As of today, our Group has invested considerable time and efforts into the development and validation of the WAV_{CNS}. The intellectual property is now protected [164], and a commercial device integrating this technology should become available in mid-2006 (NeuroSENSE^{TM} Monitor, Cleveland Medical Devices Inc., OH).

Chapter 6

A System Oriented Approach to Pharmacodynamic modeling

We exposed in details in Chapter 3 the approach to pharmacokinetic (PK) and pharmacodynamic (PD) modeling favored by pharmacologists. In terms of pharmacokinetics, the models proposed in the literature are consistent. The use of the NONMEM analysis software has enabled clinicians to improve the fit of PK models by including patients' biometric data (weight, lean body mass, age, *etc.*). The inclusion of PK models within commercial TCI systems are a testimony to their reliability and acceptance from the clinical community.

In terms of pharmacodynamics, the situation is different. The latest PD studies carried out for intravenous anesthetic drugs have yielded contradictory results and large inter-individual variability in some of the parameters. Furthermore, PD models in the literature were derived specifically for use with one particular sensor, and thus, cannot be readily used to predict the time course of an effect measured by different sensor. It is our belief that the traditional modeling approach deserves to be revisited from a system engineering perspective.

In this chapter, we propose a system oriented approach to PD modeling. This approach is detailed in the first section following a discussion of the limitation of the traditional approach. The new approach is then applied to the modeling of propofol *vs.* WAV_{CNS} using the induction data from 44 patients undergoing various ambulatory surgeries. The traditional approach is also carried out for comparison purposes.

We show that the system oriented PD approach yields better fit and predictive performance. Furthermore, the resulting PD model is independent of the sensing technology used to observe the effect. Finally, we also show that the traditional approach can lead to strongly colored residuals, which, in turn, affects the accuracy of the static dose *vs.* response characteristic of the drug.

6.1 A New System Oriented Approach to PD Modeling

This section presents in detail a PD modeling approach based on standard system engineering know-how.

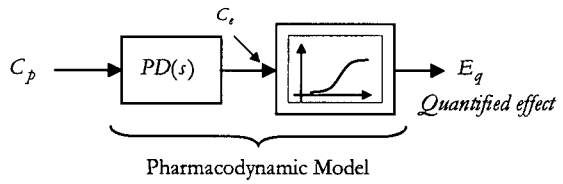


Figure 6.1: Traditional PD model block diagram.

To better grasp the fundamental differences between this system-oriented approach and the traditional pharmacological approach (also referred in the following to as *Sheiner's approach*), we first bring a critical look to the traditional approach previously discussed in Chapter 3.

6.1.1 Sheiner's Approach

The approach described in Chapter 3 can be summarized under the form of the block diagram of Fig. 6.1 where the PD model expresses the relationship between a quantified effect E_q and the plasma concentration C_p of a given drug. This relationship is thus expressed as a Wiener type model, where a Linear Time Invariant (LTI) function is followed by a non-linear element to capture the static dose *vs.* response relationship.

In this approach, $PD(s)$ is a first order transfer function with a unity gain. It expresses the drug concentration C_e within the target organ as a function of the plasma concentration. The transfer function $PD(s)$ is uniquely defined by the effect site rate constant k_{e0}, which is the rate at which the effect-site concentration equilibrates with the plasma concentration, see (3.13). The addition of this function to the classical dose/response curves was proposed by Sheiner *et al.* [114] in 1979 in an attempt to collapse the 'hysteresis' loop observed when plotting the effect *vs.* the plasma concentration during rapid administration of the drug, see Fig. 3.9.b.

The non-linear function which relates the quantified effect to the effect site concentration is handled by a Hill equation in order to capture the sigmoid characteristic of the saturation of the dose *vs.* response curve. This saturation can be particularly severe when considering an effect that can only be measured over a limited portion of the drug therapeutic window. For slowly varying plasma concentrations, $PD(s)$ can be neglected, in which case the PD model can be simplified to the classical dose *vs.* response curve expressed by the Hill equation.

While Sheiner's approach has been the standard in PD studies since 1979, PD models are known to suffer from a limited accuracy, large patient variability and strong non-linearity. In particular, the traditional PD approach suffers from three major limitations, which might explain its overall poor performance:

1. use of a *measured* input signal even though this signal is not readily available,
2. use of a limited LTI model order that fails to capture all of the system dynamics, and,
3. inherent assumption that sensor dynamics are negligible.

These limitations are discussed in more detail below.

For illustration purposes, we consider the 6 published PD studies reported in Table B.2 which involve the effect of propofol on electroencephalographic variables. These studies are recent (1997-2002) and originate from different research centers (USA, UK, Japan and The Netherlands). As such we believe that they are a good representation of the current state-of-the-art in PD modeling practices.

Predicted vs. Measured Plasma Concentration In most studies, the plasma concentration C_p is measured through blood sampling. There are different scenarios as for the use of these blood samples. Some authors (e.g., Schnider *et al.*, White *et al.*, and Ludbrook *et al.*) first derive their own PK model in order to predict the time course of the plasma concentration, which they further use with the corresponding measured effect to identify the PD parameter set. This method is referred to as the *parameteric method.* Others (Schnider *et al.*, White *et al.*, and Kuizenga *et al.*) use a connect-the-dots method to extract the time course of C_p and adjust the k_{e0} value to collapse the hysteresis loop, and further, identify the dose *vs.* response curve. In particular, Kuizenga uses linear interpolation between two consecutive measured samples during increases in blood concentration, as well as a logarithmic interpolation between measured concentrations during decreases in concentrations. This method is usually referred to as the *non-parametric method.*

While both methods seem appropriate, the blood plasma concentration of the drug is usually not available in real time. Hence, the advantage of the parametric method lies in the fact that a PK model is given in order to predict the plasma concentration time course. The method can be used to implement on-line algorithms for drug effect prediction. However, questions can arise as to the validity of PK models derived on a very limited number of patients, as compared to dedicated PK models derived in studies involving hundreds of subjects and using the latest identification methods.

As for the PD models developed by the non-parametric method (*i.e.*, where the input identification data are *measured* instead of *predicted*), one can reasonably argue that these models are probably closer to expressing the true pharmacodynamics of the drug. However, their usefulness remains limited as long as no direct real-time plasma concentration measurement is available.

In a minority of PD studies, authors use third party PK models to predict the plasma concentration time course. For instance, Kazama *et al.* have used the Gepts *et al.* propofol PK model (see Table B.1). The blood samples were then used to validate the PK model output. Cases which strayed too far from the predicted plasma concentrations were rejected from the analysis, thus yielding a more consistent PD parameter set.

It appears that the parametric approach is the only approach yielding direct applications (providing the PK parameter set is disclosed by the authors). However, the limited amount of data used to derive the PK models can be a significant factor limiting the reliability of the overall PKPD model. In that respect, the effect of the PK model in PD studies has been discussed in a 2003 study by Minto *et al.* [165]. The authors concluded that substituting the PK set used to predict the plasma concentration in the PD study

with another PK set will yield significantly different results, mostly during transient phases.

Inadequate LTI Model Structure The effect *vs.* plasma concentration 'hysteresis' phenomenon has received considerable attention since the original work of Sheiner. The idea of collapsing the hysteresis loop[1] using an additional fictitious PK effect-site compartment can be particularly attractive to clinicians and pharmacologists, as it provides a physical meaningful explanation for the effect lag, and a way to mathematically express it in the well-known framework of compartmental mamillary models.

In Sheiner's approach, one would try to collapse the 'hysteresis' loop through a first order transfer function with unity gain (3.13). We can see in Figure 6.2.c-e how the selection of the model time constant k_{e0} affects the width of the loop. As k_{e0} decreases, the loop collapses and then inverts. When k_{e0} equals the system time constant, we obtain a perfect equilibration between the measured and predicted time courses.

However, as with any identification procedure, the goodness of the fit between measured and predicted time courses depends both on the quality of the identification data, *and the selection of the model order and structure.* In the case presented in Figure 6.2, the model has a similar structure to that of the system, which explains the success of this approach. However, when the system dynamics are more involved than that of the model, the traditional approach can easily yield poor results. In particular, a mis-modelled time delay can result in a smaller time constant k_{e0}, since this constant now captures both the system first-order and time delay dynamics. This is illustrated in Figure 6.3.

Another important implication is that the residuals (*i.e.*, the error between the measured and predicted time courses) may be strongly biased. This can result in a stronger non-linear characteristic if a Hill non-linear saturation is derived to further minimize these residuals. For instance, in the example of Figure 6.3, a non-linear Hill saturation was added to improve the goodness of the fit between the measured and predicted outputs. Even though the system is linear, the model now includes a non-linear element, which is a direct consequence of not including all of the linear dynamics in the LTI part of the model.

An important consequence of this result is that PD models derived based on Sheiner's approach may only provide an adequate fit for input signals similar to those used during the identification procedure. If a different input signal is presented to the system, the model will fail in predicting adequately the system output. For instance, it is clear in the example presented in Figure 6.3 that the system output will not be accurately predicted for a slow varying low amplitude input signal due to the Hill saturation.

It appears therefore particularly important that a thorough analysis of input/output data be performed

[1]Note that the term 'hysteresis' is used here inadequately since the width of the loop depends on the rate of administration of the drug. For instance, the loop is much wider for bolus administration than for a constant drug infusion. It will then be clear to anyone well-versed in system modeling and identification that the 'hysteresis' phenomenon results mostly from unaccounted dynamics between the plasma concentration and the effect, rather than from a true non-linear hysteresis in the system. This can be easily illustrated in the example presented in Figure 6.2, where we consider a simple first order system. This system is subject to a given input signal. Using a measurement of the output, it is possible to derive an input *vs.* output plot which shows the disequilibrium generated by the system first order dynamic, see Figure 6.2.b. Even though this plot resembles to that of an hysteresis, there is no such element in the system, which is actually perfectly linear.

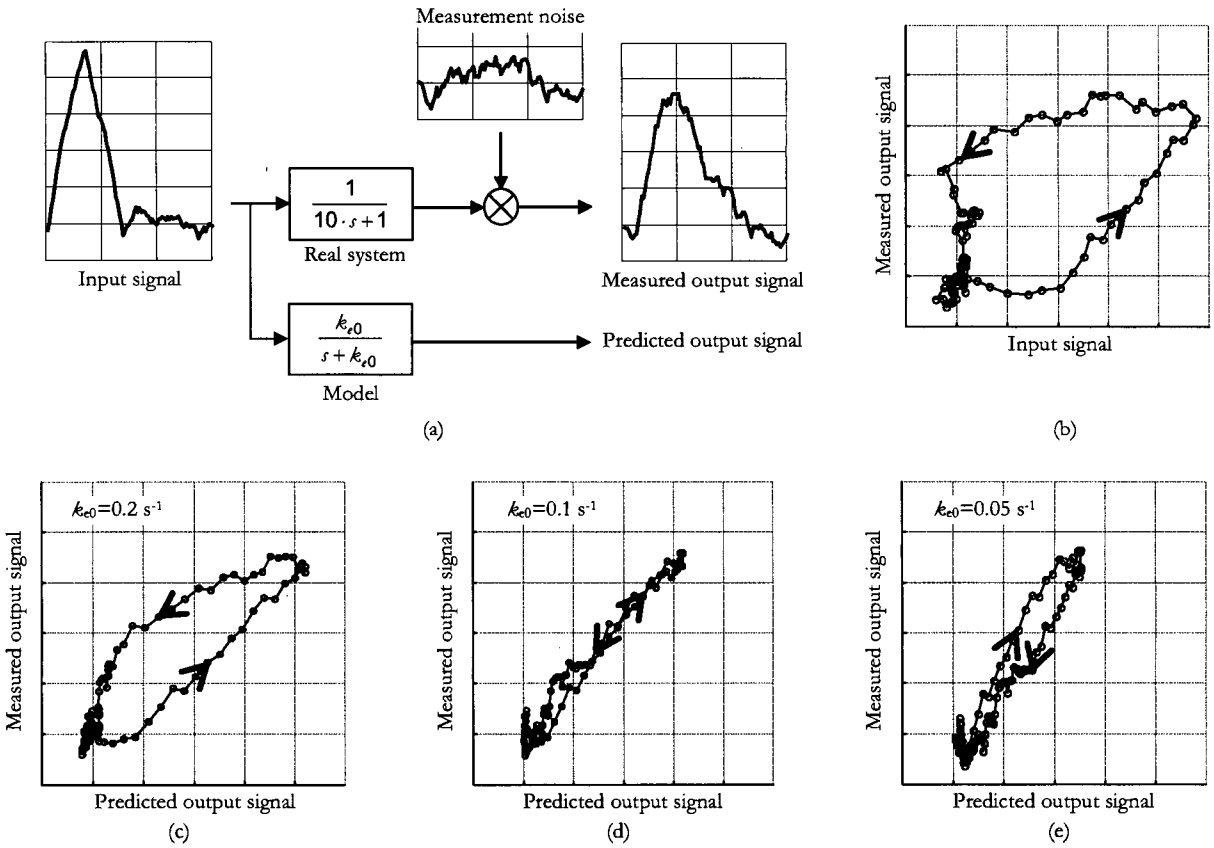


Figure 6.2: Example of the 'hysteresis' phenomenon observed between the drug plasma concentration and the corresponding effect. (a) In this example, an output signal is obtained from a first order system. (b) Plotting the output *vs.* input relationship yields the 'hysteresis' curve described by pharmacologists. By tuning appropriately the rate constant k_{e0} of a first order model, one can collapse the loop. (c) Under-compensation: the rate constant is too large. (d) Perfect equilibration. Note that the measured *vs.* predicted characteristic is linear (*i.e.*, in this case, a Hill saturation would not improve further the fit). (e) Over-compensation: the rate constant is too small, which results in an inverted loop.

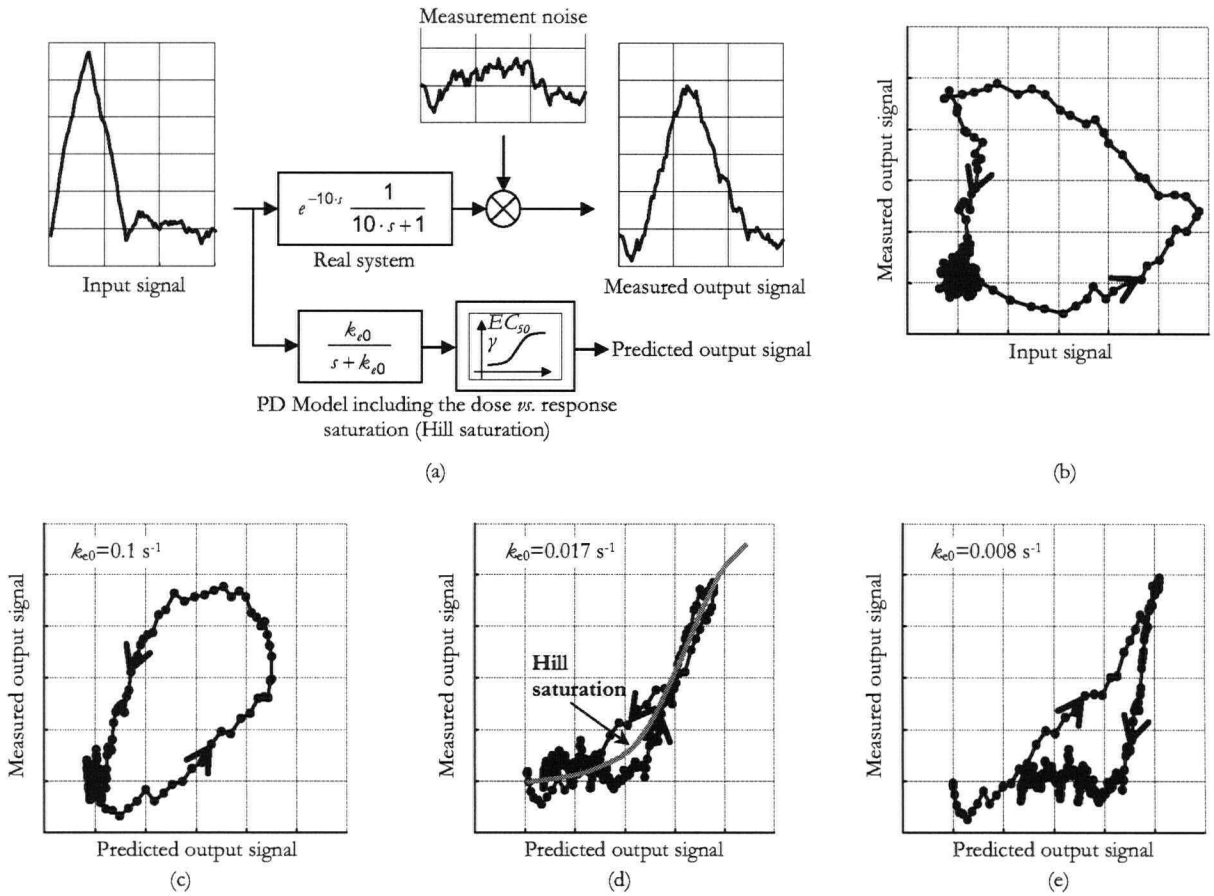


Figure 6.3: Effect of a pure time delay. (a) A pure 10 seconds time delay was added to the true system of Figure 6.2. (b) Plotting the output *vs.* input relationship yields a 'hysteresis' curve similar to that of the previous example (*i.e.*, there is no discriminating feature which can stress out the presence of a delay). (c) Under-compensation. (d) Equilibration. Note, however, that the rate constant is now about 6 times smaller than the true time constant of the system. The first order model therefore mostly captures the time delay dynamics. The model is just a coarse approximation of the true system. Even though the 'hysteresis' loop is reduced, the error between the predicted output and the measured output becomes quite large. (e) Over-compensation.

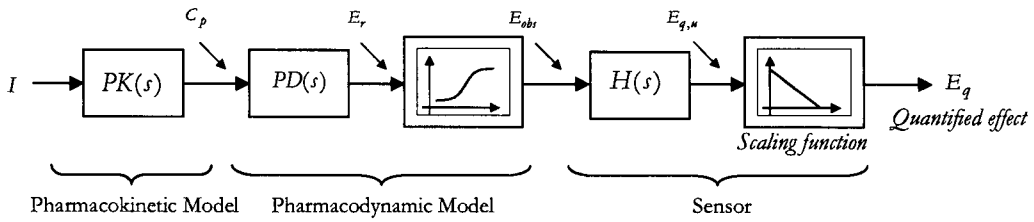


Figure 6.4: Block diagram of the proposed system oriented approach to PD modeling.

to determine the optimal LTI model structure. The non-linear saturation should only be added once all of the system linear dynamics are accounted for by the LTI part of the model, that is, once the residuals are white and show no correlation to the input signal.

While the example of Figure 6.3 illustrated the effect of a time delay unaccounted in the model structure, similar results are obtained when considering higher order systems. Note that Sheiner *et al.* in their original 1979 paper reported that a limitation of their technique is that "its kinetic simplicity may cause it to be inadequate to describe complex pharmacodynamics".

Sensor Dynamics A major part of the drug effect dynamics can be directly related to the sensor dynamic whose output filter extracts the main trend out of the raw data. Sensors typically introduce delays, low pass filtering, and sampling dynamics. They might also be responsible in part for the non-linearity of the dose *vs.* response relationship. In Sheiner's approach, these dynamics are implicitly included in the PD model. As a result, a PD model is limited to the feedback sensor used for its identification. Hence, each time a new sensor is introduced, the PD identification procedure must be repeated. Also, any change in the signal processing algorithm of the sensor will make the corresponding PD model obsolete.

Out of the 6 PD studies, 4 used the BIS index, 1 used auditory evoked potentials and 1 used the semilinear correlation index. Differences between the PD parameter sets can be explained by differences in these feedback quantities. Hence, only the 4 studies involving BIS can be reasonably compared.

This aspect of pharmacodynamics was already alluded to by Kazama *et al.* who tried to remove the influence of the BIS averaging filter by shifting in time the BIS data in order to synchronize them with the plasma concentration time course predicted by the TCI pump. Even though the resulting PD model does not exactly predict the BIS time course, it is more consistent than PD models obtained in other studies (lesser patient variability).

6.1.2 Proposed Methodology

To alleviate these limitations, we propose to revisit pharmacodynamic modeling from a more systematic system engineering point of view. As a guide for this discussion, we consider again the example of the effect of an anesthetic drug onto the brain.

While the exact mechanism of action of anesthetic drugs is still debated, it has been shown that drugs

such as barbiturates, benzodiazepines and propofol exert their pharmacological action at the level of the γ-aminobutyric acid (GABA) receptors. The GABA is one of the major inhibitory neurotransmitter in the brain. Its role is to counterbalance the influence of excitatory neurotransmitter. By mimicking its structure, anesthetic drugs bind themselves to the cell receptors, thus sending false information and shifting the excitatory balance towards relaxation. Hence, the effect of anesthetic drugs can be well understood when considering the percentage of bound GABA receptors of each individual cell in the brain. Depending on the proximity of blood vessels, and thus the bio-availability of the drug, this percentage is going to evolve in time.

Unfortunately, the effect at the cellular level cannot be readily measured. However, the combination of each individual cellular effect will result in a number of macroscopic effects which can be observed and quantified. For example, as an increasing number of receptors is occupied by the drug, loss of consciousness and memory formation can be observed. The depression of the CNS will result in a drop in heart rate and blood pressure. Similarly, changes in the EEG patterns can be observed. These EEG changes can be further quantified into a unique value (*e.g.*, the WAV_{CNS}) through the signal processing technique described in Chapter 5.

Hence, when considering the pharmacodynamics of anesthetic drugs, one can consider the effect of a given drug concentration in the blood perfusing the brain onto the EEG signal itself. We can take here the assumption that this relationship can be adequately described by a LTI transfer function followed by a non-linear function. The LTI element captures the dynamics between the percentage of bound GABA receptors and the plasma concentration. The non-linear function, in turn, models the saturation phenomenon proper to EEG changes: a low number of bound GABA receptors will have no quantifiable effects on the EEG, while no increase in the number of bound receptors will provoke any further change in the EEG when the complete suppression of cortical activity is reached.

The patient system can thus be modelled using the block diagram of Fig. 6.4. As compared to the traditional approach, the plasma concentration $C_p(s)$ is obtained using a well defined PK model derived for arterial blood concentration. Furthermore, the pharmacodynamics are now limited to the effect of the drug onto an observable physiological signal (and not its quantified correspondent). In this scheme, the sensor dynamic is thus a separate entity all together. Also, no *a priori* decision is made as to the structure of $PD(s)$, which can be any n-order plus delay transfer function:

$$PD(s) = \frac{E_r(s)}{C_p(s)} = \mathrm{e}^{-T_d \cdot s} \cdot \frac{b_m \cdot s^m + b_{m-1} \cdot s^{m-1} + \ldots + b_1 \cdot s + a_0}{s^n + a_{n-1} \cdot s^{n-1} + \ldots + a_1 \cdot s + a_0} \cdot \frac{1}{2 \cdot EC_{50}}, \quad m \leq n \tag{6.1}$$

Residual analysis should be performed to validate the selected model structure.

In this approach, the pharmacodynamic model establishes the relationship between the observable effect E_{obs} and the plasma concentration. E_{obs} is bounded between 0 (no effect) and 1 (maximal effect which can be observed in the EEG signal). The term EC_{50} represents here the steady-state plasma concentration necessary to obtain 50% of the measurable effect. Conversely to the traditional approach, the non-linear element does not express the dose *vs.* response relationship. It only models the saturation of the effect

with respect to the dose. This saturation can be expressed using the Hill relationship, which is essentially a non-linear function that can be represented as a variable steady state gain in the frequency domain, see Section 6.3.4 and Section 7.2.1.2:

$$E_{obs} = \frac{E_r^\gamma}{0.5^\gamma + E_r^\gamma} \quad \Rightarrow \quad 0 \leq E_{obs} \leq 1 \tag{6.2}$$

Note that the non-linear element is defined here entirely by the steepness coefficient γ.

Finally, the quantified effect E_q is obtained through filtering and scaling. The sensor filter $H(s)$ is essentially a unity gain element, while the scaling function is an affine function defined by two scaling parameters: E_{max} and E_0 as follows:

$$E_q = (E_{max} - E_0) \cdot E_{obs} + E_0 \tag{6.3}$$

It is interesting to note that, as compared to Sheiner's approach, the static dose *vs.* response curve is obtained by considering the non-linear function (6.2), the static gain of $PD(s)$ and the scaling function of the sensor:

$$E_{q,ss} = (E_{max} - E_0) \cdot \frac{C_{p,ss}^\gamma}{EC_{50}^\gamma + C_{p,ss}^\gamma} + E_0 \tag{6.4}$$

where $C_{p,ss}$ is the steady state blood plasma concentration of the drug.

Similarities between the traditional approach and the new system oriented approach are obvious. The traditional approach can be viewed as a simplification of the proposed approach, where the non-linear function has become an output non-linearity, and where the traditional $PD(s)$ now captures both the effect and the sensor dynamics. However, it is clear that this is an over-simplification, which results in significantly degraded modeling performances.

6.2 Application to Propofol and WAV_{CNS} Index

In this section, we derive a PD model for Propofol following both the traditional PD approach, and the new system-oriented methodology presented in Section 6.1.2. The models derived here aim at predicting the time course of the WAV_{CNS} following intravenous propofol administration.

6.2.1 Clinical Data

To identify the pharmacodynamic model of propofol, we are using here the induction data collected during the LMA study (see Annex D.2). In this study, the WAV time course during propofol induction was observed in 76 patients. The LMA insertion can provoke a strong reaction but is not as stimulating as the endotracheal intubation. The LMA study has revealed that patients that do react to the LMA insertion (jaw tone, coughing, swallowing, grimacing, etc.), had also a significantly different WAV time course. Furthermore, most of the reaction cases elicited an EMG activation, which resulted in an elevation of the WAV during the insertion itself. Hence, the LMA insertion can be considered to be a typical output

disturbance. Among the 76 patients of the study, 51 patients aged 18 to 60 years did not elicit any strong airway reaction to the insertion. To avoid having the LMA reaction acting as a disturbance in the output identification signal, only these 51 cases were considered for this analysis.

In the LMA study, the start and end of the propofol injection were recorded as well as the LMA insertion time. Even though gaseous anesthetic drugs were given once the patient's airway was instrumented, it is usually assumed that these drug have only a marginal effect in the first 2 minutes following their administration. Hence, the identification window was chosen to start at the propofol injection, and to stop 90 seconds after the LMA insertion. Three cases had insufficient data to cover the whole identification window. These cases were withdrawn from the analysis.

In order to keep consistent results, it was further decided to eliminate all cases whose WAV during induction reached values lower than 10. It is indeed possible that a saturation phenomenon can affect the WAV time course in the lower WAV values, which in turn would result in a significantly different dynamics. Two cases were excluded based on this criteria.

Finally, two more cases were withdrawn from the study due to a large EMG activation observed after the LMA insertion. These cases were characterized by strong extremity movements, but no airway reaction.

A total of 44 cases were thus compiled for the identification procedure. As compared to other published PD studies, this is the second largest study in terms of patient population. These cases were further subdivided into 4 age groups (18-29 yrs; 30-39 yrs; 40-49 yrs and 50-60 yrs) to assess if age represents a covariate factor in the PD models.

Note that a potential limitation of this study is the co-administration of a fentanyl bolus as part of the induction sequence (usually about 1 minute prior to the propofol administration). In this analysis, we neglected the effect of this fentanyl dose. Even though some synergism might have occurred, it is a well documented fact that the propofol/opioid synergism is not very acute in terms of hypnosis.

6.2.2 PD Identification - Traditional Approach

The identification procedure for the traditional PD approach can be divided into three stages. An illustrative example is given in Fig. 6.5.a.

Stage #1: Plasma Concentration Profile To identify the pharmacodynamic parameters, it is necessary to first calculate the time course of the propofol plasma concentration based on the infusion profile. To achieve this, a number of PK models can be used. We select here the parameter set from Schüttler and Ihmsen (see Section 3.1.3) since it is derived based on a large adult population (20 to 60 years) and incorporates both the age and weight as covariates. Also, one advantage of Schüttler's PK set is that the parameters can be adapted for a bolus-type administration, which is necessary when using induction data.

Stage #2: Effect Dynamic Identification Once both the input C_p and output WAV_{CNS} data are defined, the effect dynamic $PD(s)$ is identified based on a least square optimization algorithm, where the

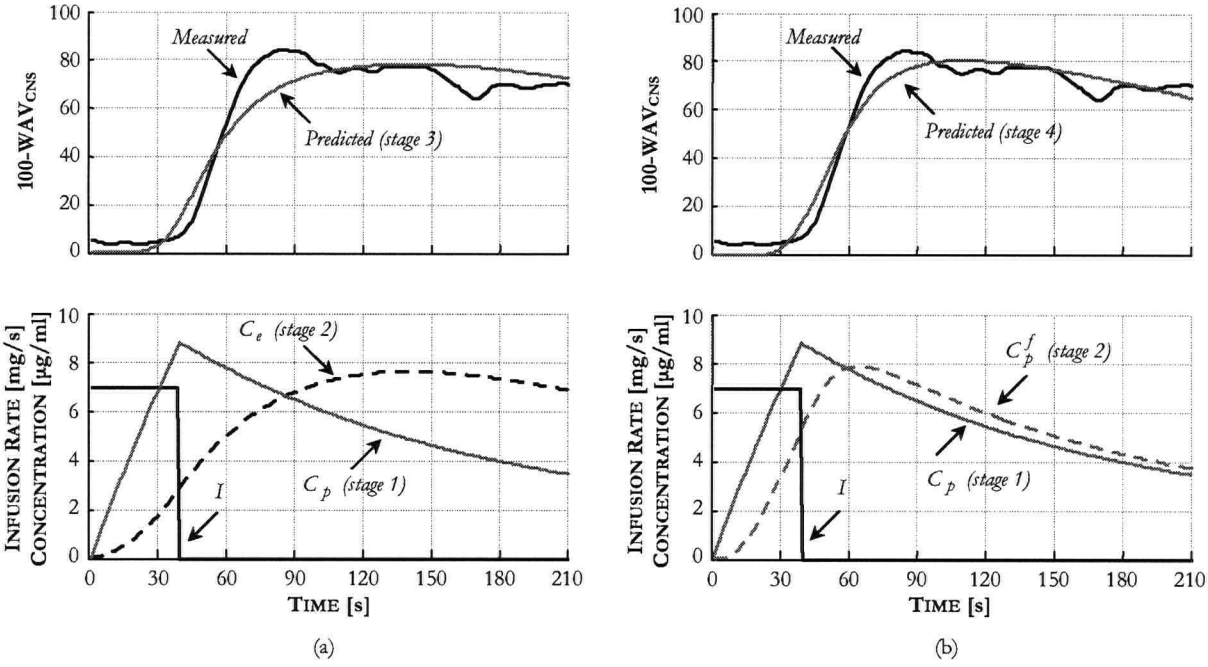


Figure 6.5: PD identification illustrative example (patient #52). (a) Traditional approach (b) System oriented approach

square of the error between the measured and predicted effects is minimized. This yields the effect rate constant k_{e0}. This method is essentially identical to that of the 'hysteresis' collapse method.

Once k_{e0} is identified, $PD(s)$ is defined such as in (3.13) and used to estimate the effect-site concentration C_e.

Stage #3: Hill Parameters Identification The Hill parameter set $\{EC_{50}, \gamma\}$ is finally identified using C_e and the measured WAV_{CNS} values. A search algorithm compares the predicted *vs.* measured WAV_{CNS} and selects the Hill parameters to minimize the root mean square (RMS) of the residuals.

During the identification analysis, the RMS of the residuals is used as a minimization criterion to optimize the fit between the predicted *vs.* measured effect. Other criteria, such as the performance indexes presented in Section 6.3.2, can also be used.

6.2.3 PD Identification - New Approach

The identification procedure for the system oriented approach is more convoluted as it involves knowledge of the sensor dynamics and the identification of a time delay. An illustrative example is given in Fig. 6.5.b.

Stage #1: Plasma Concentration Profile Similarly to the traditional approach, Schüttler's PK models are used to predict the time course of the propofol plasma concentration.

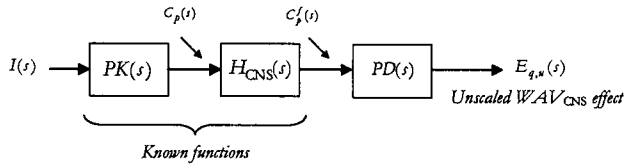


Figure 6.6: Block diagram used in stage #3: the non-linear element is omitted and the sensor filter is used as an input filter.

Stage #2: $\mathbf{WAV_{CNS}}$ Dynamics Conversely to the traditional approach, the sensor dynamics are now a separate subsystem, which can be expressed as a transfer function $H_{\text{CNS}}(s)$ and a scaling function. In that respect, a major advantage of the WAV technology is that it allows us to express the WAV_{CNS} dynamics as a *linear time invariant* transfer function. In particular, we found in Section 5.4 that:

$$H_{\text{CNS}}(s) = \frac{\omega_0^2}{s^2 + 2 \cdot \zeta \cdot \omega_0 \cdot s + \omega_0^2} \cdot \frac{1 - \mathrm{e}^{-s}}{s} \tag{6.5}$$

where $\omega_0 = 0.02$ Hz and $\zeta = 1$ (current implementation).

Stage #3: Effect Dynamic Identification Once the sensor filter $H_{\text{CNS}}(s)$ is defined, the block diagram of Fig. 6.4 is rewritten as in Fig. 6.6, where the non-linear saturation function is omitted. The only unknown element is the transfer function $PD(s)$ whose output $E_{q,u}$ can be directly obtained from the WAV_{CNS} values by applying the inverse of the scaling function. In this block diagram, we use the property of linear time invariant systems to bring the sensor filter as an input element. The transfer function $H_{\text{CNS}}(s)$ is then used to obtain a filtered version C_p^f of the plasma concentration.

Using both C_p^f and $E_{q,u}$ as input and output identification signals respectively, the parameters of $PD(s)$ can be estimated using a classical least square identification method.

One difficulty arises in the determination of the order of $PD(s)$. With data such as the LMA data, it would be difficult to identify a high order function because of the limited input excitation. In this case, we choose $PD(s)$ to be a first order plus delay function:

$$PD(s) = \frac{E_r(s)}{C_p(s)} = \mathrm{e}^{-T_d \cdot s} \cdot \frac{k_d}{s + k_d} \cdot \frac{1}{2 \cdot EC_{50}} \tag{6.6}$$

where k_d is the effect dynamics and EC_{50} is the effect site concentration that yields 50% of the maximum observable effect.

This choice is further discussed in Section 6.2.6.

Stage #4: Hill Parameters Identification Once the $PD(s)$ function is identified, the system block diagram is returned to its original form of Fig. 6.4. At this point, only the non-linear element parameters are unknown. Similarly as in the traditional approach, a search algorithm determines the Hill parameters that minimize the RMS of the residuals.

6.2.4 Adequacy of the Identification Data

When identifying the model parameters of a given system, it is usually recommended to design an experiment which will excite all the modes of the system in order to yield unbiased and rapidly converging parameter estimates. A large amplitude white noise is considered to be the ideal identification input signal, since it contains many different frequency components susceptible to stress all of the system dynamics. However, it is not always possible to subject a system to inputs such as Pseudo-Random Binary Sequences (PRBS) or filtered Gaussian signals. In the particular case of pharmacodynamics, the needs of the anesthesia procedure and surgery always supersede the needs of the identification itself. It is therefore important to assess *a priori* the adequacy of the input signal for estimating the LTI parameters of the models.

In the induction data provided by the LMA study, the propofol is given as a single bolus over a given period of time. Using Sheiner's approach, the identification uses $C_p(t)$ as an input signal to determine the time constant k_{e0}. This time course can usually be divided into two distinct parts: an initial large amplitude fast transient following the drug uptake, and a slower decreasing waveform corresponding to the initial distribution of the drug. The initial fast transient excites the system in a bandwidth roughly equivalent to 4 times the propofol administration time. This corresponds to a bandwidth of about 0.04 to 0.15 rad·s^{-1} across the whole study population. In the proposed approach, we are using a filtered version of the plasma concentration to identify the gain, the time constant and the delay of $PD(s)$ as defined in (6.6). This also roughly corresponds to an excitation bandwidth of 0.04 to 0.15 rad·s^{-1}. Note that the second part of the input transient mostly reveals the steady state characteristic of the system. However, the identification window was too short for the system to settle. Therefore, a larger error in the dc gain is expected.

The identification data provided by the LMA study appears therefore suitable to identify this system, providing that the PD dynamics are close or within the excitation bandwidth. Dynamics lower than the identification bandwidth may not be identified precisely due to the limited amount of excitation in that band.

6.2.5 Identification Results

PD models were derived for each patient following the identification methods described previously. The results are summarized for each age group in Table 6.1.

In the traditional PD approach, the models are fully defined by the three parameters $\{k_{e0}, EC_{50}, \gamma\}$ (for simplification, we assumed that E_{max}=0 and E_0=100). As for the proposed PD approach, the models are defined by the parameter set $\{T_d, k_d, EC_{50}, \gamma\}$ and the sensor dynamic $H_{\text{CNS}}(s)$. In both cases, these models must be used in conjunction with Schüttler's pharmacokinetics.

Note that, following the discussion in Section 6.2.1, the k_{e0} and k_d values are well contained within the identification bandwidth of the input signal.

PATIENT #	TRADITIONAL PD APPROACH			PROPOSED PD APPROACH			
	LTI	Hill		LTI			Hill
	k_{e0} [$s^{-1} \cdot 10^{-3}$]	EC_{50} [$\mu g/ml$]	γ	T_d [s]	k_d [$s^{-1} \cdot 10^{-3}$]	EC_{50} [$\mu g/ml$]	γ
G1: ≥18 - <30 years							
007	10.0	1.8	2.9	22	133.5	3.2	4.7
008	13.8	2.7	3.2	4	44.4	3.1	2.5
010	0.8	0.4	3.3	44	25.0	2.4	1.9
015	1.9	1.1	1.8	45	51.5	3.8	1.2
016	4.2	1.7	1.9	39	85.7	3.8	2.3
023	10.7	3.1	2.8	18	82.5	3.9	2.1
030	5.6	1.8	3.1	32	44.4	2.9	2.8
035	8.4	1.9	3.8	12	26.7	1.9	2.3
038	11.8	3.0	2.6	7	35.2	3.4	1.9
046	9.9	2.4	3.9	9	32.8	2.8	2.8
048	8.6	2.2	3.3	17	46.4	2.8	2.3
053	9.9	2.1	3.4	4	26.2	2.4	2.5
058	10.6	1.9	3.0	9	50.4	2.5	2.6
066	12.9	2.4	2.4	18	160.5	3.6	3.9
071	8.1	1.8	2.7	20	75.0	2.5	1.9
Mean	**8.5**	**2.0**	**2.9**	**20.0**	**61.4**	**3.0**	**2.5**
SD	**3.8**	**0.7**	**0.6**	**13.9**	**40.1**	**0.6**	**0.8**
G2: ≥30 - <40 years							
006	2.4	1.1	4.6	44	54.8	3.2	2.7
009	5.9	2.2	2.6	29	83.1	4.0	2.3
029	6.3	2.6	3.2	18	34.4	3.7	2.1
036	10.9	2.9	1.5	1	29.6	3.3	1.2
047	11.9	3.0	2.2	1	24.9	3.1	1.5
049	8.7	3.0	2.5	12	35.2	3.9	1.8
051	9.5	2.5	2.9	4	24.8	2.7	2.0
061	6.1	1.8	2.9	12	28.7	2.8	2.2
063	8.2	2.1	2.5	5	27.0	2.8	2.1
065	17.0	3.3	2.5	4	67.2	3.6	2.0
068	7.3	2.4	2.5	12	29.3	3.1	1.8
074	5.1	2.1	2.1	13	29.1	3.7	1.8
Mean	**8.3**	**2.4**	**2.7**	**12.9**	**39.0**	**3.3**	**2.0**
SD	**3.8**	**0.6**	**0.7**	**12.6**	**19.0**	**0.4**	**0.4**
G3: ≥40 - <50 years							
004	2.1	0.9	2.6	35	38.0	3.3	1.8
025	9.5	4.4	1.9	11	36.6	6.1	1.3
027	15.3	4.7	2.2	2	32.6	4.7	1.3
040	10.8	4.1	2.4	12	35.0	4.5	1.4
042	7.2	2.8	2.5	10	28.7	3.9	2.0
043	7.7	2.8	2.5	12	34.8	3.9	1.8
052	11.7	3.2	3.1	9	36.6	3.2	1.9
069	9.8	2.7	2.9	8	35.6	3.4	2.3
072	5.4	1.7	2.5	13	30.0	3.0	1.9
Mean	**8.8**	**3.0**	**2.5**	**12.4**	**34.2**	**4.0**	**1.7**
SD	**3.8**	**1.2**	**0.4**	**9.1**	**3.1**	**1.0**	**0.3**
G4: ≥50 - ≤60 years							
018	11.9	3.1	3.1	3	31.5	3.5	2.3
033	4.4	2.3	1.8	29	42.0	4.4	2.2
041	8.3	3.7	1.6	2	21.8	4.7	1.4
057	5.4	2.2	1.6	16	28.8	3.7	1.1
060	7.0	2.7	2.6	10	26.4	4.0	1.9
064	12.7	3.9	1.3	6	58.0	5.0	1.5
070	8.7	3.1	2.1	6	32.2	4.2	1.5
075	5.2	1.8	2.4	12	24.3	3.1	1.8
Mean	**8.0**	**2.9**	**2.1**	**10.5**	**33.1**	**4.1**	**1.7**
SD	**3.1**	**0.7**	**0.6**	**8.8**	**11.8**	**0.7**	**0.4**

Table 6.1: PD models obtained from the proposed approach with E_{max}=0 and E_0=100.

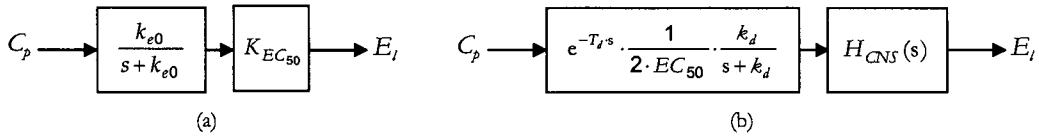


Figure 6.7: LTI part of the models. (a) In the traditional approach, the Hill equation is replaced by its static gain. (b) In the system-oriented approach, the sensor dynamics is to be explicitly included.

6.2.6 Model Validation

Any identification procedure involving *a priori* model structures, such as the one proposed by Sheiner *et al.* or the one defined in (6.6), should be followed by a discussion on the adequacy of the model for representing the system. Otherwise, the model may not represents accurately the system when subjected to inputs different than the ones used for estimating the model parameters.

Model validation can be carried out in a number of ways. It is usually recommended to test the model using a combination of tools. To give confidence in the models derived in this chapter, we present here validation results using some of these tools.

Direct Residual Analysis In pharmacodynamic modeling, the analysis of the residuals should be carried out only on the linear part of the model. In the traditional approach, the LTI part of the model is limited to a first order element and the static gain $K_{EC_{50}}$ of the Hill dose *vs.* response relationship (see Section 6.3.4). In the proposed approach, the LTI part of the model includes both $PD(s)$, as defined in (6.6), and the sensor dynamics, see Figure 6.7.

To illustrate this discussion, we consider here the pharmacodynamic models obtained for Patient #65. In this particular case, no anesthetic gas was delivered to the patient after the insertion of the LMA due to a mechanical failure. As a result, the patient progressively regained consciousness as the propofol concentration decreased in the blood plasma. About 7 minutes into this case, arousal was observed and the situation was promptly corrected.

In terms of identification data, this case provides us with both a fast and large amplitude transient following propofol administration, and a slower transient during the distribution and elimination of the propofol, see Figure 6.8. The model outputs were calculated following the parameters calculated in Table 6.1.

The residuals are usually calculated as the difference between the measured and predicted outputs, see Figure 6.9.a. Using the proposed system-oriented model, it can be noted that the peak amplitude of the residuals during the fast and large input transient is similar to the amplitude obtained during the slower transient. Considering that the system noise characteristic remains identical over the whole time window, we can reasonably infer that most of the linear dynamics have been accounted for. Conversely, the residuals obtained from the traditional approach are substantially larger during the first input transient. Direct inspection of the residuals tend therefore to stress out the inadequacy of the traditional LTI model.

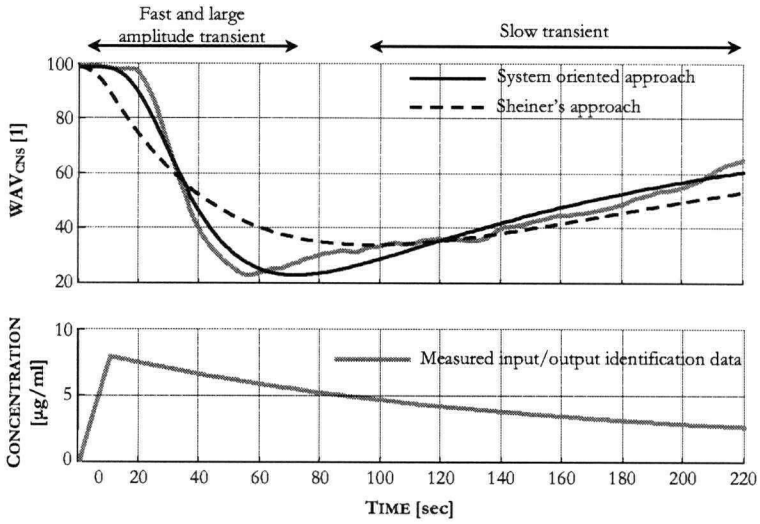


Figure 6.8: Propofol plasma concentration (bottom plot) and WAV_{CNS} time course (top plot) compared to the outputs obtained from the PD linear models (Patient #65).

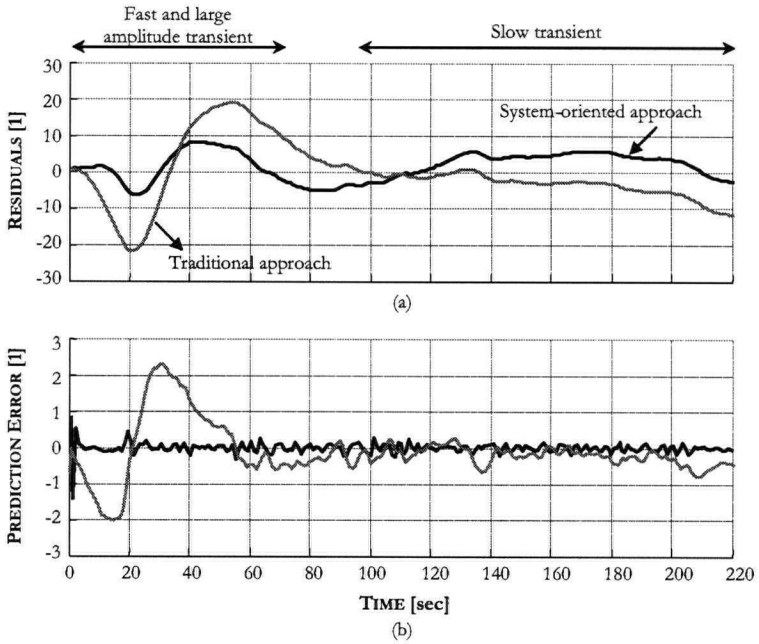


Figure 6.9: Measured *vs.* predicted WAV_{CNS} time courses for both models (Patient #65). (a) Standard residuals. (b) Prediction errors (assuming an autoregressive model).

Prediction Error Analysis A better way of expressing this is to consider the prediction error e_p. Considering that the system can be modelled by an autoregressive ARX model, we have:

$$A(q) \cdot WAV_{\text{CNS}}(k \cdot T_s) = B(q) \cdot C_p(k \cdot T_s - T_d) + e(k \cdot T_s), \tag{6.7}$$

where e is a white noise, A and B are polynomials of the delay operator q^{-1}, and T_s is the sampling time. This model expresses the fact that the output is a weighted sum of past C_p, WAV_{CNS}, and e values. The residuals ϵ are calculated as:

$$\epsilon(k \cdot T_s) = WAV_{\text{CNS}}(k \cdot T_s) - \frac{B(q)}{A(q)} \cdot C_p(k \cdot T_s - T_d) - \frac{1}{A(q)} e(k \cdot T_s), \tag{6.8}$$

In case the model $\frac{B(q)}{A(q)}$ perfectly describes the system, the residuals can be expressed as:

$$\epsilon(k \cdot T_s) = -\frac{1}{A(q)} \cdot e(k \cdot T_s), \tag{6.9}$$

Consequently, the prediction error e_p defined as $e_p(k \cdot T_s) = A(q) \cdot \epsilon(k \cdot T_s)$ should ideally have the characteristics of a white noise. If not, the model $\frac{B(q)}{A(q)}$ did not capture all of the system dynamics. The prediction error obtained from the residuals calculated from both models are presented in Figure 6.9.b. As can be observed from this plot, both prediction errors behave like white noise during the slowest part of the input transient. However, during the initial large input transient, the prediction error obtained from the traditional model is clearly biased as it shows a large oscillation.

Whiteness Test To further test the whiteness of the prediction error, it is usually recommended to calculate the autocorrelation of e_p:

$$\tilde{R}_{e_p}(\tau) = \frac{\sum_{k=1}^{N} e_p(k \cdot T_s) \cdot e_p(k \cdot T_s - \tau)}{\sum_{k=1}^{N} e_p(k \cdot T_s)^2}, \tag{6.10}$$

where N is the number of samples in e_p. Note that (6.10) is here normalized (*i.e.*, $\tilde{R}_{e_p}(0) = 1$) by the autocorrelation coefficient at lag 0 (equivalent to the mean of e_p^2). An easy way of assessing whether e_p is white is then to check whether the signal yields autocorrelation coefficients contained within the 99% confidence interval of a Gaussian distribution $\mathcal{N}(0, 1)$. The corresponding bounds for the normalized autocorrelation function can be calculated as:

$$SD_{\tilde{R}_{e_p}} = \frac{2.58}{\sqrt{N}}. \tag{6.11}$$

A large number of coefficients outside these bounds means that the residuals can be predicted by considering their past values. This is usually a sign that the model order is insufficient.

We plotted the result of this analysis in Figure 6.10.a. It clearly appears that the model order in the traditional approach is not adapted to propofol pharmacodynamics. This result suggests that the order of the A polynomial in the traditional approach is insufficient. In that respect, the proposed system-oriented approach provides much improved results.

Independence Between Prediction Error and Input Another useful test is to calculate the cross-correlation between the prediction error e_p and the input signal C_p:

$$\tilde{R}_{e_p}(\tau) = \frac{\sum_{k=1}^{N} e_p(k \cdot T_s) \cdot C_p(k \cdot T_s - \tau)}{\sqrt{\sum_{k=1}^{N} e_p(k \cdot T_s)^2 \cdot \sum_{k=1}^{N} C_p(k \cdot T_s)^2}}. \tag{6.12}$$

Similarly to (6.10), the cross-correlation is normalized by considering both the autocorrelation of the prediction error at lag 0, and the autocorrelation of the input C_p at lag 0. Ideally, this should also have the characteristics of a white noise, otherwise not all of the input contributions to the output are accounted for by the model. Similarly, this test can be checked easily by using the 99% confidence intervals for the cross-correlation function:

$$SD_{\tilde{R}_{e_p,C_p}} = 2.58 \cdot \sqrt{\frac{1}{N} \cdot \left(2 \cdot \sum_{k=2}^{N} \tilde{R}_{e_p}(k) \cdot \tilde{R}_{C_p}(k) + 1\right)} \tag{6.13}$$

The results of this analysis are presented in Figure 6.10.b. It appears that both models have captured appropriately all the contributions of the input signal. This indicates that the order of the B polynomial in the traditional approach may already be sufficient. Note, however, that the delay in this particular case is rather small (4 seconds only). A traditional model derived on a case involving a larger delay (such as for patient #15) can fail this test.

Model Validity Both the visual inspection of the residuals and the analysis of the prediction error tend to warn against the use of the traditional model for close loop control. At best, traditional models can be used to predict the time course of the drug effect, but only for inputs similar to those used in the identification process.

Conversely, the proposed system-oriented model structure, which incorporates both a first order plus delay and the sensor dynamics, yields adequate results, and accurately models the frequency response of the system. Note, however, that there is no guarantee that this model structure validated for the data collected on the patient #65 will also yield satisfactory results in other cases. Ideally, this analysis should be carried out for each individual cases. However, is is our experience that all cases presented in Table 6.1 are consistent with the example of patient #65.

6.2.7 Nominal Propofol PD Models

The models presented in Table 6.1 were obtained specifically for each patient. In order to expand on this result, we derive population-normed nominal models that can be used to predict the WAV_{CNS} time course for any surgical patient.

A usual approach would be to calculate the nominal model parameters by averaging the PD parameters over the whole study population. However, a particularly interesting result is the age dependency of some of the PD parameters. For instance, in Sheiner's approach, both the EC_{50} and the γ parameters were shown to be statistically different between the 18-39 yrs and 40-60 yrs age groups ($p < 0.05$).

