LUND UNIVERSITY

Assessing Control Performance in Closed-loop Anesthesia

Soltesz, Kristian; Dumont, Guy A.; Ansermino, J. Mark

Assessing Control Performance in Closed-loop Anesthesia

Kristian Soltesz, Guy A. Dumont *Fellow IEEE*, J Mark Ansermino

Abstract— Recently, several control systems for closed-loop anesthesia have been demonstrated both in simulation and clinical studies. A set of performance measures, proposed by Varvel et al., have constituted the standard means of comparing such systems.

This paper debates the adequacy of the Varvel measures, as applied to closed-loop anesthesia, and proposes an alternative set of measures. Key features of the proposed measures are: wide acceptance within the control community; reflection of clinical feasibility; separate measures for induction and maintenance of anesthesia; separation of outlier detection and performance evaluation. The proposed measures are descriptive, few, and easy to compute.

Index Terms— Medical control system, Performance evaluation, Drug delivery

I. INTRODUCTION

One of the objectives of clinical anesthesia is to control the consciousness level of the patient. This is typically achieved by the administration of hypnotic drugs, such as propofol. It has been shown that spectral properties of the electroencephalogram (EEG) correlate well with the depth of hypnosis, referred to as the DOH in this paper [1]. This discovery has led to the introduction of clinical devices providing surrogate measures of the DOH. The most widely used such device is the BIS monitor [1]. It, and similar products such as the NeuroSense monitor [2], provide uniformly sampled estimates, with a sample rate that is high compared to the time scale of the involved dynamics.

The availability of real-time DOH estimates has resulted in several simulation studies and clinical trials in which drug dosage is determined online by means of feedback from clinical monitors. An introduction to closed-loop controlled anesthesia systems is available in the reviews [3], [4] and [5]. It was concluded in [6] and [7] that closed-loop strategies outperforms manual dosing.

The clinical objective is initially to transition the patient from the conscious state to an adequate DOH. This corresponds to a temporal phase termed *induction* of anesthesia. Once completed, surgery can commence during the subsequent *maintenance* phase of anesthesia. Towards the end of the procedure anesthetic drug administration is halted, marking the start of the *emergence* phase of anesthesia.

Kristian Soltesz is with the LCCC Linnaeus Center and the eLLIIT Excellence Center at the Department of Automatic Control, Lund University, Lund, Sweden `kristian@control.lth.se`

Guy A. Dumont is with the Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, Canada `guyd@ece.ubc.ca`

J. Mark Ansermino is with the Department of Anesthesiology Pharmacology and Therapeutics, University of British Columbia, Vancouver, Canada `anserminos@yahoo.ca`
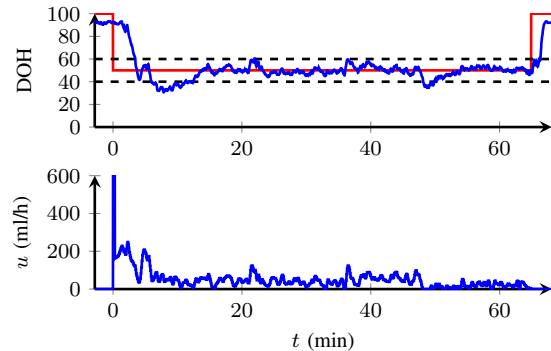
Fig. 1. Representative data from a closed-loop DOH control procedure. The DOH measurement (top, blue) and corresponding setpoint (red) are shown together with propofol infusion rate (bottom).

Figure 1 shows DOH measurements and their underlying setpoint values from a closed-loop study[1] using the NeuroSense monitor and propofol as the hypnotic drug. The DOH value of 100 corresponds to the fully conscious state, while 0 represents the fully anesthetized state. The BIS and other commercially available DOH monitors make use of the aforementioned scale[2], which is consequently used throughout this paper.

