Implicit and explicit prior information in near-infrared spectral imaging: accuracy, quantification and diagnostic value

By Brian W. Pogue*, Scott C. Davis, Frederic Leblond, Michael A. Mastanduno, Hamid Dehghani and Keith D. Paulsen

Thayer School of Engineering, Dartmouth College, Hanover, NH 03755, USA

Near-infrared spectroscopy (NIRS) of tissue provides quantification of absorbers, scattering and luminescent agents in bulk tissue through the use of measurement data and assumptions. Prior knowledge can be critical about things such as (i) the tissue shape and/or structure, (ii) spectral constituents, (iii) limits on parameters, (iv) demographic or biomarker data, and (v) biophysical models of the temporal signal shapes. A general framework of NIRS imaging with prior information is presented, showing that prior information datasets could be incorporated at any step in the NIRS process, with the general workflow being: (i) data acquisition, (ii) pre-processing, (iii) forward model, (iv) inversion/reconstruction, (v) post-processing, and (vi) interpretation/diagnosis. Most of the development in NIRS has used ad hoc or empirical implementations of prior information such as pre-measured absorber or fluorophore spectra, or tissue shapes as estimated by additional imaging tools. A comprehensive analysis would examine what prior information maximizes the accuracy in recovery and value for medical diagnosis, when implemented at separate stages of the NIRS sequence. Individual applications of prior information can show increases in accuracy or improved ability to estimate biochemical features of tissue, while other approaches may not. Most beneficial inclusion of prior information has been in the inversion/reconstruction process, because it solves the mathematical intractability. However, it is not clear that this is always the most beneficial stage.

Keywords: spectroscopy; molecular; tomography; reconstruction; near-infrared; optical

1. Introduction

Imaging using near-infrared spectroscopy (NIRS) has diversified and expanded into several fields of medical research and science, and pre-packaged biomedical instruments are now commonplace. While the rate at which new instruments have been introduced outpaces their utility, research in this field continues to expand. This expansion is driven by a rapidly increasing pace of discoveries in molecular

*Author for correspondence (brian.w.pogue@dartmouth.edu).

One contribution of 20 to a Theo Murphy Meeting Issue 'Illuminating the future of biomedical optics'.
diagnostic imaging at the pre-clinical stage, with potential to profoundly impact basic biology and clinical medicine. Most implementations of NIRS use a model-based estimation strategy for quantifying tissue parameters based upon spectral features and thus incorporate some form of prior information. For example, NIRS applications include diffuse optical tomography as well as a wide variety of other tissue interrogation approaches where biochemical information is derived using approaches in which spatial tissue sampling is more limited. One aspect that is shared by all NIRS methodologies is that they rely on the acquisition of diffuse near-infrared (NIR) light signals followed by the solution of an inverse problem in the form of either a tomography problem or a simpler type of least-squares problem, such as curve fitting. In recent years, many sources of prior information have been examined, taking several different forms and being analysed quantitatively in terms of their potential to make NIRS amenable to a level of accuracy that is closer to, if not better than, that achievable with concurrent imaging modalities that are providing functional and/or molecular tissue information. The conceptual framework examined in this study is to identify explicit prior information as a multi-parametric factor that can be inserted into the NIRS process at several stages. This is illustrated in the schematic of figure 1. The use of prior information in measuring, processing and interpreting NIRS data is examined with the goal of identifying key examples in which explicit prior information benefits the ability to quantify and/or diagnose.

The general framework illustrated in figure 1 is conceptual and can be used to systematically examine at what stages prior information improves NIRS. The NIRS process consists of the following steps: (i) data acquisition, (ii) calibration and pre-processing of the data, (iii) forward light transport modelling, (iv) inverse problem, (v) post-processing, and (vi) interpretation of the parameters. Not all of these steps are necessarily used, but this generically describes the possible steps, and into each of these steps it is feasible to involve prior information from any number of data streams. These possible streams may consist of (i) spatial structure information obtained from other imaging modalities, (ii) spectral information of the molecules being assayed, (iii) physical constraints on parameters in the model and data, (iv) subject or tissue biomarkers and/or demographic information, and (v) biophysical models to modify or interpret the NIRS process data. These are not fundamentally separate prior information features but rather general categories.

Historically, a majority of NIRS systems rely on spectral prior information as a basis for interpreting the signals in the process of using this for diagnosis. NIRS spectral features measured in most cases are broad, provide a direct transformation to physiological parameters, and the prior measurement of them is possible without any change in system hardware. Because of these characteristics, spectral priors arguably have the most effective impact on NIRS quantification for medical applications. Implementation of spatial prior information as either region definitions, soft priors or shape-based priors has been a steady area of investigation, albeit with niche conclusions for specific examples that provide synergy between NIRS and another imaging system. In the sections below, prior information streams for NIRS are identified, and the details of their implementation and potential to impact NIRS quantification and possible diagnosis are discussed. An attempt is also made to quantify the proven and potential benefits presented by different prior–NIRS convolutions.
Review. Prior information in NIR imaging

The inclusion of prior information of several different types into the NIRS imaging process has been implicit in most applications. The reason for using prior information is to improve the final interpretation of the results, through more accurate quantification, and in medical applications to improve the sensitivity or specificity of the technique. In a process manner, this can be thought of as a prior information, P, being convolved into the NIRS steps, N, resulting in a different interpretation and diagnosis, D. Most of the tests are added either to improve sensitivity or specificity, or to reach a better compromise of both, as illustrated by the receiver operating characteristic (ROC) curve.

