Dear Dr. Naylor and Professor Andersson-Engels,

Thank you for reviewing and providing comments for my PhD thesis titled “Detection of joint inflammation in rheumatoid arthritis using multispectral diffuse optical imaging”. I would like to thank both reviewers for taking the time to review the work and thank them for their highly constructive comments. I have addressed these specific queries below and made the appropriate changes within the revised document, with any substantial changes highlighted in red below. I have also made a number of spelling / grammatical corrections throughout the revised version, including those highlighted either in part 4 or annotated in the provided thesis copies.

* List of acronyms and abbreviations is needed (list of figures and tables exists, but not acronyms)

*This has been included with the corresponding acronyms updated throughout for clarity to the reader.*

* Consider removing all acronyms from the abstract and from the conclusion chapter to make the thesis easier to browse.

*Thank you, these sections have been altered accordingly.*

* It would be great if you could give a few sentences about the contributions in developing the Lambert-Bouguer Law and Beer-Lambert Law

*I have adjusted my description as follows to properly attribute the works of these Physicists in contributing to these laws.*

“In a non-scattering medium, the probability of absorption events occurring means incident light intensity (*I0*) decays exponentially with pathlength *L* through the medium, a relationship commonly described as the Lambert-Bouguer law. This behaviour was first observed in essays by French Physicist Pierre Bouguer in 1729 and later outlined in *Photometria* by Johann Heinrich Lambert, and can be summarised by the following equation,

In 1852, August Beer expanded this relationship to a more general description commonly described as the Beer-Lambert law that considers the dependence of $\mu\_a$ on the mediums individual constituent chromophores, which can be expressed according to the equation,”

* Please clarify that you mean elastic scattering when you talk about Rayleigh-type scattering

“As light travels through biological tissue, photons may also change direction due to either an inelastic scattering event, which is the theoretical basis of Raman spectroscopy and occurs around once in every 106 interactions [65], or an elastic scattering interaction, which can be either with particles smaller than λ (Rayleigh scattering) or atoms / molecules larger than λ (Mie scattering), the second of which is predominantly of interest in DOI.”

* Provide a discussion on why you selected the DA in your reconstructions. Pros and cons.

*I have rewritten the section on page 29 and 30 to include a more detailed description of why the DA was chosen, including the advantages and disadvantages when compared with alternative forward models.*

“The RTE is a highly accurate, deterministic model for light travelling through a homogeneous turbid media and this integro-differential equation can be solved numerically using, for example, Monte-Carlo (MC) simulation, where probability density functions can help calculate the trajectories of a large number of simulated photons [102]. Implementation of MC methods, however, are challenging and come with high computational expense, as both spatially and angularly discretised degrees of freedom are needed. Higher order approximations to the RTE exist that rely on more complex integral-differential equations to incorporate both directional and spatial components of light propagation. Examples include spherical harmonics PN or discrete ordinates method SN, in which the angular components of fluence in each method are either expanded into spherical harmonics [158] or discretised into a number of directions respectively [155], however implementation of these forward models is challenging and very computationally expensive, making reconstruction times undesirably long for the proposed clinical prototype device in which multiple joints are imaged during a single imaging session. A less computationally expensive, simplified spherical harmonics (SPN) approximation to the RTE has been proposed [159], in which the 3D SPN forward model consists of four coupled diffusion equations for four composite moments of fluence, the derivation and full implementation of which has been discussed in detail in the referenced works. These equations in full define the SP7 model, however they can be further simplified to either SP5, SP3 or SP1. Integration of the SPN forward model into FEM package NIRFAST was previously reported [213], however limitations of this implementation include a more under-determined inverse problem from a greater number of Nth order unknowns and an inability to experimentally measure individual composite moments using CW systems, such that the total fluence must be approximated by summing the individual composite moments [214], meaning this forward model requires significantly more testing and verification before any clinical applications.

The presence of regions of low scattering such as cerebrospinal fluid in the brain or

very short source-detector separations can invalidate assumptions used in the derivation of

the DA about isotropic fluence or the dominance of scatter over absorption and lead to errors in image reconstruction [154, 155]. Although in this sense the DA sacrifices some accuracy compared to alternative forward models of the RTE, it is significantly easier and therefore faster to solve, meaning it is commonly used as a forward model due to its robustness, flexibility and computation speed [112, 93, 190]. The DA was therefore selected in this work for several reasons, 1. it provided fast reconstruction times preferential for the extensive testing during system development and joint imaging during pilot studies, 2. it was well-integrated into the readily available open-source FEM package NIRFAST [72], and 3. this implementation had been extensively tested and verified during comparative studies with both analytical solutions or phantom measurements [113,115].”

