

## Executive Summary:

The survival chances of cancer patients can be dramatically increased by an early diagnosis. This also holds true for breast and prostate cancer, which are among the most frequent forms of cancer in Europe. However, with currently available methods, typically core needle biopsies, it is not possible to extract the necessary amount of sample material for a reliable diagnosis without running a high risk of severe side effects, not the least cancer cell seeding and risk of spread of the cancer to be diagnosed. As an alternative, fine-needle aspiration (FNA) cell sampling offers a patient friendly, minimally invasive, rapid and cost efficient tumour diagnostic procedure with minor side-effects. However, the amount of sample obtained by FNA is often sparse, and consists of aspirated cells taken out of their tissue context. It is therefore often difficult to obtain a conclusive diagnosis based upon this sample material.

In this project, we have addressed this conflict between the amount of sample material needed to obtain conclusive diagnostic information and the severe side effects that can be caused by the diagnostic sampling itself. The overall objective of this project was to develop and validate a quantitative, minimally invasive diagnostic tool for early and conclusive detection, diagnosis and monitoring of disease progression of breast and prostate cancer. A methodology was developed making use of a combination of exciting recent advances in the field of fluorescence-based microscopy, for imaging of individual FNA-sampled cells. The methodology includes advances taking the spatial resolution far beyond the fundamental limits of optical resolution, the sensitivity down to an ultimate single-molecule level, and multi-parameter detection schemes significantly increasing the fluorescence information by which these cellular images can be analysed. Apart from detecting and identifying specific protein tumour markers in the samples, tumor-specific spatial distribution patterns of the proteins within intact sample cells were also exploited. This represents a to-date almost unexploited dimension of diagnostic information. The novel optical methods were supported by state-of-the-art affinity molecule biotechnology, fluorophore chemistry, and proteomics-based biomarker identification, with the overall aim to extract a maximum amount of information out of the small amounts of sample material obtained by FNA.

Following extensive effort in optimising the sample preparation, handling and labelling, and corresponding effort to optimise the performance of the fluorescence imaging techniques, first on cultured cells as models and then on cells from clinical FNA samples, a larger number of FNA samples were analysed with a combination of ultrahigh resolution fluorescence microscopy and imaging based on multi-parameter fluorescence detection. New bioinformatic tools were developed and used to analyse the acquired images and features and classifiers were established which in an objective manner could classify the imaged samples as normal or malignant. In parallel, progress in FNA sampling technology performed within the project led to further reduced risks for cancer cell seeding and increased the ability to obtain representative diagnostic material. Taken together, we can thus demonstrate an intact chain of accomplished steps, from reliable and patient-friendly FNA sampling with minimized risk of tumor cell dissemination, over robust and reproducible sample handling and preparation, specific targeting of selected proteins with fluorophore-labelled affinity molecules, characterization of the FNA sampled cells by optical methods offering utmost sensitivity, specificity and spatial resolution, and finally bioinformatic analyses to identify the differences in the image features necessary to distinguish normal from malignant cells. This represents a novel diagnostic concept, which can combine diagnostic reliability with safe, minimally invasive diagnostic sampling and thus enable a decisive improvement of the outcome of patients suffering from breast and prostate cancer.

## **Project Context and Objectives:**

The survival chances of cancer patients can be dramatically increased by an early diagnosis. This particularly holds true for breast and prostate cancer, which are among the most frequent forms of cancer in Europe. For these forms of cancer, the dominating methodology for diagnostic sampling is based on core needle biopsies. The classification of the obtained tissue samples on the basis of morphology requires relatively large tissue samples and necessitates expert knowledge which has to be acquired over long training periods and in the end remains subject to individual judgement. With currently available methods, typically core-needle biopsies, it is not possible to extract the necessary amount of sample material for a reliable diagnosis without running a high risk of severe side effects. In particular, cancer cell seeding and spread of the cancer to be diagnosed is often caused by the sample extraction.

On the other side, fine-needle aspiration (FNA) based cell sampling technology is a patient friendly, minimally invasive, rapid and cost efficient tumour diagnostic procedure with negligible side-effects. Breast and prostate FNA is minimally invasive, inexpensive and requires no antibiotic prophylaxis, no anesthesia, and is readily accepted by more or less all patients. However, despite the overall advantages of FNA-based diagnostics for breast and prostate tumours, there are three major drawbacks that prevent a general application of FNA in clinical routine: i) FNA-based sampling in general only yields a sparse amount of diagnostic material, with aspirated cells in an extracellular fluid. The cells are taken out of their tissue context so that in particular the degree of tumor cell infiltration can not be judged, only the features of the cells themselves. ii) FNA samples are frequently non-representative by containing too few cells of the suspicious lesion. Only a very competent sampling operator can master successful lesion biopsies. iii) Even with a successful lesion biopsy it is often difficult to obtain a conclusive diagnosis based upon the frequently minimal cellular deviation of atypical or even fully transformed malignant cells compared with their normal counterparts. This demands a diagnostic competence which requires extensive training and experience. For these reasons, core needle biopsy or surgical biopsy are to date still favoured over FNA for breast and prostate cancer diagnostics. However, due to their mechanical crudeness, they are highly traumatic and combined with significant risk for tumour cell seeding in the multiple needle tracts.

The conflict between the amount of sample material needed to obtain conclusive information and the severe side effects often caused by the extraction of this material can be circumvented, if the amount of information that can be extracted from the sample could be strongly increased and if the present sampling techniques could be made far less invasive.

The overall objective of this project was to develop and validate a quantitative, minimally invasive diagnostic tool for early and conclusive detection, diagnosis and monitoring of disease and disease progression of breast and prostate cancer, with negligible sampling-related side-effects.

To increase the amount of information that can be extracted from the sample this project has taken benefit from the recent remarkable progress in fluorescence-based spectroscopy and imaging. In particular, we brought in exciting new concepts to increase the sensitivity and specificity of fluorescence detection and, not the least, to overcome the diffraction barrier in optical microscopy (increasing the spatial resolution of the cellular imaging by almost an order of magnitude). Based on these concepts and methodologies, we could put in sight a development towards novel optical diagnostic methods unimaginable just a few years ago, and to develop the means necessary to

eliminate the conflict between the amount of sample material needed and the necessary diagnostic conclusiveness.

The overall strategy to be achieved within the FLUODIAMON project was thus to exploit the combination of the unique sensitivity, specificity and spatial resolution offered by the very latest developments in fluorescence microscopy, and the minimal invasiveness of fine-needle aspiration (FNA) based cytology. Based on the fluorescence methods the project aimed at developing and applying approaches where minute FNA samples could be subject to objective analyses on a molecular or subcellular level, thereby providing sufficient diagnostic information. The strategy of the project was in particular to analyse three main categories of proteins within the FNA-sampled cells: cytoskeletal proteins, receptor proteins located in the cellular membranes, and proteins regulating the division cycles of the cells. Apart from the mere detection and identification of proteins of these categories in the sampled cells, tumor-specific spatial distributions of these proteins within the intact sample cells were also exploited. We argued that high-resolution spatial distribution patterns of proteins in intact cell may represent a wealth of information beyond the mere presence/absence of the proteins, representing an almost unexploited dimension of diagnostic information.

To reach the overall goal of the project, efforts not only in the further development of fluorescence microscopy/spectroscopy and FNA sampling technology were needed. Following FNA-sampling, the FNA samples needed to be properly prepared; and diagnostically representative proteins in the cells needed to be identified and marked using specific affinity molecules, labeled with bright and stable fluorophore marker molecules. Extraction of diagnostically significant fluorescence information from the FNA-sampled cells needed to be optimized based on iterative efforts within optical imaging and sample preparation and labeling. Finally the acquired images of the FNA-sampled cells needed to be analysed with new bioinformatic means and finally validated towards clinical data. In other words, to reach the overall objective of the project required a multidisciplinary interaction and expertise comprising clinical cytology, cancer proteomics, molecular biotechnology, fluorophore chemistry, fluorescence microscopy, nanotechnology, optics, solid state detector technology, data processing, and bioinformatics.

Given the effort required spanning over a broad range of disciplines, the main objective of the project was also divided into a set of detailed objectives, reflecting the whole chain of steps needed to reach the final diagnostic procedure aimed for. The detailed objectives of the FLUODIAMON project were:

- To improve spatial resolution of state-of-the-art light microscopy in pathology by an order of magnitude.
- To improve the sensitivity of fluorescence-based imaging of FNA acquired cells to the ultimate single-molecule level.
- To take multi-parameter fluorescence imaging of individual FNA acquired cells to its extreme in terms of information content, largely based on photon statistical approaches and parameters extracted from non-linear effects.
- To develop standardised FNA-based sampling of suspected breast and prostate cancer lesions with negligible side-effects, and with optimised needle visibility for ultra-sound guided needle positioning.

- To select already known molecular markers that will be compatible with the developed fluorescence imaging techniques and can be highly anticipated, when investigated by these techniques, to strongly correlate with malignant transformation and clinical tumor aggressiveness.
- To identify existing and develop new affinity molecules to these markers, which are highly specific and fluorophore-labeled for optimised fluorescence readout properties.
- To refine bioinformatic evaluation and data processing to find the combination of fluorescence read-out parameters that most strongly correlate with the relevant clinical parameters and yield the strongest diagnostic reliability.
- To optimise the combination of techniques and procedures, in accordance with the conclusions of the above goal (VII), to maximise sensitivity and specificity for FNA-based diagnostics of breast and prostate cancer, enabling a decisive improvement of the outcome for the patients suffering from these diseases.