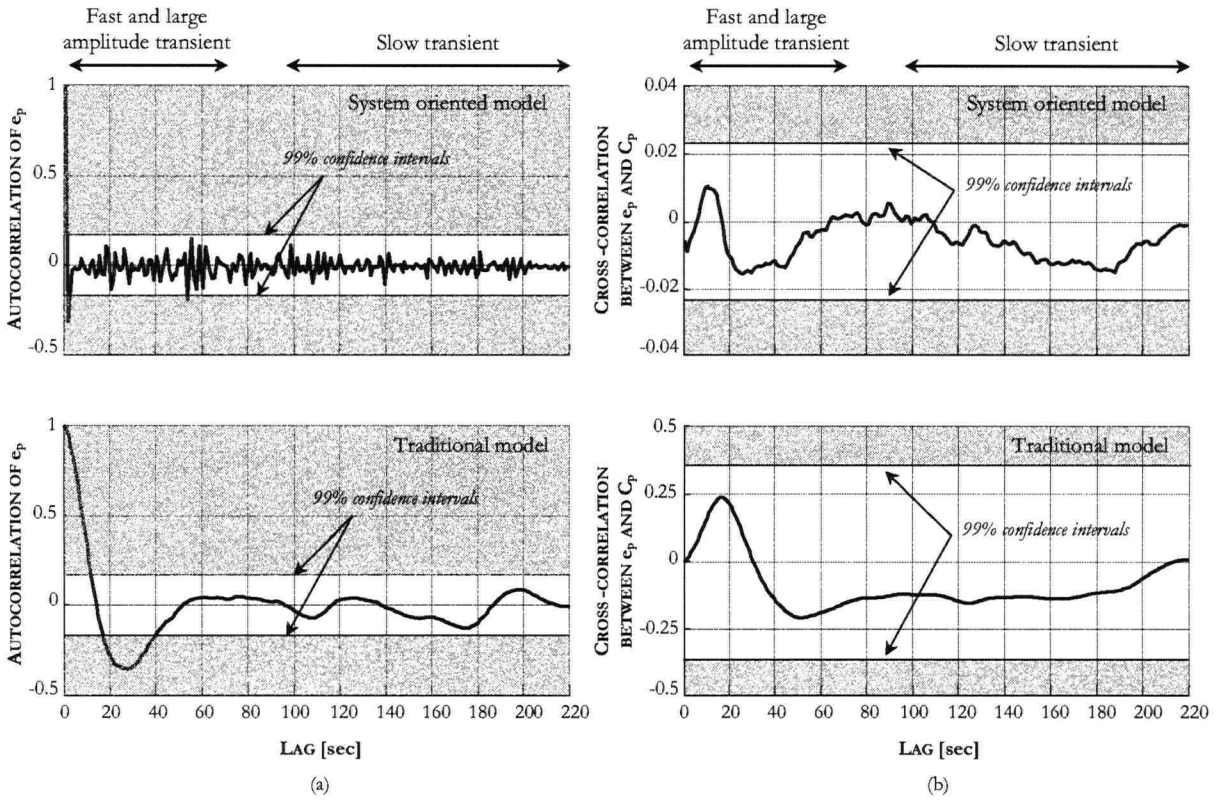


Figure 6.10: Standard residual analysis for model validation. (a) Whiteness test (note that the autocorrelation coefficient at lag 0 is not supposed to be contained within the confidence intervals). Note that the traditional PD model fails this test, indicating the need for a higher degree model. (b) Independence of the residuals with respect to the input signal C_p.

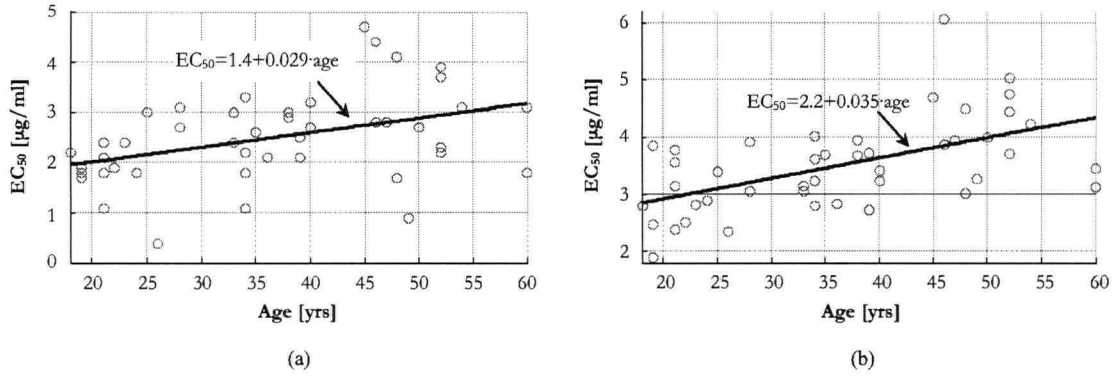


Figure 6.11: Effect of age on the EC_{50} parameter. (a) Traditional approach (b) New system-oriented approach.

	PD Parameter Set	Covariate Significance
Traditional PD		
k_{e0}	$k_{e0} = 8.4$	No covariate
EC_{50}	$EC_{50} = 1.4 + 0.029 \cdot \text{age}$	$p < 0.05$ between G1+G2 and G3+G4
γ	$\gamma = 3.5 - 0.023 \cdot \text{age}$	$p < 0.05$ between G1+G2 and G3+G4
System oriented PD		
T_d	$T_d = 26 - 0.3 \cdot \text{age}$	$p < 0.05$ between G1 and G2+G3+G4
k_d	$k_d = 81 - 0.99 \cdot \text{age}$	$p < 0.05$ between G1+G2 and G3+G4
EC_{50}	$EC_{50} = 1.94 + 0.043 \cdot \text{age}$	$p < 0.001$ between G1+G2 and G3+G4
γ	$\gamma = 3.0 - 0.027 \cdot \text{age}$	$p < 0.005$ between G1+G2 and G3+G4

Table 6.2: Population-normed PD models with age as a linear covariate.

Similarly, all of the parameters obtained from the system-oriented approach have shown a similar dependence. In particular the EC_{50} parameter was shown to increase significantly with age, see Fig. 6.11[2]. Conversely to age, neither weight nor gender could be found to be PD covariates.

Using a basic linear fit, population-normed PD parameters were derived for both approaches with age as a covariate. These models are summarized in the Table 6.2.

6.3 Discussion

In this section, we discuss the results of the results of the two identification procedures carried out to model the effect of propofol administration onto the WAV_{CNS}.

6.3.1 Clinical Adequacy of the Identification Data

We already discussed in Section 6.2.4 the adequacy of the data in terms of their suitability for the identification procedure. In this section, we discuss their suitability for deriving relevant and meaningful clinical models for propofol pharmacodynamics.

As compared to other published PD studies, it will be clear to anyone well versed in pharmacology that the identification data used here suffer from serious limitations. First, they are limited to a single bolus of propofol given within a very short period of time and in an uncontrolled manner (*i.e.*, the rate of injection was assumed constant but may have varied). As a result, the $C_p(t)$ time course may only be a rough estimate of the actual time course of the plasma propofol concentration (in a dedicated PD study, blood samples would have been obtained to check the adequacy of the PK model to capture the time course). Also, a small dose of a fast acting opioid was administered concomitantly, which might have interacted

[2]This result may be surprising. In Schnider *et al.* [166], the authors concluded that the propofol dose required to provoke unconsciousness decreases with age. As compared to Schnider's study, we are here calculating the average propofol dose which provokes a suitable anesthetic depth (WAV_{CNS}). We found that older patients actually require *more* propofol than younger patients, which is consistent with the results published by Kazama *et al.*

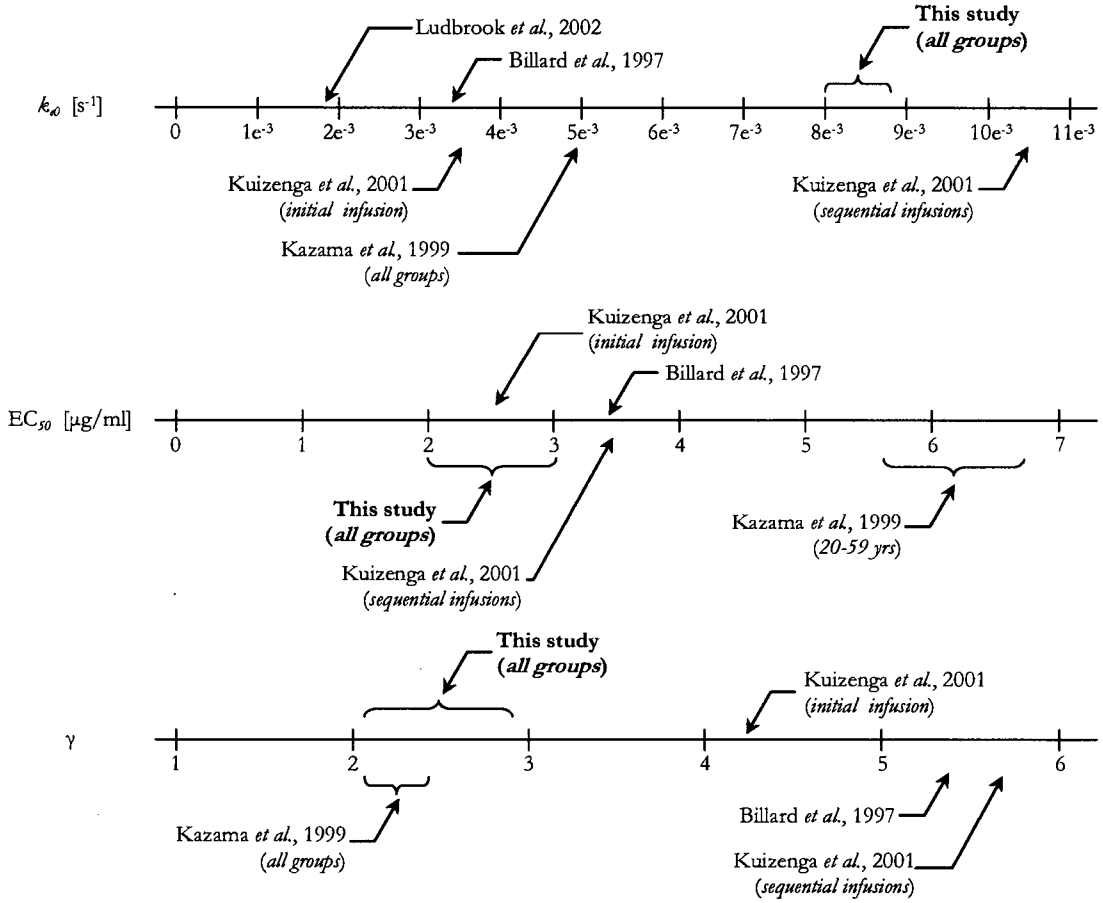


Figure 6.12: Comparison between the identified PD parameters from the traditional approach, and the parameters from literature.

synergetically with the initial propofol bolus.

It is therefore reasonable to question the reliability of the PD results presented in this chapter. In Fig. 6.12, we compare the PD parameters obtained from the traditional approach with other published parameter sets involving the bispectral index as an observed effect. The k_{e0} time constant is faster than most studies, but this could be due to the fact that WAV_{CNS} has a faster dynamic response than BIS during induction.

The EC_{50} of our study was in the same range than Kuizenga *et al.* [118] and Billard *et al.* [117] studies. However, the values were significantly less than the values reported by Kazama *et al.* [116]. However, this can be explained by the difference between the PK parameter set used here, and the PK set used by Kazama.

Finally, the steepness of the Hill characteristic is also reasonably similar to that of other studies. The stronger steepness reported by Kuizenga and Billard can be due to the switching characteristic of the BIS monitor as well as the inherent time delay of BIS (note that Kazama's study did not report such a large γ value, probably because they compensated for the BIS and the PK delay).

In conclusion, even though the identification data were not obtained specifically for a PD analysis, the consistency of our results with other studies tends to indicate that they are appropriate. It is, however, reasonable to expect a better fit and less parameter variability when using data where propofol is the sole agent and where the rate of administration is controlled.

6.3.2 System Oriented *vs.* Traditional Approach

To compare the proposed approach to the traditional approach, we define a number of performance indexes. These indexes can be classified into three categories:

i. *Performance Error (PE) based indexes* For each patient i, PE_i is defined as the error between the measured effect and the predicted effect:

$$PE_i = \{PE_{ij} = w_{ij} - \hat{w_{ij}}\}_{j=1,2,..,N_i} \tag{6.14}$$

where w_{ij} is the j^{th} measured WAV_{CNS} sample and $\hat{w_{ij}}$ is the corresponding predicted WAV_{CNS}. Note that the total number of samples N_i is different for each patient. The mean value of the performance error represents the bias between measured and predicted effect:

$$\text{bias}(PE_i) = \frac{1}{N_i} \cdot \sum_{j=1}^{N_i} PE_{ij} \tag{6.15}$$

The accuracy of the model can be assessed by calculating the root mean square of the residuals:

$$\text{RMS}(PE_i) = \sqrt{\frac{1}{N_i} \cdot \sum_{j=1}^{N_i} PE_{ij}^2} \tag{6.16}$$

Note that the PD parameter sets in Table 6.1 have been derived by minimizing for each patient the $\text{RMS}(PE_i)$ value.

Some clinicians also use the Mean Absolute Residual (MAR) value as an indicator of model accuracy:

$$\text{MAR}(PE_i) = \frac{1}{N_i} \cdot \sum_{j=1}^{N_i} |PE_{ij}| \tag{6.17}$$

The MAR indicate the expected average error between the measured effect and the model output.

ii. *Correlation coefficient* The correlation coefficient r_i^2 is defined for each individual i as:

$$r_i^2 = 1 - \frac{\sum_{j=1}^{N_i} (w_{ij} - \hat{w_{ij}})^2}{\sum_{j=1}^{N_i} \left(w_{ij} - \frac{1}{N_i} \cdot \sum_{j=1}^{N_i} w_{ij} \right)^2} \tag{6.18}$$

r_i^2 is a measure of the correlation between the measured and predicted WAV_{CNS} time courses. Both time courses are identical when $r_i^2 = 1$. Note that this value can be negative.

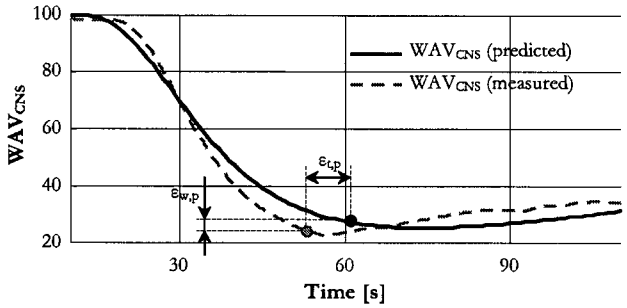


Figure 6.13: Error in peak effect (time and value).

In clinical studies, the median of r_i^2 over the study population is considered to be representative of the performance of a given model. This parameter is a meaningful performance index whenever the identification data cover a large range of plasma concentration and effect. This measure is thus particularly suitable in this analysis.

iii. *Peak Effect* The peak effect corresponds to the time t_p and value w_p of the maximum WAV_{CNS} depth immediately following propofol administration and before airway management, see Fig. 6.13. Note that to reduce sensor noise, t_p and w_p correspond to the time and WAV_{CNS} value when the index first comes within 2% of its maximum peak value.

The time to peak effect has been recently used by Minto *et al.* [165] as a clinically relevant endpoint for assessing the performance of PKPD models.

For each patients, we define two measurements:

$$\varepsilon_{i,t_p} = (t_{i,p} - \hat{t_{i,p}}), \quad \text{and} \quad \varepsilon_{i,w_p} = (w_{i,p} - \hat{w_{i,p}}) \tag{6.19}$$

Patient-specific models The performance of the PD models of Table 6.1 has been assessed using the performance indexes defined previously. The results of this analysis are summarized in the Table 6.3 for the traditional models and Table 6.4 for the system-oriented models.

As compared to the traditional PD approach, the root mean square of the residuals are reduced by about 50%. Also, the overall accuracy of the fit is significantly improved with a median r^2 of 0.968 as compared to 0.879. Of particular interest, 4 cases in the traditional approach scored a correlation coefficient inferior to 0.7 which tends to indicate that the identification procedure failed in obtaining a proper model for these cases. Probably the most dramatic improvement is the reduction of the prediction error of the time to peak effect. The system oriented PD approach also seems to yield more consistent results as shown by the reduced standard deviation in all performance indexes.

Population-normed models A similar performance evaluation was performed for the population-normed models presented in Table 6.5 and Table 6.6.

	Performance Error				Correlation			Peak Effect			
								ε_{i,t_p}		ε_{i,w_p}	
	$\overline{\text{bias}}$	$\overline{\text{RMS}}$	$\overline{\text{MAR}}$	σ(MAR)	Median	Best	Worst	$\overline{\varepsilon_{i,t_p}}$ [s]	$\sigma(\varepsilon_{i,t_p})$ [s]	$\overline{\varepsilon_{i,w_p}}$ [%]	$\sigma(\varepsilon_{i,w_p})$ [%]
≥18 - <30 yrs	0.9	10.6	8.7	6.0	0.859	0.964	0.689	-54.5	25.3	-4.1	5.3
≥30 - <40 yrs	0.8	9.4	7.8	5.2	0.895	0.952	0.779	-46.7	16.2	-3.8	4.3
≥40 - <50 yrs	0.5	8.8	7.3	4.8	0.889	0.934	0.838	-40.0	18.5	-5.5	4.6
≥50 - ≤60 yrs	0.1	9.5	8.0	5.1	0.834	0.947	0.530	-43.8	15.7	-6.6	7.3
All groups	0.7	9.7	8.0	5.4	0.879	0.964	0.530	-47.4	20.3	-4.8	5.2

Table 6.3: Goodness of fit of patient-specific models obtained from the traditional PD approach.

	Performance Error				Correlation			Peak Effect			
								ε_{i,t_p}		ε_{i,w_p}	
	$\overline{\text{bias}}$	$\overline{\text{RMS}}$	$\overline{\text{MAR}}$	σ(MAR)	Median	Best	Worst	$\overline{\varepsilon_{i,t_p}}$ [s]	$\sigma(\varepsilon_{i,t_p})$ [s]	$\overline{\varepsilon_{i,w_p}}$ [%]	$\sigma(\varepsilon_{i,w_p})$ [%]
≥18 - <30 yrs	-0.2	4.8	3.7	3.1	0.985	0.991	0.897	-10.1	8.3	-3.0	4.5
≥30 - <40 yrs	-0.3	5.1	4.2	3.0	0.974	0.995	0.833	-9.3	7.2	-3.8	2.5
≥40 - <50 yrs	-0.3	5	4.1	2.8	0.964	0.993	0.923	-7.8	7.8	-5.1	2.6
≥50 - ≤60 yrs	-0.7	6.1	5.1	3.4	0.910	0.993	0.799	-11.6	4.8	-4.5	4.2
All groups	-0.4	5.2	4.2	3.0	0.968	0.995	0.799	-9.8	7.2	-3.9	3.6

Table 6.4: Goodness of fit of patient-specific models obtained from the system-oriented PD approach.

	Performance Error				Correlation			Peak Effect			
	$\overline{\text{bias}}$	$\overline{\text{RMS}}$	$\overline{\text{MAR}}$	σ(MAR)	Median	Best	Worst	$\overline{\varepsilon_{i,t_p}}$ [s]	$\sigma(\varepsilon_{i,t_p})$ [s]	$\overline{\varepsilon_{i,w_p}}$ [%]	$\sigma(\varepsilon_{i,w_p})$ [%]
≥18 - <30 yrs	3.2	14.8	11.5	9.2	0.727	0.951	-0.127	-47.9	13.0	-2.1	7.3
≥30 - <40 yrs	3.7	12.9	10.2	7.7	0.797	0.911	0.511	-39.4	13.3	-1.9	5.8
≥40 - <50 yrs	3.9	13.2	10.9	7.4	0.783	0.900	-0.188	-35.4	13.8	-3.2	10.7
≥50 - ≤60 yrs	0.1	12.8	10.9	6.6	0.712	0.934	0.258	-37.5	9.9	-6.7	9.0
All groups	2.9	13.6	10.9	7.9	0.772	0.951	-0.188	-41.1	13.3	-3.1	8.0

Table 6.5: Goodness of fit of population-normed models obtained from the traditional PD approach.

	Performance Error				Correlation			Peak Effect			
	$\overline{\text{bias}}$	$\overline{\text{RMS}}$	$\overline{\text{MAR}}$	σ(MAR)	Median	Best	Worst	$\overline{\varepsilon_{i,t_p}}$ [s]	$\sigma(\varepsilon_{i,t_p})$ [s]	$\overline{\varepsilon_{i,w_p}}$ [%]	$\sigma(\varepsilon_{i,w_p})$ [%]
≥18 - <30 yrs	0.6	12.7	9.7	8.2	0.865	0.983	0.105	-6.0	13.9	-1.3	6.8
≥30 - <40 yrs	-0.5	10.4	8.2	6.3	0.891	0.952	0.538	-6.2	13.6	-2.6	5.1
≥40 - <50 yrs	-0.4	9.9	8.2	5.6	0.886	0.964	0.364	-7.2	13.7	-4.8	8.9
≥50 - ≤60 yrs	-4.0	11.1	9.3	5.9	0.756	0.989	0.310	-14.1	10.9	-9.2	8.8
All groups	-0.7	11.2	8.9	6.7	0.867	0.989	0.105	-7.8	13.2	-3.8	7.6

Table 6.6: Goodness of fit of population-normed models obtained from the system-oriented PD approach.

The improvement of the new PD approach in terms of predictive performance is not as significant as with the patient-specific models. As compared to in the traditional approach, an overall improvement of just about 18% in the Performance Error indexes is obtained. And while the median correlation coefficient is also higher (0.867 instead of 0.772), 10 cases in the new PD modeling approach score a correlation coefficient inferior to 0.7. When reviewing these cases, it appeared than the main source for error is the under or over-estimation of the time delay. On a more positive note, the performances in terms of peak effect were maintained.

6.3.3 Dose *vs.* Response Characteristics

The static dose *vs.* response characteristics of propofol are presented for both approaches in Fig. 6.14. These characteristics should be equivalent since they were obtained based on the same identification data. However, they are clearly different. For instance, let us consider the propofol dose *vs.* response for a 20 years old patient. In the traditional approach, we find that a WAV_{CNS} of 40 can be obtained with a plasma concentration of 2.2 $\mu g \cdot ml^{-1}$. According to the system oriented dose *vs.* response, the same effect is obtained for a concentration of 3.3 $\mu g \cdot ml^{-1}$. This is a significant difference considering that a $C_{p,ss}$ of 2.2 $\mu g \cdot ml^{-1}$ would yield a WAV_{CNS} of 68 according to the new PD approach.

This inconsistency between the two static models can be explained by considering that un-modelled dynamics in the traditional approach yield colored residuals, which, in turn, influence the derivation of the Hill parameters (see the example in Figure 6.3). This problem is particularly acute when identification data contain mostly transitory signals. Any study using the traditional approach and transitory identification data will suffer from similar limitations. A cautionary note relative to the accuracy of the dose *vs.* response relationship should be brought to the attention of the readers in this case. With the traditional approach, a better identification method would be to use only steady state data, where the plasma concentration is fixed and the effect reaches steady state. This is, however, very difficult to obtain in a clinical setting. First, no other drug should be used concomitantly in order to limit potential synergism. Further, no surgical stress or excitation is allowed in order to maintain the effect in steady state. Finally, even with the use of a TCI pump to reach a steady state plasma concentration, each experiment would take considerable time in order to bypass the transient response.

The only study achieving this is the one conducted by Kazama *et al.* in 1999. In this study, the authors induced patients using a TCI pump programmed to reach and maintain a fixed plasma concentration. Once steady state was achieved, an incremental concentration step was programmed. Even though the identification data did contain some transitory signals, the input and output signals contained large epochs of steady state data. When comparing Kazama's dose vs. response relationships to the ones presented in Fig. 6.14.b, a number of similarities become apparent. For instance, Kazama *et al.* were the only authors who established that propofol requirements increase with age, for effects compatible with an anesthetic depth suitable for surgeries. They also showed that this characteristic is inverted (*i.e.*, older patients

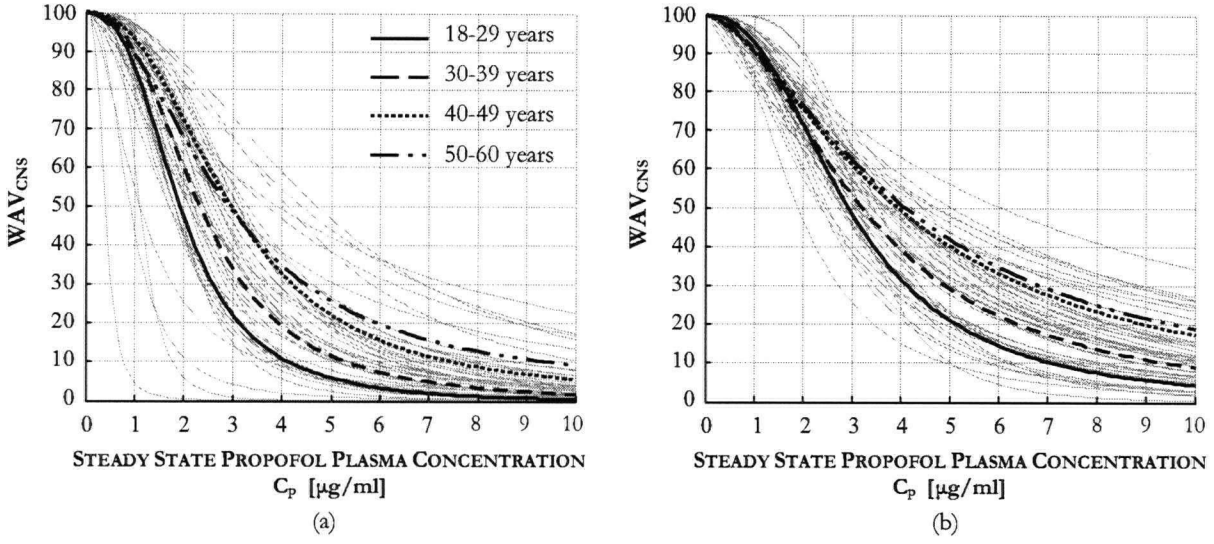


Figure 6.14: Identified WAV_{CNS} *vs.* Propofol dose-response relationships. (a) Traditional approach (b) New approach. Note that the thick lines represent averages over the corresponding age group.

become more sensitive to propofol) when only a light effect is sought. Kazama *et al.* also reported low γ values, from 2.3 to 2.0, which is consistent with our findings (from 2.5 to 1.7). Conversely, all other studies have reported much higher values (from 4.2 to 5.7). Finally, they also reported larger EC_{50} values than most studies. However, this is mainly due to the pharmacokinetic set used to drive the TCI pump. This pharmacokinetic set (from Gepts *et al.* [111]) may have over-estimated the propofol plasma concentration since it was designed for infusions (it is possible that the high rate of infusion following a step change in the pump setpoint should have been considered a bolus instead of an infusion).

6.3.4 Frequency Response Characteristic

This discussion would not be complete without presenting the frequency response of the models derived in Section 6.2.5 and represented as block diagrams in Figure 6.15.

First, we need to linearize the non-linear Hill static dose *vs.* response relationship. Let us denote by $f(x)$ the Hill function. For both approaches, $f(x)$ can be written as:

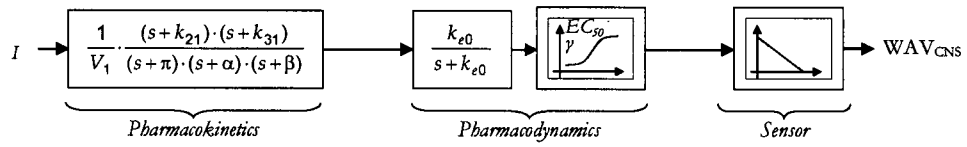
$$f(x) = \frac{x^\gamma}{X_0^\gamma + x^\gamma}, \tag{6.20}$$

where $X_0 = 0.5$ (proposed approach) and $X_0 = EC_{50}$ (traditional approach). To linearize $f(x)$, we consider small input variations $\tilde{x}$ around a given operating point $\overline{x}$. Considering further that f is a $\mathcal{L}_\infty$ function, and assuming small enough variations $\tilde{x}$, we can approximate $f(x)$ as an affine function:

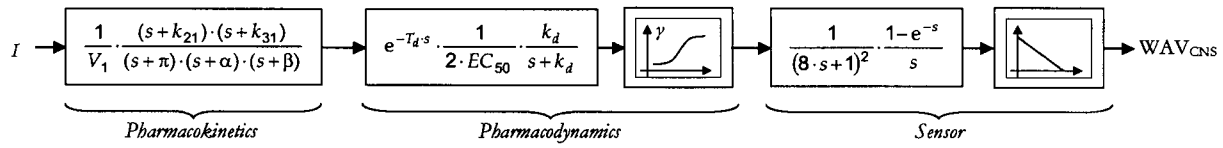
$$f(x) \approx f(\overline{x}) + K_{\overline{x}} \cdot \tilde{x}, \tag{6.21}$$

where:

$$K_{\overline{x}} = \frac{\partial}{\partial x}\left(\frac{x^\gamma}{X_0^\gamma + x^\gamma}\right)_{x=\overline{x}} = \gamma \cdot \overline{x}^{\gamma-1} \cdot \frac{X_0^\gamma}{(X_0^\gamma + \overline{x}^\gamma)^2} \tag{6.22}$$

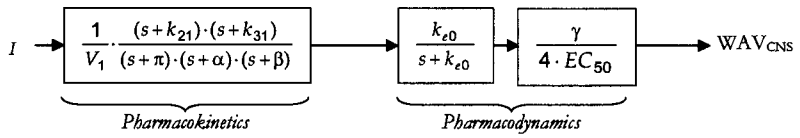


(a)

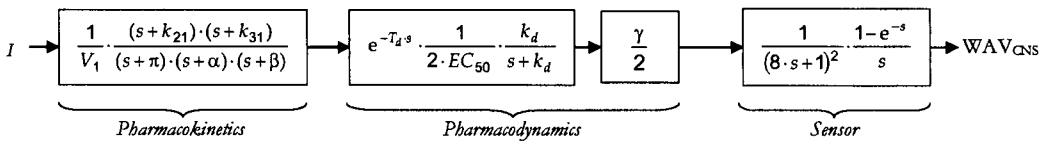


(b)

Figure 6.15: PKPD block diagrams. (a) Traditional approach. Note that the Hill saturation is characterized by both the steepness coefficient γ and the EC_{50} parameter. (b) System-oriented approach. Conversely to the traditional approach, the Hill saturation is defined uniquely by its steepness coefficient γ.



(a)



(b)

Figure 6.16: Linearized PKPD block diagrams. Note that the scaling function was removed since it does not affect the dynamic response of the system. (a) Traditional approach. (b) System-oriented approach.

The pharmacodynamic gain depends therefore directly on the operating mode $\overline{x}$ of the system. For very small ($\overline{x} \rightarrow 0$) or very large ($\overline{x} \rightarrow \infty$) doses of propofol, the gain is close to 0, indicating that variations of the propofol dose will not translate into significant changes in the observed effect. Conversely, if the system already operates in a range suitable for general anesthesia (*i.e.*, 40<WAV_{CNS} <60), the PD gain is close to its maximum.

Considering a nominal operating point set at $WAV_{CNS} = 50$, the system then operates at a steady state plasma concentration of EC_{50}. As a result, the overall pharmacodynamic gain of the system is expressed as:

$$K_{EC_{50}} = \frac{\gamma}{4 \cdot EC_{50}} \tag{6.23}$$

Equation (6.23) holds true for both traditional and system oriented models. The linearized block diagrams can then be represented as in Fig. 6.16.

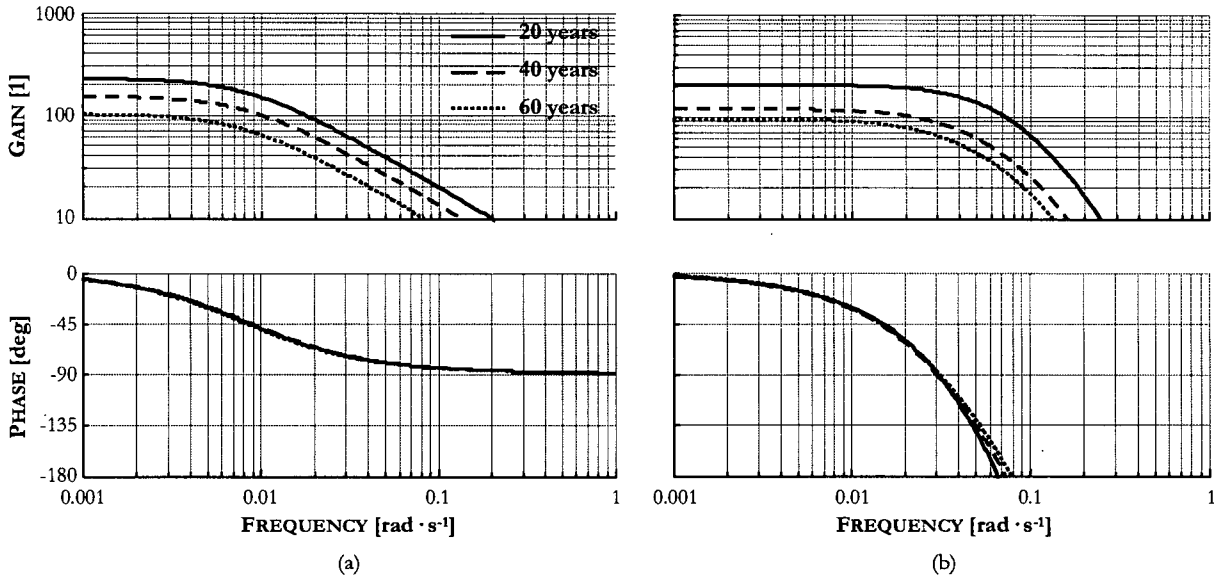


Figure 6.17: Frequency response of the PD models derived in Section 6.2.5. (a) Traditional approach (b) System-oriented approach (frequency response also models the sensor dynamics for consistency with the traditional approach).

The frequency response of the pharmacodynamic models are presented in Figure 6.17. The sensor dynamics were added to the system-oriented models for consistency with the traditional models (these models inherently contain the sensor dynamics). In both cases, age only resulted in marked steady gain differences (this was already observed in the dose *vs.* response relationships). Otherwise, age did not result in any significant differences in terms of the frequency responses.

To further compare the dynamics of the traditional *vs.* system-oriented models, we calculated the averaged frequency responses across all ages and plotted them in Figure 6.18. Beside providing a better temporal fit between measured and predicted time courses, the system oriented approach yields PD models with slightly larger bandwidth for control purposes.

This comparison also reveals that the phase margin of the system-oriented approach model drops sharply due to the time delay. Using traditional PD models for designing the controller frequency response may therefore lead to unforeseen instability if the design is too aggressive. The phase characteristic of the system-oriented model clearly motivates control designers to exercise caution when tuning the cutoff frequency of the open loop system.

6.4 Summary

Paradoxically, pharmacodynamics have been traditionally handled through static dose *vs.* response curves. However, with the advent of fast-onset short-acting drugs, a lag between effect and plasma concentration was observed and could not be accounted by a static characteristic. A model of the drug effect dynamics was required.

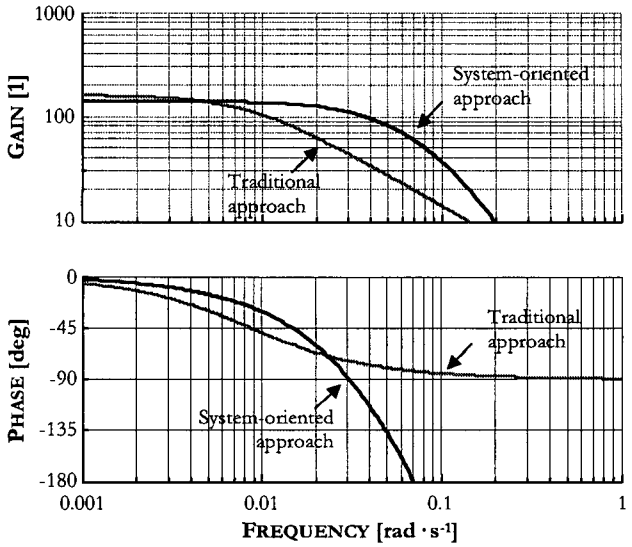


Figure 6.18: Comparison of traditional *vs.* system-oriented frequency response models (all cases averaged).

In an attempt to compensate for this lag, an addition to the mamillary pharmacokinetic compartmental model was proposed by Sheiner *et al.*. It was shown that it was possible to collapse the effect *vs.* plasma concentration 'hysteresis' by introducing a new rate constant k_{e0}. This actually resulted in the addition of a single order unity gain transfer function to the static dose *vs.* response relationship. This approach rapidly became the standard in PD studies, and has now been in use for the last 25 years.

In this chapter, we have revisited pharmacodynamic modeling using an approach based on system identification know-how. In this approach, the plasma concentration is obtained using a well defined PK model and does not necessitate blood samples. Further, the pharmacodynamic LTI function - which captures the effect dynamic - is modified in order to account for the time delay that exists between the administration of the drug and the onset of effect. Finally, the sensor dynamics is now a distinct part of the model, which allows other sensing technologies to be used in conjunction with the PD models identified using the proposed approach.

The system-oriented model yields an improved fit and more consistent performances than traditional PD models. Another advantage of the proposed approach is that the accuracy of the static dose *vs.* response relationship benefits significantly from the fact that all of the effect dynamics are now modelled through the LTI function $PD(s)$ and the sensor dynamics. As a result, even short transitory identification data (*e.g.*, induction data) can be used to identify the static dose *vs.* response relationship, thus simplifying the identification procedure in terms of clinical protocols and data acquisition (the traditional approach often requires volunteer-based studies). For instance, we show in Figure 6.19 a comparison between the static dose *vs.* response model derived in this study, and the static models derived for propofol and BIS (v.3.12) by Kazama *et al.* (note that we use here the fact that WAV_{CNS} and BIS are roughly equivalent in steady state - this equivalence was already established in Figure 5.12). The similarity between the two relationships

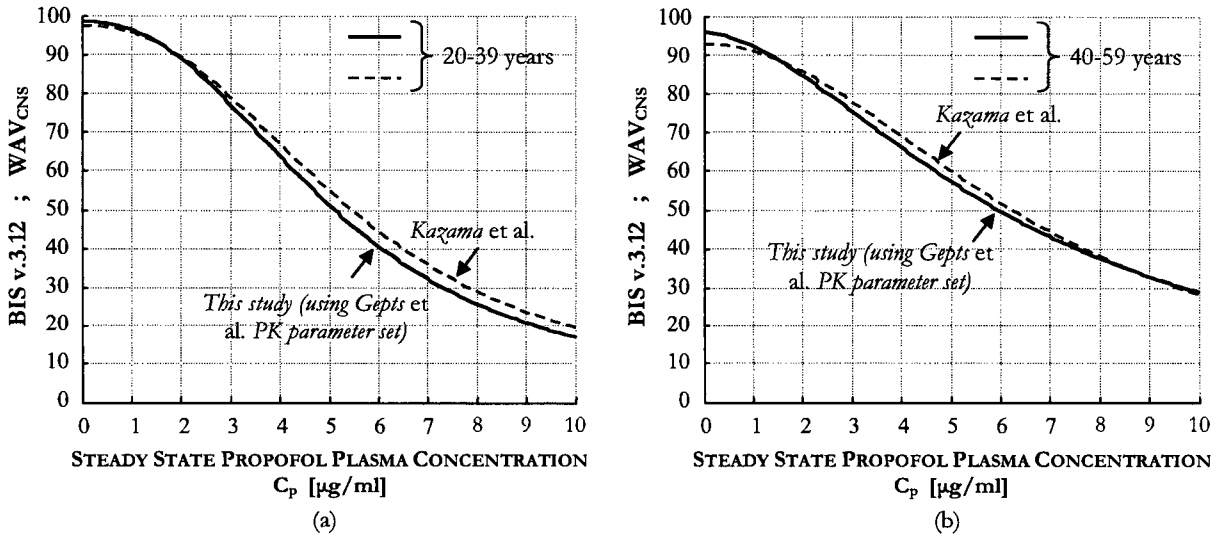


Figure 6.19: Comparison of the static dose *vs.* response relationships obtained for propofol. In their study, Kazama *et al.* used Gepts PK parameter set (see Table B.1). This set was originally derived for infusion administration. As a result, the EC_{50} found in their study are significantly different than the ones we originally found. However, this difference originates mostly from the choice of the PK set. In order to limit the influence of the PK set, we therefore re-processed all the cases using Gepts PK parameter set. We also added E_0 as an identified parameter. We found that E_0 is comprised between 0.02 (G1) and 0.06 (G4). When scaling the effect using the WAV_{CNS} scale, the E_0 values translate in a baseline WAV_{CNS} of 98 (G1) to 94 (G4). These slight differences in the overall awake baseline values were found to be also a function of age.

is remarkable, mostly when considering that Kazama's study involved for each patient an identification window of more than 90 minutes during which a TCI pump was used to target and maintain step-wise plasma concentrations (the surgery was not allowed to start during the recording of the identification data). In comparison, our method uses very limited data (2-4 minutes) which are readily available from any surgery involving limited amount of opioids and the placement of an LMA[3].

Another advantage of the proposed modeling approach concerns the reduction of system uncertainty. In particular, we will see in the next chapter (see Section 7.2.2) that the uncertainty related to the use of Sheiner's PKPD approach is higher than that of the proposed approach in the low frequency band. This implies limitation in terms of achievable performance when designing robust controllers based on quantified uncertainty.

Probably the most compelling advantage of the proposed approach is that the model structure yields non-colored residuals, indicating that it effectively models all of the system's linear dynamics. Considering the relatively poor prediction performance of population-normed models to predict the effect time course, it is likely that PD models will best be used within a close loop control framework rather than an open loop framework. Having a precise description of the frequency response of the system will then be invaluable.

[3]As compared to an endotracheal tube, the LMA limits the stimulation related to the airway management.

Chapter 7

Managing PKPD Uncertainty

Patient variability is probably the most challenging aspect in closing the loop in anesthesia. Quantifying this variability and expressing it as a *system uncertainty* is a necessity in order to prove stability when closing the loop.

The goal of this chapter is to express patient variability as a quantified uncertainty structure suitable for use with control design tools. Quantifying uncertainty allows us to assess more precisely the importance of the different sources of uncertainty in PKPD models. This can ultimately help us decide which control strategy should be implemented to reduce uncertainty and provide improved performance.

To carry out this task, a quantification and analysis method for system uncertainty is proposed and described in detail in Section 7.1. This method is based on the classical representation of uncertainty as a Nyquist disk whose radius encompasses the original uncertainty region defined by the frequency response plot of the system. This method, while conservative, appears to be the most suitable for this particular application.

In Section 7.2, we apply the proposed method to the propofol PKPD models identified in the previous chapter. In particular, we show how serious patient variability can be in terms of close loop control.

Finally, in Section 7.3, we propose a number of approaches to reduce the uncertainty to more manageable levels. Some of these approaches are rather simple and straightforward, while others require more efforts. Using the uncertainty quantification and analysis tool described in Section 7.1, we compare all methods with respect to their efficacy in reducing uncertainty *vs.* their cost in terms of design effort and practicality in a clinical setting.

7.1 Quantifying Uncertainty

When considering system uncertainty, two approaches can be envisaged:

- **Parametric (structured) uncertainty**: uncertainty can be defined by considering bounded real perturbations in some (or all) of the system parameters (gains and time constants). In this framework, it is usual to describe any given parameter τ as $\tau = \tau_0 \cdot (1 + u \cdot \Delta_s)$, where τ_0 is the nominal value of τ,

and $\Delta_s \in \mathbb{R}, \quad |\Delta_s| \leq 1$. The quantity $u = (\tau_{MAX} - \tau_{MIN})/(\tau_{MAX} + \tau_{MIN})$ is usually less than 1, and expresses how much the parameter can drift from its nominal value. In this framework, it is assumed that each parameter can take any value in their uncertainty range. Linear fractional transformations are used to represent the system under a form more suitable for robustness analysis.

- **Unstructured uncertainty**: in this framework, only the overall uncertainty in terms of the gain and phase of the system is defined. The unstructured framework is usually selected when the model structure itself is poorly defined, or when the uncertainty cannot be expressed as parametric uncertainty. Note that parametric uncertainty can also be lumped into an unstructured framework. This can lead to a more conservative design since the unstructured framework contains a large number of plants that do not belong to the original set defined by the parametric uncertainty.

When considering the PK parameters defined in Table 3.1 and expressed as in (3.10), it is possible to define for each parameter of the hybrid set $\{V_1, \pi, \alpha, \beta, k_{21}, k_{31}\}$ the corresponding parametric uncertainty bounds. This can be done using the standard error values defined Table 3.2. For instance, the PK parameters of a 30 yrs old 70 kg individual can be defined as:

$$\begin{cases} \pi = 9.4 \cdot 10^{-3} \cdot (1 + 0.40 \cdot \Delta_s^{\pi}) & [\mathrm{s}^{-1}] \\ \alpha = 4.8 \cdot 10^{-4} \cdot (1 + 0.46 \cdot \Delta_s^{\alpha}) & [\mathrm{s}^{-1}] \\ \beta = 4.0 \cdot 10^{-5} \cdot (1 + 0.51 \cdot \Delta_s^{\beta}) & [\mathrm{s}^{-1}] \end{cases} \qquad \begin{cases} k_{21} = 9.9 \cdot 10^{-4} \cdot (1 + 0.51 \cdot \Delta_s^{k_{21}}) & [\mathrm{s}^{-1}] \\ k_{31} = 7.1 \cdot 10^{-5} \cdot (1 + 0.59 \cdot \Delta_s^{k_{31}}) & [\mathrm{s}^{-1}] \\ V_1 = 9.3 \cdot (1 + 0.19 \cdot \Delta_s^{V_1}) & [\mathrm{l}] \end{cases} \tag{7.1}$$

These values were calculated considering a two-times standard error. According to this parameter set, the maximum steady state gain[1] of the PK model is $7.5 \cdot 10^6$ while the minimum gain is only $8.1 \cdot 10^4$. Even though we are considering a very limited and well defined adult patient population, note that this represents nearly a possible hundred-times difference in the steady state level of propofol plasma concentration.

This result is clearly erroneous. Parametric uncertainty assumes, indeed, that each parameter of the model can take any value from the corresponding uncertainty range. This may lead to a combination of parameters that is not possible depending on how parametric uncertainty is originally defined. For instance, the hybrid PK parameters are all function of the same θ coefficients defined in Table 3.2. Therefore, we cannot choose their values independently (*i.e.*, it is not possible to have for the same patient the minimum k_{21} and k_{31} values, while having at the same time the maximum π, α, β and V_1 values). This dependency between parameters makes here a significant difference in terms of uncertainty. For instance, when calculating all possible PK models according to the parametric uncertainty defined by Shüttler *et al.*, we find that the ratio between the maximum and minimum steady state gains is only 4.6.

In terms of the PD parameters presented in Table 6.1, parametric uncertainty may indeed be considered for k_d and EC_{50} (a student t-test reveals that these two quantities are independent). However, unless the time delay dynamics are expressed as rational LTI transfer functions (*e.g.*, using Padé approximation), the uncertainty related to the time delay cannot be directly expressed as parametric uncertainty.

[1] Steady state gain is $(k_{21} \cdot k_{31})/(V_1 \cdot \pi \cdot \alpha \cdot \beta)$.

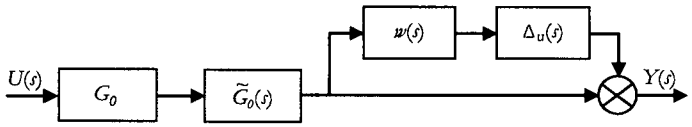


Figure 7.1: Unstructured uncertainty expressed in a multiplicative framework.

When considering PKPD uncertainty, the unstructured framework appears therefore more suitable and straightforward. We present here the multiplicative unstructured system uncertainty approach, which we will be following throughout this chapter. This approach is discussed in details in Skogestad *et al.* [167].

7.1.1 The Relative Uncertainty Framework

Let us consider a system $G(s)$ for which some uncertainty bounds exist. A common approach is to model the uncertainty using a structure such as the one of Fig. 7.1, and which can be expressed as:

$$G(j\omega) = G_0 \cdot \tilde{G}_0(j\omega) \cdot (1 + w(j\omega) \cdot \Delta_u(j\omega)), \tag{7.2}$$

where $||\Delta_u(j\omega)|| \leq 1$, $\tilde{G}_0(j\omega)$ is the normalized nominal model of the system, and $G_0 = ||G_0(j\omega = 0)||$ is the steady state gain of the system. The weight function $w(j\omega)$ quantifies the magnitude of the unstructured uncertainty, while $\Delta_u(j\omega)$ expresses this uncertainty as a unity disk in the Nyquist plot. The multiplicative uncertainty gain $w(j\omega)$ is also referred to as *relative uncertainty*.

To better understand this concept, let us consider the example of Fig. 7.2.a-b. We assume that the uncertainty bounds of $G(j\omega)$ are defined in the frequency domain and that a nominal model $G_0(j\omega)$ exists. If we map the frequency domain information into the complex Nyquist plane, the uncertainty at each frequency ω can be represented as the section of a ring centered at the origin. The outer and inner radii of the ring are defined by the uncertainty magnitudes $\tilde{G}_{\text{MIN}} = G_{\text{MIN}}/G_0$ and $\tilde{G}_{\text{MAX}} = G_{\text{MAX}}/G_0$, while the section is defined by the uncertainty angles φ_{MIN} and φ_{MAX}.

To quantify the uncertainty according to the framework of (7.2), we define the disk $\mathcal{D}$ centered on $\tilde{G}_0(j\omega)$ and of radius $l(\omega)$. This circle is the smallest circle centered on $\tilde{G}_0(j\omega)$ and which contains all of the original uncertainty surface. Following (7.2), the multiplicative weighting function $w(j\omega)$ is defined as:

$$||w(j\omega)|| = \frac{l(\omega)}{||\tilde{G}_0(j\omega)||}. \tag{7.3}$$

7.1.2 Selection of a Near Optimal Nominal Model

It is important to note that the uncertainty quantified by (7.2) considers a larger uncertainty surface than the one originally defined from the frequency domain. A limitation of the method is therefore that its representation of the frequency domain uncertainty can be very conservative, mostly if the center of the disk $\tilde{G}_0(j\omega)$ is located close to the edges of the ring section. A control design based on the uncertainty defined by the disk $\mathcal{D}(l(\omega), \tilde{G}_0(j\omega))$ may thus be unnecessarily conservative.

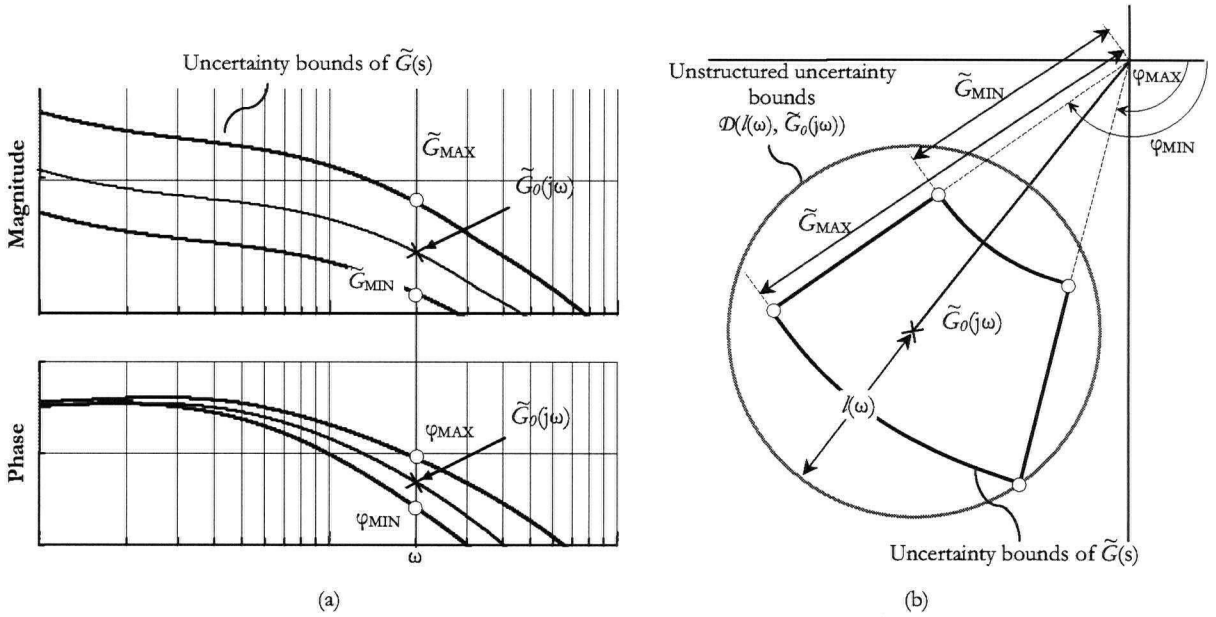


Figure 7.2: Quantifying uncertainty. (a) Uncertainty expressed in the frequency domain via Bode plots. (b) Nyquist mapping of the uncertainty at the given frequency ω.

A simple way to reducing the uncertainty is then to minimize $l(\omega)$ by selecting the best location for the center of the uncertainty disk. In other words, this leads to the optimization of the nominal model $\tilde{G}_0$, see Figure 7.3. Note that this would still be a conservative characterization of the frequency domain uncertainty.

