DETECTING BREAST CANCER TISSUE

WITH
DIFFUSE
REFLECTANCE
SPECTROSCOPY

LISANNE L. DE BOER

DETECTING BREAST CANCER TISSUE WITH DIFFUSE REFLECTANCE SPECTROSCOPY

Lisanne L. de Boer

Dit proefschrift is goedgekeurd door
de promotor(en):

prof.dr. T.J.M. Ruers

de co-promotor(en):

prof.dr. B.H.W. Hendriks
dr. M.T.F.D. Vrancken Peeters

Cover design: Douwe Hoedervanger
Lay-out: RON Graphic Power, www.ron.nu
Printed by: ProefschriftMaken || www.proefschriftmaken.nl

ISBN: 978-94-6380-327-4

DOI: 10.3990/1.9789463803274
<https://doi.org/10.3990/1.9789463803274>

© 2019 Lisanne de Boer, The Netherlands.

All rights reserved. No parts of this thesis may be reproduced, stored in a retrieval system or transmitted in any form or by any means without permission of the author. Alle rechten voorbehouden. Niets uit deze uitgave mag worden vermenigvuldigd, in enige vorm of op enige wijze, zonder voorafgaande schriftelijke toestemming van de auteur.

DETECTING BREAST CANCER TISSUE WITH DIFFUSE REFLECTANCE SPECTROSCOPY

PROEFSCHRIFT

ter verkrijging van
de graad van doctor aan de Universiteit Twente,
op gezag van de rector magnificus,
prof.dr. T.T.M. Palstra,
volgens besluit van het College voor Promoties
in het openbaar te verdedigen
op woensdag 22 mei 2019 om 14:45 uur

door

Lisanne Lotte de Boer

Geboren op 20 februari 1988
te Amsterdam

PROMOTIECOMMISSIE:

Voorzitter/secretaris prof.dr. J.L. Herek

Promotor prof.dr. T.J.M. Ruers Antoni van Leeuwenhoek Ziekenhuis
 – Nederlands Kanker Instituut

Co-promotor prof.dr. B.H.W. Hendriks
 dr. M.T.F.D. Vrancken Peeters

Leden prof.dr. M.M.A.E. Claessens
 prof.dr. S. Siesling
 dr. C. Otto
 prof.dr. K.K. Van de Vijver
 prof.dr. H.J.C.M. Sterenborg
 dr. D.M. de Bruin

Table of contents

1	Introduction	7
2	Review: <i>In vivo</i> optical spectral tissue sensing - how to go from research to routine clinical application?	23
3	Fat/Water ratios measured with diffuse reflectance spectroscopy to detect breast tumor boundaries	41
4	Using DRS during breast-conserving surgery: identifying robust optic parameters and influence of inter-patient variation	59
5	A method for co-registration of optical measurements of breast tissue with histopathology: the importance of accounting for tissue deformations	79
6	The influence of neoadjuvant chemotherapy on DRS spectra of breast surgery specimens	103
7	Fiber-optic DRS measurements of DCIS and classification of locations with mixed tissue types	123
8	Towards the use of diffuse reflectance spectroscopy for real-time <i>in vivo</i> detection of breast cancer during surgery	151
9	General discussion	177
	Appendices	195
	English summary	196
	Nederlandse samenvatting	202
	PhD portfolio	209
	Dankwoord	211
	Curriculum vitae	213

Introduction

BREAST CANCER SURGERY

Up to the 1970s, patients suffering from breast cancer were treated with a so-called Halsted radical mastectomy. During this type of surgery, the entire breast was removed, with surrounding tissue such as skin, underlying chest muscles and lymph nodes in the axilla. This mutilating procedure was abandoned with the introduction of breast-conserving treatment which aims at conserving the shape of the breast by resecting only a small portion of the breast including the tumor (breast-conserving surgery) followed by radiation therapy.[1] Nowadays, the majority of patients with small tumors undergo breast-conserving surgery after thorough research had shown that breast-conserving surgery in combination with postoperative radiation therapy results in similar survival as resection of the entire breast.[2-9] Furthermore, due to the improved efficacy of neo-adjuvant treatment regimes, more and more patients that are diagnosed with large tumors become eligible for breast-conserving surgery after successful neoadjuvant systemic treatment.[10-12] The great benefit of breast-conserving treatment opposed to a mastectomy is the improved cosmetic outcome, which results in a better quality of life for the patient.[13-19] However, there is one very important condition that has to be met during breast-conserving surgery in order to minimize the risk of a local recurrence, which is complete removal of the tumor, or in other words, the absence of positive resection margins. [20-26]

RESECTION MARGIN ASSESSMENT

Positive resection margins can be found by a pathologist during histopathological evaluation of the resected tissue which can be performed a few days after surgery. To enable this type of assessment, samples of the tissue are fixed in formalin, embedded in paraffin, and cut in one-cell layer thick slices which, after staining, can be viewed underneath the microscope. Depending on the type of staining that was used, different cell structures can be visualized which are used to distinguish the different cell types. Pathology paint, which was applied on the margin of the specimen before it was fixed, and which does not disintegrate during the processing of the tissue, allows the pathologist to recognize the margin of the specimen in the microscopy section. Subsequently, the pathologist can assess the distance between the inked margin and the malignant tissue and determines if a margin is considered 'positive' and eventually needs additional therapy. The definition of a positive margin depends on the type of malignancy, specifically whether the tissue is invasive carcinoma or ductal carcinoma in situ.

INVASIVE CARCINOMA

The definition of a positive margin in the case of invasive carcinoma has been heavily debated over the past years. It has shifted over time progressively to smaller distances between the tumor border and the edge of the resection specimen, to the currently used definition of 'no ink on tumor'. This implies that as long as pathology ink is not touching the invasive carcinoma cells, the margin is negative. In a large meta-analysis by Moran et al. it was concluded that this definition of a positive resection margin minimizes the risk of a local recurrence and can thus be used as the standard for invasive carcinoma.[27] In this study it was furthermore concluded that wider margins do not significantly lower the risk of developing a local recurrence and therefore obtaining wider margins than 'no ink on tumor' is not indicated for invasive carcinoma. In the Netherlands, there is a tendency to an even more liberal definition than 'no ink on tumor' as research has shown that re-excision of focally positive resection margins does not affect survival.[28] Focally involved margins are defined as a margin in which the tumor is touching the margin ink over a distance of up to 4 mm. A re-excision is only indicated when the margin is more than focally positive, meaning that the tumor is touching the margin ink over a distance of more than 4 mm.[29]

As the definition of a positive resection margin is different from country to country and has shifted over time, providing a number of positive resection margins is difficult, but in general the reported number of positive resection margins for invasive carcinoma range between 3.5 and 11%.[30-32]

DCIS

As previously mentioned the definition of a positive resection margin is also dependent on the type of malignancy. Some patients are diagnosed with a precursor stage of invasive carcinoma, so-called Ductal Carcinoma In Situ (DCIS). DCIS has the potential to evolve in invasive carcinoma and although that is not always the case, surgical treatment is often indicated for these patients.[33] During surgery DCIS is more difficult to detect as the volume of DCIS is often smaller than areas of invasive carcinoma, and its growth pattern is more discontinuous.[33] Due to these biological differences, the definition of a positive resection margin for DCIS is different from invasive carcinoma. A margin is considered positive when DCIS is present within a 2 mm distance from the resection margin, albeit that this definition is also under debate. [34-37] Using this definition, positive resection margins are found in a range between 4 and 52% of patients that were treated with breast-conserving surgery. [30, 38-40] Again, in the Netherlands the definition is more liberal as here no ink on DCIS is considered a negative margin.

Another reason why DCIS is clinically a more challenging disease as it is often not discovered pre-operatively but may coexist with invasive carcinoma. It then becomes apparent during the histopathologic evaluation of the resection specimen. This explains why the majority of re-operations is performed in patients with margins that were positive for DCIS.[41]

THE CONSEQUENCE OF POSITIVE RESECTION MARGINS

In current day practice, patients with positive resection margins receive additional therapy with re-excision or boost radiation therapy depending on the extent of the positive resection margin. Reexcision rates of up to 20% are reported worldwide.[42-47] Both re-excision and boost radiation therapy negatively affect the cosmetic outcome and patient satisfaction after breast-conserving surgery and induce additional morbidity.[48-57] Besides the patient discomfort and potential impact on the cosmetic outcome, secondary surgery also affects healthcare budgets.[44, 58] Additionally, in 20-70% of the patients that have a secondary surgery no tumor deposits are found in the re-excision specimen.[59-61] Because the primary motivation for choosing BCS over mastectomy is the improved cosmetic outcome the surgeon obviously wishes to avoid re-operation or boost radiation therapy and strives to obtain the best cosmetic outcome as possible by resecting the entire tumor during the initial surgery.

IN THE OPERATING THEATER...

In the operating theater, a surgeon is faced with finding the balance between resecting the entire tumor, to avoid additional therapy, while limiting the excision volume as much as possible, to have the best possible cosmetic outcome. This is a major challenge since discriminating tumor tissue from healthy tissue during surgery can be extremely difficult as it often looks and feels the same. Currently, available margin assessment technologies have limitations. For example, they are time-consuming because they lack the possibility of real-time feedback (touch imprint cytology, frozen section analysis), have limited sensitivity (intra-operative ultrasound, touch imprint cytology), damage the tissue (frozen section analysis), or require skilled personnel (intra-operative ultrasound, touch imprint cytology and frozen section analysis).[62-65]

In practice, the surgeon aims at placing his/her resection plane at a distance of 1.0 cm from the border of the tumor as this eventually will yield a microscopic margin of 1 to 5 mm in the pathological analysis.[66] This has consequences for the amount of healthy tissue that is resected. A study by Haloua et al that assessed the margin status in relation to the amount

of healthy breast tissue that was resected found that the volume of the excised tissue was 2.3 times as much as the optimal resection volume. Even in the case of a tumor involved margin, the resection volume was 1.5 times bigger than the ideal volume. Also, in 33.8% of cases, the tumor was eccentrically located in the specimen.[31]

The fact that 1) surgeons aim for a larger rim of healthy tissue than necessary (1.0 cm), 2) excision volumes are larger than necessary, 3) in one third of patients the tumor is not centrally located in the specimen and 4) a positive resection margin is found in up to 52% of patients emphasize that there is a clinical need for real-time tissue characterization of the resection plane during breast-conserving surgery.

DIFFUSE REFLECTANCE SPECTROSCOPY

Diffuse Reflectance Spectroscopy (DRS) is an optical technology that reflects the composition and morphology of tissue by measuring the interaction between light and tissue. Light from a broadband light source is used to illuminate the tissue. After the light has interacted with the tissue, the collected light is acquired with spectrometers which can resolve the spectral response in the visible and near-infrared wavelength range. The magnitude and shape of the acquired spectrum represent the absorption and scattering properties of the tissue. These optical characteristics can be used to characterize tissue and discriminate different tissue types.

The advantages of DRS are that it is non-destructive, does not require exogenous contrast with dyes, and has the potential to be performed real-time. In practice, DRS measurements can be performed by means of a camera (hyperspectral imaging) or a fiber-optic probe (point measurements). In hyperspectral imaging, hypercubes are generated, which are 3D datasets in which two dimensions contain spatial information and one dimension contains spectral information.[67] This technology has the ability to acquire the optical properties of tissue over a large field of view in a relatively short amount of time. For the fiber-optic setting, measurements are performed with optical fibers, for example, embedded in a probe or needle. One of the fibers is used for illuminating the tissue, whereas another fiber is used for detecting the light after it has interacted with the tissue. The distance between the illuminating and detecting fiber typically range from 1 to 8 mm and will influence the volume of the tissue that is measured. By bringing the tool in contact with the tissue a point measurement can be performed. Fiber-optic DRS has been investigated thoroughly for discriminating healthy tissue from tumor tissue in several organs.

There are several methods to analyze a DRS measurement. One approach is to use an analytical fit model to quantify the spectrum. Such a model is based on the diffusion theory that describes light propagation through tissue. With the known scattering and absorption characteristics of tissue constituents and the measured DRS spectrum, the fit model can

quantify the presence of these constituents in the measurement volume. Main substances that are included in the model are fat, water, oxyhemoglobin, deoxyhemoglobin. The benefit of this approach is that the derived optical parameters can be directly linked to the biology of the tissue. The main drawback of this method is that it will only be useful when all assumptions of the diffusion theory are met. An important assumption is that the photons undergo several scattering events before collection, which in practice means that the distance between the illuminating and collecting fiber should be big enough. However, when the distance between the illuminating fiber and detecting fiber becomes too large the measurement volume becomes inhomogeneous. In this case, another requirement of diffusion theory is violated which states that the measurement volume has to be homogeneous. Use of diffusion theory-based models can thus not be applied in all cases.

Another method is to analyze the optical spectra without relating them to biological components. With such an approach, datasets are analyzed using the measured intensity or other spectral features to discriminate different tissue types. This method does not require any a priori knowledge about the composition of the tissue but also does not allow quantifying a spectrum in terms of biological substances.

Previous research investigated the use of DRS in breast cancer for chemotherapy response monitoring[68-70], biopsy guidance[71, 72], and margin assessment[73, 74]. It is hypothesized that adding DRS to a surgical tool can help the surgeon in characterizing the tissue and thereby provide guidance during breast-conserving surgery. The aim of this thesis was to address some of the challenges that hamper progressing DRS into the OR room for the detection of positive resection margins during breast-conserving surgery. Specifically, we will investigate:

How to use DRS for discriminating between healthy tissue and tumor tissue (invasive carcinoma)

The first step is to investigate if DRS measurements differ between healthy tissue and tumor tissue. Then, if spectra differ, the next question is how these differences can be used to actually discriminate healthy breast tissue from tumor tissue.

Differences between *in vivo* and *ex vivo* measurements

Initial accuracy of the technology for discriminating healthy tissue from tumor tissue is often investigated in a controlled *ex vivo* setting. The ultimate goal though is to detect a positive resection margin *in vivo*, at the time of resection during surgery. It is thus important to investigate if the time of measurement (i.e. *in vivo* or *ex vivo*), affects the ability of DRS to discriminate healthy tissue from tumor tissue.

The potential influence of patient-specific characteristics

Besides the time of measuring also patient-specific characteristics might influence the

optical contrast between tumor tissue and healthy breast tissue. These patient-specific characteristics include the menopausal status of the patient, as well as treatment with chemotherapy prior to surgery. Especially the influence of neoadjuvant chemotherapy is relevant as, due to improved efficacy of chemotherapeutics, the number of patients that receive neoadjuvant chemotherapy increases.

The inhomogeneous structure of breast tissue

The fourth challenge is related to the often inhomogeneous structure of breast cancer. In the breast, different tissue types are often mixed within a small tissue volume. To be able to investigate not only 'pure' measurement locations but also locations that contain a mixture of tissue types, accurate correlation of measurement locations with histopathology is imperative. This challenge is also directly related to the development of classification models that require robust data input.

Measuring DCIS

A fifth major challenge for successfully identifying positive resection margins is the detection of DCIS. This tissue type accounts for the majority of positive resection margins and is often not discovered until microscopic evaluation by the pathologist. It is therefore important to acquire DRS measurements of areas with DCIS. This requires an accurate method of registering histopathology to optical measurements and inclusion of a large number of patients as often DCIS is missed because of its diffuse and scattered growth pattern.

Translating the technology into the clinical workflow

The last challenge considers the use of DRS technology in the clinical workflow. In the final application, we envision that DRS will be used in the operating room, *in vivo*, at the time of resection. It is therefore important to consider if DRS can measure (near) real-time and if these measurements can be classified correctly.

OUTLINE

Ultimately overcoming the previously mentioned challenges should advance the technology further towards the ultimate goal of developing a tool that can be used by the surgeon for margin assessment. The challenges are addressed throughout the chapters in this thesis. The outline of the thesis is as follows:

Chapter 2 involves a review regarding the current position of optical spectroscopy and the challenges encountered in translating DRS technology to the clinic and how to overcome these hurdles.

In Chapter 3 the results of an *ex vivo* study that focused on the detection of tumor boundaries in *ex vivo* breast specimens are presented. The first aim was to examine which optical parameters provided the best discrimination between healthy tissue and tumor tissue. The second aim was to investigate the ability of this optical parameter to detect the tumor boundary.

The difference between DRS measurements that were acquired *in vivo* and DRS measurements that were acquired *ex vivo* are researched in Chapter 4. In the analysis also patient-specific characteristics such as neoadjuvant treatment with chemotherapy and menopausal status were included.

As previously stated, attaining accurate correlation between optical measurements and histopathology is important when using the technology in inhomogeneous tissue as breast tissue. To enable evaluation of the DRS for the detection of DCIS and mixture locations a method that improved registration between the optical and histopathology is presented and evaluated in Chapter 5.

Chapter 6 reports on the assessment of the influence of chemotherapy on the DRS spectra as well as the optical parameters in a large *ex vivo* dataset.

With the use of the improved registration method in Chapter 5, it was possible to examine the spectral difference between invasive carcinoma and DCIS (Chapter 7). Next, we investigated the influence of the purity of the measurements on classification model performance. One classification model was selected and with this model, the locations with a mix of tissue types are classified.

In Chapter 8 we present the results of an *in vivo* study in which an optical biopsy needle is used to perform DRS measurements during routine breast biopsy procedures. With this dataset which was well-correlated to histopathology, a classification algorithm was developed that was used to classify the measurements that were acquired in a continuous mode in an additional five patients.

The general discussion and future perspectives are presented in Chapter 9.

REFERENCES

- [1] U. Veronesi, and S. Zurrada, "Breast cancer surgery: a century after Halsted," *Journal of Cancer Research and Clinical Oncology*, vol. 122, pp. 74-77, 1996.
- [2] J. A. Van Dongen, A. C. Voogd, I. S. Fentiman, C. Legrand, R. J. Sylvester, D. Tong, E. Van der Schueren, P. A. Helle, K. Van Zijl, and H. Bartelink, "Long-Term Results of a Randomized Trial Comparing Breast-Conserving Therapy With Mastectomy: European Organization for Research and Treatment of Cancer 10801 Trial," *Journal of the National Cancer Institute*, vol. 92, no. 14, pp. 1143-1150, 2000.
- [3] M. Morrow, E. A. Strom, L. W. Bassett, D. D. Dershaw, B. Fowble, A. Guiliano, J. R. Harris, F. O'Malley, S. J. Schnitt, S. E. Singletary, and D. P. Winchester, "Standard for Breast Conservation Therapy in the Management of Invasive Breast Carcinoma," *CA: A Cancer Journal for Clinicians*, vol. 52, no. 5, pp. 277-300, 2002.
- [4] B. Fisher, S. Anderson, J. Bryant, R. G. Margolese, M. Deutsch, E. R. Fisher, J.-H. Jeong, and N. Wolmark, "Twenty-Year Follow-up of a Randomized Trial Comparing Total Mastectomy, Lumpectomy, and Lumpectomy Plus Irradiation for the Treatment of Invasive Breast Cancer," *The New England Journal of Medicine*, vol. 347, no. 16, pp. 1233-1241, 2002.
- [5] M. M. Poggi, D. N. Danforth, L. C. Sciuto, S. L. Smith, S. M. Steinberg, D. J. Liewehr, C. Menard, M. E. Lippman, A. S. Lichter, and R. M. Altamus, "Eighteen-Year Results in the Treatment of Early Breast Carcinoma with Mastectomy versus Breast Conservation Therapy," *Cancer*, vol. 98, pp. 697-702, 2003, [[10.1002/cncr.11580](https://doi.org/10.1002/cncr.11580)].
- [6] R. Arriagada, M. G. Lê, F. Rochard, and G. Contesso, "Conservative Treatment Versus Mastectomy in Early Breast Cancer: Patterns of Failure With 15 Years of Follow-Up Data," *Journal of Clinical Oncology*, vol. 14, no. 5, pp. 1558-1564, 1996.
- [7] U. Veronesi, N. Cascinelli, L. Mariani, M. Greco, R. Saccozzi, A. Luini, M. Aguilar, and E. Marubini, "Twenty-year Follow-up of a Randomized Study Comparing Breast-Conserving Surgery with Radical Mastectomy for Early Breast Cancer," *The New England Journal of Medicine*, vol. 347, no. 16, pp. 1227-1232, 2002.
- [8] S. Litière, G. Werutsky, I. S. Fentiman, E. Rutgers, M.-R. Christiaens, E. Van Limbergen, M. H. A. Baaijens, J. Bogaerts, and H. Bartelink, "Breast conserving therapy versus mastectomy for stage I-II breast cancer: 20 year follow-up of the EORTC 10801 phase 3 randomised trial," *Lancet Oncology*, vol. 13, pp. 412-419, 2012, [[10.1016/S1470-2045\(12\)70042-6](https://doi.org/10.1016/S1470-2045(12)70042-6)].
- [9] R. J. Bleicher, K. Ruth, E. R. Sigurdson, J. M. Daly, M. Boraas, P. Anderson, and B. L. Egleston, "Breast Conservation Versus Mastectomy for T3 Primaries (>5 cm): A Review of 5,685 Medicare Patients," *Cancer*, vol. 122, no. 1, pp. 42-49, 2016, [[10.1002/cncr.29726](https://doi.org/10.1002/cncr.29726)].
- [10] H. Charfare, "Neoadjuvant chemotherapy in breast cancer," *British Journal of Surgery*, vol. 92, pp. 14-23, 2005, [[10.1002/bjs.4840](https://doi.org/10.1002/bjs.4840)].
- [11] S. E. Singletary, M. D. McNeese, and G. N. Hortobagyi, "Feasibility of Breast-Conservation Surgery After Induction Chemotherapy for Locally Advanced Breast Carcinoma," *Cancer*, vol. 69, no. 11, pp. 2849-2852, 1992.
- [12] A. Charehbili, D. B. Y. Fontein, J. R. Kroep, G. J. Liefers, J. S. D. Mieog, J. W. R. Nortier, and C. J. H. Van de Velde, "Neoadjuvant hormonal therapy for endocrine sensitive breast cancer: A systematic review," *Cancer Treatment Reviews*, vol. 40, pp. 86-92, 2014, [[10.1016/j.ctrv.2013.06.001](https://doi.org/10.1016/j.ctrv.2013.06.001)].
- [13] J. F. Waljee, E. S. Hu, P. A. Ubel, D. M. Smith, L. A. Newman, and A. K. Alderman, "Effect of Esthetic Outcome After Breast-Conserving Surgery on Psychosocial Functioning and Quality of Life," *Journal of Clinical Oncology*, vol. 26, pp. 3331-3337, 2008, [[10.1200/JCO.2007.13.1375](https://doi.org/10.1200/JCO.2007.13.1375)].
- [14] V. Arndt, C. Stegmaier, H. Ziegler, and H. Brenner, "Quality of life over 5 years in women with breast cancer after breast-conserving therapy versus mastectomy: a population-based study," *Journal of Cancer Research and Clinical Oncology*, vol. 134, pp. 1311-1318, 2008, [[10.1007/s00432-008-0418-y](https://doi.org/10.1007/s00432-008-0418-y)].

- [15] W. Janni, D. Rjosk, T. Dimpfl, K. Heartl, B. Strobl, F. Hepp, A. Hanke, F. Bergaure, and H. Sommer, "Quality of Life Influenced by Primary Surgical Treatment for Stage I-III Breast Cancer—Long-Term Follow-Up of a Matched-Pair Analysis," *Annals of Surgical Oncology*, vol. 8, no. 6, pp. 542-548, 2001.
- [16] D. Curran, J. P. van Dongen, N. K. Aaronson, G. Kiebert, I. S. Fentiman, F. Mignolet, and H. Bartelink, "Quality of Live of Early-stage Breast Cancer Patients Treated with Radical Mastectomy or Breast-conserving Procedures: Results of EORTC Trial 10801," *European Journal of Cancer*, vol. 34, no. 3, pp. 307-314, 1998.
- [17] R. Chow, N. Pulezas, L. Zhang, C. Ecclestone, A. Leahey, J. Hamer, C. DeAgelis, G. Bedard, R. McDonald, A. Bhatia, J. Ellis, E. Rakovitch, s. Vuong, E. Chow, and S. Verma, "Quality of life and symptom burden in patients with breast cancer treated with mastectomy and lumpectomy," *Supportive Care in Cancer*, vol. 24, 2016, [[10.1007/s00520-015-3027-8](https://doi.org/10.1007/s00520-015-3027-8)].
- [18] J. H. Volders, V. L. Negenborn, M. H. Haloua, N. M. A. Krekel, K. Jóźwiak, S. Meijer, and P. M. Van den Tol, "Cosmetic outcome and quality of life are inextricably linked in breast-conserving therapy," *Journal of Surgical Oncology*, vol. 115, pp. 941-948, 2016, [doi.org/10.1002/jso.24615].
- [19] M. K. Kim, T. Kim, H. G. Moon, U. S. Jin, K. Kim, J. Kim, J. W. Lee, J. Kim, E. Lee, T. K. Yoo, D.-Y. Noh, K. W. Minn, and W. Han, "Effect of cosmetic outcome on quality of life after breast cancer surgery," *European Journal of Surgical Oncology*, vol. 41, pp. 426-432, 2015, [[10.1016/j.ejso.2014.12.002](https://doi.org/10.1016/j.ejso.2014.12.002)].
- [20] S. E. Singletary, "Surgical margins in patients with early-stage breast cancer treated with breast conservation therapy," *American Journal of Surgery*, vol. 184, pp. 383-393, 2002.
- [21] C. C. Park, M. Mitsumori, A. Nixon, A. Recht, J. Connolly, R. Gelman, B. Silver, S. Hetelekidis, A. Abner, J. R. Harris, and S. J. Schnitt, "Outcome at 8 Years After Breast-Conserving Surgery and Radiation Therapy for Invasive Breast Cancer: Influence of Margin Status and Systemic Therapy on Local Recurrence," *Journal of Clinical Oncology*, vol. 18, no. 8, pp. 1668-1675, 2000.
- [22] N. Houssami, P. Macaskill, L. Marinovich, and M. Morrow, "The Association of Surgical Margins and Local Recurrence in Women with Early-Stage Invasive Breast Cancer Treated with Breast-Conserving Therapy: A Meta-Analysis," *Annals of Surgical Oncology*, vol. 21, pp. 717-730, 2014, [[10.1245/s10434-014-3480-5](https://doi.org/10.1245/s10434-014-3480-5)].
- [23] M. M. Pilewskie, M., "Margins in Breast Cancer: How Much Is Enough?," *Cancer*, 2018, [[10.1002/cncr.31221](https://doi.org/10.1002/cncr.31221)].
- [24] T. L. Huston, and R. M. Simmons, "Locally recurrent breast cancer after conservation therapy," *The American Journal of Surgery*, vol. 189, pp. 229-235, 2005, [[10.1016/j.amjsurg.2004.07.039](https://doi.org/10.1016/j.amjsurg.2004.07.039)].
- [25] N. Biglia, R. Ponzzone, V. E. Bounous, L. L. Mariani, F. Maggiorotto, C. Benevelli, V. Liberale, M. C. Ottino, and P. Sismondi, "Role of re-excision for positive and close resection margins in patients treated with breast-conserving surgery," *The Breast*, vol. 23, pp. 870-875, 2014, [[10.1016/j.breast.2014.09.009](https://doi.org/10.1016/j.breast.2014.09.009)].
- [26] A. Bodilsen, K. Bjerre, B. V. Offersen, P. Vahl, N. Amby, J. M. Dixon, B. Ejlersen, J. Overgaard, and P. Christiansen, "Importance of Margin Width in Breast-Conserving Treatment of Early Breast Cancer," *Journal of Surgical Oncology*, vol. 113, pp. 609-615, 2016, [[10.1002/jso.24224](https://doi.org/10.1002/jso.24224)].
- [27] M. S. Moran, S. J. Schnitt, A. E. Guiliano, J. R. Harris, S. A. Khan, J. Horton, S. Klimberg, M. Chavez-MacGregor, G. Freedman, N. Houssami, P. L. Johnson, and M. Morrow, "Society of Surgical Oncology - American Society for Radiation Oncology Consensus Guideline on Margins for Breast-Conserving Surgery With Whole-Breast Irradiation in Stages I and II Invasive Breast Cancer," *International Journal of Radiation Oncology Biology Physics*, vol. 88, no. 3, pp. 553-564, 2014.
- [28] E. L. Vos, S. Siesling, M. H. A. Baaijens, C. Verhoef, A. Jager, A. C. Voogd, and L. B. Koppert, "Omitting re-excision for focally positive margins after breast-conserving surgery does not impair disease-free and overall survival," *Breast Cancer Research and Treatment*, vol. 164, pp. 157-167, 2017, [[10.1007/s10549-017-4232-6](https://doi.org/10.1007/s10549-017-4232-6)].

- [29] "<https://www.oncoline.nl/borstkanker>", Integraal Kankercentrum Nederland, 2018.
- [30] S. G. Brouwer de Koning, M. T. F. D. Vranken Peeters, K. Jóźwiak, P. A. Bhairasing, and T. J. M. Ruers, "Tumor Resection Margin Definitions in Breast-Conserving Surgery: Systematic Review and Meta-analysis of Current Literature," *Clinical Breast Cancer*, 2018.
- [31] M. H. Haloua, J. H. Volders, N. M. A. Krekel, E. Barbé, C. Sietses, K. Jóźwiak, S. Meijer, and M. P. Van den Tol, "A nationwide pathology study on surgical margins and excision volumes after breast-conserving surgery: There is still much to be gained," *The Breast*, vol. 25, pp. 14-21, 2016, [[10.1016/j.breast.2015.11.003](https://doi.org/10.1016/j.breast.2015.11.003)].
- [32] A. M. Shulman, J. A. Mirrielees, G. Levenson, J. Landercasper, C. Greenberg, and L. G. Wilke, "Reexcision Surgery for Breast Cancer: An Analysis of the American Society of Breast Surgeons (ASBrS) MasterySM Database Following the SSO-ASTRO "No Ink on Tumor" Guidelines," *Annals of Surgical Oncology*, vol. 24, pp. 52-58, 2017, [[10.1245/s10434-016-5516-5](https://doi.org/10.1245/s10434-016-5516-5)].
- [33] H. J. Burstein, K. Polyak, J. S. Wong, S. C. Lester, and C. M. Kaelin, "Ductal Carcinoma in Situ of the Breast," *New England Journal of Medicine*, vol. 350, pp. 1430-1441, 2004.
- [34] M. Morrow, K. J. Van Zee, L. J. Solin, N. Houssami, M. Chavez-MacGregor, J. R. Harris, J. Horton, S. Hwang, P. L. Johnson, M. L. Marinovich, S. J. Schnitt, I. Wapnir, and M. S. Moran, "Society of Surgical Oncology–American Society for Radiation Oncology - American Society of Clinical Oncology Consensus Guideline on Margins for Breast-Conserving Surgery with Whole Breast Irradiation in Ductal Carcinoma In Situ," *Practical Radiation Oncology*, vol. 6, no. 5, pp. 287-295, 2016, [[10.1016/j.prro.2016.06.011](https://doi.org/10.1016/j.prro.2016.06.011)].
- [35] A. L. Merrill, R. Tang, J. K. Plichta, U. Rai, S. B. Coopey, M. P. McEvoy, K. S. Hughes, M. C. Specht, M. A. Gadd, and L. A. Smit, "Should New "No Ink On Tumor" Lumpectomy Margin Guidelines be Applied to Ductal Carcinoma In Situ (DCIS)? A Retrospective Review Using Shaved Cavity Margins," *Annals of Surgical Oncology*, vol. 23, pp. 3453-3458, 2016.
- [36] K. J. Van Zee, P. Subhedar, C. Olcese, S. Patil, and M. Morrow, "Relationship Between Margin Width and Recurrence of Ductal Carcinoma In Situ: Analysis of 2996 Women Treated With Breast-conserving Surgery for 30 Years," *Annals of Surgery*, vol. 262, no. 4, pp. 623-631, 2015, [[10.1097/SLA.0000000000001454](https://doi.org/10.1097/SLA.0000000000001454)].
- [37] M. L. Marinovich, L. Azizi, P. Macaskill, L. Irwig, M. Morrow, L. J. Solin, and N. Houssami, "The Association of Surgical Margins and Local Recurrence in Women with Ductal Carcinoma In Situ Treated with Breast-Conserving Therapy: A Meta-Analysis," *Annals of Surgical Oncology*, vol. 23, no. 12, pp. 3811-3821, 2016, [[10.1245/s10434-016-5446-2](https://doi.org/10.1245/s10434-016-5446-2)].
- [38] T. Shaikh, T. Li, C. T. Murphy, N. G. Zaorsky, R. J. Bleicher, E. R. Sigurdson, R. W. Carlson, S. B. Hayes, and P. Anderson, "Importance of Surgical Margin Status in Ductal Carcinoma In Situ," *Clinical Breast Cancer*, vol. 16, no. 4, pp. 312-318, 2016, [[10.1016/j.clbc.2016.02.002](https://doi.org/10.1016/j.clbc.2016.02.002)].
- [39] A. Laws, M. S. Brar, A. Bouchard-Fortier, B. Leong, and M. L. Quan, "Does intra-operative margin assessment improve margin status and re-excision rates? A population-based analysis of outcomes in breast-conserving surgery for ductal carcinoma in situ," *Journal of Surgical Oncology*, vol. 118, pp. 1205-1211, 2018, [[10.1002/jso.25248](https://doi.org/10.1002/jso.25248)].
- [40] A. Monaghan, N. Chapinal, L. Hughes, and C. Baliski, "Impact of SSO-ASTRO margin guidelines on reoperation rates following breast-conserving surgery," *The American Journal of Surgery*, vol. In press, 2018.
- [41] L. Langhans, M.-B. Jensen, M.-L. M. Talman, I. Vejborg, N. Kroman, and T. F. Tvedskov, "Reoperation Rates in Ductal Carcinoma In Situ vs Invasive Breast Cancer After Wire-Guided Breast-Conserving Surgery," *JAMA Surgery*, vol. 152, no. 4, pp. 378-384, 2017, [[10.1001/jamasurg.2016.4751](https://doi.org/10.1001/jamasurg.2016.4751)].
- [42] S. C. J. Bosma, F. Van der Leij, E. Van Werkhoven, H. Bartelink, J. Wesseling, S. Linn, E. J. Rutgers, M. J. Van de Vijver, and P. H. M. Elkhuisen, "Very low local recurrence rates after breast-conserving therapy: analysis of 8485 patients treated over a 28-year period," *Breast Cancer Research and Treatment*, vol. 156, pp. 391-400, 2016, [[10.1007/s10549-016-3732-0](https://doi.org/10.1007/s10549-016-3732-0)].

- [43] M. G. Valero, M. A. Mallory, K. Losk, M. Tukenmez, J. Hwang, K. Camuso, C. Bunnell, T. A. King, and G. Mehra, "Surgeon Variability and Factors Predicting for Reoperation Following Breast-Conserving Surgery," *Annals of Surgical Oncology*, vol. 25, pp. 2573-2578, 2018, [[10.1245/s10434-018-6526-2](#)].
- [44] R. E. Pataky, and C. R. Baliski, "Reoperation costs in attempted breast-conserving surgery: a decision analysis," *Current Oncology (Toronto, Ont.)*, vol. 23, no. 5, pp. 314-321, 2016, [[10.3747/co.23.2989](#)].
- [45] L. E. McCahill, R. M. Single, E. J. Aiello Bowles, H. S. Feigelson, T. A. James, T. Barney, J. M. Engel, and A. A. Onitilo, "Variability in Reexcision Following Breast Conservation Surgery," *Journal of the American Medical Association*, vol. 307, no. 5, pp. 467-475, 2012.
- [46] A. J. Isaacs, M. L. Gemignani, A. Pusic, and A. Sedrakyan, "Association of Breast Conservation Surgery for Cancer With 90-Day Reoperation Rates in NewYork State," *JAMA Surg*, 2016, [[10.1001/jamasurg.2015.5535](#)].
- [47] L. G. Wilke, T. Czechura, C. Wang, B. Lapin, E. Liederbach, D. P. Winchester, and K. Yao, "Repeat Surgery After Breast Conservation for the Treatment of Stage 0 to II Breast Carcinoma A Report From the National Cancer Data Base, 2004-2010," *JAMA Surgery*, vol. 149, no. 12, pp. 1296-1305, 2014, [[10.1001/jamasurg.2014.926](#)].
- [48] C. Vrieling, L. Collette, A. Fourquet, W. J. Hoogenraad, J.-C. Horiot, J. J. Jager, M. Pierart, P. M. Poortmans, H. Struikmans, B. Maat, E. Van Limbergen, and H. Bartelink, "The influence of patient, tumor and treatment factors on the cosmetic results after breast-conserving therapy in the EORTC 'boost vs. no boost' trial," *Radiotherapy and Oncology*, vol. 55, pp. 219-232, 2000.
- [49] C. Vrieling, L. Collette, A. Fourquet, W. J. Hoogenraad, J.-C. Horiot, J. J. Jager, M. Pierart, P. M. Poortmans, H. Struikmans, M. Van der Hulst, E. Van der Schueren, and H. Bartelink, "The influence of the boost in breast-conserving therapy on cosmetic outcome in the EORTC 'Boost versus No Boost' trial," *International Journal of Radiation Oncology Biology Physics*, vol. 45, no. 3, pp. 677-685, 1999.
- [50] J. H. Volders, V. L. Negenborn, M. H. Haloua, N. M. A. Krekel, K. Jóźwiak, S. Meijer, and M. P. Van den Tol, "Breast-specific factors determine cosmetic outcome and patient satisfaction after breast-conserving therapy: Results from the randomized COBALT study," *Journal of Surgical Oncology*, vol. 117, pp. 1001-1008, 2018, [[10.1002/jso.25012](#)].
- [51] S. K. Al-Ghazal, L. Fallowfield, and R. W. Blamey, "Comparison of psychological aspects and patient satisfaction following breast conserving surgery, simple mastectomy and breast reconstruction," *European Journal of Cancer*, vol. 36, pp. 1938-1943, 2000.
- [52] R. A. Cochrane, P. Valasiadou, A. R. M. Wilson, S. K. Al-Ghazal, and R. D. Macmillan, "Cosmesis and satisfaction after breast-conserving surgery correlates with the percentage of breast volume excised," *British Journal of Surgery*, vol. 90, no. 12, pp. 1505-09, 2003, [[10.1002/bjs.4344](#)].
- [53] A. Munshi, S. Kakkar, R. Bhutani, R. Jalali, A. Budrukhar, and K. A. Dinshaw, "Factors Influencing Cosmetic Outcome in Breast Conservation," *Clinical Oncology (Royal College of Radiologists)*, vol. 21, pp. 285-293, 2009, [[10.1016/j.clon.2009.02.001](#)].
- [54] M. A. Olsen, K. B. Nickel, J. A. Margenthaler, A. E. Wallace, D. Mines, J. P. Miller, V. J. Fraser, and D. K. Warren, "Increased Risk of Surgical Site Infection Among Breast-Conserving Surgery Re-excisions," *Annals of Surgical Oncology*, vol. 22, pp. 2003-2009, 2015, [[10.1245/s10434-014-4200-x](#)].
- [55] J. Heil, K. Breitkreuz, M. Golatta, E. Czink, J. Dahlkamp, J. Rom, F. Schuetz, M. Blumenstein, G. Rauch, and C. Sohn, "Do Reexcisions Impair Aesthetic Outcome in Breast Conservation Surgery? Exploratory Analysis of a Prospective Cohort Study," *Annals of Surgical Oncology*, vol. 19, pp. 541-547, 2012, [[10.1245/s10434-011-1947-1](#)].
- [56] N. Krekel, B. Zonderhuis, S. Muller, H. Bril, H. J. van Slooten, E. de Lange de Klerk, P. van den Tol, and S. Meijer, "Excessive resections in breast-conserving surgery: a retrospective multicentre study," *Breast J*, vol. 17, no. 6, pp. 602-9, 2011, [[10.1111/j.1524-4741.2011.01198.x](#)].

- [57] C. Dahlbäck, J. Manjer, M. Rehn, and A. Ringberg, "Determinants for patient satisfaction regarding aesthetic outcome and skin sensitivity after breast-conserving surgery," *World Journal of Surgical Oncology*, vol. 14, pp. 303, 2016, [[10.1186/s12957-016-1053-8](#)].
- [58] S. E. Abe, J. S. Hill, Y. Han, K. Walsh, J. T. Symanowski, L. Hadzikadic-Gusic, T. Flippo-Morton, T. Sarantou, M. Forster, and R. L. White, "Margin Re-excision and Local Recurrence in Invasive Breast Cancer: A Cost Analysis Using a Decision Tree Model," *Journal of Surgical Oncology*, vol. 112, pp. 443-448, 2015, [[10.1002/jso.23990](#)].
- [59] E. L. Vos, J. Gaal, C. Verhoef, K. Brouwer, C. H. M. Van Deurzen, and L. B. Koppert, "Focally positive margins in breast conserving surgery: Predictors, residual disease, and local recurrence," *European Journal of Surgical Oncology*, vol. 43, pp. 1846-1854, 2017, [[10.1016/j.ejso.2017.06.007](#)].
- [60] A. L. Merrill, S. B. Coopey, R. Tang, M. P. McEvoy, M. C. Specht, K. S. Hughes, M. A. Gadd, and B. L. Smith, "Implications of New Lumpectomy Margin Guidelines for Breast Conserving Surgery: Changes in Reexcision Rates and Predicted Rates of Residual Tumor," *Annals of Surgical Oncology*, vol. 23, pp. 729-734, 2017, [[10.1245/s10434-015-4916-2](#)].
- [61] R. Tang, S. B. Coopey, M. C. Specht, L. Lei, M. A. Gadd, K. S. Hughes, E. F. Brachtel, and B. L. Smit, "Lumpectomy specimen margins are not reliable in predicting residual disease in breast conserving surgery," *The American Journal of Surgery*, vol. 210, pp. 93-98, 2015, [[10.1016/j.amjsurg.2014.09.029](#)].
- [62] R. J. Gray, B. A. Pockaj, E. Garvey, and S. Blair, "Intraoperative Margin Management in Breast-Conserving Surgery: A Systematic Review of the Literature," *Annals of Surgical Oncology*, vol. 25, pp. 18-27, 2018, [[10.1245/s10434-016-5756-4](#)].
- [63] C. Reyna, and S. M. DeSnyder, "Intraoperative Margin Assessment in Breast Cancer Management," *Surgical Oncology Clinics of North America*, vol. 27, pp. 155-165, 2018, [[10.1016/j.soc.2017.08.006](#)].
- [64] B. W. Maloney, D. M. McClatchy III, B. W. Pogue, K. D. Paulsen, W. A. Wells, and R. J. Barth Jr, "Review of methods for intraoperative margin detection for breast conserving surgery," *Journal of Biomedical Optics*, vol. 23, no. 10, pp. 100901, 2018, [[10.1117/1.JBO.23.10.100901](#)].
- [65] K. Esbona, Z. Li, and L. G. Wilke, "Intraoperative Imprint Cytology and Frozen Section Pathology for Margin Assessment in Breast Conservation Surgery: A Systematic Review," *Annals of Surgical Oncology*, vol. 19, no. 10, pp. 3236-3245, 2012, [[10.1245/s10434-012-2492-2](#)].
- [66] L. A. Newman, and H. M. Kuerer, "Advances in Breast Conservation Therapy," *Journal of Clinical Oncology*, vol. 23, no. 8, pp. 1685-1697, 2005, [[10.1200/JCO.2005.09.046](#)].
- [67] G. Lu, and B. Fei, "Medical hyperspectral imaging: a review," *Journal of Biomedical Optics*, vol. 19, no. 1, pp. 010901, 2014, [[10.1117/1.JBO.19.1.010901](#)].
- [68] B. E. Schaafsma, M. Van der Giessen, A. Charehbili, V. T. H. B. M. Smit, J. R. Kroep, V. P. F. Lelieveldt, G. Liefers, A. Chan, C. W. G. M. Löwik, J. Dijkstra, C. J. H. Van der Velde, M. N. J. M. Wasser, and A. L. Vahrmeijer, "Optical mammography using diffuse optical spectroscopy for monitoring tumor response to neoadjuvant chemotherapy in women with locally advanced breast cancer" *Clinical Cancer Research*, 2014, [[10.1158/1078-0432.CCR-14-0736](#)].
- [69] W. T. Tran, C. Childs, L. Chin, E. Slodkowska, L. Sannachi, H. Tadayyon, E. Watkins, S. L. Wong, B. Curpen, A. El Kaffas, A. Al-Mahrouki, A. Sadeghi-Naini, and G. J. Czarnota, "Multiparametric monitoring of chemotherapy treatment response in locally advanced breast cancer using quantitative ultrasound and diffuse optical spectroscopy," *Oncotarget*, vol. 7, no. 15, pp. 19762-19780, 2016.
- [70] B. J. Tromberg, Z. Zhang, A. Leproux, T. D. O'Sullivan, A. E. Cerussi, P. M. Carpenter, R. S. Mehta, D. Roblyer, W. Yang, K. D. Paulsen, B. W. Pogue, S. Jiang, P. A. Kaufman, A. G. Yodh, S. H. Chung, M. Schnall, B. S. Snyder, N. Hylton, D. A. Boas, S. A. Carp, S. J. Isakoff, and D. Mankoff, "Predicting Responses to Neoadjuvant Chemotherapy in Breast Cancer: ACRIN 6691 Trial of Diffuse Optical Spectroscopic Imaging," *Cancer Research*, vol. 76, no. 20, pp. 1-20, 2016, [[10.1158/0008-5472.CAN-16-0346](#)].

- [71] B. Yu, E. S. Burnside, G. A. Sisney, J. M. Harter, C. Zhu, A. H. Dhalla, and N. Ramanujam, "Feasibility of near-infrared diffuse optical spectroscopy on patients undergoing image-guided core-needle biopsy," *Optics Express*, vol. 15, no. 12, pp. 7335-50, 2007.
- [72] R. L. P. Van Veen, A. Amelink, M. Menki-Pluymers, C. Van der Pol, and H. J. C. M. Sterenborg, "In vivo optical biopsy measurement of local optical properties of healthy and malignant breast tissue," *Physics in Medicine and Biology*, vol. 7, no. 50(11), pp. 2573-2581, 2005.
- [73] B. S. Nichols, C. E. Schindler, J. Q. Brown, L. G. Wilke, C. S. Mulvey, M. S. Krieger, J. Gallagher, J. Geradts, R. A. Greenup, J. A. Von Windheim, and N. Ramanujam, "A Quantitative Diffuse Reflectance Imaging (QDRI) System for Comprehensive Surveillance of the Morphological Landscape in Breast Tumor Margins," *PloS One*, vol. 10, no. 16, pp. e0127525, 2015, [[10.1371/journal.pone.0127525](https://doi.org/10.1371/journal.pone.0127525)].
- [74] A. M. Laughney, V. Krishnaswamy, E. J. Rizzo, M. C. Schwab, R. J. Barth, B. W. Pogue, K. D. Paulsen, and W. A. Wells, "Scatter Spectroscopic Imaging Distinguishes between Breast Pathologies in Tissues Relevant to Surgical Margin Assessment," *Clinical Cancer Research*, vol. 18, pp. 6315-6325, 2012, [[10.1158/1078-0432.CCR-12-0136](https://doi.org/10.1158/1078-0432.CCR-12-0136)].

Review: *In vivo* optical spectral tissue sensing - how to go from research to routine clinical application?

L.L. de Boer * | J.W. Spliethoff *
H.J.C.M. Sterenborg | T.J.M. Ruers

* Joint first authorship

Innovations in optical spectroscopy have helped the technology reach a point where performance previously seen only in laboratory settings can be translated and tested in real-world applications. In the field of oncology, spectral tissue sensing (STS) by means of optical spectroscopy is considered to have major potential for improving diagnostics and optimizing treatment outcome. The concept has been investigated for more than two decades and yet spectral tissue sensing is not commonly employed in routine medical practice. It is therefore important to understand what is needed to translate technological advances and insights generated through basic scientific research in this field into clinical practice. The aim of the discussion presented here is not to provide a comprehensive review of all work published over the last decades but rather to highlight some of the challenges found in literature and encountered by our group in the quest to translate optical technologies into useful clinical tools. Furthermore, an outlook is proposed on how translational researchers could proceed to eventually have STS incorporated in the process of clinical decision making.

Keywords: Fiber-optic, optical spectroscopy, tissue diagnosis, clinical translation, progress

INTRODUCTION

The interaction of light within tissue to recognize disease has been widely researched since the mid-19th century when Joseph von Fraunhofer developed diffraction grating. A large number of scientists have brought optical spectroscopy forward and enabled it to become a precise and quantitative scientific technology. In recent years, improvements made to spectrometers, software and overall design positively affected instrument characteristics such as speed, sensitivity, size, and price. These technological advances, combined with increased awareness of the potential of optical spectroscopy have led to the development of optical systems which were usable for clinical research purposes and allowed further exploration of the potential applications. In the medical field, the technology has gained interest for numerous biomedical applications for its advantages over existing conventional techniques. Optical spectroscopy at infrared and visible wavelengths avoids the use of ionizing radiation, is non-destructive, utilizes relatively inexpensive equipment and can be performed near real-time without pharmaceutical means to enhance contrast, i.e. contrast agents. Several spectroscopic techniques have been applied for tissue characterization, including; diffuse reflectance spectroscopy (DRS); autofluorescence spectroscopy (FS); elastic scattering spectroscopy (ESS); and Raman Spectroscopy (RS). All of these methods rely on the same underlying principle: tissue characterization is performed by measuring the spectral response after the tissue is illuminated with a selected spectral band of light. This spectral response contains specific quantitative morphologic, compositional, and functional information about the probed tissue, thereby enabling tissue discrimination. Especially in the field of oncology, characterization of human tissues by optical spectroscopy (spectral tissue sensing; STS) is considered to have major potential. Various STS methods already have been used for tissue assessment in several organs for years with promising results.[1-11]

HOW TO TRANSFER OPTICAL SPECTROSCOPY RESEARCH TO CLINICAL PRACTICE

Despite these technological advancements and promising results, fiber-optic STS technology has not (yet) found widespread clinical acceptance in medical practice. This gap between original research and final clinical implementation is not specific for STS. It is considered to be one of the key aspects of translational research in general, as it can take more than two decades before the findings of original research become part of routine clinical practice. [12, 13] However, awareness of the causes of the gap and potential bridging strategies could help the field of STS avoiding unnecessary delay in the evolution of this technology into clinically useful tools.

For example, research in this field is often not focused on the potential application or clinical relevance.[14] This might not come across as a problem since indeed, many transformative discoveries leading to inventions that are now used in daily clinical practice did not originate from application-directed research.[15, 16] However, in the case of STS, the field has become mature technologically and reached a point where spectral tissue sensing in its current form is ready to enter the clinical arena. Since translating novel technologies into clinical applications is a very time-consuming and expensive process, it is important to assess the clinical relevance of an application and keep this under constant review during clinical research as this will eventually be a dominating force in the process of successful implementation in the clinic.[17, 18] Furthermore, the success of an intended application will also partly depend on the extent of additional improvements in detection sensitivity and specificity in combination with instrument cost. Expensive techniques that are slightly better compared to well-established methods will be more difficult to implement in the hospital. It can be highly rewarding to perform a cost-benefit analysis in a very early stage and understand the preferences and drivers of the various stakeholders (e.g. patients, health insurances, health care providers, regulatory bodies).[16, 19] In addition to the clinical relevance, also the extent to which the clinical workflow needs to be adapted is a major factor determining the feasibility of a new method in clinical practice. Techniques that are not compatible with the existing clinical routines are likely to encounter more resistance from physicians during implementation than those that can easily be added to the accustomed process. Moreover, modification of the accepted workflow can generate secondary effects on the outcome of the procedure that may counteract the intended benefits. On the other hand, when a new method significantly improves procedure outcome or shortens procedure time, disrupting the existing clinical workflow is likely to be a smaller obstacle. Researchers should be aware of this important trade-off in an early stage of clinical research.

Besides these clinical aspects, it should be noted that optical spectroscopy is a very broad concept including many settings and configurations. The range of different optical technologies and available hardware choices lead to a complicated process of convergence upon the optimal system for each specific clinical need. For example, for reflectance spectroscopy, the distance between the source and detection fibers to a large extent defines the probing depth and spatial resolution and therefore determines the feasibility of a system for a certain clinical application. Since challenges in the biomedical field are extremely versatile, developing a universal solution applicable to each clinical problem is unlikely to be possible. As a consequence, the clinical success of optical spectroscopy, in general, is highly dependent on finding the best modality with appropriate optical geometry for a specific application.

Seeking in-depth collaboration with medical physicians may help basic and translational scientists to acquire a greater understanding of both the medical aspects and the optics of the clinical problem, and design a solution appropriate for the specific clinical problem.[19,

20] Comprehensive collaboration is not only important to elucidate the relevant details of the clinical problem[18], but also for tackling other non-scientific challenges (i.e. previously mentioned barriers such as workflow issues and clinical acceptance as well as regulation issues) that translational researchers are faced with.[21, 22] Invariably, extensive expertise is needed as technology develops from the preclinical stage towards increasingly complex and demanding phases of clinical testing, where patient safety and clinical usability become major determinants for progress. Furthermore, similar to the process of translating drug discoveries, the participation of industrial partners, especially during testing in clinical trials, can help financing the often costly trials and help to reflect on the feasibility of implementing an application in terms of cost-effectiveness.[23, 24] Therefore, close collaboration and continuous communication between researchers (basic and translational), clinicians, and industry professionals are required, even if individuals' needs, motivations, and research attitudes may differ.[21, 24-26] Because the perspectives of co-operating partners may be different and might even change during the development of an application it is difficult to set up a network of co-operating partners. However, this interactive multidisciplinary approach is the only way to properly match technology and clinical problem and give direction to clinically valuable research which can ultimately lead to an accepted clinical application.

Translating technology into clinical applications is not confined to taking into account the previously described non-technical barriers and establishing an optimal environment for translational research. Conducting translational optical research in a medical environment brings some typical challenges and potential pitfalls that are likely to be universal across the biomedical optical research field. To our knowledge, few investigators have undertaken the task of describing these key challenges. In the following sections, we attempt to provide an overview of the practical challenges that were found in literature as well as those that the authors encountered during nearly a decade of clinical testing and evaluation. The first four key challenges presented are related to the struggles encountered during the acquisition of reliable optical data and interpretation of it. The last key challenge discusses how fiber-optic STS can impact clinical decision making which is the ultimate goal of translational research.

KEY CHALLENGES

Proper study design

Obviously, the design of a study is important as it influences the quality and relevance of study results. Preclinical studies on animals or *ex vivo* measurements on human specimens may serve as a surrogate for *in vivo* measurements and usually provide useful information in the first phase of development. However, when it comes to human disease applicability, they do not necessarily reflect the *in vivo* status of human tissue for all applications. For

example, given the differences between *in vivo* and *ex vivo* tissues, validating a database from *ex vivo* data and applying it to *in vivo* data may not work as various tissue parameters (e.g. blood content, oxygenation) might change once the tissue has been removed from the patient's body.[27] On the contrary, research also suggests that these differences are limited to the blood-related wavelengths and for example, the near-infrared wavelengths are not affected by tissue status. Furthermore, *in vivo* studies gathering spectral tissue data in a well-controlled setting can be limited because this data might not be representative for a real-world situation. For instance, estimates for blood content and associated oxygenation levels might show promising differences when measured *in vivo* during surgery. However, once measured during a percutaneous biopsy procedure, these parameters do not necessarily reflect the true physiological composition of the measured tissue due to pooling of blood around the needle tip.[28] Similarly, blood present on top of the surface of a resection margin may be a major obstacle for resection margin assessment by STS.[29] All these circumstances should be taken into consideration when interpreting measurements but also in a later stage during the designing and constructing of a clinical STS tool.

Much of the research reported to date has been based on observational studies, and a large portion of the scientific knowledge in this field comes from analyzing spectroscopic data observational studies. Data from such studies may help to prove clinical feasibility and develop new trials by determining sample size requirements and optimal design but may be confounded when data used for algorithm training is also used for validation.

Once early-phase human studies have been conducted, research efforts generally move to larger clinical trials. Such studies and subsequent clinical implementation require that the equipment and analytical algorithms have been optimized and fitted to be used in the clinical routine. Especially in this stage of clinical validation, it is important that the basic principles and potential pitfalls of diagnostic test development are well understood. For instance, as clearly elaborated by M. Fitzmaurice[30], unintentional bias in selecting patients for study groups may lead to the conclusion that a new optical technique is a better or worse diagnostic tool than it really is. Furthermore, measures of test performance, such as sensitivity and specificity are not only influenced by the definition of the threshold between positive and negative results but also disease prevalence. For instance, if a new diagnostic tool is tested in a high prevalence setting, it is more likely that persons who test positive, truly have the disease than if the test is performed in a population with low prevalence. In addition, selection of patients based on visual clues may lead to serious optical bias in a dataset. For instance, in a study described by de Veld et al. on oral cancer an apparently excellent sensitivity and specificity for distinguishing a visually detected suspect lesion from normal mucosa was reported, while the clinically relevant question: can we distinguish malignant lesions from benign lesions could not be answered.[31]

Thus, researchers have to concede the restrictive nature of testing in a preclinical and in well-controlled clinical settings when it comes to real-world applicability. Eventually,

performing well-designed large-scale clinical trials will help to gain trust in the actual performance of new optical technologies.

Co-registration between spectral measurements and histology

Histology remains the gold standard for diagnosis and staging in oncology, making it, in essence, a ground truth for new technologies attempting to measure and quantify these pathologies. Once registered, histopathology provides valuable information regarding the structure and composition of tissues and is often used to understand the physiological mechanisms behind spectral contrasts.