2. Prior information: structure

(a) External shape and internal fluence

Without question, the most dominant factors in NIRS measurement accuracy are the tissue shape and source–detector coupling to the tissue [1,2]. The external shape affects the internal light transport distribution and thereby affects the amplitude of the detected light signal at points on the surface. See figure 2a for an illustration of how the shape of the tissue may influence the light distribution. A large number of papers have been published describing logistical approaches to modelling light in tissues of different shapes and source–detector arrangements [3], and it is generally thought that using as accurate a light propagation model as possible is required in order to interpret the data or accurately estimate optical properties through an inversion of the measurements. In this case, the prior information is implicit in the design of the system and the forward modelling. Prior information can be used in the pre-processing of data in order to make the data more accurately match the model. For example, the first committed step for light transport modelling is determining, for each optical measurement, the spatial locations of where light enters tissue and where it is detected. This is achieved by first reconstructing the outer surface of the specimen using, for

Figure 1. The inclusion of prior information of several different types into the NIRS imaging process has been implicit in most applications. The reason for using prior information is to improve the final interpretation of the results, through more accurate quantification, and in medical applications to improve the sensitivity or specificity of the technique. In a process manner, this can be thought of as a prior information, P, being convolved into the NIRS steps, N, resulting in a different interpretation and diagnosis, D. Most of the tests are added either to improve sensitivity or specificity, or to reach a better compromise of both, as illustrated by the receiver operating characteristic (ROC) curve.

2. Prior information: structure

(a) External shape and internal fluence

Without question, the most dominant factors in NIRS measurement accuracy are the tissue shape and source–detector coupling to the tissue [1,2]. The external shape affects the internal light transport distribution and thereby affects the amplitude of the detected light signal at points on the surface. See figure 2a for an illustration of how the shape of the tissue may influence the light distribution. A large number of papers have been published describing logistical approaches to modelling light in tissues of different shapes and source–detector arrangements [3], and it is generally thought that using as accurate a light propagation model as possible is required in order to interpret the data or accurately estimate optical properties through an inversion of the measurements. In this case, the prior information is implicit in the design of the system and the forward modelling. Prior information can be used in the pre-processing of data in order to make the data more accurately match the model. For example, the first committed step for light transport modelling is determining, for each optical measurement, the spatial locations of where light enters tissue and where it is detected. This is achieved by first reconstructing the outer surface of the specimen using, for

Phil. Trans. R. Soc. A (2011)
Figure 2. Major limitations of NIRS imaging. There are biophysical factors and internal tissue factors that limit the capability of NIRS imaging. (a,b) With respect to spatial issues, it is well known that the shape of the tissue changes the internal fluence (a) and must be accounted for in most cases, requiring prior knowledge of the shape of the tissue. The sensitivity profile of the NIR signal is illustrated in (b) and shows how the details of this internal optical structure may be lost due to the blurring and partial volume averaging that dominate the recovery of internal values. (c,d) In terms of biochemical effects, the spectral features of most biological species are broad (c), leading to an increased value with spectral prior information. Finally, at the microenvironment level (d), the NIRS data are often not sufficient to understand the binding or partition status of biochemical species, requiring biomarker or biological modelling prior information. (Online version in colour.)

example, an optical profilometer or by co-registering the coordinate system of the NIRS system with that of an imaging modality providing structural images from which the outer surface of the tissue can be extracted. Then, a light transport modelling approach must be devised and strategies designed for finding numerical or analytical solutions that are subsequently used in the scope of the inverse problem allowing interpretation of the data. Improving the fidelity of the match between model and data through the use of prior information for source and detector positioning generally relies on algorithm designs that combine geometrical hardware considerations for the NIRS system design with co-registration tools allowing registration with the coordinate system of the instrument from which the prior information is obtained. As discussed in more detail in §2b, another important factor in the modelling chain that can significantly affect the match between model and data is the light propagation model itself.

The modelling stage of the NIRS process is critical because it is reasonably well established that model–data mismatch errors can dominate the inversion process and lead to large bias errors [2,4–6]. However, measurement approaches can be used that minimize the need for accurate modelling. Such approaches can be categorized into several different data types, as outlined in table 1. These
Table 1. The types of data acquisition implemented that implicitly reduce the need for model accuracy or can utilize simpler empirical models.

<table>
<thead>
<tr>
<th>acquisition data</th>
<th>example</th>
<th>application</th>
<th>references</th>
</tr>
</thead>
<tbody>
<tr>
<td>wavelength ratio data</td>
<td>two transmitted wavelengths</td>
<td>pulse oximetry and tissue oximetry</td>
<td>[7–9]</td>
</tr>
<tr>
<td></td>
<td>spectral derivative data</td>
<td>multiple analyte analysis</td>
<td>[10–12]</td>
</tr>
<tr>
<td></td>
<td>fluorescence/excitation ratio</td>
<td>fluorescence imaging</td>
<td>[13–15]</td>
</tr>
<tr>
<td>multi-distance data</td>
<td>estimation of slopes with linear approximations</td>
<td>robust tissue spectroscopy systems insensitive to slight shape changes</td>
<td>[16–18]</td>
</tr>
<tr>
<td>small-volume sampling</td>
<td>scatter spectroscopy</td>
<td>fibre probes with size less than effective scatter distance</td>
<td>[19–22]</td>
</tr>
<tr>
<td></td>
<td>absorption spectroscopy</td>
<td>fibre probes sensitive to absorption only</td>
<td>[23,24]</td>
</tr>
<tr>
<td></td>
<td>fluorescence spectroscopy</td>
<td>fibre probes sensitive to fluorescence and not tissue</td>
<td>[25,26]</td>
</tr>
<tr>
<td>temporal signals</td>
<td>millisecond variations in tissue</td>
<td>oxygen saturation and haemodynamic sampling</td>
<td>[8,27,28]</td>
</tr>
<tr>
<td></td>
<td>microsecond sampling</td>
<td>flash photolysis analysis of biochemical changes in vivo</td>
<td>[29,30]</td>
</tr>
<tr>
<td></td>
<td>picosecond sampling</td>
<td>ultrafast signals to reduce model dependence</td>
<td>[31–35]</td>
</tr>
<tr>
<td></td>
<td>high-frequency phase shift derivative with distance</td>
<td>frequency-domain spectroscopy of tissue</td>
<td>[36,37]</td>
</tr>
<tr>
<td></td>
<td>lifetime-based signals</td>
<td>fluorophore identification or microenvironment analysis</td>
<td>[38]</td>
</tr>
</tbody>
</table>

include: (i) ratiometric or derivative data at two or more different wavelengths, (ii) multi-distance ratio or derivative data, (iii) small spatial volumes that either limit the effect of physical boundaries through scatter and/or absorption, or allow simpler empirical modelling, and (iv) temporal signals that are less sensitive to boundaries and/or more robustly insensitive to shape changes. The use of prior information about the tissue to be sampled is still essential in the design process with these systems, but can be implemented in the very first step of the NIRS process, namely data acquisition and calibration. This type of prior information is often used to eliminate the need for modelling completely, allowing simpler and perhaps more robust analysis of just the raw data or some processed version of the data. Examples of and references to these are shown in Table 1.