A first step towards the optimization of the nominal model is to determine the optimal Nyquist path for $\tilde{G}_0(j\omega)$. We define the 4 following relevant coordinates:

$$C_1 = \begin{bmatrix} x_1 = \tilde{G}_{\text{MAX}} \cdot \cos(\varphi_{\text{MIN}}) \\ y_1 = \tilde{G}_{\text{MAX}} \cdot \sin(\varphi_{\text{MIN}}) \end{bmatrix} \quad C_2 = \begin{bmatrix} x_2 = \tilde{G}_{\text{MAX}} \cdot \cos(\varphi_{\text{MAX}}) \\ y_2 = \tilde{G}_{\text{MAX}} \cdot \sin(\varphi_{\text{MAX}}) \end{bmatrix}$$
$$C_3 = \begin{bmatrix} x_3 = \tilde{G}_{\text{MIN}} \cdot \cos(\varphi_{\text{MAX}}) \\ y_3 = \tilde{G}_{\text{MIN}} \cdot \sin(\varphi_{\text{MAX}}) \end{bmatrix} \quad C_4 = \begin{bmatrix} x_4 = \tilde{G}_{\text{MIN}} \cdot \cos(\varphi_{\text{MIN}}) \\ y_4 = \tilde{G}_{\text{MIN}} \cdot \sin(\varphi_{\text{MIN}}) \end{bmatrix} \tag{7.4}$$

The coordinates C_1, C_2, C_3, and C_4 mark the four corners of the ring. Depending on the phase uncertainty, we can distinguish between the three following cases:

- $(\varphi_{\text{MAX}}(\omega) - \varphi_{\text{MIN}}(\omega)) < \pi$: in this case, it is possible to find a circle $\mathcal{C}$ that circumscribes the ring section (see Figure 7.3.a). The origin C_{cir} of $\mathcal{C}$ is located at the intersection of the perpendicular bisectors of the line segments $\{C_1, C_2\}$ and $\{C_2, C_3\}$ (or $\{C_4, C_1\}$). This can be analytically expressed as:

$$C_{cir} = \begin{bmatrix} x_{cir} = \dfrac{b'' - b'}{a' - a''} \\ \\ y_{cir} = a' \cdot \dfrac{b'' - b'}{a' - a''} + b' \end{bmatrix}, \tag{7.5}$$

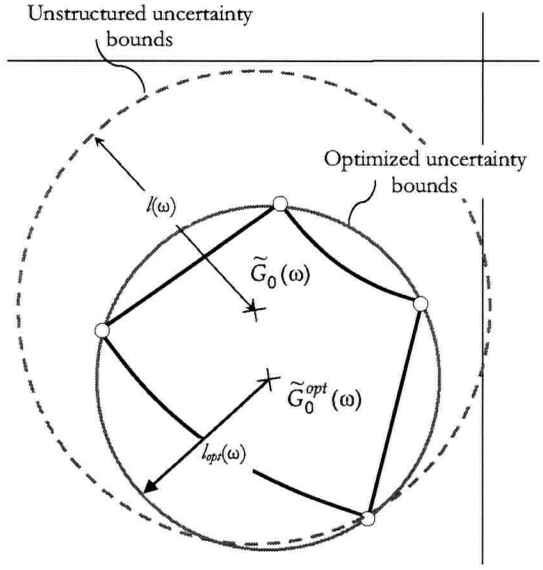


Figure 7.3: Effect of the nominal model Nyquist path with respect to the uncertainty radius: note that nominal models which are close to the uncertainty edges can significantly increase the unstructured uncertainty radius.

where:

$$\begin{bmatrix} a' \\ b' \end{bmatrix} = \begin{bmatrix} \dfrac{x_2 - x_1}{y_1 - y_2} \\ \dfrac{1}{2} \cdot \dfrac{y_1^2 - y_2^2 - x_2^2 + x_1^2}{y_1 - y_2} \end{bmatrix}, \quad \text{and} \quad \begin{bmatrix} a'' \\ b'' \end{bmatrix} = \begin{bmatrix} \dfrac{x_2 - x_3}{y_3 - y_2} \\ \dfrac{1}{2} \cdot \dfrac{y_3^2 - y_2^2 - x_2^2 + x_3^2}{y_3 - y_2} \end{bmatrix}. \tag{7.6}$$

In most cases, the circumscribing circle minimizes the uncertainty radius $l(\omega)$. However, if the phase uncertainty covers a larger span than the gain uncertainty (*i.e.*, the ring section is 'thin'), the center point of the line segment $\{C_1, C_2\}$ may yield a smaller uncertainty circle (see Figure 7.3.b). In general, we find that the optimal center is defined as:

$$\begin{cases} C_{opt} = \begin{bmatrix} x_{opt} = x_{cir} \\ y_{opt} = y_{cir} \end{bmatrix} & \text{if : } \sqrt{x_{cir}^2 + y_{cir}^2} \leq \dfrac{1}{2} \cdot \sqrt{(x_1 + x_2)^2 + (y_1 + y_2)^2} \\ C_{opt} = \begin{bmatrix} x_{opt} = (x_1 + x_2)/2 \\ y_{opt} = (y_1 + y_2)/2 \end{bmatrix} & \text{if: } \sqrt{x_{cir}^2 + y_{cir}^2} > \dfrac{1}{2} \cdot \sqrt{(x_1 + x_2)^2 + (y_1 + y_2)^2} \end{cases} \tag{7.7}$$

- $\pi < (\varphi_{\text{MAX}}(\omega) - \varphi_{\text{MIN}}(\omega)) < 2\pi$: when the phase uncertainty reaches 180 degrees, (7.5) yields the center of the circle that circumscribes the complimentary ring section (see Figure 7.3.c). The center of the minimizing circle is opposite to C_{cir} and located on the inner radius of the ring:

$$C_{opt} = \begin{bmatrix} x_{opt} = -\dfrac{\tilde{G}_{\text{MIN}}}{\sqrt{x_{cir}^2 + y_{cir}^2}} \cdot x_{cir} \\ y_{opt} = -\dfrac{\tilde{G}_{\text{MIN}}}{\sqrt{x_{cir}^2 + y_{cir}^2}} \cdot y_{cir} \end{bmatrix} \tag{7.8}$$

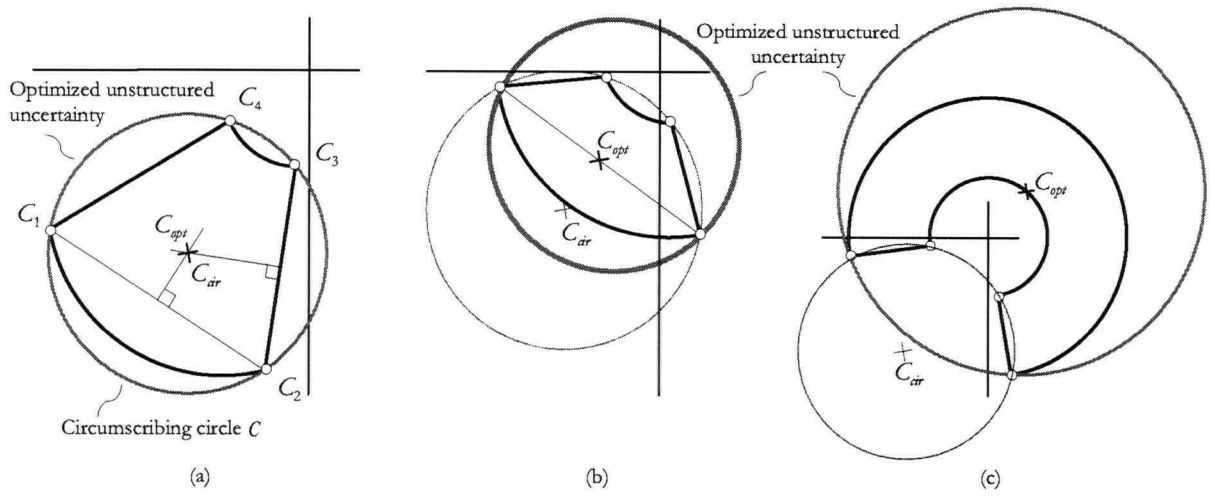


Figure 7.4: Relocation of the nominal model $\tilde{G}_0$ to the location $\tilde{G}_0^{opt}$ to reduce the uncertainty radius. (a) $(\varphi_{\text{MAX}}(\omega) - \varphi_{\text{MIN}}(\omega)) < \pi$ (large gain uncertainty relative to phase): in this case the circumscribing circle minimizes the uncertainty disc. (b) $(\varphi_{\text{MAX}}(\omega) - \varphi_{\text{MIN}}(\omega)) < \pi$ (small gain uncertainty relative to phase): in this case, the minimizing uncertainty disc center is located on the center-point of the $\{C_1, C_2\}$ segment. (c) $(\varphi_{\text{MAX}}(\omega) - \varphi_{\text{MIN}}(\omega)) > \pi$: in this case, the minimizing uncertainty disc center is located on the inner radius of the ring section, directly opposite to the center of the circle which circumscribes the complementary ring section.

- $(\varphi_{\text{MAX}}(\omega) - \varphi_{\text{MIN}}(\omega)) \geq 2\pi$: in this trivial case, the ring section becomes a complete ring. As such, the center of the circle can be located anywhere on the inner edge of the ring. For convenience sake, we adopt the following convention:

$$C_{opt} = \begin{bmatrix} x_{opt} = \tilde{G}_{\text{MIN}} \cdot \cos\big((\varphi_{\text{MIN}} + \varphi_{\text{MIN}})/2\big) \\ y_{opt} = \tilde{G}_{\text{MIN}} \cdot \sin\big((\varphi_{\text{MIN}} + \varphi_{\text{MIN}})/2\big) \end{bmatrix} \tag{7.9}$$

Note that this analysis is carried out for each frequency ω of interest. This results in the definition of an optimal Nyquist path $C_{opt}(j\omega)$.

Finding the Nyquist path describing the optimal nominal model is a first step. For control purposes, it is then necessary to express this path into a LTI transfer function which can be used in the design process. In most cases, it is not possible to find a rational LTI function that exactly describes the selected Nyquist path. Instead, a frequency domain identification can be carried out to determine a near optimal solution.

The realization of the near optimal nominal model $\tilde{G}_0^{opt}(j\omega)$ can be done in a number of ways:

- **Nominal parameters tuning**: this tuning aims at minimizing the difference between the nominal Nyquist path and the optimal path by selecting the most appropriate model parameters. This yields a realization of the same order than the original nominal model.

- **Weight function**: a LTI function $w_{nom}(j\omega)$ is tuned in order for $\tilde{G}_0^{opt}(j\omega) = w_{nom}(j\omega) \cdot \tilde{G}_0(j\omega)$ to be as close as possible from the optimal Nyquist path. This approach yields a nominal model $\tilde{G}_0^{opt}(j\omega)$ whose order is greater than that of the initial nominal model.

- **Hybrid tuning**: a hybrid realization, where a combination of both methods are used, can also be investigated. However, this already requires much more efforts in terms of the identification procedure.

Due to the large numbers of parameters in the initial PKPD nominal model, the tuning of a weight function $w_{nom}(j\omega)$ appears to be a more attractive method to optimize the nominal Nyquist path.

A general approach to identifying $w_{nom}(j\omega)$ is to minimize the difference between $\tilde{G}_0^{opt}(j\omega)$ and $C_{opt}(j\omega)$ directly in the Nyquist complex plane by minimizing the cost function $\mathcal{J}_{NYQ}$:

$$\mathcal{J}_{NYQ}^2 = \int_0^\infty \lambda_{nom}(\omega) \cdot \left(||w_{nom}(j\omega)||^2 - 2\frac{||C_{opt}(j\omega)||}{||\tilde{G}_0(j\omega)||} \cdot ||w_{nom}(j\omega)|| \cdot \cos\Delta\varphi(\omega) \right) \cdot d\omega, \tag{7.10}$$

where $\Delta\varphi(\omega) = \arg(\tilde{G}_0(j\omega)) - \arg(C_{opt}(j\omega)) + \arg(w_{nom}(j\omega))$. The parameter $\lambda_{nom}(\omega)$ is a frequency dependent weight which can be used to improve the agreement between $||\tilde{G}_0^{opt}(j\omega)||$ and $C_{opt}(j\omega)$ in the frequency band of interest (for instance, a good agreement in the controller bandwidth is more important than in the high frequency range). In practice, $\lambda_{nom}(\omega) = 1$ in the lower frequency range and $\lambda_{nom}(\omega) = 0$ for frequencies for which the system nominal phase reaches values below -2π, or well beyond the expected controller cutoff frequency.

In systems presenting a large time delay uncertainty, and whose nominal model is already of order >2, minimizing the Nyquist cost function $\mathcal{J}_{NYQ}$ may be computationally demanding as the number of parameters of $w_{nom}(j\omega)$ may be large. A more effective method is to carry out the identification by first minimizing the gain difference between $\tilde{G}_0^{opt}(j\omega)$ and $C_{opt}(j\omega)$ by tuning only the poles and zeros of $w_{nom}(j\omega)$. This can be done by minimizing the cost function $\mathcal{J}_{nom}$:

$$\mathcal{J}_{nom}^2 = \int_0^\infty \lambda_{nom}(\omega) \cdot \left| ||w_{nom}(j\omega)|| - \left|\left| \frac{C_{opt}(j\omega)}{\tilde{G}_0(j\omega)} \right|\right| \right|^2 \cdot d\omega, \tag{7.11}$$

In order to obtain also a close agreement in terms of the phase, we add to $w_{nom}(j\omega)$ a time delay $e^{j \cdot T_{opt} \cdot \omega}$ calculated to minimize the phase difference. This simplified approach can only be done if the phase of $e^{j \cdot T_{opt} \cdot \omega}$ is predominant to the phase of $w_{nom}(j\omega)$.

7.1.2.1 Realization of the Uncertainty Weight

Once the near optimal nominal model is identified, the new uncertainty radius $l(\omega)$ can be easily calculated as:

$$\begin{cases} l(\omega) = \sqrt{(x_0^{opt}(\omega) - x_1(\omega))^2 + (y_0^{opt}(\omega) - y_1(\omega))^2} & \text{if: } \varphi_{\text{MAX}}(\omega) - \varphi_{\text{MIN}}(\omega) < 2\pi \\ l(\omega) = \tilde{G}_{\text{MIN}}(\omega) + \tilde{G}_{\text{MAX}}(\omega) & \text{if: } \varphi_{\text{MAX}}(\omega) - \varphi_{\text{MIN}}(\omega) > 2\pi \end{cases}, \tag{7.12}$$

where:

$$\begin{cases} x_0^{opt}(\omega) = \|\tilde{G}_0^{opt}(\omega)\| \cdot \cos\big(\arg(\tilde{G}_0^{opt}(\omega))\big) \\ y_0^{opt}(\omega) = \|\tilde{G}_0^{opt}(\omega)\| \cdot \sin\big(\arg(\tilde{G}_0^{opt}(\omega))\big) \end{cases}. \tag{7.13}$$

The resulting uncertainty weight magnitude $\|w(j\omega)\|$ is directly calculated as the ratio between $l(\omega)$ and the new nominal model $\tilde{G}_0^{opt}(j\omega)$.

Once $\|w(j\omega)\|$ is determined, an identification procedure is carried out to express $w(j\omega)$ as a rational linear transfer function $\hat{w}(j\omega)$. This transfer function will be used in the design process to optimize the stability of the controller. The parameters of $\hat{w}(j\omega)$ are estimated by minimizing the criteria $\mathcal{J}_w$:

$$\mathcal{J}_w^2 = \int_0^\infty \Big[1 + \lambda_w(\omega) \cdot \text{sign}(\epsilon(\omega))\Big] \cdot \epsilon(\omega)^2 d\omega, \tag{7.14}$$

where $\epsilon(\omega) = ||w(j\omega)|| - ||\hat{w}(j\omega)||$, $\lambda_w >> 1$, and:

$$\begin{cases} \text{sign}(\epsilon(\omega)) = 0 & \text{if } \epsilon(\omega) \leq 0 \\ \text{sign}(\epsilon(\omega)) = 1 & \text{if } \epsilon(\omega) > 0 \end{cases} \tag{7.15}$$

Equation (7.14) represents essentially a least square minimization criterion. Note that the phase of $\hat{w}(j\omega)$ does not bear any relevance.

The parameter $\lambda_w(\omega)$ simply adds a penalty weight on the realizations $\hat{w}(j\omega)$ whose magnitude at some frequency ω is less than $||w(j\omega)||$. Note that it is sometime necessary to have $\lambda_w(\omega)$ as a function of ω when the uncertainty in the higher frequency range is large. In this case, we would choose $\lambda_w(\omega)$ and $\mathcal{J}_w$ such that:

$$\mathcal{J}_w^2 = \int_0^{10\cdot\omega_c} \Big[1 + \lambda \cdot \text{sign}(\epsilon(\omega))\Big] \cdot \epsilon(\omega)^2 d\omega, \text{ where: } \begin{cases} \lambda_w(\omega) = 0 & \text{if } \omega < 0.1 \cdot \omega_c \\ \lambda_w(\omega) >> 1 & \text{if } \omega \in [0.1 \cdot \omega_c; 10 \cdot \omega_c] \end{cases} \tag{7.16}$$

where ω_c is the cutoff frequency of the controller. This choice essentially guarantees that the realization $\hat{w}(j\omega)$ captures adequately the system uncertainty within the controller bandwidth. Assuming a large phase margin in the low frequency region, the uncertainty is less of an issue for $\omega < 0.1 \cdot \omega_c$. Finally, high frequency uncertainty is also not an issue since the cutoff frequency ω_c of the controller guarantees a low gain in this frequency range.

7.1.2.2 Comparing Uncertainty Between Systems

Note that any system presenting a relative uncertainty gain above 1 cannot be tightly controlled since the uncertainty disk contains the Nyquist origin. Therefore, there exists a particular plant for which the open loop gain can be 0, meaning that no control action can have any effect on the output.

More generally, it can be shown that Robust Stability (RS) and Robust Performance (RP) are guaranteed if:

$$\begin{cases} (\text{RS}) \Leftrightarrow ||w(j\omega) \cdot T(j\omega)|| < 1 \\ (\text{RP}) \Leftrightarrow ||w_p(j\omega) \cdot S(j\omega)|| + ||w(j\omega) \cdot T(j\omega)|| < 1 \end{cases}, \tag{7.17}$$

where $w_p(j\omega)$ is the performance weight, and S and T are the sensitivity and complimentary sensitivity functions of the close loop system. Clearly, low $||w(j\omega)||$ values will benefit to the overall performance of the system. Hence, we can compare the effect of system uncertainty between two systems S1 and S2 by comparing directly their relative uncertainty weight: a system S1 presents less uncertainty than a system S2 in a frequency bandwidth BW if $\|\hat{w}_{S1}(j\omega)\| < \|\hat{w}_{S2}(j\omega)\|$, for $\omega \in$ BW.

7.2 Application to Propofol PKPD Models

In this section, we calculate the relative uncertainty of the propofol PKPD models derived in the previous chapter. As a first step, it is necessary to clearly identify the sources of PKPD uncertainty to adequately account for them in our calculations and analysis.

7.2.1 Origins of PKPD Uncertainty

In PKPD modeling, it is customary to distinguish between two different types of uncertainty: the uncertainty caused by inter-patient variability (*i.e.*, the variability observed between different individuals), and the uncertainty originating from intra-patient variability (*i.e.*, the variability observed within one particular individual).

7.2.1.1 Inter-patient Variability

Inter-patient variability was already discussed in Chapter 3. It affects both the pharmacokinetics and pharmacodynamics of any given drug. For instance, it has been shown that age as well as weight, lean body mass, ethnicity, *etc.*, are all factors of PKPD variability in humans. Co-existing illnesses involving either the liver and/or kidneys may also significantly alter the way drugs are metabolized and eliminated from the body.

In general, 2 patients with similar physiological characteristics (age, weight, lean body mass, ASA) may have largely different PKPD parameters. For instance, patient #15 in Table 6.1 (female, 21 yrs old, 53 kg, 157 cm, ASA I) and patient #53 (female, 21 yrs old, 67 kg, 163 cm, ASA I) have significantly different PK time delay (45 sec *vs.* 4 sec), EC_{50} parameter (3.8 μg/ml *vs.* 2.3 μg/ml), and saturation characteristics (Hill steepness of 1.2 *vs.* 2.5).

Inter-patient variability can be easily characterized by considering the differences between PKPD models obtained over a large population of patients. In particular, Table 6.1 provides a good representative sample of an adult population with respect to the response to propofol administration. The parametric variability observed between the different PKPD models presented in this table is summarized in Table 7.1.

Inter-patient variability clearly plays a prominent role in the overall system uncertainty. For instance, there is a significant difference in the PK time delay and PD time constant between patients. Also, while the EC_{50} variability is more limited, there is still a 6-times difference in terms of the overall PKPD steady state gain[2].

7.2.1.2 Intra-patient Variability

Intra-patient variability expresses the variability observed in the drug response within one particular subject. This variability originates from different factors.

[2]The overall PKPD gain regroups both PK and PD steady state gains, as well as the linearized Hill gain (operating point: WAV_{CNS}=50).

Table 7.1: Propofol PKPD Inter-patient Variability

	T_d				K_d				EC_{50}				γ			
	MIN [s]	MAX [s]	$T_{d,0}$ [s]	u [%]	MIN [s^{-1}]	MAX [s^{-1}]	$K_{d,0}$ [s^{-1}]	u [%]	MIN [μg/ml]	MAX [μg/ml]	$EC_{50,0}$ [μg/ml]	u [%]	MIN [1]	MAX [1]	γ_0 [1]	u [%]
AGE GROUP																
G1	4	45	24.5	83.7	25.0	160.5	92.7	73.0	1.9	3.8	2.8	33.3	1.9	4.7	3.3	42.4
G2	1	44	22.5	95.6	24.8	83.1	53.9	54.0	2.8	4.0	3.4	17.6	1.2	2.7	1.9	38.5
G3	2	35	18.5	89.1	28.7	38.0	33.3	13.9	3.0	6.0	4.5	33.3	1.3	2.3	1.8	27.8
G4	2	29	15.5	87.1	21.8	58.0	39.9	45.4	3.1	5.1	4.1	24.4	1.1	2.3	1.7	35.3
POPULATION	*1*	*45*	*23*	*95.6*	*21.8*	*160.5*	*91.1*	*76.1*	*1.9*	*6.0*	*4.0*	*51.9*	*1.1*	*4.7*	*2.9*	*62.0*

Drug administration It is a well-documented fact that the pharmacokinetics of intravenous agents differ depending on the method of administration of the drug. Even though bolus and infusion PK models have the same steady state gain, the initial peak plasma concentration following a bolus administration is significantly over-predicted by the corresponding infusion model (see for instance Figure 3.4).

During steady state (and for small setpoint changes and/or disturbances), it is likely that the controller will administer propofol at an infusion rate inferior to 0.5 $mg{\cdot}min^{-1}{\cdot}kg^{-1}$. In this range, it is expected that the propofol pharmacokinetics will be accurately described by the infusion model. However, during large transients, the controller may have to output infusion rates above 1 $mg{\cdot}min^{-1}{\cdot}kg^{-1}$, in which case the propofol uptake and distribution may follow the behavior observed for bolus regimen. If the controller output is not constrained to infusion rates inferior to 0.5 $mg{\cdot}min^{-1}{\cdot}kg^{-1}$, the controller design must therefore account for the difference in dynamics between the bolus and infusion PK models. This difference in models can be expressed as system uncertainty by associating to each case presented in Table 6.1 the two possible PK frequency responses.

Controller Setpoint In Section 6.3.4, we commented on the fact that the Hill saturation may be viewed as a gain that is dependent on the operating point of the system. In terms of the close loop application, it is desired to maintain control over a wide range of WAV_{CNS} values (*e.g.*, from 80 to 20). As a result, for each PD model in Table 6.1, we can reduce the Hill equation as a gain K bounded between two values K_{max} and K_{min} (see Figure 7.5), and defined as:

$$K_{max} = \max\{K_{\overline{x}}, \quad \overline{x}_{min} \leq \overline{x} \leq \overline{x}_{max}\} \quad \text{and} \quad K_{min} = \min\{K_{\overline{x}}, \quad \overline{x}_{min} \leq \overline{x} \leq \overline{x}_{max}\}, \tag{7.18}$$

where:

$$\overline{x}_{min} = \frac{1}{2} \cdot \sqrt[\gamma]{\frac{E_{min}}{1 - E_{min}}} \quad \text{and} \quad \overline{x}_{max} = \frac{1}{2} \cdot \sqrt[\gamma]{\frac{E_{max}}{1 - E_{max}}}, \tag{7.19}$$

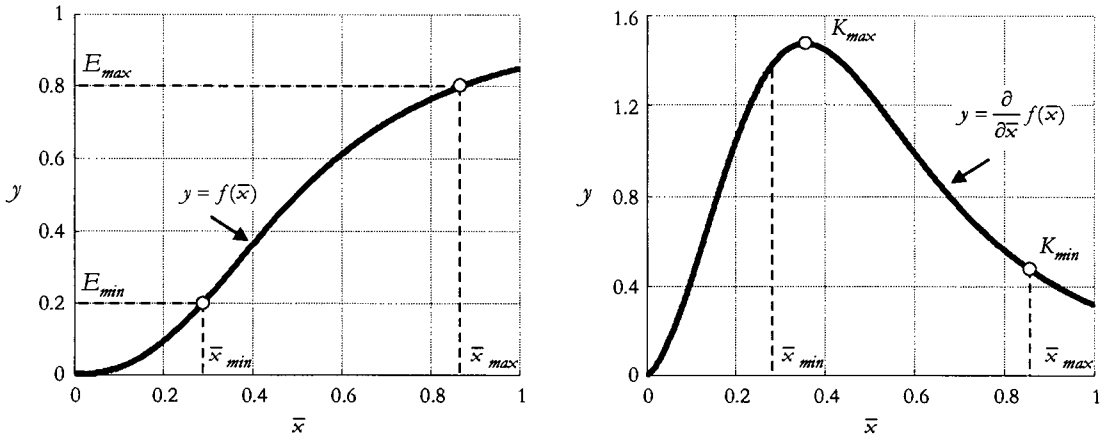


Figure 7.5: Linearization of the Hill equation.

and where $E_{min} = 0.2$ (WAV_{CNS}=80%) corresponds to the smallest desired effect (shallow sedation), and $E_{max} = 0.8$ (WAV_{CNS}=20%) corresponds to the strongest desired control setpoint[3].

The Hill saturation being simplified as a gain $K \in [K_{min}; K_{max}]$, we can therefore associate for each case in Table 6.1 two frequency responses corresponding to the minimum and maximum Hill gains.

Physiological Time Variance Physiological processes evolve in time (*e.g.*, circadian rhythms, hemodynamic changes, hormonal fluctuations, *etc.*). Alterations in these processes can affect drug distribution, elimination, and effect, which in turn can result in large intra-patient PKPD parameter variability.

In an attempt to quantify intra- *vs.* inter-patient parametric variability, we carried out in 2002 a clinical study involving 5 patients receiving electro-convulsive shock therapy (ECT)[4]. These patients received a total of 6 treatments given over a couple of months. Each treatment consisted in the administration of a single thiopental induction dose to provoke a rapid loss of consciousness before the application of the electric shock. Similarly to the LMA data, the WAV_{CNS} induction time courses were used to derive the PD model for thiopental. The PK part of the model was derived based on a published thiopental PK parameter set. The model parameters derived during these multiple repeats are presented in Table D.6 and Table D.7, and expressed as parametric uncertainty in Table 7.2.

Results from this ECT study show that intra-patient variability is limited as compared to inter-patient variability. For instance, when plotting the time courses of the WAV_{CNS} for each patient, we can clearly see that the PK time delay is consistent for each individual patient, but can vary significantly between different patients. (A representative example is plotted in Figure 7.6.a.) In this study, we found that intra-patient parametric uncertainty reaches 30% at most for the PK time delay, while inter-patient uncertainty can be as high as 70%. Similar results were found for the overall PKPD gain. These results are summarized in the Figure 7.6.b. While expected, these results also show that the WAV_{CNS} can be used to discriminate

[3]Such a deep effect may be desired prior (and during) extremely stimulating surgical acts.

[4]The protocol of this study and the identification procedure are presented in more details in the Appendix D.3.

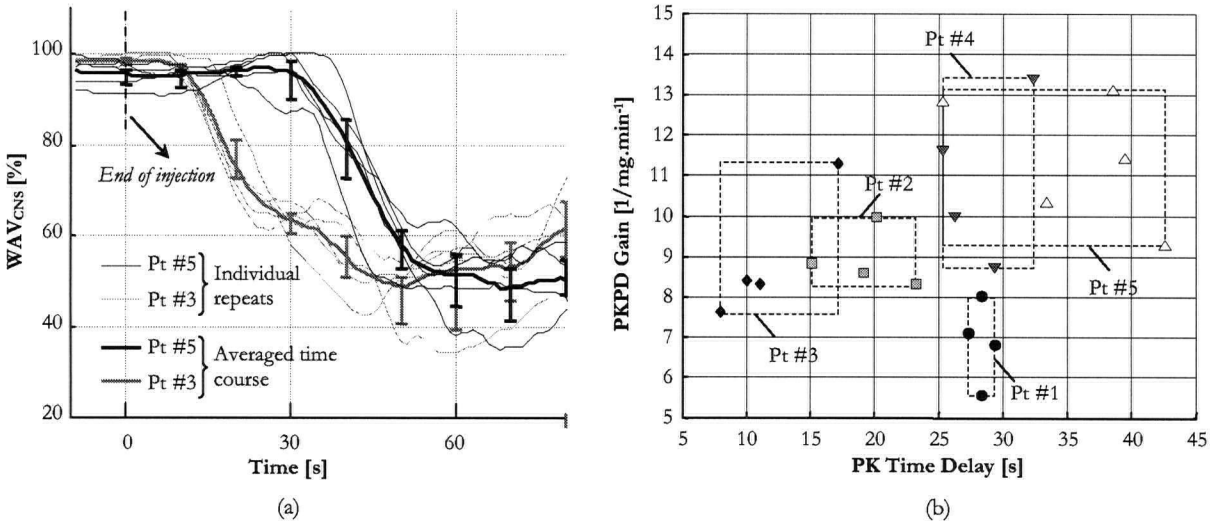


Figure 7.6: Intra- and inter-patient variability during thiopental induction. (a) WAV_{CNS} time courses for 2 patients. Each repeat is synchronized with the end of thiopental injection. (b) Identified PKPD gains and PK time delays.

Table 7.2: ECT Intra- and Inter-patient Variability

	T_d				K_d				EC_{50}				γ			
	MIN [s]	MAX [s]	$T_{d,0}$ [s]	u [%]	MIN [s^{-1}]	MAX [s^{-1}]	$K_{d,0}$ [s^{-1}]	u [%]	MIN [$\mu g/ml$]	MAX [$\mu g/ml$]	$EC_{50,0}$ [$\mu g/ml$]	u [%]	MIN [1]	MAX [1]	γ_0 [1]	u [%]
PATIENT																
#1	23	29	26.0	11.6	17.4	67.9	42.7	59.2	17.2	22.1	19.7	12.5	1.4	1.6	1.5	6.7
#2	15	23	20.5	12.2	30.5	46.2	38.4	20.5	14.5	15.9	15.2	4.6	1.5	1.8	1.7	9.1
#3	8	17	13.5	25.9	48.5	81.5	65.0	25.4	15.3	21.5	18.4	13.8	1.3	1.6	1.5	10.3
#4	25	45	35.0	28.6	26.0	39.8	32.9	20.1	16.4	21.4	18.9	13.2	1.6	2.1	1.8	13.5
#5	25	42	33.5	25.4	37.1	214.0	125.5	70.5	20.2	28.2	10.2	16.5	1.2	1.5	1.4	7.1
POPULATION	8	45	26.5	69.8	17.1	214.0	115.5	85.2	14.5	28.2	21.4	32.1	1.2	2.1	1.6	27.3

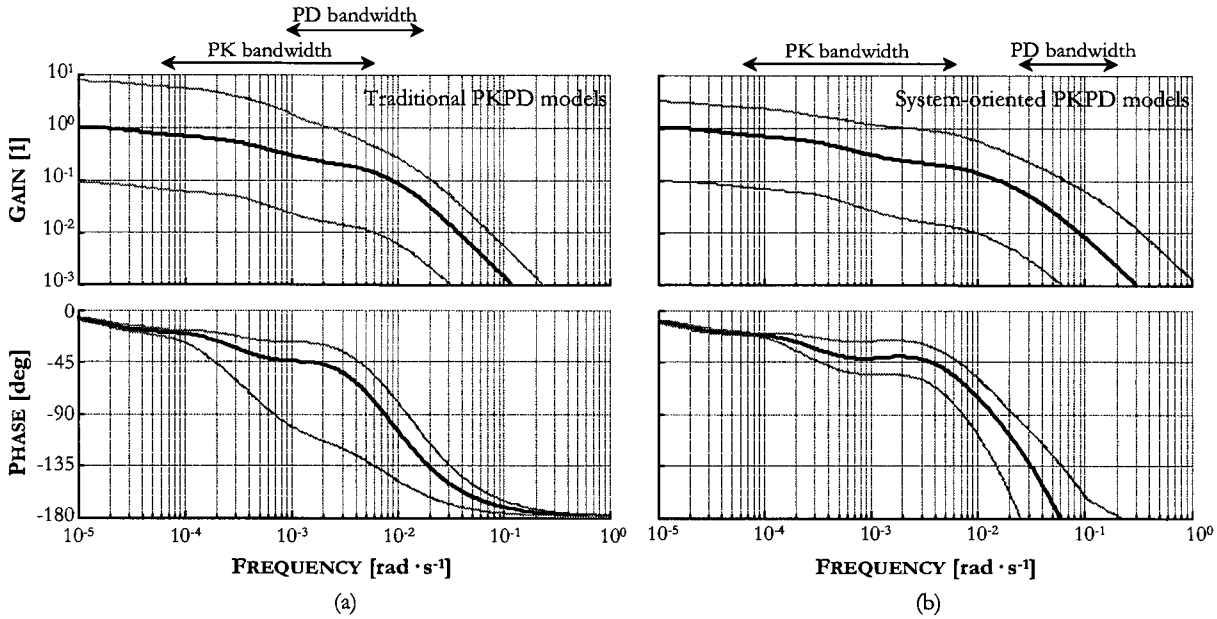


Figure 7.7: Normalized Propofol PKPD uncertainty bounds expressed in the frequency domain. The PK part of the models was derived based on the published PK parameter sets of Schüttler *et al.* (a) Based on Traditional PD models (b) Based on system-oriented PD models (does not include the sensor dynamics).

between the need of each individual in terms of drug administration.

In terms of propofol and the LMA study, intra-patient variability due to physiological time variance cannot be directly inferred from the data of Table 6.1. However, it is likely that the large number of patients considered in the propofol study captures most of the potential intra-patient parametric uncertainty due to physiological time variance (note, for instance, that the population parametric uncertainty of the LMA study is larger than that of the ECT study).

7.2.2 Propofol PKPD Uncertainty Results

In order to apply the methodology presented in Section 7.1 for the characterization of propofol PKPD uncertainty, it is first necessary to derive the uncertainty bounds in the frequency domain. To do so, we consider the PD models of Table 6.1. For each case, we derive four frequency responses corresponding to both infusion and bolus PK models, and the two Hill gains K_{MIN} and K_{MAX}. The propofol PKPD frequency bounds for the whole study population are presented in Figure 7.7. For comparison, we also included the uncertainty bounds obtained based on the PD models obtained from the traditional modeling approach. Note also that the bounds were normalized based on the steady state gain of the nominal PKPD model[5].

A direct observation is that the traditional modeling approach yields significantly more uncertainty in the lower frequency range. In particular, the uncertainty in terms of the phase is considerably larger. One explanation is that the PD model bandwidth intersects with part of the PK bandwidth (the fast distribution

[5]In this case, the nominal PKPD model was obtained by averaging the PKPD parameters over the study population.

and redistribution time constants π and k_{21} are in the same order than the PD time constant k_{e0}). As a result, both PK and PD uncertainties add up and result in the large phase uncertainty observed in the $[4 \cdot 10^{-4}, 1 \cdot 10^{-2}]$ rad·s^{-1} frequency bandwidth. Conversely, most of the PD dynamics of the proposed system-oriented approach is accounted for by the sensor dynamics which does not carry uncertainty. As a result, the identification procedure yielded a higher frequency band PD dynamic[6]. The PK and PD uncertainty do not overlap, which explains the reduced uncertainty observed in the lower frequency band. Another observation is that the steady state gain uncertainty is also larger in Sheiner's approach. This can be explained by considering the fact that the un-modelled time delay results in steeper Hill relationship in the traditional models, which in turn results in a larger $K_{\text{MAX}}/K_{\text{MIN}}$ ratio.

The relative uncertainty weights derived from the frequency response bounds are presented in Figure 7.8 for both modeling approaches.

Figure 7.8.a presents the weights obtained directly using the population-averaged PKPD nominal model $\tilde{G}_0(j\omega)$. In both cases, the weights are significantly higher than the threshold, indicating that even in the low frequency region, the complementary sensitivity function (i.e., the close loop transfer function) cannot be equal to one and has to be reduced to less than 0.5 (new approach) and 0.14 (traditional approach). This eventually would result in a 50% steady state error (resp. 86% error in the traditional model). This result clearly shows the need to improve on the selection of the nominal model in order to reduce the conservatism introduced by the methodology.

Figure 7.8.b presents the uncertainty weights in case the nominal model of the system follows the optimum Nyquist path, as defined in Section 7.1.2. These weights are significantly reduced, and actually show that the system can be robustly stabilized while guaranteing no steady state error. In this case, the controller cutoff frequency can be as high as $2 \cdot 10^{-2}$ rad·s^{-1}.

Figure 7.8.c presents the uncertainty weights obtained with the near optimal nominal model $\tilde{G}_0^{opt}(j\omega)$. These weights are significantly larger than the optimal weights. This is due to the fact that the $\tilde{G}_0^{opt}(j\omega)$ realization only approximates the optimal Nyquist path (see Figure 7.9). Also, note that the realization $\hat{w}(j\omega)$ brings additional conservatism. In case of the traditional PKPD modelling approach, it is doubtful that a robust controller presenting no steady state error can be derived.

The nominal PKPD model $G_0(j\omega)$, the weight $w_{nom}(j\omega)$ and the uncertainty weight $\hat{w}(j\omega)$ are summarized as transfer functions in the Table C.1 and Table C.4 (model #1). Note that a 4^{th} order realization was necessary for both $w_{nom}(j\omega)$ and $\hat{w}(j\omega)$ in order to obtain satisfactory results (the need for such a high order realization is motivated by considering the large frequency spread of the different PKPD time constants).

[6]Note that the time constant k_d is one order of magnitude higher than k_e0.

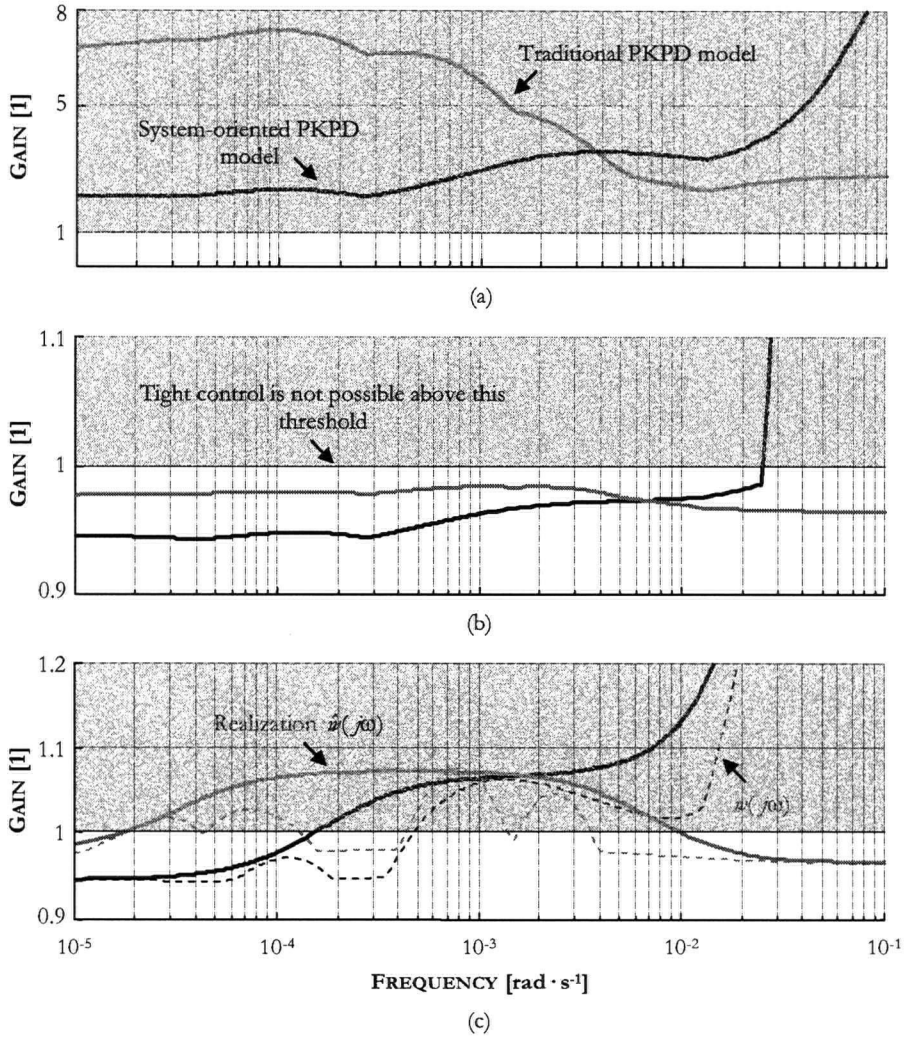


Figure 7.8: Propofol PKPD uncertainty weights for the complete adult population (drug administration: infusion and bolus ; age group: 18-60 yrs old ; WAV_{CNS} range: 80 to 20). (a) Uncertainty weights calculated based on the initial $\tilde{G}_0(j\omega)$ nominal model. (b) Uncertainty weights based on an optimized Nyquist path. (c) Near optimum uncertainty weights $\|w(j\omega)\|$ (and its realization $\|\hat{w}(j\omega)\|$) obtained based on the near optimal realization $\tilde{G}_0^{opt}(j\omega)$ of the nominal model.

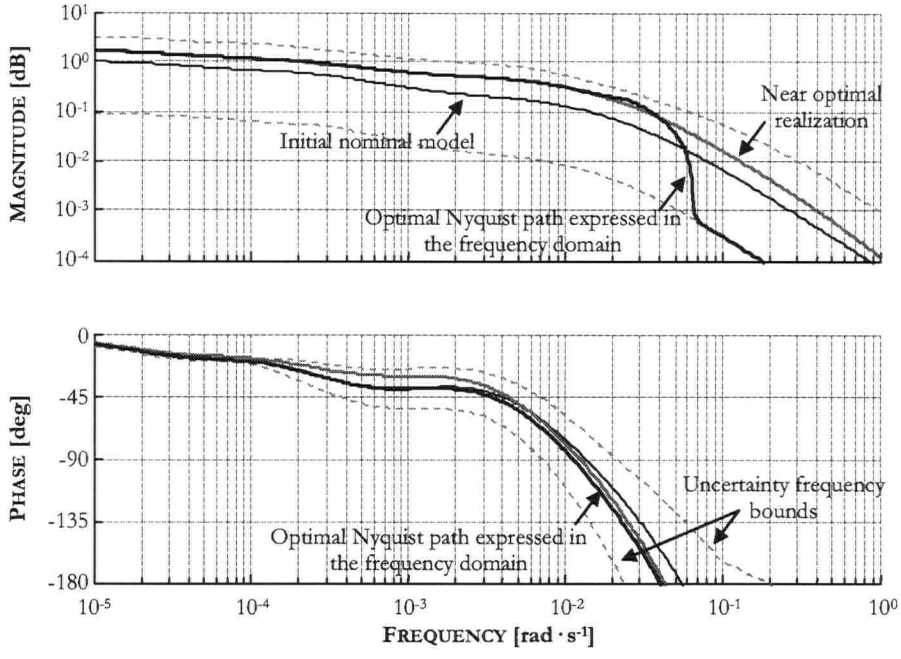


Figure 7.9: Nominal model selection in the frequency domain. Note that the rapid gain drop in the high frequency range could not be captured adequately by $w_{nom}(j\omega)$. Also, some small phase differences between the optimal model and the realization $\tilde{G}_0^{opt}$ at about $w = 1 \cdot 10^{-3}$ rad·s^{-1} results in 'bumps' which significantly increase the uncertainty gain.

7.3 Reducing Uncertainty

System uncertainty usually leads to conservative control designs in order to guarantee stability. This usually translates in limited close-loop performances. In particular, the results presented in Figure 7.8 emphasize the need to investigate methods to reduce uncertainty to more manageable levels.

In that respect, we present here 2 simple methods which can bring immediate benefits without adding significant complexity to the control design. The first method (see Section 7.3.1) aims at reducing uncertainty by simply limiting the validity of the models to a particular population subset, or operating conditions. Conversely, the second method aims at eliminating inter-patient variability by directly identifying the patient model during the initial induction phase. This self tuning can be either total or partial, in which case some inter-patient uncertainty may remain in the model. We see in Section 7.3.2 that even partial identification can provide a significant reduction in uncertainty.

Note that these methods can be combined together for optimum results.

7.3.1 Passive Methods

By 'passive methods', we refer to techniques which do not require further modeling nor identification efforts. These methods aim at reducing uncertainty by simply limiting the validity of the models to a smaller subset of population and/or operating conditions.

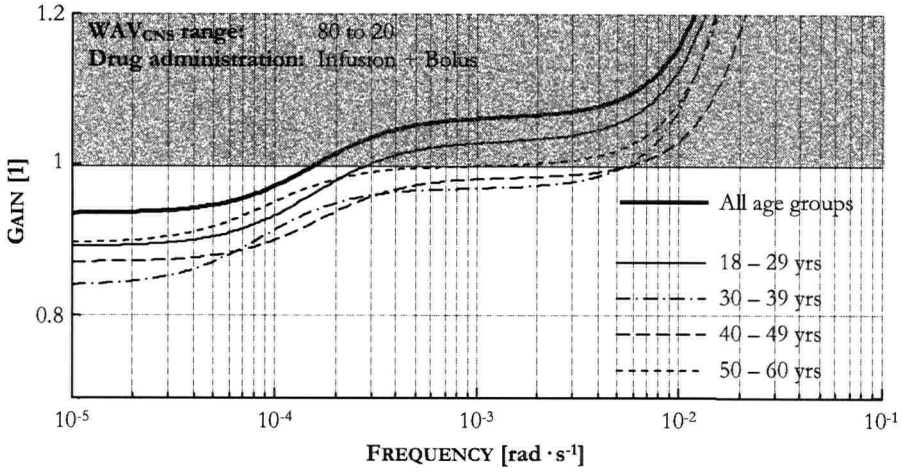


Figure 7.10: Effect of age consideration on PKPD relative uncertainty.

7.3.1.1 Using Covariates

A very simple way of minimizing inter-patient variability is to consider covariates. In particular, age has been shown to be both a PK and PD covariate. As such, considering a limited age bracket should reduce patient variability, and thus, system uncertainty.

We present in Figure 7.10 the uncertainty weights derived from the four different age groups. Each of these weights does indeed present a lower magnitude than that of the weight obtained from the whole adult population. Hence, by simply considering the age of the patient, a specific nominal model can be chosen (models #2 to #5 in Table C.1 and Table C.4) and used to derive a controller that is specific for that age group.

However, the uncertainty reduction is not significant. For instance, the uncertainty weight magnitude is reduced by only 5% in the frequency range of interest (18-29 yrs old age group). Better results were obtained for the 30-39 yrs old age group, however, it is possible that the population sample is not large enough to capture the whole inter-patient variability in this age group. This result comes as a surprise since age is usually considered to be the strongest covariate in PKPD model. Hence, this may indicate that including covariates may only have a marginal effect in reducing inter-patient variability. This very issue was raised by S.L. Shafer during the 2005 Advanced Modeling and Control in Anesthesia (AMCA, Switzerland) during his talk on PK covariates. He commented on the fact that the reduction in PK inter-patient variability by accounting for all covariates (age, weight and lean body mass) is only in the order of 5 to 10%, which is consistent with our finding.

In this analysis we selected a 10 years age bracket. It is expected that further reduction of this bracket will provide reduction in uncertainty. However, it would be necessary to acquire more cases in order to obtain more precise uncertainty bounds.

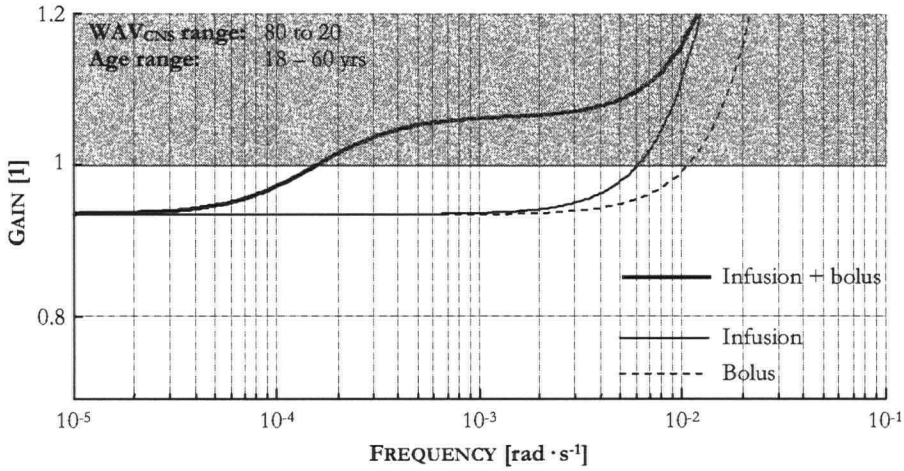


Figure 7.11: Effect of drug administration consideration on PKPD relative uncertainty. Note that the uncertainty weight for bolus administration presents only an academic interest since it is unlikely that a controller be designed to work uniquely in this mode.

7.3.1.2 Infusion *vs.* Bolus Administration

In the uncertainty defined in the frequency bounds of Figure 7.7, we included both bolus and infusion-type administrations. Considering a rather conservative (robustly stable) control design, it is expected that the control action will remain bounded in a lower range of infusion rates. The controller may therefore not be able to deliver boluses. It is also expected that additional controller constraints will limit the ability of the controller to deliver such large doses without human supervision.

Consequently, we may be able to limit the intra-patient variability by considering only infusion-type pharmacokinetics. Such a limitation of the PKPD model directly results in a smaller uncertainty weight, see Figure 7.11. In particular, the uncertainty 'bump' observed in the lower frequency range and due to differences in the drug distribution and redistribution dynamics is now eliminated.

Another advantage of limiting the drug administration to either an infusion or a bolus-based scheme is that $w_{nom}(j\omega)$ can be reduced to a second order LTI without impacting significantly the results. The corresponding PKPD nominal models, as well as the uncertainty weights, are summarized in the Table C.1 and Table C.4 (models #6 and #7).

7.3.1.3 Limiting the Control Range

Another effective way of minimizing intra-patient variability (and thus uncertainty) is to consider a limited operating range. In particular, the results presented in Figure 7.8.c were derived assuming the controller may target any WAV_{CNS} level from 80 to 20 (light sedation to deep anesthesia). It is likely that controllers tuned specifically for sedation (WAV_{CNS} range: 80-50), general anesthesia (WAV_{CNS} range: 60-30), and deep anesthesia (WAV_{CNS} range: 40-20) may have to account for less uncertainty than a general purpose controller.

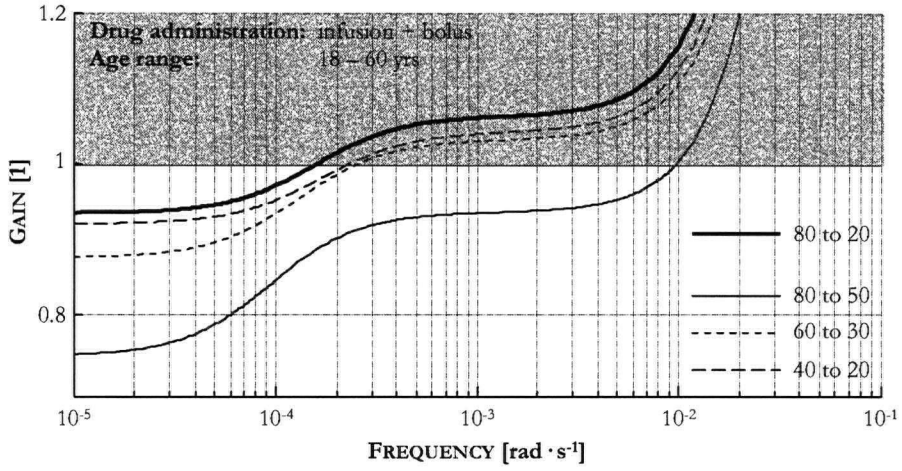


Figure 7.12: Effect of operating WAV_{CNS} range reduction on uncertainty.

This assumption is evaluated in Figure 7.12, where we present the uncertainty weights associated to each WAV_{CNS} range. In particular, it is interesting to notice that the uncertainty is significantly reduced when considering the 80-50 sedation range. This stems form the fact that the slope of the Hill saturation is quasi-constant within this range. Conversely, as the effect nears the end of the observable range, the slope of the saturation reduces considerably, which yields a large $K_{\text{MAX}}/K_{\text{MIN}}$ ratio, a major component in gain uncertainty.

The corresponding PKPD nominal models, as well as the uncertainty weights, are summarized in the Table C.2 and Table C.5 (models #8 to #10).

7.3.2 Total and Partial Self Tuning

Another recourse when faced with large system uncertainty is the online adaptation of the model parameters. This allows the controller to adapt the model parameters to the specificities of the process, while reducing significantly the uncertainty bounds. However, the difficulty of guaranteeing a proper model identification, and the high computational complexity associated to the self-tuning procedure itself, are usually considered important obstacles when implementing on-line adaptive schemes.

On-line self-tuning may therefore not be an option for this application. However, we have shown in the previous chapter that even limited induction data can be used to identify the PD parameters. Since induction is a procedure common to all anesthetic acts, sufficient identification data may always be available to identify the patient's dynamic response to the initial induction bolus. It is then likely that the information obtained during induction may be helpful to reduce some of the inter-patient variability, as it would allow the controller to focus more precisely onto the patient's PD characteristics.

This approach may substantially reduce the model uncertainty. However, as mentioned earlier, intra-patient variability originates from the inherent non-linearity of the PKPD model (infusion *vs.* bolus PK, Hill saturation), and from changes in some of the physiological processes. Therefore, even if the patient's

PKPD model is identified during induction, some uncertainty would still need to be accounted for. This section aims at quantifying the minimum intra-patient uncertainty which needs to be accounted for after self-tuning during induction.

7.3.2.1 Studying Intra-Patient Variability: the ECT Framework

The LMA data cannot be used to study the intra-patient variability which originates from changes in physiological processes. Therefore, quantifying the uncertainty due to intra-patient variability and the potential benefits of self-tuning during induction cannot be investigated based on these data.

In 2002, we carried out a clinical study aimed at providing some data on intra-patient variability. This study was principally aimed at comparing the time course of the WAV_{CNS} index during induction and following a thiopental bolus. The clinical framework which was selected to conduct this study was that of the Electro-shock Convulsive Therapy (ECT). Patients undergoing this therapy typically receive a large sedation bolus of thiopental just before the application of the electroshock. In some situations, the patients receive the treatment every couple of days for a total of 12 repeats. In other cases, the patients receive the treatment every week for as long as their condition warrants it. As a result, the ECT procedure provides a framework in which a patient will be induced multiple times in the same conditions (same drugs, dosage, time of the day, *etc.*).

The ECT study protocol, and the thiopental PKPD identification procedure, are summarized in the Annex D.3.

Study Hypothesis After ethic board review and approval, 5 patients were consented for a total of 6 repeats. After review of the induction data, this study provided 4 to 5 PKPD models for each single patient. These models were obtained within similar conditions (same drugs/dosage and anesthesia protocol) and over a 1 to 3 months period.

We hypothesize that the variability observed in the models derived from each patient is representative of the variability which can occur during the course of a specific surgery. This hypothesis is backed up by the daily clinical observation that the reaction of a patient during the induction period is a good indicator of the intra-operative requirements of that particular patient.

Study Limitations As compared to the propofol modeling study, the quality of the identification data was rather poor:

- The quality of the WAV_{CNS} induction time courses was compromised in many cases by the co-administration of succinylcholine. This NMB agent generated muscle tone (fasciculation) in some patients soon after the loss of consciousness. The corresponding increase in EMG activity elevated the WAV_{CNS}. Cases with significant fasciculation-related WAV_{CNS} increases were removed from the identification procedure.

- The data time-window was also much shorter than in the LMA study (90 to 120 seconds *vs.* 120 to 240 seconds). This short window and the use of the very fast acting thiopental drug required the use of a faster WAV_{CNS} trending filter. As a result, the induction time courses presented more measurement noise than in the LMA study. Also, in some cases, the WAV_{CNS} did not settle before the application of the ECT shock. These cases were removed from the analysis.