Now, at the end of the project we can conclude that the overall objective, as well as all the detailed objectives have been fulfilled.

## **Project Results:**

In the forefront of the project efforts was the further development and adaptation of cutting-edge optical methods so that they could extract sufficient diagnostic information out of very small amounts of diagnostic material. If only very small amounts of diagnostic material are sufficient for diagnosis, the diagnostic sampling can be made minimally invasive, and sampling-related side-effects can be reduced to an absolute minimum. However, to reach the overall objective of this project - to develop and validate a quantitative, minimally invasive diagnostic tool for early and conclusive detection, diagnosis and monitoring of disease and disease progression of breast and prostate cancer, with negligible sampling-related side-effects – required interaction between a broad range of different disciplines, as well as significant progress within each of these disciplines. In particular, substantial effort and progress were needed within the following areas, covering all steps from the initial sampling step to the final evaluation of the acquired images and the validation of the whole diagnostic procedure:

1. Sampling of suspect breast and prostate cancer lesions with minimal side effects by fine-needle aspiration (FNA)
2. Preparation and handling of the clinical FNA samples
3. Selection of protein targets to be visualized by the optical methods in the cells in the FNA samples, so that the proteins when visualized by these methods can provide specific diagnostic information based on differences between normal and cancer cells in the amounts of the proteins in the cells and their spatial distribution patterns within the cells.
4. Selection and development of affinity binders, which can bind specifically to the target proteins, and which carry a fluorophore marker molecule to enable fluorescence detection of the targeted proteins.
5. Selection and development of fluorophore marker molecules which provide a maximum of sensitivity and specificity, when used in the specific FNA sample specimens, and for the different combinations of affinity molecules-target proteins.
6. Development and optimization of high-end fluorescence methods for diagnostics of sparse amount of cells sampled by FNA, focusing in particular on sensitive detection of multiple specific proteins in individual cells, and ultrahigh resolution imaging of the spatial distribution patterns of specific proteins in the cells. For different methodologies were considered in the project: A. Ultrahigh resolution fluorescence microscopy (Nanoscopy, or stimulated emission depletion (STED) microscopy), B. Multi-parameter detection imaging (MFDi), C. Transient state (TRAST) imaging, and D. Two-photon excitation (TPX) spectroscopy.
7. Development of bioinformatic tools for the diagnostic evaluation of the protein content and the subcellular features within the imaged FNA sampled cells.

In the section below, we summarize the progress within each of these major areas involved in the project, and finally describe the outcome when integrating the results from each of these areas into the final diagnostic procedure for safe and reliable detection of early breast and prostate cancer.

1. Development of fine-needle aspiration (FNA) sampling procedures of suspect breast and prostate cancer lesions

Fine-needle aspiration (FNA) based cell sampling technology is a patient friendly, minimally invasive, rapid and cost efficient tumour diagnostic procedure. In addition, ultrasound-guided FNA makes it possible to obtain diagnostic cellular material from non-palpable breast and prostate lesions as small as 2 mm in diameter. However, there are three major drawbacks that prevent a general application of FNA in clinical routine: i) FNA-based sampling in general only yields a sparse amount of cell sample on which histopathological analysis cannot be performed. ii) FNA samples are frequently non-representative by containing too few cells of the suspicious lesion. iii) Even with a successful lesion biopsy it is often difficult to obtain a conclusive diagnosis based upon the frequently minimal cellular deviation of atypical or even fully transformed malignant cells compared with their normal counterparts. This demands a diagnostic competence which requires extensive training and experience.

For these reasons, core needle biopsy or surgical biopsy are to date still favoured over FNA for breast and prostate cancer diagnostics. However, due to their mechanical crudeness, they are highly traumatic and combined with significant risk for tumour cell seeding in the multiple needle tracts. For prostate cancer diagnostics it has been shown that core biopsy in addition to extensive bleeding and tumour cell seeding results in significant bacteriuria in between 21–53% of the patients. Such side-effects are completely negligible with the fine-needle aspiration (FNA) technique.

In this project, the Stockholm-based small company Neodynamics AB has managed to improve the FNA sampling procedures in several aspects:

First, a new FNA needle was developed with improved sample yield and surface-treated for considerably increased ultra-sound visibility. Sample yield (i.e. the amount of sample material obtained after needle aspiration) is important, given that FNA-based sampling in general only yields a very sparse and often also insufficient amount of cell sample material. Ultrasound visibility is important since upon sampling, when the needle is introduced in the tissue, the positioning of the needle prior to the aspiration relies on the guiding and visualization via ultrasound. If the needle position is not positioned at the right place, the sample is obviously not obtained from the lesion region to be investigated. Final development on the needle design was performed resulting in two prototypes (outer diameter 0.7 mm) with and without enhanced ultra-sound visibility. Large scale clinical testing was carried out in order to evaluate their performance. The new needle achieved outstanding results when compared to needles commonly used today. It yielded three times more material than standard needles of the same diameter and two times more than needles with a diameter of 0.7 mm.

Second, needle additives to facilitate needle penetration, manoeuvrability and precision in positioning were developed. In the device developed by Neodynamics AB called “Cytotest” mechanical energy is added to the sampling-needle by computer controlled oscillating movements. These oscillatory movements of the needle have been demonstrated to strongly facilitate tumor penetration, allow the direction of the needle to be changed to a larger extent upon penetration, as well as to increase the sampling phase efficacy. Also a rotating movement was added to the needle in order to mimic the way the needles are handled by trained cytologists (sampling is most often performed by non-cytopathologist clinicians e.g. urologists and radiologists who often lack direct feed-back sampling experience). Moreover, the design of the bevel section of the needle has been adapted for optimal cutting behaviour together with the oscillatory and rotational movement schemes. “Cytotest” has been

used at two breast diagnostic centers in Stockholm (Sabbatbergs Breast Center and St Göran Hospitals Breast Unit) on more than 500 breast cancer patients with excellent results. Due to concerns from the operators regarding decrease of tactility upon sampling when using the oscillatory needles, an improved handle was designed, in collaboration with the Design Programme at Kalmar University, Sweden, and successfully tested by clinical operators. The needle handle will be implemented into the “Cytotest” instrument before large-scale production is initiated.

Third, Neodynamics AB has succeeded with the development of a fully functional anti-seeding instrument as well as a specialized anti-seeding needle. The anti-seeding instrument consists of a signal processor, an amplifier as well as a patient unit. The anti-seeding needle tip acts as an electrode whereas the shaft is isolated. When radio frequency pulses are given, a current is applied between a dispersive electrode, for instance located over the chest of the patient when sampling from the breast, and the needle electrode. Thereby, the tissue around the open needle tip is heating up in a very localized fashion where the current density is highest. Starting with tests on pork livers provided the knowledge to successively improve the used pulses and their properties, an ethical permit was then acquired in order to commence with clinical studies. From the tests on patients undergoing the sampling procedure based on the anti-seeding procedure the results have been positive in every aspect. No or very little additional discomfort is experienced by the patients and the analyzed samples have not been denaturized (meaning the samples are still intact). To estimate the extent of local tumor cell-dissemination upon FNA sampling, partners Neodynamics AB and KI (Wiksell and Auer) investigated the blood droplet(s) that spontaneously penetrates the skin orifice after FNA needle retraction. They found that there is a high frequency of these blood droplets contained tumor cells. From this finding it is reasonable to assume that upon needle sampling, cancer cells can also enter the circulating blood system of the patient, and may finally also cause distant metastasis. Interestingly, applying the developed anti-seeding procedure significantly reduced the overall amount of blood and exudate passing the skin orifice after retraction of the FNA needle. Moreover, a test series on patients indicated that there were no viable suspended tumor cells in the blood droplets leaking out of the skin orifice after needle retraction.

Taken together, the development on FNA sampling techniques is likely to lead to a dramatic improvement regarding sampling related risks, sample reproducibility, quality and volume of collected cells as well as regarding safety against cell tumor dissemination. These developments also give the necessary prerequisites, covering the first steps in the diagnostic procedure developed in the project.

## 2. Preparation and handling of the clinical FNA samples

Once a patient sample is safely and reliably obtained by the FNA sampling procedures, the subsequent preparation and handling of the cellular material need to be optimised and properly adapted to the overall diagnostic procedure. In the project, the optimisation of the sample handling and preparation had to incorporate many aspects:

1. the sample preparation should not compromise the specificity of the affinity molecules against their protein targets in the cells,
2. the recorded fluorescence intensity from the fluorophore marker molecules (labelled onto the affinity molecules) versus the background level in the imaging had to be maintained, as well as the stability, reliability and reproducibility of the recorded fluorescence parameters



3. the cellular morphology and specific cellular features should not be severely affected by the sample handling and preparation,
4. the preparation should yield a proper adhesion of cells to glass cover slides and preparation formats compatible with transport,
5. the preparation protocols should be sufficiently simple, so that they could be performed at clinical sites with typically limited resources and infrastructure for cell handling.

Taken all these aspects together favoured a procedure, where immediately after sampling the patient material was subject to a three-step procedure consisting of a blood cell lysis step (to remove impurities, in particular red blood cells), a fixation step (to stabilize the material against degradation) and a cytopspin step (centrifugation of the sample when deposited on a cover glass slide, to make the cells adhere to the glass slide), thereby preparing the sample for subsequent sending, labeling and imaging.

