RfPB project plan for testing reliability and diagnostic accuracy of Fast Raman device

Summary

The RfPB Fast Raman study of diagnostic accuracy will start on September 26th, 2022 and involves the measurement of the first Mohs resection layer from a minimum of 122 patients undergoing Mohs surgery at the NUH Treatment Centre for treatment of BCC on the face – head – and neck area. This is a prospective non-inferiority single-centre study designed to determine the diagnostic accuracy of the Fast Raman instrument for the detection of residual basal cell carcinoma on tissue specimens removed through Mohs micrographic surgery. During this study, the reliability of the Fast Raman instrument, which shows to what extent it is prone to between-operator variations, will also be investigated. A statistical sample size calculation suggests that a sample size of 122 patients will provide an estimate of sensitivity of 96% with (95%CI 0.81, 0.99). In order to measure 122 patients for the study, an average of 4 patients would need to be investigated per week.

Study design

The proposed study will contain a measurement of instrument reliability and a measurement of instrument diagnostic accuracy. During the course of the study, we will measure tissue samples comprising of full Mohs tissue layers from 122 patients. With an observed BCC-positive rate at our recruitment centre (NUH Nottingham Treatment Centre) of ~22%, we expect to investigate 27 BCC-positive layers. This sample size will allow us to predict the sensitivity of the Fast Raman device of 96% with (95%CI 0.81, 0.99).

Measurement of instrument reliability

Instrument reliability will be tested by investigating the same tissue blocks on replicate times. Twenty specimens will be utilised to determine inter-user variability. To assess inter-user variability of the Fast Raman instrument, the measurement of each of the 20 samples will be performed twice, with the same specimen being loaded by a different surgeon. The two users are not allowed interfere with each other's measurements. Ideally, the specimens investigated in repeat measurements for the reliability test would be re-measured at a later date as a new measurement. This is not possible during the current study as the tissue specimen is processed and destroyed during frozen section histopathology. Therefore the repeat measurements need to be performed immediately after the initial measurements in order to minimise the delay for the patient's surgical procedure.

A measurement entails placing the tissue samples in the sample cassette, closing the cassette lid, adjusting the cassette lid tightness, placing the cassette within the instrument and starting a measurement. The cassette lid was designed to apply different amounts of pressure to different regions within the cassette. A set of nine grub screws within the lid can be tightened or loosened depending on the thickness of a tissue layer at a certain region. This can help

increase the amount of tissue surface area that is in contact with the cassette window, without overly squeezing thicker regions of the sample.
Sample positioning within the cassette will be decided by each user so that the familiarity that the Raman operator has developed with the tissue samples does not lead to bias in the positioning of the samples into the Raman cassettes for the second measurement.

Sample pre-processing for Fast Raman requires surgeons to knick the "full-face" tissue the specimens to preserve orientation, immerse them in RBC lysis buffer, blot them to remove superficial blood and then use coloured marker pens to mark the edges of the resulting tissue samples. Loading the samples into the cassette will be performed by the surgeon. After Fast Raman measurements, the surgeons will ink the tissue specimens according to the markings on the cassette coverslip and to send the samples to be processed for frozen section histopathology.

Specimen blotting for blood removal is performed as follows: The sample is blotted with tissue paper in order to remove the majority of the superficial blood. Blotting consists of applying downwards pressure with a finger on the epidermal side of the sample, with the sample being placed between two layers of tissue paper. The sample is then immersed in 1x RBC lysis solution for 10 seconds. This step aims to burst red blood cells, increasing the effectiveness of blotting. The sample is then blotted with tissue paper again by firmly pressing on the middle and around the edges on the sample's epidermal side. The surgeon will alternate between immersion in RBC lysis buffer and blotting for a minimum of 10 times and a maximum of 25 times. After a minimum of 10 blots, blotting may stop if no further blood imprints can be observed on the tissue paper or if there is risk of damaging the tissue. The specimen is then immersed in saline and blotted one more time. A more detailed description of the sample pre-procession procedure can be found in the instrument operation and sample preparation protocol (Fast Raman measurement sample preparation protocol v2.docx). The blotting paper used for each sample will be photographed and logged to investigate the effect of blotting efficacy on the Fast Raman results.

During the measurement procedure, user-induced variation can be introduced in two ways:

- Via sample pre-processing (which aims to remove superficial blood from the sample); this task is currently performed by the surgeon. As blood removal can be only performed once per specimen, this step will not be included in the intra-user variability assessment.

- Via placing the tissue sample inside the cassette (ensuring that the entire resection surface is in contact with the coverslip), tightening the cassette lid screws and placing the sample cassette within the instrument; performed by the instrument operator and the surgeon. These steps are included in the intra-user variability assessment.

In order to mitigate user-induced variation, a measurement checklist (Fast Raman measurement sample preparation protocol v2.docx) is provided which guides the user through tissue pre-processing. An accompanying video guide which shows a full measurement procedure will also be provided.