Consequently, registration of histology serves as an important and necessary validation step for novel optical spectroscopy based diagnostic techniques. However, the use of histopathology as the benchmark does suffer from a number of critical limitations. Co-registration of separate optical and physical biopsies is subject to imperfect spatial correlation of the optical reading and the tissue sample removed for histopathological analysis.[30, 32] Several methods have been employed to achieve sufficient accurate co-registration. To facilitate registration, usually, a tissue sample is removed from the measurement spot after spectral measurements are completed or a marker is left behind, such as a dye or suture. These methods are susceptible to location mismatch that can be particularly challenging in the case of micro-environmental heterogeneity. Even more difficulties in co-registration arise when the tissue site being investigated is not directly accessible to the operator. For example, in cases that *in vivo* measurements are obtained through the working channel of an endoscope or through a hollow needle at deep located tissue. In these cases, the tissue sample for histopathological evaluation is generally obtained by a second needle after retracting the fiber-optic probe used for spectral measurements. Linking spectral data with such a “best estimate” of the spectral measurement spot leads to inherent registration inaccuracies.

An approach to address discrepancies between spectral data and histology is to integrate fiber-optics into standard tissue-sampling tools, such as a core biopsy needle or endoscopic biopsy forceps[32-34], thereby linking spectral tissue sensing and biopsy functionality in a single instrument. A few examples are shown in Figure 1.

These integrated fiber-optic instruments represent a major step forward for clinical evaluation of new spectral tissue sensing techniques by greatly increasing the spatial correlation of physical biopsies with spectral measurement spots and simplifying study procedures. Beyond validation studies, the developed integrated fiber-optic tools could be clinically useful for increasing the pre-biopsy probability of obtaining representative tissue samples in diagnostic biopsy procedures (Figure 2).

Despite (almost) perfect co-registration between the optical and physical biopsy with integrated fiber-optic tools, the histopathological analysis still is fundamentally limited in accuracy. With the use of fiber-optic probes tissue volumes of approximately 1 mm³ can be

Figure 1. Integrated fiber-optics in clinical biopsy tools allow 1:1 correlation between spectral data and biopsy sample. (a) The WavSTAT biopsy forceps for optical diagnosis based on laser-induced autofluorescence spectroscopy. Image courtesy of SpectraScience Inc., San Diego, California, USA). (b) Elastic scattering spectroscopy through an integrated endoscopic tool for use in a range of applications suitable to the upper gastrointestinal (GI) tract and the colon. Image courtesy of I.J. Bigio, Boston University. (c) Added quantitative spectral functionality during routine percutaneous biopsy procedures using a fiber-optic core biopsy needle (Philips Research, Eindhoven, the Netherlands).

Figure 2. Example of spectral tissue sensing functionality integrated into a “smart” clinical instrument. Real-time tissue characterization of the tissue at the needle tip is performed during lung biopsy, thereby providing guidance to the physician. This could help to increase successful biopsy yield. Image courtesy of Clinical Cancer Research. [28]

interrogated, whereas usually, the conventional biopsy samples are larger than this. Another complicating factor is the fact that histopathological evaluation of specimens is performed on the two-dimensional histological section whereas optical measurements yield information of three-dimensional tissue volumes. Finding and orienting a representative two-dimensional histological section in a three-dimensional *ex vivo* tissue volume can,

therefore, be challenging. Especially when taken into consideration that during the processing of the tissue deformation and cutting artifacts are also common.[35]

To make the problem even more complex, histopathological tissue diagnosis by pathologists is subject to significant inter- and intra-observer variability. This is a particularly difficult problem in the diagnosis and grading of dysplasia, a premalignant lesion seen in patients at high risk for development of carcinoma.

Furthermore, in the case of STS, the acquired information from the tissue is from a different nature compared to the information provided by histopathologic assessment. For example, STS can provide quantified measures of substances present in tissue whereas the pathologist regards tissue in terms of normal or tumorous. Classifying STS measurements based on these historically developed pathologic definitions can lead to erroneous classification. This mismatch between information types requires translational scientists to fully comprehend the limitations and biases of the gold standard as it will seriously hamper clinical acceptance of a novel method.

Thus, the apparent diagnostic performance of a new technology will, therefore, depend not only on its ability to detect abnormalities but also on the (in)accuracy of the gold standard. Even if a novel technology is 100% sensitive and 100% specific, it may appear inaccurate when either the golden standard is imperfect or incorrectly applied. Therefore, careful attention should be paid to the adequacy of tissue samples, correlation with measured tissue sites, and consistency of pathology reporting terminology. Although the comparison between STS and histopathology is challenged in many ways, this should not be interpreted or used as an argument to stop the efforts of translating STS to the clinic.

Inter-patient and intra-patient variation

Characterizing and differentiating between various tissues by an optical spectroscopy system relies on the measurement of the absolute or relative differences in intensity or spectral contrasts between the tissue types of interest. However, sources of variation such as inter- and intra-patient variability may outweigh the difference between the tissue types of interest leading to hampered diagnostic performance. Tumor tissue is well known for its heterogeneity[36], but also in healthy tissues non-uniformity is a common phenomenon[37, 38]. In particular breast tissue is a well-known example of inter and intra-patient variation. [39, 40] The optical contrast between healthy and tumorous tissue can be manipulated by other sources of contrast which are related to biological processes such as menopausal status and the phase of the menstrual cycle.[41] This does not necessarily have to be the case, but researchers should be aware of these potential disturbing factors. Including patient demographics to the databases with reference measurements can circumvent the unwanted blurring of optical contrast by sources not related to the difference between normal and tumor tissue.[39] However, this might also require further extending the database with a comprehensive amount of reference measurements to comprise the influence of all patient

characteristics. Another approach is to measure optical characteristics at a distant location from the tissue site under evaluation and use these measurements as an internal reference. In this case, instead of the absolute values, relative changes between tissue sites are used. Taroni et al.[42], Laughney et al.[43] and Spliethoff et al.[44] found that such methods could be applied to account for the tissue- and patient heterogeneity in breast and lung tissue. This indicates that using the patient's measurement as its own reference is effective in reducing the influence of inter-patient variation. Such a method might be very useful in a clinical setting since heterogeneity is also reported in studies of other tissues such as prostate and nasopharyngeal tissue.[45, 46]

For biopsy procedures, STS measurements can also be performed in a continuous mode. Opposed to points measurements of the different tissue types a continuous series of measurements has the advantage of providing a full overview of the tissue characterization along the needle path and allows an evaluation of the local changes, rather than a comparison to a cohort-based reference value. Recently, Nachabé et al. demonstrated the potential of real-time tissue characterization by diffuse optical spectroscopy measurements at the tip of a needle during percutaneous interventions.[47] Such continuous measurements may be of great relevance during percutaneous procedures because they enable detection of the transition from healthy tissue to tumor based on the clinical parameters derived from STS measurements.

Thus, several investigators have shown the benefit of deriving relative diagnostic criteria by calculating differential spectral tissue parameters between the tumor and a reference tissue. The use of relative tissue parameters may help to define more effective detection criteria that are less sensitive to inter-patient variations and tissue heterogeneities.

Spectral analysis and classification methods

There is no substitute for good study design and proper data collection, but after proper data collection, pre-processing of raw spectral data is often an important subsequent step. The goal of pre-processing of raw spectral data is to apply the calibrations to remove device-specific features of the data. After preprocessing, the data can be interpreted by performing further analysis on the spectral data which ultimately enables classification of the measured tissue sites.

One method for tissue classification is to use pattern recognition by supervised machine learning methods. This method correlates spectral features directly to tissue type labels, thereby enabling identification of unknown tissue based on the measured spectra without the need for extensive data processing. Various approaches of spectral analysis have been developed, including multivariate statistical data analysis[3], partial least squares discriminant analysis[48], support vector machines and statistical learning[48], k-nearest neighbor classification[49], as well as neural network methods[50, 51]. Due to the nature of the supervised machine-learning approaches, the diagnostic performance is usually

compromised when the number of classes increases. Theoretically, discriminating tumor from surrounding healthy tissue is a two-class problem (healthy versus malignant). However, in practice tumor tissue is surrounded by many tissue types, both healthy and diseased, and this demands identification of a wider variety of tissue pathologies, including histological tumor subtypes, inflammation, and fibrosis. Having to deal with this tissue heterogeneity poses a true challenge in the development of machine learning approaches.[52, 53] To overcome this difficulty a comprehensive high-quality spectral database is required containing any healthy or diseased tissue type that is likely to be encountered during a clinical procedure.[54]

Several investigators have analyzed the spectral information in a different manner by applying mathematical models based on knowledge of light propagation in tissue to determine the scattering and absorption properties within the tissue.[1, 55, 56] Each measured spectra is, in fact, a combination of the spectral signatures of various tissue chromophores present in the probed volume. Meaningful estimation of these components can be extracted by using the known absorption spectra of these biologically relevant chromophores. The main benefit of this approach is that it uses a priori knowledge of light-tissue interaction that can help to understand the underlying biological composition and physiological processes that determine the spectral shape. Over the years, the differences in composition between malignant and normal tissues have been extensively researched by this approach. Understanding which biological substances could potentially play a role in the identification of abnormal tissue may be very useful. In our own group, we found that the biological chromophore bile is significantly higher in healthy liver tissue as opposed to colorectal liver metastases.[5, 57] In the same way breast tumor tissue has shown to contain far less fat and more water compared to surrounding healthy tissue.[58] The often-used diffusion theory has proven to be sufficiently realistic to describe light transport in many different human tissues. But it should be noticed that these diffusion theory based fit models require that the optical properties of the tissue are either homogeneous or close to homogeneous as well as sufficient source-detector distance. Care should be taken when basic assumptions of the theory are violated. Highly inhomogeneous tissue (for example layered tissue) or tissue with high absorption and low scattering (for example when a substantial amount of blood is present) may result in unrealistic estimations of the tissue composition. And, even when the criteria are met, a realistic estimation of the true composition of the tissue can only be obtained when all the absorption spectra of tissue chromophores actually present in the measurement are included in this process. However, adding absent or too many chromophores in this method may lead to erroneous estimations of the tissue characteristics.

Over the years, a variety of advanced computational methods has been developed to improve the quantitative and qualitative diagnostic capability of spectral tissue sensing. The occurrence and co-occurrence of specific spectral signatures depend on the underlying

physical, morphological, and biochemical tissue composition. Whatever method is used for interpretation of the spectral data, expert knowledge should be employed throughout the various stages of clinical research such that robust tissue discrimination algorithms can be developed.

Impacting clinical decision making

The objective of each new technology is to support clinicians in reaching a decision regarding diagnosis or treatment of a disease and ultimately fundamentally improve patient care. In order to help in the process of decision making it is important to consider at which point in the workflow the technology can play a central role, based on its strengths and weaknesses. The main advantage of STS methods that use endogenous contrast is the possibility to perform real-time tissue characterization in a non-destructive manner at multiple points in time. This enables to obtain optical measurements without obstructing pathological examination protocols and allows optical tissue sensing to be employed as a pre-pathology modality by surgeons and radiologist while performing a procedure. Furthermore, since the technology can be integrated into needles or probes, STS can be combined with other complementary modalities. In combination with the versatility of configurations that STS can have this permits a wide range of different applications. These strong qualities provide STS with the power to make a difference in cancer treatment.

As mentioned previously, spectral tissue sensing technologies have been researched for several oncological applications. The question is which applications at which point in oncological care (e.g. prevention, screening, diagnosis, treatment, management, surveillance) have the greatest potential to contribute substantially to clinical decision making. Application of STS in different stages of this process also demands different requirements in order to provide useful input for clinical decision making. For example, a screening application is likely to affect the entire care path of a patient. This can improve the prognosis for the patient tremendously since successful early detection of cancer enables curative treatment and hence, result in a considerable gain for the patient.[59] Unfortunately, the implementation of early detection methods is also challenging since the information provided by the new technology can alter the entire workflow of patient care as a consequence. Introducing an end-of-line application on the other hand, will likely change the procedure of cancer management less radically and by leaving the current workflow intact the much needed clinical acceptance can be acquired faster.

Another important consideration when preparing a tool for clinical use is the trade-off between sensitivity and specificity. Depending on the intended application, higher sensitivity or specificity is desired. Furthermore, in order to optimize the diagnostic performance, it may be necessary to adjust the decision threshold, based on the specific type of patient studied. The same diagnostic tool may need different decision thresholds

when used in high-risk patients (high disease prevalence) than when used as a screening test in the general population (low disease prevalence).[30]

Ultimately, cost-effectiveness and the impact on procedure outcome are major important considerations for successful clinical application. It is crucial to view the expected added clinical value of a new clinical method in the context of the evolving landscape of medical practice (need for increased productivity and improved patient outcomes) and the macroeconomic health care environment (cost constraints). From a clinical application perspective, it makes sense to start with the development of applications that can be used as adjunct tools to other established imaging modalities in oncology. Because STS can be integrated into familiar clinical instruments, it could be easily adopted as an “add-on” to the existing protocols with minimal impact on procedure flow. For instance, in the case of biopsy guidance, STS can be used in conjunction with conventional imaging modalities thereby exploiting the complementary strengths of each method. We anticipate that STS will also find its way to clinical settings where no good tools exist today, especially in situations where under-treatment has a significant negative impact. For example, in the case of guided surgery, STS has the potential to influence “on-table” decision-making during treatment.

OUTLOOK

Spectral tissue sensing research holds tremendous promise for the development of novel diagnostic and therapeutic tools for many oncological applications. After a quarter century of rapid advances, spectroscopic research has generated a rich body of ideas, insights, and discoveries as well as the necessary technical improvements. As in all areas of medicine, the maturation of an experimental, early-phase technological concept into a useful clinical tool is a long and complex process usually involving many years of rigorous preclinical and clinical testing and many setbacks and failures. The goal of this article is to raise awareness among the optical community about some key challenges that should be acknowledged in the process of translating STS technology towards the clinic, and thus contribute to the clinical implementation of STS. In order to realize the potential of spectral tissue sensing, there are still hurdles to overcome to bridge the gap between technological advances and clinical practice. These issues are clearly surmountable although this does require effort from the entire scientific, medical, and industrial community. Without any claim of completeness, we highlighted various challenges faced by translational researchers as they move from the proof-of-concept stage through the translational stage and into the clinical setting. We believe that a broad-based, multidisciplinary effort and in-depth collaboration between biomedical researchers and clinicians as well as the involvement of industrial partners with a clear focus on a well-defined clinical problem can help to address these key challenges.

REFERENCES

- [1] A. Cerussi, N. Shah, D. Hsiang, A. Durkin, J. Butler, and B. Tromberg, "In vivo absorption, scattering, and physiologic properties of 58 malignant breast tumors determined by broadband diffuse optical spectroscopy," *Journal of Biomedical Optics*, vol. 11, no. 4, pp. 044005-1 - 044005-16, 2006, [[10.1117/1.2337546](#)].
- [2] J. Q. Brown, L. G. Wilke, J. Geradts, S. A. Kennedy, G. M. Palmer, and N. Ramanujam, "Quantitative optical spectroscopy: a robust tool for direct measurement of breast cancer vascular oxygenation and total hemoglobin content in vivo," *Cancer Research*, vol. 69, no. 7, pp. 2919-26, 2009, [[10.1158/0008-5472.CAN-08-3370](#)].
- [3] Z. Volynskaya, A. S. Haka, K. L. Bechtel, M. Fitzmaurice, R. Shenk, N. Wang, J. Nazemi, R. R. Dasari, and M. S. Feld, "Diagnosing breast cancer using diffuse reflectance spectroscopy and intrinsic fluorescence spectroscopy," *Journal of Biomedical Optics*, vol. 13, no. 2, pp. 024012, 2008, [[10.1117/1.2909672](#)].
- [4] D. J. Evers, R. Nachabé, M. J. Vranken Peeters, J. A. Van der Hage, H. S. Oldenburg, E. J. Rutgers, G. W. Lucassen, B. H. W. Hendriks, J. Wesseling, and T. J. M. Ruers, "Diffuse reflectance spectroscopy: towards clinical application in breast cancer," *Breast Cancer Research and Treatment*, vol. 137, pp. 155-165, 2013, [[10.1007/s10549-012-2350-8](#)].
- [5] D. J. Evers, R. Nachabé, D. Hompes, F. van Coevorden, G. W. Lucassen, B. H. W. Hendriks, M.-L. F. van Velthuisen, J. Wesseling, and T. J. M. Ruers, "Optical sensing for tumor detection in the liver," *EJSO*, vol. 39, pp. 68-75, 2013, [[10.1016/j.ejso.2012.08.005](#)].
- [6] G. Zonios, A. Dimou, M. Carrara, and R. Marchesini, "In Vivo Optical Properties of Melanocytic Skin Lesions: Common Nevi, Dysplastic Nevi and Malignant Melanoma," *Photochemistry and Photobiology*, vol. 86, pp. 236-240, 2010, [[10.1111/j.1751-1097.2009.00630.x](#)].
- [7] N. Rajaram, J. S. Reichenberg, M. R. Migden, T. H. Nguyen, and J. W. Tunnell, "Pilot clinical study for quantitative spectral diagnosis of non-melanoma skin cancer," vol. 42, no. 10, pp. 876-887, 2010, [[10.1002/lsm.21009](#)].
- [8] O. M. A'Amar, L. Liou, E. Rodriguez-Diaz, A. De las Morenas, and I. J. Bigio, "Comparison of elastic scattering spectroscopy with histology in ex vivo prostate glands: potential application for optically guided biopsy and directed treatment," *Lasers in Medical Science*, vol. 28, pp. 1323-1329, 2013, [[10.1007/s10103-012-1245-6](#)].
- [9] V. T.-C. Chang, S. M. Bean, P. S. Cartwright, and N. Ramanujam, "Visible light optical spectroscopy is sensitive to neovascularization in the dysplastic cervix," *Journal of Biomedical Optics*, vol. 15, no. 5, pp. 057006, 2010, [[10.1117/1.3495730](#)].
- [10] M. P. L. Bard, A. Amelink, M. Skurichina, V. Noordhoek Hegt, R. P. W. Duin, H. J. C. M. Sterenborg, H. C. Hoogsteden, and J. G. J. V. Aerts, "Optical Spectroscopy for the Classification of Malignant Lesions of the Bronchial Tree," *Chest*, vol. 129, pp. 995-1001, 2006.
- [11] I. J. Bigio, and S. G. Bown, "Spectroscopic Sensing of Cancer and Cancer Therapy," *Cancer Biology & Therapy*, vol. 3, no. 3, pp. 259-267, 2004.
- [12] J. M. Westfall, J. Mold, and L. Fagnan, "Practice-Based Research—"Blue Highways" on the NIH Roadmap," *The Journal of the American Medical Association*, vol. 297, no. 4, pp. 403-406, 2007.
- [13] S. R. Hanney, S. Castle-Clarke, J. Grant, S. Guthrie, C. Henshall, J. Mestre-Ferrandiz, M. Pistollato, A. Pollitt, J. Sussex, and S. Wooding, "How long does biomedical research take? Studying the time taken between biomedical and health research and its translation into products, policy, and practice," *Health Research Policy and Systems*, vol. 13, no. 1, 2015, [[10.1186/1478-4505-13-1](#)].
- [14] J. P. A. Ioannidis, "Evolution and Translation of Research Findings: From Bench to Where?," *PLoS Clinical Trials*, vol. 1, no. 7, pp. e36, 2006, [[10.1371/journal.pctr.0010036](#)].
- [15] N. B. Vale, J. Delfino, and L. F. B. Vale, "Serendipity in Medicine and Anesthesiology," *Revista Brasileira de Anestesiologia*, vol. 55, no. 2, pp. 224-249, 2005.
- [16] W. F. Crowley, "Translation of Basic Research into Useful Treatments: How Often Does It Occur?," *The American Journal of Medicine*, vol. 114, pp. 503-505, 2003, [[10.1016/S0002-9343\(03\)00119-0](#)].

- [17] J. S. Fernandez-Moure, "Lost in Translation: The Gap in Scientific Advancements and Clinical Application," *Frontiers in Bioengineering and Biotechnology*, vol. 4, 2016, [[10.3389/fbioe.2016.00043](#)].
- [18] P. Rousche, D. M. Schneeweis, E. J. Perreault, and W. Jensen, "Translational neural engineering: multiple perspectives on bringing benchtop research into the clinical domain," *Journal of Neural Engineering*, vol. 5, no. 1, pp. 16-20, 2008, [[10.1088/1741-2560/5/1/P02](#)].
- [19] A. Wagstaff, "Five steps to putting innovation at the heart of cancer care," *Cancer World*, no. January-February, pp. 25-33, 2014.
- [20] N. H. G. M. Peters, S. van Esser, M. A. A. J. van den Bosch, R. K. Storm, P. W. Plaisier, T. van Dalen, S. C. E. Diepstraten, R. Weits, P. J. Westenend, G. Stapper, M. A. Fernandez-Gallardo, I. H. M. Borel Rinkes, R. van Hillegersberg, W. P. T. M. Mali, and P. H. M. Peeters, "Preoperative MRI and surgical management in patients with nonpalpable breast cancer: The MONET – Randomised controlled trial," *European Journal of Cancer*, vol. 47, pp. 879-886, 2011.
- [21] N. Fudge, E. Sadler, H. R. Fisher, J. Maher, C. D. A. Wolfe, and C. McKeivitt, "Optimising Translational Research Opportunities: A Systematic Review and Narrative Synthesis of Basic and Clinician Scientists' Perspectives of Factors Which Enable or Hinder Translational Research," *PloS One*, vol. 11, no. 8, pp. e0160475, 2016, [[10.1371/journal.pone.0160475](#)].
- [22] E. M. Meslin, A. Blasimme, and A. Cambon-Thomsen, "Mapping the translational science policy 'valley of death,'" *Clinical and Translational Medicine*, vol. 2, no. 14, 2013, [[10.1186/2001-1326-2-14](#)].
- [23] P. D. Wagner, and S. Sirvastava, "New Paradigms in Translational Science Research in Cancer Biomarkers," *Translational Research: The Journal of Laboratory and Clinical Medicine*, vol. 159, no. 4, pp. 343-353, 2012, [[10.1016/j.trsl.2012.01.015](#)].
- [24] J. P. A. Ioannidis, "Materializing research promises: opportunities, priorities and conflicts in translational medicine," *Journal of Translational Medicine*, vol. 2, no. 5, 2004.
- [25] S. C. Johnston, S. L. Hauser, and S. Desmond-Hellmann, "Enhancing ties between academia and industry to improve health," *Nature Medicine*, vol. 17, no. 4, pp. 434-436, 2011.
- [26] H. Madry, M. Alini, M. J. Stoddart, C. Evans, T. Miclau, and S. Steiner, "Barriers and strategies for the clinical translation of advanced orthopaedic tissue engineering protocols," *European Cells and Materials*, vol. 27, pp. 17-24, 2014, [[10.222B0a3r/reieCrMsa.vn0d2s7tsraa0te4g](#)].
- [27] G. M. Palmer, C. L. Marshak, K. M. Vrotsos, and N. Ramanujam, "Optimal Methods for Fluorescence and Diffuse Reflectance Measurements of Tissue Biopsy Samples," *Lasers in Surgery and Medicine*, vol. 30, pp. 191-200, 2002, [[10.1002/lsm.10026](#)].
- [28] J. W. Spliethoff, W. Prevoo, M. A. Meier, J. de Jong, D. J. Evers, H. J. C. M. Sterenborg, G. W. Lucassen, B. H. W. Hendriks, and T. J. M. Ruers, "Real-time *in vivo* tissue characterization with diffuse reflectance spectroscopy during transthoracic lung biopsy: a clinical feasibility study," *Clinical Cancer Research*, vol. 22, no. 2, pp. 357-65, 2016, [[10.1158/1078-0432.CCR-15-0807](#)].
- [29] W.-C. Lin, S. A. Toms, D. Jansen, and A. Mahadevan-Jansen, "Intraoperative Application of Optical Spectroscopy in the Presence of Blood," *IEEE Journal on Selected Topics in Quantum Electronics*, vol. 7, no. 6, pp. 996-1003, 2001.
- [30] M. Fitzmaurice, "Principles and pitfalls of diagnostic test development: implications for spectroscopic tissue diagnosis," *Journal of Biomedical Optics*, vol. 5, no. 2, pp. 119-130, 2000.
- [31] D. C. G. De Veld, M. Skurichina, M. J. H. Witjes, R. P. W. Duin, H. J. C. M. Sterenborg, and J. L. N. Roodenburg, "Clinical study for classification of benign, dysplastic, and malignant oral lesions using autofluorescence spectroscopy," *Journal of Biomedical Optics*, vol. 9, no. 5, pp. 940-950, 2004, [[10.1117/1.1782611](#)].
- [32] E. Rodriguez-Diaz, I. J. Bigio, and S. K. Singh, "Integrated optical tools for minimally invasive diagnosis and treatment at gastrointestinal endoscopy," *Robot Computer Integr Manuf*, vol. 27, no. 2, pp. 249-256, 2011, [[10.1016/j.rcim.2010.06.006](#)].

- [33] E. Rodriguez-Diaz, Q. Huang, S. R. Cerda, M. J. O'Brien, I. J. Bigio, and S. K. Singh, "Endoscopic histological assessment of colonic polyps by using elastic scattering spectroscopy," *Gastrointestinal Endoscopy*, vol. 81, no. 3, pp. 539-547, 2015, [[10.1016/j.gie.2014.07.012](#)].
- [34] T. Kuiper, Y. A. Alderlieste, K. M. A. J. Tytgat, M. S. Vlug, J. A. Nabuurs, B. A. Bastiaansen, M. Löwenberg, P. Fockens, and E. Dekker, "Automatic optical diagnosis of small colorectal lesions by laser-induced autofluorescence," *Endoscopy*, vol. 47, pp. 56-62, 2015, [[10.1055/s-0034-1378112](#)].
- [35] L. H. Chen, H. Ho, R. Lazaro, C. H. Thng, J. Yuen, W. S. Ng, and C. Cheng, "Optimum slicing of radical prostatectomy specimens for correlation between histopathology and medical images," *The International Journal for Computer Assisted Radiology and Surgery*, vol. 5, pp. 471-487, 2010, [[10.1007/s11548-010-0405-z](#)].
- [36] A. Marusyk, V. Almendro, and K. Polyak, "Intra-tumour heterogeneity: a looking glass for cancer?," *Nature Reviews: Cancer*, vol. 12, no. 5, pp. 323-334, 2012, [[10.1038/nrc3261](#)].
- [37] N. Shah, A. E. Cerussi, D. Jakubowski, D. Hsiang, J. Butler, and B. Tromberg, "Spatial variations in optical and physiological properties of healthy breast tissue," *Journal of Biomedical Optics*, vol. 9, no. 3, pp. 534-540, 2004, [[10.1117/1.1695560](#)].
- [38] T. Gabrecht, B. Lovisa, H. Van den Bergh, and G. Wagnières, "Autofluorescence bronchoscopy: quantification of inter-patient variations of fluorescence intensity," *Lasers in Medical Science*, vol. 24, pp. 45-51, 2009, [[10.1007/s10103-007-0518-y](#)].
- [39] S. Kennedy, J. Geradts, T. Bydlon, J. Q. Brown, J. Gallagher, M. Junker, W. Barry, N. Ramanujam, and L. Wilke, "Optical breast cancer margin assessment: an observational study of the effects of tissue heterogeneity on optical contrast," *Breast Cancer Research*, vol. 12, no. 6, pp. R91, 2010, [[10.1186/bcr2770](#)].
- [40] T. D. O'Sullivan, A. Leproux, J. Chen, S. Bahri, A. Matlock, D. Roblyer, C. E. McLaren, W. Chen, A. E. Cerussi, M. Su, and B. J. Tromberg, "Optical imaging correlates with magnetic resonance imaging breast density and reveals composition changes during neoadjuvant chemotherapy," *Breast Cancer Research*, vol. 15, no. R14, 2013.
- [41] B. W. Pogue, S. Jiang, H. Dehghani, C. Kogel, S. Soho, S. Srinivasan, X. Song, T. D. Tosteson, S. P. Poplack, and K. D. Paulsen, "Characterization of hemoglobin, water, and NIR scattering in breast tissue: analysis of intersubject variability and menstrual cycle changes," *Journal of Biomedical Optics*, vol. 9, no. 3, pp. 541-552, 2004, [[10.1117/1.1691028](#)].
- [42] P. Taroni, A. Pifferi, G. Quarto, L. Spinelli, A. Torricelli, F. Abbate, N. Balestreri, S. Ganino, S. Menna, E. Cassano, and R. Cubeddu, "Effects of tissue heterogeneity on the optical estimate of breast density," *Biomedical Optics Express*, vol. 3, no. 10, pp. 2411-2418, 2012.
- [43] A. M. Laughney, V. Krishnaswamy, E. J. Rizzo, M. C. Schwab, R. J. Barth, B. W. Pogue, K. D. Paulsen, and W. A. Wells, "Scatter Spectroscopic Imaging Distinguishes between Breast Pathologies in Tissues Relevant to Surgical Margin Assessment," *Clinical Cancer Research*, vol. 18, pp. 6315-6325, 2012, [[10.1158/1078-0432.CCR-12-0136](#)].
- [44] J. W. Spliethoff, L. L. De Boer, M. A. Meier, W. Prevoo, J. De Jong, T. M. Bydlon, H. J. C. M. Sterenborg, J. A. Burgers, B. H. W. Hendriks, and T. J. M. Ruers, "Spectral sensing for tissue diagnosis during lung biopsy procedures: the importance of an adequate internal reference and real-time feedback," *Lung Cancer*, vol. 98, pp. 62-68, 2016, [[10.1016/j.lungcan.2016.05.019](#)].
- [45] T. Svensson, S. Andersson-Engels, M. Einarsdóttir, and K. Svanberg, "In vivo optical characterization of human prostate tissue using near-infrared time-resolved spectroscopy," *Journal of Biomedical Optics*, vol. 12, no. 1, pp. 014022, 2007, [[10.1117/1.2435175](#)].
- [46] H. J. Nyst, R. L. P. Van Veen, I. B. Tan, R. Peters, S. Spaniol, D. J. Robinson, F. A. Stewart, P. C. Levendag, and H. J. C. M. Sterenborg, "Performance of a Dedicated Light Delivery and Dosimetry Device for Photodynamic Therapy of Nasopharyngeal Carcinoma: Phantom and Volunteer Experiments," *Lasers in Surgery and Medicine*, vol. 39, pp. 647-653, 2007, [[10.1002/lsm.20536](#)].

- [47] R. Nachabé, B. H. W. Hendriks, R. Schierling, J. Hales, J. M. Racadio, S. Rottenberg, T. J. M. Ruers, D. Babic, and J. M. Racadio, "Real-Time In Vivo Characterization of Primary Liver Tumors With Diffuse Optical Spectroscopy During Percutaneous Needle Interventions - Feasibility Study in Woodchucks," *Investigative Radiology*, vol. 50, pp. 443-448, 2015.
- [48] C. Zhu, G. M. Palmer, T. M. Breslin, J. Harter, and N. Ramanujam, "Diagnosis of Breast Cancer Using Diffuse Reflectance Spectroscopy: Comparison of a Monte Carlo Versus Partial Least Squares Analysis Based Feature Extraction Technique," *Lasers in Surgery and Medicine*, vol. 38, no. 7, pp. 714-24, 2006, [[10.1002/lsm.20356](#)].
- [49] A. M. Laughney, V. Krishnaswamy, P. Beatriz Garcia-Allende, O. M. Conde, W. A. Wells, K. D. Paulsen, and B. W. Pogue, "Automated classification of breast pathology using local measures of broadband reflectance," *Journal of Biomedical Optics*, vol. 15, no. 6, pp. 066019, 2010, [[10.1117/1.3516594](#)].
- [50] T. J. Farrell, B. C. Wilson, and M. S. Patterson, "The use of a neural network to determine tissue optical properties from spatially resolved diffuse reflectance measurements," *Physics in Medicine and Biology*, vol. 37, no. 12, pp. 2281-2286, 1992.
- [51] I. J. Bigio, S. G. Bown, G. Briggs, C. Kelley, S. Lakhani, D. Pickard, P. M. Ripley, I. G. Rose, and C. Saunders, "Diagnosis of breast cancer using elastic-scattering spectroscopy: preliminary clinical results," *J Biomed Opt*, vol. 5, no. 2, pp. 221-228, 2000.
- [52] T. Li, C. Zhang, and M. Ogiwara, "A comparative study of feature selection and multiclass classification methods for tissue classification based on gene expression," *Bioinformatics*, vol. 20, no. 15, pp. 2429-2437, 2004, [[10.1093/bioinformatics/bth267](#)].
- [53] H.-H. Hsu, *Advanced Data Mining Technologies in Bioinformatics*: Idea Group Pub., 2006.
- [54] J. Balog, T. Szaniszló, K.-C. Schaefer, J. Denes, A. Lopata, L. Godorhazy, C. Szalay, D. Balogh, L. Sasi-Szabo, M. Toth, and Z. Takats, "Identification of Biological Tissues by Rapid Evaporative Ionization Mass Spectrometry," *Analytical Chemistry*, vol. 82, no. 17, pp. 7343-7350, 2010, [[10.1021/ac101283x](#)].
- [55] G. M. Palmer, and N. Ramanujam, "Monte Carlo-based inverse model for calculating tissue optical properties. Part I: Theory and validation on synthetic phantoms," *Applied Optics*, vol. 45, no. 5, pp. 1062-1071, 2006, [[10.1364/AO.45.001062](#)].
- [56] G. Zonios, and A. Dimou, "Light scattering spectroscopy of human skin in vivo," *Optics Express*, vol. 17, no. 3, pp. 1256-1267, 2009.
- [57] R. Nachabé, D. J. Evers, B. H. W. Hendriks, G. W. Lucassen, M. Van der Voort, J. Wesseling, and T. J. M. Ruers, "Effect of bile absorption coefficients on the estimation of liver tissue optical properties and related implications in discriminating healthy and tumorous samples," *Biomedical Optics Express*, vol. 2, no. 3, pp. 600-614, 2011.
- [58] L. L. De Boer, B. G. Molenkamp, T. M. Bydlon, B. H. W. Hendriks, J. Wesseling, H. J. C. M. Sterenberg, and T. J. M. Ruers, "Fat/Water ratios measured with diffuse reflectance spectroscopy to detect breast tumor boundaries," *Breast Cancer Research and Treatment*, vol. 152, no. 3, pp. 509-518, 2015, [[10.1007/s10549-015-3487-z](#)].
- [59] R. Etzioni, N. Urban, S. Ramsey, M. McIntosh, S. Schwartz, B. Reid, J. Radich, G. Anderson, and L. Hartwell, "The case for early detection," *Nature Reviews: Cancer*, vol. 3, pp. 1-10, 2003, [[10.1038/nrc1041](#)].

Fat/Water ratios measured with diffuse reflectance spectroscopy to detect breast tumor boundaries

L.L. de Boer | B.G. Molenkamp | T.M. Bydlon | B.H.W. Hendriks | J. Wesseling
H.J.C.M. Sterenborg | T.J.M. Ruers

Purpose: Recognition of the tumor during breast-conserving surgery (BCS) can be very difficult and currently, a robust method of margin assessment for the surgical setting is not available. As a result, tumor-positive margins, which require additional treatment, are not found until histopathologic evaluation.

Methods: With Diffuse Reflectance Spectroscopy (DRS), tissue can be characterized during surgery based on optical parameters that are related to the tissue morphology and composition. Here we investigate which optical parameters are able to detect tumor in an area with a mixture of healthy and tumor tissue and hence which parameters are most suitable for intra-operative margin assessment. DRS spectra (400-1600 nm) were obtained from *ex vivo* lumpectomy specimens from healthy, tumor border, and tumor tissue.

Results: The optical parameter related to the absorption of fat and water in the Near-Infra Red (NIR) wavelength region provided the best discrimination between healthy and tumor sites resulting in a sensitivity and specificity of 100%. Per patient, the F/W-ratio clearly decreased when the measurement positions moved from grossly healthy tissue towards the tumor. A benefit of using the NIR wavelength region is the absence of light absorption by blood, making it more suitable for an intra-operative surgical application.

Conclusions: Based on the F/W-ratio, DRS produced a surgical resection plane that nearly overlapped with a 2 mm rim of healthy tissue, 2 mm being the most widely accepted definition of a negative margin. This work shows the potential for DRS to guide the surgeon during breast-conserving surgery.

Keywords: Diffuse Reflectance Spectroscopy (DRS), resection margins, breast cancer, Near-Infrared (NIR)

INTRODUCTION

Patients with early-stage breast cancer can be treated with Breast-Conserving Surgery (BCS) in combination with adjuvant radiation therapy. This results in long term survival comparable to mastectomies, with the advantage of a better cosmetic outcome and a better quality of life.[1, 2] A strict condition for successful BCS is the removal of the entire tumor with a rim of healthy tissue, hence, the absence of positive resection margins.[3-10] Unfortunately, BCS is often inadequate; 10-60% of patients treated with BCS require additional surgery to obtain negative margins.[11-17] Remarkably, the definition of a negative margin differs between institutions, ranging from 1 cm of healthy tissue around the tumor to no pathology ink touching the tumor.[18-21] Although under debate, the most widely accepted safe standard, currently used is a distance of 2 mm between the border of the tumor and the resection margin.[7, 21]

To date, there are no robust methods to help the surgeon identify tumor tissue at the resection margin during surgery. Intra-operative margin assessment by imprint cytology and frozen section analysis can reduce the number of re-excisions[22-24], but both techniques experience major disadvantages. They require additional operation time and a skilled on-site pathologist to correctly interpret the results.[25, 26] In addition, the reliability of imprint cytology is hampered by drying, cautery, and irregularity of the tissue surface[9, 27]; while frozen section analysis damages the resection specimen and is prone to artifacts and sampling errors[15, 28]. Besides these techniques, measurements of electrical impedance can also be used to detect tumor tissue at the resection margin, as is done with the MarginProbe® (Dune Medical Devices). Assessing resection specimens with this technique and, in case of a positive reading, performing additional cavity shaving, 6% of positive resection margins could be avoided.[29] However, this FDA-approved device has not (yet) been researched for *in vivo* use.

An alternative for *in vivo* margin assessment may be the use of optically-based techniques.[30-40]. One of these techniques is DRS, which measures the amount of diffusely reflected light after it has undergone multiple scattering and absorption events within the tissue. Each DRS spectrum, therefore, has a specific 'optical fingerprint' reflecting the composition and morphology of the tissue within the probing volume. DRS has been widely investigated for the purpose of margin assessment. Within the visible wavelength region oxygen saturation[30, 36, 38], β -carotene[30, 36], and light scattering[30, 36, 38, 41] have proved to be able to discriminate tumor tissue from healthy tissue. However, in this wavelength range blood is one of the predominant absorbers. The blood that is present at the tissue surface during surgery can hamper the reliability of intra-operative *in vivo* DRS measurements based on the visible wavelength region. In the NIR wavelengths, absorption of light by the blood is negligible, making this region more robust for an *in vivo* surgical application.

The first aim of this study was to determine the best optical parameters within the NIR region that would allow intra-operative discrimination of healthy and malignant breast tissue. Furthermore, previous research mainly focused on comparing measurements of 100% healthy locations with 100% tumor locations. However, at the resection border, the probed tissue volume will likely contain a mixture of these tissue types. Therefore, the second goal of this study was to investigate the ability of these optical parameters to detect tumor at a margin where a mixture of healthy and tumor tissue is present. Such an ability would enable finding an optimal surgical plane during breast-conserving surgery.

MATERIALS & METHODS

In the first part of the study, we investigated the performance of various optical parameters in recognizing tumor tissue by obtaining DRS spectra from *ex vivo* lumpectomy specimens. In the second part of the study, a custom-made grid was used to acquire DRS measurements in a structured manner on a breast tissue sample of a mastectomy specimen, thereby mimicking the use of DRS for potentially delineating a tumor. All measurements were performed at the pathology department (Dutch Cancer Institute - Antoni van Leeuwenhoek) with minimal interference in the standard pathology protocol. Lumpectomy specimens with a tumor size > 1.0 cm were included. Specimens from patients with a radiological complete or favorable response after neoadjuvant chemotherapy were excluded. The margins were inked, the specimen was cooled and hardened in the freezer, and then sliced. One of the slices (~5 mm thickness) containing both healthy tissue and tumor tissue was used for the DRS measurements. Three classes of measurement sites were defined: healthy locations (> 1 cm from the tumor), border locations (at the edge of the tumor) and tumor locations (in the middle of the tumor). These locations were chosen in consultation with a pathologist, based on the macroscopic appearance of the tissue. At each measurement location, three DRS spectra were obtained which were averaged for further analysis. To make a comparison between patients, the locations with the same classification of one patient were averaged, resulting in an average value for the healthy, border and tumor tissue for each patient.

Correlation with histopathology

The histology of the slices of the lumpectomy specimens from the first part of the study could only be obtained by cutting the lumpectomy slices into separate sections for histological staining because they were too large to fit in one cassette. Thus, multiple sections formed the one measured slice of the lumpectomy specimen. To create one microscopic image of the measured lumpectomy slice, the H&E sections of the sections were digitalized (Aperio Scanscope, Leica Biosystems, Wetzlar, Germany) and merged (Adobe Photoshop CS5) (Figure 1a). To correlate the measurement locations with histopathology, photographs were

Figure 1. Procedure for matching the overview photograph (macroscopic image) with histopathology (microscopic image). (a) Microscopic images of the sections. (b) Measurement locations located on the overview photograph. (c) Matching the microscopic image with the overview photograph which provides the histology of the measurement locations

taken of the probe at each measurement location, as well as one overview photograph of the entire specimen. With these photographs, the measurement locations were marked in the overview photograph (Figure 1b). The macroscopic overview photograph and the microscopic image were matched enabling correlation with histopathology at the measurement locations (Figure 1c).

Grid measurements

In the second part of the study, we wanted to acquire DRS spectra over a larger surface to determine a potential resection line around the tumor tissue. To this end, the tissue sample was placed in a cassette with the custom-made grid on top of it. The small holes in the grid were the exact shape of the tip of the probe strictly defining probe positions and limiting movement during the measurements. Once the measurements were done, the tissue remained in the cassette and was fixated in the same position. This ensured an accurate correlation of the histopathology with the measured spectra.

Instrumentation

The light source in the measurement set-up was a broadband Tungsten halogen light. Two spectrometers were used, one with a silicon detector resolving the visual light between 400 and 1100 nm (Andor Technology, DU420A-BRDD) and one with an InGaAs detector resolving light in the NIR region from 800 to 1700 nm (Andor Technology, DU492A-1.7).[42] The spectrometers were in contact with the tissue via an optical probe that was shaped like a pencil with a blunt end to avoid damaging the tissue. The probe tip incorporating the optical fibers was 1.75 mm in diameter with a distance between the emitting fiber and the collecting fiber of 1.5 mm. The probing volume of this optical probe is approximately 1-3 mm³.

Fit model

The measured spectra were fit with an analytical model based on diffusion theory to calculate the optical characteristics of the probed volume.[42-44] The analytical model uses known absorption and scattering characteristics of substances present in the probed volume to translate the measured spectrum into estimations of these substances. To this end the absorption coefficients of oxyhemoglobin (HbO₂), deoxyhemoglobin (Hb), fat, water, β -carotene and collagen were provided as a prior knowledge to the model, resulting in estimations of volumes (i.e. blood fraction, fat fraction, total volume of fat and water, collagen fraction), oxygen saturation (StO₂), β -carotene concentration, and the scattering parameters (i.e. the reduced scattering coefficient at 800 nm (s_{800}), Mie scattering (f_{mie})). The absorption coefficients of fat and water were not included separately in the formula for total absorption to avoid covariance since both substances absorb light over the same wavelength range. Instead, the two were combined in the total volume of fat and water (Fat+Water) and the fat fraction in that volume (Fat/(Fat+Water)) resulting in a more stable fit.[42] From these optical parameters the fat percentage and water percentage was derived and by dividing the fat percentage by the water percentage the ratio between the two was calculated (F/W-ratio).

To account for light absorption by the Patent blue dye that was injected peri-tumorally in some patients, the absorption coefficient of Patent blue was also included in the fit. For the same reason, the absorption coefficients of red, yellow, and blue pathology ink (Electron Microscopy Sciences Tissue Marking Dyes, Hatfield, Pennsylvania, USA.) were included in the fitting process, although areas with ink were avoided some small spots may have still been present at the measurement locations.

Statistics

Tumor tissue and healthy breast tissue were compared for each optical parameter using a generalized estimating equation (GEE) method which accounts for the correlation between multiple measurements obtained from one specimen.[45] To represent the within-patient

dependencies a correlation matrix was used which assumed all pairwise correlations to be equal. It was furthermore assumed that the parameters were normally distributed. The analyses were performed in Matlab software (MATLAB R2014a) using the GEEQBOX toolbox and p-values smaller than 0.01 were considered significant.

RESULTS

The first goal of this study was to determine which optical parameters provided the best discrimination between healthy tissue and tumor tissue with a special emphasis on the optical parameters in the NIR region. A total of 16 specimens were analyzed which resulted in 169 different locations used for analysis; 42 locations were from healthy tissue, 25 locations from tumor tissue, and 102 locations from the border of the tumor.

First, a collective analysis was performed, comparing all the healthy locations with tumor locations. In Figure 2 the total absorption of light in healthy and tumor tissue over both the visible and NIR wavelengths is displayed. This figure also depicts the wavelength regions where the absorption by blood (HbO_2 & Hb), fat and water take place. The lower wavelengths are predominantly influenced by absorption through blood, while the higher wavelengths are predominantly influenced by absorption through fat and water. For intra-operative use of DRS, contamination of blood at the resection margin will impair the reliability of the visual wavelengths making the NIR region more suitable for this application. In addition, the differences between healthy and malignant tissue are clearly more profound in the higher wavelengths which further emphasizes the preference for parameters such as water and fat in the NIR wavelength region.

The best discrimination between normal and tumor tissue was obtained for the optical parameters, F/W-ratio ($p < 1.0 \cdot 10^{-7}$); Mie scattering ($p = 1.3 \cdot 10^{-7}$) and StO_2 ($p = 0.004$). (Figure 3) The p-values of the other parameters were higher and therefore not significant; β -carotene ($p = 0.23$), blood ($p = 0.13$), scattering at 800 nm ($p = 0.12$), and collagen ($p = 0.02$). The best performing optical parameter was F/W-ratio, proving that fat and water content differ the most between healthy and malignant tissue.

The F/W-ratios of all healthy and tumor locations are depicted in Figure 4. All the F/W-ratios of the healthy locations were higher than 1, whereas all the F/W-ratios of the tumor locations were lower than 1. Thus, by calculating the ratio of fat divided by water, tumor tissue can be separated from normal tissue with a sensitivity and specificity of 100% if F/W-ratio < 1 is used as a threshold.

Subsequently, in a per-patient analysis, the measurements obtained at the border of the tumor were included. Figure 5a displays the average F/W-ratio of each tissue class for each specimen. The F/W-ratio of healthy tissue varied considerably between specimens. However, more importantly in nearly every specimen the F/W-ratio clearly decreased when moving

Figure 2. The green and red lines represent the total absorbance of light in healthy and tumor tissue with their 95 % confidence intervals. The data are normalized at 800 nm. The insets show the known absorption coefficients that are included in the fit model. The left grey area (within the visible wavelength region) indicates the part of the spectrum influenced by the presence of blood substances, β -carotene, and collagen, the right grey area (within the NR wavelength region) displays the wavelengths influenced by the presence of fat, water, and also collagen.

from grossly healthy tissue, across the tumor boundary, to the tumor. In only two specimens (12.5%) the F/W-ratio of the border locations was slightly higher than in the healthy locations. To account for inter-patient variability the F/W-ratio was scaled by setting the F/W-ratio in healthy tissue to 1. (Figure 5b) The F/W-ratio of the border locations of the two specimens that did not decrease have a scaled F/W-ratio higher than 1. All other scaled F/W-ratios of the border sites are below a value of 0.58. For further analysis, this value was accepted as the threshold for distinguishing where healthy tissue starts to mix with tumor tissue.

In the second part of the study, we investigated if it was possible to construct a safe surgical plane, based on the DRS measurements. To this end, 55 locations were measured

Figure 3. Boxplots with healthy and tumor locations for the significant optical parameters: Fat/Water ratio (a), Mie scattering (b), and oxygen saturation (c).

Figure 4. Bar graph displaying F/W-ratio of all healthy and tumor locations. The blue line represents a F/W-ratio of 1. The F/W-ratio of all normal locations is higher than 1, whereas the F/W-ratio of all tumor locations is lower than 1 (sensitivity and specificity of 100 %).

on a tissue sample using the custom-made grid. In Figure 6 the histology image is overlaid with dots depicting the scaled F/W-ratio of all measurement locations. Due to an artifact in the H&E section, histology of one location is not available (grey dot). The dots with a scaled F/W-ratio > 0.58 are colored green, locations with a F/W-ratio < 0.58 are colored in shades ranging from green to orange to red depending on the F/W-ratio. The red dots correlate to the area where tumor tissue is present in the microscopic image; the dots that are colored orange-yellow are located at the border of the tumor. At the orange-yellow locations, according to the microscopic image, a mixture of healthy tissue and tumor tissue is present. The difference in scaled F/W-ratio between absolutely healthy breast tissue and absolutely tumor tissue was considerable as should be noted from the bar graphs in Figure 6. Out of 54 measurement locations with available histology, the F/W-ratio of four dots did

Figure 5. The average ratios of each specimen ($n = 16$) plotted per tissue type (a). In the right graph (b), the average ratios are scaled the F/W-ratio in healthy tissue to 1. The blue line represents a threshold of 0.58.

Figure 6. Overview of grid measurements. The colors of the dots in the H&E section (a) correspond to the measurement ratio of fat/water. The 2-mm zone around the tumor is depicted with the black line. The dotted line displays where the surgical plane would have been if DRS would have been as a guide. Two dots marked with the asterisk have yellow/orange color but are located in an area suspected for tumor. The two orange dots that are situated in healthy tissue are marked with a hashtag. Histopathology is not available for one location (grey dot) due to the artifact in the H&E section. The profile of the measurement within one line, marked with a number, is displayed on the right (b). In the right bottom (c), the cassette with the grid is depicted.

not completely match with the underlying tissue (7.4%). There are two orange dots located in an area suspected for tumor (marked with an asterisk) and two orange dots seem to be situated in healthy tissue (marked with a hashtag). By connecting the dots with a scaled F/W-ratio ≥ 0.58 a surgical plane was defined that should be safe according to the DRS measurements (black dotted line). The black solid line depicted in the image represents the most widely used definition of a negative margin (2 mm). In theory, avoiding this area can prevent a positive resection margin. The potential surgical plane indicated by DRS is in accordance with the commonly used definition of a safe surgical plane, or slightly exceeds this 2 mm zone as shown in Figure 6.

DISCUSSION

Adequate *in vivo* intraoperative margin assessment during BCS is still not widely available. This makes it difficult for surgeons to judge the most optimal surgical plane. Both, the need for re-excision in the case of a positive resection margin and resecting more tissue than is strictly necessary, have a negative result on the cosmetic outcome. Several research groups have shown the feasibility of DRS in discriminating healthy and tumor tissue, thus, making it a good candidate for margin assessment. Here we bring the technology one step closer to an intra-operative clinical application by investigating the best performing optical parameters, with special emphasis on the NIR region, and their behavior around the boundary of the tumor. Previous research by our group reported that based on their absorption peaks in the higher NIR wavelength region, fat and water can be quantified by measuring a wavelength range up to 1600 nm.[42, 44] Since healthy breast tissue is fattier compared to tumor tissue and tumor tissue contains more water probably due to biological processes (increased metabolic activity, vascularization, necrosis or inflammation) DRS was expected to discriminate healthy breast tissue from tumor tissue. Indeed, the F/W-ratio discriminated healthy tissue from tumor tissue with a sensitivity and specificity of 100%, making it the best performing parameter. Previous papers, which were limited to the visible wavelength region, reported other optical parameters (scattering, β -carotene, and hemoglobin) to provide the best discrimination.[30, 36, 38]. In the present study differences in absolute absorbance between healthy and tumor tissue were seen both in the wavelength region between 400-650 nm, related to light absorption through hemoglobin and β -carotene, as well as in the wavelength range between 1000-1600 nm, related to the presence of fat and water. The absorbance difference in the NIR wavelength region was larger, which explains why the F/W-ratio was the best performing parameter. Our results, based on the absorption of fat and water in the NIR wavelength region, are in line with publications by other groups like Laughney et al. in which spatial frequency domain imaging with NIR wavelengths was performed[41], and Taroni et al. in which DRS spectra were measured up to 1100 nm[46].

An advantage of using the NIR wavelength region over the visible wavelength region is the absence of light absorption through blood. Especially in an intraoperative setting, visual wavelengths can be impaired by the contamination of blood on the resection surface making the F/W-ratio more suitable for a surgical application.

The average healthy F/W-ratios of the lumpectomy specimens varied. This is related to the heterogenic nature of healthy breast tissue.[36] However, in all measurements going from healthy to tumor tissue, a decreasing trend in the F/W-ratios was clearly present in all specimens. The F/W-ratio measured in tumor tissue was for each specimen less than the F/W-ratio measured in the healthy locations. With respect to the border locations, only two cases had a higher F/W-ratio compared to the F/W-ratio of the healthy locations. A more detailed look at the histopathology of these two specimens revealed that multiple border locations contained > 90% fat. The registration of the overview photograph and the histopathology image may have been not absolutely perfect in these cases, and consequently, the volumes measured at these locations contained only healthy tissue thus explaining the high F/W-ratio.

Since there is considerable variance in the F/W-ratio of healthy tissue between patients, setting a threshold for the F/W-ratio which is applicable in all patients may not be possible. Therefore, these ratios should be scaled using the F/W-ratio of healthy tissue as a reference. In practice, for an intra-operative setting, this will not be an obstacle since the surgeon is able to measure some healthy locations, determining a patient-specific reference, before assessing the margin.

In the second part of the study, measurements were obtained with a grid in a very controlled setting, providing the highest possible accuracy for matching with the histopathology. In Figure 6 a comparison was made between a 2 mm surgical demarcation zone, regarded as safe according to current surgical standards, and the plane as indicated by the DRS measurements of the scaled F/W-ratio. The line indicated by the DRS measurements nearly overlaps with the 2 mm zone around the tumor and more importantly does not touch the tumor anywhere. This suggests the ability of DRS to indicate a safe surgical margin based on the fat and water content of the tissue.

Contrary to margin assessment techniques currently used in the operating room, DRS cannot detect individual tumor cells because the changes in optical spectra caused by a few random tumor cells in the probed volume are too subtle. However, the need for detecting individual cells in order to improve patient care is debatable. First, because adjuvant therapies which are part of the current standard of care will tackle small deposits of tumor cells; and second, aiming for resection of random tumor cells will lead to sacrificing additional healthy tissue, thus deteriorating cosmetic outcome. Current guidelines prescribe that re-excision is indicated for positive resection margins, defined as tumor touching ink over an area of 2 mm or more. Furthermore, a 2 mm rim of healthy tissue around the tumor is recommended to obtain a negative surgical margin. Here we prove that DRS can discriminate tumor tissue

from normal tissue thereby detecting resection margins that are considered positive. Moreover, DRS can recognize the tumor border helping to guide the surgeon in placing the resection at a safe distance from the tumor to obtain negative resection margins without resecting more healthy tissue than needed.

In this study, we explored the NIR wavelength region for tissue discrimination in breast cancer. It was observed that absorption related to fat and water content in the NIR region provided the best discrimination between tumor tissue and healthy breast tissue. Using the NIR wavelengths reduces the influence of light absorption by blood making the technique more suitable for *in vivo* use during surgical procedures. We demonstrated that the scaled F/W-ratio can detect relative changes that are related to the presence of tumor near the potential resection border of the tumor. In this way, DRS shows the potential to guide surgeons during breast-conserving surgery. Further studies should concentrate on the robustness of the technique in general practice.

ACKNOWLEDGMENTS

We would like to thank the NKI pathology department for their help in the collection of the data as well as the Philips Research project members for their contribution. We also thank Jarich Spliethoff and Jasper Nijkamp for their valuable feedback. This work is supported by Philips Research, Eindhoven, The Netherlands.