### 3. Selection of protein targets

In general, the cure of cancer diseases strongly depends upon early and detailed diagnosis of type and stage of the malignant disease, typically requiring a molecular characterisation of the malignancy. Consequently, the aim of the project was to further develop and apply analyses of the minute FNA samples on a molecular or subcellular level. Partners Karolinska Institutet (KI) and Lübeck University (LUEBECK) have previously identified a substantial number of proteins in freshly taken clinical samples from tumours of the breast, prostate and others. By comparing cancer type specific protein expression with that of corresponding normal tissue, single proteins or groups of proteins were identified, highly correlated to malignant transformation and clinical tumour aggressiveness.

In this project, we wanted to extend the protein characterisation beyond the plain presence/absence, or rather deviations in the amount of these proteins in the FNA sampled cells or in the extracellular sample material. Given the sensitivity and spatial resolution offered by our fluorescence methods, we also wanted to map the spatial distributions of selected proteins within the FNA sampled cells as objective parameters specific for malignant transformations. From more fundamental biophysical studies of cells it is known that malignant cells often display an increased elasticity compared to corresponding normal cells. It is therefore reasonable to assume that the difference in the elasticity of the cells is also reflected in the spatial organisation of the cytoskeleton, and in the spatial distribution of proteins which are part of the cytoskeleton. Moreover, fundamental studies in cell biology have shown that the interaction between specific proteins in the cellular membranes can be different as a result of malignancy, and that the regulation of cell division is disturbed in cancer disease is essentially a part of the definition of cancer disease. Based on these notions, the project focused on diagnostically monitoring three different categories of proteins: cytoskeletal proteins, membrane proteins, and cell cycle regulating proteins.

The specific proteins selected to demonstrate the diagnostic procedure, and the diagnostic feasibility of the microscopic techniques, were found from extensive literature searches, and among our already identified tumour markers. The final condensed list of protein biomarkers for breast- and prostate cancer were then selected according to the following parameters:

- Presence or absence of protein markers in disease as compared to controls
- potential of changes in spatio-temporal behaviour and interaction patterns in diseased stages



- changes in the sub-cellular location (cytoplasm, nucleus, extracellular space or plasma membrane) of the protein upon disease

In the project, the diagnostic procedure was eventually narrowed down on analysing the abundance, interactions and/or the spatial distribution patterns of the following proteins in the FNA sampled cells: vimentin and tubulin (cytoskeletal proteins), IGF1R, HER1, HER2 (membrane proteins), Cyclin A, Cyclin E (cell cycle regulating proteins).

The monitoring of the presence and spatial distribution patterns in individual cells of these specifically selected proteins proved to be a successful strategy. The experience gained using these biomarkers for fluorescence imaging will help to select additional markers in the future if necessary for improving the diagnostic performance even further, and can also serve as a template for selection of proteins to diagnose other diseases with similar detection and imaging strategies.

#### 4. Selection and development of affinity binders to the protein targets

The implementation of the novel and advanced fluorescence-based imaging techniques for the diagnostic procedure, as aimed for in the project, requires that the specific proteins that were selected to be analysed also can be properly detected. In this project the strategy was to use affinity binders that bind specifically to the proteins. These affinity binders were in turn labelled with fluorophore marker molecules, which when excited by laser light in the imaging instruments emit fluorescence and thereby can be detected and analysed. An important link in the chain of processes comprising the overall diagnostic procedure is thus the selection and availability of target-selective affinity probes to which fluorophores can be decorated.

In the project, three different classes of affinity probes were used. Apart from conventional antibody molecules, available via commercial suppliers, affinity binders of two different classes were developed. At the biotechnology department of the Royal Institute of Technology, Stockholm (KTH), new affinity binders were developed based on an existing protein library technology platform denoted "affibody binding proteins". These affinity reagents, Affibodies, are characterised by a small sized scaffold, about 2 nm in diameter and consisting of 58 amino acids. 13 of these amino acids, making up a binding surface area of approx. 800 Å<sup>2</sup>, can be combinatorially randomised. Since there are 20 different amino acids in the proteins of humans, there are in principle different combinations for the set of 13 amino acids at the binding surface, and thus considerable potential to optimise the affinity and selectivity of the Affibodies towards specific target proteins. At Academisch Ziekenhuis Leiden (LUMC), heavy chain antibodies (HCAbs), were developed for the project. HCAbs represent a class of antibodies specific for camels. The HCAb fragments responsible for the target binding are significantly smaller than conventional antibodies, represent the smallest naturally occurring protein-based affinity binders known to date (Affibodies are not considered as naturally occurring), and are frequently used as binders within biomedical research, diagnosis and therapy.

To provide the necessary affinity binders to the project, an inventory of available affinity reagents to the selected protein targets was performed and finalized, and an affinity protein reagent bank (consisting of commercially available monoclonal antibodies and already selected affibody variants to certain proteins) was established for these targets. It should be borne in mind that even if there are one or several affinity binders reported to be specifically binding to a certain protein with high affinity, both the binders and the targets are proteins, and their conformations and their binding properties may change dramatically depending on the environmental conditions. Following identification of protein

targets, available affinity reagents to these proteins therefore had to be further investigated with respect to their binding properties under conditions close to those expected in the FNA sampled cells. In these investigations, the reagents were either directly labelled with fluorophore markers, or indirectly labelled via labelled, so-called secondary antibodies binding to the affinity binders to be investigated. The investigations of the binders were performed by conventional fluorescence-based confocal microscopy on cells from three selected cell lines representing normal cells and two different aggressive forms of breast cancer. As a result of these investigations, a list of affinity binders to the prioritized list of protein targets could be identified.

In parallel to these efforts, work was performed both at LUMC and KTH towards development of novel affibody and HCAb fragment binders to an interesting breast cancer marker candidate, the protein SATB1. This work resulted in the identification and validation of several new affibody and HCAb fragment variants. These were produced recombinantly (i.e. the DNAs encoding for these binders were introduced into host organisms for replication), subsequently labelled by various principles and found to selectively recognize the SATB1 target in cell samples. Data from binding studies suggest that different epitopes (here: different surface regions of the target proteins) are preferred by the two classes, opening up for a possible increased specificity in obtained localization signals. The new binders were used either separately or in combination to evaluate the presence and subcellular localization of the SATB1 target in different cell types. Although SATB1 was eventually not selected among the proteins showing the largest diagnostic potential in the diagnostic procedures of the project, the possible alternative use of the developed binders for detection of targets present in cell-lysates was demonstrated, and shows the versatility and value of having such selective binders to targets of interest.

Taken together, a set of affinity binders could be identified for the selected target proteins, well functional also under the conditions relevant for FNA sampled cells. The project has further shown that novel technologies for development of novel affinity proteins of different nature have the potential to be very valuable for future diagnostic applications. Compared to conventional, full-sized antibody binding reagents, the alternative binder protein classes used can provide advantages in terms of size, the preference to and selectivity of where on the target protein they bind, production routes (i.e. chemical synthesis) and controlled fluorescent labelling, incl. both number and position of fluorophores

## 5. Selection and development of fluorophores.

Both high-resolution microscopy (nanoscopy) and high-sensitivity, multi-parameter, fluorescence-based characterization of proteins in cells critically depends on the availability of suitable fluorophores, carrying specific chemical groups so that they can be coupled to relevant affinity proteins or to other molecules of interest. Besides spectral properties suited to the optical instrumentation, the most important properties of the label are high fluorescence quantum yield (i.e. that they offer an efficient conversion of the absorbed laser excitation into fluorescence emission) and high photostability. The University of Siegen (UNISI) has a vast experience in the area of fluorophore development, and was the main partner responsible for the fluorophore selection and optimisation.

In the project, commercially available fluorescent labels were evaluated with respect to their fluorescence properties, and tested more in detail by the various spectroscopic and imaging techniques in the project. Apart from fluorescence quantum yield and photostability, a critical aspect investigated

was the stability and reliability of the fluorescence parameters recorded from the fluorophores. In particular, the fluorescence lifetimes (decay time of the fluorescence after excitation) under different conditions and for various dye-conjugates were investigated, given that the fluorescence lifetime was an important parameter to differentiate between differently labelled affinity binders. Based on these investigations, a set of fluorophore labels could be identified as appropriate for the project, and were subsequently used in the developed imaging procedures.

In addition to testing and selection of commercially available fluorophores for the project, new chromophores were also synthesised. In particular, new hydrophilic and water-soluble fluorophores were designed and synthesized, specifically suitable for the labelling of affinity molecules such as antibodies and affibodies. The newly synthesized fluorophores displayed minimised dye aggregation and very low unspecific binding to cell-based specimens, yet exceptional properties in terms of photostability and brightness, as required for the ultrasensitive and high-resolution microscopy techniques in the project.

Taken together, highly photostable and bright fluorophores are indispensable for successful fluorescence spectroscopy/imaging in general, and in particular for the cutting-edge fluorescence-based methodologies developed in this project. The selected and developed fluorophores have well fulfilled their purpose, and will find applications not only in the context of the project but as markers for a range of different applications in bioscience.