Measurement of instrument diagnostic accuracy

The validity of the instrument entails the comparison between the diagnostic results provided by the Fast Raman instrument and the results following the histopathologic evaluation of the

frozen H&E sections, which acts as the reference standard for the study. Validity is measured using two primary parameters: sensitivity and specificity. For the current application, sensitivity of detection is the most important performance metric of the instrument. It is of utmost importance to ensure that the entirety of the tumour is removed during the surgical procedure. A lower specificity is permitted, as a false positive detection would only induce the removal of an extra tissue layer. This is preferable to having residual tumour still present in the lesion after the operation. A range of other parameters such as positive and negative predictive values and likelihood ratios are derived from the sensitivity and specificity. In order to ensure that the results of Mohs are not influenced by Fast Raman and vice-versa and that index test bias is kept to a minimum, the Raman operator will perform the index technique first, before the H&E sections are produced. Afterwards, the surgeon will perform the histopathologic diagnosis whilst not aware of the results of the index test. The reference standard will be produced at a later date by consensus histopathological assessment provided by two experienced dermatopathologists. In case of disagreement, an external dermatopathologist will investigate the discordant H&E slides and will help reach a consensus, providing the final reference standard. Further details on how the reference standard is produced can be found in the histological evaluation protocol (Histological evaluation of frozen HE.docx).

The sensitivity of the instrument will be calculated as the probability of a positive measurement for samples which have been determined as BCC positive by the reference standard. The specificity of the instrument will be calculated as the probability of a negative measurement for samples which have been determined as BCC negative by the reference standard. It was previously observed that tissue samples that contain muscle (such as samples from periocular lesions) may retain large amounts of blood that may not be removed via blotting without damaging the tissue. Such specimens continue to leak blood during the course of the Fast Raman measurement, which could result in cancer cells being covered by blood and not measured by the Fast Raman device. This is a known limitation of the sample pre-processing procedure, but it does not represent a limitation of the Fast Raman device. As there may be instances when such samples produce false negative Fast Raman results, the sensitivity and specificity of Fast Raman will be calculated twice: once when these specimens are included and once when these specimens are excluded. In doing so, we will determine the impact of such specimens on the diagnostic capability of the device.

What will be regarded as a correct Raman detection?

- In the case of a BCC negative specimen, a Raman result is correct if there are no red segments detected within the Raman map and incorrect if there is at least one red segment within the Raman map.

- In the case of a BCC positive specimen, a Raman result is correct if there is at least a single red segment detected within the Raman map and incorrect if there are no red segments within the Raman map.

In order to reduce bias when comparing the Fast Raman result to the reference standard and to ensure that the neither of the two techniques is influenced by the other, an automated software will perform the comparison. The Fast Raman results are displayed as a tissue map wherein BCC is depicted as red segments. After the Fast Raman measurement, the outline of the investigated specimens will be printed out and handed to the surgeon and the two

dermatopathologists. This map contains no diagnostic information or features other than the outline of the tissue specimens. The surgeon and the two dermatopathologists, following a histopathologic assessment, will annotate the outline map with the location where they suspect cancer cells are present at the resection surface and the approximate size of the tumours. The annotated maps will be digitised and registered to the Fast Raman result. An automated algorithm will compare the Fast Raman results to the reports that comprise the reference standard.

The validity of the instrument can be calculated on a per layer basis or on a per patient basis. The statistical analysis will be performed on a per layer basis, as it best retains the localisation of the tumour within the lesion and is therefore truer to a real-life scenario application of the technique.

Reduction of bias and data integrity

In order to ensure that the study presents a highly accurate assessment of the performance of the Fast Raman instrument in a clinical setting, the study design should attempt to minimise biases. The main biases that could be encountered in a diagnostic tests of accuracy are: selection bias, verification bias, incorporation bias, observer bias and differential reference bias.

Selection bias

Selection bias refers to bias induced by the criteria used in the selection of recruited patients. In the present study, a BCC lesion will be considered for enrolment using predefined inclusion and exclusion criteria:

Patient inclusion criteria and sample size

A cohort of 122 consenting patients undergoing Mohs surgery at NUH Treatment Centre for treatment of BCC on face – head – and neck will be included in the present study.

As part of the study, the entire first Mohs resection layer will be measured with the Fast Raman instrument followed by assessment by frozen section histopathology. All specimens smaller than 2 cm x 2 cm (size permitted by the tissue cassette) will be considered to be included in the study regardless of anatomical site. Specimens from up to three patients will be investigated via Fast Raman during each measurement day. This will ensure the recruitment target is reached while at the same time minimising possible delays caused by the Fast Raman measurement procedure. The surgical excisions will be performed by four surgeons of varying degrees of experience: Dr Sandeep Varma, Dr Ashish Sharma, Dr Anand Patel and Dr Sunita Odedra.

Participating patients with multiple BCC lesions may have more than one tissue layer included (first Mohs layer in all cases), provided the Fast Raman measurements are not expected to cause delays to the standard of care.

Patient exclusion criteria

The first patient during each surgery day will not be recruited. Their specimens will be sent directly to the histopathology lab, to reduce delays to the surgery day. After patients have

consented to participate in the study, a decision will be made of whether their specimens are suitable to be included. Patients will not be included in the study if their tissue specimens have the following characteristics:

- Patients whose tissue layers are larger than 2 cm in diameter will not be analysed, as they do not fit within the sample cassette.