- Reliable bolus PK models for thiopental have not been published. Consequently, the infusion PK models published by Stanski *et al.* [168] in 1990 are used here to predict the drug plasma concentration time course. Consequently, it is expected that the PD time constant k_d does - in part - compensate for the PK modeling error introduced when using infusion PK models.

- The thiopental induction dose was mainly titrated to provoke a rapid loss of consciousness followed by a deep sedation. As a result, the WAV_{CNS} did not describe the whole 100-0 therapeutic range. The models derived here are, therefore, limited to describe effect dynamics in the sedation range.

- Finally, due to the nature of the illness for which the ECT was administered, it is likely that these patients may have also received a pharmacological therapy directly affecting the CNS, in which case some interaction with thiopental could have occurred.

These limitations may explain the fact that the ECT study yielded more inter-patient parametric variability than the LMA study (see Table 7.2 and Table 7.1).

While it is doubtful that the PKPD models derived in this study bear much clinical interest (co-administration of succinylcholine, lack of proper thiopental bolus PK model, limited therapeutic window), they give good grounds to compare the uncertainty originating from inter-patient variability *vs.* the uncertainty caused *specifically* by intra-patient variability.

7.3.2.2 Inter- *vs.* Intra-Patient Uncertainty

A protocol similar to the one outlined in Section 7.2.2 for propofol was implemented to derive the frequency domain uncertainty bounds for the whole study population (the drug administration was limited to infusion-type regimen, and the Hill saturation was linearized in the 80 to 50 WAV_{CNS} range). The resulting bounds were then used to calculate the uncertainty weight. A similar analysis was then carried out for each individual patient using the 4 (or 5) available PKPD models in order to quantify the intra-patient uncertainty.

A significant reduction in uncertainty can be observed (a least a 60% reduction in the lower frequency range, and a 3-times reduction in the higher frequency range), see Figure 7.13. Another direct advantage of working with patient-specific models is that the controller cutoff frequency can also be increased by about half a decade (the uncertainty weight crosses the '1' threshold at least half a decade higher than the population-normed uncertainty).

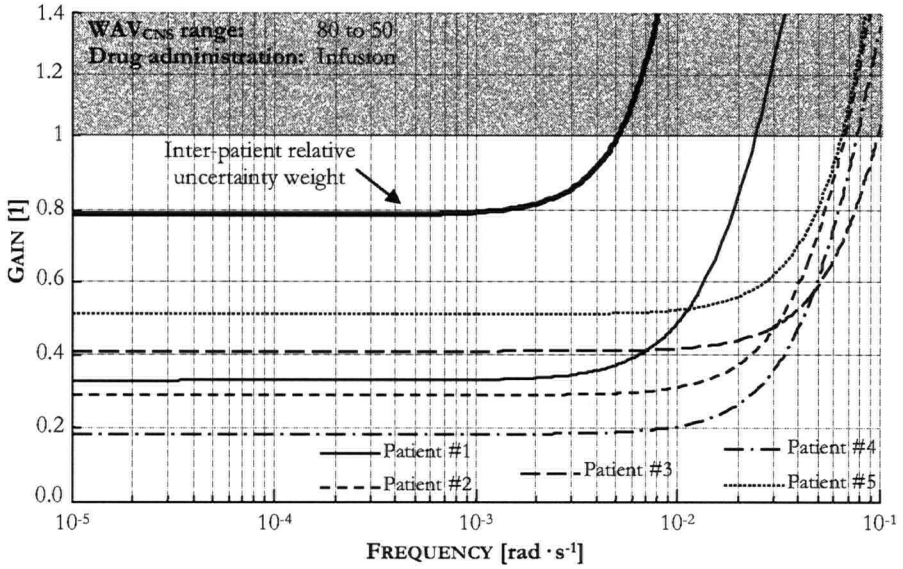


Figure 7.13: Inter- *vs.* intra-patient uncertainty (thiopental PKPD models). The patient-specific nominal PKPD models used to calculate the uncertainty weights were derived based on the 4 (or 5) PKPD models derived to calculated the uncertainty bounds. In this case, prior knowledge of each patient PKPD characteristics was necessary.

Working with patient-specific PKPD uncertainty is, however, hardly practical since it assumes that the frequency domain uncertainty bounds are already well defined (these bounds are used to calculate the nominal model and the uncertainty weight). This situation is clearly not common, since, in most cases, no PKPD information is available when patients first arrive in the operating room.

7.3.2.3 PKPD Identification during Anesthesia Induction

Since intra-patient uncertainty is significantly reduced as compared to population-normed uncertainty, it is reasonable to assume that information obtained during induction can be helpful to reduce the uncertainty.

Full PKPD Parameter Identification For instance, a simple approach which may yield immediate benefits is to first identify the PKPD model using induction data, and then replace the parameters of the generic nominal model with the newly identified PKPD parameters.

If we follow this approach for each ECT patient, we obtain the uncertainty weights plotted in Figure 7.14. These weights can be used to derive an intra-patient upper uncertainty bound. This bound corresponds to the largest uncertainty obtained when applying this method to each patient and each PKPD model.

A direct application of this analysis is that, considering an 'unknown' patient for whom a unique PKPD model is obtained during induction, the uncertainty weight that needs to be considered to guarantee a stable design should be at least equal to the intra-patient upper uncertainty bound.

As compared to the uncertainty weight obtained for the whole study population, the intra-patient upper uncertainty bound is smaller across the whole frequency range (this is to be expected since the identification adds information concerning the patient). Note, however, that the uncertainty reduction in

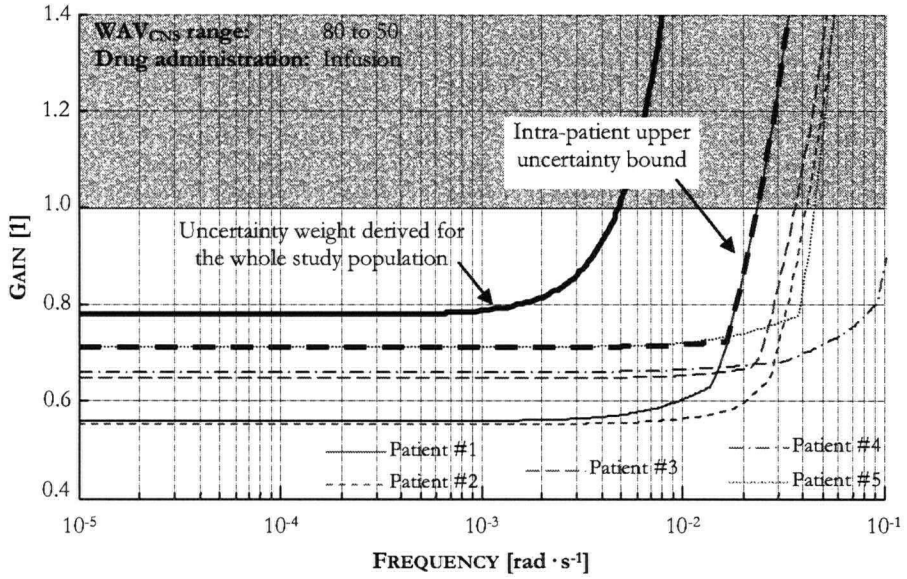


Figure 7.14: Intra-patient upper uncertainty bounds. This bound (in thick dash line) represents the minimum uncertainty weight which must be considered to guarantee stability if a PKPD model can be identified from the induction data (this model is then substituted to the generic nominal model to calculate the controller parameters).

the lower frequency band is moderate (only about 10%). The interest of this technique lies mostly in the higher frequency range where the uncertainty reduction is significant enough to allow for a wider control bandwidth (about half a decade larger).

Partial PKPD Parameter Identification Even though the new system-oriented modeling approach makes it possible to use induction data for PKPD identification, the procedure remains difficult since external factors may perturb the estimation of the model parameters. For instance, only about 50% of all ECT cases could be successfully modelled (many cases had to be discarded due to strong succinylcholine-induced fasciculation). Similarly, about 30% of the 74 LMA cases had to be discarded from the PD analysis due to increase in EMG during LMA insertion. It is also likely that the co-administration of other drugs may influence the PD steady state gain.

The method presented in the previous paragraph is still not very practical. In a normal clinical setting, we expect that only a limited number of cases will be suitable for such an analysis.

However, a parameter such as the PK time delay can be much easier to identify as compared to any other PKPD parameter. For instance, if the initial induction bolus is delivered through the infusion pump, the controller has access to the exact time at which the drug administration begins. After detecting a significant change in the WAV_{CNS} time course, the controller can then estimate the time which elapsed since the start of the syringe push. This time will not be affected by any stimulation which may occur during the LMA insertion or intubation. Also, the co-administration of other drugs should not affect significantly the PK time delay, as it is a reflection of the patient's arm-to-brain travel time.

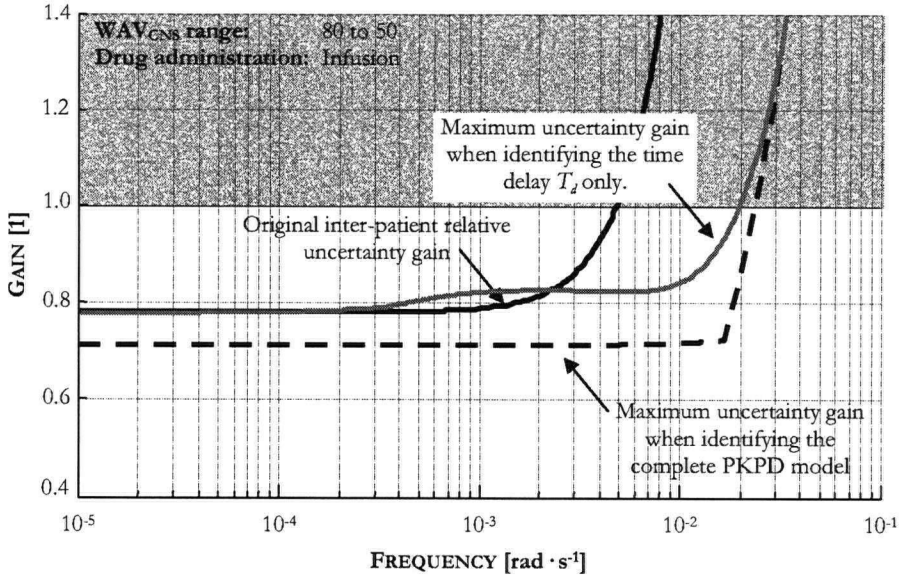


Figure 7.15: Limiting the PD identification to the sole estimation of the PK time delay can significantly reduce the uncertainty weight in the high frequency region. As expected, there is no reduction in the low frequency band. (Note that the increase in the uncertainty magnitude observed in the $[2 \cdot 10^{-4}; 2 \cdot 10^{-3}]$ rad·s^{-1} frequency range may be the result of slight differences in the way the optimization function that calculates the realization of the optimal nominal model was set.)

In Figure 7.15, we show the intra-patient upper uncertainty bounds when updating the generic nominal model with only the PK time delay identified during induction. As expected, there is no reduction of uncertainty in the low frequency range (the time delay does not affect the low frequency response of the PKPD models). The advantage of using the identified PK time delay lies once again in the high frequency range where the intra-patient upper uncertainty bounds is significantly smaller than the population-based uncertainty weight. A 3 to 5-times increase of the controller bandwidth can then achieved, provided that information on the patient's PK time delay is available. Note that another advantage of obtaining the PK time delay is that this information can also be used in the controller to compensate for the time delay and improve close loop performance (*e.g.*, using a Smith Predictor).

Note that only the identification of the PKPD steady state gain can bring improvements in the low frequency region.

7.4 Summary

In this chapter, we have investigated in details the PKPD uncertainty. This uncertainty stems from both inter- and intra-patient variability in drug disposition and effect.

Inter-patient variability can be easily understood as the variability originating from natural physiological differences between individuals. Intra-patient variability, on the other hand, has mainly two very different origins:

- **Changes in physiological processes**: all physiological processes do evolve in time. As a result, the model parameters may fluctuate in time. However, we found that the variability due to these changes is limited as compared to the inter-patient variability.

- **Linearization of the PKPD model**: another source of variability comes from the fact that the PKPD model is inherently non-linear. Linearizing the model non-linearity introduces a variable gain, which is responsible for a large part of the steady state gain uncertainty.

The approach chosen for quantifying uncertainty was based on the relative uncertainty methodology, where knowledge of the plant variability is expressed as uncertainty bounds in the frequency domain. These bounds are used to define Nyquist uncertainty disks that contain all possible plants. The ratio between the disks radii and the magnitude of the nominal plant define the relative uncertainty weight, which, in turn, imposes limitations in terms of the close loop characteristic of the control system[7].

This analysis has revealed that the variability observed in PKPD models is cause for concern. When considering a generic PKPD model, the uncertainty weight in the low frequency region already reaches 0.95 and crosses the '1' threshold at a frequency of $1.5{\cdot}10^{-4}$ rad$\cdot$s^{-1}. A sharp increase in the uncertainty weight can be further observed for frequencies $>1.0{\cdot}10^{-2}$ rad$\cdot$s^{-1}, thus effectively limiting the controller cutoff frequency to the $[1 \cdot 10^{-3}; 1 \cdot 10^{-2}]$ rad$\cdot$s^{-1} range.

When faced with performance limitations due to uncertainty, it is recommended to investigate methods to reduce the magnitude of the uncertainty weight.

- We have shown that the inter-patient variability can be simply reduced by considering the age of the patient. When dividing the study population into 10 yrs brackets, we found that the uncertainty magnitude can be reduced by 5 to 10%. This moderate improvement can be surprising at first since age is recognized as the most prominent covariate in PK and PD models. However, this result is similar to what has been reported in the PK literature, where models accounting for all PK covariates exhibit a 10% reduction in prediction error as compared to generic PK models. While reducing the age bracket further may improve on this result, a much larger sample size would then be needed.

- PK dynamics are also particularly non-linear with respect to the infusion rate. When expressing this non-linearity as system uncertainty, we observed a 10 to 15% increase in the uncertainty weight in the $[1 \cdot 10^{-4}; 1 \cdot 10^{-2}]$ rad$\cdot$s^{-1} range (which corresponds to the PK dynamic range). However, it is doubtful that controllers will be given the possibility of administrating boluses without human supervision. Hence, constraining the control action to infusion rates <1 mg$\cdot$min$^{-1}\cdot$kg^{-1} is a safety measure that also guarantees that the patient's pharmacokinetics follows the dynamic behavior of infusion-based regimen. This decrease in uncertainty can be achieved by constraining the controller action.

[7] In general, tight control is not possible for frequencies where the uncertainty weight is > 1. Also, the inverse of the uncertainty weight represents the largest magnitude of the complementary sensitivity function T. If T is above this limit for any frequency ω, there exists a set of plants for which the control design will be unstable.

- We have also shown that it is possible to further reduce the uncertainty by limiting the operating range of the controller. Indeed, a large part of the uncertainty originates from the linearized Hill saturation, which can be considered as a non-linear gain. This gain is typically bounded. However, the bounds may be quite large when considering that the system can operate within a wide output range. Limiting this range automatically limits the steady state gain uncertainty, which in turn, limits the overall system uncertainty.

Limiting the PKPD models to a certain age bracket and constraining the control action to infusion-type administrations reduce the uncertainty while not adding complexity to the controller design, see Figure 7.16.a. This should therefore be considered as part of the uncertainty strategy when dealing with the control of general anesthesia. Limiting the operating range of the controller can also be envisaged if a gain scheduling scheme (*i.e.*, where a specific controller is chosen according to the current WAV_{CNS} operating point of the system) is implemented. However, the added complexity does not justify the increased design cost of this solution, mostly when considering close loop controllers designed for general anesthesia and/or deep anesthesia (in these cases, the improvement when compared to the general purpose controller is only about 5%). Note, however, that the situation is quite different when considering the control of sedation in the ICU. Since targeted setpoints below 50 are not expected, a unique sedation-specific controller may be designed based on the relative uncertainty weight obtained for the 80 to 50 WAV_{CNS} range. In this case, the uncertainty is significantly reduced, see Figure 7.16.b. This observation leads us to conclude that the close-loop control of sedation in the ICU may be easier to achieve and will lead to better performances than the control of general anesthesia. Another implication of this result is that a stable controller designed specifically for sedation in the ICU may be unstable in the OR.

The three methods discussed above simply limit the validity of the model to a particular population of patients and/or operating conditions. A more active method is discussed in Section 7.3.2. This method makes use of the fact that the system-oriented approach to PD modelling requires limited amount of data for identification purposes. In many cases, the induction time course of the WAV_{CNS} can therefore provide useful information concerning the patient's specific dynamic behavior to the administration of an anesthetic drug. We have indeed shown that the magnitude of the uncertainty weight can be reduced by about 10% in the low frequency region by using the identified induction PD model. More importantly, we have shown that the control bandwidth can be increased by about half a decade, as the uncertainty in the high frequency region is significantly reduced.

PD identification during induction, however, may not be practical due to a large number of factors that can affect the WAV_{CNS} time course. Self-tuning of the model parameters during induction may thus be limited to specific situations that require trained human operators to assess the validity of the identification data and the derived PD parameters. An alternative to a full identification procedure is to only identify specific PKPD parameters. In particular, the sole identification of the PK time delay provides substantial improvements in terms of uncertainty reduction in the high frequency range. Note that the identification

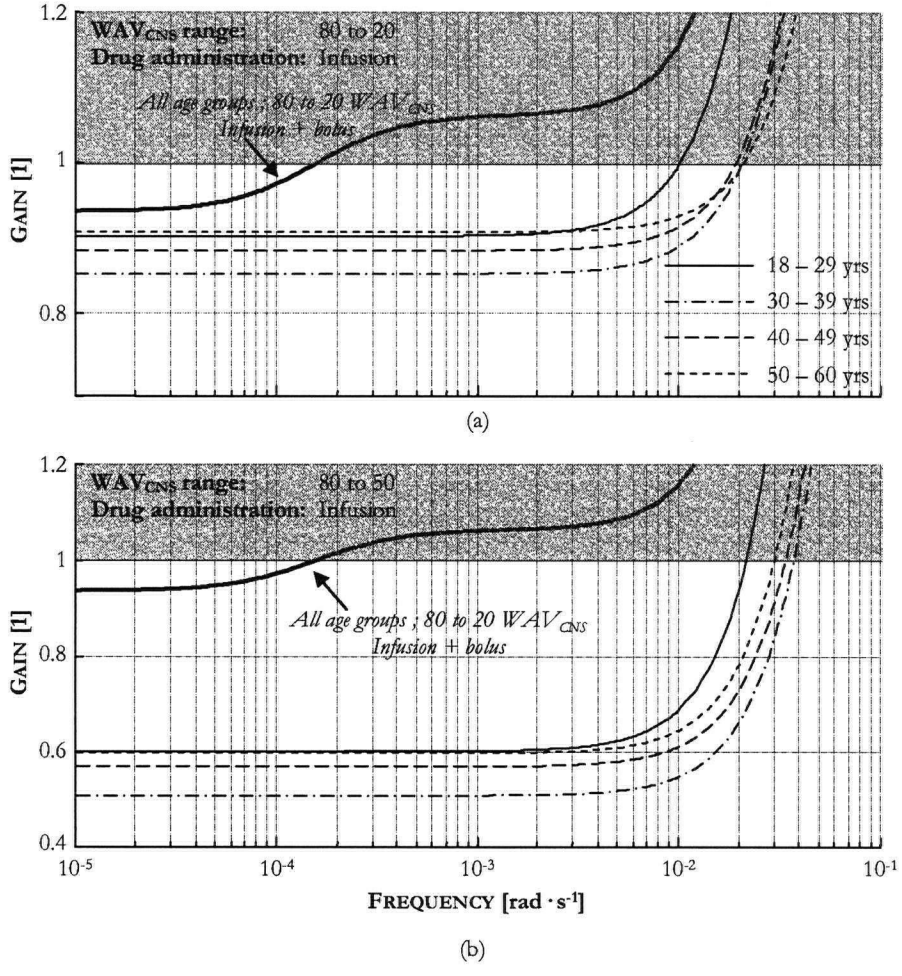


Figure 7.16: Selected uncertainty strategy. The corresponding PKPD nominal models, as well as the uncertainty weights, are summarized in the Table C.2 and Table C.5 (models #11 to #14), and Table C.3 and Table C.6 (models #15 to #18). (a) Application: general anesthesia in the peri-operative environment (the controller takes into account the age of the patient and limit the drug administration to infusions only). (b) Application: sedation in the ICU (in addition to accounting for the patient's age and limiting the drug administration to infusions only, the controller is also limited to maintain a sedation level in the 80 to 50 WAV_{CNS} range).

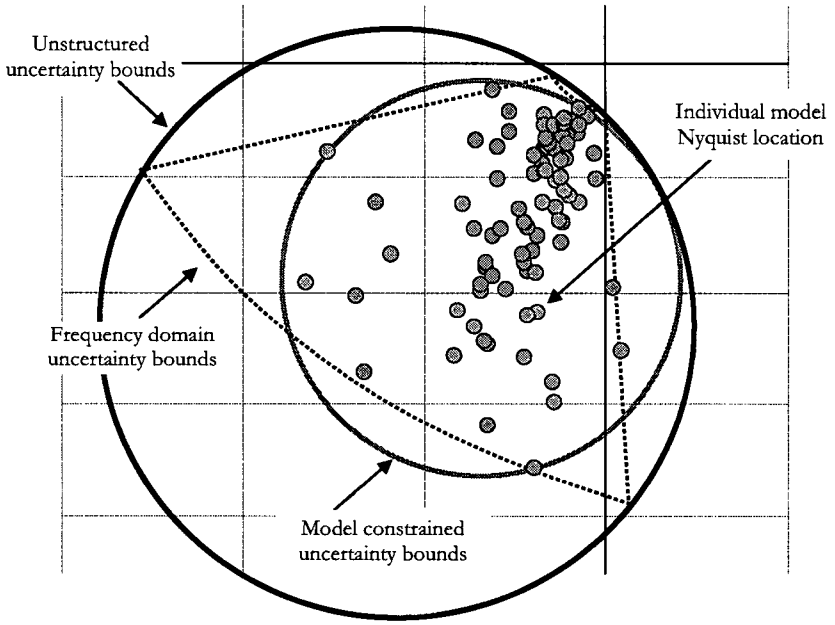


Figure 7.17: The frequency uncertainty bounds are represented as ring sections in the Nyquist plot. In the classical uncertainty analysis method followed in this chapter, the ring sections represent the system uncertainty. However, when plotting each individual PKPD model, we can notice that the Nyquist uncertainty disk can be significantly reduced. This approach introduces therefore less conservatism than the classical approach. Note that the absence of models in the outer left and right corners of the ring section may reflect the fact that PKPD models are essentially positive systems.

of this time delay is rather straightforward and can be carried out even if the patient reacts to LMA insertion, intubation, or the administration of a depolarizing NMB drug. Also, since this time delay is mostly representative of the arm-to-brain travel time, it is reasonable to assume that the co-administration of opioids during the induction procedure may not affect this parameter. If the thiopental results from the ECT study hold true for propofol, we found that identifying the PK time delay can potentially yield a 3 to 5-times increase of the controller cutoff frequency.

At this point it is important to note the strong conservatism introduced by the uncertainty methodology followed in this chapter. While this approach is usually recommended in the literature for uncertainty which cannot be expressed as parametric uncertainty, it may not be suitable for systems that present very large model variability. In this particular case, it was necessary to optimize the PKPD nominal model in order to obtain manageable uncertainty weights. The fact that such an optimization was necessary can already be construed as a warning that the method may be inadequate.

A less conservative method is to consider the Nyquist path of each individual PKPD model instead of their frequency response. At each frequency ω, it is then possible to define an uncertainty disk that contains all of the possible models, see Figure 7.17. Note that the radius of this new uncertainty circle is

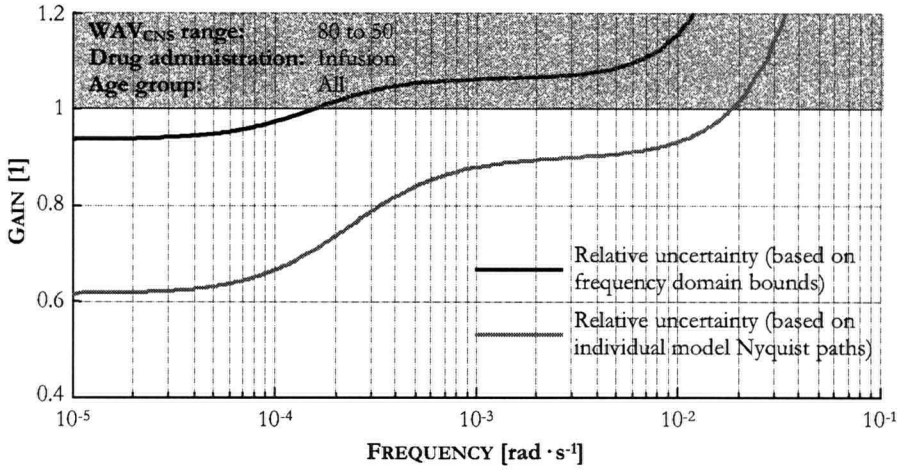


Figure 7.18: Comparison between the uncertainty weight obtained from the classical method (where the uncertainty is first expressed as frequency domain bounds), and the new approach where all the models are directly expressed into the complex domain.

much reduced as compared to the circumscribing circle that contains the uncertainty ring section defined from the frequency domain analysis. By iterating this analysis over the whole frequency range, we can define a new optimal nominal Nyquist path. This path can be used to estimate $\tilde{G}_0^{opt}$, which, in turn, leads to a new uncertainty weight. When compared with the weight obtained from the classical method, we can observe a reduction of up to 35% in the low frequency region, and about 15% in the high frequency region, see Figure 7.18. Note that the uncertainty in both cases increases sharply after $1 \cdot 10^{-2}$ rad·s^{-1}, indicating that even the less conservative approach does not not necessarily yield an improved cutoff frequency.

Since this new approach is less conservative, it is recommended to obtain a population sample large enough to guarantee that the uncertainty bounds contain all possible patients' dynamics. In case of small population samples, such as in the LMA and ECT studies, the additional conservatism introduced by the approach followed in this chapter may be necessary.

Chapter 8

SISO Control Design

In this chapter, we design and assess the performance of a single-input single-output controller for the control of the WAV_{CNS} index (see Chapter 5), using propofol as the actuating drug. The control performances are assessed through simulation, using the 44 patient models derived in Chapter 6 (see Table 6.1).

In Section 8.1, we discuss the Matlab-Simulink (MathWorks, Natick, MA) PKPD patient simulator. In particular, and to guarantee the accuracy of the simulator output, we identify both the measurement noise and modeling error in order to simulate their combined effects on the system output.

In Section 8.2, we use basic loop shaping techniques to design a PID controller. We show that, without accounting for PKPD uncertainty, the controller can easily lead to instability. Only a design based on complementary sensitivity loop shaping guarantees stability, at the cost of degraded control performance. In this section, we also identify three patient models which present very different dynamic characteristics. These three patient models are used throughout this chapter for illustration purposes.

A more efficient way of designing robust controllers is to consider an H_∞ design procedure, which makes direct use of the uncertainty weights derived in Chapter 7. This procedure is discussed in Section 8.3. The application to the design of a Propofol controller for WAV_{CNS} control is shown to yield slightly better results to those obtained from the stable PID loop shaping design, at the cost of an increased complexity in terms of the controller structure.

Furthermore, in Section 8.4, we show that uncertainty reduction obtained when considering the age group of the patients, and when identifying the PK time delay, directly translates into significant improvements in terms of control performance. We show that the proposed age-targeted control design can yield clinically relevant performance. We also show that the on-line identification of the PK time delay does result in minor performance improvements.

When interviewing clinicians on the issue of control performance, it appeared that faster settling times were desired. However, clinicians were ready to relax the overshoot requirements, therefore looking for a more aggressive design. In Section 8.5, we revisit the control design in order to improve the performances. In particular, we investigate the effect of increasing the controller gain and using a Smith Predictor structure. We show that large surgical disturbances can be rejected in about 3 min 30 sec in the majority of cases.

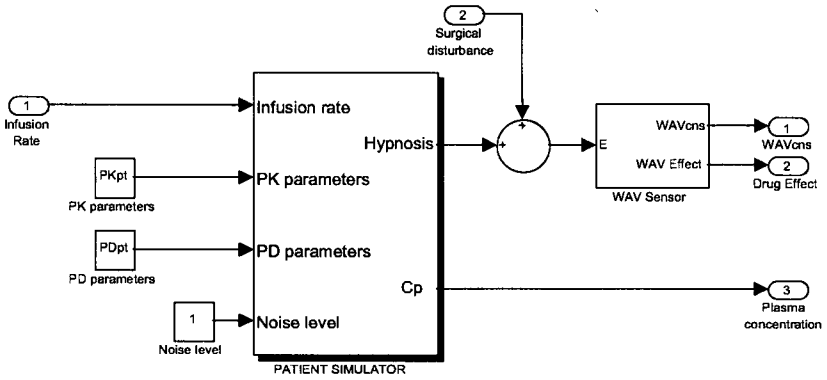


Figure 8.1: Patient Simulator block diagram.

A larger sensitivity peak and sensitivity to measurement noise are the trade-offs for these more aggressive designs.

Note that, in all cases, special care is taken for designing controllers that can handle constraints in terms of minimum and maximum infusion rates. In particular, we limit the maximum infusion rate to 0.8 $mg{\cdot}s^{-1}$ to ensure that the controller does not administer boluses for safety reasons.

8.1 Patient Simulator

Matlab-Simulink is one of the most well-known simulation platform used to test and validate control laws before implementation. In this section, we briefly present an overview of our PKPD patient simulator. We also identify the measurement and modeling noises in order to guarantee an accurate simulation of the patients' response to propofol administration when using the WAV_{CNS} sensor.

8.1.1 Simulator Structure

To simulate a patient's reaction to varying doses of propofol, a Matlab-Simulink simulator was constructed based on the PKPD model structure discussed in Chapter 6, see Figure 8.1. The inputs of the `Patient Simulator` block are the infusion rate (expressed in $mg{\cdot}s^{-1}$), the patient's PK parameters (both for infusion and bolus administration), the patient's PD parameters, and the measurement/modeling noise level. The outputs are the plasma concentration and the unfiltered hypnosis index (dimensionless value between 0 (no drug effect) and 1 (maximum drug effect)).

The `Patient Simulator` block contains three distinct subsystems, see Figure 8.2. The first subsystem calculates the plasma concentration based on the model PK parameters. The plasma concentration is then used by the second subsystem to calculate the drug effect. The third subsystem simulates the measurement and modeling noise by using a random sequence whose amplitude is user selectable (see Section 8.1.2). The simulated noise is directly added to the drug effect.

The `Pharmacokinetics` block can simulate both infusion and bolus concentration profiles, see Figure 8.3.

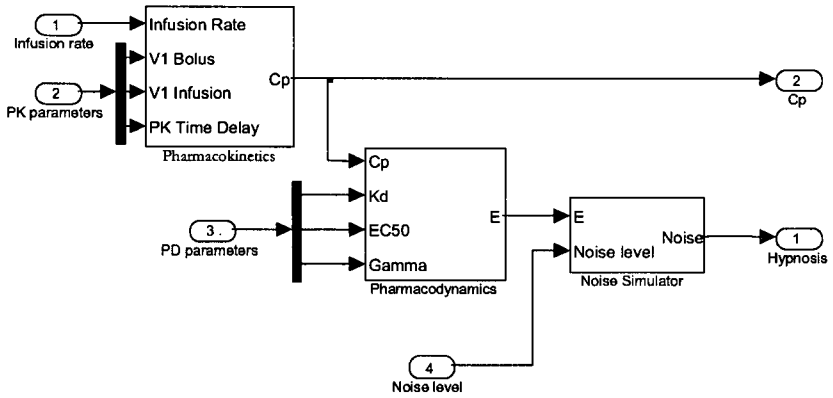


Figure 8.2: Patient Simulator block - Details level 1.

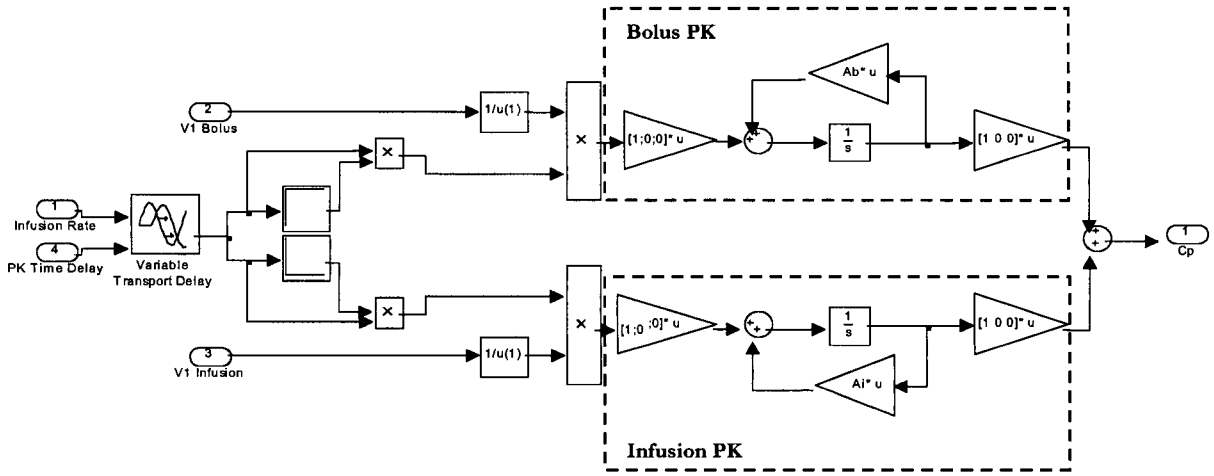


Figure 8.3: Patient Simulator - Pharmacokinetics block diagram. Depending on the infusion rate, the simulator will either output the plasma concentration corresponding to the bolus or infusion PK models.

The PK time delay is modelled through the transport delay block. A look up table is used to determine whether the infusion rate should be considered as infusion or bolus (in this simulation, an infusion rate above 1 $mg{\cdot}s^{-1}$ is considered to be a bolus).

The `Pharmacodynamics` block calculates the propofol effect based on the calculated plasma concentration, see Figure 8.4. The measurement noise is added after the Hill saturation element. As a result, the output of the patient simulator may take values higher than 1 or lower than 0 due to the noise.

The `WAV Sensor` subsystem models the WAV_{CNS} second order IIR trending filter (see Eq. (5.25)) and the affine scaling function, see Figure 8.5. A zero-order-hold is added to model the update rate of the sensor. The `WAV Sensor` block outputs both the WAV_{CNS} index (scaled from 100 to 0) and the trended drug effect (scaled from 0 to 1), which will be used as the feedback quantity for the controller.

Note that the surgical disturbance is added to the unfiltered hypnosis index. If a non-linear sensor such as BIS was to be used, (in which case the sensor output can display large abrupt changes), it would then

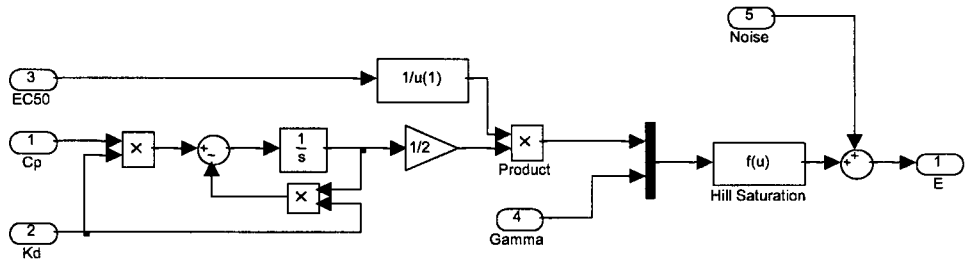


Figure 8.4: Patient Simulator - Pharmacodynamics block diagram.

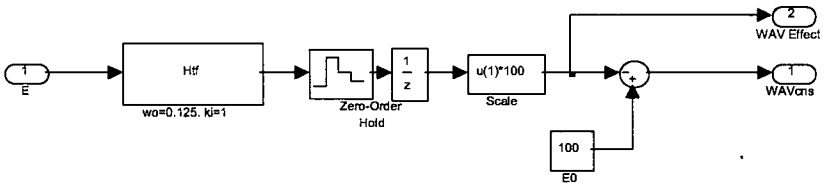


Figure 8.5: Patient Simulator - WAV Sensor block diagram.

be advisable to add the surgical disturbance just after the sensor output.

8.1.2 Noise Modeling

PKPD models present 2 distinct sources of noise: the measurement noise, *i.e.*, the high frequency variations observed in the WAV_{CNS} index (not representative of a drug or surgery-related change in the patient's state), and the modeling noise, which results from modeling errors.

Measurement Noise To quantify the measurement noise, we used the clinical EEG data obtained during the arthroscopy study. A subset of 12 arthroscopy cases presenting limited amount of electrocautery was first obtained[1]. Each case was processed using the WAV_{CNS} algorithm. The trending filter was disabled. As a result, only the un-trended WAV_{CNS} index was obtained. A wavelet-based denoising algorithm was then applied to extract the main trend from the raw WAV_{CNS} data. Patterns longer than $\approx$1 min were considered to be the result of surgical stimulation and/or drug administration. Patterns shorter than $\approx$1 min were associated to measurement noise, see Figure 8.6.

The difference between the WAV_{CNS} trend and the raw index was analyzed in the frequency domain. The frequency characteristics obtained in each case were combined and averaged, see Figure 8.7.

Modeling Error The modeling error can be assessed by using the residuals of the PD identification carried out in Chapter 6 based on the LMA data. Similarly to the previous noise characterization, we first processed each LMA case using the WAV_{CNS} algorithm without the post-processing IIR trending filter. We

[1]The use of the electrosurgical knife can create discontinuities in the WAV_{CNS} index which can falsely increase the measurement noise level.

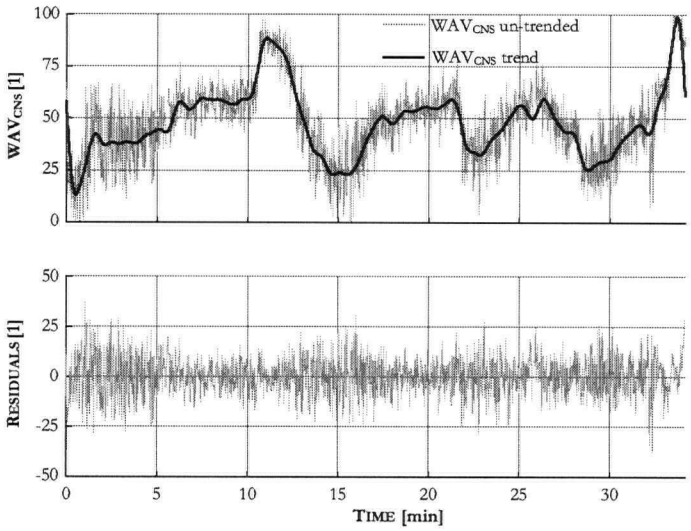


Figure 8.6: WAV_{CNS} noise measurement based on the raw index.

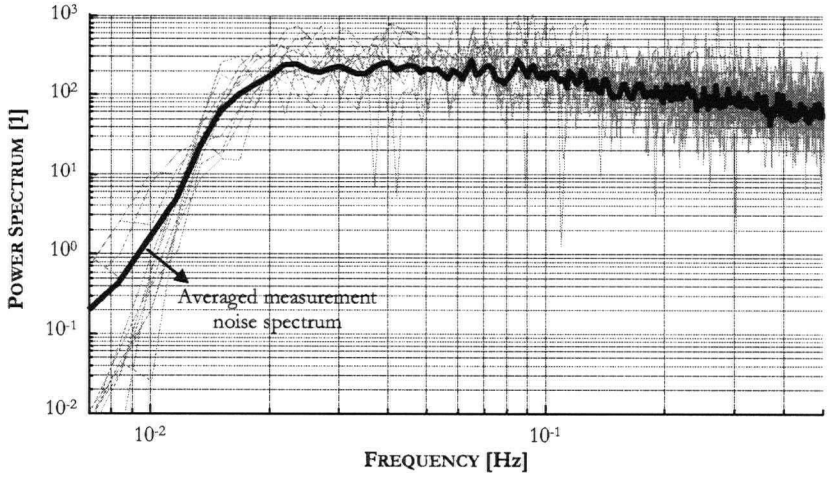


Figure 8.7: Measurement noise frequency characteristic.

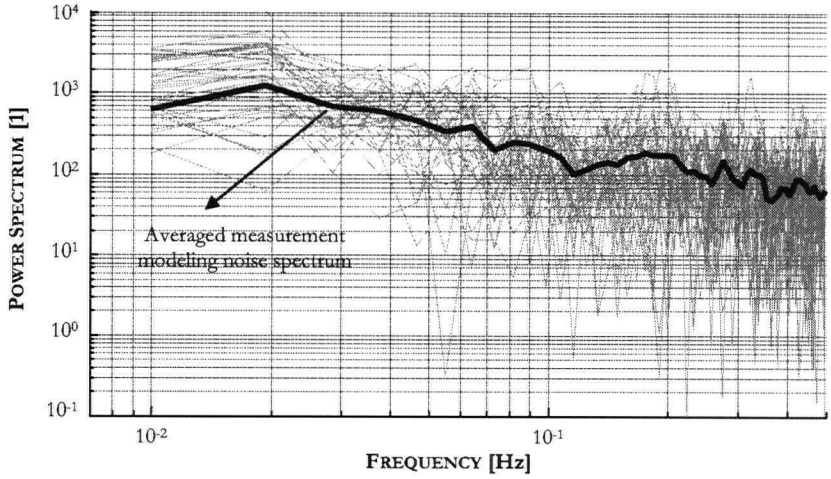


Figure 8.8: Measurement and modeling noise frequency characteristic.

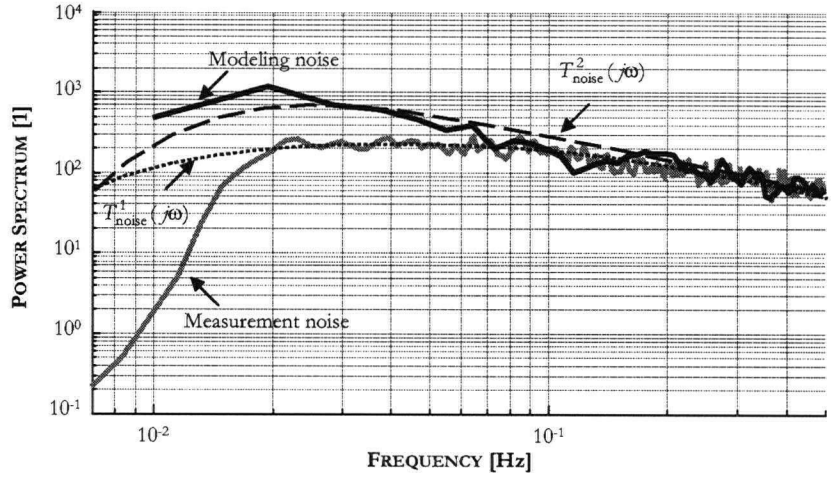


Figure 8.9: Identification of the noise characteristics. Two IIR filters were derived to capture the spectral nature of the measurement and modeling noises.

then subtracted the raw WAV_{CNS} from the predicted WAV_{CNS} time course obtained from the models. The power spectra obtained from the residuals were further combined and averaged, see Figure 8.7. Note that this noise frequency characteristic combines both modeling and measurement noises.

Simulating Measurement and Modeling Noises When plotted together, the frequency characteristics of Figure 8.7 and 8.8 reveal that both noises have very similar frequency content in the high frequency region, see Figure 8.9. It appears therefore that modeling errors add noise mainly in the lower frequency region. This increase in noise - as characterized by the power spectra - is significant. However, note that the modeling noise characterization was done using data from the the induction phase, were most of the modeling errors are expected. During smaller transients, it is likely that modeling errors will add less noise as compared to what is suggested in Figure 8.9.

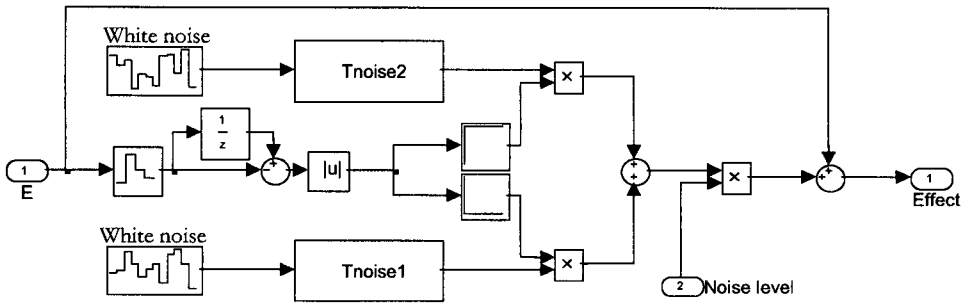


Figure 8.10: Patient Simulator - Noise Simulator block diagram. The drug effect E calculated in the `Pharmacodynamics` block is used to determine which noise profile that needs to be applied.

To simulate the measurement and modeling noises, a digital white noise generator followed by an IIR filter is used, see Figure 8.10. The IIR filter is tuned to yield a similar frequency characteristic to that of the noise. In particular, we derive here two filters, T^1_{noise}, which models only the measurement noise, and T^2_{noise}, which models both measurement and modeling noises:

$$\begin{cases} T^1_{\text{noise}}(j\omega) = 240 \cdot \left(\dfrac{s + 5 \cdot 10^{-4}}{s + 4 \cdot 10^{-2}} \right)^4 \cdot \dfrac{1}{s + 0.8} \\ T^2_{\text{noise}}(j\omega) = 240 \cdot \left(\dfrac{s + 5 \cdot 10^{-4}}{s + 5 \cdot 10^{-2}} \right)^8 \cdot \dfrac{1}{s + 0.2} \end{cases} \tag{8.1}$$

The frequency characteristics of these two filters are plotted in Figure 8.9. Note the good agreement between the measured noise profiles and the filter frequency responses.

During large amplitude transients (*e.g.*, when the WAV_{CNS} index varies more than 3 units per second), the aggressive noise filter T^2_{noise} is selected. Conversely, the noise filter T^1_{noise} is active whenever the WAV_{CNS} is slowly evolving in time.

8.1.3 Patient Simulator Output

To illustrate the output of the patient simulator, we simulated the response of patient #065 to the propofol dose given during the anesthesia induction. We then compared the simulator output to the WAV_{CNS} time course obtained during the case, see Figure 8.11.

Results show a very good agreement between measured and predicted time courses. The noise levels were also consistent with those initially measured.

8.1.4 Tests Setup and Performance Requirements

To test the proposed controllers, we propose to target a WAV_{CNS} level of 50 and analyze the time course of the index during changes in the targeted setpoint and during surgical disturbances, see Figure 8.12. Note that the level of 50 corresponds to an operating mode where the linearized Hill gain nears its maximum. As such, unstable designs will most likely exhibit an oscillatory behavior in this range.

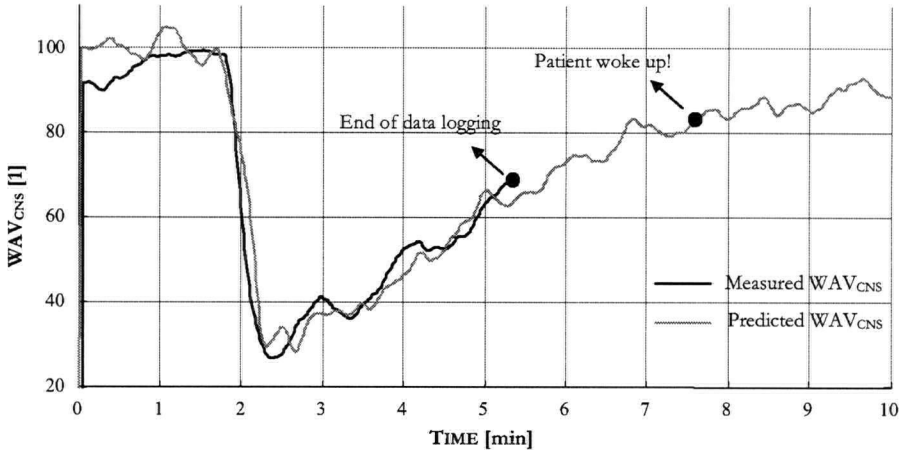


Figure 8.11: Measured *vs.* predicted WAV_{CNS} time courses (LMA patient #065). The noise filter T^2_{noise} was switched from 90 s to 170 s. The noise filter T^1_{noise} was used during the rest of the time. Note that only 5 min 20 sec of data were collected in this case. As a result of a human operator error, no inhalational anesthetic was provided to the patient after the LMA was placed. The anesthesia record mentions that the patient woke up 7 min 35 seconds after the start of the surgery. The predicted WAV_{CNS} time course does indeed indicate that the patient was above 80 by that time.

A setpoint change and disturbance amplitude of $\pm$ 20 WAV_{CNS} units is chosen. It has been our experience that the WAV_{CNS} rarely deepens or lightens for more than 20 units within a time span of a minute during surgery (besides for during induction and emergence, where the controller is not designed to have an active role).

To ensure the controller reaches steady state, the duration of each setpoint change and disturbance is set to 1 hour.

The controller performances are expressed in terms of the settling time, overshoot and maximum steady state error. The settling time is defined as the time it takes for the close loop system to settle to within ± 5 WAV_{CNS} units of the targeted setpoint. The steady state error is defined as the absolute difference between the targeted setpoint and the system's output 1 hour after either a step change in setpoint, or a step change in the surgical disturbance. Since the controllers are designed to work for a given population of patients, the performance requirements for the settling time and overshoot are defined by the performance obtained in 50% and 95% of the patient population. In order to be able to characterize the controllers performances, the measurement and modeling noises are not added during the simulations.

The control performance requirements are presented in Table 8.1. These requirements were derived based on my own interpretation of the anesthesiologists' needs. As such, they represent the point of view of a control engineer, rather than that of a clinician (the clinician's point of view on performance requirements is discussed in Section 8.5). Note that the settling time is our foremost priority and performance indicator. It is also desired that the controller design be more aggressive for disturbance rejection than for setpoint tracking[2].

[2]We are hypothesizing that, since surgical disturbances cannot be anticipated by the anesthesiologists, the controller should

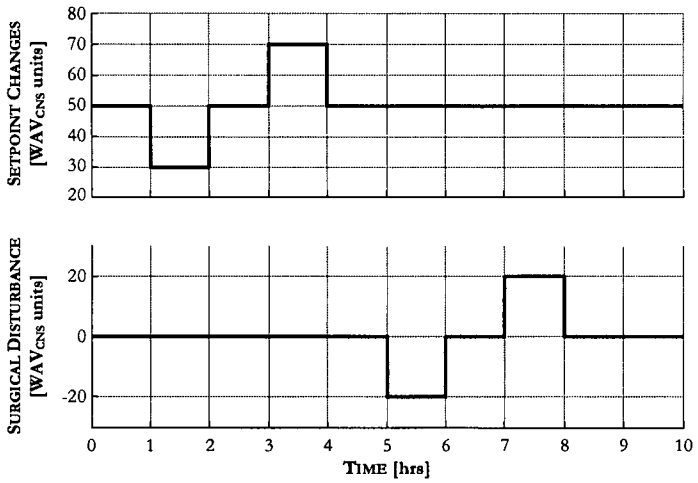


Figure 8.12: Controller test protocol.

	SETPOINT TRACKING		DISTURBANCE REJECTION	
	Desired	**Max. allowed**	**Desired**	**Max. allowed**
SETTLING TIME[1]				
50 percentile[2]	< 8 min	< 15 min	< 5 min	< 10 min
95 percentile	< 15 min	< 20 min	< 15 min	< 20 min
OVERSHOOT[3]				
50 percentile	< 2.5	< 5.0	< 5.0	< 10
95 percentile	< 5.0	< 7.5	< 7.5	< 10
SSE[4]	< 5.0			

[1] Settling time: time for the WAV_{CNS} to reach ± 5 units of its target setpoint.
[2] Maximum settling time obtained in 50% of the adult population.
[3] Expressed in WAV_{CNS} units.
[4] Maximum absolute steady state error allowed.

Table 8.1: Control Performance Requirements.

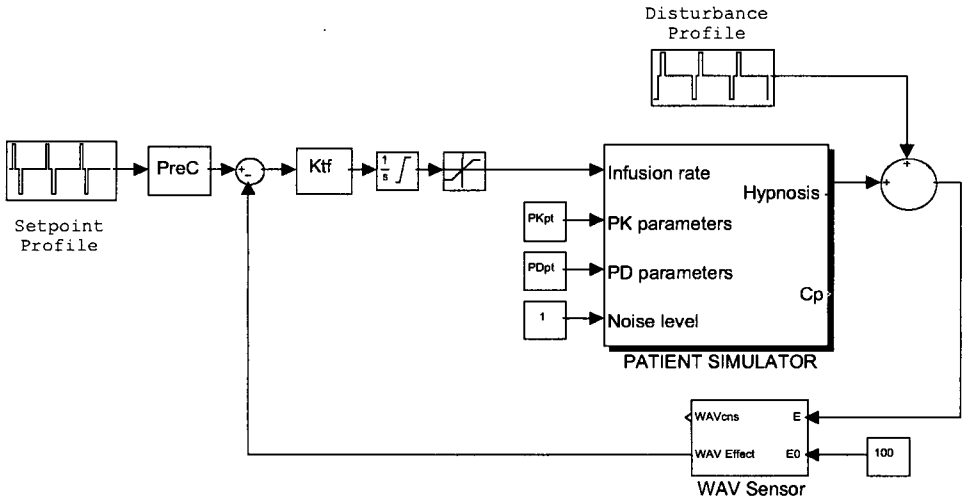


Figure 8.13: Close loop system with anti-windup and pre-filter.

8.2 PID Loop Shaping Design

In this section, we select the structure of a simple PID controller, where the derivative action is compensated through the use of a second order low pass filter to provide sufficient roll-off in the high frequency region:

$$K(s) = K_{PID} \cdot \frac{\tau_p \cdot s + 1}{s} \cdot \frac{\tau_d \cdot s + 1}{\left(\frac{\tau_d}{10} \cdot s + 1\right)^2}. \tag{8.2}$$

In the control of anesthesia, PID controllers have often been chosen by clinical researchers [123], [142] for their simplicity, good performance, and ease of design and implementation when applied to SISO linear systems.

8.2.1 First Design

In this first design, the time constants τ_p and τ_d, and the controller gain were tuned in order to obtain a satisfactory phase and gain margin (see Table 8.2). Since the controller is constrained to administer only infusion rates, the PKPD nominal model #6 (see Table C.1) is used to tune the parameters.