The question addressed in this paper is how to measure performance of a closed-loop DOH control system, based on one or several clinical data sets as the one in Figure 1. In fact, the discussion is not limited to closed-loop dosing, but valid for any dosing strategy, e.g. manual control, guided by a DOH monitor. The current practice for such performance assessment is the computation of four measures introduced by Varvel et al. [9]. These *Varvel* measures were proposed to evaluate the performance of anesthesia systems, however not the EEG-guided variety described above. Despite this, they have become the gold standard in clinical evaluation as evident from several studies in which DOH measures derived from the EEG were used to control infusion rate [10]– [8]. Each of the cited studies represents unique research groups world-wide and the list could be further expanded by including simulation studies and studies in which drugs other than propofol were used in the closed-loop setting.

This paper is organized as follows: Desirable properties of adequate performance measures for EEG-guided DOH

---

[1]The data is from a closed-loop controlled anesthesia study approved by the UBC Children's and Women's Research Ethics Board (H10-01174), Vancouver, Canada [8].

[2]The DOH scale can appear confusing in that a low value corresponds to a deeper DOH and vice versa.

(a) Closed-loop dosing regimen.
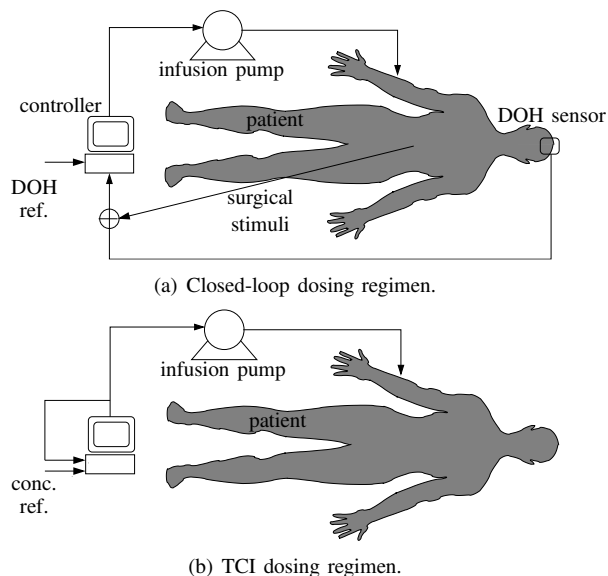

(b) TCI dosing regimen.

Fig. 2. Two control paradigms to determine drug dosing in anesthesia.

control are proposed in Section II. The Varvel performance measures and their background are introduced and evaluated against these objectives in Section III. A revised set of performance measures are proposed and discussed in Section IV. Section V provides a summary and suggested future work is outlines in Section VI.

II. OBJECTIVES

A. The General Idea

The rationale behind introducing performance measures is to map the n-dimensional DOH error[3] vector (if n samples are available) to an m-dimensional ($m \ll n$) space in which each dimension has a distinct interpretation, strongly coupled to the performance of the system. In the case of the Varvel measures, $m = 4$. Ultimately, it would be desirable to set $m = 1$ and describe the performance of the system with only one scalar. This was attempted in [6] by introducing the *global score*, discussed in Section III-D.

In the context of EEG-guided DOH control, the foremost concern is the clinical outcome. The challenge therefore becomes to find properties of the DOH error, which are strongly linked to the clinical outcome.

B. Temporal Phases of Anesthesia

The control objectives differ between the temporal phases of anesthesia, as outlined below. This suggests different performance measures for each temporal phase and calls for a systematic method to determine the transitions between subsequent phases.

1) Induction of Anesthesia: The clinical objective during induction of anesthesia is to make a *fast transition* to the DOH setpoint, with *limited overshoot* and *short settling time* [18]. This is complicated by large inter-patient variability

in drug sensitivity. However, the disturbances acting on the system are typically limited [19]. Limiting the induction phase duration is particularly motivated for anxious patients. A more rapid induction of anesthesia may also alleviate the discomfort associated with the infusion of propofol in conscious patients. Furthermore, limiting the duration of the induction phase results in an increased availability of the operating room and its staff. Meanwhile, limiting the overshoot is critical during maintenance of procedures which require spontaneous breathing, since a small DOH value generally results in apnea. To avoid overshoot is also critical in some patient groups such as the elderly, in order to avoid the associated potential for hypotension due to overdosing.

2) Maintenance of Anesthesia: During maintenance of anesthesia, it has been recommended that the DOH should lie in the 40–60 range, with a setpoint of 50 [20], [21]. The two main challenges in meeting this recommendation are large inter-patient variability in hypnotic drug sensitivity and output disturbances, introduced foremost by surgical stimulation.