* Provide a short discussion that your reconstruction is objective in terms of algorithm and initial conditions.

*The following has been added to highlight this point in the conclusion of chapter 3.*

The resulting system was safe, fast, fully-automated and straightforward to operate, such that an examination could be easily carried out by a trained non-clinician. The entire data processing pipeline described in this chapter, including mesh generation, optical data filtering and reconstruction, was implemented in an automated fashion with identical algorithm parameters and initial conditions chosen for all participants. This produced maps of the joint pathophysiology that were quantitatively comparable between subject groups and made the algorithm objective in the sense that the same outcomes would be reached if it were repeated by multiple users when using the same raw data. Some aspects of the data acquisition workflow that would benefit from further automation or computer-aided guidance in future included the joint positioning process and the thresholding to produce the image mask used during the auto-exposure routine, in order to ensure any operator subjectivity is minimised.

* In the discussion of optical imaging, provide a short paragraph comparing the different optical techniques and the strength of DOT in comparison

*The following has been added to Section 2.8*

“A number of distinct optical-based techniques were presented as potential candidates for imaging RA patients. Despite PAT and FOI both demonstrating the capability to detect inflammation with a high spatial resolution, DOT provided a unique combination of strengths that made it the most desirable approach in the context of diagnostics and monitoring. DOT is a more economically viable solution for widespread clinical implementation in comparison to PAT, with additional, expensive hardware needed in the latter for US measurement, whilst the ability to provide contrast to endogenous pathophysiological properties of DOT is highly preferable when compared to the reliance on intravenous contrast agents for FOI, particularly given the regularity of which patients are monitored for disease progression.”

* Correct Figure 3.4(b)

*This figure has been removed, as the data was not correctly calibrated for at the time and its removal is not significantly detrimental to the work.*

* In the system characterisation chapter include a small paragraph pointing out the most important parameters in order to get an accurate reconstruction (SI, limitations of the diffusion approximation, or any of the other parameters. Also add a bit of discussion on error propagation analysis in this part.

*Thank you, this is an important aspect of system design and characterisation that was considered throughout the development, and has now been discussed in more detail as follows in the conclusion of chapter 4.*

“The complexity of non-contact DOT meant that the accuracy of reconstructed image maps was dependant on many measurement parameters, which together will have combined measurement errors in a non-trivial way. A rigorous data filtering operation based on absolute intensity and Gaussian smoothing was chosen based on phantom studies, to ensure that random noise was in DOI data was kept to a minimum and the SNR was greater than 100 in all measurements. As with many optical-based devices, instrument calibration played a crucial role in accounting for a number offsets or system characteristics, including the intrinsic and extrinsic parameters for both CMOS and CCD cameras for SI, the spectral instrument response function for DOI data, light source coupling variation, the data-model offset and 3D source positions, all of which were relevant for tomographic reconstruction. These calibrations were all implemented on a one-time a basis and fixed in all cases, an approach that was validated during repeat studies, providing benefits by removing the need for regular, lengthy recalibration and additionally eliminated the introduction of random variation between datasets resulting from noise in recalibration data that would have an multiplicative impact on boundary fluence. Nevertheless, there will inevitably still be some systematic error in these calibrations, for example a degree of discrepancy between the calibrated and true source positions in 3D, which should be further minimised in future studies. The exponential relationship between boundary flux and pathlength, as described by the Lambert-Bouger law in Equation 2.6, highlights the importance of accurate of SI to produce a FEM boundary truly reflective of the pathlength taken as any error in pathlength will also have a similar impact on boundary flux. Finally, a number of assumptions required for light propagation modelling, for example the conditions placed on the interaction of light at boundary or the use of the DA, will lead to systematic errors. Attempts have been made during system design and characterisation in this chapter to quantify and minimise these potential sources of measurement error, and importantly, an emphasis has been placed on ensuring consistency in acquisition and processing settings between subjects, so that participant data in human studies is quantitatively comparable in all instances and that results of classification are only indicative of any true differences between the pathophysiological states of inflamed and non-inflamed joints.”