Note that the output of the controller is constrained between 0 and 0.8 mg·s^{-1}. In addition, an anti-windup scheme resets the integrator whenever either constraint is reached. A low pass pre-filter was also added to filter setpoint changes and limit abrupt control actions whenever such a change is initiated by the user. The pre-filter was slightly stronger than the WAV_{CNS} trending filter. As a result, we expect less overshoot during step changes in setpoint than during output disturbances. As a trade off, the settling time will be longer for setpoint tracking. The complete close loop system is presented in Figure 8.13.

compensate for them more rapidly than when a new setpoint is programmed. As a trade-off, the overshoot requirements for setpoint tracking are more stringent.

CONTROLLER	PHASE MARGIN [deg]	GAIN MARGIN [dB]	CUTOFF FREQUENCY $[10^{-3} \cdot rad \cdot s^{-1}]$
Population: 18-60 yrs			
First design	62.8	7.7	14.21
Second design	67.9	15.5	7.47
Third design	64.8	18.6	4.58

Table 8.2: PID Controller Characteristics.

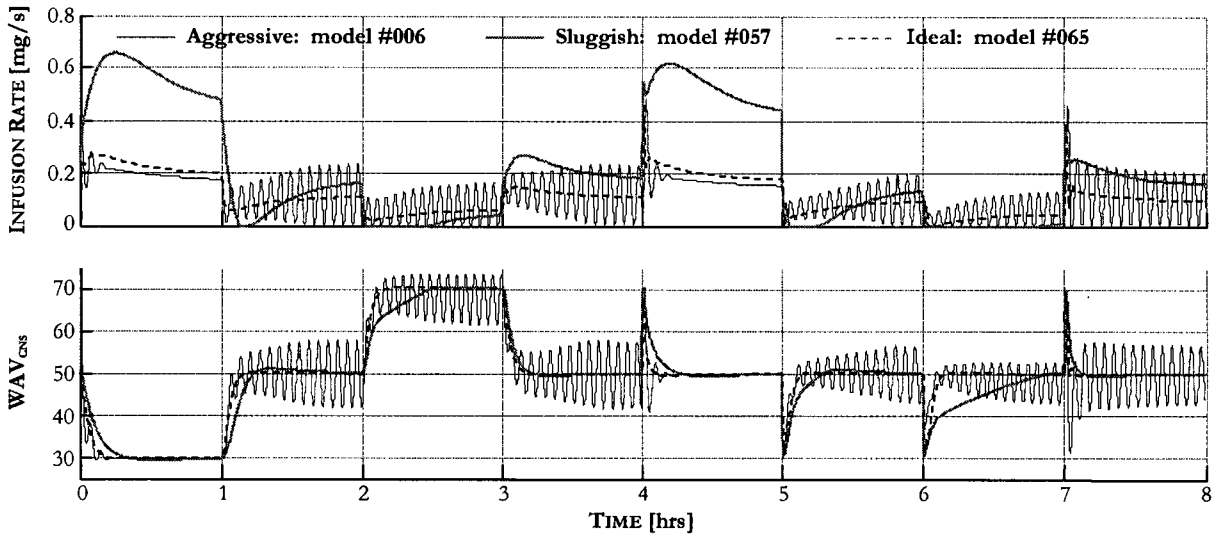


Figure 8.14: Infusion rates and WAV_{CNS} time courses in the 3 test patients using the classical PID loop shaping design method (first design). Note that patients presenting an aggressive dynamic behavior are unstable.

To further illustrate the closed-loop system output, we select 3 particular patient models presenting very different dynamic behaviors:

- **Aggressive behavior** (LMA case #006): the high gain and large phase lag of this model yields a poorly damped dynamic behavior which results in large overshoot, and even instability in control designs that are too aggressive.

- **Sluggish behavior** (LMA case #057): the low dc gain of this model makes it very sluggish to changes in infusion rates.

- **Ideal behavior** (LMA case #065): this case is a typical example of a well-behaved system (high gain and small phase lag) for which excellent results can be achieved.

The WAV_{CNS} time courses for these 3 patient models are plotted in Figure 8.14.

Out of the 44 patient models, 4 cases were unstable (#006, #007, #033, and #068) and 2 other cases (#004, and #015) exhibited a poorly damped oscillatory behavior during step changes in the output disturbance.

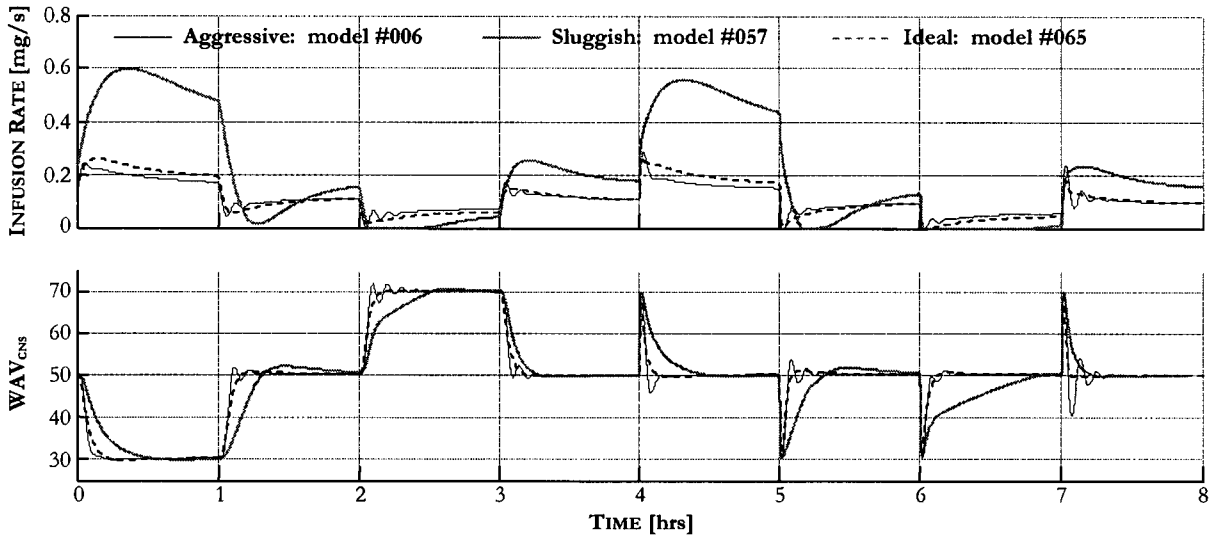


Figure 8.15: Infusion rates and WAV_{CNS} time courses in the 3 test patients using the classical PID loop shaping design method (second design). Note the poorly damped oscillations in the high gain patient. This controller is shown to be mathematically unstable.

8.2.2 Second Design

In the second design, the controller cutoff frequency was reduced by half, see Table 8.2. The resulting gain and phase margins would now be considered conservative for most designs.

After simulating the study population, all cases were found to be stable. However, 2 cases (#006, and #033) displayed poorly damped oscillations during step changes in the output disturbance.

The time course of the WAV_{CNS} index for the 3 test patients are plotted in Figure 8.15. The aggressive behavior of the first case resulted in poorly damped oscillations. Conversely, the sluggish case exhibits poor performances, mostly when the system tries to recover to the setpoint when surgical stimulation disappears.

The controller performances are further summarized in Table 8.3. While all performance requirements are met, the overshoot in one case was close to the maximum allowed of 10 WAV_{CNS} unit.

Even though all cases were stable, further investigation of the complementary sensitivity function reveals that the close loop system is potentially unstable, see Figure 8.16. This analysis is based on the augmented nominal model $\tilde{G}_0^{opt}(j\omega) = w_{nom}(j\omega) \cdot \tilde{G}_0(j\omega)$ defined in Table C.1 and Table C.4 (model #6)[3]. The instability of this design is apparent as the complementary sensitivity function crosses over the inverse of the uncertainty weight, thereby violating the Robust Stability (RS) condition (7.17).

8.2.3 Third Design

A further reduction in the frequency cutoff frequency was necessary to obtain a stable close loop system. The PID parameters of the second design were de-tuned such that the RS condition is met, see Figure 8.16. The

[3]Note that the uncertainty weight $w(j\omega)$ was derived based on the optimized nominal model $\tilde{G}_0^{opt}(j\omega)$. As such, it should only be used in conjunction with this nominal model.

	Setpoint tracking	Disturbance Rejection
Second Design		
Settling time		
50 percentile	5 min 49 s	3 min 57 s
95 percentile	10 min 54s	9 min 43 s
Worst case[1]	13 min 26s	10 min 40 s
Overshoot		
50 percentile	0.43	0.85
95 percentile	1.68	4.07
Worst case	3.68	9.57
Maximum SSE	0.72	
Third Design		
Settling time		
50 percentile	10 min 00 s	7 min 27 s
95 percentile	20 min 28s	18 min 22 s
Worst case[1]	27 min 09s	24 min 49 s
Overshoot		
50 percentile	0.98	1.49
95 percentile	2.53	2.61
Worst case	2.83	3.16
Maximum SSE	2.58	

[1] Excluding saturating cases where the infusion rate remains 0 for more than 5 minutes.

Table 8.3: Control Performance - PID Loop Shaping Design.

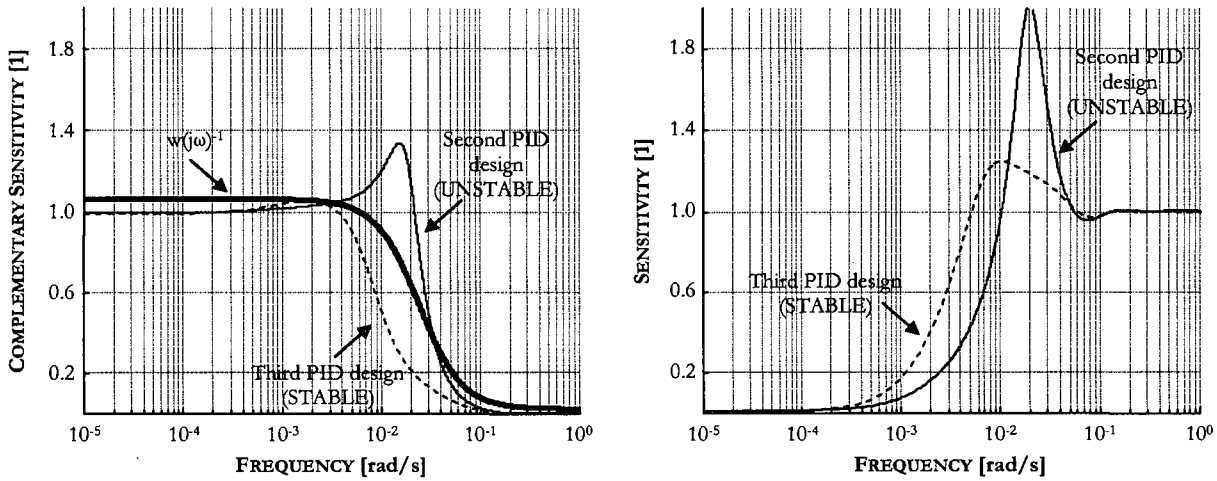


Figure 8.16: Close loop sensitivity analysis using the augmented nominal model $\tilde{G}_0^{opt}(j\omega)$. The second design is unstable since there exists a frequency range where the complementary sensitivity is greater than the inverse of the uncertainty weight.

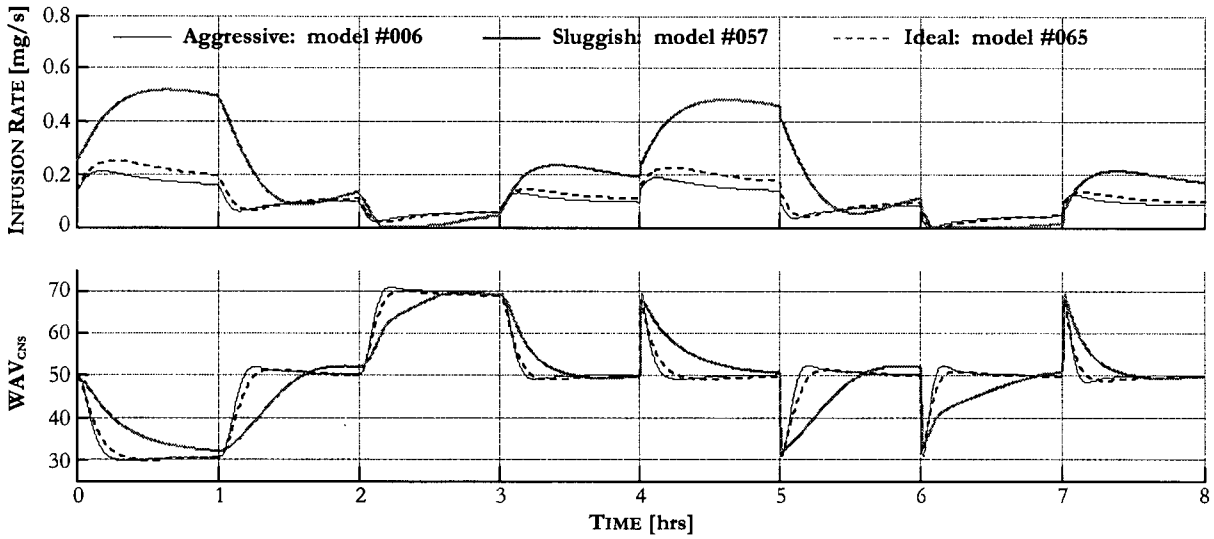


Figure 8.17: Infusion rates and WAV_{CNS} time courses in the 3 test patients using the classical PID loop shaping design method (third design). Performances are strongly reduced, in particular for patients presenting a low PKPD gain.

cutoff frequency is now $4.58{\cdot}10^{-3}\text{rad}{\cdot}\text{s}^{-1}$, which results in strongly degraded performances, see Figure 8.17 and Table 8.3. This design does not comply with the requirements of Table 8.1.

8.3 H_∞ Control Design

In recent years, a number of optimization tools for deriving controllers based on closed-loop sensitivity parametrization have been derived and are now directly available through computing platforms such as Matlab. In particular, the H_∞ design procedure has become increasingly used for designing closed-loop systems presenting large amount of uncertainty, as the optimization procedure directly accounts for the uncertainty weight.

8.3.1 H_∞ Design Principle

The basic principle of a signal-based H_∞ design approach is summarized in the Figure 8.18, where we are considering the control of a PKPD system. The objective of the control design is to tune the parameters of the controller $K(s)$ in order to minimize the outputs $z(s)$. These outputs are the weighted sensitivity, and complementary sensitivity of the closed-loop system, as well as the weighted control action.

Performance Weight The performance weight $w_p(s)$ weights the error signal $\epsilon(s)$. This weight is usually large in the low frequency range, thereby forcing the H_∞ optimization to find a controller that reduces static and low frequency control errors. Conversely, high frequency errors that cannot be properly controlled are given a low impact on the cost function by having a much reduced performance weight. It is therefore usual

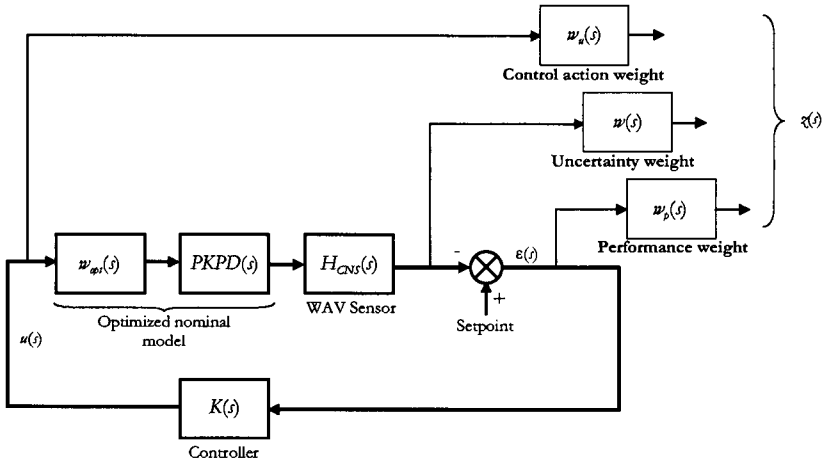


Figure 8.18: Mixed-sensitivity minimization.

to choose the performance weight as a first or second order function:

$$w_p(s) = \frac{\left(s/M^{1/2} + \omega_B\right)^2}{\left(s + A^{1/2} \cdot \omega_B\right)^2} \tag{8.3}$$

The dc gain of this function is given by the parameter $1/A$. Choosing $A = 0$ typically enforces the use of an integrator. However, numerical errors usually do not allow this choice. The parameter M on the other hand sets the high frequency weight of $w_p(s)$.

The main performance parameter remains the cutoff frequency ω_B. Increasing ω_B usually translates in a more aggressive design.

Uncertainty Weight In systems presenting large uncertainty, it is particularly important to ensure that the robust stability condition is respected. This implies that the complementary sensitivity of the closed-loop system should not be larger than the inverse of the uncertainty weight. This condition can be added to the H_∞ optimization by weighting the system output by the uncertainty weight $w(s)$ (note that the relationship between the system output and the setpoint is directly given by the complementary sensitivity function).

Control Action Weight In most designs, a weighted sensitivity and complementary sensitivity are sufficient to obtain the required closed-loop behavior. However, if we were to fix the complementary sensitivity weight to be equal to the uncertainty weight, it can then be useful to introduce the additional weight w_u to penalize large control action and guarantee sufficient roll-off in the high frequency region. For simplicity, w_u is often chosen as a constant.

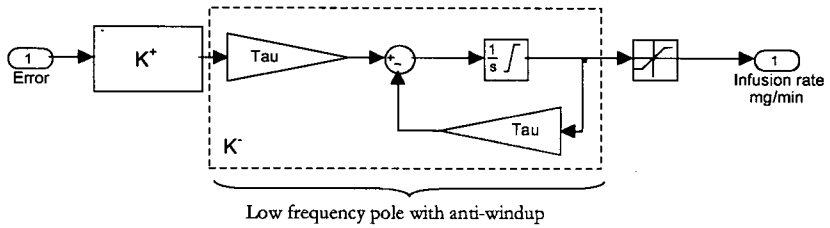


Figure 8.19: H_∞ controller implementation with anti-windup.

8.3.2 Implementation

The H_∞ design typically yields a controller whose order is equal to that of the open loop system, including the performance, uncertainty and control action weights. For this application, the open loop transfer function is a 8^{th} order system followed by a time delay modelled using a 1^{st} order Padé approximation. The uncertainty and performance weights are both 2^{nd} order transfer functions, and the control action weight is of order 0. As a result, the H_∞ controller is a 13^{th} order transfer function. While the Matlab simulator can easily handle this function, any practical implementation would have to use some order reduction technique.

H_∞ controllers usually do not have pure integrators. However, they usually have a very low frequency pole if the value of A is chosen close to 0. This low frequency pole can be used to add an anti-windup capability to the controller, see Figure 8.19. The controller $K(s)$ obtained from the H_∞ design procedure is first separated into its low frequency and high frequency components: $K(s) = K^-(s) \cdot K^+(s)$. $K^-(s)$ is a proper and monic function that contains only the lowest frequency pole. Conversely, $K^+(s)$ is a semi-proper function that contains all remaining dynamics. The anti-windup capability can easily be added to the controller by resetting the integrator of the $K^-(s)$ function whenever the saturation bounds are reached. In this case, the integrator is reset whenever the controller output is lower than 0 $mg{\cdot}s^{-1}$, of higher than 0.8 $mg{\cdot}s^{-1}$.

8.3.3 Application to Anesthesia Control

In this design, we first derive a performance weight based on the sensitivity function obtained from the third PID design. This provides us with a first performance weight.

The parameter A is chosen equal to 0.01, thereby indicating that a steady state error of 1 WAV_{CNS} unit is acceptable. We also choose to limit the sensitivity peak by having $M = 1.1$. The cutoff frequency ω_B was found to be 0.0015 $rad{\cdot}s^{-1}$:

$$w_p(s) = \frac{\left(0.95 \cdot s + 1.5 \cdot 10^{-3}\right)^2}{\left(s + 1.5 \cdot 10^{-4}\right)^2} \tag{8.4}$$

The uncertainty weight was set to be equal to the function corresponding to the model #6 in the Table C.4:

$$w(s) = 45.6 \cdot \left(\frac{s + 2.40 \cdot 10^{-2}}{s + 1.67 \cdot 10^{-1}}\right)^2 \tag{8.5}$$

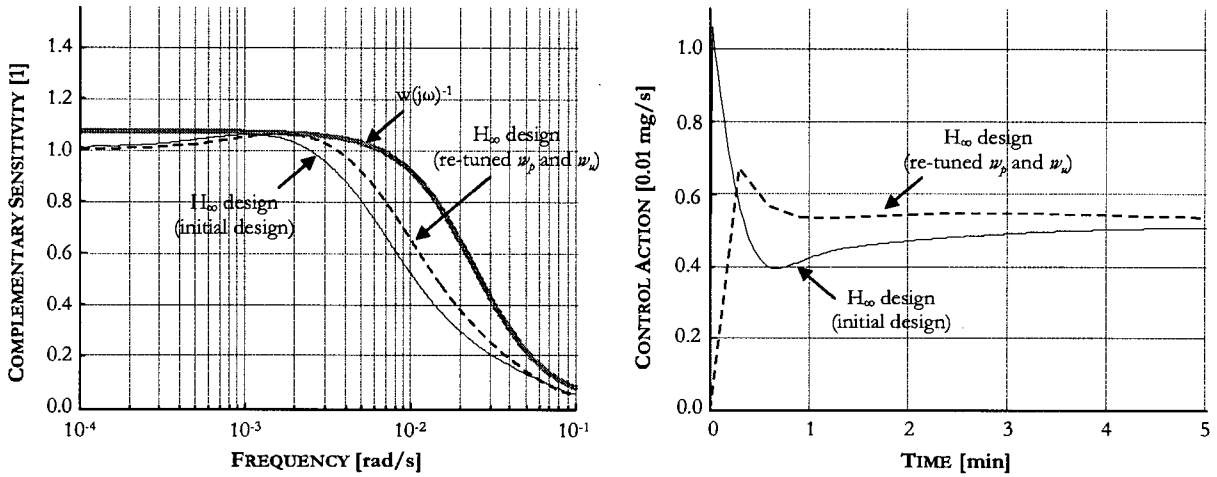


Figure 8.20: Initial and optimized H_∞ designs. (a) Closed-loop complementary sensitivity. (b) Control action profile in response to a setpoint change of 1 WAV_{CNS} unit.

CONTROLLER	PHASE MARGIN [deg]	GAIN MARGIN [dB]	CUTOFF FREQUENCY [10^{-3}·rad·s^{-1}]
Adult population	72.6	17.0	5.30

[1] Average values (each controller is specifically designed for a given PK time delay. As such, their frequency responses are all slightly different).

Table 8.4: H_∞ Controller Characteristics.

Finally, the control action weight was fixed to 1.

Using this set of weights, the H_∞ design yielded a first controller. While the complementary sensitivity remained below the inverse of the uncertainty weight, the control action to a setpoint change presented a large overshoot (≈100%), indicating high sensitivity to high frequency signals, see Figure 8.20. This design, while stable, would therefore behave badly in the presence of high frequency measurement noise.

To compensate for the shortcomings of the first design, we then increased the control action weight w_u to 3 and increased the ω_B parameter such that the complementary sensitivity is as close as possible to the instability bound defined by the uncertainty weight. This design resulted in a much improved control action profile, see Figure 8.20, while optimizing the performance with respect to the system uncertainty. This new design presents a cutoff frequency of $5.30 \cdot 10^{-3} \cdot \text{rad} \cdot \text{s}^{-1}$ and a large gain margin, see Table 8.4. As compared to the stable PID design, it also presents a larger complementary sensitivity bandwidth and a smaller sensitivity peak, see Figure 8.21.

The results obtained from the 3 test patients are plotted in Figure 8.22. Patients presenting low PKPD gain are expected to perform poorly. The controller performances are further summarized in Table 8.5.

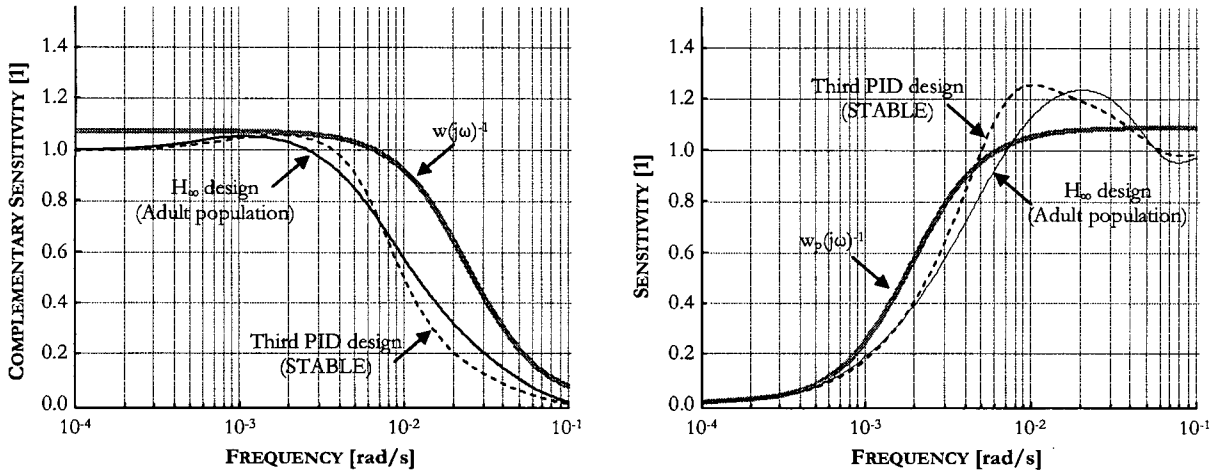


Figure 8.21: Close loop sensitivity analysis.

	SETPOINT TRACKING	DISTURBANCE REJECTION
SETTLING TIME		
50 percentile	9 min 52 s	7 min 15 s
95 percentile	20 min 17s	19 min 52 s
Worst case[1]	28 min 46s	26 min 40 s
OVERSHOOT		
50 percentile	0.77	0.65
95 percentile	2.42	2.60
Worst case	3.07	3.04
Maximum SSE	3.04	

[1] Excluding saturating cases where the infusion rate remains 0 for more than 5 minutes.

Table 8.5: Control Performance - H_∞ Control Design.

This controller does not meet the minimum requirements.

As compared to the stable PID design, this controller does perform slightly better, but at the expense of an increased complexity (13^{th} order controller *vs.* 3^{rd} order for the PID). The frequency responses of the H_∞ and the PID designs are compared in Figure 8.23. Both controllers show remarkable similarity. It is likely that the H_∞ controller can be reduced to a PID structure without significant loss in terms of closed-loop stability and/or performance.

8.4 Reducing the Uncertainty

In the previous designs, the controllers bandwidth are significantly limited by the uncertainty weight. In Section 7.3, we studied and derived methods to reduce this uncertainty. Two methods, in particular, were found to be very promising: targeting the patient's specific age group, and identifying the PK time delay in order to characterize more precisely the PKPD dynamics in the high frequency range. This section

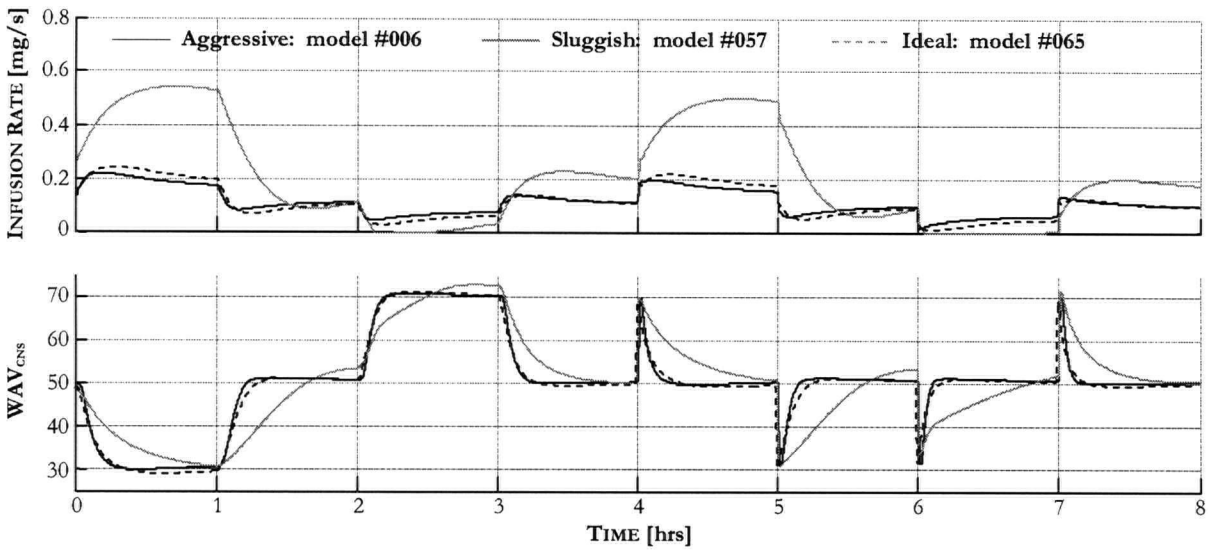


Figure 8.22: Infusion rates and WAV_{CNS} time courses in the 3 test patients. The H_∞ design robustified significantly the system which is now stable for the whole adult population. As a result, patients presenting low gains are not well compensated.

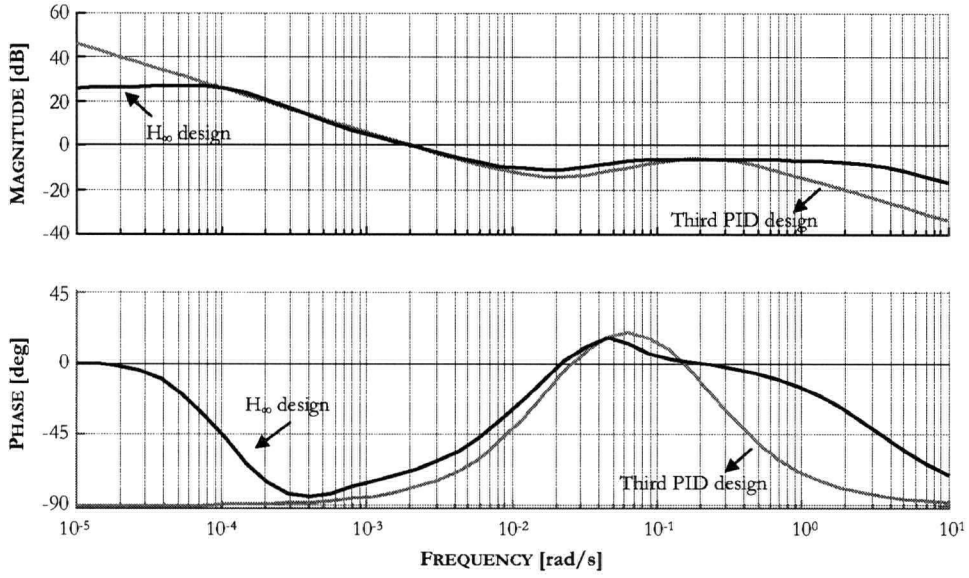


Figure 8.23: Frequency response of the H_∞ and PID controllers. Note that the PID has a faster roll-off in high frequencies, as well as a higher gain in low frequencies due to its integrator.

CONTROLLER	PHASE MARGIN [deg]	GAIN MARGIN [dB]	CUTOFF FREQUENCY [10^{-3}·rad·s^{-1}]
Age targeted design			
G1	72.1	15.5	6.18
G2	73.3	15.4	6.28
G3	73.9	15.5	5.97
G4	74.2	16.0	6.50
Age + Time delay identification[1]			
G1	62.6	14.6	7.64
G2	67.1	14.1	6.87
G3	65.1	13.4	6.86
G4	62.0	13.2	7.76

[1] Average values (each controller is specifically designed for a given PK time delay. As such, their frequency responses are all slightly different).

Table 8.6: H_∞ Controller Characteristics.

investigates the benefits brought by these two techniques towards the improvement of control performance.

8.4.1 Accounting for age

Even though the uncertainty weight reduction was moderate at best in the low frequency range, the uncertainty weight bandwidth was found to be significantly larger (by about 60%) when considering close loop infusion administration in each age group, instead of a combined type administration that targets the whole adult population (refer to Figure 7.16 and Figure 7.11). Hence, by separating the study population into the 4 age groups defined in Table 6.1, new H_∞ controllers can be derived. These controllers correspond to the PKPD models #11, #12, #13 and #14 (see Annex C).

The results obtained on the 3 test patients revealed a significant improvement in terms of settling time and steady state error, see Figure 8.24.

Applied to the 44 patient models, this control strategy results in the control performance summarized in Table 8.7. All requirements are now met.

8.4.2 Identifying the PK time delay

Another possible uncertainty improvement may be obtained by identifying the PK time delay during induction. For instance, we observed in Section 7.3.2.3 a significant increase in the bandwidth of the uncertainty weight when this parameter is identified. The ECT data and analysis carried out in Section 7.3.2.3 suggest that the poles and zeros of the uncertainty weights defined in Tables C.4, C.5, and C.6 can be multiplied up to factor 3 to 5. Assuming that the PK time delay is known, we can thus recalculate the controllers parameters. We then obtain for each patient an H_∞ controller tailored specifically to his/her PK time delay. This yields a slightly higher control bandwidth, which does result in some improvements, see Figure 8.25 and Table 8.7.

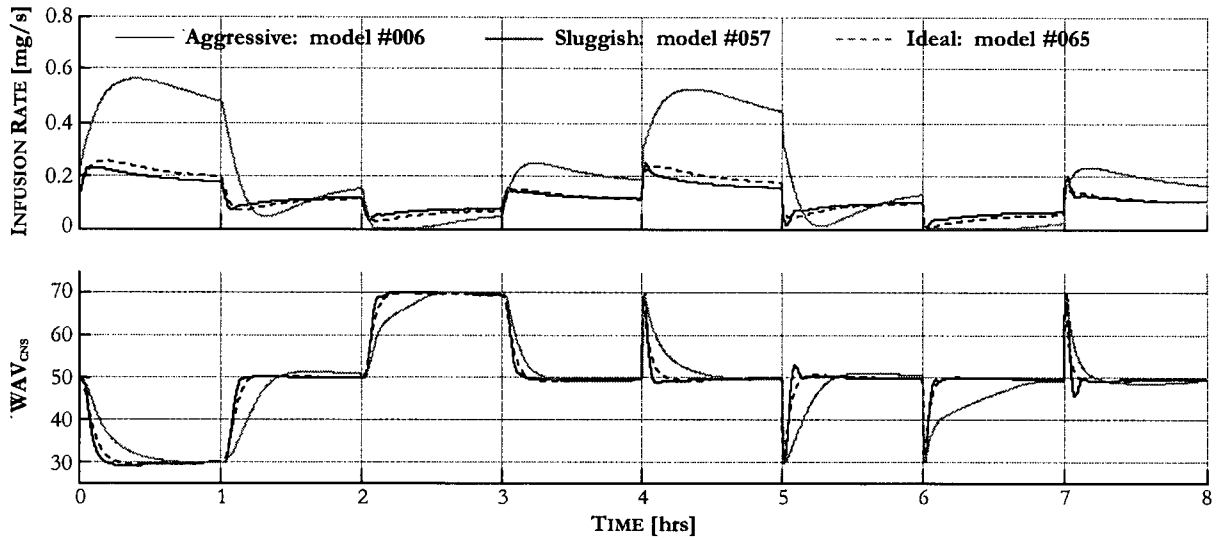


Figure 8.24: Infusion rates and WAV_{CNS} time courses in the 3 test patients. Accounting for age does result in a significant increase in performance.

	SETPOINT TRACKING	DISTURBANCE REJECTION
Age targeted design		
SETTLING TIME		
50 percentile	7 min 04 s	4 min 30 s
95 percentile	12 min 39s	10 min 53 s
Worst case[1]	16 min 17s	13 min 27 s
OVERSHOOT		
50 percentile	0.60	0.74
95 percentile	2.03	2.25
Worst case	3.14	3.63
Maximum SSE	1.69	
Age targeted design + PK ident.		
SETTLING TIME		
50 percentile	6 min 21 s	4 min 08 s
95 percentile	10 min 37s	8 min 42 s
Worst case[1]	13 min 58s	11 min 02 s
OVERSHOOT		
50 percentile	0.73	1.20
95 percentile	2.76	4.10
Worst case	3.56	5.30
Maximum SSE	1.89	

[1] Excluding saturating cases where the infusion rate remains 0 for more than 5 minutes.

Table 8.7: Control Performance - H_∞ Control Design based on Age.

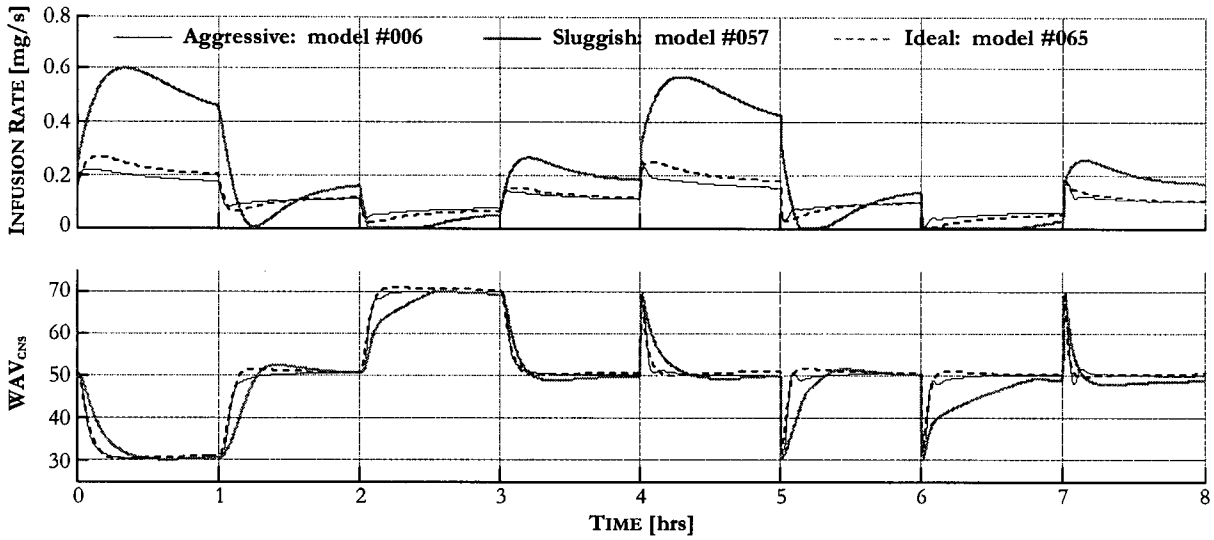


Figure 8.25: Infusion rates and WAV_{CNS} time courses in the 3 test patients. The on-line identification of the PK time delay reduces the system uncertainty and allows for a larger control bandwidth. Performances are improved at the expense of an increased complexity in the controller design.

In this design, the system's own dynamics become the limiting factor. Uncertainty is less of an issue.

8.5 The Clinicians' Point of View

For an anesthesia controller to meet the approval of the anesthesia community, it is important that its performances match the expectations of the clinicians. In that respect, the performance criteria of Section 8.1.4 were established based on a control engineering point of view, where overshoot is usually not desired. These criteria may therefore be skewed and should be revisited based on proper clinical feedback.

We have therefore asked 3 clinicians, Drs. Craig R. Ries (Vancouver General Hospital, Vancouver, BC), Mark Ansermino (British Columbia Children's Hospital, Vancouver, BC) and Don M. Voltz (University Hospital Health Services, Cleveland, OH), to share their expectations, see Table 8.8. They were asked to fill out a survey (see Annex E) indicating their preferences in terms of control performances.

The clinicians' control performance specifications were found to be rather different than those of Table 8.1. In particular, their expectations in terms of settling times are much more stringent, while the overshoot requirements are more relaxed. Their specifications are close to the performance they would be able to achieve in their practice. For instance, when a new setpoint is desired, or when a surgical stimulation has occurred, they generally respond by administrating a propofol bolus. This type of titration results indeed in a much faster reaction (within a minute), but at the expense of a large over- or undershoot. Under these circumstances, anesthesiologists would also permit a larger SSE.

The difference between our initial expectations and those of the clinicians may be due to:

- **Differences in our respective scientific cultures.** In control engineering, overshoots in industrial

	Setpoint tracking		Disturbance Rejection	
	Desired	**Max. allowed**	**Desired**	**Max. allowed**
Dr. Craig R. Ries (VGH)				
Settling time				
50 percentile	< 2 min	< 3 min	< 1 min	< 2 min
95 percentile	< 3 min	< 4 min	< 2 min	< 3 min
Overshoot				
50 percentile	< 5.0	< 10.0	< 10.0	< 15.0
95 percentile	< 10.0	< 15.0	< 15.0	< 20.0
SSE[4]	< 5.0	< 10.0	< 10.0	< 20.0
Dr. Mark Ansermino (BCCH)				
Settling time				
50 percentile	< 3 min	< 5 min	< 3 min	< 5 min
95 percentile	< 5 min	< 8 min	< 5 min	< 8 min
Overshoot				
50 percentile	< 5.0	< 12.0	< 5.0	< 12.0
95 percentile	< 8.0	< 15.0	< 8.0	< 15.0
SSE[4]	< 5.0	< 8.0	< 5.0	< 8.0
Dr. Don M. Voltz (UHHS)				
Settling time				
50 percentile	< 3 min	< 4 min 30 sec	< 3 min 30 sec	< 5 min 30 sec
95 percentile	n.d.	n.d.	n.d.	n.d.
Overshoot				
50 percentile	< 12.5	< 20.0	< 12.5	< 20.0
95 percentile	n.d.	n.d.	n.d.	n.d.
SSE[4]	< 8.0		< 8.0	
SUMMARY				
Settling time				
50 percentile	< 2 min 30 sec	< 5 min	< 2 min	< 4 min
95 percentile	< 4 min	< 6 min	< 3 min	< 6 min
Overshoot				
50 percentile	< 7.5	< 15.0	< 10.0	< 15.0
95 percentile	< 10.0	< 15.0	< 12.5	< 20.0
SSE[4]	< 8.0		< 8.0	

Table 8.8: Control Performance Requirements - Clinicians' Point of View.

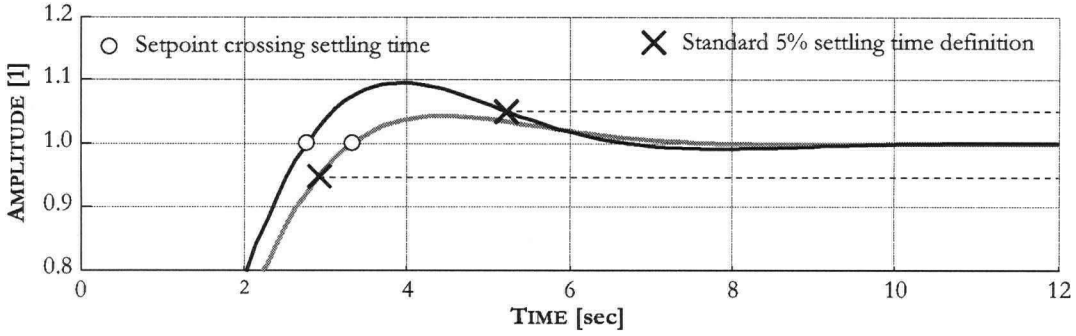


Figure 8.26: Settling time definition. In the standard definition, the settling time is defined as the time when the system settles within 5% of the setpoint. In cases with overshoot higher than 5%, this may result in much longer settling time values. For instance, in this example, one system has a significantly longer settling time. Therefore, if the overshoot is less of a concern, it may be indicated to use the setpoint crossing as settling time, in which case the two responses plotted are almost equivalent.

processes are usually associated with waste in terms of material and energy expenditure. Avoiding those is usually a priority. Conversely to control engineering, prompt action is usually recommended in a clinical situation.

- **Misinterpretation of the control specifications**. Our initial survey may not appropriately describe the settling time concept. After further discussion with Dr. Voltz, his interpretation of 'settling time' term was the time it takes for the controller to realize that a large error is present, and react to that error. This is very different to the engineering interpretation of the 'settling time' term, which is usually used to describe the time taken by the close loop system to be within 5% of the setpoint after the beginning of the disturbance (*i.e.*, by that time, the disturbance is already mostly rejected).

- **Inadequate performance specifications**. It is also possible that the performance specification usually followed in control engineering may not reflect adequately the concerns of anesthesiologists. For instance, since the overshoot is not a significant issue to clinicians, we may define the settling time as the time at which the close loop output crosses the setpoint. This would avoid situations with large settling times due to a larger-than-expected overshoot, see Figure 8.26.

Since a larger overshoot is permissible, there exist simple techniques that can be used to lower the settling time at the expense of the overshoot.

8.5.1 Increased Controller Gain

The most straightforward technique is to simply increase the controller gain, thereby increasing the cutoff frequency. This obviously translates into lower phase and gain margins, which herald larger overshoots and more oscillations.

To assess how this technique may perform, we revisited the age targeted designs of the previous section. For each age group, we selected the most aggressive patient model and increased the controller gain such

CONTROLLER	PHASE MARGIN [deg]	GAIN MARGIN [dB]	CUTOFF FREQUENCY $[10^{-3} \cdot rad \cdot s^{-1}]$
Age targeted design with increased gain			
G1	65.6	11.4	9.89
G2	71.1	13.8	7.54
G3	67.6	12.5	8.36
G4	66.6	12.5	9.74
Age targeted design + SP			
G1	70.3	18.5	9.25
G2	77.0	21.1	7.53
G3	57.5	12.2	7.93
G4	59.5	12.8	8.21
Age targeted + Time delay ident. + SP			
G1	68.5	19.2	10.1
G2	60.2	14.3	8.50
G3	56.9	12.8	8.36
G4	59.2	13.1	8.42

Table 8.9: H_∞ Controller Characteristics with Increased Gain.

that the overshoot reached $\approx$5 WAV_{CNS} units at most. The increase in gain was about 30 to 50%. Visual inspection of the close loop sensitivity functions was performed to ensure that the RS condition was not violated.

The controllers characteristics are presented in Table 8.9. As expected, the cutoff frequencies are increased at the expense of the gain margin. Note, however, that the phase margins remained well above 60 degrees.

In terms of performances, these controllers do perform better than in the previous design, see Table 8.10. The settling times are reduced by about 1 minute, while the overshoot remains well below the minimum requirements. The performance of the controller during setpoint changes can be further improved by removing the pre-filter. In this case, any change in setpoint is immediately detected by the controller which takes immediate tracking action. The settling time for setpoint tracking is significantly reduced and is equivalent to that of disturbance rejection. This obviously resulted in stronger overshoot, but is now closer to the clinicians' requirements.

8.5.2 Use of a Smith Predictor Structure

In systems with large delays, performance can also be improved by using a Smith Predictor structure that compensates for the nominal time delay. This time delay compensation allows an increase in the controller bandwidth, which results in improved performance.

The Smith Predictor (SP) makes use of the nominal model of the system in order to compensate for the delay, see Figure 8.27. The zero-delay nominal model is simulated based on the same infusion rate that is input to the system. As such, the model output represents the predicted delay-free response of the system. This response is then compared to the response with delay. The result of this comparison is a

	SETPOINT TRACKING	SETPOINT TRACKING[2]	DISTURBANCE REJECTION
Increased gain			
SETTLING TIME			
50 percentile	6 min 09 s	3 min 20 s	3 min 56 s
95 percentile	10 min 52s	4 min 51s	9 min 06 s
Worst case[1]	17 min 33s	10 min 41s	14 min 21 s
OVERSHOOT			
50 percentile	0.52	1.47	0.98
95 percentile	1.98	3.38	2.79
Worst case	2.77	5.25	5.12
Maximum SSE	3.10 1.28[2]		
Smith Predictor			
SETTLING TIME			
50 percentile	6 min 08 s	3 min 24 s	3 min 46 s
95 percentile	9 min 53s	7 min 10s	8 min 20 s
Worst case[1]	14 min 56s	11 min 51s	12 min 13 s
OVERSHOOT			
50 percentile	0.83	1.47	1.38
95 percentile	2.58	3.38	3.40
Worst case	3.60	5.25	4.65
Maximum SSE	2.05 2.18[2]		
Smith Predictor + PK time delay ident.			
SETTLING TIME			
50 percentile	5 min 45 s	3 min 20 s	3 min 37 s
95 percentile	9 min 17s	6 min 34s	8 min 09 s
Worst case[1]	11 min 50s	8 min 24s	8 min 34 s
OVERSHOOT			
50 percentile	0.79	1.89	1.70
95 percentile	2.96	4.19	3.99
Worst case	3.65	5.57	5.51
Maximum SSE	3.25 3.57[2]		

[1] Excluding saturating cases where the infusion rate remains 0 for more than 5 minutes.
[2] Without pre-filter.

Table 8.10: Control Performance - Improved H_∞ Control Designs.

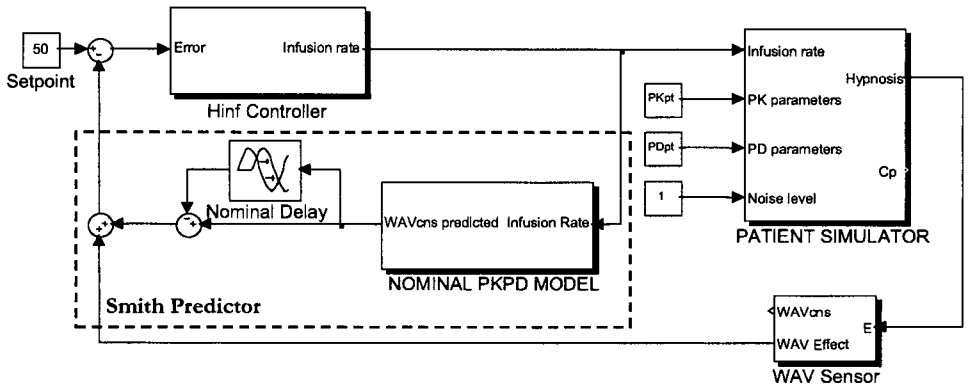


Figure 8.27: Smith Predictor control structure.

signal that represents the future system response to the control action. This signal is then added to the feedback signal. As a result, the controller can be designed based on a delay-free model, which results in added stability in the control loop that can be further used to increase the controller bandwidth. While the inherent limitation of a delayed system is still present, the increased control bandwidth usually results in increased performances.

However, in systems presenting large gain uncertainty, this technique potentially results in larger over- and undershoot. For instance, if the gain of the nominal model is an order of magnitude different than that of the real system, the SP structure may either result in under- or over-compensation, which is not desired. Because of the large uncertainty in this system, the use of a Smith Predictor would not be advised. However, the larger permissible overshoot requirements makes the use of a SP a possibility.

Age Targeted Design A SP structure was first added to the age targeted H_∞ designs of Section 8.4. The H_∞ control designs were re-iterated for each age group based on the delay-free nominal model. Visual inspection of the close loop sensitivity functions was performed to ensure that the RS condition is satisfied. To achieve optimal results, the most aggressive case in each age group was simulated. The control action weight $w_u(s)$ and the time delay ω_B of the performance weight were tuned in order to obtain a maximum under/overshoot of about 5 WAV_{CNS} units. The results obtained from the 3 test patients are presented in Figure 8.28.

The controller performances are summarized in Table 8.10. We are also presenting the performance obtained without pre-filter. Results are very similar to those obtained in the age targeted design with increased gain.

Age Targeted Design with PK Time Delay Identification One advantage of the SP structure in case the PK time delay is known, is that there is no longer a need for a distinct controller for each patient. The H_∞ control design is carried out based on the premises that there is no time delay, and that the uncertainty weight is shifted by half a decade towards higher frequencies. As far as the implementation of

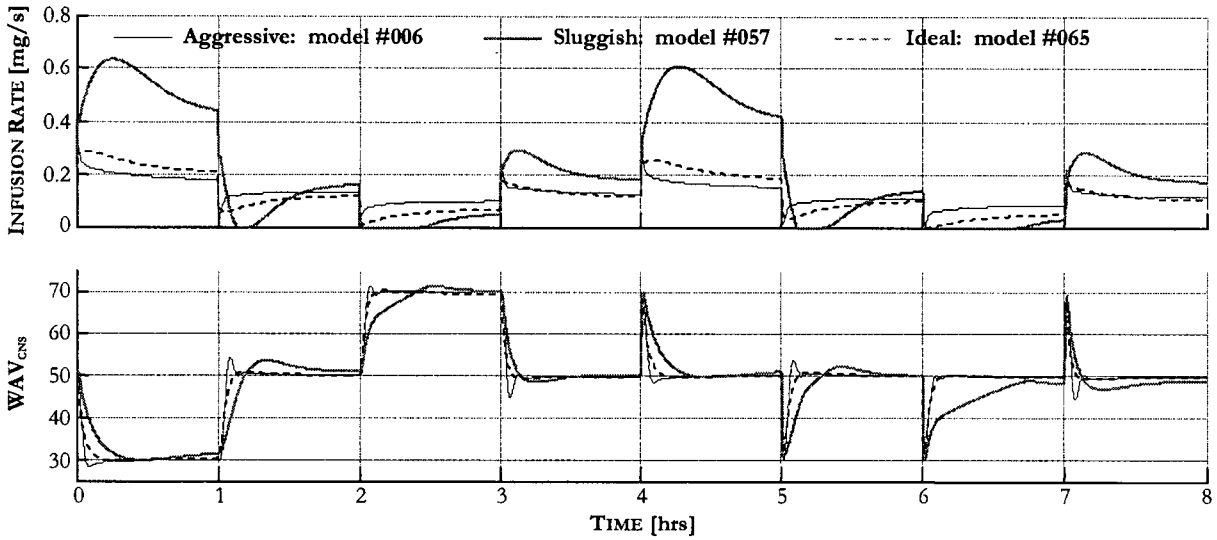


Figure 8.28: Infusion rates and WAV_{CNS} time courses in the 3 test patients. The use of a Smith Predictor allows for a faster settling time at the expense of a larger overshoot. Note that, in this simulation, the pre-filter was removed to reduce the settling time during setpoint change.

the SP is concerned, the nominal time delay is now replaced by the identified time delay.

Similarly than in the previous case, the model yielding the most aggressive response to setpoint changes and output disturbances in each age group is used to tune the performance weight such that to avoid obtaining an under- or overshoot of more than ≈ 5 WAV_{CNS} units. As per usual, visual inspection is used to ensure that the RS condition is satisfied.

The controller performances are summarized in Table 8.10. As expected, the performance improvement does not justify the increased costs and risks involved for identifying the time delay.