Approaches to analytically model NIRS light transport and allow direct inversion or fast inversion have dominated the successful commercial applications and led to several practical biological discoveries [39–42]. Many advanced solutions to the diffusion equation have been developed for regular geometries [43] and are widely used in practical systems, although contact variation or tissue shape, which cannot be accurately accounted for through modelling, can be a major problem for quantitative parameter estimation. The ideal systems use data types that are robust and less sensitive to boundary shape changes and fibre contact issues [35,44], as described in Table 1.
Figure 3. The use of spatial and spectral prior information has been widely studied. Spatial information can come from imaging systems such as X-ray, magnetic resonance imaging (MRI), ultrasound or surgical imaging. The integration of NIRS into these systems has been developed technologically, and the effect of prior spatial information upon imaging accuracy has been studied in niche cases. The application of spectral priors is thought to be quite beneficial, and has been widely adopted. In general, it is believed that using more spectral features, and sampling more wavelengths, improves the specificity of the inversion. Unfortunately, the most specific spectroscopy methods, such as Raman spectroscopy, also tend to correlate inversely with interaction cross section, leading to degraded signal-to-noise ratio. (Online version in colour.)

Systems that encode shape, location and external boundaries have been developed in many research and commercial systems. Figure 3 illustrates a handful of possible combinations, showing how inclusion of spatial information is done with additional hardware, and spectral prior inclusion is done with careful choice of the wavelength and source–detector bands.

One of the earliest hybrid systems was demonstrated by Ntziachristos et al. [45], where concurrent magnetic resonance imaging (MRI) and NIRS of breast tumours were used to improve the estimation of indocyanine green uptake concentration in tumours. The examples shown in this study stimulated substantial interest in this approach.

**(b) Internal shapes and physical limitations of light transport**

One of the largest limitations of NIRS in tissue is the physical spread of light during transmission as a result of multiple scattering effects. This multiple scattering over large distances leads to diffuse propagation and the inability to resolve structures near the size of the transport scattering length (near
Table 2. The different models that have been used for optical radiation transport modelling are listed (left column) with relative rankings (using + marks) of their applicability and utility for accurate use with different levels of prior information: −, not feasible; +, permitted but difficult; ++, permitted; and ++++, approach is ideally suited.

<table>
<thead>
<tr>
<th>Radiation transport model</th>
<th>Complex external shape</th>
<th>Complex boundary conditions</th>
<th>Refractive index variation</th>
<th>Internal heterogeneity inversion constraint</th>
<th>Hard spatial inversion constraint</th>
<th>Soft spatial inversion constraint</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boltzmann equation neutral particles</td>
<td>+</td>
<td>+++</td>
<td>+++</td>
<td>+++</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td>SPn approximation to the Boltzmann</td>
<td>+</td>
<td>+</td>
<td>++</td>
<td>+</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td>Monte Carlo stochastic diffusion—analytic</td>
<td>++</td>
<td>+++</td>
<td>++</td>
<td>+</td>
<td>+++</td>
<td>+++</td>
</tr>
<tr>
<td>diffusion—finite difference</td>
<td>++</td>
<td>+++</td>
<td>++</td>
<td>+++</td>
<td>+++</td>
<td>+++</td>
</tr>
<tr>
<td>diffusion—boundary element</td>
<td>+++</td>
<td>+++</td>
<td>++</td>
<td>−</td>
<td>+++</td>
<td>−</td>
</tr>
<tr>
<td>diffusion—finite element</td>
<td>+++</td>
<td>+++</td>
<td>+++</td>
<td>+++</td>
<td>+++</td>
<td>+++</td>
</tr>
</tbody>
</table>

1 mm). While recovery of heterogeneous regions on the surface can approach the resolution of microscopy systems, the resolution degrades by orders of magnitude with depth into the tissue, and so measurements taken over several centimetres of tissue effectively become averages over larger volumes of tissue. See figure 2b for an illustration of this effect. This phenomenon places size limits on objects that can be resolved with NIRS [46,47]. Moon et al. [46] developed a useful metric that estimates that the photon bundle path width, as illustrated in figure 2b, is approximately 20 per cent of the distance between the source and detector, when sources and detectors are arranged in transmission geometry. This suggests that the sensitivity profile (sometimes called the photon measurement density function, or photon path width) is 2 mm for a 1 cm separation, and 2 cm for a 10 cm separation. These numbers provide a rough guide to the type of resolution that could be achieved, although it is generally accepted that regularized inversion solvers can improve this by nearly four times, such that a fundamental resolution for optical tomography in a 10 cm domain might be closer to 5 mm. This improvement in resolution comes in part from multiple overlapping source–detector projections through the domain.

The forward light transport model must have the flexibility to accurately model any spatial heterogeneity. There are several approaches used to solve this problem, largely separated into deterministic models, such as the Boltzmann equation for neutral particle transport, or stochastic approaches, such as a Monte Carlo model. Table 2 summarizes the major modelling approaches used, with a relative ranking of their ability to incorporate heterogeneity. Several solutions for Monte Carlo models exist, and the ability to incorporate arbitrary geometries and arbitrary internal heterogeneities with ease has improved dramatically in recent years, as developed by Fang [48]. The inversion of Monte Carlo data is still somewhat crippled by the need to run repeated forward solutions, which takes significant amounts of time unless very regular geometries are used [49,50]. However,
the recent developments in graphics card processing units have made this a more realistic potential even for complex geometries [48,51]. A more common approach is the deterministic modelling of the diffusion equation, a second-order differential equation that is an approximation to the Boltzmann equation (a more complicated integro-differential equation). The transport equation itself can be directly solved with discrete ordinates solvers but requires substantial specialized computer resources and programming. As a consequence, the diffusion equation is simpler to solve than the transport equation in terms of both the mathematical complexity and the computational burden involved. Moreover, the diffusion equation depends on two rather than three phenomenological tissue parameters, namely, the absorption coefficient and the reduced scattering parameter. In fact, in the mathematical derivation leading to the diffusion equation, the scattering coefficient and the so-called phase function are absorbed into one parameter called the reduced scattering parameter. A critical point here, though, is that it is even more difficult, if not impossible, to devise generic approaches that would allow disentangling phase and scattering. However, accomplishing this would be the minimal step required for using the full benefits of the transport equation for forward modelling in NIRS. Moreover, the ability to incorporate accurate boundary conditions and internal heterogeneities also comes at a much greater difficulty level for approaches based on the Boltzmann equation.