REFERENCES

- [1] D. Curran, J. P. van Dongen, N. K. Aaronson, G. Kiebert, I. S. Fentiman, F. Mignolet, and H. Bartelink, "Quality of Live of Early-stage Breast Cancer Patients Treated with Radical Mastectomy or Breast-conserving Procedures: Results of EORTC Trial 10801," *European Journal of Cancer*, vol. 34, no. 3, pp. 307-314, 1998.
- [2] S. K. Al-Ghazal, L. Fallowfield, and R. W. Blamey, "Comparison of psychological aspects and patient satisfaction following breast conserving surgery, simple mastectomy and breast reconstruction," *European Journal of Cancer*, vol. 36, pp. 1938-1943, 2000.
- [3] C. C. Park, M. Mitsumori, A. Nixon, A. Recht, J. Connolly, R. Gelman, B. Silver, S. Hetelekidis, A. Abner, J. R. Harris, and S. J. Schnitt, "Outcome at 8 Years After Breast-Conserving Surgery and Radiation Therapy for Invasive Breast Cancer: Influence of Margin Status and Systemic Therapy on Local Recurrence," *Journal of Clinical Oncology*, vol. 18, no. 8, pp. 1668-1675, 2000.
- [4] B. Fisher, S. Anderson, J. Bryant, R. G. Margolese, M. Deutsch, E. R. Fisher, J.-H. Jeong, and N. Wolmark, "Twenty-Year Follow-up of a Randomized Trial Comparing Total Mastectomy, Lumpectomy, and Lumpectomy Plus Irradiation for the Treatment of Invasive Breast Cancer," *The New England Journal of Medicine*, vol. 347, no. 16, pp. 1233-1241, 2002.
- [5] J. A. Van Dongen, A. C. Voogd, I. S. Fentiman, C. Legrand, R. J. Sylvester, D. Tong, E. Van der Schueren, P. A. Helle, K. Van Zijl, and H. Bartelink, "Long-Term Results of a Randomized Trial Comparing Breast-Conserving Therapy With Mastectomy: European Organization for Research and Treatment of Cancer 10801 Trial," *Journal of the National Cancer Institute*, vol. 92, no. 14, pp. 1143-1150, 2000.
- [6] M. M. Poggi, D. N. Danforth, L. C. Sciuto, S. L. Smith, S. M. Steinberg, D. J. Liewehr, C. Menard, M. E. Lippman, A. S. Lichter, and R. M. Altomus, "Eighteen-Year Results in the Treatment of Early Breast Carcinoma with Mastectomy versus Breast Conservation Therapy," *Cancer*, vol. 98, pp. 697-702, 2003, [[10.1002/cncr.11580](https://doi.org/10.1002/cncr.11580)].
- [7] H. Houssami, P. Macaskill, L. Marinovich, and M. Morrow, "The Association of Surgical Margins and Local Recurrence in Women with Early-Stage Invasive Breast Cancer Treated with Breast-Conserving Therapy: A Meta-Analysis," *Ann Surg Oncol*, vol. 21, pp. 717-730, 2014, [[10.1245/s10434-014-3480-5](https://doi.org/10.1245/s10434-014-3480-5)].
- [8] U. Veronesi, N. Cascinelli, L. Mariani, M. Greco, R. Saccozzi, A. Luini, M. Aguilar, and E. Marubini, "Twenty-year Follow-up of a Randomized Study Comparing Breast-Conserving Surgery with Radical Mastectomy for Early Breast Cancer," *The New England Journal of Medicine*, vol. 347, no. 16, pp. 1227-1232, 2002.
- [9] S. E. Singletary, "Surgical margins in patients with early-stage breast cancer treated with breast conservation therapy," *American Journal of Surgery*, vol. 184, pp. 383-393, 2002.
- [10] M. Morrow, E. A. Storm, L. W. Bassett, D. D. Dershaw, B. Fowble, A. Guiliano, J. R. Harris, F. O'Malley, S. J. Schnitt, S. E. Singletary, and D. P. Winchester, "Standard for Breast Conservation Therapy in the Management of Invasive Breast Carcinoma," *CA: A Cancer Journal for Clinicians*, vol. 52, no. 5, pp. 277-300, 2002.
- [11] L. E. McCahill, R. M. Single, E. J. Aiello Bowles, H. S. Feigelson, T. A. James, T. Barney, J. M. Engel, and A. A. Onitilo, "Variability in Reexcision Following Breast Conservation Surgery," *Journal of the American Medical Association*, vol. 307, no. 5, pp. 467-475, 2012.
- [12] J. F. Waljee, E. S. Hu, L. A. Newman, and A. K. Alderman, "Predictors of re-excision among women undergoing breast-conserving surgery for cancer," *Ann Surg Oncol*, vol. 15, no. 5, pp. 1297-303, 2008.
- [13] R. Jeevan, D. A. Cromwell, M. Trivella, G. Lawrence, O. Kearins, J. Pereira, C. Sheppard, C. M. Caddy, and J. H. P. van der Meulen, "Reoperation rates after breast conserving surgery for breast cancer among women in England: retrospective study of hospital episode statistics," *British Medical Journal*, vol. 345, pp. e4505, 2012, [[10.1136/bmj.e4505](https://doi.org/10.1136/bmj.e4505)].

- [14] P. J. Lovrics, S. D. Cornacchi, F. Farrokhyar, A. Garnett, V. Chen, S. Franic, and M. Simunovic, "Technical factors, surgeon case volume and positive margin rates after breast conservation surgery for early-stage breast cancer," *Canadian Journal of Surgery*, vol. 53, no. 5, pp. 305-312, 2010.
- [15] V. S. Klimberg, S. Harms, and S. Korourian, "Assessing margin status," *Surgical Oncology*, vol. 8, pp. 77-84, 1999.
- [16] M. Morrow, "Margins in breast-conserving therapy have we lost sight of the big picture," *Expert Review of Anticancer Therapy*, vol. 8, no. 8, pp. 1193-1196, 2008.
- [17] N. Cabioglu, K. K. Hunt, A. A. Sahin, H. M. Kuerer, G. V. Babiera, S. E. Singletary, G. J. Whitman, M. I. Ross, F. C. Ames, B. W. Feig, T. A. Buchholz, and F. Meric-Bernstam, "Role for Intraoperative Margin Assessment in Patients Undergoing Breast-Conserving Surgery," *Annals of Surgical Oncology*, vol. 14, no. 4, pp. 1458-71, 2007, [[10.1245/s10434-006-9236-0](https://doi.org/10.1245/s10434-006-9236-0)].
- [18] M. Azu, P. Abrahamse, S. J. Katz, R. Jagsi, and M. Morrow, "What is an Adequate Margin for Breast-Conserving Surgery? Surgeon Attitudes and Correlates," *Annals of Surgical Oncology*, vol. 17, pp. 558-563, 2010, [[10.1245/s10434-009-0765-1](https://doi.org/10.1245/s10434-009-0765-1)].
- [19] S. L. Blair, K. Thompson, J. Rococco, V. Malcarne, P. D. Beitsch, and D. W. Ollila, "Attaining Negative Margins in Breast-Conservation Operations: Is There a Consensus among Breast Surgeons?," *Journal of the American College of Surgeons*, vol. 209, no. 5, pp. 608-613, 2009, [[10.1016/j.jamcollsurg.2009.07.026](https://doi.org/10.1016/j.jamcollsurg.2009.07.026)].
- [20] N. Houssami, P. Macaskill, M. L. Marinovich, J. M. Dixon, L. Irwig, M. E. Brennan, and L. J. Solin, "Meta-analysis of the impact of surgical margins on local recurrence in women with early-stage invasive breast cancer treated with breast-conserving therapy," *European Journal of Cancer*, vol. 46, pp. 3219-3232, 2010, [[10.1016/j.ejca.2010.07.043](https://doi.org/10.1016/j.ejca.2010.07.043)].
- [21] M. S. Moran, S. J. Schnitt, A. E. Giuliano, J. R. Harris, S. A. Khan, J. Horton, S. Klimberg, M. Chavez-MacGregor, G. Freedman, N. Houssami, P. L. Johnson, and M. Morrow, "Society of Surgical Oncology - American Society for Radiation Oncology Consensus Guideline on Margins for Breast-Conserving Surgery With Whole-Breast Irradiation in Stages I and II Invasive Breast Cancer," *International Journal of Radiation Oncology Biology Physics*, vol. 88, no. 3, pp. 553-564, 2014.
- [22] K. Esbona, Z. Li, and L. G. Wilke, "Intraoperative Imprint Cytology and Frozen Section Pathology for Margin Assessment in Breast Conservation Surgery: A Systematic Review," *Annals of Surgical Oncology*, vol. 19, no. 10, pp. 3236-3245, 2012, [[10.1245/s10434-012-2492-2](https://doi.org/10.1245/s10434-012-2492-2)].
- [23] T. P. Olson, J. Harter, A. Muñoz, D. M. Mahvi, and T. M. Breslin, "Frozen Section Analysis for Intraoperative Margin Assessment During Breast-Conserving Surgery Results in Low Rates of Re-excision and Local Recurrence," *Annals of Surgical Oncology*, vol. 14, no. 10, pp. 2953-2960, 2007.
- [24] C. Dener, A. Inan, M. Sen, and S. Demirci, "Intraoperative frozen section for margin assessment in breast conserving surgery," *Scandinavian Journal of Surgery*, vol. 98, pp. 34-40, 2009.
- [25] S. A. Khan, and F. Eladoumikhachi, "Optimal Surgical Treatment of Breast Cancer: Implications for Local Control and Survival," *Journal of Surgical Oncology*, vol. 101, no. 8, pp. 677-686, 2010, [[10.1002/jso.21502](https://doi.org/10.1002/jso.21502)].
- [26] M. S. Sabel, "Surgical Considerations in Early-Stage Breast Cancer: Lessons Learned and Future Directions," *Seminars in Radiation Oncology*, vol. 21, pp. 10-19, 2011, [[10.1016/j.semradonc.2010.08.002](https://doi.org/10.1016/j.semradonc.2010.08.002)].
- [27] M. Bakhshandeh, O. Tutuncuoglu, G. Fischer, and S. Masood, "Use of Imprint Cytology for Assessment of Surgical Margins in Lumpectomy Specimens of Breast Cancer Patients," *Diagnostic Cytopathology*, vol. 35, no. 10, pp. 656-659, 2007.
- [28] K. N. Holder, and I.-T. Yeh, "Intraoperative Evaluation of Margin Status," *Pathology Case Reviews*, vol. 15, no. 5, pp. 148-155, 2010, [[10.1097/PCR.0b013e3181f66188](https://doi.org/10.1097/PCR.0b013e3181f66188)].

- [29] F. Schnabel, S. K. Boolbol, M. Gittleman, T. Karni, L. Tafra, S. Feldman, A. Police, N. B. Friedman, S. Karlan, D. Holmes, S. C. Willey, M. Carmon, K. Fernandez, S. Akbari, J. Harness, L. Guerra, T. Frazier, K. Lane, R. M. Simmons, A. Estabrook, and T. Allweis, "A Randomized Prospective Study of Lumpectomy Margin Assessment with Use of MarginProbe in Patients with Nonpalpable Breast Malignancies," *Annals of Surgical Oncology*, vol. 21, pp. 1589-1595, 2014, [DOI 10.1245/s10434-014-3602-0].
- [30] C. Zhu, G. M. Palmer, T. M. Breslin, J. Harter, and N. Ramanujam, "Diagnosis of Breast Cancer Using Diffuse Reflectance Spectroscopy: Comparison of a Monte Carlo Versus Partial Least Squares Analysis Based Feature Extraction Technique," *Lasers in Surgery and Medicine*, vol. 38, no. 7, pp. 714-24, 2006, [10.1002/lsm.20356].
- [31] C. Zhu, G. M. Palmer, T. M. Breslin, J. Harter, and N. Ramanujam, "Diagnosis of breast cancer using fluorescence and diffuse reflectance spectroscopy: a Monte-Carlo-model-based approach," *Journal of Biomedical Optics*, vol. 13, no. 3, pp. 034015, 2008, [10.1117/1.2931078].
- [32] J. Q. Brown, T. M. Bydlon, L. M. Richards, B. Yu, S. A. Kennedy, J. Geradts, L. G. Wilke, M. Junker, J. Gallagher, W. Barry, and N. Ramanujam, "Optical assessment of tumor resection margins in the breast," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 16, no. 3, pp. 530-544, 2010.
- [33] A. P. Wax, J. Q. Brown, T. M. Bydlon, S. A. Kennedy, J. Geradts, L. G. Wilke, W. Barry, L. M. Richards, M. K. Junker, J. Gallagher, N. Ramanujam, and V. Backman, "Assessment of breast tumor margins via quantitative diffuse reflectance imaging." pp. 757300-757300-7, [10.1117/12.841949].
- [34] N. Ramanujam, J. Q. Brown, T. M. Bydlon, S. A. Kennedy, L. M. Richards, M. K. Junker, J. Gallagher, W. T. Barry, L. G. Wilke, and J. Geradts, "Quantitative spectral reflectance imaging device for intraoperative breast tumor margin assessment " *Engineering in Medicine and Biology Society*, pp. 6554-6556, 2009, [10.1109/IEMBS.2009.5334501].
- [35] L. G. Wilke, J. Q. Brown, T. M. Bydlon, S. A. Kennedy, L. M. Richards, M. K. Junker, J. Gallagher, W. T. Barry, J. Geradts, and N. Ramanujam, "Rapid Non-invasive Optical Imaging of Tissue Composition in Breast Tumor Margins," *American Journal of Surgery*, vol. 198, no. 4, pp. 566-574, 2009, [10.1016/j.amjsurg.2009.06.018.].
- [36] S. Kennedy, J. Geradts, T. Bydlon, J. Q. Brown, J. Gallagher, M. Junker, W. Barry, N. Ramanujam, and L. Wilke, "Optical breast cancer margin assessment: an observational study of the effects of tissue heterogeneity on optical contrast," *Breast Cancer Research*, vol. 12, no. 6, pp. R91, 2010, [10.1186/bcr2770].
- [37] D. J. Evers, R. Nachabe, M. J. Vranken Peeters, J. A. van der Hage, H. S. Oldenburg, E. J. Rutgers, G. W. Lucassen, B. H. Hendriks, J. Wesseling, and T. J. Ruers, "Diffuse reflectance spectroscopy: towards clinical application in breast cancer," *Breast Cancer Res Treat*, vol. 137, no. 1, pp. 155-65, 2013, [10.1007/s10549-012-2350-8].
- [38] Z. Volynskaya, A. S. Haka, K. L. Bechtel, M. Fitzmaurice, R. Shenk, N. Wang, J. Nazemi, R. R. Dasari, and M. S. Feld, "Diagnosing breast cancer using diffuse reflectance spectroscopy and intrinsic fluorescence spectroscopy," *Journal of Biomedical Optics*, vol. 13, no. 2, pp. 024012, 2008, [10.1117/1.2909672].
- [39] A. S. Haka, Z. Volynskaya, J. A. Gardecki, J. Nazemi, J. Lyons, D. Hicks, M. Fitzmaurice, R. R. Dasari, J. P. Crowe, and S. M. Feldman, "In vivo Margin Assessment during Partial Mastectomy Breast Surgery Using Raman Spectroscopy," *Cancer Research*, vol. 66, no. 6, pp. 3317-3322, 2006, [10.1158/0008-5472.CAN-05-2815].
- [40] M. D. Keller, E. Vargis, N. de Matos Granja, R. H. Wilson, M.-A. Mycek, M. C. Kelley, and A. Mahadevan-Jansen, "Development of a spatially offset Raman spectroscopy probe for breast tumor surgical margin evaluation," *J Biomed Opt*, vol. 16, no. 7, pp. 077006, 2011, [10.1117/1.3600708].
- [41] A. M. Laughney, V. Krishnaswamy, E. J. Rizzo, M. C. Schwab, R. J. Barth, D. J. Cuccia, B. J. Tromberg, K. D. Paulsen, B. W. Pogue, and W. A. Wells, "Spectral discrimination of breast

- pathologies in situ using spatial frequency domain imaging," *Breast Cancer Research*, vol. 15, 2013.
- [42] R. Nachabé, B. H. W. Hendriks, M. Van der Voort, A. E. Desjardins, and H. J. C. M. Sterenberg, "Estimation of biological chromophores using diffuse optical spectroscopy: benefit of extending the UV-VIS wavelength range to include 1000 to 1600 nm," *Optics Express*, vol. 18, no. 24, pp. 1432-1442, 2010.
 - [43] T. J. Farrell, M. S. Patterson, and B. A. Wilson, "A diffusion theory model of spatially resolved, steady-state diffuse reflectance for the noninvasive determination of tissue optical properties *in vivo*," *Journal of Medical Physics*, vol. 19, no. 4, pp. 879-888, 1992.
 - [44] R. Nachabé, B. H. W. Hendriks, A. E. Desjardins, M. van der Voort, M. B. van der Mark, and H. J. C. M. Sterenberg, "Estimation of lipid and water concentrations in scattering media with diffuse optical spectroscopy from 900 to 1600 nm," *J Biomed Opt*, vol. 15, no. 3, pp. 037015, 2010, [[10.1117/1.3454392](https://doi.org/10.1117/1.3454392)].
 - [45] J. Shults, and S. J. Ratcliffe, "Analysis of multi-level correlated data in the framework of generalized estimating equations via xtmultcorr procedures in Stata and qls functions in Matlab," *Stat. Interface*, vol. 2, pp. 187-196, 2009.
 - [46] P. Taroni, A. Bassi, D. Comelli, A. Farina, R. Cubeddu, and A. Pifferi, "Diffuse optical spectroscopy of breast tissue extended to 1100 nm," *J Biomed Opt*, vol. 14, no. 5, pp. 054030, 2009, [[10.1117/1.3251051](https://doi.org/10.1117/1.3251051)].

Using DRS during breast-conserving surgery: identifying robust optical parameters and influence of inter-patient variation

L.L. de Boer | B.H.W. Hendriks | F. van Duijnhoven
M.T.F.D. Vrancken Peeters-Baas | K. Van de Vijver | C.E. Loo Katarzyna Jóźwiak |
H.J.C.M. Sterenborg | T.J.M. Ruers

Successful breast-conserving surgery consists of complete removal of the tumor while sparing healthy surrounding tissue. Despite currently available imaging and margin assessment tools, recognizing tumor tissue at a resection margin during surgery is challenging. Diffuse reflectance spectroscopy (DRS), which uses light for tissue characterization, can potentially guide surgeons to prevent tumor-positive margins. However, inter-patient variation and changes in tissue physiology occurring during the resection might hamper this light-based technology. Here we investigate how inter-patient variation and tissue status (*in vivo* vs *ex vivo*) affect the performance of the DRS optical parameters. *In vivo* and *ex vivo* measurements of 45 breast cancer patients were obtained and quantified with an analytical model to acquire the optical parameters. The optical parameter representing the ratio between fat and water provided the best discrimination between normal and tumor tissue, with an area under the receiver operating characteristic curve of 0.94. There was no substantial influence of other patient factors such as menopausal status on optical measurements. Contrary to expectations, normalization of the optical parameters did not improve the discriminative power. Furthermore, measurements taken *in vivo* were not significantly different from the measurements taken *ex vivo*. These findings indicate that DRS is a robust technology for the detection of tumor tissue during breast-conserving surgery.

Keywords: Spectroscopy, Fiber optics, Tissue characterization.

INTRODUCTION

There is an unmet clinical need for intra-operative detection of tumor deposits on resection margins during breast-conserving surgery. Presently, surgeons still have to rely on their visual and tactile abilities when determining the optimal resection plane. Since tumor tissue in the breast is difficult to feel or see, the surgeon's capability to distinguish tumor tissue from normal tissue during surgery is often compromised. Leaving tumor tissue behind during surgery will increase the chance of developing local recurrence, and for this reason, in patients with tumor-positive resection margins additional measures such as re-excision or boost radiation therapy may be indicated. On the other hand, removing too much normal breast tissue will impair the cosmetic outcome. To guide the surgeon in this delicate balance, several margin assessment techniques are currently available. Traditionally used microscopic techniques, such as frozen section analysis and imprint cytology, however, proved to have limited accuracy [1]. Imaging methods, like intra-operative ultrasound, decrease the number of positive resection margins but still lack sensitivity for small tumor deposits [2]. A recently introduced margin evaluation tool is the MarginProbe® (DUNE Medical Devices, Paoli, PA, United States), which uses radiofrequency spectroscopy for positive resection margin detection. Currently, this technology is under evaluation for more widespread intraoperative use during breast-conserving surgery (BCS). First results indicate that with this device the number of re-excisions decreases although also larger volumes of tissue are resected [3, 4].

In an attempt to characterize tissue and perform margin assessment, optical technologies such as Raman spectroscopy [5, 6], fluorescence spectroscopy [7], optical coherence spectroscopy [8], and diffuse reflectance spectroscopy (DRS) [9-12], have been investigated. These optical technologies build on the principle of measuring optical responses, which are influenced by the structure and composition of the underlying tissue. Specifically, DRS with its ability to perform near real-time tissue characterization of tissue volumes of several mm³, without the need for exogenous contrast agents, seems a suitable candidate for margin assessment during breast-conserving surgery.

Much of the previously published data demonstrates the potential of DRS to discriminate normal tissue from tumor tissue with high sensitivity (74-79%) and specificity (78-93%). However, many of these studies have been conducted in a well-controlled *ex vivo* setting which does not necessarily resemble intra-operative circumstances when trying to discriminate normal tissue from tumor tissue in an individual patient.

One of the challenges that need to be addressed is the non-uniformity of breast tissue both in an individual patient and between patients. In previous publications, the heterogeneity of normal breast tissue is reported to range between 20%-40% for absorption (μ_a) and 5%-20% for scattering (μ_s) [13, 14]. Besides this intra-subject variation, also inter-subject variation is present in optical measurements of breast tissue [15, 16]. Various publica-

tions report on the influence of menopausal status on breast tissue, many of them indicating that breast tissue of pre-menopausal women has higher scattering values, water content, and concentrations of hemoglobin in comparison to post-menopausal women, whereas in the post-menopausal group lipid content is reduced [13, 14, 17, 18]. Furthermore, large inter-subject variation is found [19]. Such inter-patient variation could negatively affect the potential of optical DRS parameters to discriminate normal tissue from tumor tissue.

Another challenge may be the physiological differences between *in vivo* and *ex vivo* measurement. Due to the loss of blood flow and exposure to air, tissue physiology will be different in resected tissue. Such phenomena could impair optical parameters, especially those that are related to blood, thereby hindering one-on-one translation of *ex vivo* results to an *in vivo* setting [20, 21].

Both aspects, the inter-patient variation and changing physiology during the resection of the tissue, can hamper the use of optical spectroscopy for the detection of tumor tissue at a resection margin during breast-conserving surgery. To further investigate whether these aspects could hamper future clinical implementation of DRS in BCS we evaluated the optical parameters derived from the DRS measurements for the discrimination of normal tissue from tumor tissue and assessed the influence of inter-patient variation. Furthermore, we explore the potential of data normalization to diminish this phenomenon. Finally, we examined the influence of tissue status (*in vivo* vs *ex vivo*) on the optical measurements to investigate if the physiological changes occurring in tissue during surgery impede the use of DRS as an intra-operative tool.

MATERIALS & METHODS

Spectroscopy system

The optical system used for the spectroscopic measurements consisted of a broadband light source and two spectrometers for 1) the visual wavelength region (Andor Technology, DU420A-BRDD), and 2) the near-infrared wavelength region (Andor Technology, DU492A-1.7). The spectrometers and light source were controlled on an attached laptop with LabView and allowed measuring diffuse reflectance spectra between 400 and 1600 nm [22]. The measurement set-up was approved for use in the operating theater. The measurements were obtained through a beveled optical needle (18°) with a fiber distance of approximately 0.17 cm between the illuminating and collecting fiber at the tip of the needle (Figure 1). This needle allows to obtain optical data of tissue volumes of a few mm³.

Figure 1. Schematic image of data acquisition. A cannula with the optical needle is placed in normal tissue to acquire spectra at the first measurement location (a) and subsequently at the second measurement location (b). Afterward the second measurement location is marked with a twistmarker (c). The optical needle is then again introduced to obtain measurements of tumor tissue (d & e) and similarly, the second tumor measurement is marked with a twistmarker (f).

Data acquisition

Female patients diagnosed with an invasive carcinoma or *in situ* carcinoma, ductal or lobular, were included. The Medical Ethics Committee of the Netherlands Cancer Institute – Antoni van Leeuwenhoek approved the study and all patients signed informed consent.

Under ultrasound guidance, a coaxial 15/14G needle cannula (Invivo, Schwerin, Germany) was placed at the first measurement location. Through the cannula, the optical needle was inserted to collect *in vivo* measurements of non-malignant tissue at a distance of a few centimeters from the tumor. Once data acquisition at the first location had finished the cannula and the optical needle were moved forward to the second measurement location in normal tissue. After the spectra of the second measurement location had been acquired, the optical needle was retracted and through the same cannula, a twistmarker (BARD GmbH, Karlsruhe, Germany) was inserted to mark the second measurement location. Next, the optical needle was introduced a second time and the tip of the cannula and optical needle were positioned in the tumor. Similar to the measurements in normal tissue, two tumor locations were measured followed by the insertion of a second twistmarker (Figure 1). During the histopathologic evaluation, the pathologist, who was blinded for the outcome of the spectroscopic measurements, sampled tissue around the twistmarkers to confirm histopathology at the measurement locations (twistmarker verification). However, to facilitate straightforward correlation between histopathology and the measurement locations the study protocol was amended halfway allowing biopsies to be taken from the measurement locations (biopsy verification). Needle biopsies were obtained of the second normal measurement location and the second tumor measurement location.

Once the specimen was resected, the *ex vivo* measurements were obtained, using ultrasound to guide needle placement in normal and tumor measurement locations.

Data processing

The obtained spectra were quantified with an analytical model that was derived from the diffusion theory and has been described extensively in previous publications [22, 23]. In short, the acquired optical spectra of tissue and absorption spectra of known absorbers in the breast are provided as input to the model which uses a Trust Region algorithm to translate the obtained spectra into estimations of optical parameters (Matlab 2015a). Optical parameters evaluated by the model were; blood (%), oxygen saturation (%), scattering at 800 nm (cm^{-1}), fraction Mie scattering, β -carotene concentration, collagen fraction, fraction of fat ($[\text{fat}]/([\text{water}] + [\text{fat}])$), and total volume of water and fat ($[\text{water}] + [\text{fat}]$). The latter two were used to calculate the ratio between fat and water (F/W-ratio).

The spectra with more than 20% blood or more than 10% patent blue according to the fit results were excluded. More than 20% blood indicates pooling of blood around the tip of the optical needle whereas patent blue, a dye injected peritumoral for sentinel node detection, may confound data analysis even when included in the fitting model.

Analysis

Each optical parameter was modeled using the Generalized Estimating Equation (GEE) with clustering of all measurements obtained in the same individual subject and assuming equicorrelated structure of the covariance matrix for correlated measurements. This GEE method allows taking into account relations within individual measurements when assessing the association between optical parameters and covariates [24]. The GEE model evaluates a certain optical parameter by comparing two instances (e.g. normal vs tumor or *in vivo* vs *ex vivo*) while fixing all other covariates. For each optical parameter, the model returns a β -value and a p-value indicating significance. The β -value reflects the change of the optical parameter comparing the two instances. In example, a β -value of 10 of the optical parameter blood in the comparison of normal tissue with tumor tissue means that the average percentage blood increases with 10% between the measurements in normal tissue and the measurements in tumor tissue. Visa versa a negative β -value indicates a decrease. For each optical parameter, the β -values for the covariates can be compared.

Covariates included in these models were; 'Tissue Type' (normal or tumor tissue); 'Histological verification method' (twistmarker or biopsy); 'Tissue Status' (*in vivo* or *ex vivo*); 'Menopausal Status' (pre- or post-menopausal); 'Chemotherapy' (neo-adjuvant chemotherapy or not); and 'Hormonal Therapy' (neo-adjuvant hormonal therapy or not). Both neo-adjuvant chemotherapy and hormonal therapy were provided as covariates to the model because in literature effects of these treatments are described for both tumor tissue as well as non-tumor bearing breast tissue [25-30]. Patients who were peri-menopausal ($n = 7$) were included in the pre-menopausal group. A p-value of 0.05 or smaller was considered significant; all analyses were performed using Matlab R2015a software.

Table 1. Patient characteristics

		Twistmarker verification (n = 25)	Biopsy verification (n = 20)
Mean age (SD)		55.9 year (8.5)	58.3 year (12.1)
Menopausal status	Pre	13	6
	Post	12	14
Previous treatment	Chemotherapy	4	5
	Hormonal therapy	1	2
	None	21	17
Invasive carcinoma	Ductal	23*	17
	Lobular	3*	3

*in one patient histopathological evaluation showed a mixture of ductal and lobular carcinoma *in situ*.

For all optical parameters that were significant in the GEE analysis, median values were calculated at each measurement location and used to evaluate ROC curves and associated 95% confidence intervals in MedCalc. For calculation and comparison of the ROC, the methodology as described by DeLong et al. was used. [31]

To nullify the potential influence of inter-patient variation and thereby improve the discriminative power of the optical parameters, the optical measurements were normalized by using the patient as her own reference. Normalized data was computed by dividing the measurements from a certain location with the median of all the normal measurement of the same patient.

RESULTS

Table 1 lists the characteristics of all patients (n = 45) used for analysis. Across these patients, 1394 DRS spectra were obtained from 263 locations. The coefficient of variation of the spectra at a measurement locations was <25% for 82% of the measurement sites.

Measurement were obtained of normal tissue *in vivo* (n = 85 locations), normal tissue *ex vivo* (n = 73 locations), tumor tissue *in vivo* (n = 60 locations) and tumor tissue *ex vivo* (n = 45 locations). In incidental cases, the tumor could not be clearly identified in the *ex vivo* setting and in some cases, fewer measurements were obtained due to time restraints in the OR-schedule.

The β -values calculated in the GEE analysis are displayed in Table 2, in which the scorings of each covariate can be compared with each other for each optical parameter.

It is shown that the parameters F/W-ratio, β -carotene, blood, scattering at 800 nm, and fraction Mie scattering, significantly varied between tissue types (first column of Table 2). The optical parameters blood and scattering at 800 nm were higher for tumor tissue measurements than for normal tissue measurements. Conversely, F/W-ratio, β -carotene,

Table 2. β -values of the covariates in the GEE analysis. An asterisk indicates significance.

Parameter	Tissue Type ^a	Menopausal Status ^b	Chemo-therapy ^c	Hormonal Therapy ^c	Verification method ^d	Tissue Status ^e
F/W-ratio	-9.11*	1.56	-1.25	-3.17*	-2.40*	0.86
β -carotene (μM)	-4.37*	-0.93	0.69	1.07	-1.36	0.90
Blood (%)	3.06*	0.36	-0.33	-0.76	-0.21	0.14
Scat. 800 nm (cm^{-1})	9.19*	-3.85	0.43	-1.37	3.07	-0.31
Mie scat. fraction	-0.08*	-0.01	0.01	-0.09*	0.03	-0.01
Collagen (μM)	0.04	-0.02	-0.02	0.03	0.03	-0.03
StO ₂ (%)	-2.12	-6.21	1.79	4.69	6.55	-5.92

^aComparing tumor tissue to normal tissue, ^bComparing postmenopausal measurements to premenopausal measurements,

^cComparing neo-adjuvant therapy to no neo-adjuvant therapy, ^dComparing biopsy verification to twistmarker verification,

^eComparing *ex vivo* measurements to *in vivo* measurements.

and fraction Mie scattering were higher for normal tissue than for tumor tissue. Other optical parameters, StO₂, and collagen, were not significantly different between normal tissue and tumor tissue.

There were no significant differences between the optical measurements obtained in pre-menopausal and post-menopausal patients (second column Table 2).

F/W-ratio and Mie scattering fraction were significantly different comparing patients with hormonal therapy to those without hormonal therapy. However, the groups of patients who received neoadjuvant chemotherapy ($n = 9$) or hormonal therapy ($n = 3$) are small.

The comparison between the two verification methods (twistmarker or biopsy) was only significant for the F/W-ratio (fifth column Table 2), suggesting that measurements in the biopsy study were less fatty and/or contained more water. A more detailed analysis of the data revealed that the difference in F/W-ratio between the studies was mainly induced by the measurements of normal tissue. No specific reason for this difference could be identified.

Figure 2 illustrates the medians of all measurement locations obtained in an *in vivo* and *ex vivo* setting, represented as a boxplot as well as a scatter plot and obtained for the optical parameters that were significant in the GEE analysis. These graphs capture differences between the normal and tumor measurements but also demonstrate that there is variation in the measured values.

To evaluate the true potential of the optical parameters in this dataset to discriminate normal tissue from tumor tissue the Receiver Operator Characteristic (ROC) curves and corresponding Area Under Curve (AUC) values were calculated for the optical parameters that were significant in the GEE analysis (Figure 3). The F/W-ratio was able to discriminate normal tissue from tumor tissue with the highest value of the AUC of 0.943. The AUC values for β -carotene, blood, scattering at 800 nm and fraction Mie scattering were 0.824; 0.746; 0.715; and 0.679 respectively.

The data was normalized in an attempt to limit the influence of inter-patient variation. The ROC curves of the normalized data are depicted in Figure 4. Similar to the absolute data

Figure 2. Boxplot and scatter plot of the five parameters that were significant in the GEE analysis (F/W-ratio (a), β -carotene (b), blood (c), scattering at 800 nm (d), and fraction Mie scattering (e)). The figures comprise the medians from all measurement locations categorized based on the type of tissue (normal or tumor) and tissue status (*in vivo* or *ex vivo*).

Figure 3. ROC curves of the optical parameters that were significant in the GEE analysis (F/W-ratio (a), β -carotene (b), blood (c), scattering at 800 nm (d), and fraction Mie scattering (e)). The dotted lines around the ROC represent the 95% confidence intervals.

Figure 4. The ROC curves of the normalized data of the optical parameters (F/W-ratio (a), β -carotene (b), blood (c), scattering at 800 nm (d), and fraction Mie scattering (e)). The dotted lines represent the 95% confidence intervals.

Figure 5. The F/W-ratio ROC curves of the *in vivo* measurements (green) and the *ex vivo* measurements (red). The dotted lines represent the 95% confidence interval.

the F/W-ratio was the best discriminator in the normalized data with an AUC of 0.951. All other optical parameters performed slightly worse with AUCs of; 0.826 (β -carotene); 0.751 (blood); 0.752 (scattering at 800 nm); and 0.647 (fraction Mie scattering).

The AUCs of the absolute data and the normalized data were compared in MedCalc with a ROC comparison method [31], to investigate the benefit of normalization. The p-values for the comparison between the absolute data and the normalized data were; 0.68 (F/W-ratio), 0.96 (β -carotene), 0.91 (blood), 0.43 (scattering at 800 nm) and, 0.53 (fraction Mie scattering); indicating that there were no significant differences between the AUC of the absolute data and the normalized data.

Since tissue status (*in vivo* or *ex vivo*) could potentially be a confounding factor in the discrimination between normal tissue and tumor tissue this variable was also included in the GEE analysis. The sixth column of Table 2 comprises the results for the comparison between the *in vivo* and *ex vivo* data and presents no significant differences in all the optical parameters. As can be seen in Figure 5 the ROC curves and the 95% confidence intervals of the *in vivo* and *ex vivo* data of the F/W-ratio almost entirely overlap (p-value: 0.20) which confirms that discriminative power is not significantly different between the *in vivo* and *ex vivo* data.

DISCUSSION

In this study, we investigated the usability of optical spectroscopy for margin assessment during breast-conserving surgery by evaluating the discriminative power of optical parameters for tissue differentiation between normal breast tissue and tumor tissue. In this evaluation, we took in consideration confounding factors such as menopausal status (pre-

or post-menopausal), previous treatments (neo-adjuvant chemotherapy/hormonal therapy or not), verification method (twistmarker or biopsy) and tissue status (*in vivo* or *ex vivo*) that might have an influence on the optical parameters. ROC-curves were calculated of those optical parameters that were significant in the GEE analysis to consider the true potential to discriminate normal tissue from tumor tissue in this dataset. Additionally, the focus was on the effect of data normalization to improve discriminative power of the optical parameters, as well as on the effect of a change in tissue status (from *in vivo* to *ex vivo*).

The optical parameter F/W-ratio provided the best discrimination between normal tissue and tumor tissue in this study, which is in line with previously published results [32, 33]. Other optical parameters, i.e. β -carotene, blood, scattering at 800 nm and fraction Mie scattering displayed significance in the GEE analysis and revealed AUC values in ROC curves between 0.647 and 0.826. This demonstrates that these optical parameters have some potential to discriminate measurements of normal tissue from measurements of tumor tissue but are less distinctive than the F/W-ratio.

Experience learned that sometimes the twistmarker was difficult to find for the pathologist. To avoid problems, the measurement protocol was amended and thereafter biopsies were taken from measurement sites. The influence of this alteration in protocol was assessed by incorporating verification method as a covariate in the GEE analysis. Except for F/W-ratio, all optical parameters were comparable between data acquired in the two protocols. Since the significant difference between the twistmarker verification and the biopsy verification was only present in the measurements of normal tissue, we hypothesize that the density of the breast tissue of the women in the biopsy study might have been higher compared to the density of those women in the twistmarker study despite their comparability in age (Table 1).

Menopausal status, although highly correlated with involution of glandular tissue in the breast, did not significantly influence our results, similar to the results of Pogue et al [19]. Potentially this is caused by the fact that, although involution of glandular tissue is correlated to menopausal status this does not necessarily imply that post-menopausal women cannot have dense breast tissue. Stomper et al. reported a parenchymal breast density of >50% in 26% of cases in a group of women aged 55 years and above. [34]

We included both neo-adjuvant chemo- and hormonal therapy in the analysis since previous research found an influence of these therapies on the optical characteristics of both the normal breast tissue and tumor tissue. Whereas F/W-ratio and fraction Mie scattering were correlated with previous hormonal therapy, previous chemotherapy did not affect any of the optical parameters. The number of patients in both groups was small and any definitive conclusions on the influence of neo-adjuvant therapy on optical measurements cannot be drawn from these numbers. It should be noted that the primary goal of this study was not to assess the influence of menopausal status or neo-adjuvant therapy on the optical parameters as these factors were only included to control the differences in

optical parameters between normal and tumor tissue. In other words, incorporating them as covariates in the GEE analysis ensured the potential influence of these factors was not falsely attributed to optical differences between normal and tumor tissue. In other studies, also BMI and breast density information were used in covariate adjustment [35], however, for the patients included in this study, this information was not available. In future studies, we will record both BMI and breast density for all patients.

To improve the discriminative power of the optical parameters and nullify the influence of inter-patient variation we normalized the individual patient data. Previous work by our group used a similar approach and found an increase in both sensitivity and specificity in the individual patient data analysis compared to the cohort data analysis [33]. However, in this research, a Classification and Regression tree (CART) algorithm with Leave-One-Out (LOO) cross-validation was used which can be prone to biased parameter selection resulting in improved sensitivity and specificity that is not necessarily associated with differences in optical characteristics of normal tissue and tumor tissue. In the current study, the measurements were normalized by dividing them with the average value of normal tissue of the patient, the CART classification was omitted, and instead, ROC curves were calculated and compared. In contrast to what was expected, the normalization step did not improve or deteriorate the ability of any of the optical parameters to discriminate normal tissue from tumor tissue. This could suggest that the influence of inter-patient variation is negligible in this dataset since otherwise, the normalization step would have improved the discriminative power. However, potentially not only inter-patient variation but also intra-patient variation influences the optical measurements since the composition and structure of breast tissue types in each individual may vary as well. The performed normalization step did contribute to diminishing the influence of inter-patient variation but did not fully address the heterogeneity observed in tissue types of an individual patient, which might be a more important factor. Although research by Taroni et al. concludes that the optically derived parameters of a single point time-resolved measurement at a central position in the breast are comparable to a full scan of the breast [35], this probably did not apply to the measurements in this research. Presumably the fact that the measurement volumes of a single measurement in the study of Taroni et al. were much larger compared to the probed volume in this study can explain why in the current study using a single measurement as a reference could not account for all intra-patient variation. Obtaining measurements on a number of sites might be necessary to address the intra-patient heterogeneity with our optical measurement set-up.

As for tissue status, measuring *in vivo* or *ex vivo* did not result in significantly different measurements for any of the optical parameters. The ROC curve and confidence intervals of the F/W-ratio of the *in vivo* measurements overlap with those of the *ex vivo* measurements indicating that this optical parameter had the same outstanding performance in both settings. A direct one-on-one comparison of the *in vivo* with *ex vivo* measurements could, however, not be performed since the sites measured *ex vivo* on the tissue sample were not

exactly the same as the locations of the *in vivo* measurements. This is a clear limitation of the study.

Comparisons between *in vivo* and *ex vivo* DRS measurements of normal and tumor tissue are sparse in literature. Many studies are performed in animals comprising only measurements of normal tissue [20, 21], or data is obtained with different optical technology such as Raman [36], fluorescence [21] or electrical and impedance spectroscopy [37]. Optical differences between *in vivo* and *ex vivo* measurements that are described in these studies are mainly attributed to the change of oxygenation state of hemoglobin and blood loss of the tissue. In a study comparing site-matched *in vivo* and *ex vivo*, DRS measurements of adenocarcinoma and normal tissue in pancreatic cancer patients the measurements of the two tissue statuses were found to be qualitatively and quantitatively in agreement [38]. To our knowledge, the difference between *in vivo* and *ex vivo* measurements of breast tissue, both normal tissue, and tumor tissue has not been investigated before. Bydlon et al. did comment on the use of hemoglobin saturation as a parameter for margin assessment. Since excessive changes in oxygenated and deoxygenated hemoglobin were found post-excision, due to oxygen being consumed by metabolic active tissue immediately after excision, they concluded that hemoglobin saturation is an unreliable parameter for this purpose [39]. In this study, measurements were obtained directly after resection, which might be different to other studies where the time between resection and measurement is longer. It should also be noted that it is difficult to determine whether a measurement should be considered as *in vivo* or *ex vivo* since tissue characteristics will alter immediately when tissue is manipulated during the resection (e.g. changes in blood flow, secretion of cytokines).

Besides the previous limitations regarding the changes in protocol and the fact that the *in vivo* and *ex vivo* measurements were not acquired at the exact same measurement spot, another important limitation is the correlation with the histopathology. In the twistmarker study tissue for histopathological evaluation was obtained close to the twistmaker, however, this might not have been the exact same location. Something similar concerns the biopsy study when two different needles were used for measuring and taking the biopsy. Tissue can be classified reliably as 'normal' or 'tumor' since ultrasound imaging was used to guide the needle and confirm placement in the targeted tissue. Direct correlation between the optical measurements and the actual histopathology at a measurement site, however, was not possible. This was an important reason why the measurements were not subdivided into more detailed groups than normal and tumor. Since normal tissue consists of both fatty tissue and fibrous/glandular tissue in varying ratios, classifying all these measurement locations as normal might be considered as somewhat limited. To be able to claim that DRS can also be used for discriminating all types of tissue present in the breast a more refined classification of the measurement locations based on the direct correlation between optical measurements and histopathology is necessary. The fact that normal measurements can have different underlying tissue types may also explain the heterogeneity seen in the normal tissues as described previously in the discussion.

The results of the current study indicate that the F/W-ratio outperformed all other optical parameters in discriminating between normal tissue and tumor tissue with an AUC of 0.94. Furthermore, it was concluded that this parameter was not affected by tissue status because the performance in the *in vivo* and *ex vivo* setting was similar. For the application of DRS as a margin assessment tool, these results imply that DRS technology based on F/W-ratios is a suitable candidate for intra-operative use, even if during the resection tissue physiology will change. The data was normalized in an attempt to overcome inter-patient variation. However, the normalization of data did not result in the hypostasized enhanced performance of the optical parameters. Potentially increasing the number of normal measurement that is used for determining a reference value can help in establishing a more robust baseline measurement of normal tissue. The results of this study support the potential use of DRS during breast-conserving surgery for the detection of tumor on resection margins.

ACKNOWLEDGMENTS

We would like to thank all people that helped during the data acquisition in the OR and especially Danny Evers and Paulien Stegehuis. We also thank the members of the Philips Research project for their contribution.

REFERENCES

- [1] K. Esbona, Z. Li, and L. G. Wilke, "Intraoperative Imprint Cytology and Frozen Section Pathology for Margin Assessment in Breast Conservation Surgery: A Systematic Review," *Annals of Surgical Oncology*, vol. 19, no. 10, pp. 3236-3245, 2012, [[10.1245/s10434-012-2492-2](#)].
- [2] H. Eggemann, T. Ignatov, S. Dan Costa, and A. Ignatov, "Accuracy of ultrasound-guided breast-conserving surgery in the determination of adequate surgical margins," *Breast Cancer Research and Treatment*, vol. 145, pp. 129-136, 2014, [[10.1007/s10549-014-2932-8](#)].
- [3] M. Sebastian, S. Akbari, B. Anglin, E. H. Lin, and A. M. Police, "The impact of use of an intraoperative margin assessment device on re-excision rates," *SpringerPlus*, vol. 4, no. 198, 2015, [[10.1186/s40064-015-0801-5](#)].
- [4] J.-U. Blohmer, J. Tanko, J. Kueper, J. Groß, R. Völker, and A. Machleidt, "MarginProbe reduces the rate of re-excision following breast conserving surgery for breast cancer," *Archives of Gynecology and Obstetrics*, 2016, [[10.1007/s00404-016-4011-3](#)].
- [5] A. S. Haka, Z. Volynskaya, J. A. Gardecki, J. Nazemi, J. Lyons, D. Hicks, M. Fitzmaurice, R. R. Dasari, J. P. Crowe, and S. M. Feldman, "In vivo Margin Assessment during Partial Mastectomy Breast Surgery Using Raman Spectroscopy," *Cancer Research*, vol. 66, no. 6, pp. 3317-3322, 2006, [[10.1158/0008-5472.CAN-05-2815](#)].
- [6] K. Kong, F. Zaabar, E. Rakha, I. Ellis, A. Koloydenko, and I. Notingher, "Towards intra-operative diagnosis of tumours during breast conserving surgery by selective-sampling Raman micro-spectroscopy," *Physics in Medicine and Biology*, vol. 59, pp. 6141-6152, 2014, [[10.1088/0031-9155/59/20/6141](#)].
- [7] M. D. Keller, S. K. Majumder, M. C. Kelley, I. M. Meszoely, F. I. Boulos, G. M. Olivares, and A. Mahadevan-Jansen, "Autofluorescence and Diffuse Reflectance Spectroscopy and Spectral Imaging for Breast Surgical Margin Analysis," *Lasers in Surgery and Medicine*, vol. 42, pp. 15-23, 2010, [[10.1002/lsm.20865](#)].
- [8] A. M. Zysk, K. Chen, E. Gabrielson, L. Tafta, E. A. May Gonzalez, J. K. Canner, E. B. Schneider, A. J. Cittadine, P. S. Carney, S. A. Boppart, K. Tsuchiya, K. Sawyer, and L. K. Jacobs, "Intraoperative Assessment of Final Margins with a Handheld Optical Imaging Probe During Breast-Conserving Surgery May Reduce the Reoperation Rate: Results of a Multicenter Study," *Surgical Oncology*, vol. 22, no. 10, pp. 3356-3362, 2015, [[DOI 10.1245/s10434-015-4665-2](#)].
- [9] J. Q. Brown, L. G. Wilke, J. Geradts, S. A. Kennedy, G. M. Palmer, and N. Ramanujam, "Quantitative optical spectroscopy: a robust tool for direct measurement of breast cancer vascular oxygenation and total hemoglobin content in vivo," *Cancer Research*, vol. 69, no. 7, pp. 2919-26, 2009, [[10.1158/0008-5472.CAN-08-3370](#)].
- [10] N. Lue, J. W. Kang, C. Yu, I. Barman, N. C. Dingari, M. S. Feld, R. R. Dasari, and M. Fitzmaurice, "Portable Optical Fiber Probe-Based Spectroscopic Scanner for Rapid Cancer Diagnosis: A New Tool for Intraoperative Margin Assessment," *PloS One*, vol. 7, no. 1, pp. e30887-1 - e30887-10, 2012, [[10.1371/journal.pone.0030887](#)].
- [11] A. M. Laughney, V. Krishnaswamy, E. J. Rizzo, M. C. Schwab, R. J. Barth Jr, D. J. Cuccia, B. Tromberg, K. D. Paulsen, B. W. Pogue, and W. A. Wells, "Spectral discrimination of breast pathologies *in situ* using spatial frequency domain imaging," *Breast Cancer Research*, vol. 15, pp. R61, 2013.
- [12] B. S. Nichols, C. E. Schindler, J. Q. Brown, L. G. Wilke, C. S. Mulvey, M. S. Krieger, J. Gallagher, J. Geradts, R. A. Greenup, J. A. Von Windheim, and N. Ramanujam, "A Quantitative Diffuse Reflectance Imaging (QDRI) System for Comprehensive Surveillance of the Morphological Landscape in Breast Tumor Margins," *PloS One*, vol. 10, no. 16, pp. e0127525, 2015, [[10.1371/journal.pone.0127525](#)].
- [13] T. Svensson, J. Swartling, P. Taroni, A. Torricelli, P. Lindblom, C. Ingvar, and S. Andersson-Engels, "Characterization of normal breast tissue heterogeneity using time-resolved near-infrared

- spectroscopy," *Physics in Medicine and Biology*, vol. 50, pp. 2559-2571, 2005, [[10.1088/0031-9155/50/11/008](#)].
- [14] N. Shah, A. E. Cerussi, D. Jakubowski, D. Hsiang, J. Butler, and B. Tromberg, "Spatial variations in optical and physiological properties of healthy breast tissue," *Journal of Biomedical Optics*, vol. 9, no. 3, pp. 534-540, 2004, [[10.1117/1.1695560](#)].
 - [15] P. Taroni, G. Quarto, A. Pifferi, F. Abbate, N. Balestreri, S. Menna, E. Cassano, and R. Cubeddu, "Breast Tissue Composition and Its Dependence on Demographic Risk Factors for Breast Cancer: Non-Invasive Assessment by Time Domain Diffuse Optical Spectroscopy," *PloS One*, vol. 10, no. 6, pp. e0128941, 2015, [[10.1371/journal.pone.0128941](#)].
 - [16] R. L. P. Van Veen, H. J. C. M. Sterenborg, A. W. K. S. Marinelli, and M. Menke-Pluymers, "Intra-operatively assessed optical properties of malignant and healthy breast tissue, to determine the optimum wavelength of contrast for optical mammography," *Journal of Biomedical Optics*, vol. 6, pp. 1129-1136, 2004.
 - [17] S. Kennedy, J. Geradts, T. Bydlon, J. Q. Brown, J. Gallagher, M. Junker, W. Barry, N. Ramanujam, and L. Wilke, "Optical breast cancer margin assessment: an observational study of the effects of tissue heterogeneity on optical contrast," *Breast Cancer Research*, vol. 12, no. 6, pp. R91, 2010, [[10.1186/bcr2770](#)].
 - [18] D. R. Leff, O. J. Warren, L. C. Enfield, A. Gibson, T. Athanasiou, D. K. Patten, J. Hebden, G. Z. Yang, and A. Darzi, "Diffuse optical imaging of the healthy and diseased breast: A systematic review," *Breast Cancer Research and Treatment*, vol. 108, pp. 9-22, 2008, [[10.1007/s10549-007-9582-z](#)].
 - [19] B. W. Pogue, S. Jiang, H. Dehghani, C. Kogel, S. Soho, S. Srinivasan, X. Song, T. D. Tosteson, S. P. Poplack, and K. D. Paulsen, "Characterization of hemoglobin, water, and NIR scattering in breast tissue: analysis of intersubject variability and menstrual cycle changes," *Journal of Biomedical Optics*, vol. 9, no. 3, pp. 541-552, 2004, [[10.1117/1.1691028](#)].
 - [20] E. Salomatina, and A. N. Yaroslavsky, "Evaluation of the in vivo and ex vivo optical properties in a mouse ear model," *Physics in Medicine and Biology*, vol. 53, pp. 2797-2807, 2008, [[10.1088/0031-9155/53/11/003](#)].
 - [21] G. M. Palmer, C. L. Marshak, K. M. Vrotsos, and N. Ramanujam, "Optimal Methods for Fluorescence and Diffuse Reflectance Measurements of Tissue Biopsy Samples," *Lasers in Surgery and Medicine*, vol. 30, pp. 191-200, 2002, [[10.1002/lsm.10026](#)].
 - [22] R. Nachabé, B. H. W. Hendriks, A. E. Desjardins, M. van der Voort, M. B. van der Mark, and H. J. C. M. Sterenborg, "Estimation of lipid and water concentrations in scattering media with diffuse optical spectroscopy from 900 to 1600 nm," *J Biomed Opt*, vol. 15, no. 3, pp. 037015, 2010, [[10.1117/1.3454392](#)].
 - [23] R. Nachabé, B. H. W. Hendriks, M. Van der Voort, A. E. Desjardins, and H. J. C. M. Sterenborg, "Estimation of biological chromophores using diffuse optical spectroscopy; benefit of extending the UV-VIS wavelength range to include 1000 to 1600nm," *Optics Express*, vol. 18, no. 24, pp. 1432-1442, 2010, [[10.1364/BOE.1.001432](#)].
 - [24] J. Shults, and S. J. Ratcliffe, "Analysis of multi-level correlated data in the framework of generalized estimating equations via xtmultcorr procedures in Stata and qls functions in Matlab," *Stat. Interface*, vol. 2, pp. 187-196, 2009.
 - [25] S. Sahoo, and S. C. Lester, "Pathology of Breast Carcinomas After Neoadjuvant Chemotherapy," *Archives of Pathology and Laboratory Medicine*, vol. 133, pp. 633-642, 2009.
 - [26] T. D. O'Sullivan, A. Leproux, J. Chen, S. Bahri, A. Matlock, D. Roblyer, C. E. McLaren, W. Chen, A. E. Cerussi, M. Su, and B. J. Tromberg, "Optical imaging correlates with magnetic resonance imaging breast density and reveals composition changes during neoadjuvant chemotherapy," *Breast Cancer Research*, vol. 15, no. R14, 2013.

- [27] L. Enfield, G. Cantanhede, V. Ramalingam, A. Purushotham, J. Hebden, and A. Gibson, "Monitoring the response to neoadjuvant hormone therapy for locally advanced breast cancer using three-dimensional time-resolved optical mammography," *Journal of Biomedical Optics*, vol. 18, no. 5, pp. 056012, 2013, [[10.1117/1.JBO.18.5.056012](#)].
- [28] J.-H. Chen, H. Yu, M. Lin, R. S. Mehta, and M.-Y. Su, "Background Parenchymal Enhancement in the Contralateral Normal Breast of Patients Undergoing Neoadjuvant Chemotherapy Measured by DCE-MRI," *Magnetic Resonance Imaging*, vol. 31, no. 9, pp. 1465-1471, 2013, [[10.1016/j.mri.2013.07.014](#)].
- [29] K. Kerlikowske, A. J. Cook, D. S. M. Buist, S. R. Cummings, C. Vachon, P. Vacek, and D. L. Miglioretti, "Breast Cancer Risk by Breast Density, Menopause, and Postmenopausal Hormone Therapy Use," *Journal of Clinical Oncology*, vol. 28, no. 24, pp. 3830-3837, 2010, [[10.1200/JCO.2009.26.4770](#)].
- [30] A. Sadeghi-Naini, E. Vorauer, L. Chin, O. Falou, W. T. Tran, F. C. Wright, S. Gandhi, M. J. Yaffe, and G. J. Czarnota, "Early detection of chemotherapy-refractory patients by monitoring textural alterations in diffuse optical spectroscopic images," *Medical Physics*, vol. 42, no. 11, pp. 6130-6146, 2015, [[10.1118/1.4931603](#)].
- [31] E. R. DeLong, D. M. DeLong, and D. L. Clarke-Pearson, "Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach," *Biometrics*, vol. 44, no. 3, pp. 837-845, 1988.
- [32] L. L. DeBoer, B. G. Molenkamp, T. M. Bydlon, B. H. W. Hendriks, J. Wesseling, H. J. C. M. Sterenborg, and T. J. M. Ruers, "Fat/Water ratios measured with diffuse reflectance spectroscopy to detect breast tumor boundaries," *Breast Cancer Research and Treatment*, vol. 152, no. 3, pp. 509-518, 2015, [[10.1007/s10549-015-3487-z](#)].
- [33] D. J. Evers, R. Nachabé, H. M. Klomp, J. W. Van Sandick, M. W. Wouters, G. W. Lucassen, B. H. W. Hendriks, J. Wesseling, and T. J. M. Ruers, "Diffuse Reflectance Spectroscopy: A New Guidance Tool for Improvement of Biopsy Procedures in Lung Malignancies," *Clinical Lung Cancer*, vol. 13, no. 6, pp. 424-431, 2012, [[10.1016/j.clcc.2012.02.001](#)].
- [34] P. C. Stomper, D. J. D'Souza, P. A. DiNitto, and M. A. Arredondo, "Analysis of Parenchymal Density on Mammograms in 1353 Women 25-79 Years Old," *AJR: American Journal of Roentgenology*, vol. 167, pp. 1261-1265, 1996.
- [35] P. Taroni, A. Pifferi, G. Quarto, L. Spinelli, A. Torricelli, F. Abbate, N. Balestreri, S. Ganino, S. Menna, E. Cassano, and R. Cubeddu, "Effects of tissue heterogeneity on the optical estimate of breast density," *Biomedical Optics Express*, vol. 3, no. 10, pp. 2411-2418, 2012.
- [36] L. Raniero, R. A. Canevari, L. N. Z. Ramalho, F. S. Ramalho, E. A. P. dos Santos, R. A. Bitar, K. J. Jalkanen, H. S. Martinho, and A. A. Martin, "In and ex vivo breast disease study by Raman spectroscopy," *Theoretical Chemistry Accounts*, 2011, [[10.1007/s00214-011-1027-4](#)].
- [37] R. J. Halter, T. Zhou, P. M. Meaney, A. Hartov, R. J. Barth Jr, K. M. Rosenkranz, W. A. Wells, C. A. Kogel, A. Borsic, E. J. Rizzo, and K. D. Paulsen, "The correlation of in vivo and ex vivo tissue dielectric properties to validate electromagnetic breast imaging: initial clinical experience," *Physiological Measurement*, vol. 30, no. 6, pp. S121-S136, 2009, [[10.1088/0967-3334/30/6/S08](#)].
- [38] W. R. Lloyd, R. H. Wilson, S. Y. Lee, M. Chandra, B. McKenna, D. Simeone, J. Scheiman, and M.-A. Mycek, "In vivo optical spectroscopy for improved detection of pancreatic adenocarcinoma: a feasibility study," *Biomedical Optics Express*, vol. 5, no. 1, pp. 9-15, 2014, [[10.1364/BOE.5.000009](#)].
- [39] T. M. Bydlon, W. T. Barry, S. A. Kennedy, J. Q. Brown, J. E. Gallagher, L. G. Wilke, J. Geradts, and N. Ramanujam, "Advancing optical imaging for breast margin assessment: an analysis of excisional time, cautery, and patent blue dye on underlying sources of contrast," *PloS One*, vol. 7, no. 12, pp. e51418, 2012, [[10.1371/journal.pone.0051418](#)].