8.5.3 Faster Sensing Dynamics

One of the main limiting factor of the designs presented in Section 8.5.1 and 8.5.2 is the WAV_{CNS} trending filter, whose cutoff frequency is 0.125 rad$\cdot$s^{-1}. Shifting the pole of the IIR trending filter towards higher frequencies would then allow for a larger phase margin, which, in turn, would allow for a larger control bandwidth. However, this would also result in additional measurement and modeling noise, directly located in the bandpass of the closed-loop system. The use of a faster sensing dynamics should then be accompanied by the development of a new sensor design presenting reduced measurement noise characteristics[4].

[4]We recently investigated using the WAV_{CNS} index derived from both left and right hemispheres. Averaging the measurement from both hemispheres allowed a noise reduction of 8%. This improvement was not enough to allow for a significant increase of the filter pole.

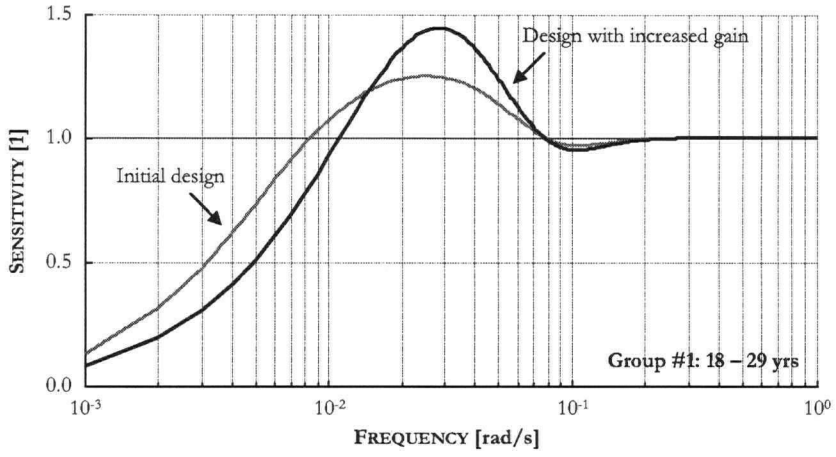


Figure 8.29: Sensitivity peaks of the initial H_∞ age targeted design, and the design involving an increased controller gain.

8.5.4 Improved Performance: Important Trade-offs

The improved performances obtained in this section do, however, present two important trade-offs: a higher sensitivity peak, and a higher sensitivity to measurement noise. These issues are briefly commented here.

Higher Sensitivity Peak While increasing the cutoff frequencies of the controllers have resulted in improved performances, attention must be brought to the fact that the sensitivity peaks are also increased, see for instance Figure 8.29. The age targeted design with increased gain presents a sensitivity peak close to 1.5 in the $[0.01, 0.08]$ rad$\cdot$s^{-1} frequency range. In comparison, the more conservative design peaks up at 1.25.

This essentially indicates that any surgical disturbance in this particular frequency range will be amplified instead of compensated. While periodic surgical stimulation is not expected, this may result (in some rare cases) in poor control performance.

Measurement Noise Interpreted as Disturbance Due to the larger controller cutoff frequency, a large quantity of the measurement/modeling noises is now interpreted by the controller as a potential output disturbance. As a result, the controller will try to compensate for this noise by constantly adjusting the propofol infusion rate.

While this is of no consequence in terms of close loop stability, the controller sensitivity in the measurement noise frequency range can slow down the close loop response when the controller saturates to its lowest or highest constraint. A typical example is to consider the response of a 'sluggish' patient to the disappearance of surgical stimulation, see Figure 8.30.a. The fact that the measurement noise is constantly activating the controller prevents it from stopping the drug delivery altogether. As a result, the patient is constantly given more drug and takes more time to settle back to the setpoint.

One possible solution to avoid this situation is to have the low frequency pole integrator reset to a

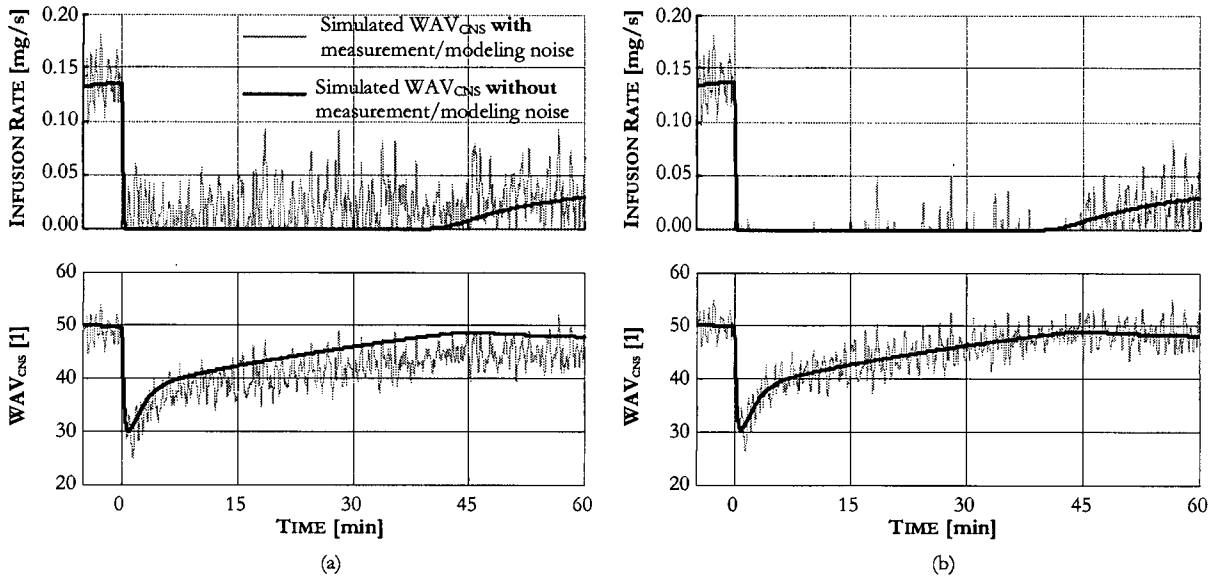


Figure 8.30: Effect of measurement/modeling noise on the closed-loop system response during saturation. (a) The age targeted design with increased gain keeps on administrating propofol, which delays the system response. (b) Using a modified anti-windup scheme where the low frequency pole integrator resets for negative values instead of 0 can be a solution.

negative value instead of 0, see Figure 8.30.b. However, this does limit the anti-windup capability of the controller, which can overshoot in high gain patients.

Having the controller act on measurement noise also implies that it becomes increasingly difficult for the anesthesiologist to gauge the validity of the controller output, since it can significantly change over short periods of time.

8.6 Summary

In this chapter, we have derived a number of controllers using both the classical PID loop shaping and H_∞ design methods. The controllers are summarized in the Figure 8.31.

Without considering the system uncertainty, the first PID design results in instability in about 10% of our study population. A standard approach is then to derate the controller sufficiently to gain more stability. A conservative design which presents a large gain and phase margin does indeed result in a stable result for the whole study population. However, the mathematical analysis of the complimentary sensitivity of the close loop system reveals that the system remains unstable. Designing a controller without accounting for uncertainty can thus lead to instability, even when adopting a conservative approach. While stable, the third PID design proposed in Section 8.2 has poor control performance, and does not meet the minimum requirements set forth in the first section.

The H_∞ design based on a mixed-sensitivity approach has the advantage of automatically accounting for the uncertainty weights defined in the Chapter 7. All controllers designed with this approach have shown

remarkable stability (fast decay ratio in cases presenting oscillatory transients, and limited overshoot). However, the general purpose H_∞ controller designed for the whole adult population did not meet our performance requirements.

Significant improvements could only be obtained by designing controllers specific to a certain age group. We found that large surgical disturbances can be compensated within 5 minutes in the majority of cases, while exhibiting low overshoot. The intra-patient variability study pursued in Chapter 7 suggested that the on-line identification of the PK time delay could bring improvements in terms of the controller bandwidth. While correct, we found that the improvement does not justify the added costs and risks of performing an on-line identification.

When asked about control performance requirements, the 3 clinicians we interviewed unanimously wished for a system responding faster to disturbances and setpoint changes, even if it meant larger overshoots and steady state error. We therefore revisited our designs to improve the performance by increasing the controllers static gain, and using a Smith Predictor structure. The results show that these solutions meet the clinicians' minimum requirements.

As we are nearing the physical limitations of the PKPD system, it becomes increasingly clear that the optimal solution will involve both a human operator and the closed-loop system. Large disturbances or setpoint changes will be dealt with by the anesthesiologist who can use bolus adminstration to fasten the response of the system. The closed-loop controller, on the other hand, will meticulously re-actualize the infusion rate to ensure limited overshoot and a low steady state error. It is therefore possible that the more conservative designs of Section 8.4 may be more favorable than the designs proposed in Section 8.5, as they would make optimal use of the synergy between the anesthesiologist and the closed-loop system to better control the patients' response to propofol administration and surgical stimulation.

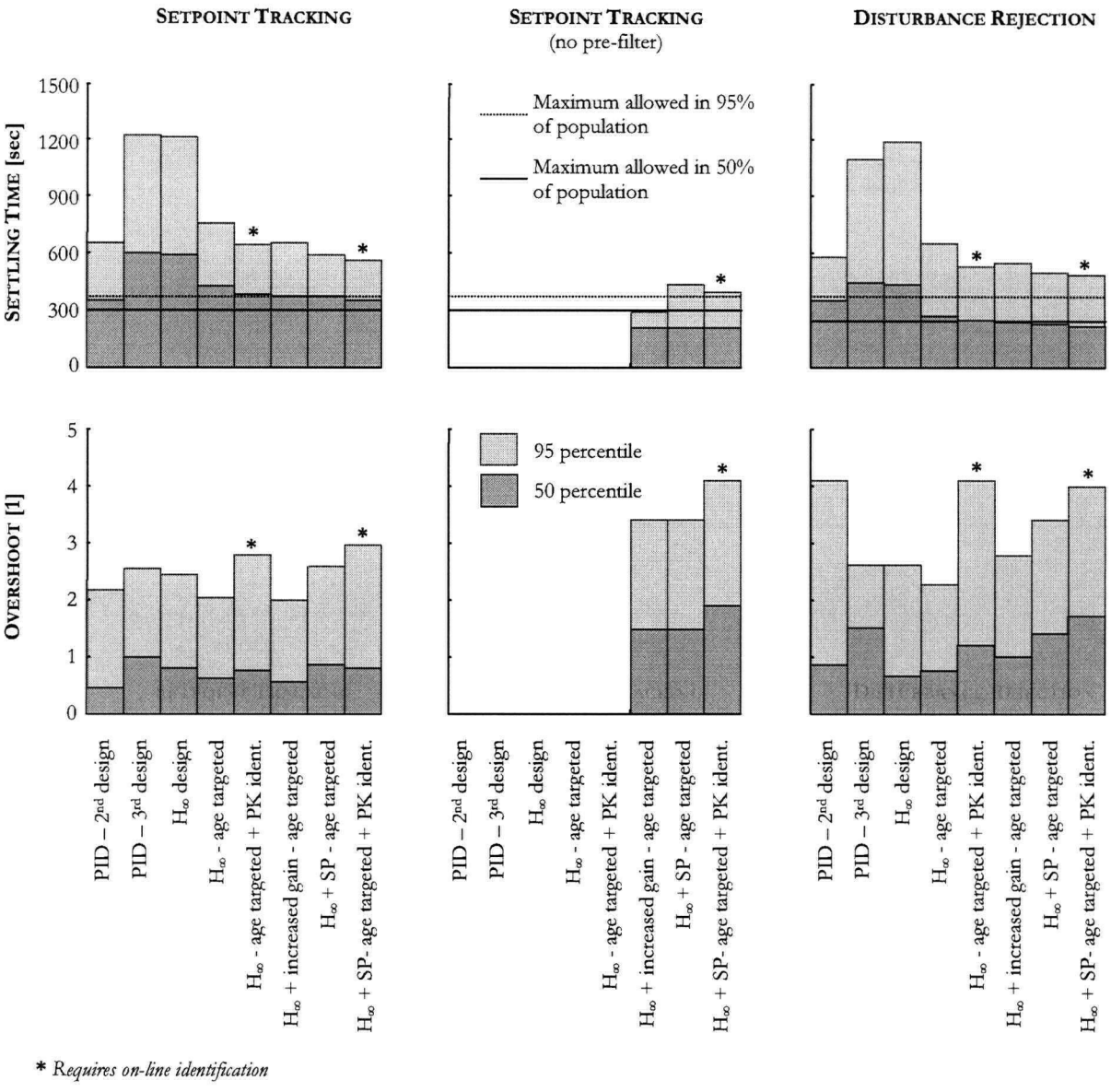


Figure 8.31: Summary of the controllers performance in terms of settling time and overshoot.

Chapter 9

Conclusion, Contributions and Recommendations

This concluding chapter reviews the work done under this Ph.D. program, emphasizing in particular the significance and contributions of this thesis. We also propose new research directions to further this work towards the development of a closed-loop system suitable for the everyday practice of clinical anesthesia in the OR and ICU.

9.1 Significance

The idea of automating anesthesia drug delivery dates back to the early 1950s. The visionary work of Bickford *et al.* already hinted at the tremendous potential of closed-loop systems for anesthesia, although, fifty years later, anesthetic drugs are still delivered manually in most of North American hospitals and clinical institutions.

Yet, the practice of anesthesia has made in recent years significant progress towards the automatization of drug delivery systems:

i. **Smart pumps**: infusion pumps driven by pharmacokinetic models have found their way into European ORs and ICUs, and are now used daily. These pumps, referred to as TCI (Target Controller Infusion), are able to calculate and constantly adjust the infusion profile needed to maintain a desired plasma concentration. While the Food and Drug Administration (FDA) has not yet approved the use of TCI for anesthesia procedures in the United States, new clinical evidence of the efficacy of these systems may reverse the FDA decision in the near future. Yet, in the absence of a real time feedback measure of the drug plasma concentration, TCI pumps simply remain open loop systems. Their success in achieving clinically relevant performance can only be explained by the reasonable predictive capability of population-normed pharmacokinetic models, and the inherent stability of PKPD systems.

ii. **New sensors**: the development of sensors that can quantify drug effect has been (and still is!) one of the major challenge in the anesthesia field. In 1996, the commercial availability of the BIS monitor to quantify hypnosis has generated a new momentum in the research community. Since then, similar monitors using different technologies have been introduced.

iii. **Faster drugs**: after the development of intravenous drugs presenting fast onset and short duration of action, drug companies are now investigating drugs with ultra-short duration of action. The fast pharmacokinetic capabilities of these drugs will make it increasingly difficult for the anesthesiologists to provide an adequate titration, as it will require dedicated and unswerving attention. The use of boluses will not be recommended as they would result in too deep and too short of an effect. Conversely, the fast dynamics and large bandwidth of these new compounds will make it increasingly easy for a closed-loop controller to achieve clinically relevant results.

iv. **Closing the loop**: finally, since the late 1990s, there has been a profound renewal of interest in the anesthesia field for automated drug delivery. In Europe, research groups from Gent (Belgium), Bern (Switzerland), Paris (France) and Glasgow (United Kingdom) have developed closed-loop systems that are used daily in their clinical facilities. For instance, Dr. Liu, an anesthesiologist from Hospital Foch (Paris, France), has already successfully closed the loop on more than 500 patients.

It is likely that closed-loop systems will find their way out of research labs, and into daily practice. The widespread use of TCI systems (actuators) and cortical monitors (sensors) are paving the way towards the anesthesiologists' acceptance of fully automated drug delivery systems. In addition, the development of new ultra-short acting compounds will motivate drug companies to investigate better ways of administrating drugs. Finally, as more clinical evidence appear in the literature as to the feasibility and efficacy of closed-loop systems, it appears likely that anesthesiologists will start wanting to benefit from these new technologies, as they will become more reliable and available.

To provide optimal control performance, *it appears essential that every aspect of the anesthesia closed-loop system be investigated from a control engineering point of view.* This thesis work is therefore set at a time when clinical researchers are seeking control engineering know-how to design and implement closed-loop applications.

9.2 Synopsis and Contributions

The main issue addressed in this thesis is the feasibility of designing a closed-loop control system providing clinically acceptable performance, while guaranteing stability in a large population of patients (18-60 yrs old, 50-110 kg). Due to the specificities of this application, it is further required that the control design be simple to implement, operate and monitor, thus ruling out any self-tuning, on-line identification, or other gain scheduling schemes.

9.2.1 Synopsis

It is essential for control engineers to become familiar with their application field before attempting to close the loop. In that respect, we first started this work by thoroughly reviewing the anesthesia field and understanding its specificities and needs. As such, the first 4 chapters were an introduction to clinical anesthesia targeting an engineering audience. The concepts described in Chapter 1, while simplistic from a clinician's point of view, presents an accurate picture of this field. In Chapter 2, the current state-of-the-art in anesthesia monitoring were reviewed. Special emphasis was made on the advantages of the new cortical monitors, and their technological limitations. The basic pharmacology principles concerning drug uptake, distribution and effect were reviewed in Chapter 3. Finally, we briefly reviewed the prior art in terms of closed-loop control in Chapter 4.

The performance of any closed-loop system is intimately related to that of the feedback sensor. We found that available anesthesia monitors, such as the BIS monitor, were not suitable for this application, as they use decisional schemes that can change the dynamic behavior of the index depending on the patient's state. We therefore investigated in Chapter 5 the use of the wavelet transform to rapidly characterize changes in the patient's state, and avoid the limitations of existing technologies. This investigation resulted in the development of the Wavelet-based Anesthesia Value (WAV) technology, which was first applied to the quantification of cortical activity (WAV_{CNS}), and autonomic activity (WAV_{ANS}).

Following this work, we used the WAV_{CNS} to derive the pharmacodynamic (PD) model of propofol in Chapter 6. This investigation first established that the traditional approach used by pharmacologists to derive PD models is severely flawed. We therefore proposed a new approach, which makes use of our new cortical sensor. We derived the PD model of 44 patients spanning the whole adult population (18-60 yrs old, 50-110 kg). These models were validated using standard system identification validation tools. Generic nominal models were further established from the 44 individual models.

The physiological response of patients to the administration of drugs is known to present a very large variability. Deriving controllers based on population-normed models may therefore lead to instability in some cases. Due this potential limitation, any closed-loop system would have to be thoroughly tested before being released. Considering the FDA decision concerning TCI pumps, it is doubtful that any amount of clinical tests will be able to convince this regulatory board of the safety of this type of system. It appears therefore essential to quantify inter- and intra-individual variability as system uncertainty in order to *mathematically prove* robust stability. This analysis is carried out in Chapter 7 using the 44 individual PD models derived in Chapter 6. We found that uncertainty is a major limiting factor of any closed-loop design. However, we also found several different ways to limit this uncertainty to more manageable levels. Depending on the operating condition of the closed-loop system, a number of nominal PKPD models have been derived. Each model was further associated with an uncertainty weight to quantify patients' variability. These models are fully disclosed in the Appendix C.

Using the preceding results, we derived in Chapter 8 a number of controllers that are mathematically

stable, using an H_∞ design procedure. We came to the conclusion that a *stable* general purpose controller designed for the whole adult population does not yield adequate performances. However, using the insight gained in Chapter 7, we found that targeting the age group of the patient, and constraining the drug administration to the 0-0.8 $mg \cdot s^{-1}$ infusion range, were sufficient to significantly limit the uncertainty. Under the assumption that the patients' age group is known and that the controller is sufficiently constrained, we found that the closed-loop system yields clinically relevant performance. In the majority of cases, large disturbance and setpoint changes can be dealt with in under 5 minutes, and under 5% overshoot.

In this thesis, we concluded that population-normed controllers can be derived and used effectively to track setpoint changes and reject disturbances. In addition, these controllers are shown to be mathematically stable, even when using a conservative measure of system uncertainty.

9.2.2 Contributions and Implications

The main contributions and implications of this thesis are summarized below:

i. **The Wavelet-based Anesthesia Value (WAV)**: the development of the WAV technology was spurred by the need for an anesthesia sensor dedicated to closed-loop control. It was particularly important to develop a sensing technology which do not rely on any decisional scheme in order to avoid the pitfall of other technologies. To that end, we have used the time-frequency localization property of wavelets to rapidly characterize changes in electrophysiological signals.

   Applied to the electroencephalogram (EEG), we have shown that this method yields remarkable results when compared to the BIS monitor. The advantage of our approach is that the EEG signal is processed the same way whether the patient is awake or comatose. As a result, the WAV_{CNS} sensor is the only index of cortical activity that is fully characterized by an LTI transfer function across its whole operating range.

   As of today, we have collected data from more than 200 anesthesia procedures and 30 volunteers during sleep studies. These data are used to further our understanding of the effect of anesthetic drugs onto cortical activity, and develop a commercial application that integrates the WAV technology. This development and commercialization effort is being spearheaded since 2003 by Ms. Zikov, a co-inventor of the WAV technology.

   We have also shown that the same technology can be applied to the quantification of the autonomic activity, through the analysis of the heart-rate variability (HRV) signal. If our preliminary results hold true, the WAV_{ANS} may very well become the first 'analgesia' index.

   The WAV technology may therefore play an important role in the near future, both from a standard monitoring and feedback sensor point of views.

   The WAV_{CNS} technology has been patented and should be made commercially available in 2006.

We would like to bring to the readers' attention that the contributions for the WAV_{CNS} are shared with Ms. Tatjana Zikov. Contributions for the WAV_{ANS}, however, are original.

ii. **System-oriented Pharmacodynamic Model**: the traditional pharmacological approach to pharmacodynamic (PD) modeling is seriously flawed. We therefore proposed a new modeling approach based on basic system identification know-how. In particular, this new approach incorporates a distinct model for the sensor dynamics. In addition, the LTI part of the PD model now includes a time delay to characterize the travel time of the drug.

The identification procedure carried out in Chapter 6 yielded white residuals when using the new system-oriented modeling approach, thereby indicating that the model is a good representation of the system dynamic behavior. We also found that the fit between measured and predicted WAV_{CNS} time courses was significantly improved with the new approach when compared to the traditional approach. Furthermore, we found that the new system-oriented models present less intrinsic uncertainty than the traditional models. This reduction in uncertainty is mainly the result of a weaker non-linear characteristic, as more of the system dynamics is now expressed in the LTI element of the model.

The third, and probably most compelling, advantage of this new modeling approach is that very limited amount of data is necessary for identifying properly the parameters of the PD model. We have shown that it is possible to use the induction data (2-3 minutes data window) to identify the PD model of propofol, and obtain similar results in terms of the dose *vs.* response relationships to those derived based on a much longer identification window (about 90 minutes). This effectively empowers anesthesiologists to identify pharmacodynamic and interaction models by using only data obtained during routine anesthesia procedures. Conversely, the traditional approach usually necessitates the enrollment of volunteers, which implies many ethical hurdles.

iii. **Intra- and Inter-individual Variability Analysis - Implication for Control**: patient variability has always been cause for concern in pharmacology. This has been, and still is, one of the main argument brought forth by opponents of automation in anesthesia.

Proven systems like TCI pumps have still not be approved by the FDA regulatory committee based on the fact that these open loop systems may not yield the expected setpoint targeted by the user. One would therefore expect that a closed-loop system may be easier to approve, providing that it remains well behaved for any patient fitting a certain profile.

However, obtaining regulatory approval will necessitate very large multi-center trials. As cases of instability are reported, the controller(s) will have to be re-designed and re-tested. This potentially involves a very long and expensive process.

A better approach is to quantify PKPD patient-variability as system uncertainty. This characterization can then be directly used in the design process to ensure stability. Risks of having unstable

cases are therefore significantly diminished, while giving the regulatory board a mathematical proof for asserting stability claims.

In parallel, the efficacy of uncertainty reduction methods can also be directly quantified. This gives valuable information to control designers as to which uncertainty reduction technique should be employed.

To my knowledge, this thesis is the first publication presenting such PKPD uncertainty analysis.

iv. **Robust Control Design**: to my knowledge, the issue of robust stability for anesthesia control has not yet been thoroughly discussed in the literature. Even though we have not yet closed the loop in the OR, our control design by itself merits consideration, as we have proven its stability and performance under very conservative assumptions.

9.3 Future Work

We propose here some directions for future research and development work.

i. **Improved $\mathbf{WAV_{CNS}}$ Sensor**: the success of closed-loop systems heavily depends on the feedback sensors. Improving their reliability, accuracy and dynamic response is therefore a priority.

In terms of the WAV_{CNS}, the following aspects should be investigated:

(a) 'Automatic' WAV_{CNS}: one inherent limitation of the WAV_{CNS} is that the index strongly reacts to the presence of ocular artifacts (OAs). Awake patients who are blinking and/or moving their eyes can have a WAV_{CNS} as low as 20. In order to solve this issue, we have developed and patented an artifact denoising algorithm to remove OAs patterns from the EEG. The so-called denoised EEG can then be used to quantify the cortical state of awake patients.

However, during anesthesia induction, the anesthesiologist must turn off the denoising process in order to preserve the EEG information. Failure to do so results in an elevated WAV_{CNS} index, which does not reflect the patient's cortical state. In that respect, all competing technologies have an automatic detection of the loss of consciousness event. In order to use the WAV_{CNS} as an anesthesia monitor to gain clinical credibility, it is then necessary to design an algorithm that can automatize the denoising switch[1].

While this may not appear as a critical issue for a closed-loop control system dedicated for the maintenance of anesthesia, this automatization issue must be resolved in order for the WAV_{CNS} to be used as a stand alone monitor, and be accepted by the clinical community.

(b) Increasing the sensor bandwidth: the WAV_{CNS} IIR trending filter is less than a decade above the fastest PKPD dynamics. Improving the dynamic response of the index will therefore result

[1]This issue has recently been resolved satisfactorily.

in improved control performance. Unfortunately, measurement noise is also quite significant. Any increase in the filter dynamic response will also increase the closed-loop noise, and make it increasingly difficult for the anesthesiologists to monitor the controller output. Hence, new methods for reducing the measurement noise should be investigated. Only by then, an increase in the sensor bandwidth may be envisaged.

It is important at this point to notice that the measurement noise may very well represent true cortical changes.

(c) Improved hardware for use in the OR: the OR is a particularly difficult environment for EEG acquisition devices. Electro-Surgical Units (ESU) create large amplitude widespread interferences. These interferences cause the saturation of the pre-amplifier stage of the EEG acquisition module. In most cases, the EEG signal cannot be retrieved, which results in discontinuities in the WAV sensor output. During periods of intense electrocautery activity, a closed-loop system may loose its feedback measure. Limiting ESU interference should therefore be given the priority in any hardware design dedicated for OR use and closed-loop control.

The WAV_{CNS} development efforts led by Ms. Zikov have already shown that a proper hardware design including a pre-amplifier notch filter centered on the ESU frequency could significantly lower the impact of ESU onto the amplifier. While total rejection of the ESU has not yet been achieved, encouraging results were already obtained.

(d) Clinical endorsement: it is imperative for the success of any closed-loop system to use feedback sensors that have been thoroughly validated and tested by the clinical community. Only trust in the sensors will enable trust in the control system. Clinical endorsement can only be obtained by conducting large scale studies where the WAV_{CNS} is used as a controlled endpoint.

The full disclosure of the WAV_{CNS} algorithm is also encouraged. It is important for clinicians to understand the main concepts used in the algorithm, and also understand its main role (*i.e.*, quantifying cortical activity). Any 'black box' solution will always be regarded with mistrust.

ii. **WAV_{ANS} Development**: in terms of the analgesia index, significant progress still need to be achieved.

The main difficulty is the design and conduct of a research protocol that allows the observation of the effect of opioids and β-blockers, as well as the effect of surgical stimulation, onto the WAV_{ANS}. This protocol should provide the clinical data needed to derive more precisely the WAV_{ANS}, and prove its clinical significance. In addition, the WAV_{ANS} project will also benefit significantly from ECG hardware that includes ESU protection.

Furthermore, one of the main hurdle when using the HRV signal is the very low time resolution of the signal itself. This may be improved by considering the HRV signal derived based on the T peak, as well as the R peak (in which case the time resolution may be increased by up to a factor 2).

Finally, the effect of respiration on the HRV signal should be thoroughly studied. We are indeed expecting that a change in the respiration pattern may induce a change in the HRV, which might then be misinterpreted as a change in the autonomic activity due to surgical stimulation or drug administration.

iii. **New PD Models**: we have shown in Chapter 6 that our new modeling approach allows the use of induction data for PD identification. These data can be obtained from virtually any elective surgery involving the use of the minimally invasive LMA. While the PD study carried out in Chapter 6 is already the second largest propofol PD study[2], it would be advised to increase the number of cases in the 40-49 and 50-60 yrs age groups in order to reach a more statistically representative result.

In parallel, a clinical study aimed at deriving a WAV_{CNS} interaction model between propofol and remifentanil should also be carried out. This study would involve a group of patients receiving a mixed propofol and remifentanil induction bolus. The PD interaction model derived from this study will be helpful to design a MISO controller (see below). It would also be the first dynamic model capturing the synergy between propofol and remifentanil.

iv. **Positive Systems**: one aspect of PKPD models which has not been discussed in Chapter 7 is that PK models are essentially describing a positive system (*i.e.*, a system whose states cannot take negative values). A potentially very interesting property of such systems is that the unstructured complex uncertainty can be reduced to a real bounded uncertainty gain. This may describe more appropriately the PKPD uncertainty and remove most of the conservatism introduced by the method used in Chapter 6.

v. **Clinical Survey of Adequate Control Performance** : the initial survey carried out in Section 8.5 has shown the large discrepancy between control engineering and clinical expectations in terms of close loop system performance. Before further designing, testing, and implementing new control laws, it will be necessary to clearly evaluate the needs of clinicians with regard to setpoint changes and surgical disturbance rejection. This may involve the definition of new control performance specifications and new testing procedures.

vi. **Patient Simulator for Control Validation**: another important project concerns the development of a Patient Simulator (PS) that can be used as a test bed for validating control laws and testing the end-user interface. This simulator should be able to synthesize EEG and ECG signals that can be used to calculate the WAV_{CNS} and WAV_{ANS} indexes.

A virtual embedded PKPD model calculates the patient's cortical and autonomic state, and generates appropriate EEG and ECG signals. The PS should also be able to simulate disturbances initiated by the user. A flow transducer attached to the infusion pump cannula can further be used to determine

[2]Kazama *et al.*'s study involved 47 patients

the infusion rate and bolus administration delivered manually by the user. This information can be input to the PS that will then re-calculate the appropriate patient response.

A variety of faults and clinical situations can thus be easily simulated. This system can also be used to test control laws, the synergy between the human operator and the close loop controller, the feedback sensors and the infusion pumps. This setup may be a necessary first step towards the validation of a control law, and its implementation and trial on human subjects.

vii. **Closed-loop Drug Delivery for ICU**: we have established in Chapter 7 that uncertainty is significantly limited when considering the WAV_{CNS} range of 50-80. This range typically corresponds to the light to deep sedation, which is commonly targeted in the ICU. Controllers designed specifically to maintain the index within this range should yield very good performance while being particularly robust to inter-patient variability. In addition, disturbances from surgical stimulation, or the co-administration of other anesthesia drugs, should be minimum. The ICU may therefore be the optimal test bed to test and validate closed-loop systems, and obtain valuable feedback information from anesthesiologists and other end-users of this system.

viii. **MISO/MIMO Control Designs**: the sole control of hypnosis (or analgesia) will not provide anesthesiologists with the full benefits of automatization. Following what has been presented in the Chapter 1, controlling anesthesia implies the control of the three endpoints that are hypnosis, analgesia and paralysis. While paralysis can be viewed as a separate issue, the control of hypnosis and analgesia must be approached within a multivariate framework to account for the opioid/anesthetic drug interaction. While the sole control of hypnosis was the goal of most of the 1990s research, the dual control of hypnosis and analgesia is now considered by many [91, 143, 169] to be next paradigm in research.

We expect that the remifentanil PKPD bandwidth (with respect to the WAV_{CNS}) may be significantly higher than that of propofol. A MISO controller designed for the control of the WAV_{CNS}, and based on the co-administration of both propofol and remifentanil, may therefore provide faster tracking and disturbance rejection, while minimizing drug dosage. During setpoint changes (or output disturbances), the controller will first use remifentanil to compensate for the high frequency transients, while propofol will only be used to compensate for the lower frequency transients, as well as dc shifts. Since remifentanil is also used to provide analgesia, it will be important to ensure that the remifentanil infusion rate settles back to the anesthesiologist's setpoint once the disturbance is compensated. The remifentanil control action will also have to be constrained.

Upon completion of the WAV_{ANS} development, a MIMO control law targeting both WAV indexes may be derived, thus achieving full closed-loop Total IntraVenous Anesthesia (TIVA).

ix. **Anesthesia Display**: before clinical trust in a closed-loop system can be earned, it is necessary to provide the anesthesiologists with an intuitive display to help them monitor the controller actions.

This display can be used as an advisory system that will help them set up a proper titration strategy. This display can also be used to predict the time needed to reach a particular setpoint.

Bibliography

[1] G. Rushman, N. Davies, and R. Atkinson, *A short history of anaesthesia: The first 150 years.* Reed Educational and Professional Publishing Ltd, first ed., Jan. 1996.

[2] R. Stoelting and R. Miller, *Basics of Anesthesia.* Philadelphia, Pennsylvania: Churchill Livingston, fourth ed., 2000.

[3] R. Kitz and L. Vandam, *Scope of Modern Anesthetic Practice*, chapter 1, pp. 3–23. Churchill Livingston, third ed., 1990.

[4] P. Knight and D. Bacon, "An unexplained death: Hannah greener and chloroform," *Anesthesiology*, vol. 96, pp. 1250–1253, May 2002.

[5] D. Brown, *Risk and outcome in anesthesia.* JB Lippincott, 2nd ed., 1992.

[6] A. Spence, "The lessons of CEPOD," *British Journal of Anaesthesia*, vol. 60, p. 753, 1988.

[7] R. Wiklund and S. Rosenbaum, "Anesthesiology," *New England Journal of Medicine*, vol. 337, pp. 1132–1141 and 1215–1219, 1997.

[8] F. Murphy, *Hazards of anesthesia*, pp. 476–484. Philadelphia, Pennsylvania: W.B. Saunders Company, 1996.

[9] A. Ross and J. Tinker, *Anesthesia Risk*, chapter 22, pp. 715–742. Churchill Livingston, third ed., 1990.

[10] J. Tracy, *Awareness in the operating room: a patient's view*, pp. 349–353. Englewood Cliffs, NJ: Prentice Hall, 1993.

[11] J. Jones, "Perception and memory during general anaesthesia," *British Journal of Anaesthesia*, vol. 73, pp. 31–37, July 1994.

[12] R. Sandin, G. Enlund, P. Samuelsson, and C. Lennmarken, "Awareness during anaesthesia: a prospective case study," *The Lancet*, vol. 355, pp. 707–711, 2000.

[13] P. Sebel, T. Bowdle, M. Ghoneim, I. Rampil, and R. Padilla, "The incidence of awareness during anesthesia: A multicenter U.S. study," *Anesthesiology*, pp. A–360, 2003.

[14] C. Lennmarken, K. Bildfors, G. Enlund, P. Samuelsson, and R. Sandin, "Victims of awareness," *Acta Anaesthesiologica Scandinavia*, vol. 43, no. 3, pp. 229–231, 2002.

[15] J. Osterman, J. Hopper, W. Heran, T. Keane, and B. Van Der Kolk, "Awareness under anesthesia and the development of posttraumatic stress disorder," *General Hospital Psychiatry*, vol. 23, no. 4, pp. 198–204, 2001.

[16] C. Prys-Roberts, "Anaesthesia: a practical or impractical construct?," *British Journal of Anaesthesia*, vol. 59, p. 1341, Nov. 1987.

[17] T. Gray and G. Rees, "The role of apnoea in anaesthesia for major surgery," *British Medical Journal*, vol. 2, pp. 891–892, 1952. in Johansen: 2000.

[18] A. Gentilini, C. Frei, A. Glattfelder, M. Morari, T. Sieber, R. Wymann, T. Schnider, and A. Zbinden, "Closed-loop control in anesthesia," tech. rep., ETH, Zürich, Switzerland, July 2000.

[19] J. Bovill, *Opioid Anesthesia*, vol. 21 of *Monographs in Anaesthesiology*, ch. 4, pp. 81–102. Amsterdam, The Netherlands: Elsevier Science Publisher, 1991.

[20] E. Lang, A. Kapila, and D. Shlugman, "Reduction of isoflurane minimum alveolar concentration by remifentanil," *Anesthesiology*, vol. 85, pp. 721–728, 1996.

[21] T. Egan, H. Lemmens, P. Fiset, D. Hermann, K. Muir, D. Stanski, and S. Shafer, "The pharmacokinetics of the new short-acting opioid Remifentanil (GI87084B) in healthy adult male volunteers," *Anesthesiology*, vol. 79, pp. 881–892, May 1993.

[22] A. Quasha, E. Eger, and J. Tinker, "Determination and application of MAC," *Anesthesiology*, vol. 53, pp. 315–334, 1980.

[23] J. Vuyk, M. Mertens, E. Olofsen, A. Burm, and J. Bovill, "Propofol anesthesia and rational opioid selection: determination of optimal EC_{50}-EC_{95} propofol-opioid concentrations that assure adequate anesthesia and a rapid return of consciousness," *Anesthesiology*, vol. 87, pp. 1549–1562, 1997.

[24] J. Vuyk, "Clinical interpretation of pharmacokinetic and pharmacodynamic propofol-opioid interactions," *Acta Anaesthesiologica Belgica*, vol. 52, pp. 445–451, 2001.

[25] C. Minto, T. Schnider, T. Short, K. Gregg, A. Gentilini, and S. Shafer, "Response surface model for anesthetic drug interactions," *Anesthesiology*, vol. 92, pp. 1603–1616, 2000.

[26] M. Morgan, "Total intravenous anaesthesia," *Anaesthesia*, vol. 38, pp. 1–9, 1983.

[27] F. Camu and B. Kay, *Why Total Intravenous Anesthesia (TIVA)*, vol. 21 of *Monographs in Anaesthesiology*, ch. 1, pp. 1–14. Amsterdam, The Netherlands: Elsevier Science Publisher, 1991.

[28] S. Suttner, J. Boldt, C. Schmidt, S. Piper, and B. Kumle, "Cost analysis of target-controlled infusion-based anesthesia compared with standard anesthesia regimen," *Anesthesia Analgesia*, vol. 88, pp. 77–82, 1998.

[29] B. Allan and I. Smith, "Cost considerations in the use of anaesthetic drugs," *Current Opinion in Anaesthesiology*, vol. 15, pp. 227–232, 2002.

[30] R. Dripps, J. Eckenhoff, and L. Vandam, eds., *Introduction to anesthesia.* Philadelphia, Pennsylvania: W.B. Saunders Company, ninth ed., 1997.

[31] V. Billard and S. Shafer, *Does the EEG Measure Therapeutic Opioid Drug Effect?*, pp. 79–95. Berlin: Springer-Verlag, 1995.

[32] D. Newton, C. Thornton, and C. Jordan, *The auditory evoked response as a monitor of anesthetic depth*, pp. 274–280. Englewood Cliffs, NJ: Prentice Hall, 1993.

[33] D. Stanski, *Monitoring Depth of Anesthesia*, chapter 30, pp. 1001–1029. Churchill Livingston, third ed., 1990.

[34] C. Pomfrett and A. Pearson, "EEG monitoring using bispectral analysis," *Engineering Science and Education Journal*, pp. 155–157, Aug. 1998.

[35] M. Rubin and H. Freeman, "Brain potential changes in man during cyclopropane anesthesia," *Journal of Neurophysiology*, vol. 3, pp. 33–42, Jan. 1940.

[36] A. Faulconer and R. Bickford, *Electroencephalography in Anesthesiology*. Springfield, Illinois: Charles C Thomas, 1960.

[37] D. Clark and B. Rosner, "Neurophysiologic effects of general anesthetics. i. the electroencephalogram and sensory evoked responses in man," *Anesthesiology*, vol. 38, pp. 564–582, June 1973.

[38] R. Bickford, "Automatic electroencephalographic control of general anesthesia," *Electroencephalography and Clinical Neurophysiology*, vol. 2, pp. 93–96, 1950.

[39] R. Bickford, "Use of frequency discrimination in the automatic electroencephalographic control of anesthesia (servo-anesthesia)," *Electroencephalography and Clinical Neurophysiology*, vol. 3, pp. 83–86, 1951.

[40] D. Soltero, A. Faulconer, and R. Bickford, "The clinical application of automatic anesthesia," *Anesthesiology*, vol. 12, pp. 574–582, Sept. 1950.

[41] J. Bellville and G. Attura, "Servo control of general anesthesia," *Science*, vol. 126, pp. 827–830, Oct. 1957.

[42] I. Pichlmayr, U. Lips, and H. Künkel, *The Electroencephalogram in Anesthesia*. Berlin: Springer-Verlag, 1984.

[43] H. Schwilden and H. Stoeckel, "Quantitative EEG analysis during anaesthesia with isoflurane in nitrous oxide at 1.3 and 1.5 MAC," *British Journal of Anaesthesia*, vol. 59, pp. 738–745, June 1987.

[44] H. Schwilden, J. Schüttler, and H. Stoeckel, "Closed-loop feedback control of methohexital anesthesia by quantitative EEG analysis in humans," *Anesthesiology*, vol. 67, pp. 341–347, Sept. 1987.

[45] H. Schwilden, H. Stoeckel, and J. Schüttler, "Closed-loop feedback control of propofol anesthesia by quantitative EEG analysis in humans," *British Journal of Anaesthesia*, vol. 62, pp. 290–296, Mar. 1989.

[46] J. Schüttler and H. Schwilden, *Feedback Control of Intravenous Anaesthetics by Quantitative EEG*, pp. 194–207. Berlin: Springer-Verlag, 1995.

[47] I. Rampil, F. Sasse, N. Smith, B. Hoff, and D. Flemming, "Spectral edge frequency - a new correlate of anesthetic depth," *Anesthesiology*, vol. 53, p. S12, Sept. 1980.

[48] W. J. Levy, "Intraoperative EEG patterns: Implications for EEG monitoring," *Anesthesiology*, vol. 60, pp. 430–434, May 1984.

[49] A. Sharma, S. Wilson, and R. Roy, "Autoregressive modeling of EEG signals for monitoring anesthetic levels," in *Proceedings of IEEE Bioengineering Conference*, pp. 39–40, 1992.

[50] A. Sharma and R. Roy, "Design of a recognition system to predict movement during anesthesia," *IEEE Transactions on Biomedical Engineering*, vol. 44, pp. 505–511, June 1997.

[51] I. Rampil, "A primer for EEG signal processing in anesthesia," *Anesthesiology*, vol. 89, pp. 980–1002, Oct. 1998.

[52] S. Bowles, P. Sebel, V. Saini, and N. Chamoun, *Effects of anesthesia on the EEG-bispectral analysis correlates with movement*, pp. 249–254. Englewood Cliffs, NJ: Prentice Hall, 1993.

[53] T. Ning and J. Bronzino, "Bispectral analysis of the rat EEG during various vigilance states," *IEEE Transactions on Biomedical Engineering*, vol. 36, pp. 497–499, Apr. 1989.

[54] L. Kearse, V. Saini, F. deBros, and N. Chamoun, "Bispectral analysis of EEG may predict anesthetic depth using narcotic induction," *Anesthesiology*, vol. 3A, p. A175, Sept. 1990.

[55] P. Sebel, S. Bowles, V. Saini, and N. Chamoun, "Accuracy of EEG in predicting movement at incision during isoflurane anesthesia," *Anesthesiology*, vol. 3A, p. A446, Sept. 1990.

[56] J. Vernon, S. Bowles, P. Sebel, and N. Chamoun, "EEG bispectrum predicts movement at incision during isoflurane or propofol anesthesia," *Anesthesiology*, vol. 77, no. 3A, p. A502, 1992.

[57] J. Muthuswamy and R. Roy, "Bispectrum analysis of EEG of a dog to determine the depth under halothane anesthesia," in *Proceedings of IEEE Bioengineering Conference*, pp. 5–6, 1993.

[58] J. Liu, H. Singh, and P. White, "Electroencephalographic bispectral index correlates with intraoperative recall and depth of propofol-induced sedation," *Anesthesia Analgesia*, vol. 84, pp. 185–189, 1997.

[59] T. Zikov, S. Bibian, G. Dumont, M. Huzmezan, and C. Ries, "A wavelet based de-noising technique for ocular artifact correction of the electroencephalogram," in *Proceedings of the 24^{th} Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2002.

[60] G. Plourde and T. Picton, "Human auditory steady-state response during general anesthesia," *Anesthesia Analgesia*, vol. 71, pp. 460–468, Nov. 1990.

[61] G. Plourde and J. Boylan, "The auditory steady state response during sufentanil anaesthesia," *British Journal of Anaesthesia*, vol. 66, pp. 683–691, 1991.

[62] C. Villemure, G. Plourde, I. Lussier, and N. Normandin, *Auditory processing during isoflurane anesthesia: a study with an implicit memory task and auditory evoked potentials*, pp. 99–106. Englewood Cliffs, NJ: Prentice Hall, 1993.

[63] D. Schwender, C. Madler, I. Keller, S. Klasing, K. Peter, and E. Pöppel, *Midlatency auditory evoked potentials indicate wakefulness during cesarean section*, pp. 333–342. Englewood Cliffs, NJ: Prentice Hall, 1993.

[64] J. Huang, Y.-Y. Lu, A. Nayak, and R. Roy, "Depth of anesthesia estimation and control," *IEEE Transactions on Biomedical Engineering*, vol. 46, pp. 71–81, Jan. 1999.

[65] A. Angel, D. Linkens, and C. Ting, "Estimation of latency changes and relative amplitudes in somatosensory evoked potentials using wavelets and regression," *Computers and Biomedical Research*, vol. 32, pp. 209–251, 1999.

[66] E. Jensen, P. Lindholm, and S. Henneberg, "Autoregressive modeling with exogenous input of middle-latency auditory-evoked potentials to measure rapid changes in depth of anesthesia," *Method in Informatics and Medicine*, vol. 35, pp. 256–260, 1996.

[67] L. Capitanio, E. Jensen, G. Filligoi, B. Makovec, M. Gagliardi, and S. Henneberg, "On-line analysis of AEP and EEG for monitoring the depth of anaesthesia," *Methods in Informatics and Medicine*, vol. 36, pp. 311–314, 1997.

[68] J. Kugler, *Elektroenzephalographie in klinik und praxis*. Stuttgart, New York, Thieme, 1981.

[69] H. Viertio-Oja, V. Maja, M. Sarkela, P. Talja, N. Tenkanen, H. Tolvanen-Laakso, M. Paloheimo, A. Vakkuri, A. Yli-Hankala, and P. Merilainen, "Description of the entropy algorithm as applied in the Datex-Ohmeda S/5 entropy module," *Acta Anaesthesiologica Scandinavia*, vol. 48, pp. 154–161, 2004.

[70] S. Kreuer, A. Biedler, R. Larsen, S. Altmann, and W. Wilhelm, "Narcotrend monitoring allows faster emergence and a reduction off drug consumption in propofol-remifentanil anesthesia," *Anesthesiology*, vol. 99, pp. 34–41, 2003.

[71] M. Luginbühl, S. Wüthrich, S. Petersen-Felix, A. Zbinden, and T. Schnider, "Different benefit of bispectral index (BIS^{tm}) in desflurane and propofol anesthesia," *Acta Anaesthesiologica Scandinavia*, vol. 47, pp. 165–173, 2003.

[72] C. Rosow and P. Manberg, "Bispectral Index monitoring," *Anesthesiology Clinics of North America: Annual of Anesthetic Pharmacology*, vol. 19, no. 4, pp. 947–966, 2001.

[73] T. Gan, P. Glass, and A. Windsor, "BIS utility study group. bispectral index monitoring allows faster emergence and improved recovery from propofol, alfentanil, and nitrous oxide anesthesia," *Anesthesiology*, vol. 87, pp. 808–815, 1997.

[74] J. Mayfield and J. Quigley, "BIS monitoring reduces phase I PACU admissions in an ambulatory surgical unit," *Anesthesiology*, vol. 91, no. 3A, pp. A–28, 1999.

[75] J. Peñuelas-Acuña, S. Oriol-Lopez, J. Castelazo-Arredondo, and C. Hernandez-Bernal, "Usefulness of bispectral index in pharmaceutical cost reduction for anesthesia," *Cirugía y Cirujanos*, vol. 71, pp. 300–303, 2003.

[76] B. Weldon, M. Mahla, M. Van Der Aa, and T. Monk, "Advancing age and deeper intraoperative anesthetic levels are associated with higher first year death rates," *ASA Meeting Abstract*, pp. A–1097, 2002.

[77] C. Lennmarken, R. Maj-Lis Lindholm, S. Greenwald, and R. Sandin, "Confirmation that low intraoperative BIS levels predict increased risk of post-operative mortality," *Anesthesiology*, pp. A–303, 2003.

[78] T. Dushane, "Cons: Monitoring the amnestic state during general anesthesia should not be a standard of care," *Anesthesiology*, pp. 509–512, Oct. 2000.

[79] J. Sleigh and J. Donovan, "Comparison of the bispectral index, 95% spectral edge frequency and approximate entropy of the EEG, with changes in heart rate variability during induction and recovery from general anaesthesia," *British Journal of Anaesthesia*, vol. 82, no. 5, pp. 666–671, 1999.

[80] A. Lehmann, J. Boldt, E. Thaler, S. Piper, and U. Weisse, "Bispectral index in patients with target-controlled or manually-controlled infusion of propofol," *Anesthesia Analgesia*, vol. 95, pp. 639–644, 2002.

[81] T. Schnider, M. Luginbuhl, S. Petersen-Felix, and J. Mathis, "Unreasonably low bispectral index values in a volunteer with genetically determined low-voltage electroencephalographic signal," *Anesthesiology*, vol. 89, no. 6, pp. 1607–1608, 1998.

[82] A. Muncaster, J. Sleigh, and M. Williams, "Changes in consciousness, conceptual memory, and quantitative electroencephalographical measures during recovery from sevoflurane- and remifentanil-based anesthesia," *Anesthesia Analgesia*, vol. 96, pp. 720–725, 2003.

[83] M. Messner, U. Besse, J. Romstöck, M. Dinkel, and K. Tschaikowsky, "The bispectral index declines during neuromuscular block in fully awake persons," *Anesthesia Analgesia*, vol. 97, pp. 488–491, 2003.

[84] B. Vivien, S. Di Maria, A. Ouattara, O. Langeron, P. Coriat, and B. Riou, "Overestimation of bispectral index in sedated intensive care unit patients revealed by administration of muscle relaxant," *Anesthesiology*, vol. 99, no. 1, pp. 9–17, 2003.

[85] R. Greif, S. Greenwald, E. Schweitzer, S. Laciny, A. Rajek, J. Caldwell, and D. Sessler, "Muscle relaxation does not alter hypnotic level during propofol anesthesia," *Anesthesia Analgesia*, vol. 94, pp. 604–608, 2002.

[86] D. Tempe and L. Satyanarayna, "Is there any alternative to the bispectral index monitor?," *British Journal of Anaesthesia*, vol. 92, pp. 1–3, 2004.

[87] X. Chen, J. Tang, P. White, R. Wender, H. Ma, and A. S. R. Kariger, "A comparison of patient state index and bispectral index values during the perioperative period," *Anesthesia Analgesia*, vol. 95, no. 6, pp. 1669–1674, 2002.

[88] N. Smith, H. Dec-Silver, T. Sanford, C. Westover, M. Quinn, F. Klein, and D. Davis, "EEGs during high-dose fentanyl-, sufentanil-, or morphine-oxygen anesthesia," *Anesthesia Analgesia*, vol. 63, pp. 386–393, Apr. 1984.

[89] J. Scott, K. Ponganis, and D. Stanski, "EEG quantitation of narcotic effect: The comparative pharmacodynamics of fentanyl and alfentanyl," *Anesthesiology*, vol. 62, pp. 234–241, Mar. 1985.

[90] J. Scott, J. Cooke, and D. Stanski, "Electroencephalographic quantitation of opiod effect: Comparative pharmacodynamics of fentanyl and sufentanil," *Anesthesiology*, vol. 74, pp. 34–42, Jan. 1991.

[91] P. Glass, "Why and how we will monitor the state of anesthesia in 2010?," *Acta Anaesthesiologica Belgica*, vol. 50, pp. 35–44, Jan. 1999.

[92] C. Pomfrett, "Heart rate variability, BIS and 'depth of anaesthesia'," *British Journal of Anaesthesia*, vol. 82, pp. 659–662, May 1999.

[93] C. Pomfrett, "Monitoring depth of anesthesia," *The Royal College of Anaesthetists. Bulletin 4*, pp. 155–157, Nov. 2000.

[94] D. Stanski and W. Watkins, *Drug disposition in Anesthesia*. Grune & Stratton, 1982.

[95] C. Prys-Roberts and C. Hug, *Pharmacokinetics of anesthesia*. Grune & Stratton, 1984.

[96] R. Fragen and M. Avram, *Barbiturates*, chapter 8, pp. 225–242. Churchill Livingston, third ed., 1990.

[97] H. Price, P. Kovnat, J. Safer, E. Conner, and M. Price, "The uptake of thiopental by body tissues and its relation to the duration of narcosis," *Clinical Pharmacology and Therapeutics*, vol. 1, pp. 16–22, 1960.

[98] D. Stanski and W. Watkins, *Pharmacokinetic Principles*, pp. 1–46. Grune & Stratton, 1982.

[99] P. Glass, S. Shafer, and J. Reves, *Intravenous drug delivery systems*, chapter 11, pp. 377–411. Churchill Livingston, fifth ed., 2000.

[100] T. Schnider, C. Minto, P. Gambus, C. Andresen, D. Goodale, S. Shafer, and E. Youngs, "The influence of method of administration and covariates on the pharmacokinetics of propofol in adult volunteers," *Anesthesiology*, vol. 88, pp. 1170–1182, May 1998.

[101] J. Schüttler and H. Ihmsen, "Population pharmacokinetics of propofol: a multicenter study," *Anesthesiology*, vol. 92, pp. 727–738, Mar. 2000.

[102] T. Egan, C. Minto, D. Hermann, J. Barr, K. Muir, and S. Shafer, "Remifentanil versus alfentanil: comparative pharmacokinetics and pharmacodynamics in healthy adult male volunteers," *Anesthesiology*, vol. 84, pp. 821–833, 1996.

[103] C. Minto, T. Schnider, T. Egan, E. Youngs, H. Lemmens, P. Gambus, V. Billard, J. Hoke, K. Moore, D. Hermann, K. Muir, J. Mandema, and S. Shafer, "Influence of age and gender on the pharmacokinetics and pharmacodynamics of remifentanil," *Anesthesiology*, vol. 86, pp. 10–23, Jan. 1997.