- Patients with samples that could not be blotted at least 10 times (as described above) due to concerns regarding tissue integrity will not be analysed.

- Patients with samples for which the quality of the histopathology section is not sufficient to allow for a definitive reference standard. In order to reduce bias, this decision will be made by agreement of three experienced dermatopathologists. To reduce bias, this decision will be made whilst blinded towards the Fast Raman results.

Verification bias

Verification bias occurs when only some of the patients who have the index test go on to have the reference standard. This is not the case for this study as all samples which will be measured with the Fast Raman device will be sent to be processed through frozen section histopathology.

Incorporation bias

Incorporation bias occurs when the reference standard is defined in part by a positive index test. This is also not applicable to the study as the reference standard will be performed regardless of the results of the index test. Moreover, the dermatopathologists that analyse the frozen section histopathology slides as part of the reference standard will be blinded towards the results of the index test.

Observer bias

Observer bias can occur when the person responsible of interpreting the results of the reference standard knows the results of the index test. In order to limit bias, the dermatopathologists will interpret the frozen section histopathology slides whilst being unaware of the Fast Raman results. The dermatopathologists will annotate a diagram representing an outline of the specimen. The diagram will then be compared with the Fast Raman result at the end of the study by an automated algorithm.

Differential reference bias

Differential reference bias occurs when patients that are investigated with the index test are then being treated with different reference standards. This is not the case for the present study, as all samples measured with the Fast Raman instrument will be processed for frozen section histopathology.

Data integrity

In order to reduce data analysis bias and show that results are not manipulated by the research team, a procedure is put in place to ensure the integrity of the data collected during the Fast

Raman diagnostic test of accuracy. The four aspects of the study that need to be scrutinized are the integrity of the files generated during the Fast Raman measurements, the inclusion of all performed measurements in the study, the reliability of the reference standard and the calculation of instrument performance.

File integrity

Within each patient folder, there are 44 files generated as part of a Fast Raman measurement. The creation date and last modified date for these files are logged by the operating system. As these files are created in sequence, in a set order, the manipulation of one or more of these files can be detected by evaluating the consistency of these two time-stamps within the dataset. A tool may be created to check whether the files generated during a measurement were indeed created (and not modified) in the correct order.
After all Fast Raman measurements are completed in a day, the files will be uploaded on a secure university online research drive (or Onedrive), where the upload date is automatically logged and cannot be modified. This process will ensure that the files are stored securely and that the files are not altered while they are stored on the drive.

Patient omission

In order to ensure that all Fast Raman measurements are included in the study and no patients are omitted, H&E report maps completed by the surgeon will be time-stamped at the time of the H&E analysis. This will show whether the Fast Raman files have been uploaded to the secure drive prior to this assessment.

Reference standard

The reference standard will be produced by assessment of the frozen H&E sections by two dermatopathologists. The dermatopathologists will each annotate a reporting map whilst blinded to the results of the Fast Raman measurements and to each other's assessment for the duration of the study. The Mohs surgeon will also annotate a reporting map, which will reflect the Mohs surgeon report. The surgeon will also be blinded towards the results of the Fast Raman measurements for the duration of the study. Therefore, the results of frozen section histopathology cannot be influenced by the results of Fast Raman measurements. Further details on how the reference standard is produced can be found in the histological evaluation protocol (Histological evaluation of frozen HE.docx).

Calculation of instrument performance

For the calculation of instrument accuracy at the end of the study, the entire dataset will be downloaded from the online drive and an automated analysis algorithm will be utilised to perform the calculation of instrument performance. In order to prevent data entry errors, an outside 3rd party will be tasked with checking the automated algorithms used for the calculation of instrument performance and will perform a replicate calculation of the instrument performance metrics.

Imperfect reference standard

Frozen section histopathology as performed for Mohs is an imperfect reference standard for our study as it does not investigate the true resection surface in one image, rather it sections through the sample in 100 μm intervals, likely being slightly biased towards providing positive diagnoses.

Surgeons are encouraged to be cautious when performing Mohs, in order to maximise probability of complete tumour removal. Therefore there may be instances where the resection surface is clear of BCC but the specimen is deemed BCC positive, such as in instances where there is uncertainty when interpreting the H&E sections.

As such, the reference standard for the study will be produced by consensus histopathological assessment provided by two experienced dermatopathologists (NUH Nottingham Treatment Centre) rather than the diagnosis produced by the Mohs surgeon. In case of disagreement, an external dermatopathologist will investigate the relevant H&E slides and will help reach a consensus, providing the final reference standard.

One concern is that there may be instances when there is no residual BCC on the resection surface of the investigated sample, but the tumour is brought to the surface by shavings performed by the Mohs technicians when processing the frozen samples. In such an instance, the instrument would correctly identify a lesion as being BCC negative, while histopathology would identify it as being BCC positive. In order to identify such possible occurrences, the frozen section H&E technicians will log the thickness of tissue that is removed from each layer before the first H&E section is obtained.