5

A method for co-registration of optical measurements of breast tissue with histopathology: the importance of accounting for tissue deformations

L.L. de Boer * | E. Kho * | J. Nijkamp | K.K. Van de Vijver | H.J.C.M. Sterenborg
L.C. ter Beek | T. J.M. Ruers

*These authors contributed equally to this work.

For the validation of novel optical diagnostic technologies, experimental results need to be benchmarked against the gold standard. Currently, the gold standard for tissue characterization is assessment of H&E sections by a pathologist. When processing tissue into hematoxylin and eosin (H&E) stained sections, the shape of the tissue deforms with respect to the initial shape when it was optically measured. In this study, we demonstrate the importance of accounting for these tissue deformations when correlating optical measurement with histopathology. We propose a method to register the tissue in the H&E sections to the optical measurements, which corrects for these tissue deformations. We compare the registered H&E sections to H&E sections that were registered with an algorithm that does not account for tissue deformations by evaluating both the shape and the composition of the tissue and using micro-computer tomography data as an independent measure. The proposed method, which did account for tissue deformations, was more accurate than the method that did not account for tissue deformations. These results emphasize the need for a registration method that accounts for tissue deformations, such as the method presented in this study, which can aid in validating optical techniques for clinical use.

Keywords: gold standard, histopathology, diffuse reflectance, registration algorithm, optical techniques, validation

INTRODUCTION

Optical technologies can potentially be used for tissue characterization in, for example, the detection of positive resection margins during cancer surgery. These technologies are based on the principle that light, which has undergone interaction with the tissue, contains information regarding the composition, morphology, and microvasculature of the tissue that can be used to discriminate different tissue types. Validation of these optical measurements is done by comparing the experimental results with the gold standard.[1] Currently, the gold standard for tissue characterization is an assessment of microscopy sections, which are stained with hematoxylin and eosin (H&E), by a pathologist. However, the tissue in the H&E sections, used for histopathological assessment, is profoundly deformed compared to the appearance of the tissue during the optical measurements. These deformations are introduced when the tissue is processed into H&E sections as explained and illustrated in Figure 1. Consequently, overlaying a digitalized H&E sections over images that reflect optical measurements is prone to registration errors.[2-5]

In our research group, we use Diffuse Reflectance Spectroscopy (DRS) for tissue characterization by means of a camera (HyperSpectral Imaging (HSI)) or fiber-optic probe (point measurements). With this optical technology, diffusely reflected light is measured after it has undergone multiple absorption and scattering events in the tissue. Therefore, a DRS measurement will contain information regarding the composition, morphology, and microvasculature of the tissue.

Registration errors between a DRS measurement and histopathology could negatively affect the assessment of the accuracy of this technology. The problem becomes even more relevant when developing classification algorithms, in which adding the correct histopathology label to each measurement is crucial. Accurate correlation between the optical measurements and histopathology is acknowledged as an important challenge encountered during clinical validation of optical technologies.[1, 6, 7] However, there is only limited literature focusing on correlating optical measurements with histopathology,[8-10] and when a method for correlation the optical data to histopathology is described, the reported registration methods do not seem to correct for tissue deformation.[11-13]

In this article, we demonstrate the importance of accounting for tissue deformations when correlating optical measurements with histopathology. We propose a registration method that takes into account these tissue deformations. For this purpose, data were obtained from fresh *ex vivo* breast tissue slices: the optical measurements (both hyperspectral images and fiber-optic probe measurements) and a white light image were acquired simultaneously. In addition, micro-computed tomography (μ CT) data was obtained right after the optical measurements. Therefore, we assumed that the tissue did not deform between the acquisition of the optical measurements, the white light image, and the μ CT data. As such, both the optical measurements and μ CT data were registered to the white light images.

Figure 1. Histopathological processing of the tissue. After arrival at the pathology department, the margins of the specimen are colored and the tissue is dissected into tissue slices (~ 3 mm thick). (a) shows the white light image of the tissue slice that is selected for optical measurements. After these measurements, the slice is imaged with the μ CT and brought back to the pathology department, fixed in formalin, and embedded in paraffin. (b) shows the white light image of the paraffin embedded tissue slice. From this paraffin embedded tissue slice, $\sim 3\mu\text{m}$ thin sections are cut, which are stained with H&E stain. (c) shows the H&E section of the tissue slice, which is available a few days after surgery. Due to the histopathological processing of the tissue, (c) is deformed in comparison to (a).

A few days after the optical measurements, histopathological information (H&E section) was available. Our proposed registration method consists of registering the H&E section to the white light image using a combination of an affine and deformable registration. We compared the correlation accuracy of our proposed registration method with the correlation accuracy after using only an affine registration method, which does not account for tissue deformations. We quantified this accuracy using the μ CT data as an independent modality. Finally, we compared the registered H&E sections using both registration methods with our optical measurements, and show that for an accurate correlation, tissue deformations should be taken into account when correlating optical measurements with histopathology.

MATERIALS & METHODS

Study procedure

Breast specimens were obtained from patients undergoing breast surgery, either breast-conserving or mastectomy, at the Antoni van Leeuwenhoek Hospital. The study was approved by the Institutional Review Board of the Netherlands Cancer Institute/Antoni van Leeuwenhoek, and according to Dutch law (WMO) no written informed consent was required. The specimens were brought to the pathology department immediately after surgery. There, after coloring of the resection margins, the specimens were frozen and sliced

in a bread-loafed manner. One tissue slice was placed in a macrocassette. Black rubber was placed underneath the tissue to ensure that measured spectra only originated from the tissue.

White light images

Prior to the DRS measurements, an overview white light image was acquired of the specimen in the macrocassette. This image was used as a reference to register the H&E image; because the tissue remained in the macrocassette between capturing this white light image and the DRS measurements, the tissue did not deform.

DRS measurements

Hyperspectral imaging

Hyperspectral images were obtained with a push-broom hyperspectral imaging setup (VLNIR CL-350-N17E, Specim, Spectral Imaging Ltd., Finland) that captures light in the near-infrared (~900-1700 nm, 256 wavelength bands, 5 nm increments) with an InGaAs sensor (320 x 256 pixels). The measurement set-up (shown in Figure 2a) and measurement calibration is described in a prior publication.[14] In short, the tissue is placed under the camera and imaged line-by-line through moving the imaged scene. Thereby a 3D hypercube is obtained that contains multiple 2D images of the tissue at different wavelengths, i.e. each pixel in the 2D image contains a full diffuse reflectance spectrum (Figure 2b). Raw hyperspectral data obtained from the tissue was normalized to a diffuse reflectance percentage relative to Spectralon® (SRT-99-100, Labsphere, Inc., Northern Sutton, New Hampshire).

Fiber-optic probe

Fiber-optic point measurements were acquired with a measurement set-up (Figure 2b) including a broadband light source, a spectrometer for the visual wavelength region (Andor Technology DU420A-BRDD) and a spectrometer for the near-infrared region (Andor Technology, DU492A-1.7). A fiber-optic probe was attached to the measurement set-up to measure diffuse reflectance spectra between 400 and 1600 nm (Figure 2d).[15, 16] The distance between the illuminating and collecting fiber at the tip of the probe was 1 mm. Before measuring, the set-up was calibrated by acquiring a white reference measurement of Spectralon® (SRT-99-100, Labsphere, Inc., Northern Sutton, New Hampshire) in a calibration cap.

To acquire a DRS measurement, the probe was brought in contact with the tissue. During the probe measurements, a custom-made grid with holes (consisting of 2 pieces and shown in Figure 2c) was placed on top of the tissue to gently fixate the tissue in the cassette and allow correlation between the measurement locations and H&E section afterward. The grid was 3D printed and transparent. In addition, the grid stabilized the probe while ensuring contact between the probe and the tissue and restraining it from moving during measurements.

Figure 2. Optical measurement set-ups. (a) The tissue was placed on the translation stage and imaged line-by-line through moving the stage. Thereby, a 3D hypercube was obtained in which each pixels contains a diffuse reflectance spectrum (b). Next, the tissue was measured with the fiber-optic probe (c). To retrieve measurement locations on the H&E section, two custom-made grids with holes of the exact size of the optical probe were used. Thereby, a diffuse reflectance spectrum was obtained (d).

Micro-computed tomography (μ CT) data

After the optical measurements, the grid was removed and the μ CT data was acquired. During the optical measurements and μ CT imaging, the tissue remained in the macrocassette and the black rubber was not removed. With the SkyScan[®] 1275B (Bruker, Kontich, Belgium), μ CT data was acquired using either 40 kV source voltage and 250 μ A source current, or 50 kV source voltage and 200 μ A source current and a 1 mm Aluminum filter. The exposure time varied between 39 and 105 ms, and one scan consisted of 1801 (0.2° increments) to 901

Figure 3. The steps required to register the H&E image to the white light (WL) image.

(0.4° increments) projections. Reconstruction of the data was performed with 30-50 μm resolution using SkyScan's NRecon software.

H&E sections

After measuring, the tissue slice was brought back to the pathology department and processed in H&E sections according to standard protocol. The side of the specimen used for the optical measurements was marked with methylene blue to ensure that the H&E section was taken from the same side. After digitalizing the H&E sections with the Aperio® ScanScope AT2 (Leica Biosystems, Wetzlar, Germany), tumor tissue was labeled in the H&E images by a pathologist. The remaining tissue in the H&E sections, fat tissue, and connective tissue, was labeled by thresholding the RGB channels of the H&E images since fat tissue is transparent and therefore white in H&E sections and connective tissue colors pink.

Registration of H&E section

The registration of H&E sections to the optical measurements required multiple steps, which are shown in Figure 3. All image registrations were made using MATLAB 2018a® (MathWorks, Natick, Massachusetts, United States). To prevent possible bias, the H&E image without the tissue labels was used throughout the process. Subsequently, after finishing the registration, the same registration steps were applied to the H&E image with the tissue labels. First, the background of both the H&E image and the white light image was manually removed (Figure 3, step 1). Second, the H&E image was rescaled using the shape of the tissue in both images such that the size of the tissue matched the size of the tissue in the white light image (Figure 3, step 2).

Affine registration

To compensate for rotation between the H&E image and the white light image, an affine registration was made between the two images. This was accomplished by using the shape of the tissue on both images and cubic interpolation (Figure 3, step 3). The registration error, which was minimized using the least squares nonlinear solver in MATLAB, was defined as the sum of the difference between the shape of the tissue on the white light image and the registered H&E image.

Affine + deformable registration

For the affine + deformable registration, an additional step was applied after the affine registration (Figure 3, step 4). First, control points were manually selected in both the H&E image (moving image) and the overview white light image (fixed image). Thereby, each control point in the H&E image had a corresponding point in the white light image. Two persons manually selected control points independently to make two registrations. In addition, the control points of both observers were combined to form the basis of one mean registration.

Figure 4 shows the control points selected by observer 1 in one representative example. First, control points were placed on distinctive features in both the H&E image and the white light image (Figure 4c and 4d). After assigning points to these features, additional points were placed in large areas of the images without any control points. These additional control points included four points in the corners of the images and a number of points surrounding the tissue. Without the additional control points, the algorithm is unable to accurately perform the local weighted mean transformation. For the deformable registration, a non-rigid local weighted mean transformation with 12 and 24 neighboring control points for the individual and mean registration respectively was applied using the built-in MATLAB function '*fitgeotrans.m*'.

Registration of optical measurements to white light images

The optical measurements were registered to the white light image to bring both histopathological and optical data into the same coordinate system. In the case of the hyperspectral images, a simple affine registration based on the shape of the tissue in both the hyperspectral image and white light image sufficed.

For the fiber-optic probe measurements, a more extensive registration was required. To retrieve the probe measurement locations in the overview white light image, we used I) a schematic image of the grid, II) a white light image of the tissue with the grid on top, and III) the overview white light image (Figure 5). First, the schematic image of the grid was registered to the white light image including the grid using a projective registration (Figure 5a). This was done by aligning the measurement locations in the four corners of the grid in both images. Second, to account for small deformations to the tissue caused

Figure 4. Representative example of control points (green and blue dots) selected by observer 1 in (a) the H&E image and (b) the white light image. The yellow squares in (a & b) correspond to the magnified squares (c & d).

by the grid, the white light image including the grid, was registered to the overview white light image using a deformable registration (similar to the registration method presented in Section 2.6.2). The calculated transformation matrices from both registrations were applied to the schematic image of the grid. With the final result (Figure 5b, IV), each measured probe location can be defined in relation to the overview white light image. All registrations were made using MATLAB 2018a® (MathWorks, Natick, Massachusetts, United States).

Data analysis

To quantify the registration performance after affine registration and affine + deformable registration, two measures were used. First, the shape of the tissue in the H&E image with respect to the shape of the tissue in the white light image was analyzed using the Dice similarity coefficient (DICE).[17] DICE is a measure for comparing the similarity of two images by measuring the overlap of the shape of the contour. Second, the tissue composition based on the H&E image was compared with the tissue composition in the μ CT data by using the point biserial correlation coefficient (r_{pb}).[18] This is a special case of the Pearsons Correlation Coefficient (PCC) suitable for calculating the correlation between a continuously measured

Figure 5. Measurement locations of the fiber-optic probe measurements. (a) A schematic image of the under grid (I) was registered to a white light image of the tissue with the grid on top (II) using a projective registration. (b) Subsequently the white light image with the grid (II) was registered to the overview white light image (III) with a non-rigid (deformable) registration. The result after both registrations is the overview white light image with a projection of the measurement locations (IV).

variable (μCT) and a dichotomous variable (H&E section). An absolute r_{pb} of 1 indicates a perfect linear correlation, whereas an r_{pb} of 0 indicates no correlation. In addition, we compared the composition of the tissue in the H&E image with the composition estimated with the optical measurements using the PCC. To compare the composition of the tissue in all modalities, we focused on discriminating fat from non-fat tissue since these tissue types can be well discriminated in all modalities.

Fat percentage in μCT data

Even though with μCT data no actual fat percentage could be obtained, a lower μCT gray value is related to a higher fat percentage.[19] Since the 3D μCT data will be correlated to 2D data (the 2D H&E image and the 2D estimated fat percentage of the optical measurements), the 3D data was reduced to a 2D image. For this purpose, first, the height of the tissue in the 3D μCT volume was determined to obtain the superficial cross-section. To prevent imaging

Figure 6. Determination of the average fat percentage within 1 mm of the tissues surface. (a) shows the side view of a μ CT image with two red lines that indicate the superficial cross section, located 0.15 mm underneath the tissue height, and the deeper cross section, located 1 mm underneath the superficial cross section. The side view corresponds to the dotted line in the 2D image (b) that represents the average gray value of the tissue within the two cross sections.

artifacts at the transition of non-tissue to tissue, the superficial cross section was defined at 0.15 mm underneath the determined tissue height as shown in Figure 6a. Subsequently, a second deeper cross section was defined at 1 mm underneath the first cross-section. Third, over this 1 mm distance, an average was calculated per pixel. Thereby, a 2D image was obtained that represents the average μ CT gray value over that 1 mm tissue thickness. An example of such a 2D image is shown in Figure 6b.

Fat percentage in H&E image

With the labels of the H&E image, the tissue in the H&E image was binary divided into fat and non-fat tissue in MATLAB. The latter group contained tumor, connective tissue, and healthy glandular ducts. The entire binary image was compared to the 2D image of the μ CT data. In addition, for comparing the fat percentage in the H&E image with the estimated fat percentage from the optical measurements, the average fat percentage of the measurement locations of the probe was calculated. This percentage was the percentage of fat pixels within a circle (diameter = 2 mm) around the centroid of the measurement location.

Fat percentage in optical measurements

Optical measurements were spectrally fitted with an analytical model that was derived from diffusion theory using a Levenberg-Marquardt nonlinear inversion algorithm in MATLAB. For the fiber-optic probe, each measurement location was fitted and for the hyperspectral measurements, each pixel in the image that contained tissue was fitted. For the optical probe, this analytical model was described elsewhere.[15, 16] For the hyperspectral camera, the analytical model used is given by [20]

$$R_d = \frac{\alpha'}{1 + 2k(1 - \alpha') + \left(1 + \frac{2k}{3}\right)\sqrt{3(1 - \alpha')}} \quad (1)$$

where $\alpha' = \frac{\mu_s'}{\mu_s' + \mu_a}$ and, $k = \frac{1 + r_d}{1 - r_d}$. r_d is the internal reflection coefficient for diffuse light and depends on the refractive index of the sample, which was calculated as $n_{\text{sample}} = n_{\text{water}} + 0.14 \cdot [\text{lipid}]$. μ_s' and μ_a are respectively the scattering and absorption coefficients and given by

$$\mu_s' = \alpha \left(\frac{\lambda}{1197 \text{ nm}} \right)^{-b}, \text{ and} \quad (2)$$

$$\mu_a = v_{WL} \left(\left(1 - \left(\frac{[\text{lipid}]}{[\text{lipid}] + [\text{water}]} \right) \right) \cdot \mu_{a, \text{water}}(\lambda) + \left(\frac{[\text{lipid}]}{[\text{lipid}] + [\text{water}]} \right) \cdot \mu_{a, \text{lipid}}(\lambda) \right), \quad (3)$$

where α is the reduced scattering at 1197 nm and b the scatter power. v_W is the fraction of water and lipid in the tissue and assumed to be 100% in the near-infrared wavelength region because water and lipid are the dominant absorbers in this wavelength region. $\mu_{a, \text{water}}$ and $\mu_{a, \text{lipid}}$ are respectively the absorption coefficients of water and lipid, and $[\text{lipid}]$ and $[\text{water}]$ correspond to the concentration of lipid and water.

RESULTS

Acquired data

Optical measurements and μCT data were acquired of 13 specimens. One specimen was excluded from further analysis because the slice was not thick enough to obtain a reliable 2D μCT image. Table 1 lists the characteristics of the tissue slices.

Registration of H&E section

Shape of the tissue: DICE

Figure 7 shows an example of an H&E image after affine registration and after affine + deformable registration (mean registration of both observers) and the effect of the transformations. The affine registration only scaled and rotated the H&E section, and hardly

Table 1. Characteristics of included specimens.

Specimen	Size (mm)	Slice thickness (mm)
	(height * width)	(depth)
1	62 * 38	3.5
2	43 * 32	2.5
3	60 * 43	3.0
4	64 * 46	6.6
5	50 * 33	2.7
6	35 * 28	2.8
7	48 * 25	3.4
8	45 * 46	3.8
9	45 * 17	3.2
10	46 * 35	4.5
11	47 * 43	5.5
12	60 * 35	3.2

any deformations are observed between the striped H&E section before (Figure 7c) and after affine registration (Figure 7d). The DICE coefficient was 0.93 for this registration. After affine + deformable mean registration (Figure 7e) however, a large transformation can be observed which increased the DICE coefficient from 0.93 (Figure 7g) to 0.97 (Figure 7h).

Composition of the tissue: r_{pb}

In addition to the DICE coefficient, we used the r_{pb} coefficient to compare the composition of the tissue in the H&E image with the composition of the tissue in the μ CT data. Inspecting the structures in the tissue that are easy to distinguish (orange arrows in Figure 7i, Figure 7j, and Figure 7k) reveals that, the composition of the tissue in the H&E section did not match the composition of the tissue in the μ CT data after only affine registration. This was corrected for after registering the H&E section with both the affine + deformable algorithm (Figure 7k). The r_{pb} coefficient increased from 0.24 to 0.53.

As can be seen in Table 2, averaged over all specimens, the r_{pb} increased substantially from the affine registration to the affine + deformable registration. Furthermore, when comparing the affine registration with the affine + deformable registration, on average over all specimens, 23% of the pixels had a different label (fat versus non-fat). Of these pixels, before the deformable registration step, 49% were labeled as non-fat, and 51% were labeled as fat.

Practical consequence of the proposed registration method: required time and operator dependence

The required time for registration depended on 1) the size of the tissue, and 2) the number of distinctive features in both images when selecting the control points. On average, the time required for removing the background in both the H&E image and the white light image was

Figure 7. Example of one of the tissue slices after affine registration (a) and affine + deformable registration (b). The second column shows the shape of the original H&E section with a striped pattern (c), and its transformation after both registrations (d & e). The third column shows the white light image (f), which was used to calculate the DICE coefficient (g & h). The fourth column shows the average μ CT intensity between the cross sections (i) and the r_{pb} coefficient with the H&E sections after both registrations (j & k). The arrow indicates a structure that is easy to distinguish on both the μ CT and the H&E image that respectively did and did not match after affine and affine + deformable registration. The values between brackets in (g, h, j & k) correspond to the DICE and r_{pb} coefficient respectively.

approximately 10 minutes. Rescaling the H&E image to the white light images and applying the affine registration was fast and took less than one minute. The most time-consuming part of the registration was the selection of control points in both images. This took around 15 minutes per observer. The registration of the optical measurement to the white light images required less than one minute for the hyperspectral images and approximately 15 minutes for the fiber-optic probe measurements.

Table 2. shows the DICE coefficient after both registrations, given as the average of all samples and the standard deviation (std). For all samples, this DICE coefficient was high after both registrations. On average, the highest DICE coefficient was obtained using the affine + deformable registration.

	Shape:		Composition:	
	DICE		r_{pb}	
	Mean	Std	Mean	Std
Affine registration	0,91	0,03	0,36	0,20
Affine + deformable registration				
Mean (#cp = 233 ± 35)	0,94	0,02	0,47	0,18
Observer 1 (#cp = 157 ± 41)	0,94	0,02	0,46	0,18
Observer 2 (#cp = 80 ± 27)	0,93	0,03	0,47	0,17

In this study, two observers, both experienced in registering H&E sections to white light images, selected control point pairs to perform the deformable registration. To assess the operator dependence of the affine + deformable registration method, the registrations of the two observers separately were compared. As shown in Table 2, the number of control point pairs varied between the registration of both observers and the mean registration (in which the control points of the two observers were combined). However, the differences between the mean registration, the registration of observer 1, and the registration of observer 2 are minimal for both the DICE and r_{pb} .

Correlation with optical measurements

The fat percentage derived from the H&E sections was compared to the fat percentage derived from the optical measurements to assess the effect of accounting for tissue deformations in the correlation between histopathology and optical measurements. A clear distinction was made between ‘unaltered’ and ‘altered’ points. In the altered points, the averaged fat percentage derived from the H&E section varied with more than 20% before and after deformable registration. The unaltered points are the remaining points. In case of a perfect correlation between optical measurements and histopathological information, the fat percentages in both modalities would be equal and therefore the PCC will equal 1.

In the representative tissue slice shown Figure 8, 7 out of 17 probe measurements are ‘altered’ points (cyan circles). Of these seven locations, the fitted fat percentage of the optical measurements was plotted against the fat percentage according to the histopathology after affine registration (Figure 8d) and after affine + deformable registration (Figure 8e). The fit results of both optical modalities have small confidence intervals and are similar. Figure 8d & 8e show that the linear correlation between the optical and the histopathological fat percentage improved substantially after adding the deformable registration for both the probe and the camera. The PCC improved from 0.08 to 0.93 and from 0.01 to 0.93, for the probe and the camera respectively.

Figure 8. Comparison of registered H&E sections with optical results in one representative specimen. (a) displays the H&E section after affine registration and (b) depicts the H&E section after affine + deformable registration. In (c) the fitted fat percentage derived from the HSI and the probe measurements is shown. In this figure, the probe measurement locations are indicated with circles and the color of circles represents the fat percentage according to the fit. In (a) and (b) the black circles indicate that the ratio of fat and non-fat tissue in both H&E sections differed less than 20%. If the ratio between fat and non-fat tissue differed more than 20% between the two registration methods the measurement location was marked with a cyan circle. For the cyan colored locations, the fat percentage as obtained with the optical measurements (as demonstrated in (c)) is plotted against the fat percentage obtained after affine registration (d) and affine + deformable registration (e). The error bars represent the standard deviation of the fitted fat percentages. The dotted lines represent the linear fitted line through the measurements.

Of all 234 locations that were measured with the fiber-optic probe, we compared the fat percentage of the optical measurements with the fat percentage according to the histopathology after affine registration and after affine + deformable registration (Table 3). We did this for the fitted fat percentages derived from both the probe and the hyperspectral measurements. Even though these fat percentages were not equal, they were similar and highly correlated with a PCC of 0.87.

Table 3. Comparison of the fat percentage according to the optical measurements and the histopathology after both registrations using the Pearson Correlation Coefficient (PCC). The altered points represent points in which the averaged fat percentage derived from the H&E section varied with more than 20% before and after deformable registration.

		Probe	HSI
		<i>PCC</i>	<i>PCC</i>
Unaltered points	Affine registration	0,90	0,86
	Affine + deformable registration	0,91	0,89
Altered points	Affine registration	0,36	0,36
	Affine + deformable registration	0,66	0,61

For the 186 ‘unaltered’ points, the PCC is not different before and after adding the deformable registration. For these points, using only affine registration, the PCC was 0.90 and 0.86, for the probe and HSI data respectively. Adding the deformable registration step produced a PCC of 0.91 for the probe measurements and 0.89 for HSI data. Thus, as the deformable registration resulted only in slightly higher PCC’s, correcting for tissue deformations was considered not necessary for these measurement locations. For the 48 ‘altered’ points, however, adding the deformable registration strongly enhanced the correlation between the optical measurements and histopathological information. The PCC increased from 0.36 to 0.61 for the HSI measurements, and from 0.36 to 0.66 for the probe measurements. These findings confirm the importance of accounting for tissue deformations when correlating optical measurements with histopathology.

DISCUSSION

Co-registration between optical measurements and histopathology is extremely important for validation of a novel optical diagnostic technology. Due to deformation of the tissue during histopathology processing, accurate correlation between optical measurements and the gold standard requires registration of H&E images to the optical measurements. Especially in tissue with an inhomogeneous character, like breast tissue, this correlation is crucial for validating the optical measurements with the current gold standard. In this manuscript, we highlight the importance of accounting for tissue deformations when correlating optical measurement with histopathology. As such, we present a methodology for making a registration between optical measurements and the gold standard using an affine + deformable registration method. In this process, the H&E images were registered to a white light image that was taken simultaneously to the optical measurements. This registration approach was compared to a simple registration method using only an affine registration that does not account of tissue deformations. The performance of both registration methods was quantified using a different independent imaging modality, μ CT.

The registrations were assessed by first, comparing the shape of the tissue by calculating the DICE for each registration, and second, by comparing the composition of the tissue by calculating the r_{pb} . Although both registration methods have high DICE coefficients (> 0.90), the affine + deformable registration provides the highest registration accuracy when evaluating both the shape (DICE) and the composition of the tissue (r_{pb}).

H&E sections after affine registration and after affine + deformable registration were compared to μ CT data to see which registration method results in the smallest registration errors between histopathology and optical measurements. We showed that accounting for tissue deformations is important for reducing the registration error between optical measurements and histopathology. It should be noted, however, that in all cases the H&E section remains a 2D image of a few cell layers thick whereas the optical measurements represent the optical properties of a volume of several mm^3 . Therefore, differences in depth will not be expressed in the H&E section, although such differences will affect the optical measurements. In addition, the thickness of the tissue within a tissue slice varies, as can be seen in Figure 6. In practice, the pathology technician will keep slicing sections from the paraffin-embedded tissue until a section is obtained that encloses the complete shape of the tissue slice. Therefore, the H&E section does not necessarily represent exactly the surface of the tissue as it was measured with the optical techniques. To account for this, we compared the registered H&E sections with μ CT data over a tissue thickness of 1 mm underneath the tissue's surface. The fact that the H&E sections are only a cell layer thick, and the height of the tissue differs over the surface, are explanations why a registration accuracy of 100% can never be obtained.

A limitation of the affine + deformable registration method as currently described is the manual input that was required, which makes it time-consuming. In addition, the method can be operator-dependent as the observers have to manually place points in both the H&E image and the overview white light image based on distinctive features that they recognize in both images.

In this study, registrations were made by two observers, with the same experience in registering H&E images to white light images. We showed that, even though the number of selected control point pairs varied between both observers, the registration accuracy was similar (Table 2). Therefore, registering the white light image with the H&E section using only the input of one of the two observers would also have resulted in an accurate registration. However, care should be taken when observers are less experienced. Also, we did not investigate how the number of neighboring control points that were used in the local weighted mean transformation impacted the registration. Optimization might be possible by using more or less neighboring points.

To reduce processing time, Naranjo *et al.* presented several registration methods to register infrared images with H&E stained sections, with the ultimate goal to automatically segment regions in the infrared data.[8] However, in their study, H&E stained sections

needed to be registered to sections of the tissue samples that are only a few cell layers thick, whereas in our study the tissue samples are ~3 mm thick. Furthermore, the total area of tissue that was registered to an H&E section in Naranjo's research was much smaller than the surface that was registered in this study. Due to these differences, their automatic key point selection method, which was based on finding the centroids of cells, was not applicable to our data. In their study, the projective transformation is reported to provide the best registration. This generalization of an affine transformation is a much less complex registration method compared to the affine + deformable method as proposed in this research. The projective transformation was sufficient in their study probably because the tissue slices were much smaller and therefore less deformation was present.

As shown in Table 2 and Figure 7, the r_{pb} increased substantially from affine registration to affine + deformable registration. Also, when comparing the labels of the H&E after affine registration and the H&E section after affine + deformable registration, on average 23% of the pixels obtained a different label. This result is confirmed by comparing the PCC of only affine registration or affine + deformable registration of the altered points (Figure 8 & Table 3). Thus, using a different registration method leads to different labeling of a substantial part of the optical measurements. This potentially has a strong effect on the maximum classification performance that can be obtained with the optical techniques.

Although, the fit results of the two optical modalities used are not exactly the same, the linear correlation between the fat percentage derived from the optical measurements and the histopathological fat percentage improved for both optical modalities after adding the deformable registration method to the affine registration. Differences between the fit results of the optical techniques can be explained by a difference in the measurement configuration: the distance between the illuminating and collecting fiber determines the measured volume with the fiber-optic probe, whereas for the hyperspectral camera the volume is mainly depended upon the tissue properties. Therefore, both configurations might measure a different volume, resulting in a difference between the obtained fat percentages. Nevertheless, the linear correlation between the fat percentages derived from the probe and the camera measurements is high with a PCC of 0.87.

In this study, we used μ CT as an independent 3D modality to quantify both registration methods. Abe *et al.* has reported on using MRI for imaging of pathology slices and showed good image quality.[21] In our study, we did attempt to quantify the registration methods with MRI using the mDIXON-Quant sequence. This sequence was chosen as it can make excellent discrimination between fat and non-fat tissue.[22-24] However, the resolution, although acquired with a relatively high-resolution voxel size of 0.6 x 0.6 x 0.9 mm, proved to be too low for quantifying the registration methods. In the study of Abe *et al.*, an animal MRI with a higher magnetic field (9.4T) was available than the clinical human wide-bore 3T MRI available in this study. For this reason, in this study, MRI data could not be used for quantifying the co-registrations and was thus not included in the analysis.

A disadvantage of using only μ CT data was that we could only discriminate fat from non-fat tissue. In which the latter included both connective tissue and tumor tissue. One of the reasons was the low contrast between tumor and healthy fibro-glandular breast tissue in the μ CT scans. This has been reported by previous research as well.[25] Other studies indicate that this differentiation should be possible.[19] However, in that specific study, biopsy specimens were imaged with a much higher resolution than the resolution that could be obtained in this study. Nevertheless, the use of μ CT did allow us to quantify the registration accuracy of H&E images based on the presence or absence of fat in the tissue.

It should be noted that this research was conducted mainly to improve the validation of optical measurements based on the current gold standard. As previously explained, the H&E section remains a 2D image of a few cell layer thick and will, therefore, no matter which registration method is used, never completely represent the volume measured with the optical modalities. In addition, histopathology can also suffer from other potential flaws arising from factors such as intra-observer variation[26-28], inter-observer variation[29, 30], under sampling[31, 32], and specimen handling and reporting[33-36]. Nevertheless, there is no way in validating the technology without considering the current gold standard. Therefore, once established as a robust diagnostic technology, the true potential of optical measurements in clinical practice should be assessed with evaluation of clinical end-points like, for breast cancer surgery, a decrease in positive resection margins or decrease in excised specimen volumes.

In conclusion, for an accurate correlation between histopathological information and optical measurements, tissue deformations should be taken into account in the registration. The proposed registration method in this study, which does account for these tissue deformations, shows a better correlation with the optical measurements. Although it is labor intensive, adding the deformable registration step can aid in validating the optical measurements with the current gold standard.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the financial support of this research by the Dutch Cancer Society (grant KWF 10747). We also thank Bruker MicroCT for providing the microCT scanner. The authors thank N.N.Y. Janssen for help with acquiring μ CT data, O. Ivashchenko for help in developing the registration algorithm and the NKI-AVL core Facility Molecular Pathology & Biobanking (CFMPB) for supplying NKI-AVL biobank material. The authors also thank all surgeons and nurses from the Department of Surgery and all pathologists and pathologist assistants from the Department of Pathology for their assistance in collecting specimens.

REFERENCES

- [1] W. A. Wells, P. E. Barker, C. MacAulay, M. Novelli, R. M. Levenson, and J. M. Crawford, "Validation of novel optical imaging technologies: the pathologists' view," *Journal of Biomedical Optics*, vol. 12, no. 5, pp. 051801, 2007, [[10.1117/1.2795569](#)].
- [2] W. K. Jerjes, T. Upile, B. J. Wong, C. S. Betz, H. J. Sterenborg, M. J. Witjes, K. Berg, R. van Veen, M. A. Biel, A. K. El-Naggar, C. A. Mosse, M. Olivo, R. Richards-Kortum, D. J. Robinson, J. Rosen, A. G. Yodh, C. Kendall, J. F. Ilgner, A. Amelink, V. Bagnato, H. Barr, L. Bolotine, I. Bigio, Z. Chen, L.-P. Choo-Smith, A. K. D'Cruz, A. Gillenwater, A. Leunig, A. J. MacRobert, G. McKenzie, A. Sandison, K. C. Soo, H. Stepp, N. Stone, K. Svanberg, I. B. Tan, B. C. Wilson, H. Wolfson, and C. Hopper, "The future of medical diagnostics: review paper," *Head & Neck Oncology*, vol. 3, no. 38, 2011, [[10.1186/1758-3284-3-38](#)].
- [3] M. N. Gurcan, L. Boucheron, A. Can, A. Madabhushi, N. Rajpoot, and B. Yener, "Histopathological Image Analysis: A Review," *IEEE Reviews in Biomedical Engineering*, vol. 2, pp. 147-171, 2009, [[10.1109/RBME.2009.2034865](#)].
- [4] V. Krishnaswamy, A. M. Laughney, W. A. Wells, K. D. Paulsen, and B. Pogue, "Scanning in situ Spectroscopy platform for imaging surgical breast tissue specimens," *Optics Express*, vol. 21, no. 2, pp. 2185-2194, 2013, [[10.1364/OE.21.002185](#)].
- [5] M. J. Piling, A. Henderson, J. H. Shanks, M. D. Brown, N. W. Clarke, and P. Gardner, "Infrared spectral histopathology using haematoxylin and eosin (H&E) stained glass slides: a major step forward towards clinical translation," *Analyst*, vol. 142, pp. 1258-1268, 2017, [[10.1039/c6an02224c](#)].
- [6] B. C. Wilson, M. Jermyn, and F. Leblond, "Challenges and opportunities in clinical translation of biomedical optical spectroscopy and imaging," *Journal of Biomedical Optics*, vol. 23, no. 3, pp. 030901, 2018, [[10.1117/1.JBO.23.3.030901](#)].
- [7] G. Lu, D. Wang, X. Qin, L. Halig, S. Muller, H. Zhang, A. Chen, B. W. Pogue, Z. G. Chen, and B. Fei, "Framework for hyperspectral image processing and quantification for cancer detection during animal tumor surgery," *Journal of Biomedical Optics*, vol. 20, no. 12, pp. 126012, 2015, [[10.1117/1.JBO.20.12.126012](#)].
- [8] V. Naranjo, E. Villanueva, G. R. Lloyd, N. Stone, F. López-Mir, and M. Alcañiz, "Stained and Infrared Image Registration as First Step for Cancer Detection," in *IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI)*, Valencia, Spain, 2014, pp. 420-423, [[10.1109/BHI.2014.6864392](#)].
- [9] M. Halicek, J. V. Little, X. Wang, Z. G. Chen, M. Patel, C. C. Griffith, M. W. El-Deiry, N. F. Saba, A. Y. Chen, and B. Fei, "Deformable Registration of Histological Cancer Margins to Gross Hyperspectral Images using Demons."
- [10] G. Lu, L. Halig, D. Wang, Z. G. Chen, and B. Fei, "Hyperspectral Imaging for Cancer Surgical Margin Delineation: Registration of Hyperspectral and Histological Images," [[10.1117/12.2043805](#)].
- [11] J. E. Phipps, D. Gorpas, J. Unger, M. Darrow, R. J. Bold, and L. Marcu, "Automated detection of breast cancer in resected specimens with fluorescence lifetime imaging," *Physics in Medicine and Biology*, vol. 63, pp. 015003, 2018, [[10.1088/1361-6560/aa983a](#)].
- [12] A. M. Laughney, V. Krishnaswamy, P. Beatriz Garcia-Allende, O. M. Conde, W. A. Wells, K. D. Paulsen, and B. W. Pogue, "Automated classification of breast pathology using local measures of broadband reflectance," *Journal of Biomedical Optics*, vol. 15, no. 6, pp. 066019, 2010, [[10.1117/1.3516594](#)].
- [13] A. M. Laughney, V. Krishnaswamy, E. J. Rizzo, M. C. Schwab, R. J. Barth, B. W. Pogue, K. D. Paulsen, and W. A. Wells, "Scatter Spectroscopic Imaging Distinguishes between Breast Pathologies in Tissues Relevant to Surgical Margin Assessment," *Clinical Cancer Research*, vol. 18, pp. 6315-6325, 2012, [[10.1158/1078-0432.CCR-12-0136](#)].
- [14] E. Kho, L. L. De Boer, K. K. Van de Vijver, F. Van Duijnhoven, M. T. F. D. Vranken Peeters, H. J. C. M. Sterenborg, and T. J. M. Ruers, "Hyperspectral imaging for resection margin assessment

- during surgery," *Clinical Cancer Research*, vol. accepted, 2019, [[10.1158/1078-0432.CCR-18-2089](#)].
- [15] R. Nachabé, B. H. W. Hendriks, A. E. Desjardins, M. van der Voort, M. B. van der Mark, and H. J. C. M. Sterenberg, "Estimation of lipid and water concentrations in scattering media with diffuse optical spectroscopy from 900 to 1600 nm," *J Biomed Opt*, vol. 15, no. 3, pp. 037015, 2010, [[10.1117/1.3454392](#)].
 - [16] R. Nachabé, B. H. W. Hendriks, M. Van der Voort, A. E. Desjardins, and H. J. C. M. Sterenberg, "Estimation of biological chromophores using diffuse optical spectroscopy; benefit of extending the UV-VIS wavelength range to include 1000 to 1600nm," *Optics Express*, vol. 18, no. 24, pp. 1432-1442, 2010, [[10.1364/BOE.1.001432](#)].
 - [17] L. R. Dice, "Measures of the Amount of Ecologic Association Between Species," *Ecology*, vol. 26, no. 3, pp. 297-302, 1945, [[10.2307/1932409](#)].
 - [18] G. V. Glass, and K. D. Hopkins, *Statistical Methods in Education and Psychology*, 3rd ed.: Allyn & Bacon, 1995.
 - [19] H. Gufler, F. E. Franke, S. Wagner, and W. S. Rau, "Fine Structure of Breast Tissue on Micro Computed Tomography: A Feasibility Study," *Academic Radiology*, vol. 18, pp. 230-234, 2011, [[10.1016/j.acra.2010.10.006](#)].
 - [20] S. T. Flock, M. S. Patterson, B. C. Wilson, and D. R. Wyman, "Monte Carlo modeling of light propagation in highly scattering tissues. I. Model predictions and comparison with diffusion theory," *IEEE Transaction son Biomedical Engineering*, vol. 36, no. 12, pp. 1162-1168, 1989, [[10.1109/TBME.1989.1173624](#)].
 - [21] H. Abe, A. Shimauchi, X. Fan, J. N. River, H. Sattar, J. Mueller, G. S. Karczmar, and G. M. Newstead, "Comparing post-operative human breast specimen radiograph and MRI in lesion margin and volume assessment," *Journal of Applied Clinical Medical Physics*, vol. 13, no. 6, pp. 267-276, 2012, [[10.1120/jacmp.v13i6.3802](#)].
 - [22] A. Mashhood, R. Railkar, T. Yokoo, Y. Levin, L. Clark, S. Fox-Bosetti, M. S. Middleton, J. Riek, E. Kauh, B. J. Dardzinski, D. Williams, C. Sirlin, and N. J. Shire, "Reproducibility of Hepatic Fat Fraction Measurement by Magnetic Resonance Imaging," *Journal of Magnetic Resonance Imaging*, vol. 37, pp. 1359-1370, 2013, [[10.1002/jmri.23928](#)].
 - [23] Y. Kise, T. Chikui, Y. Yamashita, K. Kobayashi, and K. Yoshiura, "Clinical usefulness of the mDIXON Quant method for estimation of the alivary gland fat fraction: comparison with MR spectroscopy," *British Journal of Radiology*, vol. 90, pp. 20160704, 2017, [[10.1259/bjr.20160704](#)].
 - [24] G. M. Kukuk, K. Hittatiya, A. M. Sprinkart, H. Eggers, J. Gieseke, W. Block, P. Moeller, W. A. Willinek, U. Spengler, J. Trebicka, H.-P. Fischer, H. H. Schild, and F. Träber, "Comparison between modified Dixon MRI techniques, MR spectroscopic relaxometry, and different histologic quantification methods in the assessment of hepatic steatosis," *European Radiology*, vol. 25, pp. 2869-2879, 2015, [[10.1007/s00330-015-3703-](#)].
 - [25] S.-Q. Qiu, M. D. Dorrius, S. J. De Jongh, L. Jansen, J. De Vries, C. P. Schröder, G.-J. Zhang, E. G. E. De vries, B. Van der Vegt, and G. M. Van Dam, "Micro-computed tomography (micro-CT) for intraoperative surgical margin assessment of breast cancer: A feasibility study in breast conserving surgery," *European Journal of Surgical Oncology (EJSO)*, vol. 44, no. 11, pp. 1708-1713, 2018, [[10.1016/j.ejso.2018.06.022](#)].
 - [26] S. L. Jackson, P. D. Frederick , M. S. Pepe, H. D. Nelson, D. L. Weaver, K. H. Allison, P. A. Carney, B. M. Geller, A. N. A. Tosteson, T. Onega, and J. G. Elmore, "Diagnostic Reproducibility: What Happens When the Same Pathologist Interprets the Same Breast Biopsy Specimen at Two Points in Time?," *Annals of Surgical Oncology*, vol. 24, pp. 1234-1241, 2017, [[10.1245/s10434-016-5695-0](#)].
 - [27] P. Boiesen, P.-O. Bendahl, L. Anagnostaki, H. Domanski, E. Holm, I. Idvall, S. Johansson, O. Ljungberg, A. Ringberg, G. Östberg, and M. Fernö, "Histologic Grading in Breast Cancer:

- Reproducibility Between Seven Pathologic Departments," *Acta Oncologica*, vol. 39, no. 1, pp. 41-45, 2000, [[10.1080/028418600430950](https://doi.org/10.1080/028418600430950)].
- [28] S. J. Schnitt, "Evaluation of Microscopic Margins in Patients with Invasive Breast Cancer: Technical and Interpretive Considerations," *The Breast Journal*, vol. 4, no. 3, pp. 204-208, 1998, [[10.1046/j.1524-4741.1998.430204.x](https://doi.org/10.1046/j.1524-4741.1998.430204.x)].
- [29] K. H. Allison, L. M. Reisch, P. A. Carney, D. L. Weaver, S. J. Schnitt, F. P. O'Malley, B. M. Geller, and J. G. Elmore, "Understanding diagnostic variability in breast pathology: lessons learned from an expert consensus review panel," *Histopathology*, vol. 65, no. 2, pp. 240-251, 2014, [[10.1111/his.12387](https://doi.org/10.1111/his.12387)].
- [30] J. G. Elmore, G. M. Longton, P. A. Carney, B. M. Geller, T. Onega, A. N. A. Tosteson, H. D. Nelson, M. S. Pepe, K. H. Allison, S. J. Schnitt, F. P. O'Malley, and D. L. Weaver, "Diagnostic Concordance Among Pathologists Interpreting Breast Biopsy Specimens," *Journal of the American Medical Association*, vol. 313, no. 11, pp. 1122-1132, 2015, [[10.1001/jama.2015.1405](https://doi.org/10.1001/jama.2015.1405)].
- [31] G. M. Clarke, C. M. B. Holloway, J. T. Zubovits, S. Nofech-Mozes, K. Liu, M. Murray, D. Wang, and M. J. Yaffe, "Whole-mount pathology of breast lumpectomy specimens improves detection of tumour margins and focality," *Histopathology*, vol. 69, pp. 35-44, 2016, [[10.1111/his.12912](https://doi.org/10.1111/his.12912)].
- [32] N. Houssami, and M. Morrow, "Margins in Breast Conservation: A Clinician's Perspective and What the Literature Tells Us," *Journal of Surgical Oncology*, vol. 110, pp. 2-7, 2014, [[10.1002/jso.23594](https://doi.org/10.1002/jso.23594)].
- [33] S. Persing, T. A. James, J. Mace, A. Goodwin, and B. Geller, "Variability in the Quality of Pathology Reporting of Margin Status Following Breast Cancer Surgery," *Annals of Surgical Oncology*, vol. 18, no. 11, pp. 3061-3065, 2011, [[10.1245/s10434-011-1916-8](https://doi.org/10.1245/s10434-011-1916-8)].
- [34] R. Emmadi, and E. L. Wiley, "Evaluation of Resection Margins in Breast Conservation therapy: The Pathology Perspective-Past, Present and Future," *Journal of Surgical Oncology*, vol. 2012, pp. 180259, 2012, [[10.1155/2012/180259](https://doi.org/10.1155/2012/180259)].
- [35] S. J. Schnitt, and J. L. Connolly, "Processing and Evaluation of Breast Excision Specimens: A Clinically Oriented Approach," *Anatomic Pathology*, vol. 98, no. 1, pp. 125-137, 1992, [[10.1093/ajcp/98.1.125](https://doi.org/10.1093/ajcp/98.1.125)].
- [36] S. P. Povoski, F. E. Jimenez, W. P. Wang, and R. X. Xu, "Standardized and reproducible methodology for the comprehensive and systematic assessment of surgical resection margins during breast-conserving surgery for invasive breast cancer," *BMC Cancer*, vol. 254, no. 9, 2009, [[10.1186/1471-2407-9-254](https://doi.org/10.1186/1471-2407-9-254)].

The influence of neoadjuvant chemotherapy on DRS spectra of breast surgery specimens

L.L. de Boer | E. Kho | K. Józwiak | K.K. Van de Vijver
M.T.F.D. Vranken Peeters | F. van Duijnhoven | B.H.W. Hendriks
H.J.C.M. Sterenborg | T.J.M. Ruers

In preparation

Diffuse Reflectance Spectroscopy (DRS) can discriminate different tissue types based on optical characteristics. Since this technology has the ability to detect tumor tissue several groups have proposed to use DRS for margin assessment during breast-conserving surgery. Nowadays, an increasing number of patients is being treated by neoadjuvant chemotherapy. Limited research has been published on the influence of neoadjuvant chemotherapy on the optical characteristics of the tissue. Hence, it is unclear if DRS is feasible for margin assessment in this specific group of patients. Here we investigate the effect of neoadjuvant chemotherapy on optical measurements of breast tissue. To this end, DRS measurements were performed on *ex vivo* lumpectomy specimens of patients with or without prior neoadjuvant chemotherapy. Generalized estimating equation (GEE) models were generated, comparing measurements of patients with and without neoadjuvant chemotherapy in datasets of different tissue types. As input for the GEE-models either the intensity at a specific wavelength or a fit parameter, which was derived from the spectrum, were used. In the evaluation of the intensity, no influence of chemotherapy was found since none of the wavelengths were significantly different between the measurements with and without chemotherapy in any of the datasets. These results were confirmed by the analysis of the fit parameters which showed only a significant difference for the amount of collagen in one dataset. All other fit parameters did not show significance for any of the datasets. These findings indicate that assessment of the resection margin with DRS may also be feasible in the growing population of breast cancer patients that receive neoadjuvant chemotherapy.

Keywords: chemotherapy, diffuse reflectance spectroscopy, margin assessment

INTRODUCTION

Diffuse Reflectance Spectroscopy (DRS) is an optical technology that measures the composition and morphology of tissue based on absorption and scattering of light. A DRS measurement of tissue can be used to discriminate tissue types as the optical characteristics differ between tissue types. Previous research showed the potential of this technology to provide surgical guidance during breast-conserving surgery based on the ability of DRS to discriminate healthy breast tissue from tumor tissue.[1-5] The reported accuracies for discriminating healthy breast tissue from tumor tissue reach up to 92% *in vivo*. These results have generally been obtained with measurements of women undergoing immediate breast-conserving surgery, without any pre-operative treatment with chemotherapy. However, nowadays, there is an increasing tendency to treat breast cancer patients with neoadjuvant chemotherapy before surgery.[6] Reported benefits consist of improved cosmetic outcomes, since less extensive surgery is required, buying time for testing on germline mutations that may influence the surgical plan, the possibility of monitoring therapy response and eradicating already present, but undetected micro-metastasis.[7-11]

Chemotherapy is aimed at attacking the tumor tissue, and will thus induce morphologic changes such as decreased cellularity, increased fibrosis, increased tubular features, and decreased nuclear pleomorphism.[12-14] For example, the (previous) tumor bed may contain newly formed fibrosis with fibroblasts and inflammatory cells induced by chemotherapy. [12] Besides the impact on tumor tissue, chemotherapy therapy may also affect healthy breast tissue. Since these drugs suppress ovarian function, the excretion of sex hormones is changed with results in decreased breast density. [15] This change in density appears in the healthy tissue as a reduction of lobular acini, sclerosis and the attenuation of the lobular and ductal epithelium, which can be seen under the microscope. [13, 14, 16, 17]

Currently, it is unclear whether DRS measurements will be affected by the previously described morphological changes in the tissue, and thus if DRS can be used for margin detection during surgery in patients who have been treated with neoadjuvant chemotherapy. A preliminary study on the influence of chemotherapy by our group did not find any differences between patients with and without chemotherapy, however, the number of included patients was limited ($n = 9$) as well as the number of measurements per patient.[18]

Since our ultimate goal is to use DRS technology during surgery for the detection of tumor deposits at the resection edge, it is essential to consider if the changes induced by the chemotherapy will affect the DRS measurements. If so, resection margin assessment using DRS might then be less feasible for patients after neoadjuvant chemotherapy. Secondly, such an outcome has implications for the development of classification algorithms that are intended to be used for classifying the DRS measurements. If measurements of a specific tissue type differ between untreated patients and patients treated with neoadjuvant

chemotherapy, the treatment history should be included as input for the diagnostic algorithm.

In the present study, we evaluate whether chemotherapy induces changes in the breast tissue that may hamper the use of DRS for margin detection. For this purpose, we evaluate optical measurements of surgical breast cancer specimens in a large dataset of patients with and without chemotherapy ($n = 92$).

MATERIALS AND METHODS

DRS measurements

Specimens

All measurements were performed on *ex vivo* breast samples from either breast-conserving surgeries or mastectomy procedures. Fresh specimens were brought to the pathology department where, after inking of the margins and slicing the specimen in a bread-loafed manner, one slice was provided for the optical measurements. Before measuring, the tissue was placed on top of a piece of black rubber in a macrocassette.

Optical measurements

DRS measurements were obtained between 400 and 1600 nm. Details on the measurement set up can be found in [19, 20]. In short, the light of a halogen light source was transferred to the tissue by means of the illuminating fiber that is integrated into a fiber-optic probe. Two other integrated fibers were attached to two spectrometers, one covering the visual wavelength range (Andor Technology, DU420A-BRDD, 400-1000 nm) and one covering the near-infrared wavelength range (Andor Technology, DU492A-1.7, 900-1600 nm), which both collected the photons after interaction with the tissue. The set-up was controlled with Labview software that automatically performed a white calibration and combined the output of both spectrographs to form one continuous spectrum from 400 to 1600 nm. The fiber distance between the emitting and collecting fiber was 1 mm. Raw spectra were normalized with a calibration cap with Spectralon® (SRT-99-100, Labsphere, Inc., Northern Sutton, New Hampshire) at the bottom, to calculate the relative diffuse reflectance.

Grid

During the measurements, a custom-made grid was used to ensure a robust correlation between the locations of the optical measurements and the corresponding histopathology. The grid with 64 holes comprised of two parts, an 'under-grid' that was placed directly on top of the tissue and an 'upper-grid' that together with a small extension piece fixated the

Figure 1. Schematic overview of method for correlating the histopathology with the optical measurement locations. A registration is made between the H&E section (a) and the white light overview image (c). The transformation matrix obtained after this registration procedure is applied to the H&E section with annotations (b). This way the annotated image is corrected for deformations in the H&E section induced during the processing of the tissue at the pathology department. Using a white light image with the grid (d) the measurement locations are retrieved. The retrieved measurement locations (e) can be projected on the H&E section (as both are registered to the white light image). The end result is an image including the measurement locations, that is corrected for tissue deformation, and in which each pixel is annotated (f).

probe during measurement. Before the fiber-optic measurement, two white light images were acquired of 1) the tissue in the cassette with the grid and 2) without the grid. After the optical measurements, the grid was removed and tissue was processed according to standard protocol resulting in an H&E stained section of the entire tissue slice.

Correlating DRS measurement to histopathology

The entire H&E section (Figure 1a) was evaluated by a pathologist who annotated areas with invasive carcinoma and areas with DCIS as well as glandular ducts. The section and annotations of the pathologist were digitalized and, with thresholding, in Matlab, all other pixels in the H&E image were labeled as 'fat' or 'connective', based on the fact that fat areas appear white in the H&E section whereas connective tissue will stain pink. Glandular tissue was included in the 'connective' tissue class. The result was an image in which each pixel had a label corresponding to a tissue type (fat, connective, tumor, and DCIS) (Figure 1b). The histopathology was registered with the white light image (Figure 1c) to adjust for tissue deformation which was introduced during histopathological processing. Another registration was made between the white light image and a white light image including the grid (Figure 1d) to retrieve the actual measurement locations in the white light image (Figure 1e). Combining both registrations then allowed to project the measurement locations in the annotated H&E section which was corrected for tissue deformation. (Figure 1f) This registration method is described in more detail in another paper.[21](Chapter 5) With the final result for each measurement location, an estimate of the presence of each tissue type can be made.

Optical data

Although the measurements were acquired over the wavelengths between 400 and 1600 nm, the part of the spectrum between 400 and 850 nm was excluded in the analysis as pathology ink, present on the specimens, strongly influenced the reflection measurements of the wavelengths below 850 nm. For the first analysis, the spectral analysis, the intensities measured between 850 and 1600 nm are used. For the second analysis, the fit parameter analysis, the spectra were fitted with an analytical fit model which was described in previous publications. [19, 20] This model estimated the composition of the measurement volumes based on diffusion theory using the measured spectrum and the absorption spectra of substances known to be present in the tissue. Included in the model were: oxygenated and deoxygenated hemoglobin, fat, water, and collagen. The fit parameters, derived from the spectra with the analytical fit model, which were used in the analysis were: F/W-ratio, fraction Mie scattering, Collagen, α , and β . The latter two define the amplitude (α) and slope (β) in the function that describes the reduced scattering.[19]

At each measurement location, three DRS measurements were acquired. Both for the spectral data as well as the fit parameter data the mean of the measurements was calculated for each measurement location.

Composition of measurement locations

To prevent bias in tissue selection, first, the composition of the tissue in all H&E sections of the specimens with and without neoadjuvant chemotherapy was examined by

Figure 2. Datasets used in the analyses. Locations were selected from the specimens of patient with and without chemotherapy, which consisted of >95% fat tissue (Dataset 1), >95% connective tissue (Dataset 2), a mixture of fat and connective tissue (Dataset 3) or a mixture of tumor cells and connective tissue (Dataset 4). With each dataset two analyses were performed, either based on the spectra (850-1600 nm) or fit parameters, in which the measurements of patient with and without chemotherapy are compared.

calculating the mean percentage of all tissue types (fat, connective, tumor cells, and DCIS). Subsequently, the composition of the measurement locations at which DRS measurements were performed was compared to the composition of the entire H&E sections to ensure that the measurement locations were a good representation of the tissue composition of the specimens.

Statistics

The influence of chemotherapy on the optical characteristics of the tissue was evaluated by comparing the optical data of untreated patients with data of pre-treated patients in datasets with the same histopathology. Dataset 1 and dataset 2 consisted of locations that were respectively >95% fat tissue, or >95% connective tissue. Dataset 3 consisted of measurement locations that were a mixture of fat tissue and connective tissue. The last dataset, dataset 4 comprised of all locations that contained tumor cells in combination with some connective tissue. (Figure 2) To select the measurements for each dataset the estimated percentages of fat, connective, and tumor cells were used, as described in section 'Composition of measurement locations'.

In each of the four datasets, the difference between measurements with and without chemotherapy was assessed by performing two analyses. For the first analysis, the intensities measured between 850 and 1600 nm were used, for the second analysis, the fit parameters which were derived from this part of the DRS spectrum were used.