[104] D. Hermann, T. Egan, and K. Muir, "Influence of arteriovenous sampling on remifentnail pharmacokinetics and pharmacodynamics," *Clinical Pharmacology and Therapeutics*, vol. 65, no. 5, pp. 511–518, 1999.

[105] R. Upton and Y. Huang, "Influence of cardiac output, injection time and injection volume on the initial mixing of drugs with venous blood after i.v. blous administration to sheep," *British Journal of Anaesthesia*, vol. 73, . pp. 333–338, 1993.

[106] R. Upton, G. Ludbrook, C. Grant, and A. Martinez, "Cardiac output is a determinant of the initial concentrations of propofol after short-infusion administration," *Anesthesia Analgesia*, vol. 89, pp. 545–552, 1999.

[107] T. Krejcie and M. Avram, "What determines anesthetic induction dose? It's the front-end kinetics, Doctor!," *Anesthesia Analgesia*, vol. 89, pp. 541–544, 1999.

[108] Y. Adachi, K. Watanabe, H. Higushi, and T. Satoh, "The determinants of propofol induction of anesthesia dose," *Anesthesia Analgesia*, vol. 92, pp. 656–661, 2001.

[109] L. Briggs, M. White, I. Cockshott, and E. Douglas, "The pharmacokinetics of propofol (Diprivan) in female patients (Abstract)," *Postgraduate Medical Journal*, vol. 61, no. Suppl. 3, pp. 58–59, 1985.

[110] S. Gill, E. Wright, and C. Reilly, "Pharmacokinetic interaction of propofol and fentanyl: single bolus injection study," *British Journal of Anaesthesia*, vol. 65, pp. 760–765, 1990.

[111] E. Gepts, K. Jonckheer, V. Maes, W. Sonck, and F. Camu, "Disposition kinetics of propofol during alfentanil anaesthesia," *Anaestesia*, vol. 43, no. Suppl., pp. 8–13, 1988.

[112] M. Mertens, J. Vuyk, E. Olofsen, J. Bovill, and A. Burm, "Propofol alters the pharmacokinetics of alfentanil in healthy male volunteers," *Anesthesiology*, vol. 94, pp. 949–957, 2001.

[113] M. Ausems, C. Hug, D. Stanski, and A. Burm, "Plasma concentrations of Alfentanil required to supplement nitrous oxide anesthesia for general surgery," *Anesthesiology*, vol. 65, pp. 362–373, Oct 1986.

[114] L. Sheiner, D. Stanski, S. Vozeh, R. Miller, and J. Ham, "Simultaneously modeling of pharmacokinetics and pharmacodynamics: application to d-tubocurarine," *Clinical Pharmacology and Therapeutics*, vol. 25, pp. 358–371, 1979.

[115] M. White, M. Schenkels, F. Engbers, A. Vletter, A. Burm, J. Bovill, and G. Kenny, "Effect-site modeling of propofol using auditory evoked potentials," *British Journal of Anaesthesia*, vol. 82, no. 3, pp. 333–339, 1999.

[116] T. Kazama, K. Ikeda, K. Morita, M. Kikura, M. Doi, T. Ikeda, T. Kurita, and Y. Nakajima, "Comparison of the effect-site k_{e0}s of propofol for blood pressure and EEG bispectral index in elderly and younger patients," *Anesthesiology*, vol. 90, pp. 1517–1527, 1999.

[117] V. Billard, P. Gambus, N. Chamoun, D. Stanski, and S. Shafer, "A comparison of spectral edge, delta power, and bispectral index as EEG measures of alfentanil, propofol, and midazolam drug effect," *Clinical Pharmacology and Therapeutics*, vol. 61, no. 1, pp. 45–58, 1997.

[118] K. Kuizenga, J. Proost, J. Wierda, and C. Kalkman, "Predictability of processed electroencephalography effects on the basis on pharmacokinetic-pharmacodynamic modeling during repeated propofol infusions in patients with extradural analgesia," *Anesthesiology*, vol. 95, pp. 607–615, 2001.

[119] J. Larsson and G. Wahlstrom, "The influence of age and administration rate on the brain sensitivity to propofol in rats," *Acta Anaesthesiologica Scandinavia*, vol. 42, pp. 987–994, 1998.

[120] C. Smith, A. McEwan, and R. Jhaveri, "Reduction of propofol C_{p50} by fentanyl," *Anesthesiology*, vol. 77, p. A340, 1992.

[121] H. Vinik and I. Kissin, "Rapid development of tolerance to analgesia during remifentanil infusions in humans," *Anesthesia Analgesia*, vol. 86, pp. 1307–1311, 1998.

[122] M. Struys, T. D. Smet, L. Versichelen, S. V. de Velde, R. V. den Broecke, and E. Mortier, "Comparison of closed-loop controlled administration of propofol using BIS as the controlled variable versus "standard practice" controlled administration," *Anesthesiology*, vol. 95, pp. 6–17, 2001.

[123] A. Absalom, N. Sutcliffe, and G. Kenny, "Closed-loop control of anesthesia using bispectral index," *Anesthesiology*, vol. 96, pp. 67–73, 2002.

[124] C. Frei, A. Gentilini, M. Derighetti, A. Glattfelder, M. Morari, T. Schnider, and A. Zbinden, "Automation in anesthesia," in *Proceedings of the IEEE American Control Conference*, vol. 2, pp. 1258–1263, 1999.

[125] A. Gentilini, M. Rossoni-Gerosa, C. Frei, R. Wymann, M. Morari, A. Zbinden, and T. Schnider, "Modeling and closed-loop control of hypnosis by means of bispectral index (BIS) with isoflurane," tech. rep., ETH, Zürich, Switzerland, June 2000.

[126] D. Linkens, M. Mahfouf, and M. Abbod, "Self-adaptive and self-organising control applied to nonlinear multivariable anaesthesia: a comparative model-based study," in *Proceedings of IEE Conference on Control Theory and Applications*, vol. 139, pp. 381–394, July 1992.

[127] D. Linkens, "Adaptive and intelligent control in anesthesia," *IEEE Control Systems Magazine*, pp. 6–11, Dec. 1992.

[128] D. Linkens, *Introduction to intelligent control paradigms and system modeling*, ch. 1, pp. 1–18. Bristol, PA: Taylor & Francis, 1994.

[129] M. Mahfouf and M. Abbod, *A Comparative Study of Generalized Predictive Control GPC and Intelligent Self-Organizing Fuzzy Logic Control (SOFLC) for Multivariable Anaesthesia*, ch. 4, pp. 79–132. Bristol, PA: Taylor & Francis, 1994.

[130] X.-S. Zhang, R. Roy, and J. Haung, "Closed-loop system for intravenous anesthesia by simultaneously administering two anesthetic drugs," in *Proceedings of IEEE Annual Conference of the Medecine and Biology Society*, vol. 20, pp. 3052–3055, 1998.

[131] H. Schwilden and H. Stoeckel, "Closed-loop feedback controlled administration of alfentanil during alfentanil-nitrous oxide anaesthesia," *British Journal of Anaesthesia*, vol. 70, no. 4, pp. 389–393, 1993.

[132] H. Schwilden and J. Schüttler, *Model-based Adaptive Control of Volatile Anaesthetics by Quantitative EEG*, pp. 163–174. Berlin: Springer-Verlag, 1995.

[133] G. Kenny, H. Mantzaridis, and A. Fisher, *Validation of monitoring anesthetic depth by closed-loop control*, pp. 255–264. Englewood Cliffs, NJ: Prentice Hall, 1993.

[134] G. Kenny and D. Ray, *Adaptive Control of Intravenous Anaesthesia by Evoked Potentials*, pp. 208–219. Berlin: Springer-Verlag, 1995.

[135] G. Kenny and H. Mantzaridis, "Closed-loop control of propofol anaesthesia," *British Journal of Anaesthesia*, vol. 83, pp. 223–228, 1999.

[136] A. Nayak and R. Roy, "Anesthesia control using midlatency auditory evoked potentials," *IEEE Transactions on Biomedical Engineering*, vol. 45, pp. 409–421, Apr. 1998.

[137] A. Gentilini, M. Rossoni-Gerosa, C. Frei, R. Wymann, M. Morari, A. Zbinden, and T. Schnider, "Modeling and closed-loop control of hypnosis by means of bispectral index (BIS) with isoflurane," *IEEE Transactions on Biomedical Engineering*, vol. 48, no. 8, pp. 874–889, 2001.

[138] A. Gentilini, C. Frei, A. Glattfedler, M. Morari, T. Sieber, R. Wymann, T. Schnider, and A. Zbinden, "Multi-tasked closed-loop control in anesthesia," *IEEE Engineering in Medicine and Biology Society Magazine*, vol. 20, no. 1, pp. 39–53, 2001.

[139] K. Leslie, A. Absalom, and G. Kenny, "Closed-loop control of sedation for colonoscopy using the bispectral index," *Anaesthesia*, vol. 57, pp. 690–709, 2002.

[140] A. Absalom and G. Kenny, "Closed-loop control of propofol anaesthesia using bispectral index^{tm}: performance assessment in patients receiving computer controlled propofol and munually controlled remifentanil infusions for minor surgery," *British Journal of Anaesthesia*, vol. 90, no. 6, pp. 738–741, 2003.

[141] S. Locher, K. Stadler, T. Boehlen, T. Bouillon, D. Leibundgut, P. Schumacher, R. Wymann, and A. Zbinden, "A new closed-loop control system for isoflurane using bispectral index outperforms manual control," *Anesthesiology*, vol. 101, pp. 591–602, 2004.

[142] N. Liu, A. Genty, A. Landais, T. Chazot, and M. Fischler, "Titration of propofol guided by the bispectral index (BIS): Closed-loop versus manual target controlled infusion," in *Proceedings of the 2005 American Anesthesiology Society Conference*, October 22^{nd}-26^{th} 2005.

[143] P. Glass and I. Rampil, "Automated anesthesia: Fact or fantasy?," *Anesthesiology*, vol. 95, pp. 1–2, 2001.

[144] H. Yppärilä-Wolters, M. van Gils, I. Korhonen, M. Huiku, H. Viertiö-Oja, P. Meriläinen, M. Kymäläinen, P. Takala, M. Rantanen, and A. Yli-Hankala, "A modular approach for estimating nociception during general anaesthesia," in *Proceedings of the 2005 Advanced Modeling and Control in Anesthesia*, 2005.

[145] M. Huiku, M. van Gils, I. Korhonen, M. Kymäläinen, P. Meriläinen, M. Paloheimo, M. Rantanen, P. Takala, K. Uutela, H. Viertiö-Oja, and A. Yli-Hankala, "Index of nociception: a complementary parameter for the control of anaesthesia," in *Proceedings of the 2005 Advanced Modeling and Control in Anesthesia*, 2005.

[146] V. P. V., J. Gaspoz, S. Molliex, A. Antoniadis, T. Busso, F. Roche, F. Costes, L. Q. L, J. Lacour, and J. Barthelemy, "Wavelet transform to quantify heart rate variability and to assess its instantaneous changes," *Journal of Applied Physiology*, vol. 86, pp. 1081–1091, 1999.

[147] I. Korhonen, E. Seitsonen, M. van Gils, K. Korttila, and A. Yli-Hankala, "Heart rate variability responses to skin incision in 0.8 MAC sevoflurane anaesthesia," *International Journal of Bioelectromagnetism*, vol. 5, pp. 240–241, 2003.

[148] M. Sesay, J. Vignes, M. Stckle, G. Boulard, and P. Maurette, "Spectral analysis of the ECG R-R interval permits early detection of vagal responses to neurosurgical stimuli," *Annales Françaises d'Anesthésie et de Ré-animation*, vol. 86, pp. 1081–1091, 1999.

[149] C. Burrus, R. Gopinath, and H. Guo, *Introduction to wavelets and wavelet transform.* Upper Saddle River, NJ: Prentice-Hall, 1998.

[150] G. Strang and T. Nguyen, *Wavelets and filter banks.* Wellesley-Cambridge Press, 1996.

[151] T. Zikov, "Monitoring the anesthetic-induced unconsciousness (hypnosis) using wavelet analysis of the electroencephalogram," Master's thesis, The University of British Columbia, 2002.

[152] E. John, "A field theory of consciousness," *Conscious Cognition*, vol. 10, no. 2, pp. 184–213, 2001.

[153] G. Nason and B. Silverman, "The stationary wavelet transform and some statistical applications," *Lecture Notes in Statistics*, vol. 103, pp. 281–29, 1995.

[154] R. Coifman and D. Donoho, "Translation invariant de-noising," *Lecture Notes in Statistics*, vol. 103, pp. 125–150, 1995.

[155] D. Scott, "On optimal and data-based histograms," *Biometrika*, vol. 66, pp. 605–610, 1979.

[156] G. Schmidt, P. Bischoff, T. Standl, M. Issleib, M. Voigt, and J. Schulte Am Esch, "ARX-derived auditory evoked potential index and bispectral index during the induction of anesthesia with propofol and remifentanil," *Anesthesia Analgesia*, vol. 97, no. 1, pp. 139–144, 2003.

[157] G. Schmidt, P. Bischoff, T. Standl, K. Jensen, M. Voigt, and J. Schulte Am Esch, "Narcotrend and bispectral index monitor are superior to classic electroencephalographic parameters for the assessment of anesthetic states during propofol-remifentanil anesthesia," *Anesthesiology*, vol. 99, no. 5, pp. 1072–1077, 2003.

[158] R. Anderson, G. Barr, and J. Jakobsson, "Correlation between AAI-index and BIS-index during propofol hypnosis: a clinical study," *Journal of Clinical Monitoring and Computing*, vol. 17, no. 1, pp. 325–329, 2002.

[159] S. Hagihira, M. Takashina, T. Mori, T. Mashimo, and I. Yoshiya, "Practical issues in bispectral analysis of electroencephalographic signals," *Anesthesia Analgesia*, vol. 93, pp. 966–970, 2001.

[160] K. Lundqvist, S. Bibian, C. Ries, and P. Yu, "Can the wavelet-based anesthetic value (WAV) predict airway motor response to LMA insertion?," in *Proceedings of the American Society of Anesthesiologists*, pp. A–606, 2004.

[161] Task Force of the European Society for Cardiology and the North American Society for Pacing and Electrophysiology, "Heart rate variability: standards of measurement, physiological interpretation, and clinical use," *European Heart Journal*, vol. 17, pp. 354–381, 1996.

[162] D. Bloomfield, J. B. Jr, B. Pavri, J. Han, J. Fleiss, L. Rolnitzky, and R. Steinman, "Vagal modulation of RR intervals during head-up tilt and the infusion of isoproterenol," *American Journal of Cardiology*, vol. 75, pp. 1145–1150, 1995.

[163] D. Thys and J. Narang, *Cardiac Anesthesia*, ch. Electrocardiographic monitoring, pp. 175–193. Lippincott Williams & Wilkins, 2001.

[164] S. Bibian, T. Zikov, G. Dumont, C. Ries, E. Puil, H. Ahmadi, M. Huzmezan, and B. MacLeod, "Method and apparatus for the estimation of the anesthetic depth using wavelet analysis of the electroencephalogram," *Patent Pending, #60/395,313*, January 15^{th} 2004.

[165] C. Minto, T. Schnider, K. Gregg, T. Henthorn, and S. Shafer, "Using the time of maximum effect site concentration to combine pharmacokinetics and pharmacodynamics," *Anesthesiology*, vol. 99, no. 2, pp. 324–333, 2003.

[166] T. Schnider, C. Minto, S. Shafer, P. Gambus, C. Andresen, D. Goodale, and E. Youngs, "The influence of age on propofol pharmacodynamics," *Anesthesiology*, vol. 90, pp. 1502–1516, 1999.

[167] S. Skogestad and I. Postlethwaite, *Multivariable feedback control: Analysis and Design*. Wiley, 2001.

[168] D. Stanksi and P. Maitre, "Population pharmacokinetics and pharmacodynamics of thiopental: the effect of age revisited.," *Anesthesiology*, vol. 72, no. 3, pp. 412–422, 1990.

[169] I. Kissin, "Depth of anesthesia and bispectral index monitoring," *Anesthesia Analgesia*, vol. 90, pp. 1114–1117, May 2000.

[170] A. Cheung and B. Marshall, *The Inhaled Anesthetics*, ch. 8, pp. 75–87. Philadelphia, Pennsylvania: W.B. Saunders Company, ninth ed., 1997.

[171] I. Cockshott, "Propofol (Diprivan) pharmacokinetics and metabolism - an overview," *Postgraduate Medical Journal*, vol. 61, no. Suppl. 3, pp. 45–50, 1985.

[172] I. Cockshott, L. Briggs, E. Douglas, and M. White, "Pharmacokinetics of propofol in female patients," *British Journal of Anaesthesia*, vol. 59, pp. 1103–1110, 1987.

[173] T. Kirkpatrick, I. Cockshott, E. Douglas, and W. Nimmo, "Pharmacokinetics of Propofol (Diprivan) in elderly patients," *British Journal of Anaesthesia*, vol. 60, pp. 146–150, 1988.

[174] R. Tackley, G. Lewis, C. Prys-Roberts, R. Boaden, J. Dixon, and J. Harvey, "Computer controlled infusion of propofol," *British Journal of Anaesthesia*, vol. 62, pp. 46–53, 1989.

[175] B. Marsh, W. White, N. Morton, and G. Kenny, "Pharmacokinetic model driven infusion of propofol in children," *British Journal of Anaesthesia*, vol. 67, pp. 41–48, 1991.

[176] B. Kataria, S. Ved, H. Nicodemus, G. Hoy, D. Lea, M. Dubois, J. Mandema, and S. Shafer, "The pharmacokinetics of propofol in children using three different data analysis approaches," *Anesthesiology*, vol. 80, pp. 104–122, 1994.

[177] T. Short, C. Aun, P. Tan, J. Wong, Y. Tam, and T. Oh, "A prospective evaluation of pharmacokinetic model controlled infusion of propofol in paediatric patients," *British Journal of Anaesthesia*, vol. 72, pp. 302–306, 1994.

[178] J. Schüttler, H. Schwilden, and H. Stoeckel *Personal communication to I.D. Cockshott*, 1985.

[179] E. Gepts and F. Camu *Personal communication to I.D. Cockshott*, 1985.

[180] N. Kay, J. Uppington, J. Sears, E. Douglas, and I. Cockshott *Personal communication to I.D. Cockshott*, 1985.

[181] E. Gepts, F. Camu, I. Cockshott, and E. Douglas, "Disposition of propofol administered as constant rate intravenous infusion in humans," *Anesthesia Analgesia*, vol. 66, pp. 1256–1263, 1987.

[182] J. Vuyk, T. Schnider, and F. Engbers, "Population pharmacokinetics of propofol for target-controlled infusion (TCI) in thje elderly," *Anesthesiology*, vol. 93, no. 6, 2000.

[183] G. Ludbrook, E. Visco, and A. Lam, "Relation between brain concentrations, electroencephalogram, middle cerebral artery blood flow velocity, and cerebral oxygen extraction during induction of anesthesia," *Anesthesiology*, vol. 97, no. 6, pp. 1363–1370, 2002.

[184] A. Kapila, P. Glass, J. Jacobs, K. Muir, D. Hermann, M. Shiraishi, S. Howell, and R. Smith, "Measured context-sensitive half-times of remifentanil and alfentanil," *Anesthesiology*, vol. 83, pp. 968–975, 1995.

[185] P. Glass, D. Hardman, Y. Kamiyama, T. Quill, G. Marton, K. Donn, C. Grosse, and D. Hermann, "Preliminary pharmacokinetics and pharmacodynamics of an ultra-short-acting opioid: remifentanil (GI87084B)," *Anesthesia Analgesia*, vol. 77, pp. 1031–1040, 1993.

[186] C. Westmoreland, J. Hoke, P. Sebel, C. Hug, and K. Muir, "Pharmacokinetics of remifentanil (GI87084B) and its major metabolite (GI90291) in patients undergoing elective inpatient surgery," *Anesthesiology*, vol. 79, pp. 893–903, 1993.

[187] P. Glass, T. Gan, and S. Howell, "A review of pharmacokinetics and pharmacodynamics of remifentanil," *Anesthesia Analgesia*, vol. 89, pp. Suppl. 7–14, 1999.

[188] M. Doi, R. Gajraj, and H. Mantzaridis, "Relationship between calculated blood concentration of propofol and electrophysiological variables during emergence from anaesthesia: comparison of bispectral index, spectral edge frequency, median frequency and auditory evoked potential index," *British Journal of Anaesthesia*, vol. 78, pp. 180–184, 1997.

Appendix A

Pharmacopoeia

A.1 Vapour Anesthetics

Sevoflurane This drug was initially evaluated in the 1970s, but it's only in the 1980s that its low blood solubility and lack of pungency were recognized and appreciated. The fact that this agent provides the least degree of airway irritation, makes it acceptable for the induction of anesthesia. However, it is more vulnerable to metabolism. The compound resulting from the degradation of the drug is toxic, but present only in very small quantity. Due to its slight toxicity, sevoflurane has been approved for use in the U.S. only in 1995 [170].

Isoflurane Introduced in 1981 for patient use, this drug rapidly became the most commonly administered inhalational anesthetics [2]. The advantage of isoflurane is that it is less susceptible to metabolism, hence less toxic. However, this drug is pungent and slightly irritable for the airways.

Desflurane This drug is characterized by a very low solubility, which provokes a rapid onset followed by a fast recovery (particularly useful for day-surgeries). This drug is more pungent than sevoflurane and requires special pressurized vaporizers for its administration.

A.2 Intravenous Anesthetics

Barbiturates Barbiturates were the first chemical compounds developed for clinical anesthesia (*e.g.* thiopental, hexobarbitone). Derived from barbituric acid, they have a high lipid solubility which contributes to a rapid onset and shorter duration of action. Barbiturates such as thiopental are thus used for rapid intravenous induction of anesthesia. However, their particularly long context-sensitive half time and prolonged recovery period limit their usefulness. As such, they are seldom used in infusion.

Benzodiazepines Benzodiazepines such as *midazolam*, *diazepam* or *lorazepam* were introduced in the 1970s. They are known to be powerful sedatives and amnesics. However, the long acting characteristic of diazepam and lorazepam limit their use. Midazolam is metabolized by the liver and hence is shorter acting. However the time to peak effect is rather long, also limiting its use. These drugs are thus mainly used as premedicant.

Phencyclidine Derivatives of phencyclidine such as *ketamine* differ from all other anesthetics in that their mechanism of action within the brain is different. They produce a dissociative effect resembling that of a cataleptic state where the patient is noncommunicative. Amnesia is present, as well as intense analgesia. Introduced in 1965 in the U.S.A., ketamine is still widely used as an analgesic. However, the drug tends to also increase heart rate, blood pressure and muscle tone.

Carboxylated imidazole *Etomidate* produces a rapid onset of unconsciousness. Developed in the early 1970s, it was shown that it does not depress the circulation, thus being safer than barbiturates. However, it is also characterized by its excitatory effect on the brain. As a result, the patient may exhibit spontaneous movement in up to 50% of cases.

Isopropylphenol *Propofol* was introduced in 1977. Since then, it has become the intravenous drug of choice in the anesthesia practice. One particular characteristic of propofol is its fast redistribution and metabolism. As a result it can be easily used in infusion schemes, as it provides very fast emergence (no cumulative effect). Another interesting property is that propofol suppresses laryngeal reflexes, which makes it particularly attractive for induction.

A.3 Opioids

Morphine Morphine is the compound against which all other opioids are compared. Beside its analgesic effect, morphine also causes euphoria, sedation, and a decreased ability to concentrate. While the onset of effect occurs soon after intravenous administration, its peak effect is delayed by 15 to 30 min. This difference between onset and peak effect is mainly the result of the slow equilibration between the blood and brain concentration (so-called *effect-site equilibration time*).

Another distinguishing feature of morphine is its tendency of releasing histamine, thus provoking vasodilation, hypotension, tachycardia and cutaneous flushing.

Fentanyl Fentanyl is a synthetic opioid that was first synthesized by Janssen in 1962. Much more potent than morphine, fentanyl has also a more rapid onset and shorter duration of action. However, there is still a marked delay of 6 to 7 min between the peak plasma concentration and the peak effect. The short duration of action is the result of the rapid redistribution of the drug in inactive tissues. However, a saturation of these tissues can occur, mostly when using multiple boluses or long term infusions. As a result, the plasma

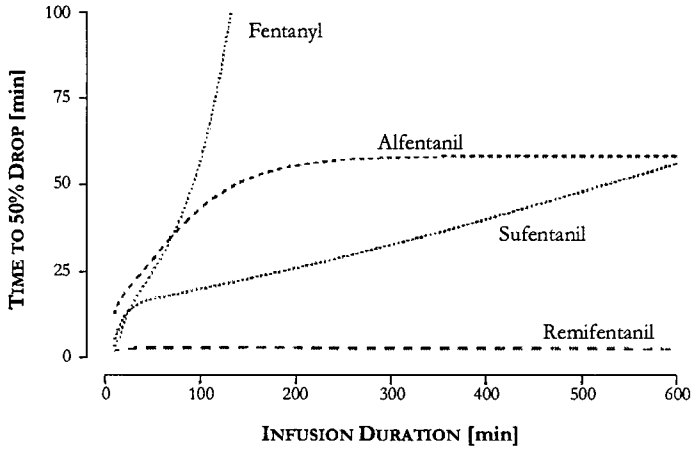


Figure A.1: Context sensitive half times of fentanyl, sufentanil, alfentanil and remifentanil (from [21]).

concentration of fentanyl reduces only slowly, which prolongs the duration of effect. This characteristic is well captured by the context-sensitive half-time curve of fentanyl, see Fig. A.1. As a result, a potential complication is the lingering effect of fentanyl, which might provoke a recurrent depression of ventilation in the post operative care unit.

As compared to morphine, the administration of synthetic opioids do not provoke histamine release. Conversely, the concurrent administration of other intravenous of inhalational anesthetics can seriously alter hemodynamic stability as they greatly potentiate the effects of these drugs.

Alfentanil Alfentanil is an analogue compound to fentanyl, with the differences that it is less potent and has a reduced duration of action. Its major advantage is its fast onset of action after intravenous administration. It takes about 1.4 min for the concentration at the effect site to equilibrate with the plasma concentration. This is particularly useful to rapidly blunt strong and brief noxious stimuli such as tracheal intubation. The elimination of alfentanil is done through hepatic metabolism. It has been shown that there can be a 10-times inter-patient variability in the systemic clearance of the drug due to large differences in enzyme activity between individuals.

Sufentanil Synthesized in 1979, sufentanil is also an analogue to fentanyl, but with a 7 times higher potency. As compared to fentanyl, the context-sensitive half-time is much lower and more constant for longer duration of infusion, see Fig. A.1. Depression of ventilation may also be more profound than with fentanyl. When used in large doses, this drug may also produces muscle rigidity which makes ventilation of the lungs more difficult. Because of its context-sensitive half-time characteristic, sufentanil is used almost uniquely for long surgeries.

Remifentanil Remifentanil is a new agent introduced in the practice in the mid 1990s. Its potency is twice that of fentanyl and its effect-site equilibration time is slightly less to that of alfentanil (about

1.1 min). A major difference with other opioids lies in its molecular structure. Its ester linkage renders this compound susceptible to hydrolysis by non-specific blood and tissue esterases. This property results in its rapid degradation in essentially inactive metabolites [21]. The main characteristics of remifentanil are: brevity of action, rapid onset, noncumulative effects in inactive tissues, and rapid recovery after termination of the infusion. As a result, its context-sensitive half-time is independent of the duration of the infusion, see Fig. A.1. Also, the risk for post-operative rebound of effect, while common with other opioids, is significantly reduced.

Remifentanil is used mostly to supplement the analgesic component of general anesthesia. It gives to anesthesiologists the ability to help their patients to recover rapidly from undesirable opioid-induced side-effects such as the depression of ventilation. Because of its brevity of action, it is usually recommended that longer acting opioids are administered at the end of the surgery to reduce post-operative pain.

Appendix B

Propofol and Remifentanil PKPD

This Annex presents a thorough review of the pharmacokinetic and pharmacodynamic studies carried out for propofol and remifentanil for the past 20 years. This review is presented under the form of tables containing the PK and PD parameters. A short description of each study complements the tables.

B.1 Propofol

B.1.1 Pharmacokinetics

Cockshott and Gepts, 1985 In 1985, a preliminary survey by Cockshott [171] reported a set of parameters obtained from three independent studies ([178], [179] and [180]). While these studies were based on a very limited number of healthy adult patients (ASA I or II) undergoing elective surgery, their results are fairly consistent as far as half-lives are concerned.

Cockshott, 1987 Following his preliminary review of propofol pharmacokinetics, Cosckshott published another set of parameters [172] obtained on 6 female patients and derived from single bolus data. His results were also consistent to those previously reported, with the exception of the central compartment volume which is larger. However, this discrepancy might be explained by the poor number of patients. For that particular parameter, Cockshott also reported a wide variability between patients.

Kirkpatrick* et al., *1988 In 1988, a study by Kirkpatrick *et al.* [173] comparing the propofol kinetics between two different age groups (12 young adults and 12 elderly patients) revealed that age is a strong covariate and must be accounted for in pharmacokinetic parameters. As compared to the younger age group, older patients experience a reduced metabolism resulting in a slower elimination phase. This observation indicates that infusion rates should be reduced to account for the increased elimination half-life.

Gepts* et al., *1988 Gepts *et al.* initiated this study to evaluate the effect of opioid administration on propofol kinetics. They first set a constant alfentanil concentration and followed the evolution of the plasma

STUDY	V_1 [l] [l/kg]*	POLES π [s^{-1}]	α [s^{-1}]	β [s^{-1}]	ZEROS k_{21} [s^{-1}]	k_{31} [s^{-1}]
Cockshott *et al.*,1985 [171]						
Schüttler *et al.*	28.9	$5.5e^{-3}$	$3.0e^{-4}$	$3.0e^{-5}$	$3.3e^{-3}$	$5.8e^{-5}$
Kay *et al.*	39.2	$5.0e^{-3}$	$2.3e^{-4}$	$4.0e^{-5}$	$3.3e^{-3}$	$5.8e^{-5}$
Cockshott *et al.*,1987 [172]	41.3	$4.0e^{-3}$	$2.6e^{-4}$	$4.1e^{-5}$	$3.3e^{-3}$	$5.8e^{-5}$
Kirkpatrick *et al.*, 1988 [173]						
Young	0.42*	$6.7e^{-3}$	$2.9e^{-4}$	$2.0e^{-5}$	$1.0e^{-3}$	$3.2e^{-5}$
Elderly	0.31*	$8.3e^{-3}$	$2.3e^{-4}$	$2.1e^{-5}$	$0.6e^{-3}$	$3.2e^{-5}$
Gepts *et al.*, 1988 [111]	16.9	$5.0e^{-3}$	$5.0e^{-4}$	$4.0e^{-5}$	$0.9e^{-3}$	$5.5e^{-5}$
Tackley *et al.*, 1989 [174]	0.32*	$4.1e^{-3}$	$4.6e^{-4}$	$4.4e^{-5}$	$1.1e^{-3}$	$5.7e^{-5}$
Gill *et al.*, 1990 [110]	23.7	$3.7e^{-3}$	$2.7e^{-4}$	$2.2e^{-5}$	$3.3e^{-3}$	$5.8e^{-5}$
Marsh *et al.*, 1991 [175]						
Adult	0.23*	$5.1e^{-3}$	$4.1e^{-4}$	$3.7e^{-5}$	$7.0e^{-3}$	$5.5e^{-5}$
Children	0.34*	$3.7e^{-3}$	$3.1e^{-4}$	$4.5e^{-5}$	$0.6e^{-3}$	$5.5e^{-5}$
Kataria *et al.*, 1994 [176]						
Children						
NONMEM method	0.44*	$5.8e^{-3}$	$6.7e^{-4}$	$4.4e^{-5}$	$1.5e^{-3}$	$7.8e^{-5}$
2-stage method	0.32*	$5.9e^{-3}$	$4.3e^{-4}$	$3.5e^{-5}$	$0.8e^{-3}$	$5.5e^{-5}$
Pooled method	0.52*	$4.4e^{-3}$	$4.5e^{-4}$	$2.9e^{-5}$	$1.0e^{-3}$	$5.3e^{-5}$
Short *et al.*, 1994 [177]						
Chinese children	0.43*	$5.7e^{-3}$	$7.4e^{-4}$	$5.7e^{-5}$	$1.8e^{-3}$	$8.2e^{-5}$
Schnider *et al.*, 1998 [100]						
Adult, 30 yr, 50 kg, 1.70 m	21.5	$16.0e^{-3}$	$14.0e^{-4}$	$4.9e^{-5}$	$3.3e^{-3}$	$5.8e^{-5}$
Adult, 30 yr, 70 kg, 1.70 m	30.1	$16.5e^{-3}$	$15.0e^{-4}$	$4.9e^{-5}$	$3.3e^{-3}$	$5.8e^{-5}$
Adult, 30 yr, 110 kg, 1.70 m	47.3	$20.0e^{-3}$	$19.0e^{-4}$	$5.3e^{-5}$	$3.3e^{-3}$	$5.8e^{-5}$
Elderly, 70 yr, 70 kg, 1.70 m	30.1	$12.3e^{-3}$	$20.0e^{-4}$	$4.9e^{-5}$	$3.3e^{-3}$	$5.8e^{-5}$
Schüttler *et al.*, 2000 [101]						
Adult, 30 yr, 50 kg	7.3	$8.9e^{-3}$	$4.2e^{-4}$	$2.8e^{-5}$	$0.8e^{-3}$	$4.8e^{-5}$
Adult, 30 yr, 70 kg	9.3	$8.7e^{-3}$	$4.3e^{-4}$	$3.4e^{-5}$	$0.8e^{-3}$	$5.8e^{-5}$
Adult, 30 yr, 110 kg	12.8	$8.5e^{-3}$	$4.4e^{-4}$	$4.5e^{-5}$	$0.8e^{-3}$	$7.4e^{-5}$

Table B.1: Propofol PK parameter sets (hybrid form).

concentration of propofol during an infusion of 6 mg/kg/hour. They obtained a first set of PK parameters which they further compared to parameters obtained without alfentanil. They concluded that propofol kinetics were not affected by the co-administration of alfentanil.

Tackley* et al., *1989 In 1989, Tackley *et al.* [174] used a TCI device to reach a desired propofol plasma concentration using the 3-compartment parameters from Cockshott. After using the TCI device on 8 patients, they proposed an optimization of Cockshott's parameters in order to reach the reference sooner. While their modified parameters indeed improved the settling time of the plasma concentration, it also resulted in a strong steady state error of about 20%.

Gill* et al., *1990 A 9-patient single-bolus injection study by Gill *et al.* [110] resulted in a rather long distribution half-life (3.1 min). However, they also obtained a large variability ($\pm$2.03 min) on this parameter. This result could be due to the method of administration of the drug and the difficulty of properly identifying the decaying rate of the distribution phase as only 5 blood samples were taken during this phase.

Marsh* et al., *1991 The first study involving children was published by Marsh *et al.* [175] in 1991 using a TCI device. It was found that pediatric patients usually have larger infusion rate requirements than adults, as their central compartment volume is much larger. Marsh and co-workers also adapted the adult pharmacokinetic parameters from Gepts [181] according to their own observations.

Kataria* et al., *1994 Later on, an extensive and well-documented study by Kataria *et al.* [176] on 53 unpremedicated children proposed three sets of parameters obtained by three different statistical approaches. The authors recommend the use of the NONMEM software (NONMEM Project, University of California, San Francisco, CA) for the analysis of pharmacokinetic data. This software allows multiple nonlinear regression analysis of population data simultaneously. In this study, they observed that, among children, weight is a strong covariate that account for one fourth of the inter-individual variability. While accounting for age as well improves the model, the actual improvement of the fit is rather small.

Short* et al., *1994 A study involving Chinese children was carried out by Short *et al.* [177] who used as a starting point the parameters published in Marsh's study. They found out that the fit could be improved by fine tuning the parameters, resulting in a slightly different set of coefficients and a larger central compartment volume.

Schnider* et al., *1998 The propofol PK during infusion regimen was then investigated by Schnider *et al.* [100] in 1998. The aim of their study was to estimate the covariates of propofol PK, and in particular, estimate the effect of the method of administration (bolus *vs.* infusion) on the PK parameters. They analyzed the blood samples of 24 volunteers from 25 to 81 years old using the NONMEM method. They

published a set of equations that allows the calculation of the PK parameters as a function of the age, weight and lean body mass.

Schüttler and Ihmsen, 2000 Finally, by far the most extensive study in this survey - covering 270 individuals and 4,112 blood samples - was carried out by Schüttler and Ihmsen in 2000 [101]. Echoing the earlier work by Kataria, Schüttler and Ihmsen also used the NONMEM software, which allowed them to quantify the influence of covariates such as age, lean body mass and gender. In short, the results they obtained were close to those of Kataria and Shafer. As compared to the earlier studies of Cosckshott, Kirkpatrick and Gill, there is a large discrepancy in half-lives and compartmental volumes. This discrepancy can be attributed to the analysis method employed in recent studies which benefit from advances in statistical analysis. However, Schüttler and Ihmsen's results concerning the elderly age population group have been seriously criticized by Vuyk *et al.* [182], who argued that the proposed model under-predicts the initial plasma blood concentration, thus leading to higher infusion rates and possibly overdosing.

B.1.2 Pharmacodynamics

Note that this review is limited to only PD studies involving quantitative EEG.

Billard* et al., *1997 The first set was published by Billard *et al.* in 1997 [117]. Propofol PD models were derived for the bispectral index, the SEF and the delta power (0.5 to 4 Hz) for three major anesthetic drugs (alfentanil, propofol and midazolam). Among other, they found that the Hill equation could not cope satisfactorily with the biphasic effect of propofol observed mostly in SEF and the delta power measures. They also noticed that the biphasic effect of propofol was less present in the bispectral measure. They concluded that BIS can be used as a measure of the EEG effects of anesthetic drugs. A limitation of this study is that the authors are using direct blood plasma measurements for $C_p(t)$ and do not provide the corresponding pharmacokinetic set.

Schnider* et al., *1999 Two years later, Schnider *et al.* [166] completed their pharmacokinetic study with a pharmacodynamic analysis of the effect of propofol on a proprietary EEG index based on semilinear canonical correlation. A particularity of this study is that they identified the biphasic characteristic of propofol using a combination of two Hill equations (one for the activation phase and one for the depression phase). The principal aim of their study was the assessment of the effect of age on the PD parameters. They found that the EC_{50} (corresponding to the concentration required to provoke a loss of consciousness in 50% of their study population) decreases with age. This result indicates that older patients are more sensitive to propofol. However, note that the maximal effect observed in this study was the loss of consciousness and not a state of un-responsiveness suitable for surgery.

STUDY	PROPOFOL PK	POP.	k_{e0} [s^{-1}]	γ	EC_{50} [$\mu g \cdot ml^{-1}$]	E_0	E_{max}
Billard *et al.*, 1997 [117]	n.d.						
Bispectral Index		12	$3.3e^{-3}$	5.3±2.1	3.40	0 %	100%
Spectral Edge Freq.		12	$3.5e^{-3}$	5.3±2.1	4.50	0 %	100%
Delta Power		12	$4.5e^{-3}$	5.3±2.1	4.60	0 %	100%
Schnider *et al.*, 1999 [166]	Schnider *et al.* [100]						
Semilinear Canonical Correlation		24	$4.3e^{-3}$	n.a.	2.1-0.014·Age	n.a.	n.a.
Kazama *et al.*, 1999 [116]	Getps *et al.* [181]						
Bispectral Index							
Group 1 (20-39 yr)		11	$(5.0 \pm 0.1)e^{-3}$	2.39	5.60±0.48	0	100
Group 2 (40-59 yr)		12	$(5.0 \pm 0.1)e^{-3}$	2.08	6.76±0.51	0	100
Group 3 (60-69 yr)		12	$(5.0 \pm 0.2)e^{-3}$	2.03	8.21±0.64	0	100
Group 4 (70-85 yr)		12	$(5.0 \pm 0.2)e^{-3}$	2.10	7.67±0.58	0	100
White *et al.*, 1999 [115]	*Authors derived their own 2-compartment PK model*						
Auditory Evoked Potentials		15	$3.7e^{-3}$	2.17	1.76	0%	100%
Kuizenga *et al.*, 2001 [118]	*Authors performed their own blood plasma concentration measurements - no PK model was published*						
Bispectral Index							
Initial infusion		8	$(3.4 \pm 1.9)e^{-3}$	4.20±3.0	2.44±1.37	n.d.	n.d.
Sequential infusions		8	$(10.5 \pm 2.2)e^{-3}$	5.67±4.6	3.44±2.46	n.d.	n.d.
Ludbrook *et al.*, 2002 [183]	*Authors derived their own 2-compartment PK model*						
Bispectral Index		7	$1.8e^{-3}$	n.d.	n.d.	n.d.	n.d.

Table B.2: Propofol effect site half-life and pharmacodynamic parametric constants.

Kazama* et al., *1999 Echoing Schnider's study, Kazama *et al.* [116] also attempted to determine whether age is a covariate of the effect half-life $T_{\frac{1}{2}}^{k_{e0}}$. To do so, they analyzed the effect of propofol infusion on the bispectral index and blood pressure of 45 patients divided into 4 age groups and undergoing elective surgery. As far as BIS is concerned, they concluded that age is not a covariate of $T_{\frac{1}{2}}^{k_{e0}}$. However, they found that the effective plasma concentration EC_{50} tends to increase with age, indicating that older patients may require a higher plasma concentration of propofol to achieve the same effect than a lower concentration in younger patients. This result goes directly against Schnider's. However, the authors refer to a study of propofol effect in rats [119] which reported that younger animals are more sensitive to propofol than old animals due to a higher propofol concentration in the the brain.

White* et al., *1999 In this study, White et al. are using Auditory Evoked Potentials (AEP) to quantify the effect of propofol onto the brain. The technique used for the derivation of the AEP is similar to that used by the A-Line Monitor (Alaris Medical Systems). The authors recruited 22 patients. However they could only successfully collapse the hysteresis loop and get a meaningful $T_{\frac{1}{2}}^{k_{e0}}$ in 15 of these. The EC_{50} reported in this study is 2.17 $\mu g \cdot ml^{-1}$ which is similar to the value reported by Schnider et al. [166]. The authors could not explain satisfactorily this rather low value. They concluded that the AEP index might be more sensitive to propofol than other electroencephalographic indexes.

Kuizenga* et al., *2001 A second set of parameters describing also the effect of propofol on the bispectral index has been proposed by Kuizenga *et al.* [118]. Their study was designed to assess the predictability of PD models. Their results are further discussed in Section 3.2.4.2.

Ludbrook* et al., *2002 The most recent propofol PD study was conducted by Ludbrook *et al.* [183] on 7 patients. Two arterial lines allowed the authors to estimate the brain propofol concentration by calculating the amount of propofol retained by the brain. They reported a much slower equilibrium between the blood and the brain as compared to other studies. They attributed this difference mainly to the method of administration (as compared to other studies, they administered propofol at a lower infusion rate, which may have limited the depression of the circulatory system, and thus, the drop in cardiac output).

During the study, propofol was infused at a rate of 110 mg/min for 5 minutes, and then 10 mg/min for another 20 minutes. This regimen created a large plasma concentration (>20 $\mu g \cdot ml^{-1}$) which in turns produced a marked electroencephalographic depression (significant burst suppression and periods of iso-electric EEG). The authors reported that burst suppression started in some patients at a concentration of 5.5 $\mu g \cdot ml^{-1}$, while most patients experienced a total EEG depression at concentrations above 15 $\mu g \cdot ml^{-1}$.

STUDY	V_1 [l] [l/kg]*	POLES π [s^{-1}]	α [s^{-1}]	β [s^{-1}]	ZEROS k_{21} [s^{-1}]	k_{31} [s^{-1}]
Egan *et al.*, 1993 [21]						
NONMEM method	7.1	$12.8e^{-3}$	$1.3e^{-3}$	$24.0e^{-5}$	$2.4e^{-3}$	$2.6e^{-4}$
2-stage method	7.0	$14.4e^{-3}$	$1.4e^{-3}$	$0.7e^{-4}$	$2.9e^{-3}$	$7.3e^{-4}$
Pooled method	7.7	$12.8e^{-3}$	$1.4e^{-3}$	$2.2e^{-4}$	$2.6e^{-3}$	$2.3e^{-4}$
Kapila *et al.*, 1995 [184]	0.12*	$13.1e^{-3}$	$1.5e^{-3}$	n.a.	$3.2e^{-3}$	n.a.
Egan *et al.*, 1996 [102]	7.1	$12.8e^{-3}$	$1.3e^{-3}$	$24.0e^{-5}$	$2.4e^{-3}$	$2.6e^{-4}$
Minto et al., 1997 [103]						
Adult, 30 yr, 50 kg, 1.70 m	4.5	$20.6e^{-3}$	$2.0e^{-3}$	$2.6e^{-4}$	$4.1e^{-3}$	$2.7e^{-4}$
Adult, 30 yr, 70 kg, 1.70 m	5.3	$18.2e^{-3}$	$1.8e^{-3}$	$2.6e^{-4}$	$3.7e^{-3}$	$2.7e^{-4}$
Adult, 30 yr, 110 kg, 1.70 m	6.2	$16.3e^{-3}$	$1.7e^{-3}$	$2.6e^{-4}$	$3.3e^{-3}$	$2.7e^{-4}$
Elderly, 80 yr, 70 kg, 1.70 m	4.1	$12.0e^{-3}$	$1.4e^{-3}$	$0.9e^{-4}$	$2.2e^{-3}$	$0.9e^{-4}$

Table B.3: Remifentanil PK parameter sets (hybrid form).

B.2 Remifentanil

B.2.1 Pharmacokinetics

Glass* et al., *1993 An initial study was published in 1993 by Glass *et al.* [185] who revealed the ultra-short acting property of remifentanil. The pharmacokinetic parameters were derived based on arterial blood samples obtained from 30 male volunteers following the administration of increasing doses of remifentanil. In their study, a two-compartment model was deemed satisfactory. Unfortunately, the PK parameter set was not published in its entirety.

Egan* et al., *1993 Following this initial report, a study by Egan *et al.* [21] on 10 adult volunteers receiving an infusion of remifentanil over 20 min at various rates (1 to 8 $\mu g \cdot kg^{-1} \cdot min^{-1}$) confirmed the promising clinical potential of this new opioid in anesthesiology. The pharmacokinetic parameters were obtained using three different statistical approaches, however, no significant improvement could be observed when using the complex NONMEM analysis or standard nonlinear regression analysis (*i.e.* the pooled method). Actually, the nonlinear regression analysis technique yielded a more accurate result for long term predictions.

Westmoreland* et al., *1993 Significant differences in pharmacokinetic parameters were obtained in the study by Westmoreland *et al.* [186]. Their NONMEM analysis was carried out on 24 patients divided into 4 groups, each group receiving a different dose of remifentanil. As compared to Egan's study, the resulting half-lives are fairly different. However, the central compartment volume is consistent with results obtained in other studies. What is remarkable, however, is that it seems to be a significant difference in the drug kinetics according to the administered dose. As the dose increases, it appears that the distribution and

elimination half-lives are increasing as well. The linearity of the pharmacokinetics of remifentanil should therefore be questioned. However, this issue was not discussed by the authors of this paper. The PK sets were incomplete and could not be translated into the hybrid form.

Kapila* et al., *1995 A study by Kapila *et al.* [184] in 1995 confirmed the earlier results from Egan *et al.* This study was carried out on 22 patients and involved a 3 hours infusion in order to measure the context-sensitive half-time of remifentanil. Their measures echoed the simulated results obtained by Egan.

Egan* et al., *1996 Before remifentanil, alfentanil was the opioid of choice in terms brevity of action and rapidity of onset. To assess the PK and PD difference between the two drugs, Egan *et al.* [102] recruited 10 volunteers to receive an infusion at various rates of the two drugs on two separate days. The NONMEM software was used to

Minto* et al., *1997 In 1997, Minto *et al.* [103] presented an extensive study to determine the covariates of remifentanil pharmacokinetics based on blood samples from 65 healthy adult patients receiving remifentanil under constant infusion rate (1 to 8 $\mu g \cdot kg^{-1} \cdot min^{-1}$). Their observations led them to conclude that age, weight and lean body mass are all factors affecting the pharmacokinetic parameters. Their analysis yielded a set of equations linking the compartmental volumes and clearances with the patient's age and lean body mass.

Glass* et al., *1999 Finally, a recent review by Glass *et al.* [187] comments on the differences in elimination half-life reported in the literature. He attributed these differences to the method of investigation used by the authors of these studies. In general, large elimination half-lives were observed for high infusion rates. For studies using smaller doses of remifentanil, the concentration of the drug in the blood might have been below the measurable threshold, resulting in a poor estimation of $T_{\frac{1}{2}}^{\beta}$.

B.2.2 Pharmacodynamics

Egan* et al., *1996 Egan et al. [102] used the spectral edge frequency to model the PD of remifentanil. They reported a large inter-individual variability, mostly in terms of the Hill parameters.

Minto* et al., *1997 Along with the pharmacokinetic study, Minto *et al.* [103] included the pharmacodynamic parameters corresponding to the effect of remifentanil on the spectral edge frequency. They also studied the influence of age on these parameters. Conversely to propofol, the effect half-life of remifentanil and the sensitivity to the drug increase with age.

STUDY	REMIFENTANIL PK	POP.	k_{e0} [s^{-1}]	γ	EC_{50} [$ng \cdot ml^{-1}$]	E_0	E_{max}
Egan *et al.*, 1996 [102]	Egan *et al.*, 1996 [102]						
Spectral Edge Freq.		10	$19.0e^{-3}$	4.3	19.9	19	13.8
Minto *et al.*, 1997 [103]	Minto *et al.*, 1997 [103]						
Spectral Edge Freq.		30					
20 yr			$12.3e^{-3}$	2.4	16.1	20	5.5
40 yr			$10.0e^{-3}$	2.4	13.1	20	5.5
60 yr			$7.6e^{-3}$	2.4	10.1	20	5.5
80 yr			$5.3e^{-3}$	2.4	7.2	20	5.5

Table B.4: Remifentanil effect site half-life and pharmacodynamic parametric constants.

Appendix C

Propofol PKPD Nominal and Uncertainty Models

This annex summarizes the results of the PKPD uncertainty work carried out in Chapter 7 for propofol. This work was based on the 44 PKPD models presented in Table 6.1 and derived from the induction data obtained during the LMA study (see Annex D.2).

Table C.1, C.2, and C.3 present the propofol PKPD nominal models (separated in both their PK and PD parts). Each model is based on a particular subset of the study population (age group), or based on a particular study condition (*e.g.*, the WAV_{CNS} range, infusion and/or bolus administration). The nominal model parameters were obtained by averaging the models over the study population corresponding to the conditions described in the first column of the tables (drug administration, age group and WAV_{CNS} range).

Table C.4, C.5, and C.6 present the weights $w_{nom}(s)$ (nominal model optimization weight) and $\hat{w}(s)$ (relative uncertainty weight). Both weights are 4^{th} order LTI transfer functions when the drug administration covers a rate of 0 up to 10 $mg{\cdot}min^{-1}{\cdot}kg^{-1}$. If the drug administration is limited to 0 to 0.5 $mg{\cdot}min^{-1}{\cdot}kg^{-1}$ (infusion) or 2 to 10 $mg{\cdot}min^{-1}{\cdot}kg^{-1}$ (bolus), the weights order can be reduced to 2 without significant differences in the overall results.