An aspect that is sometimes regarded as an advantage for using the diffusion equation is the fact that many useful analytical solutions can be derived from it, while the number of solutions associated with the transport equation is much more limited. The transport solutions also typically involve several mathematical simplifying assumptions that limit the scope of their application. However, even in the case of the diffusion-based analytical equations, solving the internal heterogeneity must be accomplished using perturbation theory approaches, where the perturbed optical backgrounds are usually defined from analytical solutions, making such approaches reliable only for simple imaging geometries (e.g. tissue slabs in transmission, semi-infinite turbid media for epi-illumination imaging). These linear perturbation approaches are associated with the other drawback that the actual optical signal contrast rarely corresponds to optical property variations that satisfy the linear approximation from which the model is derived. Nevertheless, this has formed the basis for many reconstruction algorithms, allowing fast inversion with large datasets. The inclusion of arbitrary boundaries and arbitrary internal heterogeneity is arguably best achieved with a numerical model such as the finite element approach, where the domain is solved as a discrete defined set of points, linked by a mesh of finite basis functions. This approach was introduced for NIRS about 15 years ago, and has formed the basis of several available solvers. Some of these solvers have been developed to facilitate the use of a variety of priors, such as structural priors from MRI or X-ray tomosynthesis [52,53], spectral priors and biochemical priors.

(c) Explicit prior shape inclusion applications and analysis

The incorporation of internal structures in the inversion problem has been an area of increased academic interest in recent years. The prior information approach that is the less involved from both computational and mathematical complexity standpoints consists in pre-defining the imaging domain into a small
number of specific regions. In this case, reduction of the dimensionality of the inversion is achieved by assuming the regions to be optically homogeneous. This approach is often called ‘hard prior’ information and was demonstrated in breast imaging by Ntziachristos et al. [45]. The composite approach to this consists in relaxing the homogeneity assumption in order to retain the original large dimensionality of the inversion domain. However, increased flexibility is conferred to the inversion problem by applying appropriately penalized inversion methods that are encoding the prior information but still allow optical property gradients to exist within each composite region of the domain. A common implementation of this latter approach used by several groups is to use a Laplacian-type regularization matrix that smoothes the parameter values within the pre-defined regions, but allows separate regularization between different tissue types [54,55]. This has been used in breast imaging with MRI-guided NIRS by Brooksby et al. [56] to more accurately estimate the concentrations of haemoglobin, oxygen saturation, water and scattering values in fibroglandular and adipose tissues in the normal breast. This ‘soft priors’ approach was validated in multiple tissue-simulating phantom studies as shown in figure 4 and demonstrated

Figure 4. (a) Spatial prior inclusion can improve quantitative recovery, as shown in this MRI-guided phantom study, where images show the increasing level of spatial prior information used in the reconstruction algorithm. (b) Spectral prior inclusion can improve both spatial resolution and quantitative recovery, even without explicit tissue structure, as shown in these images. (c) Combining NIRS, spatial and spectral priors maximizes recovery accuracy in large regions, here providing the most benefit for quantifying haemoglobin (HbT), oxygen saturation (StO2), water and scattering values for adipose tissue and fibroglandular tissue in vivo. (Online version in colour.)
for tumour imaging by Carpenter et al. [57]. The fluorescence tomography version of this has been adopted by several groups [58], and has been used in several pre-clinical murine tumour imaging studies of contrast agent uptake [59,60].

The posterior incorporation of imaging shapes can be done in the analysis of NIRS data, and this is perhaps the most conventional way to approach combined modality data. A commercial prototype NIRS breast imager was shown to have significant discriminatory power between benign and malignant tumours [61]. The system was designed to be used in concert with mammogram analysis, thereby allowing posterior added information, for the improvement of the diagnostic reading, similar to the way ultrasound and mammography are interpreted together [62]. This added specificity has been confirmed in the summaries of two academic studies that show promise for interpretation of NIRS imaging in posterior combination with mammography [63,64]. Poplack et al. [64] showed that the combination of NIRS and mammography could increase the positive predictive power significantly. The explicit use of X-ray structure in NIRS inversion was demonstrated by Fang et al. [65] with a hybrid X-ray tomosynthesis system. This approach could be used to increase the sensitivity and/or specificity of the tomosynthesis exam depending upon the approach to use the prior information. Similarly, NIRS breast tumour imaging to quantify response from neoadjuvant chemotherapy with a posterior comparison to MRI by Choe et al. [66] demonstrated that contrast reduction from NIRS added to the diagnostic value of MRI. Additionally, Jiang et al. [67] examined the role of defining the region of interest based upon the size of the lesion in the initial pre-treatment image, and concluded that the quantitative estimate of tumour haemoglobin is significantly affected by the posterior estimate of lesion size. Thus, while posterior analysis is perhaps the least complex combination of NIRS and imaging data, it may become the most utilized as NIRS systems enter the clinic.