Subsequently, in both analyses, these datasets were evaluated one by one via Generalized Estimating Equation (GEE) models. This statistical test is specifically suitable for modeling correlated data, such as repeat measurements in one patient or repeat measurements of tissue types in multiple patients. In such a case, incorporating within-subject and between-subject variation in the statistical analysis is important.[22] A GEE-model evaluates the association between measurements and covariates while taking into account the relations between measurements that are from one patient.[23] In the GEE-models, an equicorrelated structure was used to describe the variance and covariance between the measurements.

In the first analysis, a GEE-model was generated for each wavelength whereas in the second analysis a GEE-model was generated for each fit parameter. Covariates included in the models were 'Neo-adjuvant chemotherapy' ('yes' / 'no'), as well as 'Menopausal Status' ('premenopausal or perimenopausal' / 'postmenopausal'). Menopausal status was incorporated in the models as this factor is known to be related to breast density.[24, 25] Omitting this patient characteristic in the analysis could lead to falsely attributing differences to another covariate, e.g. neoadjuvant chemotherapy.

Dataset 3 and 4 contained measurements that were a mixture of connective tissue in combination with either fat tissue or tumor cells. Since the ratio between the two tissue types can affect the spectrum, the percentage connective tissue estimated from the annotated H&E section was added as an extra covariate in the GEE-models of these datasets to take into account this effect.

The analyses were performed in SPSS software with help of the SPSS Python plugin module that provided the possibility to iterate over either the wavelengths or fit parameters, making a GEE-model for each one of them. The α -level was set to 0.05 for all the GEE-models.

RESULTS

In this study measurements from 92 patients were used of which 62 patients did not have any neoadjuvant chemotherapy and 30 patients were pre-treated with chemotherapy. The patient characteristics are listed in Table 1. The mean age of the group of pre-treated patients was slightly younger compared to the mean age of the group of patients without chemotherapy. Furthermore, the majority of patients that were treated with chemotherapy received a combination of different types of chemotherapeutics (anthracycline, alkylating agents, taxanes, and, antimetabolites).

Composition of the tissue section

Figure 3a and Figure 3b show the mean percentages and standard deviation of the tissue types present in the entire H&E section from patients without and with chemotherapy. The mean percentage of fat is similar in both datasets. However, the amount of connective tissue

Table 1. Patient characteristics of specimens that were used in the analysis.

	No Chemo (n = 61)	Chemo (n = 30)
Age (std)	58 (10,6)	51 (11,5)
Menopausal Status		
<i>Premenopausal</i>	22 (36.1%)	13 (43.3%)
<i>Perimenopausal</i>	5 (8.2%)	3 (10%)
<i>Postmenopausal</i>	34 (55.7%)	14 (46.7%)
Chemotherapy		
<i>AC</i>	-	6 (20%)
<i>AC + P</i>	-	7 (23.2%)
<i>AC + CP + P</i>	-	3 (10%)
<i>AC + P + C</i>	-	1 (3.3%)
<i>AC + CP</i>	-	1 (3.3%)
<i>AC + T</i>	-	1 (3.3%)
<i>AC + CP + P + C</i>	-	1 (3.3%)
<i>PTC</i>	-	5 (16.7%)
<i>FEC + PTC</i>	-	1 (3.3%)
<i>TAC</i>	-	1 (3.3%)
<i>P</i>	-	2 (6.6%)
<i>AC + P + CP + CAP</i>	-	1 (3.3%)

A = Adimycine, C = Cyclofosfamide, P = Paclitaxel, CP = Carboplatin, T = Docetaxel, F = 5-fluorouracil, E = epirubicine, CAP = Capcitabine

is slightly higher in the neoadjuvant chemotherapy group, whereas the amount of tumor and DCIS is lower in this group.

Figure 3c and Figure 3d depict the mean percentages and standard deviations of the measurement locations that were measured with the fiber-optic probe. Although the standard deviations are larger, the same trends as for the composition of the entire section are seen here, thus the probe measurements can be seen as a good representation of the tissue that is present in the entire section.

Table 2 lists the number of measurements with and without chemotherapy in each of the four datasets, as well as the number of patients the measurements originated from. The datasets including measurements of pure connective tissue (dataset 2) and measurements of a mixture of connective and tumor cells (dataset 4) are relatively small. Although the mean percentage of connective tissue is rather high (Figure 3), locations of >95% connective tissue are quite rare. More often connective tissue is present in combination with fat tissue and/or connective tissue.

Intensity measured between 850-1600 nm (analysis 1)

The results of the GEE-models for the pure datasets (pure fat and pure connective tissue) are shown in Figure 4. In Figure 4a and Figure 4b, none of the wavelengths reaches a p-value of 0.05 or lower. Thus, for pure fat or pure connective tissue, there are no wavelengths that

Figure 3. Composition of HE section and the measurement locations from specimens of patients with and without chemotherapy. The upper two graphs represent the mean percentage fat tissue, connective tissue, tumor cells and DCIS cells in the entire HE section from all patients without chemotherapy (a) and all patients with chemotherapy (b). The bottom two graphs represent the mean percentage of each tissue type in the measurement locations from which DRS spectra were obtained, from patients without chemotherapy (c) and with chemotherapy (d).

Table 2. Datasets used in the two analyses.

	No neo-adjvant chemotherapy	With neo-adjvant chemotherapy
Dataset 1: Fat	222 (49)	144 (24)
Dataset 2: Connective	19 (13)	32 (9)
Dataset 3: Fat/Connective	316 (56)	236 (30)
Dataset 4: Tumor cells/Connective	43 (21)	13 (6)
Total # measurements	600	425

Number of measurements in each of the four datasets which were used for making the GEE-models. The number between brackets indicates the number of patients that the measurements originate from.

Figure 4. Result of GEE-models p-values for pure fat (dataset 1) and pure connective tissue (dataset 2). P-values calculated for the datasets of fat (a) and connective (b). None of the wavelengths are significant for either two datasets comparing the intensities measured in patients with and without chemotherapy, or comparing postmenopausal patients with pre-/perimenopausal patients.

have a significantly different intensity due to chemotherapy. Furthermore, between pre-/perimenopausal and postmenopausal women also no significant difference is seen.

In Figure 5, the results of the GEE-models for the mixture datasets are displayed. Figure 5a shows the results for the GEE-models of the measurements of connective tissue in combination with fat tissue. There was no significant difference in intensity measured at any of the wavelengths in the group with chemotherapy compared to the group without chemotherapy. Neither did the menopausal status of the patient result in a difference in intensity measured for any of the wavelengths. As explained in the methods section, in the GEE-models of the mixture datasets the percentage of connective tissue was added as an extra covariate. In the dataset including a mixture of fat and connective tissue, the percentage of connective tissue in the measurement location was significant for the wavelengths above 950 nm. In other words, the intensity measured at these wavelengths is dependent on the amount of connective tissue in the measurement. This is in line with the expectations, as there are clear biological differences between fat tissue and connective tissue that are displayed in the higher wavelength range. (Figure 5a)

In the other mixture dataset, with measurement locations that included tumor cells in combination with connective tissue (dataset 4), neoadjuvant chemotherapy also did not affect the intensity at any of the wavelengths nor did the menopausal status. In contrast to the other mixed tissue data set (dataset 3), the percentage connective tissue, or vice versa, in this case, the percentage of tumor cells, did not result in a different intensity at any of the wavelengths. This means that, unlike the clear difference between fat and connective tissue, discriminating connective tissue from tumor cells is much more difficult.

Figure 5. Result of GEE-models in p-values for the two mixture datasets, with connective in combination with fat tissue (dataset 3) and tumor cells (dataset 4). In (a) the results of the dataset consisting of measurement locations that were a mixture of fat and connective tissue (dataset 3). Chemotherapy and menopausal status was not significantly different for any of the wavelengths. The percentage connective tissue in the measurement location was significant for wavelengths above 950 nm. The results of the dataset that consisted of measurement locations with tumor cells and connective tissue (dataset 4) are displayed in (b). Here, neither chemotherapy, nor menopausal status, nor the percentage connective tissue in the measurement location resulted in a significantly different measured intensity.

Fit parameters (analysis 2)

For each of the fit parameters that were deduced from the DRS spectrum, a GEE-model was generated for each dataset. The p-values obtained for each of the fit parameters from the dataset with fat measurements (dataset 1) and the dataset composed of connective measurements (dataset 2) are listed in Table 3.

In the fat measurements (dataset 1), none of the fit parameters were significantly different due to neoadjuvant chemotherapy or menopausal status of the patient. As for the connective dataset (dataset 2), similarly, no significant differences were found for any of the fit parameters between the group of patients with or without chemotherapy treatment, or between the groups of patients that were pre-/perimenopausal or postmenopausal.

Table 4 includes the p-values for the mixture datasets, in which connective tissue is mixed with either fat tissue (dataset 3) or tumor cells (dataset 4). In the analysis of the dataset composed of connective tissue mixed with fat, none of the fit parameters were significantly different when comparing the measurements of patients with chemotherapy with the measurements of the patients without chemotherapy. Also, the menopausal status of the patient did not have a significant influence on any of the five fit parameters. The percentage of connective tissue was highly significant for all fit parameters, thus the amount

Table 3. Result in p-values of the GEE-models of fit parameters of the datasets of fat tissue (dataset 1) and connective tissue (dataset 2).

Fit parameter	Fat dataset		Connective dataset	
	Neo-adjuvant chemo	Menopausal status	Neo-adjuvant chemo	Menopausal status
F/W-ratio	0.51	0.43	0.19	0.77
Fraction Mie scat.	0.52	0.99	0.48	0.82
Collagen (μM)	0.07	0.73	0.61	0.24
α	0.49	0.19	0.12	0.94
b	0.88	0.63	1.00	0.66

Table 4. Result in p-values of the analysis of fit parameters for the mixture datasets of a combination of connective tissue with fat, or tumor cells (dataset 3 & dataset 4).

Fit parameter	Fat & Connective dataset			Tumor cells & Connective dataset		
	Neo-adjuvant chemo	Menopausal status	% connective tissue	Neo-adjuvant chemo	Menopausal status	% connective tissue
F/W-ratio	0.21	0.41	0.00*	0.11	0.40	0.36
Fraction Mie scat.	0.13	0.22	0.054*	0.64	0.32	0.39
Collagen (μM)	0.26	0.48	0.00*	0.00*	0.01*	0.03*
α	0.56	0.10	0.00*	0.87	0.32	0.34
b	0.44	0.08	0.00*	0.91	0.28	0.99

* = significant

of connective tissue in the measurement location had a clear effect on the fit parameters, which is in line with the results of the analysis with the measured intensities (analysis 1).

The analysis of the dataset with measurements of tumor cells in combination with connective tissue (dataset 4) showed that neoadjuvant chemotherapy and menopausal status had again no influence on four of the five fit parameters (F/W-ratio, fraction Mie scattering, α , and b). Also, the amount of connective tissue had no influence on these fit parameters. These findings are in line with the results of the spectral analysis in which chemotherapy, menopausal status, and connective tissue content had no influence on the spectral intensity. The fact that the fit parameter that describes the amount of collagen was related to chemotherapy and menopausal status, as well as the amount of connective tissue, may reflect differences in connective tissue content between patients with and without chemotherapy, and menopausal status.

DISCUSSION

This research was focused on investigating if DRS measurements of breast tissue are affected by neoadjuvant chemotherapy. First, the composition of tissue types of sections

from patients with and without neoadjuvant chemotherapy was analyzed, as well as the composition of the tissue volumes that were measured with DRS. The compositions were compared in order to evaluate any possible selection bias in measurement locations. Next, the DRS measurements of the *ex vivo* specimens of treated and untreated patients were compared by analyzing both the intensity data over the spectrum between 850 and 1600 nm (analysis 1) as well as the fit parameter data (analysis 2). Both analyses were performed with GEE-models that are suitable for assessing correlated data. In the GEE-models, besides treatment with chemotherapy, other confounding factors were included such as the menopausal status, and for the mixture datasets, the amount of connective tissue, as these could also affect the DRS measurement.

We found that the composition of the H&E sections from specimens from patients with neoadjuvant chemotherapy contained slightly more connective tissue and less DCIS. Both results are in agreement with our expectations as one of the effects of chemotherapy is the formation of fibrosis and thus an increase in connective tissue. The lower percentage DCIS in the neoadjuvant chemotherapy patients can be explained by the fact that DCIS is not routinely treated with neoadjuvant chemotherapy and hence DCIS will be under-presented in the group treated with neoadjuvant chemotherapy. Comparing the composition of the DRS measurement locations with the composition of the entire H&E section showed that these were similar suggesting that the probe measurements (which in essence are a subset of the sections) are a good representation of the tissue in both the patient's groups, without and with neoadjuvant chemotherapy.

Subsequently, we used the intensity data (analysis 1) to assess if there were wavelengths that were significantly different between the patients treated with neoadjuvant chemotherapy and those without neoadjuvant chemotherapy. This was done in four separate datasets that had similar histopathology. Significance was absent in all four datasets based on the measured intensity, indicating that there were no intensity differences related to neoadjuvant chemotherapy at any of the wavelengths from 850 to 1600 nm. In the analysis also the menopausal status was not significant for any of the wavelengths in any of the datasets.

In the mixture datasets consisting of connective tissue in combination with fat (dataset 3) or tumor cells (dataset 4), the percentage of connective tissue in the measurement was included as a covariate. In the dataset that included connective tissue with fat (dataset 3), this covariate was significant for wavelengths above 950 nm. However, in the other dataset in which connective tissue was mixed with tumor cells, none of the wavelengths were significant. This finding suggests that discriminating connective tissue from tumor cells is more difficult than discriminating connective tissue from fat tissue. For discriminating fat from connective, the observed wavelengths can be used. For discriminating connective tissue from tumor cells, no such clear individual wavelengths could be identified by this approach, probably due to the fact that connective/glandular tissue can resemble tumor tissue. [26, 27] Potentially, a combination of wavelengths, that individually are not significant, could discriminate connective tissue from tumor cells, however, this is not investigated with the analysis.

The fit parameters that were derived from the measured spectra by using an analytical fit model were also evaluated with GEE-models (analysis 2). Only fit parameters related to wavelengths above 850 nm were assessed. As for the datasets of pure fat tissue and pure connective tissue (dataset 1 & dataset 2), none of the fit parameters were significantly different in the groups with or without chemotherapy, and the groups of pre-/perimenopausal or postmenopausal patients. For these datasets, the results of analysis 1 and analysis 2 are the same.

In the analysis of the mixture dataset that consisted of a combination of fat tissue and connective tissue (dataset 3), the fit parameters were also not significantly different in the patient groups with and without chemotherapy, or pre-/perimenopausal and postmenopausal. The percentage of connective tissue was significant for all fit parameters, indicating that all fit parameters changed significantly if the percentage of connective tissue changed and can thus be used to discriminate fat tissue from connective tissue. For this dataset, the outcomes of analysis 1 and analysis 2 are again the same. This is also in line with the expectations as the absorption coefficient of connective tissue is different from the absorption coefficient of fat.

As for the last dataset, with measurements of tumor cells and connective tissue, four out of five fit parameters were not significant, confirming the results of the first analysis performed with the measured intensities. However, the fit parameter representing the amount of collagen in a measurement location appeared significant for all three covariates (chemotherapy, menopausal status, and % connective tissue). The fact that this parameter was directly related to the amount of connective tissue, may suggest that the observed differences in chemotherapy status and menopausal status may be related to differences in connective tissue content in the measurement locations. It should also be noted that the fiber distance of 1 mm might be at the border of what can be used in diffusion theory.

There is a large number of publications on the use of Diffuse Optical Spectroscopy (DOS) for monitoring therapy, and even predicting therapy response. [10, 28-31] However, their main focus is on comparing measurements of individual patients, pre and post-chemotherapy, whereas we are interested in comparing the group of patients with and without neoadjuvant chemotherapy. The results presented in this paper show that there is hardly any effect of chemotherapy on the optical measurements. This does certainly not imply that chemotherapy response monitoring is not possible using DRS. It merely means that the contrast between healthy tissue and tumor tissue is not altered as a consequence of the chemotherapy. In other words, due to the chemotherapy, the composition of the tissue might change, but the tissue types itself do not have a significantly different spectrum from the spectra measured in patients without chemotherapy.

There is a study by Laughney et al. which uses spectral frequency domain imaging for margin assessment in which five patients were included that had neoadjuvant chemotherapy. They found that measurement of tumor tissue that had been treated with neoadjuvant chemotherapy had an increased scattering slope, which was suggested to be related to cell shrinkage. [32] This is different from our results as no difference was found

for the scattering parameters. In another study by Keller et al spectra of patients with and without neo-adjuvant chemotherapy were measured and compared. [33] They were able to discriminate normal measurements of patients with neoadjuvant chemotherapy from normal measurements of patients without chemotherapy, suggesting that there might be an influence of the chemotherapy. However, in both studies, the analysis was based on a very limited number of patients ($n = 5$ and $n = 3$).

To the author's knowledge, this is the first research in which the effect of chemotherapy is assessed using a large number of patients. Despite a large number of patients, the datasets of pure connective tissue and connective tissue in combination with tumor cells can still be considered as small, thus conclusions should be drawn with some reservation. The response according to pathology and/or radiology was not included as a covariate in the GEE-models as the definition of response can be complicated and subjective.[12] Furthermore, patients were not subdivided based on chemotherapy regime since, in current day practice, many patients are treated with a combination chemotherapeutics that have different mechanisms of action thus all types of reactions are potentially present in each patient.

Since measurements were taken from *ex vivo* samples after some processing at the pathology department, there was contamination of pathology ink in the measurements. By excluding all wavelengths below 850 nm and fit parameters of this wavelength region influence of pathology ink was reduced as much as possible. However, as a consequence, the impact of chemotherapy on the visual wavelength range could not be investigated. Another limitation of the study is that we did not consider patients that were treated with neoadjuvant hormonal therapy. From the literature, it is known that this type of therapy can also induce tissue changes.[34-36] However, the number of patients with neo-adjuvant hormonal therapy that were included during the time of this study was very limited which hindered including this group of patients in the analyses.

In conclusion, we used two different analysis (intensity at a specific wavelength, and fit parameter data) to assess the influence of chemotherapy on optical characteristics in four different datasets. Additionally, also the influence of menopausal status on the measurements was assessed. The measured intensity was not significantly different in patients with or without chemotherapy in any of the four datasets. Also, comparing the intensity in the measurements of the group of patients that were postmenopausal to those that were pre-/perimenopausal also no significant differences were found in any of the wavelengths. These data were confirmed by the analysis of the fit parameters. All fit parameters in the different datasets were not affected by neoadjuvant chemotherapy, except for the fit parameter that reflects the amount of collagen which showed a significant difference in only one dataset that was comprised of measurements of tumor cells in combination with connective tissue. From these findings, it can be concluded that the impact of neoadjuvant chemotherapy on the optical characteristics of breast cancer tissue and healthy breast tissue is negligible. This means that also after neoadjuvant chemotherapy, margin assessment by DRS during breast-conserving surgery seems feasible.

REFERENCES

- [1] D. J. Evers, R. Nachabé, M. J. Vranken Peeters, J. A. Van der Hage, H. S. Oldenburg, E. J. Rutgers, G. W. Lucassen, B. H. W. Hendriks, J. Wesseling, and T. J. M. Ruers, "Diffuse reflectance spectroscopy: towards clinical application in breast cancer," *Breast Cancer Research and Treatment*, vol. 137, pp. 155-165, 2013, [[10.1007/s10549-012-2350-8](#)].
- [2] L. L. De Boer, B. G. Molenkamp, T. M. Bydlon, B. H. W. Hendriks, J. Wesseling, H. J. C. M. Sterenborg, and T. J. M. Ruers, "Fat/Water ratios measured with diffuse reflectance spectroscopy to detect breast tumor boundaries," *Breast Cancer Research and Treatment*, vol. 152, no. 3, pp. 509-518, 2015, [[10.1007/s10549-015-3487-z](#)].
- [3] L. L. De Boer, T. M. Bydlon, F. Van Duijnhoven, M.-J. T. F. D. Vranken Peeters, C. E. Loo, H. A. O. Winter-Warnars, J. Sanders, H. J. C. M. Sterenborg, and T. J. M. Ruers, "Towards the use of diffuse reflectance spectroscopy for real-time in vivo detection of breast cancer during surgery," *Journal of Translational Medicine*, vol. 16, no. 367, 2018, [[10.1186/s12967-018-1747-5](#)].
- [4] Z. Volynskaya, A. S. Haka, K. L. Bechtel, M. Fitzmaurice, R. Shenk, N. Wang, J. Nazemi, R. R. Dasari, and M. S. Feld, "Diagnosing breast cancer using diffuse reflectance spectroscopy and intrinsic fluorescence spectroscopy," *Journal of Biomedical Optics*, vol. 13, no. 2, pp. 024012, 2008, [[10.1117/1.2909672](#)].
- [5] B. S. Nichols, C. E. Schindler, J. Q. Brown, L. G. Wilke, C. S. Mulvey, M. S. Krieger, J. Gallagher, J. Geradts, R. A. Greenup, J. A. Von Windheim, and N. Ramanujam, "A Quantitative Diffuse Reflectance Imaging (QDRI) System for Comprehensive Surveillance of the Morphological Landscape in Breast Tumor Margins," *PloS One*, vol. 10, no. 16, pp. e0127525, 2015, [[10.1371/journal.pone.0127525](#)].
- [6] H. Charfare, "Neoadjuvant chemotherapy in breast cancer," *British Journal of Surgery*, vol. 92, pp. 14-23, 2005, [[10.1002/bjs.4840](#)].
- [7] H. Cain, I. R. Macpherson, M. Beresford, S. E. Pinder, J. Pong, and J. M. Dixon, "Neoadjuvant Therapy in Early Breast Cancer: Treatment Considerations and Common Debates in Practice," *Clinical Oncology*, vol. 29, pp. 642-652, 2017, [[10.1016/j.clon.2017.06.003](#)].
- [8] J. A. Van der Hage, C. C. J. H. Van de Velde, and S. J. S. D. Mieog, "Preoperative chemotherapy for women with operable breast cancer (Review)," *Cochrane Database of Systematic Reviews*, no. 2, pp. CD005002, 2007, [[10.1002/14651858.CD005002.pub2](#)].
- [9] A. Charehbili, D. B. Y. Fontein, J. R. Kroep, G. J. Liefers, J. S. D. Mieog, J. W. R. Nortier, and C. J. H. Van de Velde, "Neoadjuvant hormonal therapy for endocrine sensitive breast cancer: A systematic review," *Cancer Treatment Reviews*, vol. 40, pp. 86-92, 2014, [[10.1016/j.ctrv.2013.06.001](#)].
- [10] A. E. Cerussi, V. W. Tanamai, D. Hsiang, J. Butler, R. S. Mehta, and B. J. Tromberg, "Diffuse optical spectroscopic imaging correlates with final pathological response in breast cancer neoadjuvant chemotherapy," *Philosophical Transactions of the Royal Society A*, vol. 369, pp. 4512-4530, 2011, [[10.1098/rsta.2011.0279](#)].
- [11] L. M. Arthur, A. K. Turnbull, L. R. Khan, and J. M. Dixon, "Pre-operative Endocrine Therapy," *Current Breast Cancer Reports*, vol. 9, pp. 202-209, 2017, [[10.1007/s12609-017-0255-6](#)].
- [12] E. Provenzano, V. Bossuyt, G. Viale, D. Cameron, S. Badve, C. Denkert, G. MacGrogan, F. Penault-Llorca, J. Boughey, G. Curigliano, J. M. Dixon, L. Esserman, G. Fastner, T. Kuehn, F. Peintinger, G. Von Minckwitz, J. White, W. Yang, and W. F. Symmans, "Standardization of pathologic evaluation and reporting of postneoadjuvant specimens in clinical trials of breast cancer: recommendations from an international working group," *Modern Pathology*, vol. 28, pp. 1185-1201, 2015, [[10.1038/modpathol.2015.74](#)].
- [13] S. Kennedy, M. J. Merino, S. M. Swain, and M. E. Lippman, "The Effects of Hormonal and Chemotherapy on Tumoral and Nonneoplastic Breast Tissue," *Human Pathology*, vol. 21, pp. 192-198, 1990.

- [14] F. Aktepe, N. Kapucuoğlu, and I. Pak, "The effects of chemotherapy on breast cancer tissue in locally advanced breast cancer," *Histopathology*, vol. 29, pp. 63-67, 1996.
- [15] J.-H. Chen, K. Nie, S. Bahri, C.-C. Hsu, F.-T. Hsu, H.-N. Shih, M. Lin, O. Nalcioglu, and M.-Y. Su, "Decrease in Breast Density in the Contralateral Normal Breast of Patients Receiving Neoadjuvant Chemotherapy: MR Imaging Evaluation," *Radiology*, vol. 255, no. 1, pp. 44-52, 2010, [[10.1148/radiol.09091090](https://doi.org/10.1148/radiol.09091090)].
- [16] F. Fan, "Evaluation and Reporting of Breast Cancer after Neoadjuvant Chemotherapy," *The Open Pathology Journal*, vol. 3, pp. 58-63, 2009.
- [17] S. Sahoo, and S. C. Lester, "Pathology of Breast Carcinomas After Neoadjuvant Chemotherapy," *Archives of Pathology and Laboratory Medicine*, vol. 133, pp. 633-642, 2009.
- [18] L. L. De Boer, B. H. W. Hendriks, F. Van Duijnhoven, M.-J. T. F. D. Vranken Peeters-Baas, K. Van de Vijver, C. E. Loo, K. Józwiak, H. J. C. M. Sterenberg, and T. J. M. Ruers, "Using DRS during breast conserving surgery: identifying robust optical parameters and influence of inter-patient variation," *Biomedical Optics Express*, vol. 7, no. 12, pp. 5188-5200, 2016, [[10.1364/BOE.7.005188](https://doi.org/10.1364/BOE.7.005188)].
- [19] R. Nachabé, B. H. W. Hendriks, A. E. Desjardins, M. van der Voort, M. B. van der Mark, and H. J. C. M. Sterenberg, "Estimation of lipid and water concentrations in scattering media with diffuse optical spectroscopy from 900 to 1600 nm," *J Biomed Opt*, vol. 15, no. 3, pp. 037015, 2010, [[10.1117/1.3454392](https://doi.org/10.1117/1.3454392)].
- [20] R. Nachabé, B. H. W. Hendriks, M. Van der Voort, A. E. Desjardins, and H. J. C. M. Sterenberg, "Estimation of biological chromophores using diffuse optical spectroscopy: benefit of extending the UV-VIS wavelength range to include 1000 to 1600 nm," *Optics Express*, vol. 18, no. 24, pp. 1432-1442, 2010.
- [21] L. L. De Boer, E. Kho, J. Nijkamp, K. K. Van de Vijver, H. J. C. M. Sterenberg, L. C. Ter Beek, and T. J. M. Ruers, "A method for co-registration of optical measurements of breast tissue with histopathology: the importance of accounting for tissue deformations," *Journal of Biomedical Optics*, vol. submitted, 2019.
- [22] J. A. Hanley, A. Negassa, M. D. d. Adwardes, and J. E. Forrester, "Statistical Analysis of Correlated Data Using Generalized Estimating Equations: An Orientation," *American Journal of Epidemiology*, vol. 157, no. 4, pp. 364-375, 2003, [[10.1093/aje/kwf215](https://doi.org/10.1093/aje/kwf215)].
- [23] J. Shults, and S. J. Ratcliffe, "Analysis of multi-level correlated data in the framework of generalized estimating equations via xtmultcorr procedures in Stata and qls functions in Matlab," *Stat. Interface*, vol. 2, pp. 187-196, 2009.
- [24] P. Taroni, G. Quarto, A. Pifferi, F. Abbate, N. Balestreri, S. Menna, E. Cassano, and R. Cubeddu, "Breast Tissue Composition and Its Dependence on Demographic Risk Factors for Breast Cancer: Non-Invasive Assessment by Time Domain Diffuse Optical Spectroscopy," *PloS One*, vol. 10, no. 6, pp. e0128941, 2015, [[10.1371/journal.pone.0128941](https://doi.org/10.1371/journal.pone.0128941)].
- [25] K. M. Blackmore, J. A. Knight, J. Walter, and L. Lilge, "The Association between Breast Tissue and Optical Content and Mammographic Density in Pre- and Post-Menopausal Women," *PloS One*, 2015, [[10.1371/journal.pone.0115851](https://doi.org/10.1371/journal.pone.0115851)].
- [26] E. Kho, L. L. De Boer, K. K. Van de Vijver, F. Van Duijnhoven, M. T. F. D. Vranken Peeters, H. J. C. M. Sterenberg, and T. J. M. Ruers, "Hyperspectral imaging for resection margin assessment during surgery," *Clinical Cancer Research*, vol. accepted, 2019, [[10.1158/1078-0432.CCR-18-2089](https://doi.org/10.1158/1078-0432.CCR-18-2089)].
- [27] P. Taroni, A. M. Paganoni, F. Ieva, A. Pifferi, G. Quarto, F. Abbate, E. Cassano, and R. Cubeddu, "Non-invasive optical estimate of tissue composition to differentiate malignant from benign breast lesions: A pilot study," *Scientific Reports*, vol. 7, pp. 40683, 2017, [[10.1038/srep40683](https://doi.org/10.1038/srep40683)].
- [28] O. Falou, H. Soliman, A. Sadeghi-Naini, S. Iradji, S. Lemon-Wong, J. Zubovits, J. Spayne, R. Dent, M. Trudeau, J. F. Boileau, F. C. Wright, M. J. Yaffe, and G. J. Czarnota, "Diffuse Optical Spectroscopy Evaluation of Treatment Response in Women with Locally Advanced Breast Cancer Receiving Neoadjuvant Chemotherapy," 2012.

- [29] T. D. O'Sullivan, A. Leproux, J. Chen, S. Bahri, A. Matlock, D. Roblyer, C. E. McLaren, W. Chen, A. E. Cerussi, M. Su, and B. J. Tromberg, "Optical imaging correlates with magnetic resonance imaging breast density and reveals composition changes during neoadjuvant chemotherapy," *Breast Cancer Research*, vol. 15, no. R14, 2013.
- [30] M. V. Pavlov, T. I. Kalganova, S. Lyubimtseva, V. I. Plekhanov, G. Y. Golubyatnikov, O. Y. Ilyinskaya, A. G. Orlova, P. V. Subochev, D. V. Safonov, N. M. Shakhova, and A. V. Maslennikova, "Multimodal approach in assessment of the response of breast cancer to neoadjuvant chemotherapy," *Journal of Biomedical Optics*, vol. 23, no. 9, pp. 091410, 2018, [[10.1117/1.JBO.23.9.091410](#)].
- [31] A. Sadeghi-Naini, E. Vorauer, L. Chin, O. Falou, W. T. Tran, F. C. Wright, S. Gandhi, M. J. Yaffe, and G. J. Czarnota, "Early detection of chemotherapy-refractory patients by monitoring textural alterations in diffuse optical spectroscopic images," *Medical Physics*, vol. 42, no. 11, pp. 6130-6146, 2015, [[10.1118/1.4931603](#)].
- [32] A. M. Laughney, V. Krishnaswamy, E. J. Rizzo, M. C. Schwab, R. J. Barth Jr, D. J. Cuccia, B. Tromberg, K. D. Paulsen, B. W. Pogue, and W. A. Wells, "Spectral discrimination of breast pathologies *in situ* using spatial frequency domain imaging," *Breast Cancer Research*, vol. 15, pp. R61, 2013.
- [33] M. D. Keller, S. K. Majumder, M. C. Kelley, I. M. Meszoely, F. I. Boulos, G. M. Olivares, and A. Mahadevan-Jansen, "Autofluorescence and Diffuse Reflectance Spectroscopy and Spectral Imaging for Breast Surgical Margin Analysis," *Lasers in Surgery and Medicine*, vol. 42, pp. 15-23, 2010, [[10.1002/lsm.20865](#)].
- [34] J. A. Harvey, R. J. Santen, G. R. Petroni, V. E. Bovbjerg, M. E. Smolkin, F. S. Sheriff, and J. Russo, "Histologic changes in the breast with menopausal hormone therapy use: correlation with breast density, estrogen receptor, progesterone receptor, and proliferation indices," *Menopause*, vol. 15, no. 1, pp. 67-73, 2008, [[10.1097/gme.0b013e318054e29a](#)].
- [35] T. J. Anderson, J. M. Dixon, M. Stuart, T. Sahmoud, and W. R. Miller, "Effect of neoadjuvant treatment with anastrozole on tumour histology in postmenopausal women with large operable breast cancer," *British Journal of Cancer*, vol. 87, pp. 334-338, 2002, [[10.1038/sj.bjc.6600435](#)].
- [36] W. R. Miller, J. M. Dixon, L. Macfarlane, D. Cameron, and J. Anderson, "Pathological features of breast cancer response following neoadjuvant treatment with either letrozole or tamoxifen," *European Journal of Cancer*, vol. 39, pp. 462-468, 2003.

Fiber-optic DRS measurements of DCIS and classification of locations with mixed tissue types

L.L. de Boer | E. Kho | K.K. Van de Vijver | M.T.F.D. Vranken Peeters
F. van Duijnhoven | B.H.W. Hendriks | H.J.C.M. Sterenborg | T. J.M. Ruers

In preparation

Intra-operative margin assessment can aid in the detection of positive resection margins during breast-conserving surgery. Diffuse Reflectance Spectroscopy (DRS) that can measure the optical characteristics of tissue could be useful for detecting these positive resection margins. Although previous results are promising, two challenges are yet to be solved. The first challenge considers the detection of DCIS (ductal carcinoma in situ) at the resection margins, as this type of malignancy accounts for the majority of positive resection margins. The second challenge involves the classification of locations that are a mixture of tissue types. Here we investigate 1) if DCIS can be detected by DRS, and 2) how accurate mixture locations can be classified by a classification model, as well as how the performance of this model depends on the purity of the measurements used for the development of the model. For this purpose, measurements were obtained from 107 *ex vivo* breast surgery specimens and were correlated to histopathology. Measurements of DCIS locations were compared with measurements of healthy tissue and invasive carcinoma to assess the ability of DRS to detect DCIS. For the detection of mixed locations, classification models were developed with different definitions of 'healthy' and 'tumor'. As for defining the healthy locations, the percentage of connective tissue that was allowed in the measurement volume ranged between 0% and 100%. For the definition of tumor locations, the percentage of fat tissue that was allowed in the measurement volume ranged between 0% and 90%. Regarding our first research question related to the detection of DCIS, it was found that the differences between healthy tissue and invasive carcinoma are similar to the differences between healthy tissue and DCIS. And, there was no difference in the DRS measurements of invasive carcinoma cells with connective tissue, and DCIS cells with connective tissue. With regard to the second challenge we found that, all classification models with 0-20% fat tissue allowed in the measurements defined as 'tumor', and 0-20% of connective tissue allowed in the measurements defined as 'healthy', showed good performance on the test dataset that included other 'healthy' and 'tumor' measurements. When using these models on the locations with a mixture of tissue types, the model in which no fat was present in the 'tumor' locations and no connective tissue was present in the 'healthy' locations outperformed the other models with less pure data. The accuracy of the correct classification of the measurement locations was dependent on the percentage of tumor cells in the measurement location. Nevertheless, all measurement locations with >75% of tumor cells were correctly classified. When assessing the performance on invasive carcinoma measurement locations compared to DCIS measurement locations, it was concluded that the percentage of tumor

cells was more important for correct classification of a measurement location than whether these locations contained invasive carcinoma cells or DCIS cells. These results indicate that DRS may be a feasible technology for margin assessment in breast conserving surgery. Even in difficult cases where DCIS or mixed locations will be present at the resection edge, DRS may perform well. The next step will be to test the technique *in vivo* during the surgical procedure.

Keywords: resection margin assessment, diffuse reflectance spectroscopy, DCIS, classification models

INTRODUCTION

The majority of patients that are diagnosed with a breast tumor is treated with breast-conserving surgery that aims at resecting a subvolume of the breast that includes the entire tumor. During surgery, the surgeon tries to ensure that the entire tumor has been removed, but since tumor tissue can be difficult to see or feel, it can be challenging to judge if indeed all tumor tissue is resected. Histopathological evaluation, which takes place a few days after surgery, has to confirm whether all tumor tissue is removed adequately. If a so-called positive resection margin is found, this is an indication that tumor tissue might be left behind.

The definition of a positive resection margin has been under heavy debate over the past years and nowadays varies from country to country as well as between institutes and surgeons.[1, 2] For invasive carcinoma the 'no ink on tumor' margin is the most widely accepted definition of an adequate margin. Regarding DCIS, a positive margin is defined as the presence of tumor within a distance of 2 mm from the margin. In comparison to invasive carcinoma, DCIS is more difficult to detect, as pockets of DCIS tend to be smaller.

Of all patients treated with breast-conserving surgery, up to 10-20% require re-excision due to a positive resection margin.[1, 3-7] This has great implications for both the patient, in terms of cosmetic outcome and discomfort, as well as for healthcare budgets.[3] In the current clinical setting, there is no technology that can provide real-time feedback on the margin status, thereby helping surgeons in achieving adequate margins without sacrificing too much healthy tissue.

Diffuse Reflectance Spectroscopy (DRS) has the ability to measure tissue characteristics based on the interaction between light and tissue. This technology is non-destructive, does not rely on exogenous contrast agents and can be performed real-time. There have been previous publications in which DRS spectra are used for margin assessment.[8-11] These papers are based on the extensive literature showing a high accuracy of DRS measurements for differentiating healthy breast tissue from tumor tissue. However, for successful application of the technology in clinical practice still, several fundamental issues need to be resolved.

First of all specific data on whether DRS is able to recognize DCIS is scarce. In some papers, DCIS measurements are included, however, the number of DCIS measurements is often small and the wavelength range that was included during the measurements was often limited to the visual wavelengths.[9, 12-14]

Second, breast tissue can be very inhomogeneous. For example, although healthy tissue predominantly consists of fat tissue, connective tissue can be present as well. Meanwhile, connective tissue is also present in breast cancer tumors, which comprise of tumor cells surrounded by connective tissue in various ratios. In addition, in the case of a positive resection margin, an area with tumor cells and accompanied connective tissue might be also surrounded by fat tissue. Thus, a fiber-optic measurement acquired from breast tissue, also when taken from the resection margin, may consist of a mixture of tissue types. At this

stage, it is still unclear how accurate the measurements that are taken from such mixed locations can be classified, and how the classification accuracy depends on the purity of the data that is used for developing the classification model.

To resolve the current uncertainties the first aim of this paper is to evaluate if DRS is able to detect DCIS. For this purpose measurement of DCIS are compared with the measurements of healthy tissue as well as the measurements of invasive carcinoma. The second goal is to determine if DRS measurements of locations with a mixture of tissue types can be correctly classified. In order to answer this research question, we first investigated how the classification performance of a model depends on the purity of the measurements that were used for developing the model. By changing the definition of 'healthy' and 'tumor', more measurements that were a mixture of tissue types were included in the dataset used for developing the classification model. The performance of that classification model was then tested by classifying a portion of the 'healthy' and 'tumor' measurements (test dataset) that were not used for model development. Classification models with good performance on these test datasets were subsequently used to classify all other locations that were not defined as 'healthy' or 'tumor' (mix dataset) to assess what the performance of the classification models was on measurements of mixtures of tissue types.

MATERIALS & METHODS

Measurements on tissue slices

Fiber-optic diffuse reflectance spectra were acquired from slices from resection specimens of 107 patients who had undergone breast surgery, either mastectomy or breast-conserving surgery, in the Netherlands Cancer Institute. The study was approved by the hospital and according to Dutch guidelines and confirmed by the Local Ethics Committee, no informed consent had to be acquired.

Directly after surgery, at the pathology department, a pathologist inked the margins and cut the specimen in slices with a thickness between 3 and 10 mm. One of the slices from the middle of the specimen was placed in a macrocassette with a piece of black rubber underneath it, to ensure the spectral data originated from the tissue only. Subsequently, a custom made grid, that was placed on top of the specimen, defined the fiber-optic measurement locations and enabled accurate correlation with histopathology afterward.

Measurement set-up

The measurement set up consisted of two spectrometers (Andor Technology DU420A-BRDD and Andor Technology, DU492A-1.7) covering the wavelength range between 400 and 1600 nm. A blunt probe was attached to this measurement set-up with fibers that allowed to obtain a DRS measurement with a fiber distance of 1 mm between the illuminating and

collecting fiber. A calibration cap with a Spectralon reference standard at the bottom was used to white-calibrate the system prior to each measuring session. This measurement set-up has been described in detail elsewhere.[15, 16]

Histopathology and registration with optical measurements

After measuring, the tissue was further processed according to standard protocol in H&E sections that were later digitalized using the Aperio® ScanScope AT2 (Leica Biosystems, Wetzlar, Germany).

A pathologist assessed each H&E section, and circumscribed areas with invasive carcinoma, DCIS, and glandular ducts. The other tissue types, connective tissue and fat tissue stain respectively pink or white in the H&E section and were annotated automatically by digital thresholding the image of the H&E section. Eventually, the glandular ducts were included in the connective tissue class. The result was an image in which each pixel has a histopathology label; IC, DCIS, Connective or Fat.

The procedure was described in detail elsewhere[17] (Chapter 5), in short, first, the image of the H&E section (without annotations) was registered to an overview white light image of the tissue in the macrocassette. The transformation matrix derived from this registration allowed correcting for tissue deformations in the H&E section that occurred during the histopathology processing. This transformation matrix was applied to the image of the H&E section including the tissue type labels. The result was an image in which each pixel is labeled, that is corrected for deformations. With a second white light image of the tissue with the grid on top, it was possible to retrieve the measurement locations in the labeled image and thus finding the tissue labels of the pixels at each measurement location.

Additionally, the measurement locations that were located in areas that were defined as invasive carcinoma or DCIS were assessed a second time by a pathologist to have a more detailed estimation of the percentage tumor cells in each of those measurement locations. The result was a scoring of the percentage of fat, connective tissue, invasive carcinoma cells, and DCIS cells in each of the locations that were measured.

Extraction of spectral features from spectra

Although spectra were acquired between 400 and 1600 nm only the wavelengths above 850 nm were used in the analysis as the wavelengths below 850 nm were affected by pathology paint that was present on the specimen at the time of measuring. Second, the number of features in the measurements was reduced by quantifying a number of visible spectral features in specific wavelength ranges of the spectrum, specifically: I) spectral slopes, II) dips in the spectra and III) the wavelength of these dips. (Figure 1) Areas that were interesting were selected by visual inspection of the spectra of 'pure' measurement locations. (Figure 1a) The 'pure' spectra were based on locations with >95% fat tissue, >95% connective tissue, and locations with >90% tumor cells. The slopes in these areas were

Figure 1. Reduction of features. The mean spectra and standard deviations of pure measurement locations are used to select wavelength areas (a). In (b) an example of how a spectral features was calculated from the 'dip' between 1100 nm and 1280 nm. The parts of the spectra that are colored red are the first and last 25% of the wavelength range. The black dots represent the maxima in these areas. Three types of spectral features are derived from the spectra, the slope (I), the maximum difference (II), and the wavelength of the maximum difference (III). The slopes of 12 areas derived from a typical fat spectrum (c) and a typical connective spectrum (d). The 'dips' of 4 areas in a typical fat spectrum (e) and a typical connective spectrum (f)

determined by first calculating the maximum intensity in the first 25% of the wavelength range, as well as the maximum intensity in the last 25% of the wavelength range.

Then, between these two maxima, the slope was calculated and used as a spectral feature (I) (Figure 1b). In some areas were a 'dip' was seen in the spectra another two spectral

Table 1. Overview of areas selected for spectral feature reduction.

Area	Slope (I)	Maximum difference (II)	Wavelength of max. difference (III)
865-1070 nm	x		
900-1050 nm	x	x	x
930-1048 nm	x	x	x
990-1052 nm	x		
1060-1126 nm	x		
1100-1265 nm	x	x	x
1130-1155 nm	x		
1210-1242 nm	x		
1265-1440 nm	x		
1300-1365 nm	x		
1335-1400 nm	x		
1365-1440 nm	x		
1385-1450 nm	x		
1385-1600 nm	x	x	x
1440-1600 nm	x		
1500-1600 nm	x		

features were derived from the selected wavelength range. The first was the maximum difference between the slope and the measured spectrum (II). The second spectral feature was the wavelength of the maximum difference between the slope and measured spectrum (III). All spectral features that were acquired from the spectra are listed in Table 1. In total 24 spectral features were derived from each measured spectrum. (Figure 1c, Figur 1d, Figure 1e & Figure 1f)

Measurements of DCIS

To investigate the ability of DRS to detect DCIS, we compared the spectra of ‘pure’ invasive carcinoma, ‘pure’ DCIS and healthy tissue. For this purpose, we selected three groups, one group with measurements in which 95% or more of the volume included connective tissue with DCIS, one group of measurements that for 95% or more included invasive carcinoma with connective tissue, and one group of measurements of ‘pure’ healthy tissue. This last group includes all locations without invasive carcinoma or DCIS which thus consist of fat tissue and/or connective tissue. First, we analysed the differences between measurements of invasive carcinoma and healthy tissue and the differences between the measurements of the DCIS locations and healthy tissue. Second, the differences in the measurements of the invasive carcinoma and DCIS locations are compared. For the comparisons, a Wilcoxon ranked sum test is used to evaluate either 1) the measured intensities at every single wavelength or 2) the spectral features derived from the spectrum as described in the section ‘Extraction of spectral features from spectra’. The α -level for all statistical tests was set to 0.05.

Development of classification models

To assess how model performance depends on the purity of the measurements used for the development of the model, different definitions of 'healthy' and 'tumor' were used. For the healthy locations the definition was that 1) no tumor cells were present, and 2) the percentage of connective tissue at the measurement location did not exceed $\max_{\text{con}}$, in which $\max_{\text{con}}$ was set to either 0%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% or 100%. Similarly, locations were defined as 'tumor' if the following two criteria were met, 1) presence of tumor cells and 2) the percentage of fat tissue at the measurement location did not exceed $\max_{\text{fat}}$, in which $\max_{\text{fat}}$ was set to either 0%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% or 90%. There were no models generated with a $\max_{\text{fat}}$ of 100%, as based on this definition tumor locations were allowed to contain 100% fat tissue and thus these locations would meet the definition of both 'healthy' and 'tumor'.

The classification models were built with a linear support vector machine (SVM) as classifier, that was weighted based on the portion of locations that were defined as 'healthy' and 'tumor'. First, five-fold cross-validation was used to split the data in a dataset used for developing the classification model and a test dataset. Subsequently, a second five-fold cross-validation step was used to split the development dataset a second time in a dataset for training the model and a dataset for optimizing the SVM. (Figure 2)

Since for each combination of $\max_{\text{fat}}$ and $\max_{\text{con}}$ a classification model was constructed, the total number of classification models was 110 (10 possibilities of $\max_{\text{fat}}$ and 11 possibilities of $\max_{\text{con}}$).

Measurements from one patient were either used for training, or optimizing, or testing. Performance of the models was evaluated with the sensitivity, specificity, and Matthews Correlation Coefficient (MCC) of the test data. The classification models were assessed on overfitting, by comparing the performance of the model on the test dataset with the performance of the model on the dataset used for developing the model (dataset for training plus optimizing dataset).

Classification of mixed locations

All measurements that did not meet the definition of either 'healthy' or 'tumor', were labeled as 'mix' (mix dataset). Since the definitions for locations that were used for model development (those locations defined as 'tumor' or 'healthy') changed, also the composition of the dataset with mixture locations changed. Thus, performance on mixture locations also depended on which locations are used for model development. Only mix measurements of patients that were in the test dataset were classified by the model. To assess the performance of the classification models on the mixture dataset the classification output ('healthy' or 'tumor') was plotted against the percentage of tumor cells in the measurement volume (0%, 1-25%, 26-50%, 51-75%, 76-100%).

Figure 2. Flow chart of classification model development. The colored boxes represent data that is classified by the optimized support vector machine. Model performance is evaluating by calculating the sensitivity, specificity and MCC of the test dataset (light green). To check for overfitting, these metrics are compared with the sensitivity, specificity and MCC of the dataset used for model development (turquoise). The mix dataset (orange) that is classified by the model only consists of mix measurements from patients that were not used in the training data or optimizing data.

RESULTS

Measurements of DCIS

To investigate the ability of DRS to detect DCIS, we analyzed the spectra of tumor locations that consisted either of IC or DCIS with surrounding connective tissue and compared those with the measurements of pure healthy measurement locations. In Figure 3 an example of an H&E section of both a location with invasive carcinoma (Figure 3a) and DCIS (Figure 3b) are displayed.

In both locations, the estimated amount of connective tissue was 50%. To evaluate a possible bias with regard to connective tissue content between invasive carcinoma and DCIS samples we first analyzed the histological composition of both sets of locations. All locations that were included in the analysis are represented by a bar in either Figure 3c

Figure 3. Comparison of invasive carcinoma with DCIS. Image of an H&E section of a measurement location consisting of 50% tumor cells and 50% connective tissue (a). In (b) an example of an H&E section of a measurement location with 50% DCIS cells, and 50% connective tissue. The composition of all measurement locations that included invasive carcinoma cells in combination with connective tissue depicted in (c), and the composition of all measurement locations that included DCIS cells in combination with connective tissue are displayed in (d). A few locations had some inflammation or scattered fat cells (<5%) which are represented in gray ('other'). In (e) mean spectra and standard deviation of invasive carcinoma with connective tissue (red), DCIS with connective tissue (magenta), and all healthy locations without any invasive carcinoma or DCIS (grey). The results of the Wilcoxon ranked sum test over all wavelengths are displayed in (f). None of the wavelengths is reaches significance for the comparison of invasive carcinoma versus DCIS locations, whereas for the comparisons of both invasive carcinoma and DCIS with all healthy locations the majority of the wavelengths are significant.

(invasive carcinoma) or Figure 3d (DCIS). In total 69 locations ($n = 32$ patients) consisted of invasive carcinoma with some connective tissue surrounding the tumor cells whereas there were 26 locations ($n = 11$ patients) that consisted of only DCIS and connective tissue. To assure that the groups were comparable, the mean percentage and standard deviation of the percentage of connective tissue in the measurement locations were calculated for both groups. In the invasive carcinoma group 59.4% (25.0%) was connective tissue, and in the DCIS measurements on an average 51.0% (22.6%) connective tissue was present.

The mean spectra of the invasive carcinoma locations, DCIS locations, and healthy locations are depicted in Figure 3e. The spectra of the invasive carcinoma locations and DCIS locations are very similar, and standard deviations overlap almost entirely. The two mean spectra of the malignant tissue types show distinct differences with the mean spectrum of healthy tissue.

The results of the Wilcoxon ranked sum test for the comparisons of the measurements of invasive carcinoma versus healthy (red line), DCIS versus healthy (magenta line), and invasive carcinoma versus DCIS (grey line) are depicted in Figure 3f. The red line and magenta line reach both below the α -level (dashed line) for many of the same wavelengths. This result indicates that the spectral difference between invasive carcinoma locations and healthy locations is comparable to the spectral difference between DCIS locations and healthy locations. The intensity differences between invasive carcinoma and DCIS locations were also compared with a Wilcoxon ranked sum test. The result of this is displayed with the black line, which was not significant for any of the wavelengths. (Figure 3f) Thus, the measured intensities do not significantly differ between the invasive carcinoma and DCIS measurements for any of the wavelengths, and discriminating the two tissue types based on the measured intensity at a certain wavelength will not be possible.

In addition, the same statistical test was used with respect to the spectral features, again comparing invasive carcinoma locations with healthy measurements, comparing DCIS with healthy measurements, and comparing invasive carcinoma locations with DCIS locations. The Wilcoxon ranked sum test was performed with each spectral feature and the results are listed in Table 2. The significant spectral features are marked with an asterisk (*). When comparing the spectral features of invasive carcinoma with healthy tissue measurements 19 of the 24 spectral features were significantly different. Similarly, 22 spectral features were significantly different when comparing the spectral features of DCIS measurements with the spectral features in healthy tissue locations.

The majority of the spectral features are highly significant in both comparisons. On the other hand, none of the spectral features are significantly different in the invasive carcinoma measurements compared to the DCIS measurements.

These data show that DRS measurements of areas with DCIS combined with connective tissue are not significantly different from the spectra measured in areas of invasive carcinoma combined with connective tissue. Also, the difference between measurements of healthy

Table 2. P-values of Wilcoxon ranked sum test on spectral features. Features marked with an asterisk (*) are significant.

Area	Slope (I)			Maximum difference (II)			Wl. of max. difference (III)		
	IC vs H	DCIS vs H	IC vs DCIS	IC vs H	DCIS vs H	IC vs DCIS	IC vs H	DCIS vs H	IC vs DCIS
865-1070 nm	0.00*	0.00*	0.12						
900-1050 nm	0.00*	0.00*	0.12	0.00*	0.00*	0.08	0.00*	0.00*	0.22
930-1048 nm	0.00*	0.00*	0.06	0.00*	0.00*	0.07	0.01*	0.49	0.30
990-1052 nm	0.00*	0.00*	0.27						
1060-1126 nm	0.00*	0.00*	0.10						
1100-1265 nm	0.00*	0.00*	0.06	0.01*	0.00*	0.37	0.00*	0.00*	0.96
1130-1155 nm	0.00*	0.00*	0.19						
1210-1242 nm	0.00*	0.00*	0.11						
1265-1440 nm	0.00*	0.10	0.25						
1300-1365 nm	0.00*	0.00*	0.24						
1335-1400 nm	0.00*	0.00*	0.28						
1365-1440 nm	0.10	0.00*	0.29						
1385-1450 nm	0.72	0.00*	0.26						
1385-1600 nm	0.00*	0.00*	0.08	0.99	0.00*	0.14	0.00*	0.00*	0.55
1440-1600 nm	0.22	0.00*	0.39						
1500-1600 nm	0.53	0.00*	0.38						

IC = invasive carcinoma, H = healthy, DCIS = ductal carcinoma in situ.

tissue and invasive carcinoma is comparable to the difference between measurements of healthy tissue and DCIS.

According to these results, for the remainder of this study, we will combine the data of the locations of invasive carcinoma with DCIS and will express, the percentage of invasive carcinoma cells and DCIS cells by the percentage of tumor cells.

Classification model performance

The second goal of this research was to investigate if DRS measurements of locations with a mixture of tissue types could be classified correctly. In order to do so, first, the impact of using locations that are a mixture of tissue types in the classification model development is assessed. For this purpose, classification models are generated with different definitions of 'healthy' and 'tumor'. Through cross-validation, the models are tested with part of the measurements that were defined as 'healthy' or 'tumor' (test dataset). The average performances of all the classification models with the different definitions of 'healthy' and 'tumor' are displayed in Figure 4. In each graph, one square belongs to one model with a combination of $\max_{\text{fat}}$ and $\max_{\text{con}}$. The model based on a $\max_{\text{fat}}$ of 90% and a $\max_{\text{con}}$ of 100% could not be generated as almost all data was used for development of the model, and too little data was left to test the model. The left side of the figure show respectively the sensitivity (Figure 4a), specificity (Figure 4c) and MCC (Figure 4e) of the test data, and

Figure 4. Performance of classification models in terms of sensitivity ((a) & (b)), specificity ((c) & (d)), and MCC value ((e) & (f)). The left part of the figure displays the results on the test data ((a), (c), & (e)), whereas the right part of the figure displays the results on the dataset used in model development ((b), (d), & (f)). The dashed line indicated the models with max_{fat} between 0 and 20%, and a max_{con} between 0 and 20%. The model with $\text{max}_{\text{fat}} = 90\%$ and $\text{max}_{\text{con}} = 100\%$ could not be generated as there was too little data left for testing.

the right side of the figure show the sensitivity (Figure 4b), specificity (Figure 4d) and MCC (Figure 4f) of the data that was used for developing the model. Since the sensitivity, specificity, and MCC values on the test dataset and model development dataset are almost similar for all models, it can be concluded that none of the models suffer from overfitting.

The best performance is obtained when the definition of tumor measurements and healthy measurements is most strict ($\max_{\text{fat}} = 0\%$ and $\max_{\text{con}} = 0\%$). In this scenario, no fat tissue is allowed in the measurements defined as 'tumor', and no connective tissue is allowed in the measurements defined as 'healthy'. With both allowing more fat in the definition of tumor and more connective in the definition of healthy, the performance of the models deteriorate. However, in general, the classification models in which, up to 20% connective tissue is allowed in the healthy measurements, and up to 20% of fat tissue in the healthy measurements show good performance. The MCC is 0.96, 0.95, 0.91 and 0.90 respectively for the models with $\max_{\text{fat}} = 0\%$ & $\max_{\text{con}} = 0\%$, $\max_{\text{fat}} = 0\%$ & $\max_{\text{con}} = 20\%$, $\max_{\text{fat}} = 20\%$ & $\max_{\text{con}} = 0\%$, and $\max_{\text{fat}} = 20\%$ & $\max_{\text{con}} = 20\%$. Sensitivity for the classification model in the same order was 0.97, 0.96, 0.95, and 0.89. And similarly, the specificity was 0.99, 0.99, 0.97, and 0.99 for the classification models in the previously used order. These models correspond to the corners of the dashed box as displayed in Figure 4.