## 6. Development and optimisation of high-end fluorescence methods for diagnostics of FNA-sampled cells

In the project, several fluorescence-based methods were developed and finally evaluated for use in a procedure to diagnostically classify minute amounts of FNA-sampled material. The strategy was to provide a combined procedure, with the complementary information obtainable from super-resolution imaging on the one side and new, extremely sensitive and specific fluorescence-based optical imaging techniques on the other, forming a sound basis for an extensive analysis of FNA aspirates and a final highly reliable diagnostic classification.

Along this strategy, four different fluorescence-based methods were developed and modified towards diagnostic use in this context. The development and results obtained for each of these methodologies are outlined in the following.

### 6.1 Super-resolution, stimulated emission depletion (STED) microscopy

Diffraction precludes the formation of focal spots smaller than about half of the wavelength, i. e. approx. 200 nm for visible light. This has been one of the most influential limits in optical physics, and in biomolecular and cellular imaging, i.e. imaging has not been possible to perform with a resolution much better than the wavelength of the light by which it is viewed. However, following the invention of the first subdiffraction-resolution (STED) microscope, this “classical” resolution limit has been possible to overcome. A whole new scientific community has emerged which is devoted to improving fluorescence imaging, to push it to its (resolution) limits and to introduce it into biomedical applications. MPIBPC has been continuously improving the resolution in STED microscopy for many years now, and is the pioneer group in the field. In this project, we have set as a goal to introduce ultrahigh-resolution microscopy for diagnostic purposes, and the further development and adaptation

of STED microscopy, or nanoscopy, for diagnostic evaluation of FNA-sampled cells has been a major activity.

In the course and for the purposes of the project, a STED microscope design based on a novel supercontinuum light source (a laser emitting light over a continuous range of wavelengths in the visible range (approx. 400-700nm)) was implemented at MPIBPC during the first year of the project and at KTH during the second year. Following its publication, the microscope has received considerable attention (not just within the consortium) and has now been duplicated in several labs throughout Europe. The experience gained in the construction of the supercontinuum based STED system has emphasized the key role of the light source for the further improvement and dissemination of STED microscopy. The dissemination of the STED microscopy technique and its adoption in biomedical applications, has previously been somewhat impeded by the fact that expensive and maintenance-intensive laser sources were necessary for its implementation. Following the success of the supercontinuum based STED system, this threshold has been drastically reduced. The new, relatively inexpensive laser source utilized in both of these setups marks a substantial improvement not only in terms of cost but also in terms maintenance efforts and alignment required for the operation of the STED microscope. Despite its greatly facilitated assembly, the instruments provide spatial resolutions in both colours of 25–45 nm, depending on the operating conditions. The developed STED instruments provide all the necessary functionality, including dual-colour ultrahigh resolution capabilities, clearly representing two of the most significant readout parameters for our diagnostic procedure. In terms of robustness and usability they are close to commercial STED instruments currently on the market, but offer higher resolution and true dual-colour capabilities. Moreover, the cost of the instruments is almost an order of magnitude lower than that of current commercial instruments. These instruments may thus serve as a blueprint for a future commercial instrument.

Following establishment of the STED instruments, proof-of-concept studies on cultured cancer cells were performed at MPIBPC and KTH to establish proper cell preparation procedures, and to identify the proteins, affinity molecules and fluorophore label molecules most appropriate and functional for the STED acquisition and analyses. A major benefit with the high resolution that we wanted to exploit in the project was to detect and characterize features in the spatial distribution patterns of certain proteins in the FNA sampled cells, typical for malignant development. Indeed, in the proof-of-concept studies performed by STED on cultured cells we could verify the ability to resolve cancer-specific protein distribution patterns by nanoscopy within individual cells. This was further evaluated for selected target proteins on cultured cells representing different stages of malignant development, and the feature differences in the images could be transformed into quantifiable parameters.

## 6.2 Transient State (TRAST) imaging

The information content of the nanoscopy images can be complemented by also encoding spectroscopic information, e. g. the environment-sensitive population of light-induced, long-lived, transient states of the fluorophore marker molecules. Upon laser light exposure, a fluorophore marker can be excited to a higher electronic state, from which it subsequently is relaxed back to the ground state. In this relaxation, fluorescence is often emitted, which can be detected and characterised with respect to intensity, wavelength, polarisation and fluorescence lifetime (i.e. how long the fluorophore is in the excited state before it relaxes back to the ground state). Excitation of fluorophores can also lead to the generation of long-lived, transient states (by trans-cis isomerisation, intersystem crossing to a triplet state, or photoinduced electron transfer). These states are generated with a low probability but because they are relatively long-lived, they can still be considerably populated under continuous

excitation. Due to their long lifetimes they have more time to interact with the immediate fluorophore environment, making them highly environment sensitive. Their kinetics can thus change considerably due to smaller changes in e. g. accessibility of quencher molecules or microviscosities, caused by e. g. biomolecular interactions, for which practically no change would be seen in traditional fluorescence parameters (since the fluorescence lifetimes are in the range of nanoseconds, i.e. one thousand to one million times shorter than for the long-lived transient states).

In this project, a concept for transient state (TRAST) microscopy developed at KTH, and opening for the possibility to monitor transient photo-induced non-fluorescent states of fluorescent molecules on a massive parallel scale and at arbitrary concentrations, was further developed towards diagnostic applications. This concept can add considerable sensitivity to existing fluorescence microscopic readouts, providing also an additional dimension of fluorescence information that has previously been almost fully overlooked in fluorescence microscopy. In the first half of the project, the use of the technique was established and optimised under more general conditions, new image-extraction procedures were developed, and new strategies to explore the environmental sensitivity of the readout were further demonstrated. The TRAST concept was established on several different microscope modalities. In the last project period, TRAST monitoring has been further developed for cellular analyses. The TRAST approach was demonstrated as a tool to image and spatially resolve the metabolism within individual live cells (O<sub>2</sub> consumption, redox status), and to detect differences in metabolic activity typically occurring in cancer cells.

### 6.3 Multi-parameter fluorescence imaging (MFDi)

In order to be able to extract sufficient diagnostic information out of sparse, or even individual FNA-sampled cells, it is necessary not only to monitor species of one target protein in the cells, but of several different proteins at the same time. This is strongly facilitated if the proteins can be characterized and identified with respect to several independent parameters. HHU has developed a new general strategy based on multiparameter fluorescence detection (MFD) to register and quantitatively analyse fluorescence images. In multiparameter fluorescence imaging (MFDi) pulsed excitation with typically at least two different lasers is used for generation of fluorescence in the sample. The fluorescence is divided based on the spectral properties (wavelength) and polarization into several detection channels, and is detected by so-called time-correlated single-photon counting (TCSPC). Thereby, each registered fluorescence photon is characterised by its time with respect to the laser pulse, its time from the start of the experiment, the detection channel where it is registered, and the spatial coordinates from where the fluorescence photon originated in the sample. Using this data acquisition scheme, fluorescence can be detected in an eight-dimensional parameter space.

With several independent parameters registered from individual molecules the ability to identify different molecular species is facilitated via their characteristic set of parameter values, as a kind of “bar-coding”. Prior to the project, HHU showed that by using multiple fluorescence dimensions, and specific dye labelling schemes detection and identification of 16 different compounds in a mixture could be achieved at the single molecule level. This provided the starting point for this project, with the aim that similar strategies would offer strongly improved detection and identification capabilities for different protein species also in FNA-sampled cells.

During the first part of the project, MFDi in combination with laser scanning microscopy (LSM) was established at HHU and KTH. After exploring the capabilities of the technique in a feasibility study using test samples of plant cells available in Düsseldorf, fluorescence assays were further developed on the dedicated model cell lines of the project. To be able to detect and identify several different

proteins in the cells, it is important that the fluorescence signal from all species is bright, the background is low, and that the set of recorded fluorescence parameters from the different protein species are sufficiently different to allow identification. Finally, the fluorescence parameters need to be robust and stable. Identification of the different protein species in the cells could in principle be achieved by using affinity molecules, decorated with fluorophore labels differing in their combinations of fluorescence emission wavelength and fluorescence lifetime. However, much effort was required to define the conditions for and verify that the fluorescence parameters remained stable under different conditions. Here, both sample (dye-conjugation, cellular preparation and media etc) as well as instrumentation aspects had to be optimised. Regarding the instrumentation, in particular the detectors time-response had to be considered, and specific detectors developed by partner B&H could finally provide the solution, offering the required stability in their time-response needed for the fluorescence lifetime recordings. Finally, an important part of the establishment of MFDi for the cellular analyses was the data processing. Here, software packages were further developed and modified to extract and display the multitude of fluorescence parameters from the extensive raw data recorded from the cells.

Taken together, following optimisation of instrument hard-ware, soft-ware as well as sample preparation protocols, final MFDi measurement routines could be established on cultured cells, allowing detection and identification of multiple significant molecular or cytohistological features in the cells. This formed the basis for the subsequent application of MFDi on multiply stained FNA patient samples (see below). This also demonstrates that the developed acquisition and analysis procedures match the prerequisites for further adoption of the MFDi technique as a common diagnostic tool of cellular samples.

#### 6.4 TPX technology for ultrasensitive detection and identification of pathological protein profiles from cytological aspiration material

The TPX technology, developed at the Laboratory of Biophysics (UTURKU), is extremely well suited to measure concentrations of target molecules from minute samples with high precision and sensitivity. TPX is a particle-based assay. The particle is coated with a target-specific antibody, by which target molecules attaches to the particle. Thereafter additional antibodies labelled with fluorophore reporters bind to the target molecules attached to the particles, making the particles fluorescent with an intensity dependent on the concentration of the target molecule. A specific feature of the TPX method is that it is based on (two-photon) excitation of the fluorophore reporters and laser trapping of the particles using the same laser source. Using specifically antibody-coated microparticles, these are trapped, and then the target (protein) of interest gets concentrated to the surface of the microparticles by binding to the antibodies. When a second, fluorescently marked antibody attaches to the microparticle bound targets, the concentration of the target is revealed by the intensity of the particle.