Model	Model Characteristics	Pharmacokinetic Model	Pharmacodynamic Model
#1	**Administration:** infusion + bolus **Age group:** 18-60 yrs old **WAV $_{CNS}$ range:** 80 - 20	$\frac{1}{15.90} \cdot \frac{(s+1.16\cdot10^{-3})(s+4.71\cdot10^{-5})}{(s+8.34\cdot10^{-3})(s+3.56\cdot10^{-4})(s+3.03\cdot10^{-5})}$	$e^{-14.8\cdot s} \cdot \frac{5.71}{(s+44.6\cdot10^{-3})}$
#2	**Administration:** infusion + bolus **Age group:** 18-29 yrs old **WAV $_{CNS}$ range:** 80 - 20	$\frac{1}{17.97} \cdot \frac{(s+1.16\cdot10^{-3})(s+4.52\cdot10^{-5})}{(s+7.15\cdot10^{-3})(s+3.49\cdot10^{-4})(s+2.89\cdot10^{-5})}$	$e^{-20.0\cdot s} \cdot \frac{10.63}{(s+61.3\cdot10^{-3})}$
#3	**Administration:** infusion + bolus **Age group:** 30-39 yrs old **WAV $_{CNS}$ range:** 80 - 20	$\frac{1}{15.78} \cdot \frac{(s+1.17\cdot10^{-3})(s+4.69\cdot10^{-5})}{(s+8.38\cdot10^{-3})(s+3.56\cdot10^{-4})(s+3.02\cdot10^{-5})}$	$e^{-12.9\cdot s} \cdot \frac{4.79}{(s+39.0\cdot10^{-3})}$
#4	**Administration:** infusion + bolus **Age group:** 40-49 yrs old **WAV $_{CNS}$ range:** 80 - 20	$\frac{1}{15.15} \cdot \frac{(s+1.17\cdot10^{-3})(s+4.88\cdot10^{-5})}{(s+9.09\cdot10^{-3})(s+3.61\cdot10^{-4})(s+3.15\cdot10^{-5})}$	$e^{-12.4\cdot s} \cdot \frac{3.39}{(s+34.2\cdot10^{-3})}$
#5	**Administration:** infusion + bolus **Age group:** 50-60 yrs old **WAV $_{CNS}$ range:** 80 - 20	$\frac{1}{14.12} \cdot \frac{(s+1.17\cdot10^{-3})(s+4.92\cdot10^{-5})}{(s+9.09\cdot10^{-3})(s+3.61\cdot10^{-4})(s+3.15\cdot10^{-5})}$	$e^{-10.5\cdot s} \cdot \frac{3.14}{(s+33.1\cdot10^{-3})}$
#6	**Administration:** infusion **Age group:** 18-60 yrs old **WAV $_{CNS}$ range:** 80 - 20	$\frac{1}{9.44} \cdot \frac{(s+0.85\cdot10^{-3})(s+6.20\cdot10^{-5})}{(s+9.16\cdot10^{-3})(s+4.29\cdot10^{-4})(s+3.69\cdot10^{-5})}$	$e^{-14.8\cdot s} \cdot \frac{5.71}{(s+44.6\cdot10^{-3})}$
#7	**Administration:** bolus **Age group:** 18-60 yrs old **WAV $_{CNS}$ range:** 80 - 20	$\frac{1}{24.71} \cdot \frac{(s+1.48\cdot10^{-3})(s+3.22\cdot10^{-5})}{(s+7.51\cdot10^{-3})(s+2.83\cdot10^{-4})(s+2.37\cdot10^{-5})}$	$e^{-14.8\cdot s} \cdot \frac{5.71}{(s+44.6\cdot10^{-3})}$

Table C.1: PKPD Propofol nominal models - part I

Model	Model Characteristics	Pharmacokinetic Model	Pharmacodynamic Model
#8	**Administration:** infusion + bolus **Age group:** 18-60 yrs old **WAV $_{CNS}$ range:** 80 - 50	$\frac{1}{15.90} \cdot \frac{(s+1.16\cdot10^{-3})(s+4.71\cdot10^{-5})}{(s+8.34\cdot10^{-3})(s+3.56\cdot10^{-4})(s+3.03\cdot10^{-5})}$	$e^{-14.8\cdot s} \cdot \frac{7.98}{(s+44.6\cdot10^{-3})}$
#9	**Administration:** infusion + bolus **Age group:** 18-60 yrs old **WAV $_{CNS}$ range:** 60 - 30	$\frac{1}{15.90} \cdot \frac{(s+1.16\cdot10^{-3})(s+4.71\cdot10^{-5})}{(s+8.34\cdot10^{-3})(s+3.56\cdot10^{-4})(s+3.03\cdot10^{-5})}$	$e^{-14.8\cdot s} \cdot \frac{5.98}{(s+44.6\cdot10^{-3})}$
#10	**Administration:** infusion + bolus **Age group:** 18-60 yrs old **WAV $_{CNS}$ range:** 40 - 20	$\frac{1}{15.90} \cdot \frac{(s+1.16\cdot10^{-3})(s+4.71\cdot10^{-5})}{(s+8.34\cdot10^{-3})(s+3.56\cdot10^{-4})(s+3.03\cdot10^{-5})}$	$e^{-14.8\cdot s} \cdot \frac{3.88}{(s+44.6\cdot10^{-3})}$
#11	**Administration:** infusion **Age group:** 18-29 yrs old **WAV $_{CNS}$ range:** 80 - 20	$\frac{1}{10.69} \cdot \frac{(s+0.85\cdot10^{-3})(s+5.94\cdot10^{-5})}{(s+7.78\cdot10^{-3})(s+4.24\cdot10^{-4})(s+3.52\cdot10^{-5})}$	$e^{-20.0\cdot s} \cdot \frac{10.63}{(s+61.3\cdot10^{-3})}$
#12	**Administration:** infusion **Age group:** 30-39 yrs old **WAV $_{CNS}$ range:** 80 - 20	$\frac{1}{9.38} \cdot \frac{(s+0.85\cdot10^{-3})(s+6.17\cdot10^{-5})}{(s+9.21\cdot10^{-3})(s+4.29\cdot10^{-4})(s+3.67\cdot10^{-5})}$	$e^{-12.9\cdot s} \cdot \frac{4.79}{(s+39.0\cdot10^{-3})}$
#13	**Administration:** infusion **Age group:** 40-49 yrs old **WAV $_{CNS}$ range:** 80 - 20	$\frac{1}{9.00} \cdot \frac{(s+0.85\cdot10^{-3})(s+6.42\cdot10^{-5})}{(s+10.0\cdot10^{-3})(s+4.33\cdot10^{-4})(s+3.84\cdot10^{-5})}$	$e^{-12.4\cdot s} \cdot \frac{3.39}{(s+34.2\cdot10^{-3})}$
#14	**Administration:** infusion **Age group:** 50-60 yrs old **WAV $_{CNS}$ range:** 80 - 20	$\frac{1}{8.39} \cdot \frac{(s+0.85\cdot10^{-3})(s+6.47\cdot10^{-5})}{(s+10.7\cdot10^{-3})(s+4.34\cdot10^{-4})(s+3.88\cdot10^{-5})}$	$e^{-10.5\cdot s} \cdot \frac{3.14}{(s+33.1\cdot10^{-3})}$

Table C.2: PKPD Propofol nominal models - part II

MODEL	MODEL CHARACTERISTICS	PHARMACOKINETIC MODEL	PHARMACODYNAMIC MODEL
#15	**Administration:** infusion **Age group:** 18-29 yrs old **WAV $_{CNS}$ range:** 80 - 50	$\frac{1}{10.69} \cdot \frac{(s+0.85\cdot10^{-3})(s+5.94\cdot10^{-5})}{(s+7.78\cdot10^{-3})(s+4.24\cdot10^{-4})(s+3.52\cdot10^{-5})}$	$e^{-20.0\cdot s} \cdot \frac{14.66}{(s+61.3\cdot10^{-3})}$
#16	**Administration:** infusion **Age group:** 30-39 yrs old **WAV $_{CNS}$ range:** 80 - 50	$\frac{1}{9.38} \cdot \frac{(s+0.85\cdot10^{-3})(s+6.17\cdot10^{-5})}{(s+9.21\cdot10^{-3})(s+4.29\cdot10^{-4})(s+3.67\cdot10^{-5})}$	$e^{-12.9\cdot s} \cdot \frac{6.77}{(s+39.0\cdot10^{-3})}$
#17	**Administration:** infusion **Age group:** 40-49 yrs old **WAV $_{CNS}$ range:** 80 - 50	$\frac{1}{9.00} \cdot \frac{(s+0.85\cdot10^{-3})(s+6.42\cdot10^{-5})}{(s+10.0\cdot10^{-3})(s+4.33\cdot10^{-4})(s+3.84\cdot10^{-5})}$	$e^{-12.4\cdot s} \cdot \frac{4.77}{(s+34.2\cdot10^{-3})}$
#18	**Administration:** infusion **Age group:** 50-60 yrs old **WAV $_{CNS}$ range:** 80 - 50	$\frac{1}{8.39} \cdot \frac{(s+0.85\cdot10^{-3})(s+6.47\cdot10^{-5})}{(s+10.7\cdot10^{-3})(s+4.34\cdot10^{-4})(s+3.88\cdot10^{-5})}$	$e^{-10.5\cdot s} \cdot \frac{4.40}{(s+33.1\cdot10^{-3})}$

Table C.3: PKPD Propofol nominal models - part III

Model	Model Characteristics		$w_{nom}(s)$	$\tilde{w}(s)$
#1	**Administration:** **Age group:** **WAV $_{CNS}$ range:**	infusion + bolus 18-60 yrs old 80 - 20	$e^{-7 \cdot s} \cdot 42.2 \cdot 10^{-3} \cdot \frac{(s+51.0)(s+6.60 \cdot 10^{-4})(s+1.27 \cdot 10^{-4})(s+7.10 \cdot 10^{-5})}{(s+8.90 \cdot 10^{-1})(s+0.97 \cdot 10^{-3})(s+0.95 \cdot 10^{-4})(s+9.50 \cdot 10^{-5})}$	$9.46 \cdot \left(\frac{(s+3.90 \cdot 10^{-2})(s+1.60 \cdot 10^{-4})}{(s+1.16 \cdot 10^{-1})(s+1.70 \cdot 10^{-4})} \right)^2$
#2	**Administration:** **Age group:** **WAV $_{CNS}$ range:**	infusion + bolus 18-29 yrs old 80 - 20	$e^{-3 \cdot s} \cdot 61.0 \cdot 10^{-3} \cdot \frac{(s+1.34)(s+7.10 \cdot 10^{-4})(s+1.45 \cdot 10^{-4})(s+7.50 \cdot 10^{-5})}{(s+0.45 \cdot 10^{-1})(s+1.10 \cdot 10^{-3})(s+1.10 \cdot 10^{-4})(s+9.90 \cdot 10^{-5})}$	$81.8 \cdot \left(\frac{(s+3.10 \cdot 10^{-2})(s+1.40 \cdot 10^{-4})}{(s+2.75 \cdot 10^{-1})(s+1.50 \cdot 10^{-4})} \right)^2$
#3	**Administration:** **Age group:** **WAV $_{CNS}$ range:**	infusion + bolus 30-39 yrs old 80 - 20	$e^{-8 \cdot s} \cdot 87.1 \cdot 10^{-3} \cdot \frac{(s+20.8)(s+7.10 \cdot 10^{-4})(s+1.30 \cdot 10^{-4})(s+7.10 \cdot 10^{-5})}{(s+9.90 \cdot 10^{-1})(s+1.10 \cdot 10^{-3})(s+0.99 \cdot 10^{-4})(s+9.30 \cdot 10^{-5})}$	$5.12 \cdot \left(\frac{(s+3.40 \cdot 10^{-2})(s+0.86 \cdot 10^{-4})}{(s+0.78 \cdot 10^{-1})(s+0.92 \cdot 10^{-4})} \right)^2$
#4	**Administration:** **Age group:** **WAV $_{CNS}$ range:**	infusion + bolus 40-49 yrs old 80 - 20	$e^{-5 \cdot s} \cdot 83.9 \cdot 10^{-3} \cdot \frac{(s+18.7)(s+6.60 \cdot 10^{-4})(s+1.34 \cdot 10^{-4})(s+7.30 \cdot 10^{-5})}{(s+8.90 \cdot 10^{-1})(s+0.99 \cdot 10^{-3})(s+0.99 \cdot 10^{-4})(s+9.90 \cdot 10^{-5})}$	$4.26 \cdot \left(\frac{(s+4.90 \cdot 10^{-2})(s+1.70 \cdot 10^{-4})}{(s+1.02 \cdot 10^{-1})(s+1.80 \cdot 10^{-4})} \right)^2$
#5	**Administration:** **Age group:** **WAV $_{CNS}$ range:**	infusion + bolus 50-60 yrs old 80 - 20	$e^{-4 \cdot s} \cdot 68.1 \cdot 10^{-3} \cdot \frac{(s+31.1)(s+7.10 \cdot 10^{-4})(s+1.33 \cdot 10^{-4})(s+7.50 \cdot 10^{-5})}{(s+9.90 \cdot 10^{-1})(s+1.10 \cdot 10^{-3})(s+1.10 \cdot 10^{-4})(s+9.10 \cdot 10^{-5})}$	$4.22 \cdot \left(\frac{(s+3.70 \cdot 10^{-2})(s+0.95 \cdot 10^{-4})}{(s+0.76 \cdot 10^{-1})(s+1.00 \cdot 10^{-4})} \right)^2$
#6	**Administration:** **Age group:** **WAV $_{CNS}$ range:**	infusion 18-60 yrs old 80 - 20	$e^{-7 \cdot s} \cdot 65.4 \cdot 10^{-3} \cdot \frac{(s+2.59)(s+4.50 \cdot 10^{-2})}{(s+7.90 \cdot 10^{-2})(s+5.90 \cdot 10^{-2})}$	$45.6 \cdot \left(\frac{(s+2.40 \cdot 10^{-2})}{(s+1.67 \cdot 10^{-1})} \right)^2$
#7	**Administration:** **Age group:** **WAV $_{CNS}$ range:**	bolus 18-60 yrs old 80 - 20	$e^{-7 \cdot s} \cdot 105.5 \cdot 10^{-3} \cdot \frac{(s+1.60)(s+4.50 \cdot 10^{-2})}{(s+7.90 \cdot 10^{-2})(s+5.90 \cdot 10^{-2})}$	$13.5 \cdot \left(\frac{(s+4.10 \cdot 10^{-2})}{(s+1.55 \cdot 10^{-1})} \right)^2$

Table C.4: PKPD Propofol uncertainty weights - part I

Model	Model Characteristics		$w_{nom}(s)$	$\hat{w}(s)$
#8	**Administration:** **Age group:** **WAV $_{CNS}$ range:**	infusion + bolus 18-60 yrs old 80 - 50	$e^{-7\cdot s}\cdot 127.1\cdot 10^{-3}\cdot\frac{(s+12.8)(s+6.70\cdot 10^{-4})(s+1.24\cdot 10^{-4})(s+7.10\cdot 10^{-5})}{(s+8.90\cdot 10^{-1})(s+0.94\cdot 10^{-3})(s+0.95\cdot 10^{-4})(s+9.30\cdot 10^{-5})}$	$8.96\cdot\left(\frac{(s+3.60\cdot 10^{-2})(s+0.85\cdot 10^{-4})}{(s+1.11\cdot 10^{-1})(s+0.95\cdot 10^{-4})}\right)^2$
#9	**Administration:** **Age group:** **WAV $_{CNS}$ range:**	infusion + bolus 18-60 yrs old 60 - 30	$e^{-7\cdot s}\cdot 77.8\cdot 10^{-3}\cdot\frac{(s+25.5)(s+6.80\cdot 10^{-4})(s+1.31\cdot 10^{-4})(s+7.30\cdot 10^{-5})}{(s+8.90\cdot 10^{-1})(s+0.99\cdot 10^{-3})(s+0.99\cdot 10^{-4})(s+9.70\cdot 10^{-5})}$	$9.73\cdot\left(\frac{(s+3.60\cdot 10^{-2})(s+1.20\cdot 10^{-4})}{(s+1.10\cdot 10^{-1})(s+1.30\cdot 10^{-4})}\right)^2$
#10	**Administration:** **Age group:** **WAV $_{CNS}$ range:**	infusion + bolus 18-60 yrs old 40 - 20	$e^{-7\cdot s}\cdot 62.2\cdot 10^{-3}\cdot\frac{(s+43.0)(s+7.10\cdot 10^{-4})(s+1.26\cdot 10^{-4})(s+7.30\cdot 10^{-5})}{(s+9.90\cdot 10^{-1})(s+1.10\cdot 10^{-3})(s+0.97\cdot 10^{-4})(s+9.50\cdot 10^{-5})}$	$9.53\cdot\left(\frac{(s+3.40\cdot 10^{-2})(s+1.60\cdot 10^{-4})}{(s+1.03\cdot 10^{-1})(s+1.70\cdot 10^{-4})}\right)^2$
#11	**Administration:** **Age group:** **WAV $_{CNS}$ range:**	infusion 18-29 yrs old 80 - 20	$e^{-3\cdot s}\cdot 94.7\cdot 10^{-3}\cdot\frac{(s+19.5)(s+3.70\cdot 10^{-2})}{(s+9.90\cdot 10^{-1})(s+5.90\cdot 10^{-2})}$	$4.75\cdot\left(\frac{(s+2.80\cdot 10^{-2})}{(s+6.40\cdot 10^{-2})}\right)^2$
#12	**Administration:** **Age group:** **WAV $_{CNS}$ range:**	infusion 30-39 yrs old 80 - 20	$e^{-8\cdot s}\cdot 135.1\cdot 10^{-3}\cdot\frac{(s+15.3)(s+3.30\cdot 10^{-2})}{(s+9.90\cdot 10^{-1})(s+5.90\cdot 10^{-2})}$	$3.14\cdot\left(\frac{(s+4.20\cdot 10^{-2})}{(s+8.06\cdot 10^{-2})}\right)^2$
#13	**Administration:** **Age group:** **WAV $_{CNS}$ range:**	infusion 40-49 yrs old 80 - 20	$e^{-5\cdot s}\cdot 130.1\cdot 10^{-3}\cdot\frac{(s+2.23)(s+4.50\cdot 10^{-2})}{(s+1.90\cdot 10^{-1})(s+5.90\cdot 10^{-2})}$	$8.25\cdot\left(\frac{(s+5.10\cdot 10^{-2})}{(s+1.56\cdot 10^{-1})}\right)^2$
#14	**Administration:** **Age group:** **WAV $_{CNS}$ range:**	infusion 50-60 yrs old 80 - 20	$e^{-3\cdot s}\cdot 105.6\cdot 10^{-3}\cdot\frac{(s+5.43)(s+4.10\cdot 10^{-2})}{(s+2.90\cdot 10^{-1})(s+5.90\cdot 10^{-2})}$	$6.14\cdot\left(\frac{(s+6.20\cdot 10^{-2})}{(s+1.61\cdot 10^{-1})}\right)^2$

Table C.5: PKPD Propofol uncertainty weights - part II

Model	Model Characteristics		$w_{nom}(s)$	$\hat{w}(s)$
#15	**Administration:** **Age group:** **WAV $_{CNS}$ range:**	infusion 18-29 yrs old 80 - 50	$e^{-4\cdot s}\cdot 313.2\cdot 10^{-3}\cdot\frac{(s+0.35)(s+4.10\cdot 10^{-2})}{(s+0.75\cdot 10^{-1})(s+5.90\cdot 10^{-2})}$	$11.8\cdot\left(\frac{(s+2.50\cdot 10^{-2})}{(s+1.12\cdot 10^{-1})}\right)^2$
#16	**Administration:** **Age group:** **WAV $_{CNS}$ range:**	infusion 30-39 yrs old 80 - 50	$e^{-9\cdot s}\cdot 402.2\cdot 10^{-3}\cdot\frac{(s+1.17)(s+3.70\cdot 10^{-2})}{(s+3.10\cdot 10^{-1})(s+5.50\cdot 10^{-2})}$	$4.77\cdot\left(\frac{(s+3.30\cdot 10^{-2})}{(s+1.02\cdot 10^{-1})}\right)^2$
#17	**Administration:** **Age group:** **WAV $_{CNS}$ range:**	infusion 40-49 yrs old 80 - 50	$e^{-5\cdot s}\cdot 420.0\cdot 10^{-3}\cdot\frac{(s+0.34)(s+4.10\cdot 10^{-2})}{(s+0.99\cdot 10^{-1})(s+5.90\cdot 10^{-2})}$	$4.68\cdot\left(\frac{(s+3.40\cdot 10^{-2})}{(s+0.98\cdot 10^{-1})}\right)^2$
#18	**Administration:** **Age group:** **WAV $_{CNS}$ range:**	infusion 50-60 yrs old 80 - 50	$e^{-4\cdot s}\cdot 317.0\cdot 10^{-3}\cdot\frac{(s+0.48)(s+4.10\cdot 10^{-2})}{(s+0.89\cdot 10^{-1})(s+5.90\cdot 10^{-2})}$	$6.90\cdot\left(\frac{(s+3.30\cdot 10^{-2})}{(s+1.13\cdot 10^{-1})}\right)^2$

Table C.6: PKPD Propofol uncertainty weights - part III

Appendix D

Clinical Studies

In the course of this Ph.D. work, 3 major clinical studies were conducted at the UBC site of the Vancouver General Hospital. Each study was conducted after being approved by the Ethic Review Board, and after obtaining patients' consent. This appendix presents the protocol specificities, population demographics, and results of each study.

D.1 Arthroscopy Observational Study

Initial Motivation: Validation and improvement of the WAV_{CNS} algorithm.

Clinical Question: Is the WAV_{CNS} a clinical alternative to BIS?

Timeline: March 2002 to April 2002

D.1.1 Protocol

With institutional ethics board approval and patient consent, 25 patients were studied during general anesthesia for various ambulatory knee surgeries (arthroscopies, menisectomies and anterior cruciate ligament repair).

The raw EEG signal was acquired at a sampling rate of 128 S/s using the BIS A-1050 monitor with the updated software v.3.4, and standard BIS sensors. Processed variables such as the BIS, MEF, SEF, EMG, and impedance information were recorded using the monitor's serial interface. (These variables are calculated by the BIS Monitor using proprietary algorithms.) All BIS data were stored onto a portable computer.

The WAV_{CNS} was calculated online based on the raw EEG data obtained from the A-1050, and based on the implementation outlined in Section 5.2.2. The awake PDF reference was obtained from a 10-minute awake EEG signal acquired from 5 adult volunteers, and using the same A-1050 monitor (this was done to account for differences in frequency characteristics between the A-1050 and the Crystal Monitor). The

artifact de-noising algorithm was de-activated after the initial induction bolus. The post-processing IIR trending filter was replaced by a 30-second averaging filter in order to allow for better comparison with the BIS. Both the BIS and WAV_{CNS} were displayed in real time on the computer screen.

In all cases, induction was initiated by a propofol bolus ($3.6\pm1.15\ \mu g \cdot kg^{-1}$). Maintenance of anesthesia was sustained by anesthetic vapour (sevoflurane or desflurane) and supplemented with nitrous oxide or an infusion of propofol. Note that only anesthetic drugs provoking a concentration dependent change in the EEG such as described in Fig. 5.5 were used. Opioid boluses were given when necessary. The anesthetic regimen was left entirely to the anesthesiologist's discretion, who remained blinded to both the BIS and WAV_{CNS}. Note that muscle relaxants were not administered and airway control was ensured via laryngeal masks in all cases. An overview of the patients' data and titration are presented in Table D.1.

All relevant surgical and anesthetic events were logged by an anesthesia resident. During the induction phase, patients were asked to count down from 100. The loss of count event was logged once patients stopped counting. In the setting of general anesthesia, this event is usually taken as an indicator of Loss Of Consciousness (LOC). Motor movements and airway reactions during the LMA insertion were also recorded. Finally, during the emergence, the anesthesiologist stimulated the patient to provoke a reaction. Once a reaction was observed, the ability to respond to a verbal command was assessed by calling each patient by name.

Five cases were rejected from the analysis (3 cases for incomplete data, 1 case for poor signal quality and 1 case for inaccuracies in the events log).

D.1.2 Results

The results of this study are presented in Section 5.2.3.

D.1.3 Comments and Unexpected Results

One particularly interesting result concerns the difference between the Movers and Non-Movers groups at induction, see Figure D.1. While the averaged BIS time course is similar in both cases, the WAV_{CNS} behavior during induction is clearly different between the two groups. This could indicate that the WAV_{CNS} can actually predict whether a patient is going to react to the LMA insertion. This could mean that the WAV_{CNS} is sensitive enough during induction to give the anesthesiologist an advance warning if the patient is not in an adequate hypnotic depth suitable for airway manipulation. This observation was at the core of the LMA Study presented in Annex D.2, and the propofol PD modeling work discussed in Chapter 6.

D.2 Laryngeal Mask Airway Study

Initial Motivation: Confirmation of the arthroscopy study result.

Clinical Question: Can the WAV_{CNS} predict reaction to LMA insertion?

Case	Patient				Titration		
	Gender	Age	Weight	ASA	Premed.	Induction	Maintenance
4	F	25	71	I	Suf. 12.5	Prop. 400	Prop. infusion ; Des. (0.7 MAC) + N_2O ; bolus of morphine
5	M	31	90	I	Suf. 12.5	Prop. 500	Prop. infusion ; Des. (0.9 MAC) + N_2O ; bolus of morphine
6	M	43	68	I	Suf. 7.5	Prop. 400	Des. (0.8 MAC) ; bolus of Suf.
8	M	28	87	II	Midaz. 2	Prop. 250	Prop. infusion ; Des. (0.8 MAC) ; bolus of Fent.
9	M	21	99	I	Fent. 100	Prop. 300	Prop. infusion ; Des. (1.2 MAC) + N_2O ; bolus of Fent.
10	F	30	55	I	Midaz. 1 + Fent. 75	Prop. 250	Prop. infusion ; Des. (0.9 MAC)
11	F	34	66	I	Fent. 100	Prop. 250	Prop. infusion ; Des. (0.8 MAC) ; bolus of Suf.
12	M	34	103	I	Fent. 100	Prop. 300	Prop. bolus/infusion ; Des. (1.4 MAC) + N_2O
13	F	25	77	I	Midaz. 1.5 + Suf. 15	Prop. 250	Prop. bolus/infusion ; Sevo. (0.7 MAC) ; bolus of Suf.
14	M	19	80	I	Midaz. 2 + Suf. 7.5	Prop. 280	Prop. infusion ; Sevo. (0.8 MAC) ; bolus of Suf. ; Naloxone
15	M	57	80	I	Midaz. 2	Prop. 200 + Remif. 150	Prop. infusion ; Des. (0.7 MAC) ; bolus of Fent.
16	M	45	85	I	Midaz. 2	Prop. 250 + Remif. 150	Prop. bolus/infusion ; Des. ; bolus of Fent.
17	F	24	60	II	Midaz. 1 + Fent. 50	Prop. 170	Prop. bolus/infusion ; Thio. bolus ; Sevo. (0.8 MAC) ; bolus of Fent.
18	M	57	85	I	Midaz. 1.5 + Fent. 50	Prop. 250	Prop. infusion ; Sevo. (0.8 MAC)
19	M	47	108	I	Fent. 75	Prop. 290	Sevo. (1.5 MAC) ; bolus of Fent. ; Naloxone
20	M	55	88	I	Fent. 75	Prop. 280	Prop. bolus ; Sevo.
22	M	64	104	I	Suf. 10	Prop. 270	Prop. bolus ; Sevo. (1.2 MAC)
23	M	56	113	II	Suf. 5	Prop. 270	Prop. bolus ; Sevo. (1.5 MAC)
24	F	33	50	I	Midaz. 1.5 + Suf. 2.5	Prop. 200 + Remif. 150	Prop. bolus/infusion ; Sevo. (1.2 MAC) ; bolus of Suf.
25	M	39	100	I	*none*	Prop. 300	Prop. bolus/infusion ; Sevo. (0.8 MAC)

Suf.: Sufentanil; Midaz.: Midazolam; Prop.: Propofol; Fent.: Fentanyl; Remif.: Remifentanil
Sevo.: Sevoflurane; Des.:Desflurane; Thio.: Thiopental
Values in parenthesis indicate the mean MAC value during maintenance.

Table D.1: Patients' data and titration. Doses are expressed in μg (sufentanil, fentanyl, remifentanil) or mg (propofol).

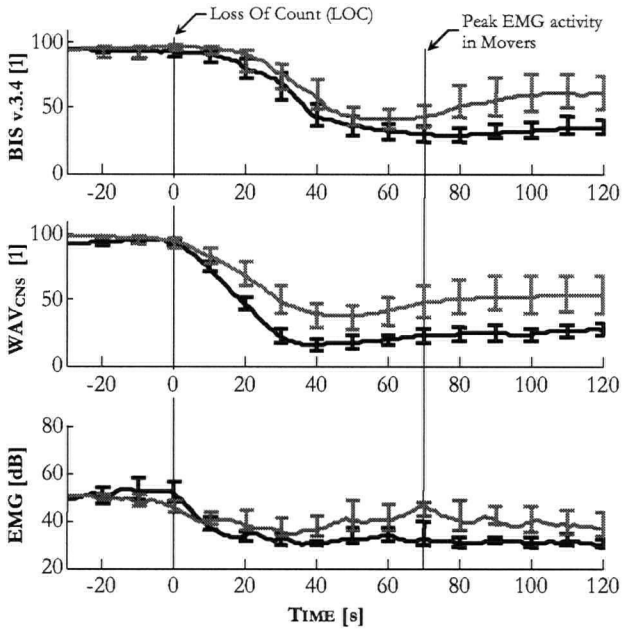


Figure D.1: Comparison between the Movers and Non-Movers groups.

Timeline: October 2003 to April 2004

D.2.1 Protocol

Following the arthroscopy study, it was hypothesized that the WAV_{CNS} time course could be an early indicator of reaction to the airway management. The use of an LMA instead of a endotracheal tube allows the observation of patient movements during the LMA placement, and therefore provides an accurate observable endpoint.

With ethic board approval, and after obtaining patient consent, 76 patients (18-65 yrs old, ASA I-III) were enrolled to study the WAV_{CNS} time course during induction, and the incidence of reaction to airway management. Table D.2 and Table D.3 present the study population demographics.

Prior to induction, a 2-channel EEG recording was initiated using a Crystal Monitor 16 (CleveMed, Cleveland, OH). The electrode positions were Fp1-AT1[1] (channel 1) and Fp2-AT2 (channel 2). Fpz was used as the referential electrode for the two bipolar montages. The 2 signals were acquired at a sampling rate of 480 S/s and resolution of 16 bits.

Each patient first received a bolus of fentanyl (0.5-1.5 μg/kg) followed by a bolus of propofol (2-3 mg/kg). The exact drug dosages were left to the anesthesiologist discretion[2]. Both the start and end of the propofol syringe push were recorded by the study monitor. During propofol administration, each patient was asked to count upward from 1 until loss of consciousness (LOC) occurred. The time to LOC was recorded by the study monitor.

[1]AT: Anterior Temporal (BIS eye level electrode position).

[2]This study was not a controlled study

Patient #	ASA	Demographics				Dosage		Grade	Comments
		Weight [kg]	Height [cm]	Age	Gender	Fentanyl [μg]	Propofol [mg]		
001									Bad signal quality
002	n.c.	80	175	21	Male	120	200	3	
003	n.c.	78	175	30	Male	100	200	1	
004	n.c.	80	177	49	Male	75	200	1	
005	n.c.	68	165	53	Female	75	150	4	
006	n.c.	73	180	34	Male	75	200	1	
007	n.c.	100	178	21	Male	75	200	1	
008	n.c.	59	168	28	Female	75	180	1	
009	n.c.	71	178	34	Male	100	200	1	
010	1	90	190	26	Male	100	200	2	
011	1	64	n.c.	27	Female			5	Multiple attempts
012	1	89	176	36	Male	150	200	3	
013	1	70	170	19	Female	75	150	3	
014	1	83	180	22	Male	80	200	3	
015	1	53	157	21	Female	75	150	2	
016	1	90	185	19	Male	75	200	2	
017	1	66	170	25	Female	50	150	4	Time of insertion outside protocol range
018	1	63	165	60	Female	50	180	2	
019	1	138	196	20	Male	125	300	2	Time of insertion outside protocol range
020	1	92	187	28	Male	50	200	4	Medical student
021	1	68	173	20	Male	75	300	3	Dose outside protocol range
022	1	85	181	32	Male	100	200	3	
023	1	60	162	28	Female	75	200	1	Dose outside protocol range
024	1	56	162	27	Female	75	200	4	Dose outside protocol range
025	2	66	168	46	Female	75	200	1	Time of insertion outside protocol range
026	1	82	182	38	Male	50	200	2	
027	1	99	182	45	Male	100	300	2	Medical student
028	1	86	173	30	Female	50	200	4	Medical student
029	1	91	189	35	Male	100	300	1	Medical student
030	1	78	170	24	Male	100	200	1	Medical student
031	1	77	173	67	Male	50	150	2	Medical student
032	1	70	170	38	Male	50	180	3	Medical student
033	1	91	185	52	Male	100	220	3	Medical student
034	2	68	160	46	Female	100	150	3	
035	1	68	164	19	Female	50	200	2	Time of insertion outside protocol range
036	1	85	180	38	Male	100	250	1	
037	1	88	165	27	Male	100	250	3	Time of insertion outside protocol range
038	1	70	170	25	Male	75	220	2	Dose outside protocol range
039	2	69	163	38	Female	75	210	3	Time of insertion outside protocol range
040	1	77	173	48	Male	100	250	1	Dose outside protocol range
041	2	100	182	52	Male	125	300	2	Time of insertion outside protocol range
042	2	98	189	47	Male	125	280	1	
043	1	79	172	46	Male	100	220	1	
044	3	108	193	59	Male	150	300	1	
045	1	70	176	20	Male	100	210	1	
046	1	81	180	23	Male	125	240	1	
047	1	86	168	33	Female	125	250	1	
048	1	83	178	18	Male	150	250	1	
049	1	75	176	38	Male	100	220	2	

Table D.2: LMA Study population (patients #001 to #049).

Patient #	ASA	Demographics				Dosage		Grade	Comments
		Weight [kg]	Height [cm]	Age	Gender	Fentanyl [μg]	Propofol [mg]		
050									Bad signal quality
051	1	91	178	39	Male	50	250	1	
052	1	96	193	40	Male	50	270	1	
053	1	67	163	21	Female	50	200	2	
054	1	86	180	53	Male	50	250	2	
055	1	90	177	21	Male	75	270	1	
056	1	84	183	39	Male	75	250	3	
057	1	97	176	52	Male	125	200	1	
058	1	88	183	22	Male	125	200	1	
059	1	88	180	38	Male	125	200	3	
060	1	95	183	50	Male	140	250	1	
061	1	78	178	34	Male	50	200	1	
062	1	83	186	48	Male	100	200	3	
063	1	65	163	36	Female	90	180	1	
064	2	56	173	52	Female	50	150	1	
065	1	58	157	34	Female	50	165	1	
066	1	59	162	21	Male	50	150	1	
067	1	72	175	36	Male	50	220	2	
068	1	77	180	33	Male	50	220	2	
069	1	77	176	40	Male	50	220	2	
070	2	100	180	54	Male	75	250	1	
071	1	72	176	19	Male	75	180	1	
072	1	97	190	48	Male	100	200	1	
073	1	64	170	56	Female	50	150	3	
074	1	105	185	39	Male	100	300	2	
075	3	95	176	60	Male	75	190	2	
076	1	75	170	23	Female	75	170	1	

Table D.3: LMA Study population (patients #050 to #076).

	GROUP NR (n=37)	GROUP NR (n=13)	p
Age [yrs]	36	36	$\approx$0.5
Weight [kg]	80	77	>0.2
Fentanyl [μg/kg]	1.08	1.16	>0.2
Propofol [mg/kg]	2.68	2.35	<0.0002

Table D.4: LMA study population - *demographics and dosage summary.*

Forty (40) seconds after LOC, jaw tone was assessed by the attending anesthesiologist and the LMA was inserted. Comments about jaw tone were logged in by the study monitor. During the LMA insertion, any observed movement from the patient was also logged in.

The EEG recording was discontinued after the start of the surgery. The EEG data were analyzed *post hoc* to derive the WAV_{CNS}.

Among the 76 recorded cases, 26 cases had to be removed from the study population (2 cases were cancelled due to computer malfunction, 3 cases presented large EEG artifacts during induction, 6 cases did not fit the dosage protocol, 6 cases were removed due to delayed LMA insertion, and 9 cases were removed due to the use of a Proseal LMA instead of the classic LMA model[3]). After review of the intra-operative events, the remaining 50 cases were classified into two groups (Non-Reactors NR; Reactors R) depending on jaw tone, and the severity and type of movement. Patients with facial activity and moderate jaw tone during LMA insertion were classified into the R group (n=13). Patients with mild jaw tone and some extremity movements were classified into the NR group (n=37). About one third of the study population reacted to the LMA insertion. This figure is consistent with clinical observations. The study demographics and anesthetic dosing regimen are summarized in Table D.4.

D.2.2 LMA Results

There were no patient demographic differences (age, weight, gender) between the groups. However, the propofol induction dose was significantly larger in Group NR (2.68±0.31 mg/kg) than in Group R (2.35±0.23 mg/kg).

The slope of the WAV_{CNS} descent was found to be an early predictor of LMA reaction. The bottom plot in Figure D.2 shows the average slope of the WAV for both groups. At LOC+20 seconds, a gradient steeper than -2 s^{-1} is characteristic of Group NR but not of Group R. Using this parameter to predict reaction to LMA insertion, we obtain an accuracy of 84% (77% sensitivity; 86% specificity).

A conclusion of this study is that the WAV_{CNS} is the first non-evoked EEG variable to be shown to predict motor response to LMA insertion during induction of general anesthesia [188]. The dynamic response of the WAV_{CNS} may therefore allow clinicians to assess the degree of the hypnotic effect of propofol at induction of anesthesia, and the likelihood of reaction.

[3]The Proseal LMA model is more stimulating than the classic LMA model.

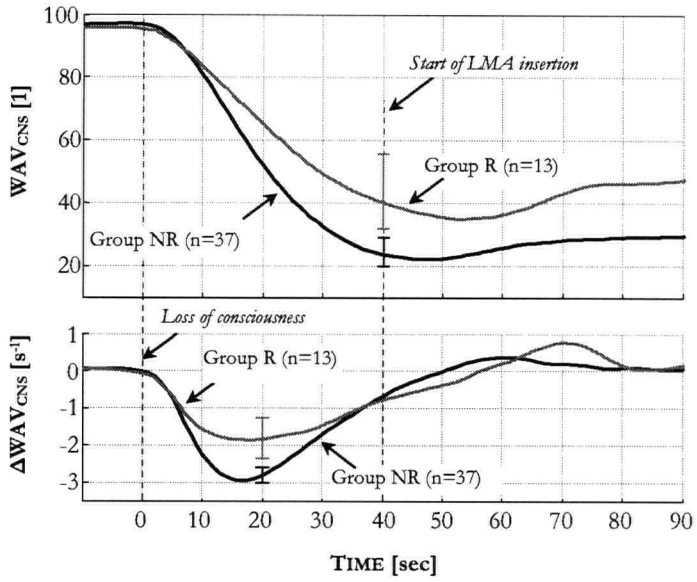


Figure D.2: Average time courses of WAV_{CNS} and ΔWAV_{CNS} for both groups. ΔWAV_{CNS} is the rate of change of the WAV_{CNS} index. Note the significant difference between the two groups *before* the LMA insertion.

This result is, however, surprising. The time course of the index seems to predict whether patients *are going to* react to the insertion, even though the stimulation has not yet started. Also, the WAV_{CNS} is a depth of hypnosis index and not an analgesic index. In this respect, it is counter-intuitive to expect the index to react, yet alone predict, a reaction to an external stimulation.

One explanation for this is that the propofol dosage of the two study groups are significantly different. The reactor group received less propofol than the non-reactor group. This may provides an explanation for the fact that the patients in this group reacted, and that the time course of the WAV_{CNS} was not as deep as in the non-reactor group. Hence, even though the average propofol dose difference between the two groups is only 0.33 mg/kg, the WAV_{CNS} seems to capture adequately this difference in the propofol dosing regimen. The relationship between the induction propofol dose and the WAV_{CNS} value 40 seconds after LOC is presented in Figure D.3.

This result encouraged us to use the induction time course of the WAV_{CNS} for propofol PD modeling.

D.3 Electro-Convulsive Shock Therapy Study

Initial Motivation: Assessment the intra- and inter-patient variability of hypnotic depth indexes such as BIS XP and WAV_{CNS} during induction.

Clinical Question: Can the BIS XP be useful in predicting ECT seizure duration?

Timeline: January 2002 to November 2002

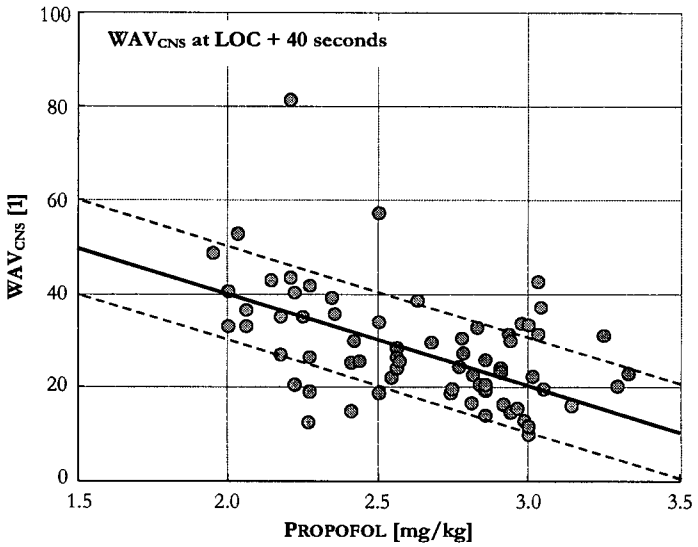


Figure D.3: Correlation between WAV_{CNS} value 40 seconds after LOC and the administered propofol dose (note that the study population was extended to patients receiving propofol doses outside the protocol range).

The ECT procedure provides a unique opportunity to study intra- and inter-patient variability in terms of the anesthetic dose *vs.* response relationship. Although the study population was limited, the identified PKPD time delays and gains were shown to be fairly consistent for each patient. This result corresponds to daily clinical observations and substantiates the fact that the WAV_{CNS} is an adequate endpoint for automatic or manual titration, in that it differentiates between patients' responses to anesthetic drugs, hence capturing each patient individual pharmacological needs.

D.3.0.1 ECT Study

The ECT treatment is proposed to patients suffering from depression and other bipolar illnesses. Each treatment consists in a single electro-shock to induce a localized seizure. It is believed that the ECT seizure can restore the chemical balance in the central nervous system. The therapy usually involves a series of repeats administered over a few weeks span.

The anesthetic regimen consists in a single bolus of thiopental for sedation followed by a bolus of succinylcholine to prevent injury during the seizure. While patients receive an induction dose calculated according to their particular physionomy and medical history, the dose is usually fixed during the therapy (small variations of the total dose have occurred for patients #1 and #2 after re-evaluation of these patients by attending anesthesiologists). This procedure is thus unique in providing a framework to study both intra- and inter-patient variability.

Study Population After the study protocol was reviewed by the UBCH ethic board, 5 patients scheduled to receive at least 6 ECT treatments were consented. The patients' data are summarized in Table D.5.

Each patient was monitored using the BIS XP Monitor (Aspect Medical Systems Inc.). A proprietary

Table D.5: ECT Study Population

PATIENT	#1	#2	#3	#4	#5
Gender	M	F	F	M	F
Age [yr]	55	41	66	74	73
Weight [kg]	96	97	76	76	41
Height [cm]	160	172	155	178	152
Thiopental [mg]					
repeat #1	350	250	150	200	100
repeat #2	350	225	150	200	100
repeat #3	300	225	150	200	100
repeat #4	275	225	150	200	100
repeat #5	275	225	150	200	100
repeat #6	275	225	150	na[1]	100

[1] Patient withdrew from treatment before completion of study

interface software was used to download onto a laptop computer the raw EEG signal and other processed parameters. During the procedure, we logged in the time corresponding to the end of thiopental injection and the time corresponding to the application of the ECT shock.

The raw EEG signal acquired during the procedure was processed off-line to calculate the WAV_{CNS}. The WAV_{CNS} algorithm provides a measurement of the patient state every second. The raw index was trended using a 2^{nd} order low pass filter with a cutoff frequency of 0.037 Hz to extract the overall trend (a filter based on a 15 s averaging window yields similar results in terms of the inherent latency introduced by the filter between the input and the output signals).

Identification The dose/response relationship of thiopental can be expressed by the PKPD model presented in Fig. 6.15.b. Due to the limited identification window and fast onset of action of thiopental, the WAV_{CNS} trending filter was replaced by a faster dynamics:

$$H_{\text{CNS}}(s) = \frac{1}{(4.35 \cdot s + 1)^2} \cdot \frac{1 - e^{-s}}{s} \tag{D.1}$$

The identification can only be carried out from the start of injection up to the application of the ECT shock. This usually corresponds to a window of 1 to 2 minutes during which the WAV_{CNS} follows an induction-type time course, see Fig. D.4.a-b. Cases where the WAV_{CNS} does not reach a valley value before the ECT is applied were eliminated from the analysis, see Fig. D.4.c. Similarly, cases exhibiting a strong EMG activity due to succinylcholine-induced fasciculation were eliminated from the analysis, see Fig.D.4.d.

Due to the scarcity of the data, only an attempt to identify the parameter set $\{T_d, k_d, EC_{50}, \gamma\}$ can be made. To do so, the pharmacokinetic part of the model was replaced using the PK models proposed by

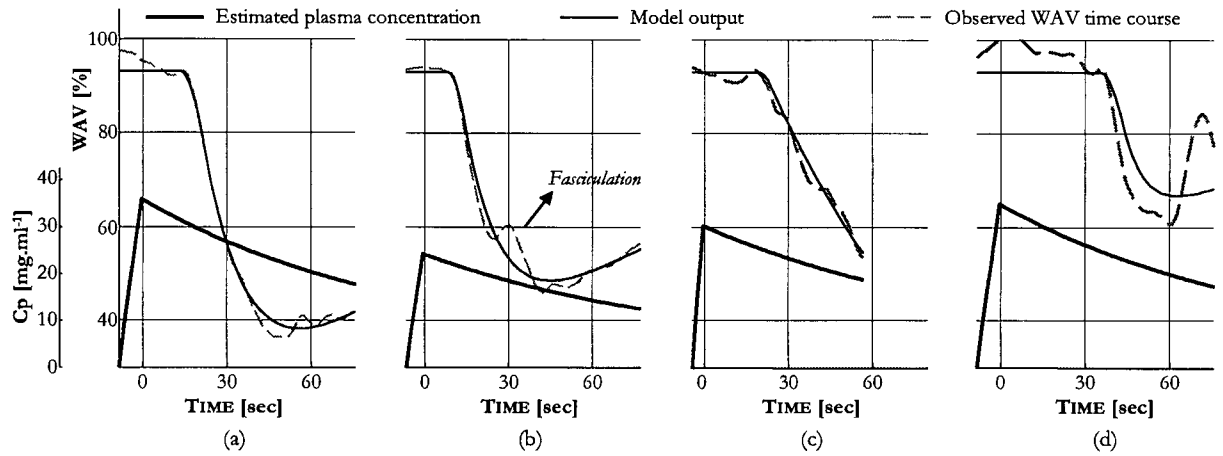


Figure D.4: Identification procedure for 4 different cases. The thiopental blood plasma concentration C_p is estimated through the 3-compartment PK model proposed by Stanski *et al.* This model is weight and age dependent. (a) Successful identification. (b) Successful identification despite limited fasciculation. (c) Identification unsuccessful due incomplete data (the index does not reach its peak value). (d) Unsuccessful identification due to a large fasciculation disturbance.

Stanski *et al.* [168]. These models are age and weight dependent and were derived based on infusion data. As such, their applicability is limited to infusion-based titrations. Using these models to estimate the blood plasma concentration following an induction bolus is bound to yield some modeling error. However, it has been shown that this error will eventually be compensated by the identified PD dynamics [165]. As a result, the overall integrated PKPD model will appropriately describe the dose/response dynamics.

Results In Fig. 7.6, the individual time courses of the WAV_{CNS} for two patients are overlayed and synchronized at the end of thiopental injection. The averaged WAV_{CNS} time courses are also represented. We observe that the WAV_{CNS} is remarkably consistent from repeat to repeat. Fig. 7.6 also emphasizes the large difference between these two patients in terms of the pharmacokinetic time delay. One patient reacts quickly to the injected anesthetic, while there is a large delay in the other patient. Similar observations were made for the 3 other patients.

In order to better quantify intra- and inter-patient variability, the identification procedure discussed in the previous section was carried out. The results are presented in Table D.7. While the identification also stresses the large difference between patients in terms of the pharmacokinetic time delay, a similar result concerning the overall PKPD gain is revealed. This gain can vary by as much as 55% from a median value of 13.3, while the maximum variation observed within one patient is only of 27%.

The results concerning the identified time constant τ are more difficult to interpret since this parameter partially models the true pharmacodynamic equilibration time constant, but also compensates for the modeling error introduced by the use of infusion-based PK models. As a result, it has been asserted that it does not bear any physical meaning.

The observed time delay and gain variations can be explained by considering that each patient reacts

Table D.6: Thiopental pharmacokinetic models

	PHARMACOKINETIC PK MODELS (*models are normalized,* i.e. *static gain is 1*)	PK GAIN
Patient 1	$\frac{6.48 \cdot 10^{-4} \cdot s^2 + 8.92 \cdot 10^{-7} \cdot s + 5.51 \cdot 10^{-11}}{s^3 + 1.09 \cdot 10^{-2} \cdot s^2 + 3.77 \cdot 10^{-6} \cdot s + 5.51 \cdot 10^{-11}}$	3.39
Patient 2	$\frac{6.48 \cdot 10^{-4} \cdot s^2 + 8.92 \cdot 10^{-7} \cdot s + 5.51 \cdot 10^{-11}}{s^3 + 1.15 \cdot 10^{-2} \cdot s^2 + 3.82 \cdot 10^{-6} \cdot s + 5.51 \cdot 10^{-11}}$	3.36
Patient 3	$\frac{6.48 \cdot 10^{-4} \cdot s^2 + 8.92 \cdot 10^{-7} \cdot s + 5.51 \cdot 10^{-11}}{s^3 + 1.03 \cdot 10^{-2} \cdot s^2 + 3.74 \cdot 10^{-6} \cdot s + 5.51 \cdot 10^{-11}}$	4.29
Patient 4	$\frac{6.48 \cdot 10^{-4} \cdot s^2 + 8.92 \cdot 10^{-7} \cdot s + 5.51 \cdot 10^{-11}}{s^3 + 9.94 \cdot 10^{-3} \cdot s^2 + 3.71 \cdot 10^{-6} \cdot s + 5.51 \cdot 10^{-11}}$	4.46
Patient 5	$\frac{6.48 \cdot 10^{-4} \cdot s^2 + 8.92 \cdot 10^{-7} \cdot s + 5.51 \cdot 10^{-11}}{s^3 + 9.98 \cdot 10^{-3} \cdot s^2 + 3.72 \cdot 10^{-6} \cdot s + 5.51 \cdot 10^{-11}}$	7.94

The model input is expressed in [mg/min] while the output is expressed in [μg/ml].

differently to the anesthetic drug. For instance, Patient #5 was an older and lighter patient. As a result, a reduced cardiac output may have resulted in an increased arm-to-brain travel time. Furthermore, the patient's small weight explains the rather limited amount of thiopental used for induction, which in turn explains the large PKPD gain. In general, differences in PKPD characteristics can be attributed to differences in age, lean body mass, genetic factors, disease state, drug tolerance and sensitivity, and eventually, drug addictions.

A major limitation of this study is the limited number of patients. We expect that a study targeting a specific population in terms of age and weight will yield similar results in terms of intra-patient variability, but would also show a reduced inter-patient variability.

Table D.7: Thiopental identified pharmacodynamic models

	Pharmacodynamic PD Models							
	Repeat		#1	#2	#3	#4	#5	#6
Patient 1	T_d	[s]	23	28	27	29	28	na^1
	k_d	$[s^{-1} \cdot 10^{-3}]$	na^2	17.4	26.1	36.6	67.9	na^1
	EC_{50}	[μg/ml]	na^2	17.2	19.5	20.4	22.1	na^1
	γ	[1]	na^2	1.6	1.6	1.6	1.4	na^1
	E_0	[1]	98.0	95.8	94.8	95.3	96.0	na^1
Patient 2	T_d	[s]	23	19	18	21	20	15
	k_d	$[s^{-1} \cdot 10^{-3}]$	30.5	41.7	na^2	na^2	46.2	34.4
	EC_{50}	[μg/ml]	15.4	15.9	na^2	na^2	15.3	14.5
	γ	[1]	1.5	1.6	na^2	na^2	1.8	1.5
	E_0	[1]	100.0	100.0	100.0	100.0	92.4	94.0
Patient 3	T_d	[s]	13	10	17	8	10	11
	k_d	$[s^{-1} \cdot 10^{-3}]$	na^2	na^1	48.5	59.9	48.9	81.5
	EC_{50}	[μg/ml]	na^2	na^1	15.3	21.5	16.9	19.7
	γ	[1]	na^2	na^1	1.6	1.5	1.3	1.5
	E_0	[1]	99.6	100.2	99.8	98.5	98.9	98.7
Patient 4	T_d	[s]	29	25	26	32	45	na^3
	k_d	$[s^{-1} \cdot 10^{-3}]$	30.5	26.0	30.9	39.8	na^2	na^3
	EC_{50}	[μg/ml]	20.7	16.4	21.4	17.5	na^2	na^3
	γ	[1]	1.6	1.7	1.9	2.1	na^2	na^3
	E_0	[1]	95.5	97.1	98.1	99.9	99.7	na^3
Patient 5	T_d	[s]	42	38	25	23	39	33
	k_d	$[s^{-1} \cdot 10^{-3}]$	214.0	111.5	37.1	na^2	79.5	68.5
	EC_{50}	[μg/ml]	28.2	22.8	20.2	na^2	22.8	23.3
	γ	[1]	1.3	1.5	1.3	na^2	1.3	1.2
	E_0	[1]	92.7	100.3	99.1	99.0	94.5	97.8

The time delay T_d and time constant K_d are expressed in seconds.
[1] WAV_{CNS} did not settle before ECT shock
[2] Strong fasciculation during induction
[3] Patient withdrew from treatment before completion of study

Appendix E

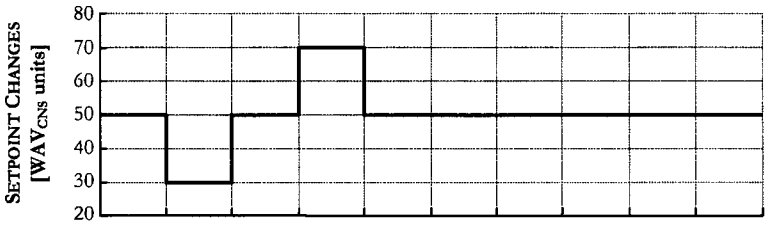
Survey

To assess whether the controller performances are adequate from a clinical point of view, the following survey was sent to Dr. Craig Ries (Vancouver General Hospital, Vancouver, BC), Dr. Mark Ansermino (BC Children Hospital, Vancouver, BC), and Dr. Don Voltz (University Hospitals of Cleveland, Cleveland, Ohio).

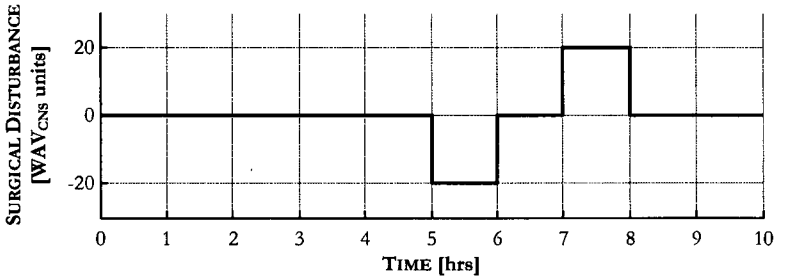
I am in the final stages of designing a controller that adjusts a propofol titration to maintain a desired WAV setpoint. I will simulate the controller response on the 44 patient models that were derived during the LMA study.

To test the controller performance, I will subject the system to the following test:

- once the system has reached a WAV of 50, I will first make a series of setpoint changes that will instruct the controller to target different WAV level (50 to 30, then 30 to 50, then 50 to 70 and finally 70 to 50). Each change will last 1 hour. Ideally, we would like the WAV to follow the exact trajectory plotted in the following figure.



- I will then subject the system to a series of surgical disturbance. These disturbances are modelled as a sudden change in the WAV index:



In this case, the surgical disturbance will first have the WAV to decrease by 20 units (corresponding to a lack of surgical stimulation) for 1 hour. Surgical stimulation will then follow, which will make the WAV jump by 20 units, etc.... Ideally, the controller will see the surgical disturbance and adapt the propofol infusion consequently. Eventually, we would like the WAV output to quickly regain the setpoint of 50.

The controllers that will be designed in the Chapter 8 will be designed for a given population of patients (they will not be specific to each individual). They will yield performances that can be expressed in terms of the settling time, overshoot and steady state error. These terms are defined below:

Settling time: time taken by the controller to settle within 5 units of the targeted WAV.

Overshoot: the amount the system output response proceeds beyond the desired response.

Steady state error: the system output 1 hour after a setpoint step change or a surgical disturbance.

In order for me to conclude whether the controller performances are suitable for clinical use, I would appreciate if you could fill in the following table.

Figure E.1: Survey I.