UNRESOLVED INSTRUMENTATION PROBLEMS FOLLOWING CLINICAL TRIALS USING NEAR INFRARED SPECTROSCOPY

Andrew J. Macnab, Roy E. Gagnon, and Faith A. Gagnon
British Columbia Children’s Hospital, Intensive Care Unit, 4480 Oak Street, Vancouver, British Columbia V6H 3V4, Canada
(Paper JBO-164 received Nov. 5, 1996; revised manuscript received July 10, 1997; accepted for publication July 28, 1998.)

ABSTRACT

Near infrared spectroscopy (NIRS) clinical trials conducted over a seven year period have identified instrument engineering problems related to fiber optic failure, electromagnetic interference, chromophore algorithms, and computational software. These problems have caused confusion amongst clinicians at the bedside, rejection of large volumes of data, repeated reanalysis of data, and a significant diversion of project resources away from clinical studies and into engineering solutions. This article summarizes previously published studies and presents new data which, together, emphasize the need for improvements in NIRS technology. Instrument designers need to be aware of the need for these improvements if NIRS is to serve clinicians better during research designed to rationally define clinical management protocols. © 1998 Society of Photo-Optical Instrumentation Engineers.

Keywords near infrared spectroscopy; algorithms; cytochrome oxidase; blood flow; haemoglobin.

1 INTRODUCTION

When near infrared spectroscopy (NIRS) was introduced to the fields of neonatology and pediatrics, there was a great deal of excitement in the clinical community. One early report stated that, “Near infrared spectrophotometry provides valuable quantitative data at the bedside for the management of sick infants and for exploring the pathophysiology of damage to the brain.” ¹ Subsequently, many clinical studies using the new NIRS technology demonstrated that NIRS had the potential to provide important qualitative and quantitative information.²⁻⁷ However, with further research use of the equipment, it has become apparent that there are unresolved problems to be addressed before NIRS becomes a routine monitoring method in the clinical setting.⁸

Since 1989, British Columbia’s Children’s Hospital has used a NIRO-500 (Hamamatsu Photonics KK, Hamamatsu, Japan) spectrometer in various clinical settings to evaluate patterns of change in the concentration of oxygenated haemoglobin (HbO₂) and deoxygenated haemoglobin (Hb), and patterns of change in the redox status of cytochrome C oxidase (cytochrome a,a₃). Our NIRO-500 uses wavelengths at 777, 828, 849, and 910 nm, and its engineering details and principles have been published previously.⁹ Our studies have used mechanically ventilated, oxygen-dependent neonates, children undergoing heart surgery or cardiopulmonary bypass, and patients in pediatric intensive care recovering from cardiac surgery, head injury, or sepsis. During the period of study there were 12 co-investigators involved in eight studies; all the investigators were initially unfamiliar with the theoretical and practical principles of living tissue near infrared spectroscopy. The engineering support for these studies was limited to the facilities and to the staff generally available in a biomedical engineering department in a major hospital, and it did not include any expertise in photonics.

The expectation of our clinicians, who spend the majority of their time attending patients, was that they would be able to collect and analyze a variety of physiological data without requiring more than a basic knowledge of the technical details of the NIRS hardware and software. Because the spectrometer in use was factory produced, commercially available, and internationally distributed it was generally believed that any major shortcomings in the technology would have been resolved previously in the university/industrial laboratory setting. The degree to which equipment developers would rely upon clinical results and clinician experience to validate and improve the devices was not apparent to the clinicians. Consequently, when technical problems interfered with the expected outcome of the clinical experiments, the long term result in our environment has been a waning of enthusiasm

Address all correspondence to Dr. A. J. Macnab. Tel: (604) 875-2729; Fax: (604) 875-2728.

1083-3668/98/$10.00 © 1998 SPIE
from clinical colleagues once they realized (a) the
time commitment necessary for acquiring the de-
tailed technical background to enable them to trust
their data collections, (b) the complexities of data
analysis and interpretation, and (c) the limited clini-
cal relevance of the technology because of the in-
ability to display real time results at the bedside.
Those reports that have reached the literature tend
to be qualified by the caveat that NIRS continues to
hold great promise rather than that it delivers what
clinicians believe they actually need.

Our clinical NIRS experience has been compli-
cated by the need to address problems of fractured
fiber optic bundles, electromagnetic interference,
correct chromatophore algorithms, and flawed
computation software. Resolution of these issues
has been an ongoing process and is not yet satisfac-
torily concluded. As with other clinical investiga-
tors, we have had to discard substantial amounts of
our data, and are tending to favor animal models
for many of our current studies.