In this project, the strategy was to develop the TPX technology to monitor the presence of nonlocalised proteins and their concentrations in extracellular aspiration material, providing complementary information to that obtained from the fluorescence-based imaging methods outlined above. Following demonstration that TPX can be used for high sensitivity detection of relevant protein targets, using a well-known assay for PSA (prostate specific antigen), the target protein 14-3-3 (a secretable tumor marker protein, expected to be found in the extracellular fluid) was chosen from the FLUODIAMON list of protein targets and screening of binders towards this protein was carried out. In contrast to microscopy or cell surface protein quantification, high sensitivity soluble protein quantification with TPX requires the use of two antibodies against different (and spatially well



separated) surface regions (epitopes) of the target protein, which made the antibody selection extra demanding. Nonetheless, the development of antibodies for the 14-3-3 target could at least be taken to a point where a screening assay for the target could be demonstrated, and the TPX technology was thereafter demonstrated for both soluble and cell surface targets, and more particularly in an assay detecting two different cancer-specific proteins; Cyclin-D and HER2. The assay was demonstrated with both clinical (epithelial) samples and model cell-lines.

Taken together, The TPX technology assay development within the project has demonstrated that the assay should be readily integrable to the FNA sampling protocol of the project. A small aliquote of the original sample could be set aside for TPX analysis, providing complementary diagnostic information to that obtained from the cellular imaging.

## 7. Combined diagnostic procedure performed on FNA-sampled material

In the instrument development, the different modalities were developed with the consideration to be used either in a parallel, simultaneous approach within one integrated instrument, or the combination of techniques would be achieved by a sequential use of the techniques on the FNA samples.

Several setups were established in the project enabling a parallel use of techniques. In particular, at MPIBPC, a STED instrument was refined towards the integration of the STED (Nanoscopy) and MFDi techniques. The combined STED/MFDi microscope was equipped to record fluorescence lifetime traces in addition to the two colour channels already available. Alternatively, it could be used to implement dual-channel imaging with a fluorescence lifetime based separation of two fluorescent labels. This latter approach to two-colour imaging does not require a second STED beam, so it is insensitive to drift and thus optimally suited for analysing co-localisation of protein target molecules, and the combination of the lifetime based separation with the two-wavelength STED approach implemented earlier allows a total of three tumour markers to be imaged simultaneously on the cells.

However, it was realised that although the FNA samples gave only a limited amount of diagnostic material, they still contained a sufficient number of cells to be divided on several glass slides. Given this possibility, it was concluded that the diagnostically most significant combination of accessible parameters could be obtained by performing different imaging modalities separately, and sequentially. Thereby, the conditions for each of the measurement techniques can be optimised independent of each other, and one can avoid the compromises that follow from a combined imaging with two or more of the modalities, on the same cells, and at the same time.

As outlined above, a sample handling and preparation protocol was defined taking into account that FNA samples are far more complex and heterogeneous than standardised cells from cultured cell lines, and may contain blood cells, fatty tissue and other contaminations. Extensive effort was required to yield optimal fluorescence characteristics and to ensure that the cellular handling does not compromise any of the most significant readout parameters of the different procedures used to image the cells. The optimisation of the sample handling and preparation had to incorporate many aspects, including proper specificity of several affinity molecules against their molecular targets, fluorescence intensity versus background, stability, reliability and reproducibility of the recorded parameters, maintenance of cellular morphology and specific cellular features, proper adhesion of cells to glass cover slides and preparation formats compatible with transport, as well as sufficient ease of preparation at clinical sites with typically limited resources and infrastructure for cell handling. Taken all these aspects together favoured a procedure including fixation of the sampled cells for stability and



reproducibility reasons. The advantages of using fixed samples more than well compensated for the additional information potentially retrievable from unfixed, or even live cells. To include fixation in the preparation of the FNA samples essentially also defined STED and MFDi as the combination of techniques most suitable for the diagnostic characterization of the FNA-sampled cells.

With the FNA sample handling and preparation properly optimised, with a combination of MFDi and Nanoscopy identified as the diagnostically most useful combination of imaging methods providing uniquely retrievable diagnostic information, and finally with a set of parameters and targets identified from the initial studies by STED and MFDi on cultured cells and then on cells from a few FNA samples as likely targets to provide high diagnostic power, tests were finally performed on a larger set of clinical FNA samples. Clinical FNA sampling from patients with suspect cancer lesions was performed in Stockholm and a few times also in Uppsala (by partner KI, with on-site sample handling and preparation support by KTH). Samples were obtained from suspect cancer lesions both from the breast and from the prostate, but with the major part of the samples obtained from the breast. After reserving a major part from each FNA sample for the routine diagnostic assessments, the remaining part of the samples were directly taken for preparation, distributed on up to 12 different glass slides, and then passed on for imaging, locally to KTH, and to HHU and MPIBPC. From the FNA samples, high-resolution images and spectroscopic data were acquired at KTH, HHU and MPIBPC at four different platforms using the different and complementary techniques.

At KTH nanoscopic imaging was performed on the membrane protein targets HER1 and IGF1R. The proteins HER1 belong to the epidermal growth factor receptor (EGFR) family, and amplification or over-expression of its gene is believed to play an important role in the pathogenesis and progression of breast cancer. The Insulin-like Growth Factor 1 Receptor (IGF-1R) is a receptor protein found in the cell surface membranes of human cells. It is activated by a hormone called Insulin-like growth factor 1 (IGF-1) and by a related hormone called IGF-2. Also the IGF-1R is implicated in several cancers, not the least breast, prostate, and lung cancers. Both membrane proteins HER1 and IGF1R can interact with a broad range of different molecules, and also form complexes with each other in the membranes. Our strategy with the nanoscopic characterisation of these proteins on the surfaces of the FNA-sampled cells was that a malignancy may not only be manifested by the amount of these proteins in the cells, but also by their spatial distribution and possible co-localization patterns, reflecting the degree of interaction between the proteins on the cell surfaces. Indeed, the two studied proteins showed significant differences in terms of distribution within the membrane of the cells, HER1 tended to form pools with high density of proteins, while there were also other areas in the membranes where no proteins were found. IGF1R were more evenly spread out into separate areas. While the fine details of these distributions cannot be detected in the diffraction limited confocal image, the high resolution data could later be used for the diagnostic analysis. Several features could be identified already by visual inspection.

At KTH, imaging was also performed with MFDi, to study the same two membrane proteins (HER1 and IGF1R), and in addition also HER2. HER2 has a similar structure as HER1. Amplification or over-expression of the gene for HER2 occurs in approximately 30% of breast cancers. HER2 is strongly associated with increased disease recurrence and a worse prognosis. HER2 can form complexes (heterodimers, i.e. associate to other molecules) and is considered to be the preferred dimerisation partner of the other EGF-like receptors, such as HER1. The major strategy with the MFDi measurements at KTH was to analyse not only contents of the target proteins in the FNA-sampled cells, but also the extent of interaction, or complex formation, between the proteins in the membranes of the cells. With fluorescently labeled affinity molecules directed towards the target proteins, the fluorescence parameters of the fluorophores will change (in particular the fluorescence

intensity and the fluorescence lifetime) if two proteins are close ( $< 100\text{\AA}$  ( $10 \times 10^{-9}\text{m}$ )) to each other. In this sense, the MFDi measurements provide complementary information to the STED measurements above, which could provide estimates of molecular interactions from high-resolution co-localization patterns. Due to a large sample heterogeneity, it was difficult from the MFDi measurements to identify clear patterns of parameter changes between cells from benign and malignant FNA samples. However, when subject to more advance image analyses, more distinctive features could indeed be indentified.

At MPIBPC (Göttingen) nanoscopy was performed on two FNA samples from each patient. The samples were stained for the tumour markers vimentin and tubulin 1c, respectively, using the established immunofluorescence protocols. In addition, a stain towards the cell nuclei (DAPI) was applied to all samples in order to allow for the counting of cells not expressing the respective tumour markers. Vimentin is a cytoskeletal protein, which plays a significant role in supporting and anchoring the position of the organelles in the cytosol. The dynamic nature of vimentin is important when offering flexibility to the cell. Altered expression and/or various modifications of vimentin have been found to correlate with certain cancer forms, and for some cancers vimentin is also established as a biomarker. In this project vimentin was included as one of the proteins to be monitored due to this record, but also for being a cytoskeletal protein. In general, biophysical studies have shown that cancer cells can display an increased elasticity compared to normal cells. This is likely to be reflected in differences between normal and malingnant cells with respect to the spatial arrangement of the cytoskeletal proteins. With ultrahigh resolution STED, we aimed at reaching the sufficient resolution to resolve these differences. As a second cytoskeletal protein also tubulin 1c was studied, based on the same ideas as for vimentin. In parallel with the STED imaging, the cells were also measured with a standard wide-field microscope to obtain complementary additional information. While STED microscopy could provide ultrahigh resolution structural information (but no expression levels/cell numbers), the wide-field microscopy images could allow counting of cells expressing the marker/ not expressing the marker. All images were sent to UH for quantitative image analysis.