(d) Summary of benefits and limits

Several approaches were presented for which the inclusion of prior spatial information can increase the biological information content that can be derived from data acquired with NIRS instruments. Significant gains can be realized by using methodologies where the impact of tissue shape, source and detector coupling and light transport are minimized by an appropriate choice of data type, while still ensuring that sufficient tissue sampling is achieved to provide useful spatial resolution and quantification. Moreover, three approaches for inclusion of spatial information have been extensively examined to include internal tissue structures into the inverse problem, referred to here as (i) hard prior, (ii) soft prior and (iii) posterior use of the structural information. Of these, the one that is less technically involved is the third one, since it does not require that the prior information be fed directly to the NIRS reconstruction algorithm, thereby eliminating any potential reconstruction bias or the propagation of registration errors in the NIRS images. However, a significant limitation of this approach is that the information it can convey is limited by the intrinsic spatial resolution of diffuse optical imaging, namely from a few millimetres to a few centimetres depending on the interrogated volume size. In other words, although a tomography dataset might encode the information associated with a small area of optical contrast (i.e. smaller than the spatial resolution), NIRS images

Phil. Trans. R. Soc. A (2011)
might not reveal it, although contrast can be resolved on the structural image. Alternatively, ‘hard prior’ and ‘soft prior’ methods can potentially reveal the biological information contained in the NIRS dataset precisely because the prior spatial information is directly encoded in the inverse problem. Of those two approaches, the method based on soft priors is preferable because of the reduced likelihood that spatial biases will be introduced during the inversion process. The implementation of soft prior constraints requires that it be tested carefully and the approach be well calibrated, otherwise the use of hard priors is perhaps a safer implementation. Yet, even with hard priors, it is critical to know that the regions identified are accurate, otherwise it is easy to recover false values in the regions.

3. Prior information: molecular and biochemical limits

(a) Molecular and biochemical limits

The inclusion of spectral fitting into data analysis has been ubiquitous throughout most of NIRS’s development owing to its ability to recover concentrations of biochemical species rather than relative values or absorption coefficients. Systems with multiple wavelengths have the ability to directly measure the spectral constituent contributions to the signal independently, and it often does not require additional hardware. The value of spectral fitting has been established in many studies, and in bulk tissue spectroscopy without model-based inversion, this is often done either at the pre-processing stage or at the post-processing stage, without a significant difference in quantification between the two approaches. This is because of the linear relationship between the concentration of chromophores and optical absorption coefficients measured by NIRS in a bulk tissue measurement paradigm. However, in the NIRS imaging problem based on matrix inversion, the explicit inclusion of the pre-determined spectra in the inversion algorithm, rather than in a post-processing step, has been shown to dramatically improve the quantification and final interpretation. This is generally considered to be important because the spectral constraint inversion limits the solution space of the ill-posed problem [68–70] and thereby improves accuracy spatially. This can be seen in figure 4b where a phantom containing higher blood concentration was imaged with high noise, resulting in blotchy image features when spectral information was applied after the absorption coefficient images had been recovered. The inclusion of the spectral prior constraint directly in the imaging algorithm improved the recovery significantly for this noisy dataset [71]. When compared with a spatial prior implementation, it was determined that spectral prior information was superior in terms of quantitative accuracy, but that the inclusion of both provided the most accurate quantification, in the setting of breast imaging [56,69]. Some optimally fitted values for breast tissues are shown in figure 4c.

The type and quality of prior information should be a key consideration in the design phase of NIRS systems. In addition to the obvious design considerations necessary for building hybrid systems, the use of an optimal wavelength set for particular chromophores should also factor strongly in the system design [72,73]. While this is a subject of ongoing analysis, it is probable that systems that can adaptively change wavelengths or map to key wavelengths will allow more accurate quantification of chromophores in tissue.

Phil. Trans. R. Soc. A (2011)
Spectral constraints in fluorescence imaging have been widely adopted in both research and commercial luminescence NIRS systems. Prior knowledge of the luminescent spectra can be used to separate the signals from the analyte of interest from contaminating background molecules or to image multiple analytes simultaneously. Major commercial systems incorporate algorithms to fit images at multiple wavelength bands to recover fluorophore concentrations, or use the approach to suppress or remove non-specific background signals [74]. These approaches are essential tools for low-concentration imaging, or multiple analyte imaging in vivo. Additionally, a large body of work has been produced to either correct for or exploit the distortion of luminescent spectra by the absorbing spectra of the tissue. This can correct quantification of signals that would otherwise be incorrect [75–77] or be used to make an intractable problem solvable, such as in the case of bioluminescence tomography where spectral constraints are used to reduce the ill-posedness of the problem [77–82]. However, an inherent limitation of such approaches is that quantification improvements heavily rely on prior knowledge not only of the absorption spectra of the tissue chromophores but also of the magnitude of their relative contributions. In cases where this information is not available, modelling errors can propagate into the inversion process, potentially negating any gains attributable to the spectral constraints. Research attempting to quantify this effect and minimize its impact in applications such as diffuse fluorescence tomography is ongoing.

The inclusion of physiological limits in the fitting or estimation process is often essential in the presence of noise, either from stochastic processes (e.g. photon-detection shot noise) or from modelling errors (e.g. neglecting to account for the presence of a chromophore in spectral fitting applications). Such constraints are often included in fitting algorithms without mention, yet it can be the key factor in obtaining accurate estimates, especially when multiple overlapping species are being estimated, such as in endogenous NIRS imaging. Most systems explicitly incorporate this in the inversion process, requiring a constrained inversion algorithm. The importance of this problem can be traced to the fact that, for any inversion problem, there generally exist several equivalent solutions that match the data for a given set of convergence criteria determined by a residual expression (e.g. the two-norm of the difference between experimental data and model) or a generalized norm that includes explicit prior information. In tomography, this problem is often referred to as ill-posedness and, because of noise, leads to the paradoxical situation that solutions that fit the measurements perfectly are always non-physical. Often, these non-physical solutions will, for example, include negative values for parameters that can only be positive (e.g. concentration of molecules) and lead to blood concentrations beyond or below what is physiologically possible. In such cases, inversion algorithms must be implemented where prior information in the form of bound constraints (e.g. non-negativity, lower and upper limits for recovered parameters) are introduced.