2 FIBER OPTIC FRACTURES

2.1 THE PROBLEM THAT WE ENCOUNTERED

During a series of studies to investigate changes in
cerebral blood flow in preterm neonates in response
to noxious stimuli, our data became increasingly
noisy, and on subsequent analysis proved mean-
less. Our spectrometer was continuing to collect
data even though we discovered that more than
90% of the fibers in the receiving fiberoptic bundle
were fractured, and that light from an external in-
candescent source could not be conducted. We
devised a repair method to affix the terminal interface
and patient optode to the remaining length of intact
cable with an internal/external composite conical
strain relief system fabricated from heat-fit shrink-
able tubing.10 To date, this repaired fiberoptic cable
has been used for more than 70 h of data collection
with satisfactory results. Occasional inspection of
the bundle by illuminated microscope has indicated
that there are no fractured fibers.

2.2 THE PROBLEM THAT REMAINS

More importantly, the issue we addressed, but
were unable to solve, was that of the spectrometer
not having a fault signal to warn the clinician that
the receiving fiberoptic bundle is defective. There is
a paradox to such a signaling system. The spec-
trometer must be capable of warning whenever ins-
sufficient light is being returned because the sam-
ping media are too dense or the interoptode spacing is too great, but at the same time the instru-
ment must not exclude physiologically possible low
intensity signals. Our impression is that the chief
characteristic of light intensity transmitted via a
fractured cable is its rapid variability in signal
strength. However, designing a warning system
based on this characteristic alone could potentially
exclude sampling of tissues in which rapid variabil-
ity was occurring for physiological reasons rather
than being due to noise.

3 ELECTROMAGNETIC INTERFERENCE

3.1 THE PROBLEM THAT WE ENCOUNTERED

In the third year of our clinical studies, we began to
have intermittent faults during data collections
which, because they were few in number, were eas-
ily corrected manually. Over several months, the
number of data misreadings progressively in-
creased to between 10% and 15% of the data collec-
tion per subject. Because of the intermittent nature
of the problem it was not clear whether the spec-
trometer, supporting computer, communication
cable, testing site, test subject, or a combination of
factors was the source of interference. We used the
anechoic chamber and benchtop methods of testing
electromagnetic interference (EMI) for our micro-
gravity qualified somatosensory evoked potential
(SSEP) device (a modified Quantum 84, Cadwell In-
dustries, Kennewick, WA) to diagnose our near in-
frared spectrometer.11 Although we did not do
anechoic chamber testing of the spectrometer or its
supporting laptop computer (Datatrain DPC 3816),
both underwent physical inspection and observa-
tion of their responses under various operational
conditions (including shielding, ferrite bead sam-
ping, cable shortening, gasketing, and proximity
and orientation to ac cables/dc motors/medical in-
struments). Incorporated in these processes were
lessons learned from the anechoic chamber testing
of the SSEP device. There were 23 inspections in
five categories pertaining to enclosures, circuit
boards, interior cables, and exterior cables. The op-
erational response tests searched for EMI suscepti-
bility, as well as for conducted, radiated, and tran-
sient emissions.11

We were unable to determine whether the spec-
trometer or the laptop computer was the cause of
the problem or whether the devices combined
could pass published EMI standards.12 Our solution
was the least expensive: we replaced the laptop. To
date there has been no further interference causing
data misreadings when using the spectrometer with
its new computer (Zenith 80386sx). Nor have there
been any EMI problems between the former laptop
and the ultrasound and SSEP devices which it now
services. The issues of site and patient EMI sources
were ruled out by our testing.

Clinicians are not trained to anticipate EMI prob-
lems in the equipment they use. We believe that the
growing use of consumer market laptop computers
to support these and other biomedical instruments
will lead to an increase in the number of unex-
pected EMI events.