Classification of mixed tissue locations

To assess the accuracy of the classification models for the mixture locations (i.e. locations that were not considered 'healthy' or 'tumor', and were thus labeled as 'mix'), these locations were classified separately. As the results of the test dataset that were previously presented showed that the models with a $\max_{\text{fat}}$ ranging between 0% and 20% and a $\max_{\text{con}}$ ranging between 0% and 20% had good performance on the test data, four models in this range were used for classifying these mixed data. In Figure 5 the performances of the four models used for classifying the mixture data are depicted (Figure 5a: $\max_{\text{fat}} = 0\%$ and $\max_{\text{con}} = 0\%$, Figure 5b: $\max_{\text{fat}} = 0\%$ and $\max_{\text{con}} = 20\%$, Figure 5c: $\max_{\text{fat}} = 20\%$ and $\max_{\text{con}} = 0\%$, Figure 5d: $\max_{\text{fat}} = 20\%$ and $\max_{\text{con}} = 20\%$). These are the models in the corners of the dashed box in Figure 4. Each graph displays the mean percentage of locations that are classified as healthy (green bars above the x-axis) and that is classified as tumor (red bars below the x-axis). The locations are grouped based on the percentage of tumor cells in the location (0%, 1-25%, 26-50%, 51-75%, and 76-100%), which are displayed as separate bars. In the ideal situation all locations with 0% tumor tissue are depicted as a green bar, and all locations with tumor cells, thus all other bars, are depicted in red.

The model generated with a $\max_{\text{fat}}$ of 0% and $\max_{\text{con}}$ of 0% (Figure 5a), classified an average of 20.7% of measurement locations that did not have any tumor cells as tumor (left bar). Measurements in the group 1-25% tumor cells are classified as tumor in 35.7% of cases on an average. The group of measurements that contains 26- 50% tumor cells is classified as tumor in 54.8% of cases on an average, and for the group with 51-75% tumor cells on an

Figure 5. Results of classifying the mixture locations with different $\max_{fat}$ and $\max_{con}$. In Figure 5a: $\max_{fat} = 0\%$ and $\max_{con} = 0\%$. In Figure 5b: $\max_{fat} = 0\%$ and $\max_{con} = 20\%$. In Figure 5c: $\max_{fat} = 20\%$ and $\max_{con} = 0\%$. In Figure 5d: $\max_{fat} = 20\%$ and $\max_{con} = 20\%$. Each bar in a graph represents a group of measurement locations with either 0% tumor cells, 1-25% tumor cells, 26-50% tumor cells, 51-75% tumor cells, or 76-100% tumor cells. The height of the green bars depict the average percentage of measurement locations classified as healthy, the red bars depict the average percentage of measurement locations that are classified as tumor.

average 69.9% of cases are classified as tumor. All locations with a 76 to 100% tumor cells were classified as tumor. Thus, with an increasing percentage of tumor cells, the number of locations are classified as tumor increases.

Figure 5b shows the effect of allowing more connective tissue in the healthy locations. Using this model the average percentage of locations without tumor cells that are classified as tumor is higher (33.7%), compared to the model with a $\max_{con}$ of 0% (20.7%). At the same time, in the groups with some tumor cells, more locations are classified as healthy. Comparing Figure 5a with 5b the average percentages of locations that were classified as healthy in the groups 1-25% tumor cells, 26-50% tumor cells, and 51-75% tumor cells, increased with respectively 2.5% (from 64.3% to 66.8%), 19.7% (from 45.2% to 64.9%), and 26.6% (from 30.1% to 56.7%). Despite, the decline in performance of the locations with 1-75% tumor cells, still, none of the locations that contained between 76% and 100% of tumor cells were classified as healthy. The changed definition of 'healthy' also led to a decrease in the average percentage of locations that have no tumor cells and are classified as healthy.

In Figure 5c the effect of allowing more fat tissue in the definition of a tumor location is assessed on the classification of the mix dataset. In this scenario, all locations that contain 76 to 100% tumor cells are used in the model development and will thus not fall in the mixture dataset. Compared to the model with $\max_{\text{fat}}$ of 0% and $\max_{\text{con}}$ of 0% (Figure 5a), more locations with tumor cells are classified as tumor. In the group with 1-25% tumor cells, the average percentage that is classified as tumor increased from 35.7% to 41.4%. Similarly, in the group with 26-50% tumor cells the average percentage increases from 54.8% to 76.1%, and in the group of 51%-75%, the average percentage raised from 69.9% to 100%. However, also the average percentage of locations with 0% tumor cells that are classified as tumor is 33.7%, whereas this was only 20.7% in the model with $\max_{\text{fat}} = 0\%$.

In the scenario in which $\max_{\text{fat}}$ is set to 20% and $\max_{\text{con}}$ is set to 20% (Figure 5d), similar to Figure 5c the group of measurements with 75-100% tumor cells is no longer present in the mix dataset. The average percentage of locations without tumor cells that are classified as healthy increased in comparison to Figure 5a (44.8% compared to 20.7%). On the other hand, the average percentages of locations with some tumor cells that are classified as tumor, decreases for the group with 1-25% tumor cells (35.7% versus 21.1%), decreases slightly for the group with 26-50% tumor cells (54.8% versus 53.3%), and decreases severely for the measurement locations with 51-75% tumor cells (69.9% versus 50.0%). As can be seen in the figure, the standard deviation of this last group is quite large.

Altogether the performance of this last model seems to suffer from the impurity of both the healthy locations as well as the tumor locations. This is an indication that using more mixture locations in the development of the classification model does not improve the correct classification of locations that have a mixture of tissue types. From these data it is concluded that for model development, the purest measurements ($\max_{\text{fat}} = 0\%$ and $\max_{\text{con}} = 0\%$) provided the best model for classifying both the test dataset as well as the mix measurements dataset.

DCIS vs IC

To examine if the detection of tumor tissue depends on the actual tissue type, the measurements with invasive carcinoma were split from the measurements with DCIS. All measurements with either one of these tissue types were classified using the model with $\max_{\text{fat}} = 0\%$ and $\max_{\text{con}} = 0\%$. Figure 6 depicts the classification results on measurements with invasive carcinoma (a) or DCIS (b). Both graphs show the same trend, with an increasing amount of tumor cells (either invasive carcinoma or DCIS) the average percentage of locations that are classified as tumor increases. In both groups, locations with more than 76% of tumor cells are always classified correctly. Similarly, locations with 51-75% of DCIS cells were always classified correctly, however, it should be noted that this group was very small with only three locations.

Figure 6. The classification results of the classification model with $\max_{fat} = 0\%$ and $\max_{con} = 0\%$ separated based on tissue type. The average percentage of locations that are classified as healthy of tumor for each of the groups with increasing percentage of invasive carcinoma (a), and DCIS (b).

These results confirm that DCIS can be classified correctly by the classification model, similar to invasive carcinoma. The classification accuracy is more dependent on the volume that is occupied by tumor cells, than the actual tissue type.

DISCUSSION

The aim of this study was first, to investigate if DRS measurements would be able to detect DCIS. The second goal was to assess the performance of a classification model for locations of mixtures of tissue types. In order to do so, the first step was to assess how model performance depended on the purity of the measurements used for developing the model. In other words, if using more mixture locations in the model development phase would affect the ability of the model to correctly classify the measurements. Thus, the performance of the models was first evaluated by comparing the output on test dataset that was labeled as 'healthy' or 'tumor'. Subsequently, the models with good performance with the test data were used to classify the measurements that were considered 'mix', and which were thus not used in model development.

As for the research question concerning the detection of DCIS, the measurements of two groups with either connective tissue and invasive carcinoma cells or connective tissue and DCIS cells were compared. The means percentage of connective tissue in both groups was comparable. It was found that the spectral differences that are seen between invasive carcinoma and healthy tissue are also seen in the comparison between DCIS locations and healthy tissue. The wavelengths that were significantly different between healthy tissue measurements and invasive carcinoma measurements are almost the same as the wavelengths that were significantly different between healthy tissue measurements and DCIS measurements. In the analysis of the spectral features also the majority of the spectral features that were significant for the comparison between healthy and invasive carcinoma

were also significant for the comparison between healthy and DCIS. Furthermore, the measurements of locations with invasive carcinoma were compared with the measurements of locations with DCIS. It was found that there were no wavelengths that had a significantly different intensity comparing the invasive carcinoma measurements with the DCIS measurements. Neither was any of the spectral features that were derived from the spectrum between 850 and 1600 nm significant. The fact that DCIS and invasive carcinoma are optically similar is not surprising as both comprise irregular cells with often less cytoplasm and atypical nuclei. For a pathologist, in situ carcinoma turns into invasive if the myoepithelial cells of the outer ductal layer are no longer preserved.[18] This might not be detectable with diffuse reflectance explaining why these tissue types could not be separated based on the DRS measurements. It was therefore concluded that the measurements of invasive carcinoma and DCIS should be combined in one class. There have been a few reports on DRS measurements of DCIS.[9, 12-14] In these publications measurements were acquired over a smaller wavelength range, typically enclosing the visual wavelengths, and often the analysis was performed with optical parameters derived from the spectrum. It is therefore not possible to compare the results of these studies with the results of our analysis.

In total 109 classification models were developed and used for classifying data as the model with $\max_{\text{fat}}$ 90% and $\max_{\text{con}}$ 100% could not be generated. The performance of the classification models on the test dataset was compared with the performance of the classification models on the dataset that was used for developing the model (Figure 4). The sensitivity, specificity, and MCC were comparable for both datasets indicating that the classification models are not suffering from overfitting. We found that when an increase in the percentage of fat tissue was allowed in the tumor locations ($\max_{\text{fat}}$), and the percentage connective tissue in the healthy locations ($\max_{\text{con}}$), the performance of the classification models deteriorates. This outcome is in line with expectations. Healthy tissue mainly consists of fat, sometimes combined with healthy ducts (i.e. glandular tissue) and thus connective tissue. Tumor tissue, on the other hand, is always a combination of tumor cells with to some connective tissue. By allowing more fat in the measurements labeled as tumor it becomes more alike healthy tissue and by allowing more connective tissue in the measurements that were labeled as healthy it will more and more resemble tumor tissue. However, increasing both the $\max_{\text{fat}}$ and $\max_{\text{con}}$ to 20% still allowed good discrimination between the healthy measurements and tumor measurements that were present in the test data (i.e. other locations that were labeled as 'healthy' or 'tumor').

Subsequently, four models were used to classify the measurements that were considered a mix (Figure 5). The model with both the $\max_{\text{fat}}$ and $\max_{\text{con}}$ set to 0% clearly showed that with an increasing percentage of tumor cell in the measurement volume the average percentage of locations that are classified as tumor increased. (Figure 5a) On the other hand, an average of 20.7% of locations without any tumor cells was falsely classified as tumor. The model in which the percentage of connective tissue that was allowed in the

healthy locations was set to 20% ($\max_{\text{con}} = 20\%$) (Figure 5b) showed worse performance than the model in which both $\max_{\text{con}}$ and $\max_{\text{fat}}$ were set to 0%. In this model, both the average percentage of locations without tumor tissue that was classified as tumor increased, while at the same time the average percentage of locations with some tumor (1-25%, 26-50%, and 51-75%) that were classified as tumor decreased. This could be explained by the fact that by allowing more connective tissue in the definition of healthy, more measurement locations with some tumor tissue (i.e. a combination of tumor cells with connective tissue) in the measurement volume will resemble 'healthy'. It suggests that the presence of connective tissue is used by the classification model as an indication of the presence of tumor. The decrease in the average percentage of locations without tumor cells that are classified as healthy can be explained by the fact that locations with a high percentage of fat tissue and up to 20% of connective tissue are included in the test dataset and are no longer present in the mixture dataset. Thus, locations with 1 to 20% connective tissue and 80 to 99% fat tissue are used in the development of the classification models and therefore not present in the mix dataset. These locations have relatively a lot of fat which probably makes it easy to classify them correctly as 'healthy'. In that case, these locations would have contributed to increasing the average percentage of locations without any tumor cells that are classified as healthy. It is concluded that this model is not to be favored over the model with both $\max_{\text{con}}$ and $\max_{\text{fat}}$ set to 0%.

Besides the impact of changing the definition of 'healthy' also the impact of a changed definition of 'tumor' was assessed with the mixed dataset. In this case, in which $\max_{\text{fat}}$ set to 20% (Figure 5c), the average percentage of locations with some tumor that were classified as tumor increased. Noteworthy, in this scenario, all locations with 76%-100% tumor cells were included in the test dataset. The negative effect of changing the definition of 'tumor' was that also the average percentage of locations without tumor cells that are classified as tumor increased substantially from 20.7% to 33.7%. The result of allowing more fat tissue in the definition of tumor is that locations that have tumor cells but also some fat tissue are more similar to what is defined as tumor. As a result, the average percentage of locations with some tumor cells but also some fat that are classified as tumor increased. The increase of locations without tumor cells that are classified as tumor can be explained since some healthy locations will also have connective tissue, which are then identified by the model as tumor since tumor cells are in general surrounded by connective tissue.

Finally, the classification model in which both the $\max_{\text{fat}}$ and $\max_{\text{con}}$ are set to 20% was assessed. (Figure 5d) Results show that this classification model performs poor on the mixed dataset, especially compared to the classification with the purest definition of 'healthy' and 'tumor' (i.e. $\max_{\text{fat}}$ and $\max_{\text{con}}$ set to 0%).

It can thus be concluded that the best classification results for the mixed locations are obtained by using measurements without any connective tissue for the 'healthy' class, and measurements without any fat for the 'tumor' class in developing a classification model.

The effect of changing the definitions of 'healthy' and 'tumor' on the classification of the mixture measurements was similar to the trends seen in the performance on the classification of the test dataset. It should be noted, however, that in the analysis alterations of the definition of both 'healthy' and 'tumor' affect which data is used for developing the model, and which data is included in the mix dataset. Increasing $\max_{\text{fat}}$ and $\max_{\text{con}}$ results in a smaller mix dataset, and more measurement locations that are used in the development of the models. The locations that switch from mix dataset to the development dataset are those locations that differ from pure locations maximally 20%. Most likely, these locations are more easy to classify by the classification models than locations in which the ratio is more fifty-fifty. However, since we were especially interested in the classification performance of the models for the mix locations the test dataset was not combined with the mix dataset. This way the difference in performance between the models became more apparent.

Lastly, the performance of the classification model ($\max_{\text{fat}} = 0\%$ and $\max_{\text{con}} = 0\%$) was assessed for invasive carcinoma and DCIS separately (Figure 6). The performance of the classification model for the measurement locations that contain either invasive carcinoma cells or DCIS cells is very comparable. Both graphs show the same trend, which is that classification accuracy improves with an increasing percentage of tumor cells in the measurement volume. The conclusion is that the percentage of the measurement volume that is taken by either invasive carcinoma cells or DCIS cells is more important for the ability to correctly classify a measurement location than the specific tissue type.

Before generating the classification models, the data was reduced by deriving so-called spectral features for the measured spectra, which were calculated from wavelength areas that visually seemed different between pure fat tissue, pure connective tissue, and tissue with >90% tumor cells. Based on the reduced data, good classification model performance was obtained, without overfitting. However, the actual contribution of each of the single spectral features was not investigated. Further reduction of the features might be possible by detecting redundant spectral features. Another limitation of this study is that only the near-infrared wavelength range was included in the analysis. Potential differences between healthy and malignant that emerge in the visible wavelength range could not be evaluated.

The classification model performance of the classification models presented in this paper is very similar to previous studies in which also tumor could be discriminated with high accuracy from healthy tissue.[9, 19-21] However, this is the first study in which also mixed tissue locations are evaluated after a robust correlation with histopathology was made. The results here indicate that fat can be easily discriminated from all other tissue types. However, connective tissue and tumor tissue have very similar optical characteristics, probably because part of the tumor tissue is, in fact, connective tissue (see Figure 3c and Figure 3d). As healthy tissue can also consist of connective tissue, this explains why a portion of the healthy locations is classified as tumor. Other groups have also reported that connective/glandular tissue has increased water content[22], and this can hamper

discrimination of connective/glandular tissue from tumor tissue. Vice versa some of the locations with 1 to 75% tumor cells were classified as healthy. If besides the tumor cells the other part of the measurement volume is fat tissue, such a 'tumor' location can be easily interpreted as healthy by the classification algorithm as up to 99% can be fat tissue. The other factor that might influence these results is that although the best possible correlation was made between the DRS measurements and the histopathology there still might be some inaccuracies. For example, a different tissue type is present underneath a superficial layer of another tissue type, could result in an error of the estimates of the percentages of each tissue type in the measurement locations.

The majority of the publications using DRS for margin assessment evaluated the entire margin by using multi-channel probes and classifying an entire margin opposed to a single measurement.[8, 10] This is a different setting from the fiber-optic probe measurements used in this research in which spatial information is absent. When using an imaging set-up one measurement is an image of a margin, whereas in the case of a fiber-optic probe measurement one single spectrum is measured that is the sum of the diffuse reflectance of multiple tissue types that were present in the probed volume. Assessing the influence of measurements that are a mixture of tissue types is therefore important. The number of publications that assessed point measurements of measurement locations with a mixture of tissue types is limited. A study by Kennedy et al. evaluated the margin in a point based method but excluded the mixed tissue locations because the fractional composition of the measurement locations was not specified by histopathology.[12] A publication by Pappo did include locations with a combination of tumor tissue and healthy tissue.[23] In this paper, the MarginProbe is used similarly to the point measurements in our study. The sensitivity and specificity for locations with more than 75% tumor tissue were 1.00 and 0.87 respectively. However, including the mixed measurement locations sensitivity and specificity dropped both to 0.70.

Our ultimate goal is to use DRS technology for improving surgical outcome by providing margin assessment based on the optical measurements. In the study, we have shown that classification results are dependent on the percentage of tumor cells in the measurement volume. Although some locations without any tumor are classified as tumor, and some locations with tumor cells are classified as healthy, overall performance on the locations with a high percentage of tumor cells is good. Especially, considering the fact that some registration errors could also affect these results, as well as the previously described possibility of having different tissue type present in depth.

To assess the clinical value of the technology this outcome should be related to the volume of tumor cells that is present in the case of a positive resection margin. In other words, are positive resection margins comparable to the measurements in the group with 1-25% tumor cells, or the group with 76-100% tumor cells? In addition, it is important to recognize that the definition of a positive resection margin has shifted over the years and

is still under debate. [24] Thus, the actual impact of this technology in clinical practice can only be assessed by performing a large clinical study in which measurements are acquired at the true resection margin. This allows investigating the minimal amount of tumor that is clinically important in relation to the measurement volume. Such information can also be helpful in determining what the most suitable fiber distance between the collecting and emitting fiber should be, as the measurement volume can be altered by changing this distance.

The proposed study can also reveal how often at the resection margin locations are encountered that have a lot of connective tissue and might, therefore, be difficult to classify as healthy. Furthermore, there is a difference between what is considered a positive resection margin for DCIS[25], compared to the definition of a positive resection margin for invasive carcinoma[26]. Although both definitions have shifted towards a more liberal definition the definition of a positive resection margin for DCIS is more strict than the definition for invasive carcinoma. As discriminating the two tissue types with DRS seems impossible, this might have implications for what is detected by DRS as a positive resection margin. This should also be investigated in the previous suggested clinical trial.

In conclusion, in this study, we have shown that the spectral differences between healthy tissue and either invasive carcinoma or DCIS are the same. Furthermore, there is no significant difference in the measurements of invasive carcinoma with connective tissue and measurements of DCIS with connective tissue. With regard to the classification model development, it was concluded that using a model based on pure measurement locations (no connective tissue allowed in the definition of 'healthy' and no fat tissue allowed in the definition of 'tumor') showed good performance on both the test dataset (with locations labeled as 'healthy' or 'tumor') as well as the locations with a mixture of tissue types ('mix' dataset). Furthermore, with an increasing percentage of tumor cells in the measurement volume, a location with a mixture location was more often correctly classified as tumor. Importantly, the classification accuracy of locations with more than 75% tumor cells was excellent. Finally, the percentage of tumor cells is more important for correct classification of the measurements, than whether those locations contain invasive carcinoma or DCIS. These results thus show that detection of areas of DCIS at the resection margin is possible with DRS. Further research should be directed to measuring DRS at the resection margin to assess the ability of the technology to detect a positive resection margin.

REFERENCES

- [1] M. G. Valero, M. A. Mallory, K. Losk, M. Tukenmez, J. Hwang, K. Camuso, C. Bunnell, T. A. King, and G. Mehra, "Surgeon Variability and Factors Predicting for Reoperation Following Breast-Conserving Surgery," *Annals of Surgical Oncology*, vol. 25, pp. 2573-2578, 2018, [[10.1245/s10434-018-6526-2](#)].
- [2] S. L. Blair, K. Thompson, J. Rococco, V. Malcarne, P. D. Beitsch, and D. W. Ollila, "Attaining Negative Margins in Breast-Conservation Operations: Is There a Consensus among Breast Surgeons?," *Journal of the American College of Surgeons*, vol. 209, no. 5, pp. 608-613, 2009, [[10.1016/j.jamcollsurg.2009.07.026](#)].
- [3] R. E. Pataky, and C. R. Baliski, "Reoperation costs in attempted breast-conserving surgery: a decision analysis," *Current Oncology (Toronto, Ont.)*, vol. 23, no. 5, pp. 314-321, 2016, [[10.3747/co.23.2989](#)].
- [4] L. E. McCahill, R. M. Single, E. J. Aiello Bowles, H. S. Feigelson, T. A. James, T. Barney, J. M. Engel, and A. A. Onitilo, "Variability in Reexcision Following Breast Conservation Surgery," *Journal of the American Medical Association*, vol. 307, no. 5, pp. 467-475, 2012.
- [5] A. J. Isaacs, M. L. Gemignani, A. Pusic, and A. Sedrakyan, "Association of Breast Conservation Surgery for Cancer With 90-Day Reoperation Rates in NewYork State," *JAMA Surg*, 2016, [[10.1001/jamasurg.2015.5535](#)].
- [6] S. C. J. Bosma, F. Van der Leij, E. Van Werkhoven, H. Bartelink, J. Wesseling, S. Linn, E. J. Rutgers, M. J. Van de Vijver, and P. H. M. Elkhuzen, "Very low local recurrence rates after breast-conserving therapy: analysis of 8485 patients treated over a 28-year period," *Breast Cancer Research and Treatment*, vol. 156, pp. 391-400, 2016, [[10.1007/s10549-016-3732-0](#)].
- [7] L. G. Wilke, T. Czechura, C. Wang, B. Lapin, E. Liederbach, D. P. Winchester, and K. Yao, "Repeat Surgery After Breast Conservation for the Treatment of Stage 0 to II Breast Carcinoma A Report From the National Cancer Data Base, 2004-2010," *JAMA Surgery*, vol. 149, no. 12, pp. 1296-1305, 2014, [[10.1001/jamasurg.2014.926](#)].
- [8] A. M. Laughney, V. Krishnaswamy, E. J. Rizzo, M. C. Schwab, R. J. Barth Jr, D. J. Cuccia, B. Tromberg, K. D. Paulsen, B. W. Pogue, and W. A. Wells, "Spectral discrimination of breast pathologies *in situ* using spatial frequency domain imaging," *Breast Cancer Research*, vol. 15, pp. R61, 2013.
- [9] M. D. Keller, S. K. Majumder, M. C. Kelley, I. M. Meszoely, F. I. Boulos, G. M. Olivares, and A. Mahadevan-Jansen, "Autofluorescence and Diffuse Reflectance Spectroscopy and Spectral Imaging for Breast Surgical Margin Analysis," *Lasers in Surgery and Medicine*, vol. 42, pp. 15-23, 2010, [[10.1002/lsm.20865](#)].
- [10] J. Q. Brown, T. M. Bydlon, S. A. Kennedy, M. L. Caldwell, J. E. Gallagher, M. Junker, L. G. Wilke, W. T. Barry, J. Geradts, and N. Ramanujam, "Optical Spectral Surveillance of Breast Tissue Landscapes for Detection of Residual Disease in Breast Tumor Margins," *PloS One*, vol. 8, no. 7, pp. e69906, 2013, [[10.1371/journal.pone.0069906](#)].
- [11] B. S. Nichols, A. Llopis, G. M. Palmer, S. S. McCachren, O. Senlik, D. Miller, M. A. Brooke, N. M. Jokerst, J. Geradts, R. Greenup, and N. Ramanujam, "Miniature spectral imaging device for wide-field quantitative functional imaging of the morphological landscape of breast tumor margins," *Journal of Biomedical Optics*, vol. 22, no. 2, pp. 026007, 2017, [[10.1117/1.JBO.22.2.026007](#)].
- [12] S. Kennedy, J. Geradts, T. Bydlon, J. Q. Brown, J. Gallagher, M. Junker, W. Barry, N. Ramanujam, and L. Wilke, "Optical breast cancer margin assessment: an observational study of the effects of tissue heterogeneity on optical contrast," *Breast Cancer Research*, vol. 12, no. 6, pp. R91, 2010, [[10.1186/bcr2770](#)].
- [13] D. J. Evers, R. Nachabé, M. J. Vranken Peeters, J. A. Van der Hage, H. S. Oldenburg, E. J. Rutgers, G. W. Lucassen, B. H. W. Hendriks, J. Wesseling, and T. J. M. Ruers, "Diffuse reflectance

- spectroscopy: towards clinical application in breast cancer," *Breast Cancer Research and Treatment*, vol. 137, pp. 155-165, 2013, [10.1007/s10549-012-2350-8].
- [14] S. K. Majumder, M. D. Keller, F. I. Boulos, M. C. Kelley, and A. Mahadevan-Jansen, "Comparison of autofluorescence, diffuse reflectance, and Raman spectroscopy for breast tissue discrimination," *J Biomed Opt*, vol. 13, no. 5, pp. 054009, 2008, [10.1117/1.2975962].
 - [15] R. Nachabé, B. H. W. Hendriks, A. E. Desjardins, M. van der Voort, M. B. van der Mark, and H. J. C. M. Sterenborg, "Estimation of lipid and water concentrations in scattering media with diffuse optical spectroscopy from 900 to 1600 nm," *J Biomed Opt*, vol. 15, no. 3, pp. 037015, 2010, [10.1117/1.3454392].
 - [16] R. Nachabé, B. H. W. Hendriks, M. Van der Voort, A. E. Desjardins, and H. J. C. M. Sterenborg, "Estimation of biological chromophores using diffuse optical spectroscopy: benefit of extending the UV-VIS wavelength range to include 1000 to 1600 nm," *Optics Express*, vol. 18, no. 24, pp. 1432-1442, 2010.
 - [17] L. L. De Boer, E. Kho, J. Nijkamp, K. K. Van de Vijver, H. J. C. M. Sterenborg, L. C. Ter Beek, and T. J. M. Ruers, "A method for co-registration of optical measurements of breast tissue with histopathology: the importance of accounting for tissue deformations," *Journal of Biomedical Optics*, vol. In preparation, 2019.
 - [18] J. Makki, "Diversity of breast carcinoma: histological subtypes and clinical relevance," *Clinical Medicine Insights: Pathology*, vol. 8, pp. 23-31, 2015, [10.4137/CPath.S31563].
 - [19] L. L. De Boer, T. M. Bydlon, F. Van Duijnhoven, M.-J. T. F. D. Vranken Peeters, C. E. Loo, H. A. O. Winter-Warnars, J. Sanders, H. J. C. M. Sterenborg, and T. J. M. Ruers, "Towards the use of diffuse reflectance spectroscopy for real-time in vivo detection of breast cancer during surgery," *Journal of Translational Medicine*, vol. 16, no. 367, 2018, [10.1186/s12967-018-1747-5].
 - [20] L. L. De Boer, B. G. Molenkamp, T. M. Bydlon, B. H. W. Hendriks, J. Wesseling, H. J. C. M. Sterenborg, and T. J. M. Ruers, "Fat/Water ratios measured with diffuse reflectance spectroscopy to detect breast tumor boundaries," *Breast Cancer Research and Treatment*, vol. 152, no. 3, pp. 509-518, 2015, [10.1007/s10549-015-3487-z].
 - [21] Z. Volynskaya, A. S. Haka, K. L. Bechtel, M. Fitzmaurice, R. Shenk, N. Wang, J. Nazemi, R. R. Dasari, and M. S. Feld, "Diagnosing breast cancer using diffuse reflectance spectroscopy and intrinsic fluorescence spectroscopy," *Journal of Biomedical Optics*, vol. 13, no. 2, pp. 024012, 2008, [10.1117/1.2909672].
 - [22] P. Taroni, A. M. Paganoni, F. Ieva, A. Pifferi, G. Quarto, F. Abbate, E. Cassano, and R. Cubeddu, "Non-invasive optical estimate of tissue composition to differentiate malignant from benign breast lesions: A pilot study," *Scientific Reports*, vol. 7, pp. 40683, 2017, [10.1038/srep40683].
 - [23] I. Pappo, R. Spector, A. Schindel, S. Morgenstern, J. Sandbank, L. T. Leider, S. Schneebaum, S. Lelcuk, and T. Karni, "Diagnostic Performance of a Novel Device for Real-Time Margin Assessment in Lumpectomy Specimens," *Journal of Surgical Research*, vol. 160, no. 2, pp. 277-281, 2010, [10.1016/j.jss.2009.02.025].
 - [24] S. G. Brouwer de Koning, M. T. F. D. Vranken Peeters, K. Jóźwiak, P. A. Bhairosing, and T. J. M. Ruers, "Tumor Resection Margin Definitions in Breast-Conserving Surgery: Systematic Review and Meta-analysis of Current Literature," *Clinical Breast Cancer*, 2018.
 - [25] M. Morrow, K. J. Van Zee, L. J. Solin, N. Houssami, M. Chavez-MacGregor, J. R. Harris, J. Horton, S. Hwang, P. L. Johnson, M. L. Marinovich, S. J. Schnitt, I. Wapnir, and M. S. Moran, "Society of Surgical Oncology–American Society for Radiation Oncology - American Society of Clinical Oncology Consensus Guideline on Margins for Breast-Conserving Surgery with Whole Breast Irradiation in Ductal Carcinoma In Situ," *Practical Radiation Oncology*, vol. 6, no. 5, pp. 287-295, 2016, [10.1016/j.prro.2016.06.011].

- [26] M. S. Moran, S. J. Schnitt, A. E. Guiliano, J. R. Harris, S. A. Khan, J. Horton, S. Klimberg, M. Chavez-MacGregor, G. Freedman, N. Houssami, P. L. Johnson, and M. Morrow, "Society of Surgical Oncology - American Society for Radiation Oncology Consensus Guideline on Margins for Breast-Conserving Surgery With Whole-Breast Irradiation in Stages I and II Invasive Breast Cancer," *International Journal of Radiation Oncology Biology Physics*, vol. 88, no. 3, pp. 553-564, 2014.

Towards the use of diffuse reflectance spectroscopy for real-time *in vivo* detection of breast cancer during surgery

L.L. de Boer | T.M. Bydlon | F. van Duijnhoven | M.T.F.D. Vranken Peeters
C.E. Loo | G.A.O. Winter-Warnars | J. Sanders | H.J.C.M. Sterenborg
B.H.W. Hendriks | T.J.M. Ruers

Background: Breast cancer surgeons struggle with differentiating healthy tissue from cancer at the resection margin during surgery. We report on the feasibility of using Diffuse Reflectance Spectroscopy (DRS) for real-time *in vivo* tissue characterization.

Methods: Evaluating the feasibility of the technology requires a setting in which measurements, imaging and pathology have the best possible correlation. For this purpose an optical biopsy needle was used that had integrated optical fibers at the tip of the needle. This approach enabled the best possible correlation between optical measurement volume and tissue histology. With this optical biopsy needle, we acquired real-time DRS data of normal tissue and tumor tissue in 27 patients that underwent an ultrasound-guided breast biopsy procedure. Five additional patients were measured in continuous mode in which we obtained DRS measurements along the entire biopsy needle trajectory. We developed and compared three different Support Vector Machine (SVM) based classification models to classify the DRS measurements.

Results: With DRS malignant tissue could be discriminated from healthy tissue. The classification model that was based on eight selected wavelengths had the highest accuracy and Matthews Correlation Coefficient (MCC) of 0.93 and 0.87, respectively. In three patients that were measured in continuous mode and had malignant tissue in their biopsy specimen, a clear transition was seen in the classified DRS measurements going from healthy tissue to tumor tissue. This transition was not seen in the other two continuously measured patients that had benign tissue in their biopsy specimen.

Conclusions: It was concluded that DRS is feasible for integration in a surgical tool that could assist the breast surgeon in detecting positive resection margins during breast surgery.

Trail registration: NIH U.S. National Library of Medicine – [clinicaltrials.gov](https://clinicaltrials.gov/ct2/show/study/NCT01730365), NCT01730365. Registered: 10/04/2012 <https://clinicaltrials.gov/ct2/show/study/NCT01730365>

Keywords: Breast cancer surgery, intraoperative margin assessment, optical technology, real-time

INTRODUCTION

The current primary treatment of breast cancer includes a multimodal approach with a combination of surgery and radiotherapy, and depending on the subtype and the extent of the disease, systemic treatment. Optimal surgical treatment is achieved when all tumor tissue is resected, and thus histopathological evaluation of the resection specimen reveals no tumor-positive margins. However, resecting too much healthy tissue compromises cosmetic outcome. Since tumor-positive resection margins are associated with a higher recurrence rate, these patients require additional treatment with boost radiotherapy or re-excision surgery.[1, 2] Along with the impact of additional treatment on healthcare budgets, both boost radiotherapy and secondary surgery impair cosmetic outcomes[3], increase morbidity[4, 5] and can affect quality of life[6-8]. The surgeon is thus balancing between completely resecting the tumor and, sparing as much healthy tissue as possible.[9] Performing surgery while doing justice to both is difficult since visually recognizing tumor tissue on the surgical margin is extremely challenging. In addition, all currently available intra-operative margin assessment techniques have their own pitfalls, such as reduced sensitivity, the requirement of skilled personnel, are labor-intensive, and can be operator dependent.[10-12] Therefore, breast surgeons are in need of a robust margin assessment tool that can assist them real-time in defining the optimal resection plane to ensure clear margins, and will also help them to minimize the resected specimen volume.[13-18]

Diffuse reflectance spectroscopy (DRS), a light-based technology, has shown promising results for discriminating normal breast tissue from tumor tissue and may address the need for a margin assessment device.[19-22] The principle behind DRS is that light interacts with tissue through scattering and absorption. The absorption is related to the chemical composition of the tissue whereas the scattering is related to the subcellular morphology. The reflected light, detected after tissue interaction, has an altered spectrum compared to the incoming light. Thus, the diffuse reflectance spectrum represents aspects of the composition and subcellular morphology of the measured tissue. Ultimately, incorporating DRS technology in instruments, such as a surgical knife, could potentially provide the surgeon with additional information that reflects the histopathology of the tissue at the resection margin. The surgeon can use this information as guidance to determining the optimal resection place for excising a breast tumor.

There have been some publications on the feasibility of DRS for breast biopsy and surgery applications. [23-26] However, many of these publications struggled with correlating the exact tissue volume measured by DRS *in vivo* to the proper location in the histopathology sections processed post-operatively. This is an important factor as the 'gold standard' for the evaluation of surgical margins is microscopic assessment by a pathologist. Thus, a mismatch between the optical measurements and the histopathology hampers the development of robust classification algorithms and validation of the technology. For the

purpose of developing a reliable database of DRS measurements, we developed a special biopsy needle with embedded optical fibers. [27] This tool enables DRS measurements and subsequent biopsy of the same tissue volume as measured spectroscopically. Although it is tempting to see this study as an attempt to perform spectroscopy guided biopsy, this was explicitly not the purpose of this study. The robust dataset gathered in this setting can be used for developing classification models and validating DRS technology for tissue characterization. The paper describes how we developed and tested three predictive classification models, based on different types of input data, to accurately classify the DRS measurements. Furthermore, we investigated the feasibility of acquiring *in vivo* DRS measurements and classifying these. To this end, DRS measurements were performed continuously along an entire needle trajectory during ultrasound-guided breast biopsy procedures, and classified based on a classification model.

METHODS

Study design

Patients suspected of having breast cancer (after palpation, X-ray and US-imaging) that required diagnostic biopsies were asked to participate in this observational study that was approved by the Institutional Review Board of the Netherlands Cancer Institute. Written informed consent was obtained from all patients prior to the biopsy procedure. Patients with suspected sensitivity to light (e.g. patient who have had photodynamic therapy) were excluded, as well as patients that had received prior chemotherapy, endocrine therapy or radiation therapy recently (i.e. within 5 years). Patients were also excluded with breast implants and those that needed a stereotactic breast biopsy. In all patients, a biopsy was obtained after the last measurement in tumor tissue. The biopsy specimen was colored with pathology ink on the distal side to indicate which side had been in contact with the fibers during measurements. The biopsy specimens were used for diagnostic assessment and to confirm histopathology of the tumor measurement location by evaluating the first 2 mm of the side of the biopsy specimen that had been in contact with the optical fibers. Patients were only included when the pathology results of the lesion (including the other diagnostic biopsy specimens) indicated the lesion was an invasive breast tumor.

Spectrometers & optical biopsy needle

DRS measurements were obtained with a specially designed optical biopsy needle (Figure 1). This optical biopsy needle, which had integrated optical fibers, combined the ability to measure DRS spectra with biopsy functionality.[27] The 14G optical biopsy needle had one 100 μ m fiber for illumination and two 200 μ m fibers for collecting the light (Invivo); with a

Figure 1. Biopsy needle with integrated optical fibers. (A) In the initial phase, the tissue is in contact with the fibers. (B) When the release button is pressed the cutting mechanism extends forward while the fibers retract. The tissue can now enter the biopsy cavity. (C) When pressing the release button further the outer stylet will extend forward, thereby cutting the tissue in the biopsy cavity from its surrounding. (D) Photograph of the biopsy needle with an extended inner stylet, with the cutting mechanism protruded similar to situation B. (E) Example of an H&E stained section of the biopsy specimen. The side of the specimen that was not in contact with the fibers (in this case the left side) is marked with red pathology ink directly after retrieving the biopsy specimen from the cavity.

20 mm cavity for the tissue biopsy. The collecting fibers were placed next to each other, and the distance between the illuminating and collecting fibers was 1.36 mm, which resulted in a penetration depth of approximately 1-2 mm. The optical biopsy needle was attached to two spectrometers that resolved light in the visual wavelength range (DU420A-BRDD, Andor Technology) and the near-infrared wavelength range (DU492A-1.7, Andor Technology). After measuring, the spectra of the two spectrometers were stitched together, to form a continuous spectrum between 400 and 1600 nm[28]. In the full spectrum and selected wavelengths classification models the first 100 nm was removed since this wavelength range was highly affected by noise which may influence the machine learning algorithms that were used for the development of the models.

DRS point measurements

In 27 patients, measurements were acquired in a point-based manner, thus the measurement needle was first held still in normal tissue, and subsequently in the breast tumor where a biopsy was taken. At each measurement location, three DRS measurements (10 spectra per measurement) were acquired and averaged. Performing a single measurement took approximately 10 seconds. In this measurement time a total of 30 high-quality DRS spectra were obtained over the full wavelength range which is necessary for building a database that can be used for classification model development. During the time the DRS measurements were obtained an US-image was made with the needle tip in view. These US-images were evaluated by a radiologist to confirm correct positioning of the needle in either normal breast tissue or tumor tissue.

DRS continuous measurements

In five additional patients, DRS measurements were obtained in a continuous mode. Here the measurements were obtained along the entire needle trajectory starting in the normal tissue, continuing through the transition zone of normal-tumor, and ending in the tumor. To enable real-time acquisition of the data adjustments were made to the settings to increase the acquisition rate. The framerate in continuous mode was approximately one spectrum per second. The integration time was set to 0.35s for all measurements along the needle trajectory. Prior to measuring the clocks of the US device and the laptop controlling the DRS set-up were synchronized thus allowing the US-images to be registered to the DRS measurements. At one location at a distance from the tumor, and at the final measurement location (also the location of the biopsy specimen) the needle was kept still to ensure sufficient data of both healthy tissue and tumor tissue (similar to the point measurements). At these locations, 10 DRS measurements were acquired, as well as an US-image.

Pre-processing point measurements

At each point measurement location, three spectra were obtained which were averaged to calculate a mean spectrum of each measurement location. Subsequently, all spectra were normalized with the standard normal variate (SNV) method.[29] Outlier detection was performed to ensure that the classification models were developed with reliable data.[30] A cut-off of 3 times the standard deviation was chosen as a threshold.

Development of classification models

For development of the first classification models only the point-based measurements were used as for these measurement US images and histopathology provided information on the nature of the tissue in front of the needle during a measurement. For the continuous data, this information was not available for each measurement in the trajectory.

To ensure a balanced dataset, only patients which had both the normal and tumor measurement available after outlier detection were used. All classification models were constructed with perClass (Academic version 5.0, PR Sys design) in Matlab (2015a, the MathWorks). Figure 2 depicts a schematic flowchart of the approach to building the classification models.

Fit parameter classification models

The input for the first classification model was optical fit parameter data. To calculate these fit parameters from the measured DRS spectra, the measurements were quantified using an analytical fit model based on diffusion theory. This can be considered as a feature reduction method in which the measured spectra is translated into chemically or physiologically meaningful parameters.[31] In order to do so, the fit model required the absorption spectra of substances present in tissue, including: blood, fat, water, β -carotene, collagen, and

Figure 2. Schematic overview of training and testing of the classification model. In the inner loop the SVM is optimized (5 fold cross-validation), this optimized model is subsequently used with the test dataset. The outer loop is performed 100 times. The sensitivity, specificity, and accuracy are averaged over all iterations.

bilirubin. The fit then optimized the parameters in such a way that the modelled spectrum matched the measured spectrum. The optical fit parameters generated by the fit model were: amount of blood (%), oxygen saturation (StO₂), total amount of fat plus water, fraction of fat, scattering at 800 nm, α and b (from the formula describing the reduced scattering, $\mu'_s = \alpha\lambda^{-b}$) amount of bilirubin, fraction of Mie scattering (in relation to the total scattering), amount of β -carotene, and amount of collagen. The amount of water was calculated from the optical parameters describing the total amount of fat and water and the fat fraction. The amount of water together with the amount of fat allowed deriving the ratio between fat and water for each measurement location.[21, 28]

Different combinations of fit parameters were used as input for this classification model. To limit the required computational effort, these combinations were formed by combining fit parameters that had shown the ability to discriminate between normal and tumor tissue previously, i.e. blood, StO₂, scattering at 800 nm, fraction Mie scattering, β -carotene, collagen and, the ratio between fat and water (F/W-ratio). With these seven fit parameters combinations (c) were made that consisted of either one or multiple fit parameters with

one combination including all seven fit parameters ($c = 127$, $c = \sum_{k=1}^7 \frac{n!}{((n-k)!k!})$) with

n the number of fit parameters to choose from, and k the number of elements in the combination). As the F/W-ratio previously proved to be an excellent discriminator,[21, 32] only fit parameter combinations that included the F/W-ratio were used as input for the first classification model. The final set of combinations consisted of 64 possibilities.

Full spectrum classification model

The input for the second classification model was the full wavelength spectrum without any feature reduction. Thus, each wavelength between 500 and 1600 nm (i.e. the full spectrum) was used as input to the model, resulting in 1100 features for the classification model.

Selected wavelengths classification model

The input for the third classification model consisted of a limited number of wavelengths. The selection of these wavelengths was based on the results of a two-sided Wilcoxon rank sum test ($\alpha = 0.05$). All normal and all tumor measurements were used in the test. For each wavelength, the Wilcoxon rank sum test assesses whether two samples of observations (in this case the normal measurements and the tumor measurements) are from the same distribution. This statistical test was used to identify wavelength regions that were significantly different between the normal and tumor spectra (mean p-value below 0.05). In each of these regions, the wavelength with the lowest p-value was selected as a wavelength for the selected wavelengths model.

Classifier

A linear support vector machine (SVM) formed the classifier in the classification model. [33, 34] This machine learning technique constructed an optimal separating linear hyperplane between two classes in a higher dimensional space by creating the biggest margin between measurements of two classes.

Bootstrap sampling and cross-validation

To avoid selection bias, a bootstrapping technique was used to randomly select 17 subsets of (not necessarily different) patients as training data for model development. On average, both the normal and tumor measurements of 11.9 (+/- 1.3) unique patients were used as training data in each of the 100 iterations. The remaining patients that were not selected for training were used for testing the model. Subsequently the SVM was optimized with a 5-fold cross-validation of the training data to find the optimum for the regularization parameter C . The unseen test data was then classified with the optimized SVM model. This process was performed a hundred times in order to test the generalizability of the classification model. Measurements provided to the classification model were classified as either 'normal' or 'tumor'. With this binary output the sensitivity, specificity, and accuracy in each bootstrap iteration were calculated for discriminating normal measurements from tumor measurements. By averaging these model performance parameters over the 100 iterations the mean sensitivity, mean specificity, mean accuracy, and mean Matthews Correlation Coefficient (MCC) were determined, and these were used to compare the performance of different classification models. Besides the binary output, the classification models could also generate the probability of a measurement being either 'normal' or 'tumor'.

Classification of continuous data

The continuous data was preprocessed, by normalizing it using the SNV method. Spectra were not averaged and no outlier detection was performed. Along the needle trajectory, US-images were captured of the biopsy needle while it was positioned in healthy tissue during the first measurement, and when the needle was at the final measurement location, targeting to be in the tumor. For the classification of the continuous data, all point measurements were used to build another three classification models with the different input data (i.e. fit parameter, full spectrum, and selected wavelengths) (Figure 3). No measurements of the continuously acquired data were used in the development of any of the models, they were only provided to the classification model to be classified. Therefore these needle trajectory measurements were not labeled based on US-imaging or pathology.

Figure 3. Schematic overview of the classification model development with point measurements to classify continuous measurements. Again, either the fit parameters data, full spectrum data or, selected wavelengths data is used as input for the classification model development.

Table 1. Patient characteristics of point measurements dataset and continuous measurements dataset.

Patient characteristics	Point measurements (n = 21)	Continuous measurements (n = 5)
Mean age (std)	53.4 year (12.1)	60.6 year (10.2)
Menopausal status		
<i>Premenopausal</i>	8	1
<i>Perimenopausal</i>	2	0
<i>Postmenopausal</i>	10	4
<i>Unknown</i>	1	0
Mean tumor size (US imaging) (std)	25.5 mm (11.1)	37.2 mm (25.1)
Cancer type*		
<i>Invasive ductal carcinoma</i>	19	3
<i>Invasive lobular carcinoma</i>	2	1
<i>Mucinous adenocarcinoma</i>	0	1

*The histopathology was based on the histopathology of all biopsy specimens taken in that patient

RESULTS

In total 32 patients were measured and had unambiguous pathology results. Of these patients, 27 formed the point measurement dataset and 5 the continuous dataset. From the 27 point-based measurement patients, in one patient the biopsy specimen was absent and two patients had biopsy specimens that were clearly damaged during processing of the tissue. In these three cases careful evaluation of the US images by a radiologist revealed that the needle tip was certainly placed a few millimeters inside the tumor and therefore these patients were still included in the analysis. Four patients were excluded from the analysis because 1) the side of the biopsy specimen that had been in contact with the fibers during the measurement consisted of healthy tissue over the extent of a few millimeters and 2) according to the radiologist the needle was moved between the measurement and biopsy. No patients were excluded because the tumor was too close to the skin, thus prohibiting the acquisition of measurements of healthy tissue.

In the procedure of outlier detection, two measurement locations were detected. An explanation for the first outlier might be that the needle tip was in a pool of blood during the normal measurements, which was confirmed by the high blood content according to the fit parameters. As for the second outlier, the histopathology of this measurement location showed benign tissue in the biopsy specimen. The patients to which these locations belonged were also excluded to ensure a balanced dataset.

Thus, in total 6 patients were excluded from the point measurement dataset. The remaining 21 patients, in whom point measurements were obtained, were included for further analysis. The patient characteristics of both patient datasets seem similar and are summarized in Table 1.

Table 2. Performance (mean accuracy, sensitivity, specificity, and MCC with standard deviations) of classification models.

Type of data used as input for the model	Mean accuracy (std)	Mean sensitivity (std)	Mean specificity (std)	Mean MCC (std)
<i>Fit parameters</i>				
<i>F/W-ratio & collagen</i>	0.85 (0.16)	0.72 (0.33)	0.99 (0.03)	0.74 (0.30)
<i>F/W-ratio</i>	0.85 (0.16)	0.71 (0.34)	0.99 (0.04)	0.72 (0.31)
<i>Full spectrum</i>	0.92 (0.06)	0.94 (0.10)	0.89 (0.11)	0.84 (0.12)
<i>Selected wavelengths</i>	0.93 (0.06)	0.95 (0.07)	0.91 (0.14)	0.87 (0.11)

Classification models based on fit parameters

In total 64 classification models that were based on combinations of fit parameters were built. The two fit parameter combinations that generated the two classification models with the highest accuracies are listed in Table 2. The fit parameter combination of F/W-ratio and collagen was the combination that resulted in the classification model with the best performance, with a mean accuracy, sensitivity, specificity and MCC of 0.85 (0.16), 0.72 (0.33), 0.99 (0.03), and 0.74 (0.30) respectively. The second best performing classification model was based on the F/W-ratio alone, which had a slightly lower sensitivity compared with the combination of the F/W-ratio and collagen.

Classification model based on full spectrum

The mean accuracy, sensitivity, specificity, and MCC of the model based on the full spectrum were 0.92 (0.06), 0.94 (0.10), 0.89 (0.11), and 0.84 (0.12) respectively (Table 2). Compared to the fit parameter model, the full spectrum model had better accuracy, sensitivity, and MCC, whereas the specificity of the fit parameter model was better. This indicates that the full spectrum classification model is useful for detecting all tumor tissue at the cost of classifying some normal tissue as tumor. With the fit parameter classification model, less normal tissue will be incorrectly classified as tumor, but also, less tumor tissue will be detected.

Classification model based on selected wavelengths

A third classification model was developed using a selection of wavelengths that were significantly different between normal and tumor spectra according to the Wilcoxon rank sum test ($\alpha = 0.05$). Figure 4 shows the results of the Wilcoxon rank sum test. The grey parts of the graph represent wavelength areas in which the p-value was lower than 0.05. From these areas the wavelength with the lowest p-value was selected for the selected wavelengths model (vertical dashed lines). The selected wavelengths were: 501 nm, 916 nm, 973 nm, 1145 nm, 1211 nm, 1371 nm, 1424 nm, and 1597 nm.

The classification model based on these wavelengths was tested similarly to the models based on the fit parameters and the full spectrum. The mean accuracy, sensitivity, specificity, and MCC of this model was 0.93 (0.06), 0.95 (0.07), 0.91 (0.14), and 0.87 (0.11) respectively

Figure 4. P-values of Wilcoxon rank sum test. Results of a two-sided Wilcoxon rank sum test ($\alpha = 0.05$) for each wavelength between normal and tumor measurements. The grey wavelength ranges indicate that over these wavelengths there is a significant difference between normal and tumor. The vertical dashed lines represent the wavelengths with the lowest p-value in each grey area. The eight selected wavelengths were: 501 nm, 916 nm, 973 nm, 1145 nm, 1211 nm, 1371 nm, 1424 nm, and 1597 nm.

(Table 2). Compared to the fit parameter model this model has improved mean sensitivity, but reduced specificity. Despite the decrease in mean specificity, the mean accuracy and MCC of the selected wavelengths model is higher in comparison to the fit parameter model. The classification model after feature selection also outperforms the full spectrum model as the mean accuracy, sensitivity and specificity are slightly higher. The MCC of the classification model based on a selection of wavelengths was the highest with the lowest standard deviation compared to the other models.

To ensure the improvement of model performance was related to the actual wavelengths in the set of selected wavelengths, the model performance was compared to the model performance of a subset of wavelengths that had the maximum p-value from wavelength ranges with p-values of > 0.5 . The selected wavelengths for this model were: 602 nm, 681 nm, 951 nm, 1018 nm, 1095 nm, 1174 nm, 1230 nm, 1397 nm, and 1503 nm. As for the model with these eight selected wavelengths the mean accuracy, mean sensitivity, and mean specificity was 0.49 (0.12), 0.48 (0.25), and 0.50 (0.27) respectively. The MCC also displayed weak performance of this model with a mean value of 0.61 and a standard deviation of 0.32.

Classifying continuous data

The data from the five patients that were measured in continuous mode were tested on the three classification models (fit parameters, full spectrum, and selected wavelengths). To make these classification models consistent with the previous models, the fit parameter

Figure 5. Classification of continuous data. In the left part of the image, US images taken along the needle trajectory at 'normal' and 'tumor'. The middle of the image includes the outcomes of the classification algorithms, where the x-axis is the measurement number ($\neq$ distance) and the y-axis is the probability of a measurement being normal (> 0.5) or tumor (< 0.5). The green and red arrows indicate the locations where the needle was kept still. The histopathology of the part of the biopsy specimen that was in contact with the needle is displayed on the right side.

model was developed with the same fit parameters as in the previous model (F/W-ratio & Collagen) and similarly for the selected wavelengths model the same eight wavelengths (501 nm, 916 nm, 973 nm, 1145 nm, 1211 nm, 1371 nm, 1424 nm, and 1597 nm) were used. The results of the classification of the continuous data are represented in Figure 5.

In the left part of the figure for each patient the US images at two locations along the needle trajectory ('normal' and 'tumor') are shown. The histopathology of the part of the biopsy specimen that was in touch with the fibers at the last measurement location is displayed in the right side of the figure. The black bars represent a distance of 1 mm in the histopathology image. The graphs in the center of the figure show the output of the classification models in terms of probabilities for each measurement. A probability of > 0.5 indicates a measurement is classified as 'normal breast tissue' by the model, whereas a probability of < 0.5 implies 'tumor'. The x-axis represents the measurements in time, not in distance. In patient 1, there are some measurements missing because these were accidentally not saved during the procedure.

The histopathologic evaluation by the pathologist revealed that there was tumor ('mucinous adenocarcinoma', or 'invasive ductal carcinoma') in the biopsy specimen of patient 1, 2 and 3. In the case of patient 4 and 5, the side of the biopsy specimens touching the fibers did not contain malignant tissue according to the pathologist. In all three patients that had invasive carcinoma in their biopsy specimen (patient 1, patient 2, and patient 3) there is a distinct decrease in probability visible in the classified DRS measurements taken along the trajectory from healthy tissue to tumor tissue. Furthermore, the first measurements of patient 1 and patient 2 are classified as normal tissue (probabilities close to one), and the final measurements are classified as tumor (probabilities close to zero) by all three models. In both these patients, the probability of the final measurement of the trajectory calculated by the full spectrum model and the selected wavelengths model are closer to zero than the output of the fit parameter model, indicating more certainty of the classification. Along the trajectory of patient 1, there is one outlier (measurement #10), which displays a distinct decrease in probability for all classification models. The measurements of patient 3 are classified as normal at the beginning of the trajectory and as the needle progressed to the tumor, the probabilities, clearly and consistently over all models, decreased. However, at the end of the trajectory, none of the three models classified the final measurements as tumor, whereas according to the biopsy specimen the needle was placed in tumor tissue.

Two patients did not have tumor tissue in the first 2 mm of the biopsy specimen that was in contact with the optical fibers (patient 4 and 5). In both cases, the outcomes of the classification models classified all measurements in the needle trajectory as normal tissue. As for patient 4, the probability of the fit parameter model does show a decrease that was not seen in the output of the other two classification models. The transition from normal tissue to tumor, seen in the patients with a malignancy (patient 1, patient 2, and patient 3), is not consistently present in the classification output of all three models in patient 4 and patient 5, whom had no malignancy in their biopsy specimens

DISCUSSION

A large amount of evidence from *in vivo* and *ex vivo* studies around the world has proven that DRS can be a highly powerful tool for clinical use to discriminate tissue types. However, the technology has not been integrated with a surgical tool for real-time margin assessment. With the goal of moving towards a real-time classification tool for surgical margin assessment during breast surgery, we aimed at developing a classification model to accurately predict the type of tissue in front of the DRS tool, as well as showing the feasibility of real-time use of this technology. To reach these goals a custom made optical biopsy needle was used enabling DRS measurements and histopathology to be assessed on the same tissue volume which is inevitable for developing a robust classification model. It should be noted that this research was conducted with this set-up as a step towards developing a surgical tool that can guide the surgeon, rather than to improve the yield of breast biopsy procedures. This research differs from previous work as it measures DRS over a broader wavelength range extending into the near-infrared wavelengths. Furthermore, to the authors' knowledge, it is the first publication to test the feasibility of *in vivo* continuous DRS data acquisition, with a frame rate of approximately one measurement per second, which is more similar to how data will be acquired in the surgical setting.

We first used the point-based measurements to determine the performance of the classification models that were based on different input data. We found that if the fit parameter data was used as input, the combination of the F/W-ratio and collagen resulted in a model with the highest accuracy and MCC (0.85 and 0.74) compared to other combinations of fit parameters. Besides the fit parameter model, two other models were developed using the full spectrum of wavelengths or a selection of wavelengths as input. The full spectrum model had a better sensitivity compared to the fit parameter model (0.94 versus 0.72), whereas the fit parameter model had a higher specificity (0.99 versus 0.89), suggesting that the full spectrum model is more suitable for detecting tumor tissue, while the fit parameter model has less misclassifications of normal tissue. Although not statistically tested, the classification model based on a subset of selected wavelengths seems to outperform the other two models with the highest accuracy and MCC (0.93 and 0.87).

We developed three classification models (fit parameters, full spectrum, and selected wavelengths) based on all available point-measurements specifically to classify the continuous measurements. Importantly, none of the continuous measurements were used for the development of a classification model, they only served as test data to be classified by the classification model (Figure 3). In all five patients, the first measurements were classified as normal tissue by the classification models, this is expected considering the fact that the needle trajectory starts in normal tissue going towards the suspected tumor tissue. In patients 1, 2, 4, and 5, the classified DRS measurements of the final measurement locations are in agreement with the pathological outcome. In patient 1, however, there is a measurement

(#10) in the trajectory that is classified as 'tumor'; this appears to be a 'false positive' since the distance from this location to the lesion is quite far. In the surgical pathology report following lumpectomy for this patient it was noted that there was a focus of DCIS 1.5 mm from the tumor. It could be possible that this smaller lesion was in the trajectory of the needle, explaining the decrease in probability.