Too large a DOH value could result in awareness with recall [22], which is likely to cause the patient considerable psychological stress, while excessively small DOH values may result in hypotension and an increase in long term mortality [23]. Variability in the DOH, which induces significant changes in blood pressure, are potentially harmful for the patient. In order to maintain hemodynamic stability, it is therefore of interest to quantify such variability.

3) Emergence from Anesthesia: It is desirable to minimize the *emergence duration* for the same reason as the induction duration should be kept short. Drug administration is halted during the emergence phase of anesthesia, but generally accepted models suggest that dosing history from the maintenance phase (and possibly also the induction phase) of anesthesia influence the emergence phase duration [24].

C. Pre-filtering and Artifact Removal

EEG-based DOH measurements are sensitive to electromagnetic noise at the sensor. The raw DOH output of existing monitors therefore contains significant noise power beyond the bandwidth of the controlled system. It is common practice to (low-pass) filter the signal prior to using it for control.

The approach taken in this paper is to assume that the DOH signal used for performance evaluation has been subject to efficient artifact removal and filtered such that measurement noise is removed, without significantly affecting clinically relevant characteristics. A fair comparison between systems is only possible if these assumptions are met, and doing so is the responsibility of the control engineer.

III. THE VARVEL PERFORMANCE MEASURES

A. Clinical Background

The Varvel measures were introduced to assess the estimation performance of target controlled infusion (TCI)[4]

---

[3]The DOH error is defined as the difference between DOH setpoint and measurement.

[4]In [9] the term computer-controlled infusion pump (CCIP) is used analogously with TCI.

(a) Systematic small error.
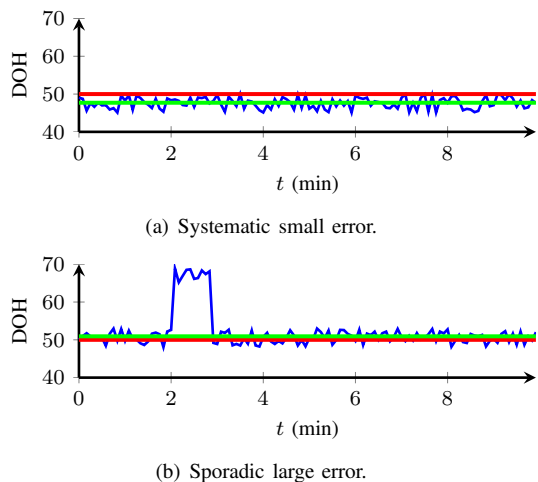

(b) Sporadic large error.

Fig. 3. Two tentative maintenance phase DOH profiles (blue), with corresponding setpoint (red) and resulting median error (green).

systems [25]. The schematic drawing of a TCI system is shown in Figure 2(b). The drug infusion rate is updated based on the output of an open-loop estimator, referred to as a pharmacokinetic (PK) model [24], relating drug infusion rate to *estimated* plasma concentration[5]. Being open-loop, the TCI control scheme is sensitive to disturbances and model errors.

The objective of the Varvel measures is to assess the fit between measured (blood sample) and estimated (PK model output) plasma concentrations in the above setting. This can be conducted either for an individual or a population. The individual case will be considered first.

B. Statistical Background

The Varvel measures are based on the *median* of the relative error and its modulus, respectively. The original motivation for use of the median rather than the mean were the asymmetric appearance of the data sets available to the authors of [9]; most PK estimation errors were close to the median, with few but distant outliers. This motivation, justified in the intended context, results in an unfortunate choice when adopted for EEG-guided DOH control. As mentioned in Section II-B.2, DOH values in the 40–60 range are considered adequate. This suggests that a maintenance phase like the one in Figure 3(a) should be desirable to the one in Figure 3(b), while basing the error metric on the *median* has the opposite outcome. It can be argued that median-based measures provide inherent filtering and artifact removal as outlined in Section II-C. However, they make no distinction between artifacts, noise and momentary large errors which indeed reflect the DOH of the patient.

