Author’s response to reviews

Title: The use of a machine-learning algorithm that predicts hypotension during surgery in combination with personalized treatment guidance: study protocol for a randomized controlled clinical trial

Authors:

Marije Wijnberge (m.wijnberge@amc.uva.nl)

Jimmy Schenk (j.schenk@amc.uva.nl)

L.E. Terwindt (l.e.terwindt@amc.uva.nl)

M. Mulder (marijnmulder@live.nl)

M.W. Hollmann (m.w.hollmann@amc.uva.nl)

A.P. Vlaar (a.p.vlaar@amc.uva.nl)

D.P. Veelo (d.p.veelo@amc.uva.nl)

B.F. Geerts (b.f.geerts@amc.uva.nl)

Version: 1 Date: 30 Jul 2019

Author’s response to reviews:

Reviewer reports:

Reviewer #2: The protocol is comprehensive with no clear areas missing. I do though have some comments below, which I believe would aid clarity for the reader.

Authors: Thank you for your thorough review of our trial protocol manuscript. We incorporated your comments.

Reviewer #2: The number of patients to be recruited for Phase A is not clear. As it states that overall 100 patients are to be recruited and the sample size calculation show that 30 patients are needed for each arm, I am assuming 40 patients are recruited to study any potential Hawthorne effect.
Authors: You are correct. In the literature there is little known about a normal baseline time weighted average (TWA) in hypotension. Therefore, we decided to include a baseline group to; I - be able to calculate the TWA spent in hypotension in our academic center and use this to recalculate and confirm (or expand) our sample size for phase B and; II - to test whether our control group is a good representative sample of our study population by comparing the control group to the baseline group.

We altered the manuscript to make this more apparent. See pages:

Page 3. (Abstract): ‘This study is divided into two parts; in phase A baseline TWA data from 40 patients will be collected prospectively. A device (HemoSphere) with HPI software will be connected but fully covered. Phase B is designed as a single-center randomized controlled trial in which 60 patients will be randomized.’

Page 6. (Methods): ‘Phase A consists of prospective data collection in 40 patients to gain insight in the normal TWA in our study population. Phase A data are collected to check our sample size for phase B and to verify if the control group is a representative sample. Phase B is a single-center randomized controlled (1:1) superiority trial including 60 patients.’

Reviewer #2: - Will this design actually be able to successfully study the Hawthorne effect? My understanding from reading the manuscript is that these patients and the clinical team, even though they don't know the purpose, will still know the patient is on a research study. This means that the patient may well receive additional observations etc, which may contribute to improved outcomes. Also will the sample size be enough to study this with any degree of certainty? There is a lack of information generally about this phase of the study.

Authors: Thank you for your comment. We chose for this study type because data on normal baseline ‘TWA’ in current literature is sparse. We decided to include a baseline group to; I - be able to calculate the time weighted average spent in hypotension in our academic center and use this to recalculate and confirm (or expand) our sample size for phase B and; II - to test whether our control group is a good representative sample of our study population by comparing the control group to the baseline group. We called the latter a ‘Hawthorne’ effect but you are correct, our study design is not primary aimed to study this effect. We powered our study on our primary
endpoint (TWA in hypotension difference between intervention and control group) and will use phase A to verify whether our first sample size analysis was accurate.

We have to be present during phase A for multiple reasons; I - we have to connect the device and; II - we aim to collect ‘normal baseline treatment data’ including treatment drug choice, treatment dose and time to treatment. We are not able to extract these things (accurately) from our electronic patient database (EPD).

We have altered our manuscript to clarify the purpose of phase A:

Page 7 (Methods): These data will be used to verify our sample size calculation for phase B and to study whether our control group is representative for the study population by comparing the baseline group versus the control group. During this phase of the study the treating anesthesiologist and anesthesia nurse will not be informed about the aim of the study or endpoints measured.

Page 12 (Methods): The baseline data collection enables us to calculate the normal TWA spent in hypotension in our hospital and will be used to verify our sample size analysis.

Page 13 (Methods): The baseline data collection enables us to calculate the normal time weighted average spent in hypotension in our hospital and will be used to verify the representativeness of our control group. We will compare the TWA in the baseline group (phase A) to the TWA in the control arm (phase B).

Reviewer #2: - How are the authors adjusting for important baseline covariates that may be associated with the primary outcome (e.g. baseline blood pressure)? Is the randomisation stratified by any baseline variables?

Authors: We did not stratify the randomization. Study subjects will be randomly assigned to either the intervention or control arm. As published in the CONSORT 2010 statement, random
assignment prevents selection bias, and any differences in baseline characteristics are the result of chance rather than bias.1,2

References:


Reviewer #2: - When reporting the sample size calculation, it is useful for context to report the expected control group TWA.