The output by the classification models of the final measurement location was not in accordance with histopathologic evaluation in one patient (patient 3). In this case, the outputs of the classification models show a decrease in probability, but never reach the threshold, and the final measurement location is classified as normal by all classification models. It could be possible that the histopathology evaluation of this patient has been compromised as the removal of the biopsy specimen from the needle cavity was difficult and, since the specimen was fragmented, part of it might have been left behind. Overall, in four out of five continuous mode patients, the classification models were able to discriminate tumor tissue from normal tissue, although the fit parameter model was least convincing with probabilities closer to 0.5. In three out of five patients malignant tissue was present in the biopsy specimen and in these patients a decrease in probability of the classified measurements is also seen along the needle trajectory. This decrease is absent in the other two patients that had healthy tissue in their biopsy specimens. The fact that a decrease can be detected is an important result when considering DRS as a margin assessment tool. In a way, this trajectory can be seen as a line that at some point crosses the optimal resection plane that is perpendicular to this line. Thus, being able to detect the upcoming tumor could provide the surgeon with viable information for guidance.

A limitation of the continuous measurements is that a biopsy was only available from the final measurement location of the presumed tumor area while no histopathology was taken along the needle trajectory. However, all breast tumors were clearly visible on the US images and could confirm that the needle was positioned in normal tissue from the start of the measurements. Nevertheless, some uncertainty will still exist on the precise location of the tissue border where normal tissue ends and tumor tissue starts. It should furthermore be noted that the x-axes of the graphs in Figure 5 are related to time opposed to distance, and thus these graphs therefore, display a change over time. Since the needle was not moved with constant speed along the needle trajectory, it was not possible to display the measurements as a function of the distance.

In literature many different methodologies are used for classifying reflectance spectra of breast tissue, for example logistic regression[20, 22], classification and regression trees[35, 36], artificial neural network[24], hierarchical cluster analysis[24], k-nearest neighbor[37], linear discriminant analysis[35], and support vector machines [35, 38-41]. In this study, a linear SVM classifier was chosen, as this classifier is relatively insensitive to overfitting.[42-46] Possibly a polynomial kernel SVM would have provided better results, however, because the number of patients in the study is limited for machine learning, a linear, less complicated,

classifier was chosen. For similar reasons the bootstrap sampling was preferred to leave-one-patient-out cross-validation, even though bootstrap methods can have the tendency to be pessimistic. By extending the number of measurements the accuracy of the classification of the DRS measurements will likely improve, and more sophisticated machine learning algorithms can be used.

The SVM classification model was developed with the input of either fit parameters (fit parameter model), or all wavelengths in the spectrum (full spectrum model), or some selected wavelengths (selected wavelengths model). A previous publication comparing the classification accuracy for discrimination of breast cancer of an SVM model based on physical parameter data (equivalent to fit parameters model), with the accuracy of an SVM classification model based on empirical data (equivalent to the selected wavelengths model) reported similar results to this study.[38] The main advantage of using the fit parameters is that these parameters can provide insight into the physical and structural features that contribute to discrimination.[38, 43] However, if fit parameters cannot be estimated accurately, for example, because the tissue has a layered structure, accuracy of classification models based on fit parameters will be lower.[43] This can explain why the accuracy of the fit parameter model was lower compared to the other two models.

We found that the performance of the selected wavelengths model is slightly better compared to the performance of the full spectrum model, which is not surprising since removing redundant wavelengths is often reported to be beneficial for classification performance.[42] There are many ways to select or reduce features, such as partial least squares [38], maximum representation and discrimination feature[20], or principal component analysis (PCA)[24, 38-40]. In this study, a Wilcoxon rank sum test is performed to find wavelengths that are significantly different between normal and tumor tissue. This method has been described for feature selection in previous publications, although in many cases this statistical test was preceded by PCA.[40, 46] The advantage of the Wilcoxon rank sum test is that the selection of wavelengths is based on true spectral differences between tissue types. The disadvantage is that wavelengths that are not discriminated according to this statistical test are excluded in the model development, although they could have discriminative power in combination with each other.

The wavelengths that were eventually selected are located in wavelength areas that are related to the absorption of light by fat, water, and to a lesser extent, blood. This result is in line with previous publications by others and our own group in which these substances also contributed to discriminating healthy tissue from tumor tissue.[32, 38]

The DRS measurements in this study were obtained during breast biopsy procedures to provide a correlated dataset (DRS data and histopathology) and test the feasibility of real-time data acquisition. This setting is obviously different than the surgical setting where the goal is to classify DRS measurements of the resection margin. In that situation, the influence of air exposure will likely affect the visual wavelength range due to differences in oxygenated and de-oxygenated blood which have different optical absorption characteristics; whereas

the near-infrared wavelength range, with predominant absorption characteristics from fat and water, will likely be less affected by the surgical setting. Furthermore, the resection margin can also be influenced by cauterization which was absent in the measurements obtained in the biopsy setting, or extravascular blood on the resection surface. With regard to the results in this study, this might imply that in the selected wavelengths model the first wavelength that was selected (501 nm) cannot be used. The accuracy of DRS measurements for the detection of tumor intra-operatively at the resection margin, should be investigated in a study in which DRS measurements (including also the near-infrared wavelengths) are acquired at the true resection margins, preferably in the surgical workflow.

Using DRS as a clinical margin assessment tool also requires that measurements can be acquired and classified in real-time. In the continuous dataset, each spectrum required 0.35s to be acquired. If necessary this acquisition time could be decreased by a factor of 4 by increasing the fiber diameter from 200 μ m to 400 μ m. As for the classification, this was not performed real-time in this study. However, once a classification model is defined the tissue can be classified in real-time as this requires little computational power.

Another important factor to consider is the influence of ambient light that might be different in the surgical resection field compared to the setup during a biopsy. Part of this challenge is overcome by the fact that a fiber is used which has to be in contact with the tissue instead of a non-contact configuration. Therefore, only light that falls in the acceptance angle of the fiber will be recorded by the spectrometer. However, in clinical practice, this might mean that very bright light sources that are in close proximity of the fiber-optic probe have to be dimmed to ensure interference with the DRS measurements is prevented.

In this paper, we demonstrate the feasibility that DRS measurements can be acquired real-time and that a predictive classification model can be built to classify the measurements as normal or tumor tissue. The classification model based on a selection of wavelengths discriminated normal tissue from tumor tissue with the highest accuracy and MCC of 0.93 and 0.87, respectively. This performance may be sufficient for the application of detecting positive resection margins during breast-conserving surgery. The needle trajectory measurements show that DRS measurements can be acquired real-time and that these measurements can be classified accurately. Furthermore, the transition from normal tissue to tumor tissue was seen in the continuous DRS measurements.

Our current results indicate that integration of DRS in a surgical tool or knife could be useful for characterizing breast tissue *in vivo* and aiding surgeons in detecting positive resection margins during surgery. The next step is to investigate the feasibility of real-time DRS acquisition and classification on resection margins and investigate the impact of a DRS guided tool on surgical outcomes.

ACKNOWLEDGEMENTS

We would like to thank Jarich Spliethoff and Niels Langhout for their input and help during the measurement procedures. We would also like to acknowledge Christian Reich, Manfred Mueller, Vishnu Pully, Walter Bierhoff, Axel Winkel, and Arnold van Keersop for their contribution in the development of the biopsy needle and fit model software. Assistance provided by the other radiologists and radiology department personnel was greatly appreciated, as well as the assistance of the surgeons and nurse practitioners with patient inclusion. In particular we would like to express many special thanks to all involved patients.

REFERENCES

- [1] N. Houssami, P. Macaskill, M. L. Marinovich, J. M. Dixon, L. Irwig, M. E. Brennan, and L. J. Solin, "Meta-analysis of the impact of surgical margins on local recurrence in women with early-stage invasive breast cancer treated with breast-conserving therapy," *European Journal of Cancer*, vol. 46, pp. 3219-3232, 2010, [[10.1016/j.ejca.2010.07.043](https://doi.org/10.1016/j.ejca.2010.07.043)].
- [2] M. C. Smitt, K. Nowels, R. W. Carlson, and S. S. Jeffrey, "Predictors of reexcision findings and recurrence after breast conservation," *International Journal of Radiation Oncology, Biology, Physics*, vol. 57, no. 4, pp. 979-985, 2003, [[10.1016/s0360-3016\(03\)00740-5](https://doi.org/10.1016/s0360-3016(03)00740-5)].
- [3] S. Collette, L. Collette, T. Budiharto, J.-C. Horiot, P. M. Poortmans, H. Struikmans, W. Van den Bogaert, A. Fourquet, J. J. Jager, W. Hoogenraad, R.-P. Mueller, J. Kurtz, D. A. L. Morgan, J.-B. Dubois, E. Salamon, R. Mirimanoff, M. Bolla, M. Van der Hulst, C. C. Wárlám-Rodenhuis, and H. Bartelink, "Predictors of the risk of fibrosis at 10 years after breast conserving therapy for early breast cancer - A study based on the EORTC trial 22881-10882 'boost versus no boost'," *European Journal of Cancer*, vol. 44, pp. 2587-2599, 2008, [[10.1016/j.ejca.2008.07.032](https://doi.org/10.1016/j.ejca.2008.07.032)].
- [4] A. Bodilsen, K. Bjerre, B. V. Offersen, P. Vahl, N. Amby, J. M. Dixon, B. Ejlersen, J. Overgaard, and P. Christiansen, "Importance of Margin Width in Breast-Conserving Treatment of Early Breast Cancer," *Journal of Surgical Oncology*, vol. 113, pp. 609-615, 2016, [[10.1002/jso.24224](https://doi.org/10.1002/jso.24224)].
- [5] M. A. Olsen, K. B. Nickel, J. A. Margenthaler, A. E. Wallace, D. Mines, J. P. Miller, V. J. Fraser, and D. K. Warren, "Increased Risk of Surgical Site Infection Among Breast-Conserving Surgery Re-excisions," *Annals of Surgical Oncology*, vol. 22, pp. 2003-2009, 2015, [[10.1245/s10434-014-4200-x](https://doi.org/10.1245/s10434-014-4200-x)].
- [6] J. H. Volders, V. L. Negenborn, M. H. Haloua, N. M. A. Krekel, K. Jóźwiak, S. Meijer, and P. M. Van den Tol, "Cosmetic outcome and quality of life are inextricably linked in breast-conserving therapy," *Journal of Surgical Oncology*, vol. 115, pp. 941-948, 2016, [doi.org/10.1002/jso.24615].
- [7] E. Hau, L. Browne, A. Capp, G. P. Delaney, C. Fox, J. H. Kearsley, E. Millar, E. H. Nasser, G. Papadatos, and P. H. Graham, "The impact of breast cosmetic and functional outcomes on quality of life: long-term results from the St. George and Wollongong randomized breast boost trial," *Breast Cancer Research and Treatment*, vol. 139, pp. 115-123, 2013, [[10.1007/s10549-013-2508-z](https://doi.org/10.1007/s10549-013-2508-z)].
- [8] J. F. Waljee, E. S. Hu, P. A. Ubel, D. M. Smith, L. A. Newman, and A. K. Alderman, "Effect of Esthetic Outcome After Breast-Conserving Surgery on Psychosocial Functioning and Quality of Life," *Journal of Clinical Oncology*, vol. 26, pp. 3331-3337, 2008, [[10.1200/JCO.2007.13.1375](https://doi.org/10.1200/JCO.2007.13.1375)].
- [9] F. MacNeill, and A. Karakatsanis, "Over surgery in breast cancer," *The Breast*, vol. 31, pp. 284-289, 2017, [[10.1016/j.breast.2016.10.023](https://doi.org/10.1016/j.breast.2016.10.023)].
- [10] F. A. Angarita, A. Nadler, S. Zerhouni, and J. Escallon, "Perioperative measures to optimize margin clearance in breast conserving surgery," *Surgical Oncology*, vol. 23, pp. 81-91, 2014, [[10.1016/j.suronc.2014.03.002](https://doi.org/10.1016/j.suronc.2014.03.002)].
- [11] K. Butler-Henderson, A. H. Lee, R. I. Price, and K. Waring, "Intraoperative assessment of margins in breast conserving therapy: A systematic review," *The Breast*, vol. 23, pp. 112-119, 2014, [[10.1016/j.breast.2014.01.002](https://doi.org/10.1016/j.breast.2014.01.002)].
- [12] J. J. Keating, C. Fisher, R. Batiste, and S. Singhal, "Advances in Intraoperative Margin Assessment for Breast Cancer," *Current Surgery Reports*, vol. 4, no. 15, 2016, [[10.1007/s40137-016-0136-3](https://doi.org/10.1007/s40137-016-0136-3)].
- [13] C. M. O'Kelly Priddy, V. A. Forte, and J. E. Lang, "The Importance of Surgical Margins in Breast Cancer," *Journal of Surgical Oncology*, vol. 113, pp. 256-263, 2016, [[10.1002/jso.24047](https://doi.org/10.1002/jso.24047)].
- [14] E. R. St John, R. Al-Khudairi, H. Ashrafiyan, T. Athanasiou, Z. Takats, D. J. Hadjiminias, A. Darzi, and D. R. Leff, "Diagnostic Accuracy of Intraoperative Techniques for Margin Assessment in Breast Cancer Surgery," *Annals of Surgery*, vol. 265, no. 2, pp. 300-310, 2017, [[10.1097/SLA.0000000000001897](https://doi.org/10.1097/SLA.0000000000001897)].
- [15] J. Heil, K. Breitkreuz, M. Golatta, E. Czink, J. Dahlkamp, J. Rom, F. Schuetz, M. Blumenstein, G. Rauch, and C. Sohn, "Do Reexcisions Impair Aesthetic Outcome in Breast Conservation

- Surgery? Exploratory Analysis of a Prospective Cohort Study," *Annals of Surgical Oncology*, vol. 19, pp. 541-547, 2012, [[10.1245/s10434-011-1947-1](#)].
- [16] A. M. Zysk, K. Chen, E. Gabrielson, L. Tafta, E. A. May Gonzalez, J. K. Canner, E. B. Schneider, A. J. Cittadine, P. S. Carney, S. A. Boppart, K. Tsuchiya, K. Sawyer, and L. K. Jacobs, "Intraoperative Assessment of Final Margins with a Handheld Optical Imaging Probe During Breast-Conserving Surgery May Reduce the Reoperation Rate: Results of a Multicenter Study," *Surgical Oncology*, vol. 22, no. 10, pp. 3356-3362, 2015, [[DOI 10.1245/s10434-015-4665-2](#)].
- [17] J. C. Bolger, J. G. Solon, S. A. Khan, A. D. K. Hill, and C. P. Power, "A comparison of intra-operative margin management techniques in breast-conserving surgery: a standardised approach reduces the likelihood of residual disease without increasing operative time," *Breast Cancer*, vol. 22, no. 3, pp. 262-268, 2015, [[DOI 10.1007/s12282-013-0473-3](#)].
- [18] N. Cabioglu, K. K. Hunt, A. A. Sahin, H. M. Kuerer, G. V. Babiera, S. E. Singletary, G. J. Whitman, M. I. Ross, F. C. Ames, B. W. Feig, T. A. Buchholz, and F. Meric-Bernstam, "Role for Intraoperative Margin Assessment in Patients Undergoing Breast-Conserving Surgery," *Annals of Surgical Oncology*, vol. 14, no. 4, pp. 1458-71, 2007, [[10.1245/s10434-006-9236-0](#)].
- [19] B. S. Nichols, C. E. Schindler, J. Q. Brown, L. G. Wilke, C. S. Mulvey, M. S. Krieger, J. Gallagher, J. Geradts, R. A. Greenup, J. A. Von Windheim, and N. Ramanujam, "A Quantitative Diffuse Reflectance Imaging (QDRI) System for Comprehensive Surveillance of the Morphological Landscape in Breast Tumor Margins," *PLoS One*, vol. 10, no. 16, pp. e0127525, 2015, [[10.1371/journal.pone.0127525](#)].
- [20] M. D. Keller, S. K. Majumder, M. C. Kelley, I. M. Meszoely, F. I. Boulos, G. M. Olivares, and A. Mahadevan-Jansen, "Autofluorescence and Diffuse Reflectance Spectroscopy and Spectral Imaging for Breast Surgical Margin Analysis," *Lasers in Surgery and Medicine*, vol. 42, pp. 15-23, 2010, [[10.1002/lsm.20865](#)].
- [21] L. L. De Boer, B. G. Molenkamp, T. M. Bydlon, B. H. W. Hendriks, J. Wesseling, H. J. C. M. Sterenborg, and T. J. M. Ruers, "Fat/Water ratios measured with diffuse reflectance spectroscopy to detect breast tumor boundaries," *Breast Cancer Research and Treatment*, vol. 152, no. 3, pp. 509-518, 2015, [[10.1007/s10549-015-3487-z](#)].
- [22] Z. Volynskaya, A. S. Haka, K. L. Bechtel, M. Fitzmaurice, R. Shenk, N. Wang, J. Nazemi, R. R. Dasari, and M. S. Feld, "Diagnosing breast cancer using diffuse reflectance spectroscopy and intrinsic fluorescence spectroscopy," *Journal of Biomedical Optics*, vol. 13, no. 2, pp. 024012, 2008, [[10.1117/1.2909672](#)].
- [23] R. L. P. Van Veen, H. J. C. M. Sterenborg, A. W. K. S. Marinelli, and M. Menke-Pluymers, "Intra-operatively assessed optical properties of malignant and healthy breast tissue, to determine the optimum wavelength of contrast for optical mammography," *Journal of Biomedical Optics*, vol. 6, pp. 1129-1136, 2004.
- [24] I. J. Bigio, S. G. Bown, G. Briggs, C. Kelley, S. Lakhani, D. Pickard, P. M. Ripley, I. G. Rose, and C. Saunders, "Diagnosis of breast cancer using elastic-scattering spectroscopy: preliminary clinical results," *J Biomed Opt*, vol. 5, no. 2, pp. 221-228, 2000.
- [25] J. Q. Brown, L. G. Wilke, J. Geradts, S. A. Kennedy, G. M. Palmer, and N. Ramanujam, "Quantitative optical spectroscopy: a robust tool for direct measurement of breast cancer vascular oxygenation and total hemoglobin content in vivo," *Cancer Research*, vol. 69, no. 7, pp. 2919-26, 2009, [[10.1158/0008-5472.CAN-08-3370](#)].
- [26] S. Kennedy, M. Caldwell, T. Bydlon, C. Mulvey, J. Mueller, L. Wilke, W. Barry, N. Ramanujam, and J. Geradts, "Correlation of breast tissue histology and optical signatures to improve margin assessment techniques," *Journal of Biomedical Optics*, vol. 21, no. 6, pp. 066014, 2016, [[10.1117/1.JBO.21.6.066014](#)].
- [27] J. W. Spliethoff, W. Prevoo, M. A. Meier, J. de Jong, D. J. Evers, H. J. C. M. Sterenborg, G. W. Lucassen, B. H. W. Hendriks, and T. J. M. Ruers, "Real-time in vivo tissue characterization with diffuse reflectance spectroscopy during transthoracic lung biopsy: a clinical feasibility study," *Clinical Cancer Research*, vol. 22, no. 2, pp. 357-65, 2016, [[10.1158/1078-0432.CCR-15-0807](#)].

- [28] R. Nachabé, B. H. W. Hendriks, M. Van der Voort, A. E. Desjardins, and H. J. C. M. Sterenborg, "Estimation of biological chromophores using diffuse optical spectroscopy: benefit of extending the UV-VIS wavelength range to include 1000 to 1600 nm," *Optics Express*, vol. 18, no. 24, pp. 1432-1442, 2010.
- [29] R. J. Barnes, M. S. Dhanoa, and S. J. Lister, "Standard Normal Variate Transformation and De-trending of Near-Infrared Diffuse Reflectance Spectra," *Applied Spectroscopy*, vol. 43, no. 5, pp. 772-777, 1989.
- [30] J. E. Burger, "Hyperspectral NIR Image Analysis," Unit of Biomass Technology and Chemistry, Swedish University of Agricultural Sciences, Umeå, 2006.
- [31] R. Nachabé, B. H. W. Hendriks, A. E. Desjardins, M. van der Voort, M. B. van der Mark, and H. J. C. M. Sterenborg, "Estimation of lipid and water concentrations in scattering media with diffuse optical spectroscopy from 900 to 1600 nm," *J Biomed Opt*, vol. 15, no. 3, pp. 037015, 2010, [[10.1117/1.3454392](#)].
- [32] L. L. De Boer, B. H. W. Hendriks, F. Van Duijnhoven, M.-J. T. F. D. Vranken Peeters-Baas, K. Van de Vijver, C. E. Loo, K. Jóźwiak, H. J. C. M. Sterenborg, and T. J. M. Ruers, "Using DRS during breast conserving surgery: identifying robust optical parameters and influence of inter-patient variation," *Biomedical Optics Express*, vol. 7, no. 12, pp. 5188-5200, 2016, [[10.1364/BOE.7.005188](#)].
- [33] V. Vapnik, *The nature of statistical learning theory*: Springer, 2013.
- [34] E. E. Osuna, R. Freund, and F. Girosi, *Support Vector Machines: Training and Applications*, Massachusetts Institute of Technology, 1997.
- [35] R. Nachabé, D. J. Evers, B. H. W. Hendriks, G. W. Lucassen, M. Van der Voort, E. J. Rutgers, M. J. Vranken Peeters, J. A. Van der Hage, H. S. Oldenburg, J. Wesseling, and T. J. M. Ruers, "Diagnosis of breast cancer using diffuse optical spectroscopy from 500 to 1600 nm: comparison of classification methods," *Journal of Biomedical Optics*, vol. 16, no. 8, pp. 087010, 2011, [[10.1117/1.3611010](#)].
- [36] N. Ramanujam, J. Q. Brown, T. M. Bydlon, S. A. Kennedy, L. M. Richards, M. K. Junker, J. Gallagher, W. T. Barry, L. G. Wilke, and J. Geradts, "Quantitative spectral reflectance imaging device for intraoperative breast tumor margin assessment " *Engineering in Medicine and Biology Society*, pp. 6554-6556, 2009, [[10.1109/IEMBS.2009.5334501](#)].
- [37] A. M. Laughney, V. Krishnaswamy, P. Beatriz Garcia-Allende, O. M. Conde, W. A. Wells, K. D. Paulsen, and B. W. Pogue, "Automated classification of breast pathology using local measures of broadband reflectance," *Journal of Biomedical Optics*, vol. 15, no. 6, pp. 066019, 2010, [[10.1117/1.3516594](#)].
- [38] C. Zhu, T. M. Breslin, J. Harter, and N. Ramanujam, "Model based and empirical spectral analysis for the diagnosis of breast cancer," *Optics Express*, vol. 16, no. 19, pp. 14961-14978, 2008.
- [39] J. S. Soares, I. Barman, N. C. Dingari, Z. Volynskaya, W. Liu, N. Klein, D. Plecha, R. R. Dasari, and M. Fitzmaurice, "Diagnostic power of diffuse reflectance spectroscopy for targeted detection of breast lesions with microcalcifications," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 110, no. 2, pp. 471-6, 2013, [[10.1073/pnas.1215473110](#)].
- [40] T. M. Breslin, F. Xu, G. M. Palmer, C. Zhu, K. W. Gilchrist, and N. Ramanujam, "Autofluorescence and Diffuse Reflectance Properties of Malignant and Benign Breast Tissues," *Annals of Surgical Oncology*, vol. 11, no. 1, pp. 65-70, 2004, [[10.1245/ASO.2004.03.031](#)].
- [41] G. M. Palmer, C. Zhu, T. M. Breslin, F. Xu, K. W. Gilchrist, and N. Ramanujam, "Monte Carlo-based inverse model for calculating tissue optical properties. Part II: Application to breast cancer diagnosis," *Applied Optics*, vol. 45, no. 5, pp. 1072-1078, 2006.
- [42] S. K. Majumder, N. Ghosh, and P. K. Gupta, "Support vector machine for optical diagnosis of cancer," *Journal of Biomedical Optics*, vol. 10, no. 2, pp. 024034, 2005, [[10.1117/1.1897396](#)].
- [43] B. H. W. Hendriks, A. J. R. Balthasar, G. W. Lucassen, M. Van der Voort, M. Mueller, V. V. Pully, T. M. Bydlon, C. Reich, A. T. M. H. Van Keersop, J. Kortsmit, G. C. Langhout, and G.-J. Van Geffen,

- "Nerve detection with optical spectroscopy for regional anesthesia procedures," *Journal of Translational Medicine*, vol. 13, no. 380, 2015, [[10.1186/s12967-015-0739-y](#)].
- [44] E. Widjaja, W. Zheng, and Z. Huang, "Classification of colonic tissues using near-infrared Raman Spectroscopy and support vector machines," *International Journal of Oncology*, vol. 32, pp. 653-662, 2008.
- [45] M. C. Skala, G. M. Palmer, K. M. Vrotsos, A. Gendron-Fitzpatrick, and N. Ramanujam, "Comparison of a physical model and principal component analysis for the diagnosis of epithelial neoplasias in vivo using diffuse reflectance spectroscopy," *Optics Express*, vol. 15, no. 12, pp. 7863-7875, 2007, [[10.1364/OE.15.007863](#)].
- [46] P. A. Valdés, A. Kim, F. Leblond, O. M. Conde, B. T. Harris, K. D. Paulsen, B. C. Wilson, and D. W. Roberts, "Combined fluorescence and reflectance spectroscopy for in vivo quantification of cancer biomarkers in low- and high-grade glioma surgery," *Journal of Biomedical Optics*, vol. 16, no. 11, pp. 116007, 2011.

General discussion

This thesis is focused on using diffuse reflectance spectroscopy for margin assessment during breast-conserving surgery. The technology in itself has proven to be able to characterize different tissue types but has not yet reached broad clinical acceptance (Chapter 2). We hypothesize that DRS technology could help the surgeon in performing optimal surgery by avoiding positive resection margins as well as providing the possibility to spare more healthy tissue. Therefore we would ultimately like to develop a tool that can improve breast-conserving surgery by intra-operatively detecting positive resection margins as well as helping in avoiding unnecessary resection of healthy tissue. To achieve this goal in this thesis several challenges, as proposed in the introduction, were addressed in this project and will be discussed one by one in the following paragraphs, among some general comments.

HOW TO USE DRS FOR DISCRIMINATING BETWEEN HEALTHY TISSUE AND TUMOR TISSUE

The first question was if DRS measurements differed between healthy tissue from tumor tissue and how optical differences could be used for discriminating these tissue types. It was found that several fit parameters were useful for discriminating tumor tissue from healthy tissue. Of these fit parameters, the one describing the ratio between the amount of fat and water (F/W-ratio) had the best performance (Chapter 3 & Chapter 4). In the first *ex vivo* study, we showed that based on this parameter we were able to detect the tumor boundary in one specimen as there was a transition seen in the F/W-ratio around the tumor. Besides the fit parameters, also the spectral data (i.e. the intensity at a certain wavelength) allowed discriminating tumor tissue from healthy tissue (Chapter 8). In addition, also spectral features such as slopes and dips that are derived from certain wavelength ranges between 850 and 1600 nm were used. (Chapter 7) The literature on the optical difference between healthy and tumor tissue is mainly limited to the visual wavelength range. Within this wavelength range, oxygen saturation, β -carotene, and scattering are reported to be able to discriminate healthy tissue from tumor tissue.[1-3] By including the near-infrared (NIR) wavelength range, up to 1600 nm, also differences in the amount of fat and water could be measured. The fact that the F/W-ratio provided such excellent discrimination between tumor tissue and healthy tissue can be explained by the nature of healthy breast tissue and tumor tissue. Healthy breast tissue consists predominantly of fat, whereas in tumor tissue an increased water concentration is present. These results are supported by other papers that also included a higher wavelength range.[4, 5] The other great benefit of using the NIR wavelength range is that blood, which is an important absorber in the visual wavelength range, does not affect these wavelengths. Thus, the effect of extravascular blood, which is unavoidable during surgery and could contaminate the optical measurement, is avoided. The fact that the F/W-ratio is independent of the visual wavelengths makes this fit parameter suitable for a surgical application.[6]

DIFFERENCES BETWEEN *IN VIVO* AND *EX VIVO* MEASUREMENTS

The next question was to assess if discrimination between healthy tissue and tumor tissue was different for *ex vivo* and *in vivo* measurements. In Chapter 4 the *in vivo* measurements are compared with the *ex vivo* measurements and there were no significant differences in any of the fit parameters due to tissue status at the time of measuring. Comparisons between the *in vivo* DRS measurements and the *ex vivo* DRS measurements in literature are sparse, and either performed in pancreatic cancer patients[7] or healthy tissue from animals. [8, 9] It should be noted that comparing *in vivo* measurements with *ex vivo* measurements is also difficult in how to exactly define ‘*in vivo*’ and ‘*ex vivo*’. Probably from the time the tissue is manipulated during surgery and once it is exposed to air, tissue characteristics will alter continuously. Therefore the interpretation of the results in this study should be that the measurements over time are relatively stable and especially that the optical fit parameters that allowed to discriminate tumor tissue from healthy breast tissue did not change after resection of the tissue. In this study also the menopausal status of the patient and treatment with neoadjuvant chemotherapy were included, which initially indicated that there was no influence of these two patient-specific characteristics.

THE POTENTIAL INFLUENCE OF PATIENT-SPECIFIC CHARACTERISTICS

The effect of chemotherapy and menopausal status was researched in Chapter 4 and in Chapter 6. In Chapter 6 the analysis was performed with a much larger patient number. In the analysis in this chapter not only the fit parameter data were included, but also the full spectrum was evaluated. The results of this study indicated that neither the full spectrum nor the fit parameter data, except for one fit parameter, was significantly affected by chemotherapy. For menopausal status, no influence on the data was found. The main conclusion from this work is that, although chemotherapy has an effect on tissue, as slightly more connective tissue was present in the patients after chemotherapy, it did not result in significantly different DRS measurements. This means that also in patients treated by neoadjuvant chemotherapy, DRS could still be feasible for margin assessment. This is one of the few papers that investigated neoadjuvant chemotherapy in a large patient group. Thus far, studies that specifically assessed the influence of chemotherapy on optical measurements both only evaluated patient numbers. [4, 10]

The majority of publications that included DRS measurements of patients with neoadjuvant chemotherapy are focused on therapy monitoring.[11-15] In these papers, the measurements acquired pre-treatment are compared with the measurements acquired post-therapy in the same patient. This is an entirely different focus than our question which was if DRS measurements can be used for margin assessment in patients that had chemotherapy

similarly as to the situation in which patients did not have chemotherapy. However, the fact that in therapy monitoring studies the same fit parameters are used for discrimination in the pre-treatment data compared to post-treatment data underlines our findings that the influence of chemotherapy does not impair DRS measurements in discriminating healthy tissue from tumor tissue.

THE INHOMOGENEOUS STRUCTURE OF BREAST TISSUE

Breast tissue is known for its inhomogeneous nature, which makes a robust correlation with histopathology imperative. A registration method is presented that accounts for deformations of the tissue that are introduced during processing at the histopathology department. This registration method was compared with a registration method without correction of the deformations, by validating both with an independent modality (micro-computed tomography). The proposed method that corrected for deformations provided an improved correlation between the optical measurements and histopathology (Chapter 5). Histopathology correlation is extremely important for validating any optical technology since an assessment of microscopy sections by a pathologist is considered the gold standard or ground truth for tissue characterization. Although the importance of correct correlation is acknowledged by many[16, 17], registering *ex vivo* measurements is often performed in a very rudimental manner. The registration method presented in this thesis is labor intensive but shows a distinct improvement over the rudimental registration method.

With the improved registration method, we were able to estimate the composition of the measurement locations that were acquired *ex vivo*. This information enabled us to evaluate how the purity of the measurement locations used for the development of a classification model, influenced the performance of the model. (Chapter 7) It was found that a classification model developed with up to 20% fat tissue in the locations defined as 'tumor', and up to 20% of connective tissue in the locations defined as 'healthy' still resulted in an adequate performance of the model (0.90 Matthews Correlation Coefficient). Using this model to classify the mixture locations showed that as the percentage of tumor cells in the measurement decreased, the percentage of locations classified as 'healthy' increased. On the other hand, some locations without any tumor cells were classified as 'tumor'. This can probably be explained by the fact that connective tissue is always present to some extent in the tumor locations. At the same time, connective tissue can be present in healthy tissue, thereby decreasing the optical contrast between tumor and healthy.

MEASURING DCIS

The next challenge we addressed was the detection of DCIS with DRS. Since DCIS has a diffuse growth pattern, detection of this tissue type is more challenging, and thus leads to the majority of positive resection margins. Using the improved registration method allowed us to compare the DRS measurements of DCIS and invasive carcinoma. It turned out that DRS measurements of both tissue types were not significantly different (Chapter 7). Since there was no difference found in the DRS measurements of invasive carcinoma surrounded by connective tissue and DCIS surrounded by connective tissue, the two tissue types were combined in one class called malignant. Comparing this result to the work of others that researched the use of DRS for the detection of DCIS is difficult as in these papers measurements were acquired over a smaller wavelength range, typically the visual wavelengths and often the analysis was performed with fit parameters derived from the spectrum.[2, 10, 18, 19] Another complicating factor is the definition of a DCIS location. In our case, we used pure locations in our analysis. Thus, locations that were a mixture of fat tissue, connective tissue, and either invasive carcinoma or DCIS cells, were not used in the comparison. Such locations were considered a mixture of tissue types and were on that ground included in the mixture dataset.

TRANSLATING THE TECHNOLOGY INTO THE CLINICAL WORKFLOW

The last challenge was the application of the technology in the *in vivo* setting. In an *in vivo* study, we showed that we were able to acquire these measurements continuously and that we were able to classify them with a high accuracy of 0.93 (Chapter 8). Similar to the result in Chapter 3, a clear transition could be seen in the DRS measurements, this time along the needle trajectory that reached from healthy tissue to the tumor. The continuous acquisition of DRS measurements has not been reported on before. There have been publications on the integration of optical technology in needle applications for breast biopsies but their main focus was on improving the biopsy yield of these procedures.[20, 21] This is a different goal compared to the goal of the current study which was to show the feasibility of the technology for real-time acquisition and correct classification of these DRS measurements. With the optical biopsy needle used in this setting, it was possible to achieve optimal correlation with histopathology which is essential in developing an intra-operative margin assessment tool.

Besides the interpretation of the results presented in this thesis, there are some more general points to discuss. These will be addressed in the following part of the discussion.

WAVELENGTH RANGE

Other groups have also reported on DRS for breast margin assessment[3, 22-24]. Measurements are predominantly acquired in the visual wavelength range, and analysis is often performed with fit parameters derived from the spectra. The fit parameters and wavelengths, from the visual wavelength range, that were useful for discriminating healthy tissue and tumor tissue in our analysis are similar to the results of others. Extending the measurements to the longer wavelength region (NIR) allowed us to use the F/W-ratio, or wavelengths that are related to the absorption of fat and water, which proved to be very useful. It should be noted that the F/W-ratio was extremely powerful in the first *ex vivo* studies (Chapter 3 & 4), whereas in the *in vivo* study (Chapter 8) this parameter had a slightly lower accuracy. In fact, in this study, the classification model based on a selection of wavelengths, of which some were still clearly related to fat and water absorption, showed better performance compared to the classification model based on the fit parameters. There are a few possible explanations for this result. First, in the *ex vivo* study of Chapter 3, locations were selected in consultation with a pathologist. To make sure the healthy tissue measurements were indeed healthy; areas of macroscopic pure fat tissue were selected, potentially creating bias by avoiding locations that macroscopically were not certainly healthy. Secondary, in the *ex vivo* study presented in Chapter 3 the average F/W-ratios were used in the analysis. If the number of locations in pure fat tissue is much higher than the number of locations in which fat tissue is combined with some connective tissue, the average F/W-ratio will likely be similar to the F/W-ratio measured in pure fat. In Chapter 4 both *in vivo* and *ex vivo*, two measurements were acquired of healthy tissue and tumor tissue. The healthy measurements were acquired a few centimeters of the tumor (*in vivo*) or from macroscopically healthy tissue (*ex vivo*). In the results, one can see that the for the *ex vivo* measurement the F/W-ratio performs slightly better compared to the *in vivo* measurements, probably due to the selection bias as previously explained. In the *in vivo* healthy locations, potentially more connective tissue could have been present which has lower contrast to tumor tissue than the contrast between pure fat tissue and tumor tissue.

MEASUREMENT SET UP AND HISTOPATHOLOGY CORRELATION

In this project both *in vivo* and *ex vivo* measurements were performed. The *ex vivo* measurement set-up had advantages and disadvantages over the *in vivo* measurement set-up. The *ex vivo* setting allowed performing more measurements per specimen, in a relatively controlled manner. The main disadvantage, however, is that the visual wavelength range can be hampered by pathology ink that is applied to the margins before measurements are performed. This influence can be diminished to some extent by including the absorption

spectra of these paints in the analytical fit model, as this will then provide an estimation of the amount of pathology ink present in these measurements. However, when using the measured intensities for an analysis, quantification of the influence of pathology ink on the data is difficult. In case of the *in vivo* measurements, no pathology ink was present yet which is beneficial. But in this setting, correlating the measurements with histopathology is more challenging. The optical biopsy needle (Chapter 8) resolved this problem, but could only be used for taking one biopsy per needle. Furthermore, as these measurements were performed during a routine biopsy procedure, the time allowed for measuring was limited. It is thus a balance between acquiring a large dataset with accurate pathology correlation at the cost of pathology ink affecting the visual wavelength range, or performing these measurements *in vivo* in a smaller number of patients with also constraints on the number of measurements per patient. Especially in the case of model development of classification models the value of a large dataset and the value of reliable histopathology labeling of the data should not be underestimated.

NORMALIZATION

In Chapter 4 we hypothesized that normalizing the measurements to the patient's own reference would improve the discriminative power of the fit parameters. However, this normalization procedure did not provide better results for any of the fit parameters, suggesting that inter-patient variation is negligible. This is a different conclusion opposed to the conclusion from the first *ex vivo* study (Chapter 3). Importantly, in the first analysis for each patient, average F/W-ratios were calculated for normal tissue, tissue at the border of the tumor, and tumor tissue. The comparison between tumor tissue and healthy tissue was performed with these average F/W-ratios. Possibly intra-patient variation induced some of the variation in F/W-ratio in normal tissue between patients, which was then corrected for by normalization. In the second study, only one measurement of healthy tissue was available and the intra-patient variation could not be assessed. This one measurement of healthy tissue might not have been sufficient to function as a 'reference'. It should also be noted that the distribution of connective tissue is not homogenous over the breast, as can be seen in, for example, mammography. Most of the glandular tissue is in the center of the breast whereas there is often a rim of pure fat tissue close to the skin. The actual location of the measurement that is used as a reference could thus have an impact on this normalization method.

FIT PARAMETERS VERSUS MEASURED INTENSITY

Over the course of this project, also focus was put on the analysis of the intensity measured per wavelength opposed to the fit parameters that were derived from the spectrum. As explained in the introduction, both methods have advantages and disadvantages. However, especially when using complex machine learning algorithms overfitting can be a threat to an accurate assessment of the performance of the model. A possible approach to circumvent overfitting is to reduce the number of features used in machine learning. In a sense, fitting the measurement spectrum can be seen as a way of feature reduction, similar to selecting a number of wavelengths. However, when using wavelength selection as a feature reduction method, different wavelengths can be selected whereas feature reduction with fit parameters will always result in the same type of information. The wavelength selection method that was used in the majority of analysis in this thesis was a Wilcoxon ranked sum test with the spectra of different tissue classes. This is a relatively simple method that is useful in pointing out the wavelengths that significantly differ between tissue types but does not account for the potential discriminative power of the combination of wavelengths that on their own are not able to discriminate tissue types, but become valuable if used together. In Chapter 7 another way of feature reduction was introduced. In this case, so-called spectral features were calculated by extracting the slopes and quantifying the dips of certain wavelength ranges. The selection of the ranges that were used was based on visual assessment of the 'pure' spectra. This method could be improved by having an algorithm select these areas. In addition, there was no analysis performed to evaluate the actual contribution of each of the 24 selected spectral features, or combination of features, to discriminating tissue types. Potentially some of these spectral features can be discarded since they are redundant.

MEASURING AT THE TIME OF RESECTION

A difference between other groups using diffuse reflectance measurements for breast margin assessment and this project is that the focus of their research is mainly directed towards imaging the entire margin via a scanning mechanism[23-25]. The advantage of this approach is that an entire margin can be imaged at once; however, the most important drawback is that such an application can only be performed after resection. Adjustment of the resection plane at the time of resection is not possible anymore. And once a positive resection margin is detected the surgeon has to return to the patient and retrieve the positive margin in the surgical wound, a translation step which can be prone to errors. Therefore in this project, we aim at the development of a tool that can be used during surgery, enabling the surgeon to change the resection plane when indicated by the DRS

measurement. Another benefit of being able to perform measurements intra-operatively in the surgical wound is that it allows assessing the tissue that is still inside the patient. This can be particularly beneficial in case the tumor grows discontinuously.[26]

THE GOLD STANDARD

An important topic throughout this thesis, but also for validation of optical measurements in general, is the correlation with histopathology. The correlation with histopathology became more reliable as the project progressed. For the *in vivo* measurements, this was achieved by the use of the optical biopsy needle that allowed sampling the tissue directly after the DRS measurement of the tissue (Chapter 8). For the *ex vivo* measurements, a registration method was developed that corrected for tissue deformation introduced during the processing of the tissue at the pathology department, a reported on in Chapter 5. Although the correct labeling of the optical measurement with histopathology is inevitable for validation of DRS measurements, some comments can be made regarding the value of histopathology. One of the issues encountered is that what is seen by the pathologist is more than just a combination of substances. The pathologist assesses the margin by looking at the coherence of cells and other tissue characteristics. Optical technology will measure the composition and morphology of a small measurement volume, but cannot take into account how this volume is related to the rest of the surrounding tissue as can be done by a pathologist. Related to this topic, is the fact that pathological assessment is subjective to a certain extent.[27-31] Regarding the detection of positive resection margins by the pathologist also other issues play a role such as undersampling and processing of the tissue that can lead to overseeing positive resection margins.[32-35]

Another way of validating a technology would be to look at different metrics such as survival and local recurrence. The beauty of these outcome measures is that they are a better representation of the effect of the new technology on the clinical outcome; however, for measuring this robustly very large patient studies are required including hundreds of patients. Such studies also allow investigating how DRS technology is utilized by the surgeon. This was not researched during this project but is important for clinical acceptance of a technology.

OTHER TECHNOLOGIES RESEARCHED FOR MARGIN ASSESSMENT

Compared to margin assessment technologies currently used in the clinic, margin assessment with DRS has clear benefits. The fact that measurements can be performed in real-time, that they are non-destructive and do not require skilled personnel for interpretation favors DRS

over other technologies such as imprint cytology, frozen section analysis, and ultrasound-guided surgery.[36-39] However, DRS is not the only 'new' technology investigated for intra-operative margin assessment. A recent paper of Shipp et al used a combination of auto-fluorescence and Raman spectroscopy to detect the microscopic residual tumor at the surface of the excised breast tissue.[40] The reported sensitivity and specificity was 95% and 82% respectively. Yet, measuring an uncut specimen took 12 to 24 minutes which is considered long and will not fit in the clinical workflow. Another 3D scanning instrument based on the same combined techniques was developed by Thomas et al, but this device has not been tested on breast cancer specimens.[41]

Optical coherence tomography (OCT) is another label-free imaging technology based on interferometry of light that is currently investigated for intra-operative margin assessment.[42] The sensitivity and specificity of OCT are suitable for intra-operative use (91.7% and 92.1%) but the main drawback is that the interpretation of the images requires expert knowledge. Although the images of the *in vivo* measurements appear similar to histopathology images, interpretation of the images is still difficult and highly depends on the training of the operator of the device.

Currently, also fluorescence lifetime measurements are researched for margin assessment. Accuracy for discrimination between cancerous, fibrous, and adipose tissue was high with 97% using automated classification[43]. This technology has not been used intra-operatively for breast applications, although it has been used *in vivo* during head and neck surgery.[44]

Electrical impedance measurements have also been researched for margin assessment. [45] At present, the MarginProbe has FDA approval for intra-operative use. Based on electrical impedance measurements it has a sensitivity of 70.4% and a specificity of 58.5%.[46, 47] Although, sensitivity and specificity are not as high as of the other technologies, results of the first trials show that use of the device leads to a decrease in the number of positive resection margins while excision volumes did not increase. [48, 49]

DETECTION LIMIT

The tumor border was detected in both the *ex vivo* and the *in vivo* measurements (Chapter 3 & Chapter 8). An important question is how the minimal amount of tumor present in the measurement volume relates to the detectability of a positive resection margin. Thus, in case of a positive resection margin, is the sensitivity of DRS high enough to be able to detect this. It should be noted that this question is strongly related to the definition of a positive resection margin. In Chapter 7 it was found that mixed locations that consist of a combination of fat tissue and connective tissue with tumor cells were sometimes difficult to classify. The classification accuracy was dependent first, on which locations were used for developing the

classification model, in other words, the purity of the locations. And second, the percentage of tumor cells in the measurement volume of the locations that were classified by the model. With an increasing percentage of tumor cells, more locations were classified as 'tumor'. It is safe to say that finding a single tumor cell in the DRS measurement volume of a few mm³ will not be possible with DRS. However, such extreme sensitivity might not be necessary. Over the course of the last years, the definition of a positive resection margin has shifted. A positive resection margin nowadays is more seen as an indication of the tumor burden which is left behind. Similarly, a negative resection margin is not considered as an absolute measure of complete eradication of the tumor.[26, 32, 35] For example, research shows that up to 46% of patients with focally positive resection margins or 'close' margins ('no ink on tumor', but within 1 mm distance of the margin) had residual disease in their re-excision.[45] Nevertheless, it is acknowledged that when these tumor deposits are small enough (as in case of a focally positive margin), they can be left behind since they will be eradicated by the postoperative radiotherapy which is standard of care in current clinical practice.[35, 50] As for DCIS, which is more difficult to detect, a more conservative definition is used in the guidelines. But similar to the definition of a positive resection margin for invasive carcinoma, this definition also became more liberal over the course of time.[51, 52] This could also imply that it is not obligatory for DRS to be able to distinguish invasive carcinoma from DCIS as long as DCIS is detected as 'tumor'.

THE VALUE OF MARGIN ASSESSMENT

One could criticize that finding positive resection margins has very limited value in clinical practice as the overall survival of patients that are treated for early-stage breast cancer is already very good.[53, 54] However, it should be noted that in up to 20% of patients secondary surgery is performed[53, 55-59], which, would not have been necessary if the resection had been optimal during the first surgery. Secondary, intra-operative guidance could also improve cosmetic outcome as more healthy tissue can be spared and decrease the number of patients that are treated with boost radiotherapy that is given in extension to the already standard radiotherapy treatment. Providing adequate intra-operative margin assessment may not fundamentally change the survival of patients, but can have a major impact on their lives, as well as health care resources[56, 60, 61].

MEASURING AT THE RESECTION MARGIN

Ultimately, the final application we envision is that DRS measurements of the resection margin provide the surgeon with real-time information that allows him/her to perform

optimal surgery. However, the majority of the measurements in this thesis were acquired from slices. It is uncertain whether these measurements on slices can be directly translated to measurements on the resection margin as this had not been tested. In a small *ex vivo* study, measurements were obtained from the actual resection margin. However, due to the difficulty in correlating these measurements with histopathology, it was not possible to include them in this thesis. A larger study is imperative to assess if there are effects in the resection margin measurements that are not present in the measurements from slices, that factually measure 'inside' the specimen. Another aspect that was not researched in this project is the real-time classification of the DRS measurements. We have shown that DRS measurements can be acquired continuously in real-time and that we could classify them correctly afterward, but we have not done this real-time yet. Lastly, it is important to investigate how a surgeon would use the tool intra-operatively, which was also not included in this thesis.

OUTLOOK

Although overall survival of patients diagnosed with early-stage breast cancer is very good there is still room for improvement, in terms of decreasing the number of re-excision, excision volumes, and avoiding boost radiotherapy that is given additionally to the standard radiotherapy. Fiber-optic DRS, by means of a surgical tool, seems a potential candidate to guide a surgeon during breast-conserving surgery towards the optimal resection plane. The next step should be to assess the sensitivity and specificity of the technology when used in a setting identical to the envisioned application, i.e. not *ex vivo*, not on slices, but with the probe on the intended resection margin, during surgery. Thus, an *in vivo* study should be designed aiming at acquiring measurements at the resection margin while at the same time ensuring measurements can be accurately correlated to histopathology. The focus should also be put on the real-time classification of the measurements and implementation in the clinical workflow. Ultimately, the final step would be to assess the value of DRS in terms of the decrease in positive resection margins and influence on resection specimen volumes.

REFERENCES

- [1] C. Zhu, G. M. Palmer, T. M. Breslin, J. Harter, and N. Ramanujam, "Diagnosis of Breast Cancer Using Diffuse Reflectance Spectroscopy: Comparison of a Monte Carlo Versus Partial Least Squares Analysis Based Feature Extraction Technique," *Lasers in Surgery and Medicine*, vol. 38, no. 7, pp. 714-24, 2006, [10.1002/lsm.20356].
- [2] S. Kennedy, J. Geradts, T. Bydlon, J. Q. Brown, J. Gallagher, M. Junker, W. Barry, N. Ramanujam, and L. Wilke, "Optical breast cancer margin assessment: an observational study of the effects of tissue heterogeneity on optical contrast," *Breast Cancer Research*, vol. 12, no. 6, pp. R91, 2010, [10.1186/bcr2770].
- [3] Z. Volynskaya, A. S. Haka, K. L. Bechtel, M. Fitzmaurice, R. Shenk, N. Wang, J. Nazemi, R. R. Dasari, and M. S. Feld, "Diagnosing breast cancer using diffuse reflectance spectroscopy and intrinsic fluorescence spectroscopy," *Journal of Biomedical Optics*, vol. 13, no. 2, pp. 024012, 2008, [10.1117/1.2909672].
- [4] A. M. Laughney, V. Krishnaswamy, E. J. Rizzo, M. C. Schwab, R. J. Barth Jr, D. J. Cuccia, B. Tromberg, K. D. Paulsen, B. W. Pogue, and W. A. Wells, "Spectral discrimination of breast pathologies *in situ* using spatial frequency domain imaging," *Breast Cancer Research*, vol. 15, pp. R61, 2013.
- [5] P. Taroni, A. Bassi, D. Comelli, A. Farina, R. Cubeddu, and A. Pifferi, "Diffuse optical spectroscopy of breast tissue extended to 1100 nm," *J Biomed Opt*, vol. 14, no. 5, pp. 054030, 2009, [10.1117/1.3251051].
- [6] J. W. Spliethoff, W. Prevoo, M. A. Meier, J. de Jong, D. J. Evers, H. J. C. M. Sterenborg, G. W. Lucassen, B. H. W. Hendriks, and T. J. M. Ruers, "Real-time *in vivo* tissue characterization with diffuse reflectance spectroscopy during transthoracic lung biopsy: a clinical feasibility study," *Clinical Cancer Research*, vol. 22, no. 2, pp. 357-65, 2016, [10.1158/1078-0432.CCR-15-0807].
- [7] W. R. Lloyd, R. H. Wilson, S. Y. Lee, M. Chandra, B. McKenna, D. Simeone, J. Scheiman, and M.-A. Mycek, "In vivo optical spectroscopy for improved detection of pancreatic adenocarcinoma: a feasibility study," *Biomedical Optics Express*, vol. 5, no. 1, pp. 9-15, 2014, [10.1364/BOE.5.000009].
- [8] E. Salomatina, and A. N. Yaroslavsky, "Evaluation of the *in vivo* and *ex vivo* optical properties in a mouse ear model," *Physics in Medicine and Biology*, vol. 53, pp. 2797-2807, 2008, [10.1088/0031-9155/53/11/003].
- [9] G. M. Palmer, C. L. Marshak, K. M. Vrotsos, and N. Ramanujam, "Optimal Methods for Fluorescence and Diffuse Reflectance Measurements of Tissue Biopsy Samples," *Lasers in Surgery and Medicine*, vol. 30, pp. 191-200, 2002, [10.1002/lsm.10026].
- [10] M. D. Keller, S. K. Majumder, M. C. Kelley, I. M. Meszoely, F. I. Boullos, G. M. Olivares, and A. Mahadevan-Jansen, "Autofluorescence and Diffuse Reflectance Spectroscopy and Spectral Imaging for Breast Surgical Margin Analysis," *Lasers in Surgery and Medicine*, vol. 42, pp. 15-23, 2010, [10.1002/lsm.20865].
- [11] A. E. Cerussi, V. W. Tanamai, D. Hsiang, J. Butler, R. S. Mehta, and B. J. Tromberg, "Diffuse optical spectroscopic imaging correlates with final pathological response in breast cancer neoadjuvant chemotherapy," *Philosophical Transactions of the Royal Society A*, vol. 369, pp. 4512-4530, 2011, [10.1098/rsta.2011.0279].
- [12] O. Falou, H. Soliman, A. Sadeghi-Naini, S. Iradji, S. Lemon-Wong, J. Zubovits, J. Spayne, R. Dent, M. Trudeau, J. F. Boileau, F. C. Wright, M. J. Yaffe, and G. J. Czarnota, "Diffuse Optical Spectroscopy Evaluation of Treatment Response in Women with Locally Advanced Breast Cancer Receiving Neoadjuvant Chemotherapy," 2012.
- [13] T. D. O'Sullivan, A. Leproux, J. Chen, S. Bahri, A. Matlock, D. Roblyer, C. E. McLaren, W. Chen, A. E. Cerussi, M. Su, and B. J. Tromberg, "Optical imaging correlates with magnetic resonance imaging breast density and reveals composition changes during neoadjuvant chemotherapy," *Breast Cancer Research*, vol. 15, no. R14, 2013.