C. Definitions

The Varvel performance measures are $MDPE$, $MDAPE$, $DIVERGENCE$ and $WOBBLE$, defined,

---

[5]Some TCI systems target the effect site of the drug (the human brain) rather than the blood plasma, but the principle remains unchanged.

explained and discussed below. The sample vector of measured plasma concentrations is denoted C_m and the corresponding estimates are in C_p. The times corresponding to entries of C_m and C_p are in t (unit: h). All vectors have N elements.

The (relative) percentage performance error PE is defined as:

$$PE = 100\frac{C_m - C_p}{C_p}. \tag{1}$$

The division in (1) is to be interpreted as element-wise, making PE a vector quantity. PE is not reported as one of the Varvel measures, but rather used as the basis for computing them.

Applying the definition (1) to DOH measurements results in less error penalty whenever the DOH setpoint is closer to the fully awake state. However, as indicated in Section II-B.2, the risk of the patient being aware requires particularly tight control in this region.

1) Median Performance Error (MDPE): is the median value of the PE between all samples:

$$MDPE = \text{median}(PE). \tag{2}$$

As such it captures the bias of the estimator, but not the variability.

2) Median Absolute Performance Error (MDAPE): As opposed to $MDPE$, the $MDAPE$ does not convey any information on the bias of the estimator, but rather yields a representative error magnitude.

$$MDAPE = \text{median}(|PE|) \tag{3}$$

3) DIVERGENCE: is the slope of a linear regression of $|PE|$ against t:

$$DIVERGENCE = \frac{t^T|PE| - N\bar{t}\overline{|PE|}}{t^Tt - N\bar{t}^2}, \tag{4}$$

where the bars denote the mean operator. The unit of $DIVERGENCE$ is %/h and it aims at describing whether the error increases or decreases over time. An unstable control system will result in a positive $DIVERGENCE$. Beyond concluding instability, the $DIVERGENCE$ is of little practical use. Three different DOH profiles are plotted in solid together with a common setpoint in Figure 4. The respective $DIVERGENCE$s of these profiles are -15, 0 and 15 %/h, proportional to the slopes of the dashed lines. It is of arguable clinical significance when during the maintenance phase the error spike occurs, yet its temporal location strongly affects the $DIVERGENCE$, while leaving all other Varvel measures unaffected.

4) WOBBLE: was introduced to capture variability of the estimator. It is defined as the median absolute deviation between PE and $MDPE$:

$$WOBBLE = \text{median}(|PE - MDPE|). \tag{5}$$

$WOBBLE$ measures variability in the DOH. As such it is strongly affected by filtering, as discussed in Section II-C. In order to compare systems in terms of $WOBBLE$ it is
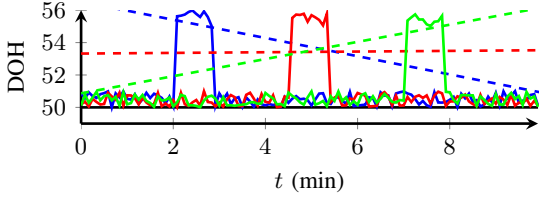
Fig. 4. The solid lines show tentative DOH profiles, with the horizontal black line as setpoint. The divergence for each of these profiles is the slope of the correspondingly colored dashed line.

Fig. 5. Graphical illustration of induction duration ID (green) and overshoot OS (magenta), using the clinical data set from Figure 1.

therefore essential that the same filtering be used. The BIS monitor uses a proprietary filtering algorithm and switches between several filters [26], making it challenging to meet this requirement.

D. The Global Score (GS)

The Global Score (GS) is not one of the Varvel measures, but was introduced in [6] as an attempt to score EEG-guided DOH control systems with one scalar:

$$GS = \frac{MDAPE + WOBBLE}{\text{fraction of time DOH} \in (40, 60)}. \quad (6)$$

Apart from characterizing performance by means of a single scalar, the GS takes the clinical feasibility bounds from [20] into account. Expanding the numerator of (6) results in

$$100(\text{median}|PE| + \text{median}|PE - \text{median}(PE)|). \quad (7)$$

While the idea of one scalar performance measure is appealing, it is hard to intuitively interpret (7) and the clinical relevance of the GS has never been established.

