Author's response to reviews

Title: Ophthalmodynamometry for ICP prediction and pilot field test on Mt. Everest

Authors:

Henry W Querfurth (henry_querfurth@brown.edu)
Philip Lieberman (philip_lieberman@brown.edu)
Steve Arms (swarms@microstrain.com)
Steve Mundell (swmundell@microstrain.com)
Michael Bennett (michael.bennett@caritaschristi.org)
Craig van Horne (craig.vanHorne@caritaschristi.org)

Version: 5 Date: 10 September 2010

Author's response to reviews: see over
Dear Drs. Alam and Serrador:

Thank you for forwarding the reviewer comments on our paper as well as the opportunity to improve it. I also appreciate the recent clarification of my questions before this re-submission. We have addressed every comment from each reviewer in this substantial revision as follows.

Text changes: insertions are highlighted, deletions are not indicated

1) There is a pre-review request from the Associate editor to provide more validation study on controls, specifically looking at day-to-day variability. We did look at this in supplement fig. 2A, in a Bland-Altman plot of variability in control subject A. It consists of 10 repeated measurements that were taken over 3 days (we now mention the 3 days in our revision). In addition, 3 normal volunteers were tested over as many days and the variability (95% CI) turned out to be similar to the inter-observer repeatability. This result is now stated in the text on page 7.

2) Reviewer FS.

   We agree that the linear correlation is not sufficient validation of the vODM technique. However, Fig. 1 is intended as a correlation of an approximate method to predict ICP (vODM) to a maximally accurate or’ true’ one (invasive monitor). This defines it as a calibration issue. Bland-Altman analysis of method comparison studies are more generally applied to evaluate agreements between two indirect, slightly different methods of measuring the same quantity (Bland and Altman Lancet 1986). They are also used in validation repetition studies, as we have done in the supplement Fig. 2A and B. As for ROC curves, these are used to assess the accuracy of a test (TP+TN/P +N) based on a cut-off point. It is commonly used in a binary classification context as the discrimination threshold is varied (Altman BMJ 1994). Our calibration curve is intended for absolute ICP prediction, rather than to convey the accuracy in identifying whether someone has elevated ICP or not.

   Therefore, in order to further address validation we have included 2 new analyses of the data presented in original Fig.1. The first is to additionally plot the
95% confidence limits (new Fig.1A). The next, as requested, is an ROC curve of this data (Fig.1B), only now based on setting a binary discrimination ICP value at 15 mm Hg. As shown, this technique is fairly accurate (area under curve=0.89) to identify those as having elevated vs. normal ICP. This new result is discussed in the legend, statistics and results (pg 10) sections.

This added analysis also adds strength to the findings of our second objective, the pilot study. Since the ROC identified vODM as accurate to discriminate normal from high ICP (>15mm Hg), the finding that predicted ICPs at altitudes of 3445m and higher were well over 15mm, compared to the predicted ICP at sea level (10mm), has extra validity.

The minor concern of the length of the article is well taken. We in fact considered separating the 2 objectives. The problem arose with independent colleagues who as draft reviewers all advised that the validation studies be juxtaposed to the pilot clinical test because interpretation of the latter depended so much on the former. We note the reviewer also acknowledged this inseparable link in the preceding comment (‘knowing the accuracy to interpret the second objective’). Moreover, the medical- neurologic community which stands to gain from further development of this technique are both hospital-based (objective 1, ICU/office patients) and field–based EMTs, paramedics (objective 2, pilot study), further linking these objectives. As reported here for the first time, the data showing that vODM tracks ICP changes between and often within the same patient(s) under pathological and physiological settings (ICU and field, respectively), is another reason to keep the objectives combined.

In deference to this concern, we have shortened the abstract and conclusion sections. We thank this reviewer for these helpful comments.

3) Reviewer TG.

It is pointed out that our group and a German one published ‘very similar’ results and that the replication study herein is ‘unoriginal’ representing only a ‘small improvement’. However, an alternative view can be based on the following major differences. Our previous study only involved 6 patients having heterogeneous diagnoses. The current study measured 12 patients, a more homogeneous group consisting largely of hydrocephalics. More importantly, the previous study was conducted unblinded, it was a pilot. The current study was performed blinded to the actual ICP and thus represents a major advancement in the validation of this technique. Additional statistics were provided in our submission (and newer results, see below) that back up the validation and calibration data. It is important to present them together so that the high altitude findings (which is now the pilot study) can be interpreted with added confidence. We agree the improvements to the instrumentation sound minor but they proved operationally huge in the ICU and harsh environment. It is also essential to understand that the German group used a suction cup ODM machine, altogether a
different device which is likely more uncomfortable to the awake patient and cannot be transported into field.

We completely agree that 95% confidence lines be plotted. They are now shown in Fig.1A. Moreover, we now provide an ROC curve in Fig 1B. It establishes the accuracy of the technique to predict whether a given individual has elevated ICP (cutoff is >15 m Hg) or is normal. The area under the curve is 0.89 and its 95% limits are 0.73 to 1.05, p<.01. This is discussed in the legend, statistics and results (pg 10) sections.

Regarding the comment that small numbers of subjects were measured at most elevations, it should have been made clearer by us that in stating ‘convenience set’ we meant a variable dataset. This study did not involve moving the same group of people from sea level to camp 2. That design proved impractical and conditions were primitive or harsh. We emphasize this phase was a Pilot study and acknowledge these major limitations in our revised conclusion on pg 14. Still, some subjects were each measured at several elevations (Fig.3B). Moreover, if other control subjects at sea level that were outside the study (Supplement Fig.2 and additional ones, see pg.7) are included, that would make n >10 volunteers at sea level, 3445m and 5335m. The upward rise in predicted ICP into the >15mm Hg ‘abnormal’ range is significant at both altitudes. We also agree that measurements performed during descent are desirable and hope to do these in a future study (cited on pg 14).

In deference to the concern that the abstract and conclusion sections are too long, we have tried to shorten them at the same time expanding a bit about the speculative retinal blood flow variable on pg 13. Admittedly, it is an unknown research area. We thank this reviewer for his helpful comments.

Thank you, sincerely yours

Henry Querfurth MD PhD