JOURNAL OF BIOMEDICAL OPTICS • OCTOBER 1998 • VOL. 3 NO. 4 387
3.2 THE PROBLEM THAT REMAINS

Hospital biomedical engineering departments are not equipped to measure and rectify such problems. EMI problems occur because instrument designers have the dilemma of whether to package the measurement, conversion, storage, retrieval, comparison, and display components as one complete unit or as individual modules supplied by the purchaser. A disadvantage of the modular approach is that EMI problems between modules cannot be detected beforehand. Although commercial, aerospace, and military EMI standards do exist, there is no guarantee that individual modules meeting these standards will pass EMI testing when linked as a single unit. Also, it cannot be assumed that once a device passes EMI testing it will continue to do so throughout its lifetime, nor can it be assumed that every production copy of the device will pass EMI testing to the same degree. Because real-time EMI sampling does not exist, manufacturing to comply with standards is the only practical quality assurance method available. Therefore, we have proposed that spectrometers and supporting computers intended for the clinical NIRS setting should be tested to a level comparable to aerospace EMI standards.\textsuperscript{11,13} Such standards are more stringent than those for commercial devices but less so than for military hardware. Improved standards will reduce the chance occurrence of EMI, allow clinicians to more readily identify physiological anomalies in the extremely weak biological signals being sampled, and enable clinicians to use NIRS simultaneously with other medical devices with greater assurance that any differences in the collected data are physiological rather than instrumental.

4 CHROMOPHORE ALGORITHMS

4.1 THE PROBLEM THAT WE ENCOUNTERED

The potential for NIRS to provide data on changes in the redox state of cytochrome c oxidase (cyt-aa\textsubscript{3}) is an important consideration for clinicians. Studies of the pattern of change of cytochrome using newly developed spectrometers contained important differences when compared with earlier studies.\textsuperscript{14} Discussion ensued which considered the experimental methodology, technical differences between the devices, and computational differences, adding fuel to the debate over whether any existing clinical spectrometer could in reality accurately measure changes in cytochrome concentration within living tissue.

To resolve the issue of whether a fundamental error existed we decided to compare 13 previously published algorithms used to convert the collected optical density data into chromophore concentrations. We were able to show that these 13 algorithms\textsuperscript{15} all produced identical results from the same data set (0.9996 coefficient of correlation); however, by forcing decimal point and wavelength misidentification errors upon the testing routine the cyt-aa\textsubscript{3} patterns of change could be made to resemble those of the optical densities. This led to the conclusion that a microchip encoding error likely existed in the device used for the earlier studies. Matcher et al.\textsuperscript{16} used simulated optical densities derived from diffusion theory formulas to show that the differences in cyt-aa\textsubscript{3} patterns of change seen between the devices from the different manufacturers could be entirely due to computational errors. When this error occurred, the algorithm used began with published concentration coefficients, which scale the proportions of chromophores present at each wavelength, rather than beginning with each chromophore’s spectral absorption value for each wavelength.\textsuperscript{17} Using absorption coefficients avoids the possibility of repeating any error made in the calculation of the published concentration coefficients.

At the bedside, clinicians need to distinguish between normal and abnormal physiological patterns without concern for hardware/software artifacts.

4.2 THE PROBLEM THAT REMAINS

While an algorithm problem such as that previously described hinders clinical trials and complicates the interpretation of data, it does not prevent studies being done. Since instrument designers cannot know whether such a problem will arise, it is desirable that they provide a capability by which clinicians can retrieve the most primary level of raw data collected. For example, in the case of clinical spectrometers, emitted and received photon counts (or their respective light intensity values) should be given as the root rather than optical density ratios. Of course concentration values/trends should continue to be the first choice of data displayed, but the ability to check unprocessed primary data must also exist. Without such checks clinicians cannot attribute their results to differences in protocol, instrumentation, physiology, analysis, or error.

5 COMPUTATION SOFTWARE

5.1 THE PROBLEM THAT WE ENCOUNTERED

Throughout the course of our clinical trials the operation of the spectrometer and use of its attendant software have not been intuitively obvious to the clinicians using it. To assess whether this was an important problem we limited the training and assistance given to one group of first time users. The main result was that the complexity and length of time required for postprocessing and spreadsheet analysis of data resulted in these clinicians postponing both activities. Their hope was that they could overcome the learning curve associated with the processing methods by doing several analyses together and by scrutinizing a larger pool of data. Consequently they continued to collect data with-
out the assurance that it was valid. Once postprocessing was performed, in many cases these data proved unusable, but it was not possible to distinguish whether the source of the problem was the clinician’s technique or one of the ergonomic, procedural, fiber optic, EMI, or algorithm problems previously described. At the bedside, entire data collection sets were also lost when software routines were terminated incorrectly; the wording of the bootup instructions which require yes/no responses is such that some answered “yes,” where, by the same logic, others answered “no.” Our experience suggests that the technical software required to support NIRS instruments in the clinical setting needs to be developed and tested in close collaboration with the clinicians who will be using it.