E. Population Measures

It was concluded in [9] that the population distribution of individually computed $MDPE$s and $DIVERGENCE$s are quite symmetric, while the corresponding $MDAPE$s and $WOBBLE$s are slightly asymmetric. Consequently, different approaches for combining the individual measures to correspond to populations were discussed. The two-stage approach defines the population measures as the mean values[6] taken over the population. A modification of two-stage approach is the pooled-data approach, where each individual is weighted by the reciprocal number of available samples. This is further elaborated in the variance-weighted approach, in which individuals are weighted by the reciprocals of the variances of the measures. Details of these pooling strategies are found in [9], in which the authors recommend the variance-weighted approach but conclude that they all yield similar results.

---

[6]Pooled measures for standard error (SE) of $MDPE$, $MDAPE$, $DIVERGENCE$ and $WOBBLE$ are also provided, but omitted from this discussion as they are not commonly computed in publications evaluating closed-loop controlled anesthesia systems.
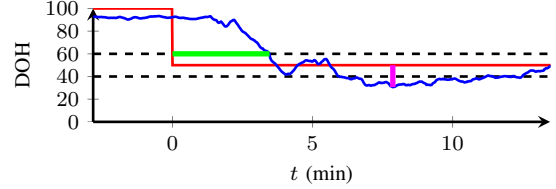
IV. PROPOSED PERFORMANCE MEASURES

This section outlines a set of performance measures more adequate for the evaluation of EEG-guided DOH control than the Varvel measures. As mentioned in Section II-B, it is relevant to use different measures of performance during the induction, maintenance and emergence phases of anesthesia. In fact, most closed-loop studies reporting the Varvel measures, do so for maintenance phase data only.

Here it is assumed that a vector Y, holding N DOH samples is available, together with the corresponding setpoint vector R. The (not necessarily uniformly sampled) time stamps of entries in Y and R are stored in T (unit: s). The vectors are assumed to be adjusted so that $T_1 = 0$ corresponds to the start of hypnotic drug administration. To simplify notation, the sections of $\{T, R, Y\}$ corresponding to the temporal phase of anesthesia being treated will be denoted $\{t, r, y\}$, each holding n elements.

A. Induction Phase Measures

Two performance measures, reflecting the objectives in Section II, are proposed.

1) Induction Phase Duration (ID): The definition of the induction phase duration (ID) proposed here is adopted from [6], where it was defined as the time elapsed from the start of hypnotic drug administration to the moment when the DOH falls to and remains under 60 for 30 s. This definition does not punish large drug boluses which rapidly achieve the desired goal but result in an excessive overshoot. Furthermore, it only considers the DOH set-point of 50. To account for this, staying below 60 is replaced by staying within $r \pm 10$ and an additional measure for characterizing the overshoot is introduced.

2) Overshoot (OS): The percentual overshoot (OS) is defined as

$$OS = 100 \cdot \min_k \frac{r_k - y_k}{E_0 - r_k}, \quad (8)$$

It is common that the maximum overshoot occurs after the end of induction, as defined in Section IV-A.1. It was consequently elected to extend the evaluation of (8) to include a subsequent 10 min period. The definitions of ID and OS are illustrated in Figure 5, using the clinical data set shown in Figure 1. The DOH in absence of hypnotic drug is E_0, typically $90 < E_0 \le 100$. If the E_0 has been identified for the particular patient, the obtained value could be used in favor of the default $E_0 = 100$.

B. Maintenance Phase Measures

1) Integrated Error (IE): The trapezoidal approximation of the error time-integral is introduced to replace the $MDPE$:

$$IE = \sum \frac{t_{k+1} - t_k}{t_n - t_1} \frac{(r_{k+1} - y_{k+1}) + (r_k - y_k)}{2}. \quad (9)$$

The IE is normalized with respect to the maintenance phase duration. As opposed to the median, the IE punishes outliers (linearly). Furthermore, it is used as minimization criterion in existing controller synthesis strategies, of which [27] presents one example within PID control.