Four FNA samples from each patient were sent to HHU for MFDi characterization, and treated according to the established immunofluorescence protocol. One sample was stained against the six selected tumour biomarkers in parallel by means of three fluorescence labelled affibodies specific for the membrane receptors human epidermal growth factor receptors 1 and 2 (HER1, HER2) and insulin-like growth factor 1 receptor (IFG1R) and three combinations of primary antibodies targeted to CyclinA, CyclinE and vimentin and corresponding secondary antibodies (i.e. antibodies binding to the primary antibodies). One sample was incubated only with secondary antibodies, thus giving the possibility to monitor unspecific binding e.g. caused by adhesive components from the extracellular matrix. Another sample was stained using only the three affibodies. As the fluorescent labels for the affibodies were chosen in a way that they were spectrally fully resolvable, the datasets acquired from these samples could be used to validate the results obtained from the fluorescence lifetime analysis performed on the six fold stained samples (i.e. by comparing presence/absence of fluorophores the extent to which the different fluorophores influenced each other could be estimated, and either exploited as an information source or calibrated for). Eventually one sample was prepared unstained to investigate the contribution of autofluorescence. By these MFDi experiments with up to six-fold protein tagging the expression and distribution of several proteins could be monitored at the same time in individual cells. In cancer diagnostics in general, the over- or sometimes under-expression of various protein biomarkers have been found to provide important information. However, by looking at the expression levels of not only one protein, but of several proteins as a kind of protein expression profile, the diagnostic reliability can in most cases likely be significantly increased. Despite the

variable staining on FNA samples several distinct features regarding the expression patterns of the markers under examination could be clearly identified already by visual inspection, and by subsequent bioinformatic analyses these observations, and feature differences between normal and malignant cells could be further quantified.

Taken together, samples from 97 patients were altogether sampled at the different imaging sites as described above. Table 1 summarises the statistics for all the patient samples included in the final analysis (the number of samples sometimes differ from the number of patients because there were a few cases where samples were taken from the same patients but from different breasts). Several features indicative of malignant development could be observed in both the MFDi and STED images already by some visual inspection. However, for a more quantitative analysis all image data from the FNA samples were sent to partner UH for bioinformatic analysis and verification.

## 8. Bioinformatic tools for diagnostic evaluation and final clinical verification

The modern imaging technologies developed and used in this project are capable of measuring intracellular molecules with a spatial resolution unimaginable by most researchers in the field 10 years ago. The high-resolution and high-content of the images requires advanced computational methodologies (bioinformatics) to translate the image data into knowledge of clinical benefit. In the project, novel and accurate microscope technologies were developed and applied to FNA-sampled cells from suspect breast and prostate cancer lesions. The major challenge at this point was the development and implementation of computational methods capable of preprocessing and analyzing the image data systematically and fast. Before this project there were no published computational methods for preprocessing and analyzing STED and MFDi images at the scale required in clinical settings.

The computational analysis of image data and integration of the extracted information from the images was done by Systems Biology Laboratory at University of Helsinki. Firstly, this project resulted in 1,486 STED and MFDi images from 63 patients. By using a collection of microfeature detection methods, we were able to detect hundreds to thousands of feature locations from a single image. Each of these locations, in turn, provided several quantitative values that were used in a prediction step. In order to be able to analyze such a large amount of data, we used a computational framework called Anduril (Freely available at <http://www.anduril.org>. Analysis, scripts and results can be found at <http://csbl.fimm.fi/pub/FLUODIAMON>) that allowed us to perform all mathematical operations in a systematic fashion. After developing the management system for systematic use of the images, we developed novel preprocessing methods for MFDi and STED images.

The main objective of the bioinformatics analysis was to identify markers that predict whether the cells obtained with FNA were normal or cancerous cells. To this end we used strong prediction methods, such as Random Forest and ensemble classifier, developed in the machine learning community. Our efforts resulted in clear improvement in diagnosing patients with the FNA procedure.

Taken together, the development of computational methods for cutting edge, high-resolution and multiparameter microscopy has resulted in 1) freely available software for analyzing large amounts of images in a systematized fashion, 2) novel methods that are widely usable in various experimental settings and microscope techniques, and 3) clinically important predictors that can aid in breast cancer patient diagnosis.

## 9. Overall diagnostic procedure and its validation

Combining the efforts and progress within the different areas as outlined above, an overall diagnostic procedure could be established in line with the initial intentions. An intact chain of accomplished steps, from reliable and patient-friendly FNA sampling with minimized risk of tumour cell dissemination, over robust and reproducible sample handling and preparation, specific targeting of selected proteins with fluorophore-labelled affinity molecules, characterization of the FNA sampled cells by optical methods offering utmost sensitivity, specificity and spatial resolution, and finally bioinformatic analyses to fish out the differences in the image features necessary to distinguish normal from malignant cells, has led to the demonstration of a novel diagnostic concept. This concept can combine diagnostic reliability with safe, minimally invasive diagnostic sampling and thus enable a decisive improvement of the outcome of patients suffering from breast and prostate cancer.

For validation, the data from the imaging configurations for the different patient FNA samples were compared to the diagnoses set for these patients clinically (collected and compiled by KI). The comparison showed that we managed to create six independent classifiers capable to detect cancer from the different image data in minimally 55% of the cases, and maximally in 80% of the cases. It can thus be concluded that each of the six independent diagnostic classifiers can yield significant predictive strength. As a next step, in order to further improve the classification accuracy, several classifiers can be combined. However, at the moment we only have 20 patient samples that have been imaged by all of the four imaging configurations (STED@KTH, STED@MPIBPC, MFDi@KTH, MFDi@HHU), and a full combination of classifiers can only be made for this slightly limited group of patient samples. Still, our early results confirm that a clear improvement in accuracy can be obtained, though this need be tested more thoroughly with more patient samples in the near future.

Complementary clinical validation was also focused on the collection, preparation and analysis of clinical material of formalin-fixed and paraffin-embedded (FFPE) benign and malignant breast cancer specimens as well as of normal adjacent breast tissue. These FFPE specimens are already characterized by routine histopathological assessment and can thus serve as a clinical validation by immunohistochemistry and immunofluorescence of those markers that were selected by the consortium for fluorescence imaging of the FNA samples. Specimen and the corresponding clinical data of 245 patients representing benign and malignant breast cancer tissue were selected. Together with experienced histopathologists, these specimens were evaluated and the most representative areas were used for generating a tissue microarray (TMA) that allows for standardized high-throughput immunohistochemistry analysis and marker validation. Establishment of staining protocols has been successful for most of the markers and has been followed by processing of TMA image analysis at UH. While prioritizing breast cancer specimens, sample selection for paraffin-embedded prostate cancer specimens has also been finalized and FFPE material of 53 prostate carcinomas is available for TMA construction and immunohistochemistry and immunofluorescence.

Taking these validation paths together, the developed classifiers for the images from the patient FNA samples have been directly validated against the set clinical diagnoses for the same patients and the relevance of the selected markers have been verified on a large number of relevant patient material. The diagnostic procedures have been validated and shown to provide an attractive strategy for future clinical diagnostic schemes.

## Potential Impact:

The FLUODIAMON project is highly interdisciplinary. The expertise needed for this project comprises clinical cytology, cancer proteomics, molecular biotechnology, fluorophore chemistry, fluorescence microscopy, nanotechnology, optics, solid state detector technology, data processing, and bioinformatics. The level and width of competence required cannot be found on a national level but requires a European initiative. In particular, the project has demonstrated that by applying analyses based on novel state-of-the-art optical microscopy techniques, even quite small amounts of diagnostic material is sufficient for reliable diagnoses of breast and prostate cancer. Thereby, the amount of diagnostic material obtainable via minimally invasive fine-needle aspiration (FNA) is sufficient and more invasive core-needle or surgical biopsies can be avoided.

In the following sections we will discuss the possible implications of this major result. However, we will also briefly describe the potential impact of the results within the various individual disciplines that were engaged in this project.

### 1. Fine-needle aspiration techniques

The FLUODIAMON has supported three major advancements in FNA sampling technology performed by partner Neodynamics AB. First, a new FNA needle has been developed with a higher sample yield (i.e. the amount of sample material retrieved from each needle sampling) and improved surface treatment for higher ultrasound visibility (the operator typically uses ultrasound imaging to guide the position of needle tip to the suspect lesion of interest). Second, by adding mechanical energy to the needle during the sampling (longitudinal oscillatory movements in the needle direction as well as rotational movements of the needle) the penetration of the needle has been considerably improved, the operator can to a larger extent maneuver the needle during sampling, and the sampling can be performed by a larger group of clinicians (also those with less training in clinical cytology). Third, an anti-seeding procedure has been established and clinically tested. By applying radio-frequency current pulses passing the needle, the surrounding tissue in the needle tract will be coagulated and potential tumor cells that might have been spread from the cancer lesion by the sampling will be killed.