The supporting software provided for the NIRO-500 is not optimized for clinical NIRS applications. This is because it does not calculate an average cerebral blood flow (CBF) value, correlation values are not given, data streams are displayed separately, and these data cannot be superimposed or displayed graphically at the maximum screen size. We found it necessary to create our own C++ language coding of trend display, as well as automated CBF analysis routines to display cerebral transit time, correlation analysis, data selection criteria, and flow profiles at the bedside. These four elements, singly and in combination, raised concerns that previous publications describing CBF computations\(^{18-20}\) were flawed; for example, that physiologically improbable CBF values would occur in the data stream within half a second of realistic values.

The CBF formula expresses the ratio between the quantity of HbO\(_2\) carried to the brain by the arterial supply and the concentration of HbO\(_2\) present in the brain’s arterial and venous pathways as measured by NIRS. This formula does not contain any exponents, but the published methodology describes using a fourth order polynomial curve fitting technique to determine first derivative instantaneous CBF values for half second interval cerebral transit periods of less than 30 s duration. This polynomial redundancy is unnecessary from a mathematical point of view but was incorporated to better accommodate the physiology of short term blood flow through the brain.\(^{21}\) When we used this same logic but extended it to include a range of polynomials equal to the possible number of half second intervals in the transit period (typically 2–14), we found that a number of errors occurred, including compiler math overflow during any attempts beyond sixth or seventh order polynomials.

To identify which part of the CBF formula was causing the errors we did an analysis of the least squares curve fitting residuals for the change in cerebral HbO\(_2\) concentration versus elapsed time and of the quantity of HbO\(_2\) delivered to the brain versus elapsed time. In the case of concentration change, the residuals remained static after first order, and for the delivered quantity they were static after second order. This process indicated that the quantity of HbO\(_2\) delivered to the brain during CBF interventions was in fact nonlinear and therefore not correctly implemented by the matrix pseudo-inverse and -derivative processes. We found that the derivative’s required fractional exponents could be avoided by expressing the quantity of HbO\(_2\) delivered to the brain as the numerator, the cerebral concentration change as the denominator (the inverse of that initially published\(^{18-20}\)), and then inverting the value obtained from the subsequent polynomial equation’s first derivative. When this was done, none of the errors previously encountered occurred.

5.2 THE PROBLEM THAT REMAINS

The above findings suggest that there may be errors in NIR publications incorporating the 1989 version of the CBF formulas. It also highlights a major problem in creating fully automated real-time NIRS software. The Fick principle for measuring CBF by NIRS creates, but does not measure, a cerebral transit period. This is the period that begins when the ratio of oxygenated/deoxygenated hemoglobin is suddenly altered and ends when the rate of change of this ratio slows as the blood having the altered HbO\(_2\) ratio begins to depart the NIR scan region. Hence there is a limited region in the data stream within which the CBF formula can be applied. Our attempts to identify the transit period data set have been confounded by signal noise and signal variability. Comparison of results between other technologies, such as positron emission tomography, may not be appropriate since NIRS is the only means of measuring CBF over a brief time course. At present, no technology can claim to be the “gold standard” for measurement of CBF; however, the noninvasive and continuous use attributes of NIRS make it a likely candidate for this status. The need for NIRS to delineate the CBF cerebral transit period indicates a need for the development of clinical NIRS devices that can collect data at intervals considerably shorter than half a second, so that iterative statistical sampling can be used reliably to isolate the end of the transit period.

6 CONCLUSION

The true value of NIRS is its potential to make continuous, noninvasive, near real-time measurements of cerebral oxygenation that can be used in the clinical setting to influence decision making on patient management. In spite of inherent difficulties, NIRS and related technologies are currently being used in a wide range of clinical and laboratory applications. In this way further progress with NIRS...
seems inevitable and it brings closer the application of this technology in a manner which is truly relevant for the clinician.

Acknowledgments
The authors gratefully acknowledge financial support for their NIRS research from the Hospital for Sick Children External Grants Foundation, Toronto, the British Columbia Lung Association, and the British Columbia Health Research Foundation, Vancouver.

REFERENCES
12. ANSI C63.2, American national standard specifications for electronic noise and field strength instrumentation, 10 kHz–1 GHz, The Institute of Electrical and Electronics Inc. (1987).