2) Integrated Absolute Error (IAE): Since the IE only conveys information about the average error, its sole use as performance measure can be misleading; an error distribution which is balanced with respect to zero results in a small IE. A measure replacing the $MDAPE$ is therefore needed and readily obtained by taking the modulus of the sample-wise error in (9), yielding the integrated absolute error, IAE:

$$IE = \sum \frac{t_{k+1} - t_k}{t_n - t_1} \frac{|r_{k+1} - y_{k+1}| + |r_k - y_k|}{2}. \quad (10)$$

3) Variability Index (VI): Variability in the DOH can be quantified by the relative difference between the IAE and the IE:

$$VI = \frac{IAE - IE}{IAE}. \quad (11)$$

The resulting variability index (VI) does not need to be explicitly reported as it is readily computable from the reported IE and IAE.

4) Percentage of Time Outside Adequate Range: Based on the discussion in Section II-B.2, it is of interest to report how well the system manages to keep the DOH within the clinically feasible range, i.e., within 10 units from the setpoint. Since the sign of the error is of clinical significance, it is justified to give separate measures for the percentage of time during maintenance that the DOH error $(r - y)$ exceeds +10 (E^+) and -10 (E^-), respectively. In case of sparse or nonuniform sampling, linear interpolation between consecutive samples can be used to determine the time instants when the maintenance phase control error crosses ± 10, respectively.

C. Emergence Phase Measure

1) Emergence Phase Rise Time (ER): The emergence phase of anesthesia is defined to begin when administration of the hypnotic drug is terminated. The duration of the emergence phase can be characterized by the 63 % (or $1 - e^{-1}$) rise time, commonly reported in other control applications. This time is defined as that between end of hypnotic drug administration, at which instance the DOH setpoint was r_1 and the frist time at which the (adequately filtered) DOH y_k exceeds $r_1 + (1 - e^{-1})(E_0 - r_1)$. If the awake baseline level E_0 introduced in Section IV-A.2 is not known for the considered patients, the default $E_0 = 100$ may be used. The ER is illustrated in Figure 6, using the data from Figure 1.
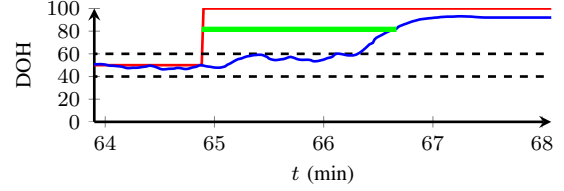

Fig. 6. The 63 % emergence phase rise time ER (green) for the data set in Figure 1.
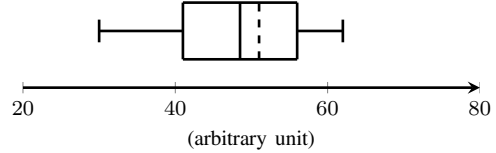

Fig. 7. This modified box plot shows an example of how the worst case (whisker), standard deviation (box), mean (solid line) and median (dashed line) of ID, OS, IE, IAE, (VI), $E^{\pm}$ and ER could be reported in a comprehensive manner.

D. Population Measures

The previous section proposed a set of performance measures aimed at EEG-guided DOH control. Having evaluated these measures for each individual in a study population, the question remains how to statistically combine these into the corresponding measures for the entire population. As reviewed in Section III-E, several options were presented in [9]. Since all the performance measures introduced in Section III-E are unaffected by the duration (in terms of time or samples) of the underlying data, it becomes natural to weigh each case equally in the population statistic.

Which statistic to use depends on the clinical aim. The *median* of each measure is adequate if the aim is to evaluate a system which should perform well in the majority of cases, but where a few cases of poorer performance are acceptable. At the other end of the scale, the *worst case* of each measure could be reported. In terms of conservatism, the *mean* lies between these two.

In order to produce comparability between studies it is suggested that at least the population *mean±standard deviation* of ID, OS, IE, IAE, (VI), $E^{\pm}$ and ER be reported. A convenient way to report these statistics together with worst case and median is the use of the modified boxplot shown in Figure 7.

V. SUMMARY

It has been argued that the Varvel measures are poorly suited for evaluating EEG-guided DOH control systems and a set of performance measures has been proposed. These measures characterize the induction (ID, OS), maintenance (IE, IAE, (VI), $E^{\pm}$) and emergence (ER) phases of anesthesia. Furthermore, a comprehensive representation (modified box plot) of population statistics was proposed.