Endogenous NIRS in vivo breast imaging requires sparse sampling to cover the large volume of tissue, creating the ill-posedness problem discussed previously. From a practical standpoint, prior information has become commonplace in diffuse optical breast imaging, and is quintessential for the recovery of prognostically valuable images. Figure 5 gives an illustrative example of a reconstruction of data from a hybrid MRI–optical imaging system. In this example, both the hardware design and the reconstruction algorithm use
spatial and spectral prior information. This system incorporates MRI images by collecting optical and MRI data simultaneously, providing a high-resolution spatial image from which to generate a boundary mesh. Additionally, adipose and fibroglandular tissue can be separated and assigned to specific nodes in the optical domain. By generating the optical solution space from the MRI images, the datasets are easily co-registered and a ‘hard priors’ reconstruction can be applied. Figure 5a shows an axial MRI image, which would be used to separate tissue types. The imaging system also uses spectral prior information in two ways. The six imaging wavelengths are optimized in the NIR window to maximize sensitivity to spectral features. As a result, each of the five chromophores seen in figure 5b–g can be more robustly decoupled during image reconstruction. The reconstruction algorithm in this case also constrains the solution to be physiologically valid, as discussed earlier. In this example, the use of spectral and spatial priors reduces the number of unknowns significantly and allows the optical solution to more accurately quantify the tissue properties.

Posterior or post-processing analysis of spectra has been suggested in the interpretation of pathlength based upon assumed concentrations of absorbers, such as water, in tissue [83]. While this approach is not widely used, it is possible that such post-processing based upon spectra could provide easier or more accurate quantification.

(b) Summary of benefits and limits

For the situation of single source–detector measurements it is unclear whether prior knowledge in pre- versus post-processing is more beneficial, and it seems probable that the two are substantially similar, owing to the linearity of the data during processing. For imaging applications where matrix inversion is involved,
which is ill-posed, there seems to be a significant benefit in the inversion step, owing to the magnitude of the regularization parameter relative to the matrix diagonal values. The evidence to date would support the fact that spectral prior information is perhaps even more important than spatial structure, but this is probably case-specific. More development and analysis in this area is warranted.

4. Prior information: patient demographics and biomarkers

(a) Patient demographics and biomarkers

Few studies have investigated incorporating demographic information about patients in NIRS; however, there are some niche applications where it may prove to be beneficial. NIRS data in breast imaging has shown strong correlations with demographic factors such as age, body mass index (BMI), radiographic breast density and menopausal/hormonal status. This correlation implies that there could be some benefit in using this prior information, perhaps to make systems simpler or more reliable. Examples of this approach include the use of radiographic density or BMI to estimate lipid concentration in the breast. Radiographic density is also highly correlated to optical scattering, indicating that this dominant factor affecting optical signals may be derived from a clinical radiograph. In the event that this was shown to be feasible, this would reduce or eliminate the need for accurate frequency- or time-domain NIRS measurement systems required for recovering optical scattering properties. Simpler, less expensive continuous-wave systems would be adequate to quantify absorbing and/or luminescent NIR signals.

The use of explicit biomarker prior information was recently demonstrated with a hybrid MRI–NIRS system [84]. This study demonstrated that incorporating fat/water values estimated with MRI fat–water imaging into the NIRS image recovery algorithm improved the accuracy of haemoglobin and oxygen saturation fitting. This hybrid information combination has been largely untested but will probably have increasing relevance as systems are developed to have mutual, complementary information sets. Additional data that would be synergistic with NIRS include blood haematocrit level and concentration of tracers in the blood.

Risk prediction of breast cancer is a recent example of how the intersection of patient demographics with NIRS has the potential for clinical impact. Lilge and colleagues (e.g. [85]) have systematically examined the potential for NIRS to quantify the parenchymal density of the breast. The density is a known correlate to the risk of incidence of breast cancer, and so NIRS is being examined as a research tool to help understand the nature of the correlation, and to determine if NIRS measurement systems could provide a low-cost screening platform for population risk analysis. In this paradigm, NIRS may be used to test and analyse the demographic community with a known genetic predisposition to high incidence of cancer, particularly younger women in whom higher numbers of mammograms might have adverse effects [86]. In addition to parenchymal density, biochemical and structural changes related to NIRS data have been shown to be associated with the risk of cancer [87], further indicating that NIRS data combined with patient demographic information might provide diagnostic information about breast cancer risk. This work is still ongoing, but serves as a key example of the potential synergy that might exist with the right application.

Phil. Trans. R. Soc. A (2011)
Table 3. Some of the major biophysical models used in conjunction with NIRS, with applications and key references.

<table>
<thead>
<tr>
<th>biophysical model examples</th>
<th>application</th>
<th>reference(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>molecular diffusion and degradation models</td>
<td>dissolved oxygen metabolism interpretation</td>
<td>[88–91]</td>
</tr>
<tr>
<td></td>
<td>drug photobleaching and transport</td>
<td>[92]</td>
</tr>
<tr>
<td></td>
<td>drug partitioning and localization</td>
<td>[93]</td>
</tr>
<tr>
<td></td>
<td>molecular consumption during therapy</td>
<td>[94]</td>
</tr>
<tr>
<td>compartment rate model</td>
<td>pharmacokinetic analysis of drugs</td>
<td>[95,96]</td>
</tr>
<tr>
<td>vascular convection model</td>
<td>blood flow quantification</td>
<td>[96–100]</td>
</tr>
<tr>
<td></td>
<td>drug vascular leakage and kinetic identification</td>
<td>[101–105]</td>
</tr>
<tr>
<td>viscoelastic model</td>
<td>interpreting response to tissue pressure and</td>
<td>[106–112]</td>
</tr>
<tr>
<td></td>
<td>displacement</td>
<td></td>
</tr>
<tr>
<td></td>
<td>vascular resistance estimation</td>
<td>[113,114]</td>
</tr>
<tr>
<td>receptor, ligand and activatable probe models</td>
<td>enzyme activation</td>
<td>[115–117]</td>
</tr>
<tr>
<td></td>
<td>vascular permeability versus cellular binding</td>
<td>[118–120]</td>
</tr>
<tr>
<td></td>
<td>receptor internalization and renewal modelling</td>
<td>[121,122]</td>
</tr>
</tbody>
</table>

(b) Summary of benefits and limits

Strong correlations exist between NIRS data and demographics. Datasets from different diagnostic modalities may be exchanged or used to reduce uncertainty in quantification, when applied a priori rather than analysed posterior. The prior implementation of demographic or biomarker data is largely undeveloped in the NIRS prior implementation.