Authors: Based on previous data we decided a TWA change of 0.38 to be clinically relevant. We expected our baseline TWA to be 0.60.

We have added this to page 12: Sample size calculation (phase B)

Difference in primary outcome will be compared using the T-test or the Mann-Whitney U test, based on normality. A statistician performed the sample size analysis. Based on previously published blood pressure data during surgery it was estimated that our control group would have a TWA of 0.60 and a difference of 0.38 or larger between two study groups will be clinically relevant. An effect size of 0.74 has been calculated by dividing the estimated difference of 0.38 (mean experimental group – mean control group) by the standard deviation of 0.51. A sample size of 30 in each group in the randomized phase will have 80% power to detect an effect size of 0.74 using a two-group t-test with a 0.05 two-sided significance level. Sample size has been calculated using R 2017.

Reviewer #2: - The secondary outcomes seem to be missing incidence and time spent in hypotension when compared to clinicatrials.gov.
Authors: Thank you again for your thoroughness. Since they are all measures of hypotension we did not report them individually. We have now added them for completeness.

Page 10 (Methods): The secondary outcome measures include incidence of hypotension, time in hypotension, the % time in hypotension and the AUC of a MAP <65 mmHg. The above-mentioned parameters including TWA will also be assessed for hypertension (defined as MAP >100 mmHg for at least 1 minute) and for the HPI alarms. For hypertension and HPI alarm the area above the curve (AAT) will be calculated instead of the AUC, see figure 4. We will assess the treatment behavior of hypotension and HPI. This includes treatment choice (i.e. vasopressor, fluids, inotropes, position changes), treatment dose, time to treatment and feasibility of working with HPI based on the number of protocol violations.

Reviewer #2: - Can AE/SAEs be defined? Is this all adverse events irrespective of causality? There is the potential risk of over-treatment using a predictive algorithm, so it will be important to obtain safety data in relation to the treatment given to treat potential hypotension.

Authors: We will report all adverse events irrespective of causality to our local medical ethical committee. Serious adverse events will be will reported directly, all other events will be reported annually. Two months after the last patient inclusion we will go through all electronic patient files to note all adverse events and serious adverse event.

We cover the potential harm of overtreatment in two ways;

1. Cumulative treatment dose during surgery
2. Hypertension (in TWA, incidence, total time and % time spent in hypertension)

We have clarified this in the manuscript,
Page 11: Safety

All adverse and serious adverse events, irrespective of causality, will be collected and reviewed by the Principal Investigator and reported to the medical ethics committee of the AMC Amsterdam. Adverse events are defined as any undesirable experience occurring to a subject during the study, whether or not considered related to the experimental intervention. All adverse events reported spontaneously by the subject or observed by the investigator or his staff will be recorded.

Serious adverse events are defined as any untoward medical occurrence or effect that: results in death; is life threatening (at the time of the event); requires hospitalization or prolongation of existing inpatients’ hospitalization; results in persistent or significant disability or incapacity; is a congenital anomaly or birth defect; or any other important medical event that did not result in any of the outcomes listed above due to medical or surgical intervention but could have been based upon appropriate judgement by the investigator. An elective hospital admission will not be considered as a serious adverse event.

We cover the potential harm of overtreatment by assessing the cumulative treatment dose during surgery and by assessing the amount of hypertension (in TWA, AUC, incidence, total time and % time spent in hypertension). We will compare the outcomes between the control and intervention group.

Insurance is provided for all participating subjects by the AMC Amsterdam.

Reviewer #2: - I don't find figure 2 (algorithm) straightforward to understand. For example following the arrows start to HPI 50-85% 'Diagnose Cause' then to HPI>85% or MAP<65 'Start treatment'. I don't understand how HPI 50-85% flows into HPI>85%?

Authors: Thank you for your feedback. We want the treating anesthesiologist to open the secondary screen and try to diagnose the underlying cause when HPI is between 50-85%, in order to start treatment within two minutes when the HPI becomes higher then 85%. We added this step to the flowchart (figure 2) for the anesthesiologist to become familiar with the secondary screen and with diagnosing the underlying cause of the impeding hypotension.

All anesthesiologists receive training on how to interpret the flowchart, how to work with the HPI monitor and how to interpret the secondary screen. Therefore, we are confident that the algorithm will be clear to the treating anesthesiologist.
Reviewer #2: There seems to be a typographical error on the acceptance for registration to clinicaltrials.gov. The trial was submitted in November 2017 and accepted in December 2018. The clinicaltrials.gov record states that it was last updated in December 2017, so I assume there is an error in the acceptance date.

Authors: Thank you for your thoroughness. We have altered this date to December 2017. See page 4.