An enhanced performance assessment could be obtained by including information from the drug administration rate

and other signals routinely logged in the operating room, e.g. blood pressure. This aspect lies outside the scope of this paper but is discussed further in Section VI.

VI. FUTURE WORK

Neither the Varvel, nor the proposed measures, take the control signal (hypnotic drug dose) into account. In case of a closed-loop controlled system, episodes of zero dose indicate possible previous overdosing. Furthermore, poor controller robustness and measurement noise sensitivity are typically visible in the control signal. A rapidly varying control signal also adds unecessary wear to the actuator of the system.

It was concluded in [28] that a simultaneous "triple low" combination of low DOH value, blood pressure, and drug dose[7] significantly increased mortality. As all three mentioned signals are typically available, it should be feasible to assemble a *triple low* measure.

As further future work, it would be of value to conduct a blinded study in which clinicians are asked to rate recorded DOH profiles, in order to compare these ratings with the proposed measures. This would serve to enforce the validity of the proposed measures and possibly suggest modifications of the underlying heuristics. Such a study might in other words refine the definitions of the proposed performance measures, while maintaining their qualitative purposes.

REFERENCES

[1] J. Liu, H. Singh, and P. White, "Electroencephalographic bispectral index correlates with intraoperative recall and depth of propofol-induced sedation," *Anesthesia and Analgesia*, vol. 84, no. 1, pp. 185 – 189, January 1997.

[2] T. Zikov, S. Bibian, G. Dumont, M. Huzmezan, and C. Ries, "Quantifying cortical activity during general anesthesia using wavelet analysis," *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 4, pp. 617 – 632, April 2006.

[3] R. Chilcot, "A review of the control of depth of anaesthesia," *Transactions of the institute of measurement and control*, vol. 2, pp. 38–45, January 1980.

[4] G. A. Dumont, "Closed-loop control of anesthesia - a review," in *8th IFAC Symposium on Biological and Medical Systems*, Budapest, Hungary, August 2012.

[5] S. Bibian, C. R. Ries, M. Huzmezan, and G. Dumont, "Introduction to automated drug delivery in clinical anesthesia," *European Journal of Control*, vol. 11, no. 6, pp. 535–557, 2005.

[6] N. Liu, T. Chazot, A. Genty, A. Landais, A. Restoux, K. McGee, P.-A. Laloe, B. Trillat, L. Barvais, and M. Fischler, "Titration of propofol for anesthetic induction and maintenance guided by the bispectral index: Closed-loop versus manual control: A prospective, randomized, multicenter study," *Anesthesiology*, vol. 104, no. 4, pp. 686–695, April 2006.

[7] N. Liu, T. Chazot, S. Hamada, A. Landais, N. Boichut, C. Dussaussoy, B. Trillat, L. Beydon, E. Samain, D. I. Sessler, and M. Fischler, "Closed-loop coadministration of propofol and remifentanil guided by bispectral index: a randomized multicenter study," *Anesthesia & Analgesia*, vol. 112, no. 3, pp. 546–557, March 2011.

[8] K. van Heusden, G. A. Dumont, K. Soltesz, C. L. Petersen, A. Umedaly, N. West, and J. M. Ansermino, "Clinical evaluation of robust PID control of propofol anesthesia in children," *IEEE Transactions on Control System Technology*, 2013, accepted.

[9] J. R. Varvel, D. L. Donoho, and S. L. Shafer, "Measuring the predictive performance of computer-controlled infusion pumps," *Journal of Pharmacokinetics and Biopharmaceutics*, vol. 20, no. 1, pp. 63–94, February 1992.

[10] A. Absalom, N. Sutcliffe, and G. Kenny, "Closed-loop control of anesthesia using bispectral index," *Anesthesiology*, vol. 96, no. 1, pp. 67–73, January 2002.

[11] G. D. Puri, B. Kumar, and J. Aveek, "Closed-loop anaesthesia delivery system (CLADS) using bispectral index: a performance assessment study," *Anaesthesia and intensive care*, vol. 35, pp. 357–362, June 2007.

[12] Y. Sawaguchi, E. Furutani, G. Shirakami, M. Araki, and K. Fukuda, "A model-predictive hypnosis control system under total intravenous anesthesia," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 3, pp. 874 – 887, February 2008.