5. Prior information: biophysical models

(a) Biophysical models

Biophysical models to interpret the NIRS data have been used widely for a range of applications. While these could also be incorporated into the NIRS process at earlier steps, it is most common at the post-processing or interpretation stage. Examples of biophysical models are listed in table 3. Tissue oxygen extraction models have perhaps been most widely incorporated into NIRS endogenous imaging of oxygen saturation because of the potential to estimate tissue consumption of oxygen in key organs such as the brain or in diseases such as cancerous tumours. Generally, biophysical model information that could be included is in three different temporal domains: (i) static information (i.e. fat, water, blood volumes), (ii) harmonic information (i.e. blood pulsation or breathing or hormonal cycling), or (iii) dynamic inflow/outflow data (i.e. contrast agent injection and clearance). The latter two are illustrated in figure 6.

Post-processing of fluorescence data has been widely analysed to study pharmacokinetics in vivo as a way to improve data interpretation and extract additional information from the measured signals. Temporal fitting can be used with compartmental models to estimate the pharmacokinetics of injected dyes [101,123,124], drug production or bleaching rates [125,126], or drug-binding rates.

Phil. Trans. R. Soc. A (2011)
Figure 6. The type of data collected for transmitted or emitted signals are illustrated as either (a) harmonic or (b) dynamic transient. If these patterns are known and can be modelled with an explicit numerical function, then it can be incorporated into the prefiltering, inversion or post-processing of the data, leading to output of the model biophysical parameters instead of raw data such as absorption, fluorescence yield or scattering. (Online version in colour.)

In vivo [118]. Temporal analysis was used to illustrate how fluorescence signals can identify different organs in vivo by Hillman et al. [102]. This is a post-processing approach to include prior information about the molecules. Incorporation of the drug kinetics in the forward or inverse photon propagation models has been shown in just one or two studies [95], and there is a promise that including an accurate kinetic model in the inversion problem may improve accuracy in quantification. More complex kinetic or tissue modelling procedures have been studied by some groups [103]. Additionally, several studies have indicated that better knowledge of the drug concentration a priori would improve the ability to define the best acquisition parameters [127], indicating that inclusion in the data acquisition or system design level could also be beneficial.

Incorporating viscoelastic tissue modelling in NIRS is an approach that has seen some recent activity. Dynamic measurements induced by changes in pressure applied to the tissue can provide a rich dataset for robust, often ratiometric, interpretation of oxygen saturation and total haemoglobin values [106–111]. These mechanical models are almost uniformly applied in post-processing or forward modelling [112], owing to the complexity of solving the viscoelastic problem in irregular domains. While they may introduce additional uncertainty in the inversion process, if implemented in the right way, they have the potential to reduce the ill-posedness of the problem and thus offer more accurate diagnostic performance.

Model-based interpretation of injected tracers is one area where there may be substantial gains, which would allow NIRS to provide fundamentally new information that cannot be obtained by other modalities. The key factor in this is that NIRS imaging allows quantification of more than one analyte in vivo so multiple species can be used to reference against one another. This has been
shown for vascular permeability studies [119] and for drug-binding studies [118].

In this type of analysis, the binding rate constant of the targeted probe is directly related to the drug affinity and the concentration of binding sites within the tissue. Thus, by combining NIRS measurements or images with compartment-binding modelling, it is feasible to quantify molecular receptor expression and affinity in vivo. The processing of this information has been in post-processing to date, but explicit incorporation of these linear models into the inversion algorithm with a temporal dataset is readily done. An example of explicit incorporation of a forward model is in the way spectral constraints are added into a multiple-wavelength inversion problem, to reduce the ill-posedness of the inversion and provide more accurate recovery properties. A similar approach could be tried with forward physical models, in the inversion problem.

(b) Summary of benefits and limits

Proven benefits have been widespread in post-processing of NIRS data. Inclusion of biophysical modelling into forward NIRS modelling is probably somewhat beneficial. However, similar to spectral constraints, there is probably a major benefit in inclusion into the inverse algorithms.

6. Prior information: impact on diagnosis

Ultimately, any convolution of prior information data streams with NIRS must be assessed for diagnostic performance. This is a critical component of the innovation process often missing from NIRS studies. The prior–NIRS convolution must be assessed against other convolution schemes as well as against clinical and pre-clinical diagnostic standards. An important consideration in this analysis is the trade-off between performance and technical complexity. The value of modest gains in diagnostic power realized through extensive, difficult and time-consuming imaging strategies must be considered carefully.

Figure 7 illustrates how a receiver operating characteristic (ROC) curve can be used to compare the diagnostic performance of different prior implementations that can be used for MRI-guided NIRS imaging. In this example, a coronal MRI image of a human breast lightly compressed between two plates (figure 7a) was segmented and used to produce a finite element mesh containing two distinct tissue regions, an adipose and a fibroglandular region. Each compression plate is lined with eight optical source–detector fibres for transmitting and receiving light through the tissue, as illustrated by grey rectangles in figure 7b. The ROC curve analysis began by generating simulated test domains from the finite element mesh. A suspicious circular legion was numerically added to the mesh in one of two locations, either near the centre of the tissue, as shown in figure 7b, or a specified distance from one edge, as illustrated in figure 7c. The lesion simulates a region of gadolinium (Gd) enhancement in the MRI image, but may not necessarily be malignant. The radius of this lesion was varied over the range illustrated in figure 7b,c, and the optical absorption contrast was varied from 0 to 5 : 1 for each lesion size. Lesions with zero contrast in optical absorption represented an MRI false positive. For each lesion size/contrast, forward data were generated numerically and noise added to the measurements. A total of 600 datasets were considered in this example.
Figure 7. (a) A T1-weighted coronal MRI image of a breast compressed between two plates. Tissue types were segmented and used to generate a two-tissue-region finite element mesh for NIRS modelling as shown in (b,c). Sixteen fibre probes contact the tissue surface in a transmission geometry, as illustrated in (b). Regions that simulate suspicious lesions defined by Gd-MRI were added numerically to the mesh, simulating a positive reading in the MRI image. The size, position and optical contrast of these lesions were varied over a wide range (the lesion sizes and positions are illustrated in (b,c)). The ROC curve in (e) demonstrates the improvement in diagnostic performance when the hard prior approach is used (red curve) versus the approach that uses the MRI information only in the interpretation step of the imaging process (blue curve). If the segmentation of the fibroglandular region is incorrect, as shown in (d), the hard prior approach (red curve) loses much of its diagnostic advantage over the other method (blue curve), as shown in (f). (Online version in colour.)