- [14] M. V. Pavlov, T. I. Kalganova, S. Lyubimtseva, V. I. Plekhanov, G. Y. Golubyatnikov, O. Y. Ilyinskaya, A. G. Orlova, P. V. Subochev, D. V. Safonov, N. M. Shakhova, and A. V. Maslennikova, "Multimodal approach in assessment of the response of breast cancer to neoadjuvant chemotherapy," *Journal of Biomedical Optics*, vol. 23, no. 9, pp. 091410, 2018, [[10.1117/1.JBO.23.9.091410](#)].
- [15] A. Sadeghi-Naini, E. Vorauer, L. Chin, O. Falou, W. T. Tran, F. C. Wright, S. Gandhi, M. J. Yaffe, and G. J. Czarnota, "Early detection of chemotherapy-refractory patients by monitoring textural alterations in diffuse optical spectroscopic images," *Medical Physics*, vol. 42, no. 11, pp. 6130-6146, 2015, [[10.1118/1.4931603](#)].
- [16] B. C. Wilson, M. Jermyn, and F. Leblond, "Challenges and opportunities in clinical translation of biomedical optical spectroscopy and imaging," *Journal of Biomedical Optics*, vol. 23, no. 3, pp. 030901, 2018, [[10.1117/1.JBO.23.3.030901](#)].
- [17] W. A. Wells, P. E. Barker, C. MacAulay, M. Novelli, R. M. Levenson, and J. M. Crawford, "Validation of novel optical imaging technologies: the pathologists' view," *Journal of Biomedical Optics*, vol. 12, no. 5, pp. 051801, 2007, [[10.1117/1.2795569](#)].
- [18] D. J. Evers, R. Nachabé, M. J. Vranken Peeters, J. A. Van der Hage, H. S. Oldenburg, E. J. Rutgers, G. W. Lucassen, B. H. W. Hendriks, J. Wesseling, and T. J. M. Ruers, "Diffuse reflectance spectroscopy: towards clinical application in breast cancer," *Breast Cancer Research and Treatment*, vol. 137, pp. 155-165, 2013, [[10.1007/s10549-012-2350-8](#)].
- [19] S. K. Majumder, M. D. Keller, F. I. Boulos, M. C. Kelley, and A. Mahadevan-Jansen, "Comparison of autofluorescence, diffuse reflectance, and Raman spectroscopy for breast tissue discrimination," *J Biomed Opt*, vol. 13, no. 5, pp. 054009, 2008, [[10.1117/1.2975962](#)].
- [20] B. Yu, E. S. Burnside, G. A. Sisney, J. M. Harter, C. Zhu, A. H. Dhalla, and N. Ramanujam, "Feasibility of near-infrared diffuse optical spectroscopy on patients undergoing image-guided core-needle biopsy," *Optics Express*, vol. 15, no. 12, pp. 7335-50, 2007.
- [21] R. L. P. Van Veen, A. Amelink, M. Menke-Pluymers, C. Van der Pol, and H. J. C. M. Sterenborg, "Optical biopsy of breast tissue using differential path-length spectroscopy," *Physics in Medicine and Biology*, vol. 50, pp. 2573-2581, 2005, [[10.1088/0031-9155/50/11/009](#)].
- [22] B. S. Nichols, C. E. Schindler, J. Q. Brown, L. G. Wilke, C. S. Mulvey, M. S. Krieger, J. Gallagher, J. Geradts, R. A. Greenup, J. A. Von Windheim, and N. Ramanujam, "A Quantitative Diffuse Reflectance Imaging (QDRI) System for Comprehensive Surveillance of the Morphological Landscape in Breast Tumor Margins," *PloS One*, vol. 10, no. 16, pp. e0127525, 2015, [[10.1371/journal.pone.0127525](#)].
- [23] B. S. Nichols, A. Llopis, G. M. Palmer, S. S. McCachren, O. Senlik, D. Miller, M. A. Brooke, N. M. Jokerst, J. Geradts, R. Greenup, and N. Ramanujam, "Miniature spectral imaging device for wide-field quantitative functional imaging of the morphological landscape of breast tumor margins," *Journal of Biomedical Optics*, vol. 22, no. 2, pp. 026007, 2017, [[10.1117/1.JBO.22.2.026007](#)].
- [24] N. Lue, J. W. Kang, C. Yu, I. Barman, N. C. Dingari, M. S. Feld, R. R. Dasari, and M. Fitzmaurice, "Portable Optical Fiber Probe-Based Spectroscopic Scanner for Rapid Cancer Diagnosis: A New Tool for Intraoperative Margin Assessment," *PloS One*, vol. 7, no. 1, pp. e30887-1 - e30887-10, 2012, [[10.1371/journal.pone.0030887](#)].
- [25] A. M. Laughney, V. Krishnaswamy, E. J. Rizzo, M. C. Schwab, R. J. Barth, B. W. Pogue, K. D. Paulsen, and W. A. Wells, "Scatter Spectroscopic Imaging Distinguishes between Breast Pathologies in Tissues Relevant to Surgical Margin Assessment," *Clinical Cancer Research*, vol. 18, pp. 6315-6325, 2012, [[10.1158/1078-0432.CCR-12-0136](#)].
- [26] S. N. Goldstein, "Controversies in Pathology in Early-Stage Breast Cancer," *Seminars in Radiation Oncology*, vol. 21, pp. 20-25, 2011, [[10.1016/j.semradonc.2010.08.003](#)].
- [27] S. L. Jackson, P. D. Frederick, M. S. Pepe, H. D. Nelson, D. L. Weaver, K. H. Allison, P. A. Carney, B. M. Geller, A. N. A. Tosteson, T. Onega, and J. G. Elmore, "Diagnostic Reproducibility: What Happens When the Same Pathologist Interprets the Same Breast Biopsy Specimen at Two

- Points in Time?," *Annals of Surgical Oncology*, vol. 24, pp. 1234-1241, 2017, [[10.1245/s10434-016-5695-0](#)].
- [28] P. Boiesen, P.-O. Bendahl, L. Anagnostaki, H. Domanski, E. Holm, I. Idvall, S. Johansson, O. Ljungberg, A. Ringberg, G. Östberg, and M. Fernö, "Histologic Grading in Breast Cancer: Reproducibility Between Seven Pathologic Departments," *Acta Oncologica*, vol. 39, no. 1, pp. 41-45, 2000, [[10.1080/028418600430950](#)].
- [29] S. J. Schnitt, "Evaluation of Microscopic Margins in Patients with Invasive Breast Cancer: Technical and Interpretive Considerations," *The Breast Journal*, vol. 4, no. 3, pp. 204-208, 1998, [[10.1046/j.1524-4741.1998.430204.x](#)].
- [30] K. H. Allison, L. M. Reisch, P. A. Carney, D. L. Weaver, S. J. Schnitt, F. P. O'Malley, B. M. Geller, and J. G. Elmore, "Understanding diagnostic variability in breast pathology: lessons learned from an expert consensus review panel," *Histopathology*, vol. 65, no. 2, pp. 240-251, 2014, [[10.1111/his.12387](#)].
- [31] J. G. Elmore, G. M. Longton, P. A. Carney, B. M. Geller, T. Onega, A. N. A. Tosteson, H. D. Nelson, M. S. Pepe, K. H. Allison, S. J. Schnitt, F. P. O'Malley, and D. L. Weaver, "Diagnostic Concordance Among Pathologists Interpreting Breast Biopsy Specimens," *Journal of the American Medical Association*, vol. 313, no. 11, pp. 1122-1132, 2015, [[10.1001/jama.2015.1405](#)].
- [32] C. M. O'Kelly Priddy, V. A. Forte, and J. E. Lang, "The Importance of Surgical Margins in Breast Cancer," *Journal of Surgical Oncology*, vol. 113, pp. 256-263, 2016, [[10.1002/jso.24047](#)].
- [33] M. J. Wright, J. Park, J. V. Fey, A. Park, A. O'Neill, L. K. Tan, P. I. Borgen, H. S. Cody, K. J. Van Zee, and T. A. King, "Perpendicular Inked Versus Tangential Shaved Margins in Breast-Conserving Surgery: Does the Method Matter?," *Journal of the American College of Surgeons*, vol. 204, pp. 541-549, 2007, [[10.1016/j.jamcollsurg.2007.01.031](#)].
- [34] R. Tang, S. B. Coopey, M. C. Specht, L. Lei, M. A. Gadd, K. S. Hughes, E. F. Brachtel, and B. L. Smit, "Lumpectomy specimen margins are not reliable in predicting residual disease in breast conserving surgery," *The American Journal of Surgery*, vol. 210, pp. 93-98, 2015, [[10.1016/j.amjsurg.2014.09.029](#)].
- [35] N. Houssami, and M. Morrow, "Margins in Breast Conservation: A Clinician's Perspective and What the Literature Tells Us," *Journal of Surgical Oncology*, vol. 110, pp. 2-7, 2014, [[10.1002/jso.23594](#)].
- [36] R. J. Gray, B. A. Pockaj, E. Garvey, and S. Blair, "Intraoperative Margin Management in Breast-Conserving Surgery: A Systematic Review of the Literature," *Annals of Surgical Oncology*, vol. 25, pp. 18-27, 2018, [[10.1245/s10434-016-5756-4](#)].
- [37] C. Reyna, and S. M. DeSnyder, "Intraoperative Margin Assessment in Breast Cancer Management," *Surgical Oncology Clinics of North America*, vol. 27, pp. 155-165, 2018, [[10.1016/j.soc.2017.08.006](#)].
- [38] B. W. Maloney, D. M. McClatchy III, B. W. Pogue, K. D. Paulsen, W. A. Wells, and R. J. Barth Jr, "Review of methods for intraoperative margin detection for breast conserving surgery," *Journal of Biomedical Optics*, vol. 23, no. 10, pp. 100901, 2018, [[10.1117/1.JBO.23.10.100901](#)].
- [39] K. Esbona, Z. Li, and L. G. Wilke, "Intraoperative Imprint Cytology and Frozen Section Pathology for Margin Assessment in Breast Conservation Surgery: A Systematic Review," *Annals of Surgical Oncology*, vol. 19, no. 10, pp. 3236-3245, 2012, [[10.1245/s10434-012-2492-2](#)].
- [40] D. W. Shipp, E. A. Rakha, A. A. Koloydenko, R. D. Macmillan, I. O. Ellis, and I. Nottingher, "Intraoperative spectroscopic assessment of surgical margins during breast conserving surgery," *Breast Cancer Research*, 2018, [[10.1186/s13058-018-1002-2](#)].
- [41] G. Thomas, T.-Q. Nguyen, I. J. Pence, B. Caldwell, M. E. O'Connor, J. Giltneane, M. E. Sanders, A. Grau, I. M. Meszoely, M. Hooks, M. C. Kelley, and A. Mahadevan-Jansen, "Evaluating feasibility of an automated 3-dimensional scanner using Raman spectroscopy for intraoperative breast margin assessment," *Scientific Reports*, vol. 7, pp. 13548, 2017, [[10.1038/s41598-017-13237-y](#)].
- [42] S. J. Erickson-Bhatt, R. M. Nolan, N. D. Shemonski, S. G. Adie, J. Putney, D. Darna, D. T. McCormick, A. J. Cittadine, A. M. Zysk, M. Marjanovic, E. J. Chaney, M. G.L., F. A. South, K.

- A. Cradock, Z. G. Lui, M. Sundaram, P. S. Ray, and S. A. Boppart, "Real-time imaging of the resection bed using a handheld probe to reduce incidence of microscopic positive margins in cancer surgery," *Cancer Research*, vol. 75, no. 18, pp. 3706-3712, 2015, [[10.1158/0008-5472.CAN-15-0464](#)].
- [43] J. E. Phipps, D. Gorpas, J. Unger, M. Darrow, R. J. Bold, and L. Marcu, "Automated detection of breast cancer in resected specimens with fluorescence lifetime imaging," *Physics in Medicine and Biology*, vol. 63, pp. 015003, 2018, [[10.1088/1361-6560/aa983a](#)].
- [44] J. D. Meier, H. Xie, Y. Sun, Y. Sun, N. Hatami, B. Poirier, L. Marcu, and D. G. Farwell, "Time-resolved laser-induced fluorescence spectroscopy as a diagnostic instrument in head and neck carcinoma," *Otolaryngology and Head and Neck Surgery*, vol. 142, no. 6, pp. 838-844, 2010, [[10.1016/j.otohns.2010.02.005](#)].
- [45] J. M. Dixon, and C. W. J. Cartlidge, "Excision margins in breast conserving therapy," *Breast Cancer Management*, vol. 6, no. 3, pp. 97-99, 2018, [[10.2217/bmt-2017-0026](#)].
- [46] E. R. St John, R. Al-Khudairi, H. Ashrafiyan, T. Athanasiou, Z. Takats, D. J. Hadjiminas, A. Darzi, and D. R. Leff, "Diagnostic Accuracy of Intraoperative Techniques for Margin Assessment in Breast Cancer Surgery," *Annals of Surgery*, vol. 265, no. 2, pp. 300-310, 2017, [[10.1097/SLA.0000000000001897](#)].
- [47] F. Schnabel, S. K. Boolbol, M. Gittleman, T. Karni, L. Tafra, S. Feldman, A. Police, N. B. Friedman, S. Karlan, D. Holmes, S. C. Willey, M. Carmon, K. Fernandez, S. Akbari, J. Harness, L. Guerra, T. Frazier, K. Lane, R. M. Simmons, A. Estabrook, and T. Allweis, "A Randomized Prospective Study of Lumpectomy Margin Assessment with Use of MarginProbe in Patients with Nonpalpable Breast Malignancies," *Annals of Surgical Oncology*, vol. 21, pp. 1589-1595, 2014, [[DOI 10.1245/s10434-014-3602-0](#)].
- [48] J. Coble, and V. Reid, "Achieving clear margins. Directed shaving using MarginProbe, as compared to a full cavity shave approach," *The American Journal of Surgery*, vol. 213, pp. 627-630, 2017, [[10.1016/j.amjsurg.2016.12.019](#)].
- [49] A. Kupstas, W. Ibrar, R. D. Hayward, D. Ockner, C. Wesen, and J. Falk, "A novel modality for intraoperative margin assessment and its impact on re-excision rates in breast conserving surgery," *The American Journal of Surgery*, vol. 215, pp. 400-403, 2018, [[10.1016/j.amjsurg.2017.11.023](#)].
- [50] M. Morrow, E. A. Strom, L. W. Bassett, D. D. Dershaw, B. Fowble, A. Guiliano, J. R. Harris, F. O'Malley, S. J. Schnitt, S. E. Singletary, and D. P. Winchester, "Standard for Breast Conservation Therapy in the Management of Invasive Breast Carcinoma," *CA: A Cancer Journal for Clinicians*, vol. 52, no. 5, pp. 277-300, 2002.
- [51] K. J. Van Zee, P. Subhedar, C. Olcese, S. Patil, and M. Morrow, "Relationship Between Margin Width and Recurrence of Ductal Carcinoma In Situ: Analysis of 2996 Women Treated With Breast-conserving Surgery for 30 Years," *Annals of Surgery*, vol. 262, no. 4, pp. 623-631, 2015, [[10.1097/SLA.0000000000001454](#)].
- [52] G. E. Ekatah, A. K. Turnbull, L. M. Arthur, J. Thomas, C. Dodds, and J. M. Dixon, "Margin width and local recurrence after breast conserving surgery for ductal carcinoma in situ," *European Journal of Surgical Oncology*, vol. 43, pp. 2029-2035, 2017, [[10.1016/j.ejso.2017.08.003](#)].
- [53] S. C. J. Bosma, F. Van der Leij, E. Van Werkhoven, H. Bartelink, J. Wesseling, S. Linn, E. J. Rutgers, M. J. Van de Vijver, and P. H. M. Elkhuizen, "Very low local recurrence rates after breast-conserving therapy: analysis of 8485 patients treated over a 28-year period," *Breast Cancer Research and Treatment*, vol. 156, pp. 391-400, 2016, [[10.1007/s10549-016-3732-0](#)].
- [54] E. Vos, A. Jager, C. Verhoef, A. C. Voogd, and L. B. Koppert, "Overall survival in patients with a re-excision following breast conserving surgery compared to those without in a large population-based cohort," *European Journal of Cancer*, vol. 51, pp. 282-291, 2015, [[10.1016/j.ejca.2014.12.003](#)].

- [55] M. G. Valero, M. A. Mallory, K. Losk, M. Tukenmez, J. Hwang, K. Camuso, C. Bunnell, T. A. King, and G. Mehra, "Surgeon Variability and Factors Predicting for Reoperation Following Breast-Conserving Surgery," *Annals of Surgical Oncology*, vol. 25, pp. 2573-2578, 2018, [[10.1245/s10434-018-6526-2](#)].
- [56] R. E. Pataky, and C. R. Baliski, "Reoperation costs in attempted breast-conserving surgery: a decision analysis," *Current Oncology (Toronto, Ont.)*, vol. 23, no. 5, pp. 314-321, 2016, [[10.3747/co.23.2989](#)].
- [57] L. E. McCahill, R. M. Single, E. J. Aiello Bowles, H. S. Feigelson, T. A. James, T. Barney, J. M. Engel, and A. A. Onitilo, "Variability in Reexcision Following Breast Conservation Surgery," *Journal of the American Medical Association*, vol. 307, no. 5, pp. 467-475, 2012.
- [58] A. J. Isaacs, M. L. Gemignani, A. Pusic, and A. Sedrakyan, "Association of Breast Conservation Surgery for Cancer With 90-Day Reoperation Rates in NewYork State," *JAMA Surg*, 2016, [[10.1001/jamasurg.2015.5535](#)].
- [59] L. G. Wilke, T. Czechura, C. Wang, B. Lapin, E. Liederbach, D. P. Winchester, and K. Yao, "Repeat Surgery After Breast Conservation for the Treatment of Stage 0 to II Breast Carcinoma A Report From the National Cancer Data Base, 2004-2010," *JAMA Surgery*, vol. 149, no. 12, pp. 1296-1305, 2014, [[10.1001/jamasurg.2014.926](#)].
- [60] S. E. Abe, J. S. Hill, Y. Han, K. Walsh, J. T. Symanowski, L. Hadzikadic-Gusic, T. Flippo-Morton, T. Sarantou, M. Forster, and R. L. White, "Margin Re-excision and Local Recurrence in Invasive Breast Cancer: A Cost Analysis Using a Decision Tree Model," *Journal of Surgical Oncology*, vol. 112, pp. 443-448, 2015, [[10.1002/jso.23990](#)].
- [61] N. H. Lester-Coll, C. E. Rutter, and S. B. Evans, "Cost-effectiveness assessment of lumpectomy cavity boost in elderly women with early stage estrogen receptor positive breast cancer receiving adjuvant radiotherapy," *Radiotherapy and Oncology*, vol. 119, pp. 52-56, 2016, [[10.1016/j.radonc.2016.01.026](#)].

Appendices

English summary

Nederlandse samenvatting

PhD portfolio

Dankwoord

Curriculum vitae

English summary

Intra-operative tools that can accurately evaluate the resection margin during breast cancer surgery are lacking. As a consequence positive resection margins are found during histopathological evaluation when the pathologist assesses the microscopy sections from the margins of the specimen. A positive resection margin is an indication that tumor tissue was left behind during surgery. Patients in which positive resection margins are present require re-excision or additional boost radiation therapy. This causes discomfort and affects cosmetic outcome. Besides the impact of positive resection margins on the patients, the additional treatments will also affect healthcare budgets.

Diffuse reflectance spectroscopy (DRS) is an optical technology that is based on the absorption and scattering characteristics of the tissue. A DRS measurement will contain information regarding the composition and morphology of the tissue, which can be used to discriminate different tissue types. With the ability to measure the optical characteristics of tissue, DRS is a potential candidate to use for margin assessment during breast-conserving surgery and thereby guide the surgeon towards an optimal resection. The goal of this thesis was to address some of the challenges that hamper the translation of this optical technology to the clinic.

Chapter 2

Chapter 2 reviews the current position of optical spectroscopy as a technology that has proven to be able to discriminate different tissue types, but has not reached broad clinical acceptance yet. Some of the key challenges faced by translational researchers are presented and addressed. This chapter does not only concern the application of DRS for margin assessment during surgery, but the more general use of spectroscopic tissue sensing. In this chapter also an outlook is proposed on how translational researchers could proceed to succeed in implementing optical technologies for clinical decision making.

Chapter 3

Chapter 3 involves the first measurements that were performed on *ex vivo* breast tissue specimens at the pathology department. Measurements were acquired of healthy tissue, tissue at the border of the tumor, and tumor tissue. After measuring, once the microscopy section of the tissue was available, a simple registration between the optical measurements and the H&E (hematoxylin and eosin stain) microscopy section was made to be able to confirm that measurements indeed were taken from healthy tissue, tumor tissue, or at the border of the two. The optical measurements were quantified with an analytical fit model that is based on the diffusion theory to extract the fit parameters from the measured spectra. The fit parameters were analyzed and several optical fit parameters were significantly

different in normal tissue and tumor tissue. The fit parameter that describes the ratio between fat and water (F/W-ratio) provided the best discrimination between these tissue types. Although also other fit parameters were useful in discriminating healthy tissue from tumor tissue these were all outperformed by the F/W-ratio.

To account for inter-patient variation the F/W-ratio was scaled by the F/W-ratio of all healthy measurements in the specimen. This scaled F/W-ratio decreased in the border measurements of all but two patients, which is an indication that a transition from normal tissue to tumor tissue could be detected by the F/W-ratio. Based on the measurements of all patients a threshold was defined that could separate tumor tissue from healthy tissue. With this threshold, the measurements acquired in the specimen of one patient were assessed which demonstrated that with DRS the tumor boundary could be detected.

Chapter 4

Chapter 4 reports on how the change in physiology due to resection of the tissue (*in vivo* versus *ex vivo*) may impact DRS measurements, as well as the influence of inter-patient variation. This chapter elaborates on if these factors hamper the use of optical spectroscopy for the detection of tumor tissue and if normalization of the measurements using the patient as its own reference can improve the accuracy.

In this study measurements of both tumor tissue and normal tissue were acquired in 45 patients before (*in vivo*) and after resection (*ex vivo*) of the specimen using a fiber-optic needle. To confirm the needle had been correctly placed in either healthy tissue or tumor tissue during the measurements in the first half of the study a twistmarker was left behind at the measurement location which allowed the pathologist to sample and evaluate the tissue around the measurement location. In the second half of the study, a biopsy was taken with a separate biopsy needle after the DRS measurements were acquired, which was then assessed by the pathologist. Similar to the previous *ex vivo* study the measurements were quantified with the analytical fit model to extract the fit parameters from the DRS spectra. The analysis with the fit parameters was performed with a Generalized Estimating Equation (GEE) model which is a statistical test that allows taking into account relations between individual measurements when assessing the association between a variable (e.g. a fitparameter) and a covariate (e.g. tissue type 'healthy' or 'tumor'). Besides, the tissue type ('healthy' or 'tumor') and tissue status ('*in vivo*' or '*ex vivo*'), patient characteristics such as menopausal status ('pre-/perimenopausal' or 'post menopausal') and treatment with chemotherapy or hormonal therapy prior to surgery ('yes' or 'no') were also included in the analysis as a covariate to be able to assess the influence of inter-patient variation.

It was found that some of the assessed fit parameters were significantly different in healthy tissue and tumor tissue (F/W-ratio, β -carotene (μM), blood (%), reduced scattering at 800 nm, and fraction Mie scattering). Again the F/W-ratio provided the best discrimination between normal and tumor tissue for both the *in vivo* as well as the *ex vivo* measurements.

None of the fit parameters was significantly affected by the moment at which measurements were acquired ('*in vivo*' or '*ex vivo*'). Furthermore, comparing the results of the F/W-ratio for the *in vivo* measurements with the results in the *ex vivo* measurements no difference in performance was found.

As for the menopausal status, no effect on any of the fit parameters was found, neither was there an effect of neoadjuvant chemotherapy although the number of patients which received chemotherapy prior to surgery was limited.

In an attempt to limit the influence of inter-patient variation the fit parameter data were normalized by using the normal measurements of a patient as a reference. However, this did not improve the discriminative power of the optical fit parameters in distinguishing healthy tissue from tumor tissue.

Chapter 5

In Chapter 5 the focus is on how to accurately correlate the optical measurements acquired *ex vivo* with the histopathology of the microscopy section (H&E). The validation of optical technology with this current golden standard is very important but difficult. Processing of the tissue at the pathology department causes deformations in the tissue that can hinder the labeling of the optical measurements with the correct tissue type. Incorrect labeling can lead to inaccurate assessment of the performance of the technology. In this chapter, we present a method that accounts for the deformation of the tissue during pathology processing, and subsequently compare this method (affine + deformable registration) with another registration method (only affine registration) in which no correction for the tissue deformations is applied. The H&E section with correction for tissue deformation and the H&E section without correction for the tissue deformations are both evaluated by comparing the shape and composition of the tissue with an independent modality (micro-computer tomography).

It was found that the H&E section registered with the method that accounted for tissue deformation was more accurate than the method that did not account for these deformations. The result of the affine + deformable registration method was also evaluated by comparing the fat percentage in the measurement volume according to the H&E section with the fat percentage in the DRS measurement. Measurement locations in which the averaged fat percentage according to the H&E section before and after deformable registration changed more than 20% were selected. This group of measurement locations with an altered pathology label ('altered' points) was compared to the locations in which the fat percentage in the H&E section did not change after applying the deformable registration step ('unaltered' points). It was found that in this 'unaltered' group the correlation between the optical measurements and the registered H&E section as registered with the affine registration was not better or worse compared to the correlation between the optical measurements and the registered H&E section that was registered including the affine +

deformable registration method. However, in the 'altered' group the correlation with the optical measurements strongly enhanced after including the deformable registration step. It was concluded that tissue deformations should be taken into account when making a registration between optical measurements and microscopy sections. And, although the presented registration method can be considered as labor intensive, this deformable registration method is helpful in validating the optical measurements.

Chapter 6

With an increasing number of patients receiving chemotherapy prior to their surgery, it is important to research how this might affect the optical measurements. If optical characteristics of tissue alter due to the pre-operative treatment this could hamper the use of DRS for margin assessment in patients that were treated with this neoadjuvant chemotherapy. The previous analysis regarding this topic was performed with a dataset of a limited number of patients (Chapter 4). In the analysis in Chapter 6, a large number of patients was included ($n = 92$). Again a GEE model is used for analysis, but this time datasets are constructed of measurements with similar histopathology (i.e. datasets of 'fat', 'connective', 'connective tissue combined with fat', and 'tumor cells combined with connective'). Besides the pre-operative treatment with chemotherapy also menopausal status was included as a covariate in the models. For the datasets with a mixture of tissue types ('connective combined with fat' and 'connective combined with tumor cells') the percentage connective tissue was also added as a covariate.

The analysis was performed with both the intensities measured at each wavelength between 850 and 1600 nm as well as the fit parameters derived from the spectra. None of the wavelengths had a significantly different intensity measured in the patients with and without chemotherapy in any of the four datasets. These results were confirmed by the analysis of the fit parameters. The fit parameters were not affected by neoadjuvant chemotherapy, except for the fit parameter (collagen) which showed a significant difference in only one dataset that was comprised of measurements of tumor cells in combination with connective tissue. Regarding the influence of the menopausal status, the influence was considered negligible, as similar to the results of neoadjuvant chemotherapy, none of the wavelengths was significant for any of the datasets between the two menopausal status groups. And again, only the collagen fit parameter was significant in the dataset comprised of measurements of tumor cells with connective tissue.

Noteworthy, the percentage of connective tissue in the measurement location did influence the optical measurements in the dataset in which the measurements consisted of a combination of fat tissue with connective tissue as both in the analysis of the intensities as well as the fit parameters significance was found. However, in the dataset in which connective tissue was combined with tumor cells, this was not the case. In this analysis the covariable was not significant. This indicated that discriminating between fat and connective tissue

is possible with the DRS measurements whereas discriminating between tumor cells and connective tissue seems more difficult.

It was concluded that the impact of neoadjuvant chemotherapy on the optical characteristics of breast tissue and breast cancer tissue as measured with DRS is limited and that margin assessment with DRS seems also feasible in patients after neoadjuvant chemotherapy.

Chapter 7

In Chapter 7 two challenges that might influence the ability of DRS technology to detect positive resection margins are addressed. The first challenge considers the detection of ductal carcinoma in situ (DCIS) which can be a precursor of invasive carcinoma. Being able to also identify areas of DCIS is important as this type of malignancy accounts for the majority of positive resection margins. The second challenge involves the accurate classification of locations that are a mixture of tissue types. Breast tissue is inhomogeneous in the sense that healthy tissue predominantly consists of fat tissue, but there can be connective tissue present as well. Meanwhile, connective tissue is also present in breast cancer tumors, which comprise of tumor cells surrounded by connective tissue in various ratios. Additionally, in the measurements that are acquired at the resection margin tumor cells with accompanied connective tissue might be surrounded by fat tissue. For successful resection margin assessment, it is important to consider if measurement locations with a mixture of tissue types can be classified accurately.

To research both these questions, DRS measurements were obtained from 107 *ex vivo* breast surgery specimens. The optical measurements were correlated with histopathology using the registration method that accounts for tissue deformation which was presented in Chapter 5. Furthermore, instead of using the fit parameters or the intensities measured at each wavelength, so-called spectral features were derived from the spectra and used in the analysis. The spectral features described slopes and dips from specific wavelength ranges in the spectrum.

With regard to the detection of DCIS, the results indicated that this tissue type differs from healthy tissue measurements similarly to the difference between invasive carcinoma and healthy tissue. Comparing the measurements of DCIS with the measurements of invasive carcinoma no significant differences were found.

As for the classification of the mix locations, first, we assessed the influence of using different definitions of 'healthy' and 'tumor' on the performance of the classification model. Locations that were not defined as 'healthy' or 'tumor' were considered 'mix' and were only classified by the classification models, and not used for developing the models. Allowing up to 20% of connective tissue in the definition of 'healthy' and allowing up to 20% of fat tissue in the definition of 'tumor' provided excellent performance in the classification of 'healthy' and 'tumor' locations. However, for the classification of the 'mix' locations, the classification

model with the purest data (no connective in 'healthy' and no fat in 'tumor') was best. With an increasing percentage of tumor cells in the measurement volume classification performance improved, and in case of more than 75% tumor cells in a measurement location classification accuracy was excellent.

The results also showed that the percentage of tumor cells in the measurement volume was more important for correct classification than whether those locations contained invasive carcinoma or DCIS. Based on the results it was concluded that detection of areas of invasive carcinoma or DCIS at the resection margin is possible with DRS.

Chapter 8

In this chapter, the results of an *in vivo* study are presented. DRS technology was applied in the clinical workflow during routine breast biopsy procedures. An optical biopsy needle was used that allowed to perform the optical measurements and subsequently, using the same needle, take a biopsy of the tissue that was measured. In 27 patients that underwent an ultrasound-guided breast biopsy procedure optical measurements were acquired from a location at a distance from the suspect lesion and from a location in the lesion. Additionally, in five patients measurements were obtained continuously along the entire biopsy needle trajectory.

Subsequently, classification models were developed to classify the DRS measurements as normal or healthy. The models were based either on the fit parameter data, or the full DRS spectrum, or a subset of wavelengths from the spectrum. The wavelengths were selected by choosing one wavelength with the lowest p-value from wavelength ranges that were significantly different in the spectra from healthy tissue compared to the spectra of tumor tissue. The performance of the classification model that was based on these selected wavelengths demonstrated the best performance.

The continuous measurements of the needle trajectory showed the transition from healthy tissue to tumor tissue. These measurements were classified with the classification models and results showed that in four out of five patients the output of the model based on a selection of wavelengths was in agreement with the pathological outcome. In one patient the output of the classification models indicated normal tissue whereas the biopsy specimen did contain tumor tissue. However, since the specimen was fragmented this might have hampered correct correlation with the tissue volume that was actually measured.

It was concluded that the DRS measurements can be classified with high accuracy based on a selection of wavelengths. Furthermore, the continuous measurements showed that DRS measurements can be acquired real-time and that these measurements can be classified correctly.

Nederlandse samenvatting

Intra-operatieve hulpmiddelen die gebruikt kunnen worden voor betrouwbare marge evaluatie tijdens borst chirurgie zijn momenteel niet beschikbaar. Het gevolg daarvan is dat tijdens het histopathologisch onderzoek van de microscopie coupes door de patholoog positieve resectie marges worden gevonden. Een positieve resectie marge kan erop duiden dat er tumor in het lichaam van de patiënt is achter gebleven en deze patiënten hebben daarom een re-excisie of extra radiotherapie nodig. Dit is oncomfortabel voor de patiënt en beïnvloed het cosmetisch resultaat. Naast de impact die een extra behandeling heeft op de patiënt, drukken extra behandelingen ook op gezondheidszorg budgetten.

Diffuse reflectie spectroscopie (DRS) is een optische techniek die gebaseerd is op de absorptie- en verstrooiingseigenschappen van weefsel. Een DRS meting bevat informatie over de compositie en morfologie van het weefsel, en op basis daarvan is het mogelijk om verschillende weefseltypes te onderscheiden. Met het vermogen om onderscheid te kunnen maken tussen verschillende weefsels is DRS een potentiële optie om marges te evalueren tijdens borst chirurgie en daarmee de chirurg te helpen in het uitvoeren van een optimale resectie. In deze thesis worden een aantal vraagstukken onderzocht die het gebruik van deze technologie voor resectie marge evaluatie tijdens borst chirurgie in de weg staan.

Hoofdstuk 2

Hoofdstuk 2 bevat een review over de huidige positie van optische spectroscopie als technologie die aantoonbaar weefsels kan onderscheiden, maar nog niet breed wordt toegepast in de kliniek. Een aantal van de belangrijkste uitdagingen waar onderzoekers die translationeel onderzoek doen mee te maken krijgen worden gepresenteerd en besproken. Dit hoofdstuk gaat niet exclusief over het gebruik van DRS voor marge evaluatie, maar behandelt het gebruik van spectroscopie in de medische oncologie in het algemeen. Daarnaast wordt besproken hoe translationeel onderzoekers in het optische onderzoeksveld verder te werk zouden kunnen gaan om ervoor te zorgen dat optische technologie succesvol ingezet kan worden in de kliniek.

Hoofdstuk 3

Hoofdstuk 3 gaat over de eerste metingen die verkregen werden van *ex vivo* borst preparaten op de pathologie afdeling. Er werden metingen gedaan op gezond weefsel, weefsel op de rand van de tumor en van de tumor zelf. Na afloop van de metingen, wanneer de microscopische coupes van het weefsel beschikbaar waren, werd een simpele registratie gemaakt tussen de optische metingen en de H&E (hemotoxyline & eosine) microscopie coupe om te bevestigen dat de metingen inderdaad waren verkregen in gezond weefsel, de tumor, of op de rand van de tumor. De optische metingen werden gekwantificeerd met

een analytisch fit model, dat is gebaseerd op de diffusie theorie, om de fit parameters van het gemeten spectrum te bepalen.

De fit parameters werden geanalyseerd en een aantal fit parameters waren significant anders in gezond weefsel ten opzichte van tumor weefsel. De fit parameter die de ratio tussen vet en water beschrijft (F/W-ratio) was het best in staat om gezond weefsel van tumor weefsel te onderscheiden. Ondanks dat andere fit parameters ook bruikbaar waren, was geen enkele fit parameter zo goed als de F/W-ratio. Deze F/W-ratio werd geschaald met behulp van alle F/W-ratio's gemeten in gezond weefsel om het effect van inter-patient variatie weg te nemen. De geschaalde F/W-ratio verminderde in de grens metingen in vergelijking tot de gezond metingen bij alle patiënten op twee na. Dit is een aanwijzing dat met behulp van de F/W-ratio een grens gebied rond de tumor gedetecteerd kan worden. Door analyse van de metingen van alle patiënten werd een drempelwaarde bepaald die bruikbaar was voor het uit elkaar houden van gezond weefsel en tumor weefsel. Met behulp van deze drempelwaarde werden de metingen van het preparaat van één patiënt beoordeeld. De resultaten lieten zien dat het in dit preparaat mogelijk was om de grens van de tumor te zien.

Hoofdstuk 4

In hoofdstuk 4 wordt onderzocht of het uit het lichaam halen van het weefsel invloed heeft op de DRS metingen (dus verschil in *in vivo* en *ex vivo* metingen) en wat de invloed is van inter-patient variatie. Het hoofdstuk gaat in op of het contrast tussen tumor weefsel en gezond weefsel beïnvloed wordt door het moment van meten, en er wordt uitgezocht of het normaliseren van de optische metingen op basis van gezond metingen van de patiënt helpt in het onderscheiden van tumor weefsel en gezond weefsel.

In de studie werden met behulp van een fiber-optische naald metingen gedaan in 45 patiënten voordat het weefsel uit het lichaam gesneden was (*in vivo*) en van het resectie preparaat van de patiënt (*ex vivo*) na afloop van de operatie. Om te kunnen bevestigingen dat daadwerkelijk gezond weefsel of tumor weefsel gemeten was tijdens de metingen werd in de eerste helft van het aantal geïnccludeerde patiënten een twistmarker op de meet locaties achter gelaten. De patholoog onderzocht in het resectie preparaat vervolgens het weefsel rond deze twistmarker. In het tweede deel van de studie werd gelijk na het meten met een bioptnaald op de meet locatie een stukje weefsel gebiopteerd dat later werd beoordeeld door de patholoog.

Vergelijkbaar met de eerdere *ex vivo* studie werden de metingen gefit met het analytische fit model om de fitparameters in de meting te berekenen. De analyse van de fit parameters werd gedaan met een 'Generalized Estimating Equation' (GEE) model. Dit is een statistische test die het mogelijk maak het verband tussen een variabele (in dit geval een fit parameter) en een covariabele (bijvoorbeeld weefsel type, 'gezond' of 'tumor') te onderzoeken waarbij rekening wordt gehouden met de relatie tussen individuele metingen.

Naast het weefseltype ('gezond' of 'tumor') en moment van meten ('*in vivo*' of '*ex vivo*') werden ook covariabelen aan het model toegevoegd die gerelateerd waren aan patiënt karakteristieken zoals menopausale status ('pre-/perimenopausaal' of 'post menopausaal'), en behandeling met chemotherapie of hormonale therapie voorafgaand aan de operatie ('ja' of 'nee').

De resultaten lieten zien dat sommige van de onderzochte fit parameters significant verschillend waren in gezond weefsel en tumor weefsel (F/W-ratio, β -caroteen (μM), bloed(%), gereduceerde verstrooiing op 800 nm en de fractie Mie verstrooiing). Wederom gaf de F/W-ratio het beste het verschil tussen gezond weefsel en tumor weefsel in zowel de *in vivo* metingen als de *ex vivo* metingen. Geen van de fit parameters was significant verschillend in de groep metingen die *in vivo* dan wel *ex vivo* verkregen waren. Daarnaast bleek het vermogen van de F/W-ratio om onderscheid te maken in de *in vivo* metingen en in de *ex vivo* metingen vergelijkbaar.

Er was geen effect van de menopausale status van de patiënt op de fit parameters, net als dat er geen invloed was van chemotherapie op de fit parameters, maar in deze laatste analyse zaten maar een beperkt aantal patiënten.

Door de metingen te normaliseren met de gezonde metingen van een patiënt werd getracht om een eventuele invloed van inter-patient variatie te verwaarlozen. De genormaliseerde fit parameters waren niet beter in staat om onderscheid te maken tussen gezond en tumor dan de niet-genormaliseerde fit parameters.

Hoofdstuk 5

In dit hoofdstuk ligt de focus op het betrouwbaar correleren van *ex vivo* verkregen optische metingen met de histopathologie van de microscopie coupes (H&E). Het valideren van optische metingen met deze gouden standaard is erg belangrijk maar moeilijk. Tijdens het verwerken van het weefsel op de pathologie afdeling vervormd het weefsel. Deze vervorming van het weefsel is een belangrijke reden waarom het complex is om te bepalen welke weefseltypes aanwezig waren in de meetvolumes waarvan DRS metingen zijn verkregen. Het gevolg van het incorrect labelen van meetlocaties met histopathologie kan leiden tot een onbetrouwbare validatie van de technologie. In dit hoofdstuk wordt een methode gepresenteerd die rekening houdt met de vervorming van het weefsel tijdens het verwerken ('affine + deformable' registratie methode). Vervolgens wordt deze registratie methode vergeleken met een registratie methode waarin vervorming niet wordt meegenomen ('affine' registratie methode) en wordt het resultaat van beide registraties geëvalueerd door deze op basis van de vorm en compositie van het weefsel te vergelijken met een onafhankelijke modaliteit (micro-computer tomografie).

De uitkomsten laten zien dat de H&E coupe die geregistreerd was met de methode die rekening houdt met vervorming nauwkeuriger was dan de methode waarin de vervorming niet werd meegenomen. Het resultaat van de registratie met vervorming werd ook

beoordeeld door het vet percentage in de meetlocatie van de H&E coupe te vergelijken met het vet percentage zoals dat gemeten was met DRS. Meetlocaties waarin het gemiddelde vet percentage in de H&E coupes meer dan 20% afweek tussen de registratie zonder vervorming ('affine') en de registratie met vervorming werden geselecteerd ('affine + deformable'). Deze groep van meetlocaties waarin het pathologie label veranderde werden vergeleken met de groep van overige meetlocaties waarin het pathologie label niet veranderde naar aanleiding van het gebruik maken van de registratie methode die corrigeert voor weefselvervorming. Voor deze groep metingen met onveranderd histopathologie label was de correlatie tussen de optische metingen en de H&E coupe geregistreerd met de 'affine' registratie methode net zo goed als de correlatie tussen de optische metingen en de H&E coupe die was geregistreerd met de 'affine + deformable' registratie methode. Dit in tegenstelling tot de correlatie tussen de optische metingen en de H&E coupes voor de meetlocaties waarin het histopathologie label wel veranderde. In deze groep met meetlocaties werd de correlatie aanzienlijk beter door rekening te houden met de vervorming van het weefsel.

Er werd geconcludeerd dat bij het registreren van de H&E coupe met de optische metingen rekening moet worden gehouden met weefselvervorming. Daarnaast werd geconcludeerd dat ondanks dat de gepresenteerde methode arbeidsintensief is deze wel behulpzaam is voor het valideren van de optische metingen.

Hoofdstuk 6

Omdat het aantal patiënten dat voorafgaand aan hun operatie chemotherapie krijgt steeds groter wordt is het belangrijk om te onderzoeken wat de invloed is van chemotherapie op de DRS metingen. Indien de optische karakteristieken van weefsel veranderen als gevolg van de pre-operatieve behandeling, kan dit het gebruik van DRS voor resectie marge evaluatie in patiënten na neoadjuvante chemotherapie bemoeilijken. De eerdere analyse aangaande deze onderzoeksvraag werd gedaan met een dataset met een beperkt aantal patiënten (Hoofdstuk 4). In de analyse in Hoofdstuk 6 werd een groot aantal patiënten geïncludeerd. In deze analyse werd wederom een GEE model gebruikt, echter in plaats van alle metingen bij elkaar te voegen werden vier datasets geconstrueerd van metingen met allemaal dezelfde pathologie (te weten de volgende datasets: 'vet', 'Bindweefsel', 'Bindweefsel gecombineerd met vet' en 'Bindweefsel gecombineerd met tumorcellen'). Zowel neoadjuvante chemotherapie als menopausale status werden in het model geïncludeerd als covariabelen. In de analyse van de datasets waarin bindweefsel gecombineerd was met of vet of tumorcellen werd ook het percentage bindweefsel als covariabele aan het model toegevoegd.

De analyse werd uitgevoerd met zowel de intensiteiten gemeten op iedere golflengte tussen 850 en 1600 nm als de fit parameters verkregen uit de spectra. Geen van de gemeten intensiteiten op de verschillende golflengtes was significant verschillend in de metingen van patiënten met en zonder behandeling van chemotherapie. Dit resultaat werd bevestigd

in de analyse van de fit parameters. De fit parameters werden ook niet beïnvloed door de neoadjuvante chemotherapie, op één fit parameter (collageen) die significant verschillend was in één van de datasets. Wat betreft de invloed van menopausale status op de optische metingen werd geconcludeerd dat deze verwaarloosbaar was omdat vergelijkbaar met de resultaten wat betreft neoadjuvante chemotherapie geen van de golflengtes significant verschillend was in de vergelijking van de verschillende menopausale status groepen. En tevens was alleen de collageen fit parameter significant verschillend in de pre-/perimenopausale groep ten opzichte van de post menopausale status groep.

Het percentage bindweefsel in de meetlocaties had wel invloed op de optische metingen in de datasets waarin bindweefsel gecombineerd was met vet. Deze covariabele was significant in zowel de analyse met de gemeten intensiteiten als de analyse met de fit parameters. Tegelijkertijd werd in de andere dataset met een mix van twee weefseltypes ('bindweefsel gecombineerd met tumorcellen') het tegenovergestelde resultaat gevonden. In deze analyse was de covariabele niet significant. Dit is een aanwijzing dat het moeilijker is om onderscheid te maken tussen bindweefsel en tumorcellen met DRS dan om onderscheid te maken tussen bindweefsel en vet.

Er werd geconcludeerd dat de impact van neoadjuvante chemotherapie op de optische karakteristieken zoals deze gemeten werden in gezond borstweefsel en tumor weefsel beperkt was en dat marge evaluatie met DRS in patiënten na neoadjuvante chemotherapie haalbaar is.

Hoofdstuk 7

In hoofdstuk 7 wordt onderzoek gedaan naar twee uitdagingen die het detecteren van positieve resectie marges door DRS in de weg kunnen staan. De eerste uitdaging is gerelateerd aan het kunnen detecteren van ductal carcinoma in situ (DCIS) wat een voorloper van een invasief carcinoom kan zijn. Het is belangrijk om ook gebieden van DCIS te kunnen detecteren omdat het merendeel van de positieve resectie marges wordt veroorzaakt omdat er DCIS in de marge van het resectie preparaat zit. De tweede uitdaging is het correct kunnen classificeren van locaties die een mix zijn van meerdere weefseltypes. Borstweefsel is erg inhomogeen omdat ondanks het feit dat het voornamelijk uit vet weefsel bestaat daarin ook bindweefsel aanwezig kan zijn. Tegelijkertijd is bindweefsel ook aanwezig in borstkanker, die gekarakteriseerd worden door tumorcellen met daarom heen bindweefsel. Daarnaast is het zo dat in metingen die verkregen worden op de resectiemarge tumorcellen en bindweefsel aanwezig is dat omringd wordt door vet weefsel. Om succesvol een marge te kunnen beoordelen is het daarom belangrijk om te onderzoeken of meetlocaties die uit een combinatie van weefsel types bestaan correct geïdentificeerd kunnen worden.

Om deze beide onderzoeksvragen te kunnen beantwoorden werden DRS metingen gedaan op 107 *ex vivo* borst preparaten. De optische metingen werden gecorreleerd met histopathologie zoals beschreven in Hoofdstuk 5. In deze analyse werden geen fit

parameters of intensiteiten bij bepaalde golflengtes gebruik, maar zogenaamde spectrale features. Deze spectrale features zijn een alternatieve manier om de meting te kwantificeren door het spectrum te beschrijven aan de hand van hellingen en 'dips' in specifieke golflengtegebieden.

Met betrekking tot het kunnen detecteren van DCIS lieten de resultaten zien dat dit weefsel type van gezond weefsel verschilt op dezelfde manier als invasief carcinoom van gezond weefsel verschilt. Daarnaast werden er geen significante verschillen gezien tussen de metingen van DCIS en invasief carcinoom.

Met betrekking tot het classificeren van de mix locaties werd eerst onderzocht wat de invloed was van het aanpassen van de definitie van 'gezond' en 'tumor' op de prestaties van het classificatie model. Meetlocaties die buiten de definitie van 'gezond' en 'tumor' vielen werden samengevoegd in de dataset 'mix', en deze werden niet gebruikt voor het ontwikkelen van de modellen, maar alleen om geclassificeerd te worden door de modellen. Classificatie modellen presteerde goed wanneer tot maximaal 20% bindweefsel in de 'gezond' metingen aanwezig was en tot maximaal 20% vet in de 'tumor' metingen. Echter, voor het classificeren van de 'mix' locaties was het model met de puurste data (geen bindweefsel in 'gezond' en geen vet in 'tumor') het best. Wanneer er meer tumorcellen in de meting aanwezig waren werd het classificatie model accurater, en wanneer er meer dan 75% tumorcellen aanwezig waren was het resultaat van het classificatie model perfect.

De resultaten lieten ook zien dat het percentage tumor cellen in het meetvolume belangrijker was voor het correct classificeren van de metingen dan of deze locaties bestonden uit invasief carcinoom cellen of DCIS cellen. Gebaseerd op deze resultaten werd geconcludeerd dat het mogelijk is om met DRS gebieden van invasief carcinoom of DCIS op de resectie marge te detecteren.

Hoofdstuk 8

In dit hoofdstuk worden de resultaten van een *in vivo* studie gepresenteerd. DRS werd gebruikt in de klinische workflow tijdens een procedure waarbij een biopt werd afgenomen bij patiënten met een tumor in de borst. Daarvoor werd een optische bioptnaald gebruikt die het mogelijk maakte om DRS metingen te doen en vervolgens met deze zelfde naald een biopt te nemen van het gemeten weefsel. In 27 patiënten die een biopsie procedure ondergingen werden optische metingen verkregen van een locatie op een afstand van de tumor en van een locatie in de tumor. Daarnaast werden in vijf extra patiënten continue metingen gedaan langs het hele traject dat de bioptnaald aflegde.

Vervolgens werden er classificatie modellen gemaakt om de metingen te classificeren als 'gezond' of 'tumor'. Deze classificatie modellen werden gebaseerd op of de fit parameter data, of de intensiteiten gemeten over alle golflengtes, of op basis van de gemeten intensiteiten bij een paar golflengtes. Deze golflengtes werden geselecteerd door in golflengte gebieden die significant verschilden tussen gezonde metingen en tumor

metingen, de golflengte te kiezen met de laagste p-waarde. Het model dat gemaakt was met deze geselecteerde golflengtes presteerde het best.

De continue metingen verkregen van het traject van de naald werden ook geclassificeerd met de modellen en het resultaat liet zien dat in vier van de vijf patiënten de output van het classificatie model gebaseerd op de geselecteerde golflengtes in overeenstemming was met de histopathologie. In één patiënt gaf het model aan dat de naald in gezond weefsel was geplaatst terwijl er volgens de patholoog tumor in het biopt zat. Dit biopt was echter beschadigd wat een mogelijke verklaring voor het gevonden resultaat kan zijn.

Er werd geconcludeerd dat de DRS metingen met grote betrouwbaarheid geclassificeerd konden worden op basis van een selectie van golflengtes. Daarnaast was het mogelijk om real-time DRS metingen te verkrijgen en deze vervolgens correct te classificeren.

PhD portfolio

PHD PERIOD: 2013-2019

Phd Supervisors: T.J.M. Ruers, B.H.W. Hendriks, M.-J.T.F.D. Vrancken Peeters

LIST OF PUBLICATIONS

L. L. De Boer, B. G. Molenkamp, T. M. Bydlon, B. H. W. Hendriks, J. Wesseling, H. J. C. M. Sterenborg, and T. J. M. Ruers, "Fat/Water ratios measured with diffuse reflectance spectroscopy to detect breast tumor boundaries," *Breast Cancer Research and Treatment*, vol. 152, no. 3, pp. 509-518, 2015, [10.1007/s10549-015-3487-z].

L. L. De Boer, J. W. Spliethoff, H. J. C. M. Sterenborg, and T. J. M. Ruers, "Review: *in vivo* optical spectral tissue sensing—how to go from research to routine clinical application?," *Lasers in Medical Science*, vol. 32, no. 3, pp. 711-719, 2016, [10.1007/s10103-016-2119-0].

L. L. De Boer, B. H. W. Hendriks, F. Van Duijnhoven, M.-J. T. F. D. Vranken Peeters-Baas, K. Van de Vijver, C. E. Loo, K. Jóźwiak, H. J. C. M. Sterenborg, and T. J. M. Ruers, "Using DRS during breast conserving surgery: identifying robust optical parameters and influence of inter-patient variation," *Biomedical Optics Express*, vol. 7, no. 12, pp. 5188-5200, 2016, [10.1364/BOE.7.005188].

J. W. Spliethoff, L. L. De Boer, M. A. Meier, W. Prevoo, J. De Jong, T. M. Bydlon, H. J. C. M. Sterenborg, J. A. Burgers, B. H. W. Hendriks, and T. J. M. Ruers, "Spectral sensing for tissue diagnosis during lung biopsy procedures: the importance of an adequate internal reference and real-time feedback," *Lung Cancer*, vol. 98, pp. 62-68, 2016, [10.1016/j.lungcan.2016.05.019].

L. L. De Boer, T. M. Bydlon, F. Van Duijnhoven, M.-J. T. F. D. Vranken Peeters, C. E. Loo, H. A. O. Winter-Warnars, J. Sanders, H. J. C. M. Sterenborg, and T. J. M. Ruers, "Towards the use of diffuse reflectance spectroscopy for real-time *in vivo* detection of breast cancer during surgery," *Journal of Translational Medicine*, vol. 16, no. 367, 2018, [10.1186/s12967-018-1747-5].

E. Kho, L. L. De Boer, K. K. Van de Vijver, F. Van Duijnhoven, M. T. F. D. Vranken Peeters, H. J. C. M. Sterenborg, and T. J. M. Ruers, "Hyperspectral imaging for resection margin assessment during surgery," *Clinical Cancer Research*, vol. accepted, 2019, [10.1158/1078-0432.CCR-18-2089].

L. L. De Boer, E. Kho, J. Nijkamp, K. K. Van de Vijver, H. J. C. M. Sterenborg, L. C. Ter Beek, and T. J. M. Ruers, "A method for co-registration of optical measurements of breast tissue with histopathology: the importance of accounting for tissue deformations," *Submitted to Journal of Biomedical Optics*, 2019.

L. L. De Boer, E. Kho, K. K. Van de Vijver, M. T. F. D. Vranken Peeters, F. Van Duijnhoven, B. H. W. Hendriks, H. J. C. M. Sterenborg, and T. J. M. Ruers, "Fiber-optic DRS measurements of DCIS and classification of locations with mixed tissue types," *In preparation*, 2019.

L. L. De Boer, E. Kho, K. Jóźwiak, K. K. Van de Vijver, M. T. F. D. Vranken Peeters, F. Van Duijnhoven, B. H. W. Hendriks, H. J. C. M. Sterenborg, and T. J. M. Ruers, "The influence of neoadjuvant chemotherapy on DRS spectra of breast surgery specimens," *In preparation*, 2019.

PRESENTATIONS

2014 Chirurgendagen, Veldhoven
2014 European Society of Surgical Oncology, Liverpool
2014 European Photonics Industry Consortium Biophotonics Workshop, Rotterdam
2016 European Society of Surgical Oncology, Krakow
2017 SPIE Photonics West, San Francisco
2017 Varioscopic Erasmus MC, Rotterdam
2017 European Photonics Industry Consortium Biophotonics Workshop, Amsterdam
2018 SPIE Photonics West, San Francisco
2018 Gordon Research Seminars, Boston
2019 SPIE Photonics West, San Francisco

POSTERS

2013 European Cancer Congress Amsterdam
2014 European Society for Medical Oncology, Madrid
2018 Gordon Research Conference, Boston

TRAINING

2013 OOA retreat
2015 Biophotonics Summerschool, Ven, Sweden
2016 Scientific writing and presenting
2017 Good Clinical Practice
2018 Grant writing

Dankwoord

De afgelopen jaren heb ik de mogelijkheid gehad om veel te leren, zowel vakinhoudelijk als op persoonlijk vlak. Ik ben erg veel dank verschuldigd aan de mensen die het mogelijk gemaakt hebben om dit proefschrift te schrijven en aan de mensen die mij vanaf de zijlijn aanmoedigden.

Prof. Dr. Ruers, beste Theo, bedankt voor deze promotie en de kansen die je mij hebt gegeven om mezelf op zoveel vlakken te ontwikkelen. Ik heb veel geleerd van het optimisme en de voortvarendheid waarmee jij te werk gaat. Daarnaast heb ik het erg gewaardeerd dat er veel ruimte was voor discussie, en dat daarbij ook vaak gelachen werd.

Prof. Dr. Hendriks, beste Benno, ondanks dat de belangen van het bedrijfsleven en het ziekenhuis soms weleens uiteenliepen heb ik onze samenwerking altijd als erg prettig ervaren. Jouw feedback was altijd 'spot on' en heeft mij vaak aan het denken gezet. Je hebt mij veel kennis gegeven over hoe complex het eigenlijk is om de stap te zetten van onderzoek naar klinische praktijk.

Dr. Vrancken-Peeters, beste Marie-Jeanne, als mammachirurg ben jij eigenlijk het startpunt van het project. Jouw klinische blik op de vraagstelling zette het klinische probleem in perspectief en heeft me veel bijgebracht over de voorwaarden waaraan een oplossing moet voldoen om in de kliniek toepasbaar te zijn. En ondanks een overvolle agenda wilde je altijd mee denken over het project.

Geachte leden van de promotiecommissie, hartelijk dank voor de tijd die jullie in het bestuderen en beoordelen van dit proefschrift hebben gestoken. Prof. Dr. Sterenborg, beste Dick, bedankt voor je dagelijkse begeleiding en het delen van een schat aan ervaring in dit onderzoeksgebied.

Beste collega's en oud-collega's, jullie bijdrage aan deze promotie zit hem zowel in de inhoud als in de vorm. Enerzijds hebben jullie kritische vragen en de gevoerde discussies eraan bijgedragen dat de kwaliteit van het werk beter is geworden. Anderzijds was er ook tijd en ruimte voor gezelligheid waardoor ik met zo veel plezier naar mijn werk ging. Daarnaast wil ik ook graag alle mensen van Philips bedanken voor hun ondersteuning en waardevolle inbreng in dit project en niet te vergeten alle mensen van de pathologie, radiologie, klinische fysica en de OK voor alle inzet en praktische hulp.

Lieve paranimfen, dank voor jullie hulp bij mijn promotie! Anne-Sophie, jou ken ik vanaf de eerste dag van de introductie bij Technische Geneeskunde. Ik denk met een glimlach op mijn gezicht terug aan borrels, feestjes, vakanties en etentjes met jou en de rest van 17 die daarop volgden. Wie van ons had dat ooit verwacht toen we nog lieve sjaarsjes waren? Esther, met jou heb ik het laatste deel van mijn promotie met heel erg veel plezier samengewerkt. Dank voor al je inspanningen, onder andere tijdens al die ellselange meetdagen, en ook alle lol daar omheen.

Lieve vrienden en familie, jullie zijn met te veel om met naam en toenaam te noemen, en het risico op die manier iemand te vergeten wil ik niet lopen. Deze alinea wijd ik aan jullie met een korte lofzang over jullie bijdrage. Stuk voor stuk hebben jullie allemaal een steentje of kei bijgedragen. Ik heb met jullie genoten, gelachen, gegeten, gedronken, gefeest, (vakantie) gevierd, gefietst, gerend, gepraat, gediscussieerd en gedanst. Jullie waren er in opgewekte en verdrietige tijden, met zowel woorden als daden. Dank daarvoor! Jullie bijdrage mag niet onderschat worden want zonder jullie waren de afgelopen jaren een stuk minder leuk geweest.

Lieve Eric, deze alinea gaat geen recht doen aan waarvoor ik je allemaal zou moeten bedanken. Hoe ook te beginnen als iemand je op de wereld heeft gezet en groot heeft gebracht. Laat ik toch een poging wagen: ik ben jou en Marja, heel erg dankbaar voor een onbezorgde jeugd met veel vrijheid en veel uitdaging om dingen te leren en begrijpen. Dit heeft er ongetwijfeld aan bijgedragen dat ik ben gaan promoveren en nu dit proefschrift ga verdedigen. Ik heb mij altijd volop gesteund gevoeld en ben blij dat we het behalen van deze titel, waarin jij ook een groot aandeel hebt, samen kunnen vieren.

Lieve Ivo, waar ik soms twijfelde had jij rotsvast vertrouwen in mij en een goede afloop. Dit oneindige vertrouwen gecombineerd met jouw onbezorgde blik op de wereld heeft mijn bijzonder goed geholpen de afgelopen jaren. Jouw steun kwam in vele vormen door bijvoorbeeld een spreekwoordelijk duwtje in de rug, door te helpen met praktische zaken of door wat bemoedigende woorden en niet te vergeten: woordgrappen. Deze laatste categorie ook vaak in herhaalde vorm en onder het motto "10 uit 8", maar altijd uit een goed hart en dat is eigenlijk het enige dat telt. (Jouw beurt. Ik voel dat hier meer in zit.)

Curriculum vitae

Lisanne de Boer was born on February 20th in 1988 in Amsterdam. She grew up in Purmerend where she went to primary and secondary school. With both an interest in Medicine as well as physics she went to Enschede to study Technical Medicine at the University Twente. After finishing her bachelor she specialized in her masters in 'Robotics and Imaging'. During her master program she fulfilled five internships in four different hospitals. The focus of these internships was on both clinical activities and research. Her final internship was in the Netherlands Cancer Institute where she proceeded with a PhD project.

After her PhD she will work as a postdoctoral researcher in the Image-guided surgery research group.