[13] T. D. Smet, M. M. R. F. Struys, M. M. Neckebroek, K. V. den Hauwe, S. Bonte, and E. Mortier, "The accuracy and clinical feasibility of a new bayesian-based closed-loop control system for propofol administration using the bispectral index as a controlled variable," *Anesthesia and Analgesia*, vol. 107, no. 4, pp. 1200 – 1210, October 2008.

[14] T. M. Hemmerling, S. Charabati, C. Zaouter, C. Minardi, and P. A. Mathieu, "A randomized controlled trial demonstrates that a novel closed-loop propofol system performs better hypnosis control than manual administration," *Canadian Journal of Anesthesia*, vol. 57, pp. 725–735, August 2010.

[15] W. M. Haddad, K. Y. Volyanskyy, J. M. Bailey, and J. J. Im, "Neuroadaptive output feedback control for automated anesthesia with noisy EEG measurements," *IEEE Transactions on Control System Technology*, vol. 19, pp. 311–326, March 2011.

[16] M. Janda, O. Simanski, J. Bajorat, B. Pohl, G. F. E. Noeldge-Schomburg, and R. Hofmockel, "Clinical evaluation of a simultaneous closed-loop anaesthesia control system for depth of anaesthesia and neuromuscular blockade," *Anaesthesia*, vol. 66, pp. 1112–1120, December 2011.

[17] N. Liu, M. Le Guen, F. Benabbes-Lambert, T. Chazot, B. Trillat, D. I. Sessler, and M. Fischler, "Feasibility of closed-loop titration of propofol and remifentanil guided by the spectral M-entropy monitor," *Anesthesiology*, vol. 116, no. 2, pp. 286–295, February 2012.

[18] S. Bibian, "Automation in clinical anesthesia," Ph.D. dissertation, University of British Columbia, Vancouver, Canada, July 2006.

[19] K. van Heusden, J. M. Ansermino, K. Soltesz, S. Khosravi, N. West, and G. A. Dumont, "Quantification of the variability in response to propofol administration in children," *IEEE Transactions on Biomedical Engineering*, 2013, accepted.

[20] S. D. Kelly, *Monitoring level of consciousness during anesthesia and sedation*, Aspect Medical Systems Inc., Newton, USA, 2003.

[21] G. Agrawal, S. Bibian, and T. Zikov, "Recommended clinical range for wav$_{CNS}$."

[22] W. H. D. Liu, T. A. S. Thorp, S. G. Graham, and A. R. Aitkenhead, "Incidence of awareness with recall during general anaesthesia," *Anaesthesia*, vol. 46, no. 6, pp. 435–437, June 1991.

[23] T. G. Monk, V. Saini, B. C. Weldon, and J. C. Sigl, "Anesthetic management and one-year mortality after noncardiac surgery," *Anesthesia & Analgesia*, vol. 100, no. 1, pp. 4–10, January 2005.

[24] H. Derendorf and B. Meibohm, "Modeling of pharmacokinetic/pharmacodynamic (pk/pd) relationships: Concepts and perspectives," *Pharmaceutical Research*, vol. 16, pp. 176–185, February 1999.

[25] A. Absalom and M. M. R. F. Struys, *An overview of TCI & TIVA*, 2nd ed. Academia Press, 2007.

[26] C. M. Ionescu, R. Hodrea, and R. de Keyser, "Variable time-delay estimation for anesthesia control during intensive vare," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 2, pp. 363–369, February 2011.

[27] O. Garpinger and T. Hägglund, "A software tool for robust PID design," in *Proc. 17th IFAC World Congress, Seoul, Korea*, July 2008.

[28] D. I. Sessler, J. C. Sigl, S. D. Kelley, N. G. Chamoun, P. J. Manberg, L. Saager, A. Kurz, and S. Freenwald, "Hospital stay and mortality are increased in patients having a "triple low" of low blood pressure, low bispectral index, and low minimum alveolar concentration of volatile anesthesia," *Anesthesiology*, vol. 116, pp. 1195–1203, June 2012.

---

[7]The results in [28] are restricted to volatile anesthetics, but it is likely that they carry over to e.g. TIVA with propofol.