The synthetic optical data were then used to recover and interpret absorption images using one of two prior-imaging convolutions. The first method applied the structural prior in the interpretation step, while the second approach considered the hard-prior method, which incorporated the MRI information directly in the image reconstruction step. In the former approach, the optical data were used to perform a ‘no-priors’ image formation. Once the image was recovered, the average value in the region defined by the lesion in the corresponding MRI image...
was extracted as the diagnostic parameter. The hard-prior approach, on the other hand, encodes the internal structure of the tissue in optically homogeneous regions. Thus, the value recovered in this lesion was extracted directly from the image.

The value of optical absorption in the lesion region was used as the diagnostic parameter in the ROC curve analysis for both methods shown in figure 7e. This result demonstrates an improvement in ROC curve performance when the prior was applied in the reconstruction step as opposed to the interpretation step. Area-under-the-curve metrics were 0.85 and 0.73 for the hard-prior and interpretation-only approaches, respectively. However, this analysis assumed perfect image segmentation of the fibroglandular region. Errors in segmenting this region will affect the performance of the hard-prior approach but should not affect the approach that uses the MRI information only in the interpretation step. To explore to what degree the performance is degraded, the segmentation of the fibroglandular region used in the hard-prior image reconstruction was dilated, as shown in figure 7d, and the analysis repeated. ROC curves for the two approaches under imperfect segmentation conditions are shown in figure 7f, and demonstrate a significant decrease in diagnostic performance for the hard-prior approach due to the segmentation error. Analyses such as this can help determine what imaging strategies are most beneficial under different experimental conditions.

The strategy described above is an effective way to evaluate and compare different implementations of prior information. ROC curve analysis is the preferred method for assessing diagnostic performance in most cases (though specific applications, such as glucose monitoring, have their own standard diagnostic scales). Other types of analyses are commonly reported for NIRS imaging, such as spatial resolution, contrast resolution and contrast detail analysis. While useful for assessing system performance, these analyses should be considered an intermediate step on the path towards a full diagnostic assessment using an ROC curve, which also accounts for biological variability. Establishing truth can be challenging and must be accomplished using accurate gold standards of conventional medicine. Also, the metrics assessed should match the objectives of the diagnostic test. In some cases, the aim of the NIRS technique is to quantify rather than detect diseased tissue, and thus an assessment of detectability would be inappropriate. This is especially true for paradigms in which structural imaging modalities are used to identify suspicious lesions and a prior–NIRS convolution strategy is used to characterize and diagnose the abnormalities. In most cases, a comprehensive ROC curve study should be the ultimate goal of the NIRS researcher to determine whether adoption of the technique in the clinic or research setting is warranted.

7. Conclusions

The observations of improved quantification or improved diagnostic values are a complex range of niche examples. A few applications show clear and significant benefits, others demonstrate only modest improvements in diagnosis, and many have not been tested using the conventions of diagnostic assessment. Table 4 provides a listing of the prior information types from figure 1, along with the NIRS process steps. The table includes a qualitative ranking of which prior information

Phil. Trans. R. Soc. A (2011)
Table 4. Listing of the prior information options and NIRS process steps outlined in figure 1, with relative scoring of the proven benefit, where notation is as follows: −, no benefit; +, possible benefit; ++, demonstrated some benefit; and ++++, demonstrated major benefit.

<table>
<thead>
<tr>
<th>NIRS process step</th>
<th>spatial structures</th>
<th>molecular spectra</th>
<th>parameter limits</th>
<th>biomarkers and demographics</th>
<th>biophysical models</th>
</tr>
</thead>
<tbody>
<tr>
<td>data acquisition</td>
<td>+ + +</td>
<td>+ +</td>
<td>+</td>
<td>−</td>
<td>+++</td>
</tr>
<tr>
<td>pre-processing</td>
<td>+ +</td>
<td>+ + +</td>
<td>−</td>
<td>−</td>
<td>++</td>
</tr>
<tr>
<td>forward model</td>
<td>+ + +</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>+</td>
</tr>
<tr>
<td>inversion model</td>
<td>+ +</td>
<td>+ + +</td>
<td>+ +</td>
<td>+</td>
<td>++</td>
</tr>
<tr>
<td>post-processing</td>
<td>−</td>
<td>+ +</td>
<td>++</td>
<td>++</td>
<td>+ +</td>
</tr>
<tr>
<td>interpretation</td>
<td>+ + +</td>
<td>−</td>
<td>−</td>
<td>+ + +</td>
<td>+ +</td>
</tr>
</tbody>
</table>

Some streams have benefits in each of the NIRS steps, using − to indicate no clear benefit, + to indicate possible benefit, ++ to indicate probable benefit with some demonstrated evidence, and ++++ to indicate a major demonstrated benefit in quantification or diagnosis.

While these scores are admittedly relative and somewhat subjective, this chart is an initial attempt to think about this framework for determining how the benefit from prior information might be maximized in the NIRS process, and how systems can be designed to maximize quantitative accuracy and ultimately diagnostic value in terms of the sensitivity and specificity trade-off.

This work was funded by the National Cancer Institute research grants RO1CA069544, RO1CA109558, K25CA138578 and R01CA120368.

References


Review. Prior information in NIR imaging

4553


Phil. Trans. R. Soc. A (2011)