* Please explain Eq.5.1 better

*The variables for this calculation have been more clearly defined to allow it to be reproduced and given in the context of an example for tHb when comparing groups of joints between subjects.*

* In the patient chapter you tested many ways to analyse the data in a relatively small data set. I would have preferred you tested all techniques on one dataset and then evaluated on a completely independent set. Please add a small discussion on this.

*and*

* It would also be preferred to have a brief discussion about any potential benefits/complications obtained by combining the two techniques of analysis used.

*The reviewer correctly identifies the need for assessing the performance of a diagnostic test in a machine learning scheme, through which any benefits of combining multiple features potentially from both techniques could be properly assessed without the risk of overfitting. These two points have been addressed in the following discussion added to the conclusion of chapter 6.*

“A large number of extracted features were considered in this chapter, from either 3D images of the joint pathophysiology using multispectral DOT or raw DOI transmission images at individual wavelengths. AUC values from ROC analysis provided useful insight into the variation of these distinct features and their individual capabilities for classifying inflamed and non-inflamed joints. For example, this analysis highlighted the potential benefit of considering spatial analysis of DOI transmission images as opposed to absolute transmitted signal. The natural extension of this work, when a larger cohort of patients is available, is the implementation of a more complex ML scheme, in which data is partitioned into appropriate training, testing and validation sets in order to fully assess the potential diagnostic accuracy of these features. In this way, optimal weightings and thresholds can be selected using the training data based on their performance when classifying unseen test data, with generalisation of this performance then verified on a separate validation set [226]. In such a scheme, the benefits of combining multiple features that provide distinct discriminatory value can be properly assessed without the risk of over-fitting. FFT features based on DOI measurements were subject to a very distinct processing pipeline compared to the additional SI data and more complex reconstruction algorithms required to recover pathophysiological parameters using DOT. These two techniques may therefore contain orthogonal information that when combined could improve diagnostic accuracies. One drawback of this approach is that the benefits of simplicity for data acquisition, speed and low cost of DOI transmission imaging on its own would no longer be relevant when tomographic data is additionally incorporated, so a significant increase in diagnostic accuracy would have to be demonstrated to make this approach worthwhile.”

* In general look at improving figure and table captions. Especially important are legends to Table 5.1, Figure 1.3 and Figure 5.23.

*These and a number of other captions including Figures 1.1, 1.6, 2.12, 2.7, 3.7, 5.6 and Table 1.2, have been expanded upon in more detail, based on comments in manuscript. Figure 1.3 has also been reorientated to be made more clear.*

* Keep the order of StO2, tHb, SA and H2O the same in all figures e.g. Fig 5.22

*This has been rearranged for all relevant figures and tables.*

* Standardise the reference style throughout the reference list

*This has been done, with all references having been standardised to have initials only.*

Alterations based on comments in the hard copies of the manuscript.

* Pg 20 with decay constant µa known as the absorption coefficient having units mm-1, commonly chosen as such within the literature in place of their SI unit equivalent as they provide a sensible numerical scaling of values.
* Pg 23 Importantly, when bound to oxygen molecules at up to four potential binding sites,
* Pg 26 The later has been commonly implemented in joint imaging, presumably as it should provide a better dynamic range for this application as a result of a smaller variation in the pathlength when considering source-detector pairs for an array of detectors on an opposing boundary of a cylindrical-like object when compared to reflectance mode, when sources and detectors are placed on a common boundary.
* Pg 34 Through optimising a wavelength set selecting a combination that simultaneously minimise k and maximise R,
* Pg130 The null hypothesis that the median values between different subjects were equal was rejected for all features except minimum SA, with the majority of p-values much less than 0.05, meaning a statistical difference was seen between different participants in recovered pathophysiological parameters in these cases. In contrast, the null hypothesis that the median values were between different subjects was accepted for all features for variability between fingers, with the exception of minimum tHb.
* Pg138 Following the release of the cuff pressure, the total optical transmission increases, typically returning to a slightly lower value during recovery than initial baseline at rest. This may be a result of the large level of blood pooling occurring during venous occlusion and would be expected to return to baseline values given a longer total monitoring time.
* Pg 143 In future work, extending these healthy studies to a larger cohort, in which volunteer enrolment is controlled for age and sex matching with arthritis patient demographics, would be beneficial in order to minimise any volunteer bias and provide a control group dataset more comparable with patient studies.
* Pg 171. Labels added to this Figure 6.16.

I hope I have addressed all issues raised by the reviewers and would like to thank them again for their constructive comments.

Yours faithfully,

Daniel Lighter