These advancements are likely to have a large impact on future diagnostics. The new FNA sampling technology contribute with considerable improvements on several of the most important aspects of diagnostic sampling: increased amount of sample material with the same degree of invasiveness (same type of needle etc), higher precision of positioning and penetrability is likely to increase the fraction of obtained samples that indeed are from the region of interest (and not from some neighboring tissue, yielding non-representative sample material), and not the least the risk of tumor cell dissemination is strongly reduced. These features make the FNA sampling technology very attractive for breast and prostate cancer diagnostics (in particular given that the small amounts of samples can be analyzed and yield conclusive diagnostics, as discussed below). However, it is also worth pointing out that the needle development can also to a large part be applied on core-needles as well. For the case core-needle would still be needed, mechanical oscillations and anti-seeding procedures can strongly improve the features in a similar way also for these larger diameter needles. The developed needle technology is also not only limited to breast and prostate cancer diagnostics. There is a range of different forms of cancer, where invasive sampling is to be avoided, but where FNA harnessed with the new technology can be considered as a possible way for diagnostic sampling. For several cancer



forms, for instance aggressive pancreas and ovarian cancers, it is not advisable to perform diagnostic sampling since these cancer forms are notorious for their tendency to form metastases and any form of dissemination or spread is to be avoided. The new FNA sampling technology may provide an opening for diagnostic sampling and treatment monitoring also of these highly malignant cancer forms. In general, the needle development within this project can have a considerable impact on all forms of diagnostic sampling, of cancer or infectious diseases, where risk of spread, accessibility, or risk of damage of vital tissues is an issue. Finally, the technology for diagnostic needle sampling also seems to have implications for treatment. On the one hand, if invasiveness, discomfort and risks with the sampling can be strongly minimized or even eliminated, it is also possible to perform repetitive sampling, providing continuous feed-back on treatment regimes and thereby facilitate individualized treatment regimes with possible better outcomes. Moreover, the anti-seeding can also directly be used as a treatment modality. As an alternative to surgical excision, Neodynamics AB has developed an instrument to treat breast cancer in a revolutionary cost-effective way. In a similar way as for the anti-seeding procedure used for diagnostic sampling, this so-called PRFA (preferential radio-frequency ablation) technique uses radio waves to locally heat up tumour tissue and denaturize cancer cells. It is minimally invasive, and designated to treat both benign and malignant tumours in an early stage when they have not evolved beyond walnut size. Furthermore the research of Neodynamics AB has shown that this treatment preferentially targets cancerous cells, resulting in the ability to treat tumour strands as well as to leave healthy fatty tissue unaffected. Two groups of patients are currently evaluated as target groups for the PRFA technique. First, PRFA is evaluated as an adjuvant treatment, where the patients are treated three weeks prior to planned partial or radical mastectomy (surgical removal of the breast). Second, PRFA is evaluated as an alternative to surgery in elderly inoperable patients. For both groups, the results this far are very promising and PRFA may thus open for a new attractive mode of surgical treatment.

## 2. Affinity molecule development

In the FLUODIAMON project, a set of affinity binders were identified for a selected group of target proteins, well functional also under the conditions relevant for FNA sampled cells. The project has shown that novel technologies for development of novel affinity proteins of different nature can be very valuable for future diagnostic applications. Compared to conventional, full-sized antibody binding reagents, the alternative binder protein classes used can provide advantages in terms of size, the preference to and selectivity of where on the target protein they bind. The two groups of affinity binders developed in the project (camel antibodies and affibodies) are considerably smaller than conventional antibodies. The small size of the marker molecules make them very valuable reagents for the optical techniques, where the high resolution offered may even make the actual size of the binder an issue (if the optical resolution is high, the size of the binder may eventually be the factor that limits the resolution of the image). The small size may also allow binding to sites on the proteins not accessible by larger binders due to steric reasons (i.e. that there is not enough space for a larger binder at the binding site on the protein surface). Since at least affibodies can be designed such that the position of the fluorophore marker molecule(s) can be specified to desired locations at the affibodies (e.g. via chemical synthesis routes), also the fluorescence properties can be modulated and optimised via the design of the affibodies. This controlled fluorescent labelling, including both number and position of fluorophores, was taken benefit of in the project, where the fluorescence properties of the samples are very important for the overall performance of the high-end fluorescence imaging methods used.

Taken together, the affinity molecule development in the project has demonstrated the feasibility of using novel class of affinity molecules for high-end optical imaging and diagnostics, and that the specific properties of the affinity binders can offer specific advantages compared to conventional antibody binders. The combined use of high-end fluorescence microscopy and these novel groups of affinity binders will have applicability for a wide range of cellular imaging, extending well beyond breast and prostate diagnostics.

### 3. Fluorophore development

In the FLUODIAMON project the access to high-performance fluorophores has been critical to the development of the diagnostic imaging procedures. The demands on the fluorophore marker molecules are for both high-resolution microscopy (nanoscopy) and high-sensitivity, multi-parameter, fluorescence-based characterization of proteins considerably higher than in more conventional fluorescence imaging and spectroscopy. Besides carrying specific chemical groups so that the fluorophores can be coupled to relevant affinity proteins or to other molecules of interest, and spectral properties suited to the optical instrumentation, the most important properties of the labels are high fluorescence quantum yield (i.e. that they offer an efficient conversion of the absorbed laser excitation into fluorescence emission) and high photostability (i.e. that the fluorophores can survive strong light exposure by the lasers). Our project has demonstrated the feasibility and strength of using high-end fluorescence imaging methods for breast and prostate cancer diagnostics. However, the applicability of these imaging methods goes far beyond this particular application. No matter the application, the performance rely on the use of appropriate fluorophores, and the developed fluorophores in this project, and the experience coming from the development and use of these, will be highly useful in a range of applications in diagnostics, cell biology and biosciences in general.

### 4. Cancer proteomics

It was not within the scope of this project to identify new tumour markers, specific for early stage prostate and breast cancer disease. Instead, the strategy was to demonstrate the feasibility of the microscopic techniques based on already identified tumour markers, expected to reflect specific cellular deviations, both with respect to cytomorphology, spatial distribution, and expression levels. In a sense, the project merged knowledge about tumour-specific markers with the use of highly sensitive, high resolution imaging methods by which the markers can be detected and identified down to a single molecule level, and by which their molecular status can be analysed in the context of individual sample cells. In particular, the project has shown for a limited number of selected protein targets that apart from plain deviations in the amount of the proteins in the FNA sampled cells, the diagnostic strength can be considerably increased by further extending the protein characterisation to include also high-resolution spatial distribution and high-sensitivity expression level profiling of several proteins in the individual cells at the same time. The impact this project will have in the context of protein biomarkers thus essentially lies in the notion that not only expression levels (the amount) of certain individual proteins matters. Additional valuable diagnostic information can be obtained from analysing expression profiles of multiple proteins and the spatial distribution patterns of proteins. The proteins analysed in this way need not necessarily be specifically up- or down-regulated. Also other “every-day” proteins may be of interest to analyse in the cells if they show a different spatial distribution patterns upon malignant transformation of the cells. In the project, we have introduced a



novel concept of “tumor biomarker” focusing primarily on its spatial distribution within the cells and how its expression is related to other proteins expressed in the same individual cells. The concept has been demonstrated to work, and several proteins have been identified as possible biomarkers in this sense. Our results suggest new aspects to consider when identifying new tumor markers in the future.

## 5. Development and optimization of high-end fluorescence methods for diagnostics of FNA-sampled cells

For the diagnosis of breast and prostate cancer today’s standard procedures for diagnostic sampling are core needle or surgical biopsy. However, due to their mechanical crudeness, these procedures are highly traumatic and encumbered with a significant risk for tumour cell seeding in the (multiple) needle tracts. In contrast, fine-needle aspiration (FNA) offers a minimally invasive, extremely patient-friendly extraction of cellular material. However, the amount of material obtained by FNA is often too sparse for a conclusive diagnosis, using presently available methodologies. A major strategy of this project was to overcome these drawbacks of FNA, by maximising the information obtainable from the sparse FNA-sampled material, using and further developing cutting-edge methods for fluorescence-based sub-cellular imaging. Four different methodologies were developed in the project, and although only two of those were combined in the final characterization of the FNA sampled cells, the results for all four methodologies will have an impact on future cellular imaging and diagnostics.

### 5.1. Stimulated emission depletion (STED) microscopy

In this project, ultrahigh-resolution fluorescence microscopy was introduced and adapted for diagnostic purposes on individual sample cells. New STED instruments have been designed and established that are relatively cheap, easy to operate, and with true dual-colour high-resolution imaging capabilities. The design of the new STED instruments has considerably lowered the threshold to start with STED imaging both economically and in terms of the knowledge and experience required from the operators. The demonstration that ultrahigh resolution optical imaging can be applied for subcellular diagnostics opens for a range of diagnostic applications of STED and related techniques. In this project we have shown its capabilities for breast and prostate cancer diagnostics, but the range of diseases where high-resolution protein distribution patterns in the sampled cells carry disease specific information is most likely quite broad. The ability to optically image cells and subcellular features in a far field microscope (i.e. without being confined to surfaces or interrogation of the sample at very close distances) is indeed a major step in cell biology. The development of the STED technique within the project has lead to the first demonstration of its use for cancer diagnostics, but also paves the way for future applications of ultrahigh resolution microscopy in cell biology and biomedicine in general.

### 5.2. Transient state (TRAST) imaging

TRAST imaging addresses information not retrievable from the normal fluorescence parameters (fluorescence lifetime, intensity, wavelength and polarisation). In the project, several realisations of the technique in various setups have shown its broad applicability, and the ability to retrieve information coupled to low frequency molecular encounters or metabolic states have been demonstrated. Cancer cells often display a different metabolism than their non-malignant counterparts, and this seems to be reflected in the transient state parameters recorded by TRAST imaging of the cells. This may open for entirely new aspects and new parameters to diagnose cancer, and to follow effects of treatment. Since the metabolic state, and the regulation of the metabolism is one of

the first things to be influenced in cells upon disease and toxic influence, the ability to monitor the metabolic state and the metabolic regulation of cells is also likely to open for applications in toxicology, drug tests and several other applications in fundamental cell biology.

### 5.3. Multi-parameter fluorescence imaging (MFDi)

In this project, a new general strategy based on multiparameter fluorescence detection (MFD) to register and quantitatively analyse fluorescence images was further developed for imaging purposes (MFDi). The necessary hard-ware and soft-ware were developed and the sample preparation procedures required were established to be able to extract sufficient diagnostic information out of sparse, or even individual FNA-sampled cells. Thereby not only one target protein could be monitoring in the cells, but several different proteins at the same time. The developed MFDi technology has been demonstrated as a useful diagnostic imaging method, and a similar use of MFDi for subcellular diagnostics can most likely be extended to a broader range of diseases in the future. Given the wealth of information contained in the data registered by MFDi (each detected fluorescence photon is characterised with respect to its spatial coordinates of origin, arrival time after the laser excitation pulse, how it has been filtered by wavelength and polarisation filters prior to detection, time intervals to neighbouring photons, and the overall time scale of detection, altogether corresponding to an eight-dimensional parameter space), the imaged cells can be viewed with respect to a manifold of aspects. This also opens for a broad use of MFDi in cell biology in general. Parts of the hardware and the concepts for fluorescence data acquisition and processing included in MFDi can also be used outside of the context of MFDi. Partner B&H has developed and adapted several instrument components as well as concepts of this kind for commercial use in this project.

### 5.4 TPX technology

The TPX-technology, originally developed by the UTURKU group and commercialized by Arctic Diagnostics Ltd is extremely well suited to measure concentrations of target molecules from minute samples with high precision and sensitivity. The work in the FLUODIAMON project has been building the technology and methodological basis for TPX with cellular samples. The TPX technology assay development within the project has demonstrated that the assay should be readily integrable to the FNA sampling protocol of the project. A small aliquote of the original sample could be set aside for TPX analysis, providing complementary diagnostic information to that obtained from the cellular imaging. In addition, the assay development within the project has also set the ground for developments of the TPX technology towards a new method for detection of MRSA and the development of TPX assay for bacterial counting have further paved the way towards quantitative and sensitive TPX assays of cellular matrices. Along with the already commercialized pathogen detection assays these results highlight the potential and versatility of the TPX technology.

## 6. Bioinformatics

Bioinformatic tools were developed to analyze the large number of images taken by STED and MFDi on FNA-sampled cells from suspect breast and prostate cancer lesions. The high-resolution and high-content of the images required advanced computational methodologies(bioinformatics) to translate the image data into clinically useful information. Before this project there were no published computational methods for preprocessing and analyzing STED and MFDi images at the scale required in clinical settings.

With the goal to identify markers that predict whether the imaged FNA-sampled cells were normal or cancerous, partner UH developed computational methods based on strong prediction methods developed in the machine learning community. The efforts resulted in clear improvement in diagnosing patients with the FNA procedure. The development of computational methods for cutting edge, high-resolution and multiparameter microscopy has resulted in 1) freely available software for analyzing large amounts of images in a systematized fashion, 2) novel methods that are widely usable in various experimental settings and microscope techniques, and 3) clinically important predictors that can aid in breast cancer patient diagnosis. In the project, in particular the study of microfeatures has emerged as a novel powerful approach in the analysis of images, opening new possibilities in all studies, where the resolution is large enough for imaging subcellular features. The outcome of the bioinformatic analysis demonstrates the power of multiple marker and/or ultrahigh resolution spatial distribution pattern analyses of proteins on a subcellular level as a new path for early, sensitive cancer detection, but the developed image analysis tools also have a more general applicability.

## 7. Overall diagnostic procedure

In this project, we have developed and successfully demonstrated a joint methodology for objective, quantitative sub-cellular diagnosis of early breast and prostate cancer disease. By applying analyses based on novel state-of-the-art optical microscopy techniques much more information can be extracted and even quite small amounts of diagnostic material is sufficient for reliable diagnoses of breast and prostate cancer. Thereby, the amount of diagnostic material obtainable via minimally invasive fine-needle aspiration (FNA) is sufficient and more invasive core-needle or surgical biopsies can be avoided. The potential impact following from this successful development of the methodology can not be overestimated.

- It can strongly contribute to an increased cure and survival for breast and prostate cancer patients.
- It should be able to provide valuable support guiding decisions whether therapeutic interventions are to be undertaken, and if so, when and to what extent.
- It should minimise the risk of cancer cell seeding, patient discomfort, and other side effects such as pain and infections.

Outside of breast and prostate cancer diagnostics and progression monitoring, the methodology should be applicable also to other diagnostic fields facing the same principal conflict between minimal invasiveness and the information content needed for a clear diagnosis.

As a major, overall result of this project, we have obtained a diagnostic methodology for breast and prostate cancer that combines diagnostic reliability with minimal invasiveness and a negligible risk for sampling-related side effects, in particular cancer cell seeding and spread. This combination is impossible to obtain with other presently available methodologies. The ability to obtain conclusive information with minimal invasiveness and without risk of cancer cell seeding and spread also enables safe and reliable repetitive sampling and thereby quantitative monitoring, analysis and prediction of disease progression over time, e. g. for follow-up of pre-malignant lesions or for evaluation of response to therapy. The developed methodology can therefore also be expected to form the basis for important diagnostic tools supporting and guiding possible therapeutic interventions. It is also reasonable to predict that the proposed methodology would offer significantly improved and

profoundly new possibilities for monitoring and analysing, on a subcellular level, molecular mechanisms for tumour development and effects of various therapies. The project results thus clearly can be expected to contribute to the development of new therapies and pharmaceuticals, and in particular therapies based on molecular and cellular mechanisms.

In the forefront of the potential impact that this project may bring about is the fact that breast and prostate cancer belong to the most frequently occurring malignancies in women and men, respectively, in the western world. When possible, detection of these diseases at the earliest stage results in a 15 year survival rate of more than 90 %. In contrast, survival of breast cancer patients exhibiting advanced stages has not improved during the last 20 years. Thus, for improving survival of patients suffering from breast and prostate cancer, the most important task is to detect, diagnose and treat malignancies as early as possible. The benefit to patients that patient-friendly and secure sampling methods for breast and prostate cancer would bring about, offering the necessary diagnostic sensitivity, specificity and reliability, is enormous.

Further, this project improves one of the most important quality aspects in the diagnostics and treatment of patients with breast and prostate cancer - that sample extraction may cause severe side effects and even worsen the outcome. At this stage, we have developed, demonstrated and clinically verified the methodology. Already at this stage, the procedure is objective, and with additional refinements of our combined procedure it is likely that the diagnostic sensitivity and specificity can be even improved compared to today's standards.

## **Dissemination activities and exploitation of results**

The activities within FLUODIAMON have throughout the project generated a continuous flow of dissemination activities.

In total FLUODIAMON have to-date resulted in 24 peer-reviewed articles, all published in well recognised scientific journals. Ten additional publications are either already submitted or soon to be submitted.

Moreover 7 PhD theses and 4 MSc theses have to date been achieved with support from the project. In addition, several PhD students will graduate in the next 1-2 years, and their PhD theses are also to a large extent a fruit of and supported by the project.

Among dissemination activities one can also mention several open access software packages for classification and image analysis, interviews in newspapers and TV, exhibitions, and well over 50 invited oral presentations and posters at international conferences and workshops.

On the commercial side, 10 specific patents or patent applications, related to needle sampling technology, fluorophores and detector and imaging technology, have been either generated or have directly profited from the activities and results within the project. A corresponding development of several different commercial products within these areas has also commenced as a result of the project.

In conjunction to the last consortium meeting, a workshop was arranged in Stockholm on the 2nd of December 2011 (“Ultra-high resolution and ultra-sensitive fluorescence methods for objective sub-cellular diagnosis of early disease and disease progression in breast and prostate cancer”). This workshop was very well attended, and brought researchers from the full range of disciplines covered within the FLUODIAMON project. Finally, last but not least, the progress within the project now allows us to disseminate the results also to a wider audience. From a patient perspective, considerable benefits have been demonstrated. A first seminar with breast cancer patients has been arranged by partner KI to disseminate these results. Given the positive and encouraging results from the large-scale assessment and clinical verification, a larger meeting with breast cancer patients and clinicians is already planned together with the major patient organisations in Sweden during the autumn 2012, and similar activities will also be planned in Germany.

Taken together, the FLUODIAMON projects has yielded a very broad range of results, which has been disseminated scientifically, on the commercial side, and not the least due to the large impact on future cancer diagnostics, towards patient organisations and clinicians.



**List of Websites:**

project website:

[www.biomolphysics.kth.se/fluodiamon](http://www.biomolphysics.kth.se/fluodiamon)

Contact details of the coordinator:

Prof. Jerker Widengren

KTH, Dept Applied Physics / Exp Biomol Physics

Albanova Univ Center

106 91 Stockholm

phone: +46-8-55378030

fax: +46-8-55378216

email: [jwideng@kth.se](mailto:jwideng@kth.se)