DEVELOPMENT OF PAPER-BASED MALE FERTILITY DIAGNOSTICS FOR MOBILE HEALTH SCREENING

by

Tian Kong

A thesis submitted in conformity with the requirements
for the degree of Master of Applied Science

Department of Mechanical and Industrial Engineering
University of Toronto

Development of Paper-based Male Fertility Diagnostics for Mobile Health Screening

Tian Kong

Master of Applied Science

Mechanical and Industrial Engineering Department
University of Toronto

2018

Abstract

Mobile health technologies are emerging for population health screening applications. Current male fertility diagnostics lag these other health arenas in many aspects. This thesis provides novel methods to aid in and commercialize paper-based male infertility diagnostics. First, to quantify cell concentration and viability a smartphone-based colourometric imaging method was developed that corrects for environmental lighting conditions and provides quantitative measurements from paper-based tests. A fluid delivery system for the colourimetric paper-based assay was also developed. Clinical testing revealed that the device had false-positives from bacteria in some semen samples. To mitigate the false-positive issues, a companion method was developed to purify semen prior to testing. Collectively these works advance the technology and applicability of portable male fertility diagnostics in aid of fertility challenges worldwide. The Appendix reports additional recent work on the development and testing of a deep learning model to correlate single cell morphology parameters to DNA integrity analysis.

Acknowledgments

First and foremost, my deepest gratitude goes to Professor David Sinton, my supervisor who provided me with constant encouragement and guidance during my master study. Without his consistent and illuminating instruction, I would not have been able to develop the academic competence and research acumen.

I also would like to sincerely appreciate the generous support from my fellow lab group members: Professor Biao Zhang, Dr. Jae Bem You, Dr. Yihe Wang and Dr. Christopher McCallum. Their working ethics and positive attitudes will have a profound impact on my future life.

Finally, I would like to thank my beloved family and friends for their trust, altruistic love and unconditional support.

Table of Contents

List of Tables

List of Figures

Chapter 1 Foreword

1.1   Motivation

The sensing of solution concentration is an important topic of research, especially in medical pre-diagnosis. Numerous products have been invented for patients at home to monitor health condition indicators such as concentration of glucose, achieving ease of use and disposal. Compared to sophisticated analysis devices in the lab, the most widely used tools, like chemical sensing strips, can only provide a qualitative result. However, the rising awareness of global health issues and the demand for more accurate data for clinical analysis require new solutions toward this matter.

Due to the growth in high-end smartphones, fine camera peripherals, better mobile computing, and substantial image processing power are now more widely accessible with less cost. Therefore, the integration of cheap mobile technology with paper-based microfluidic chips is a growing research direction, especially in the context of colourimetric detection of concentrations on a smartphone. In my work I am pursuing the measurement of live and motile sperm concentrations using existing sensing strip technology on a smartphone.

Infertility is a growing global concern, about half of which is due to the male factor. At the population level, recent geographical and demographic trends are exhibiting reduced semen quality. Delayed pregnancy, aging, and modern lifestyle, including obesity and smoking are a few of the factors driving the global infertility cases. Examples of environmental and physical causes are exposure to pollutants, prostate cancer, and diabetes. Psychologically, men can be self-conscious regarding their fertility potential, and avoid visiting a urologist for infertility testing. Also, when male infertility is undetected, a couple's infertility burden can inaccurately be assumed to be female-factor.

The global trend of rising male infertility motivates immediate attention and highlights the need for affordable and accessible semen diagnosis. To find effective medical solutions to this worldwide infertility problem, deployable, cost-effective, point-of-care diagnostic devices are needed to monitor male fertility indicators in a large scale, especially for resource limited areas.

1.2   Thesis Overview

This section outlines the primary contributions of my thesis studies.

Chapter 2 introduces and overviews the conventional techniques for male fertility analysis, current technologies for home-based analysis, the evolution of paper-based devices and mobile imaging and colour calibration.

Chapter 3 presents a smartphone colour sensing method that can provide quantitative data for further analysis while retaining the advantages of qualitative paper-based methods such as portability and low cost. This method involves utilizing the continuous LED light at the back of the phone as the dominant light source for imaging and integrating black-white calibration with ΔRGB for a single-colour shift analysis. I demonstrate that the adoption of this method can compensate for ambient lighting conditions and the significant variations between common smartphone models. Lastly, this imaging method is demonstrated in simple paper-based yeast viability quantification application.

Chapter 4 presents the achievements in further developing, testing and debugging a paper-based male fertility diagnostic. In preparation for use with a smartphone, our groups previous paper-based device is optimized for smartphone health diagnosis with modifications on chip design, better choice of colourimetric assay and further customization and characterization. The development and fabrication of two fluid delivery interface designs are introduced to substitute pipetting and to simplify the end users experience. In clinical testing I showed paper-based approach suffered from false positives with bacteria in semen samples being registered, chemically, as motile sperm. I developed a solution to this false positive issue, but the solution increased the cost and user requirements making the technology overall less attractive from a commercial perspective.

Chapter 5 summarizes the primary works presented in this thesis, describes the significance of each work in the progress of commercialization of the paper-based male fertility device, explains why it has not been commercialized yet and presents the next steps for improving both diagnostics and machine learning for male fertility studies.

Appendix D presents the study on the correlation between sperm morphology and DNA integrity at single cell level. The procedures of experiment include functionalization of the substrate, sample preparation for DNA stain and acrosome stain and cell imaging under confocal microscope. A five-layer deep learning neural network was utilized to create a non-linear regression model with significantly better accuracy than regular linear regression.

2 Introduction

2.1 Paper-based Male Fertility Diagnostics

2.1.1 Background and Motivation

Infertility is a growing global public health issue with significant psychological, social and economic impacts, affecting more than 70 million couples worldwide.[1–3] In Canada, one in six couples experiences infertility, a 2-fold increase since 1992.[1] Male infertility is responsible for about 50% of the cases internationally and it in particular solely contributes to about 30% of infertility cases(Fig. 1b).[4–6] This deterioration in fertility is a phenomenon in many different areas around the world, for instance, the total fertility rate, defined as births per woman, in all European Union, East Asia and Pacific and North America are below the limit (2.1) to sustain their current population levels (Fig. 1c).[7]
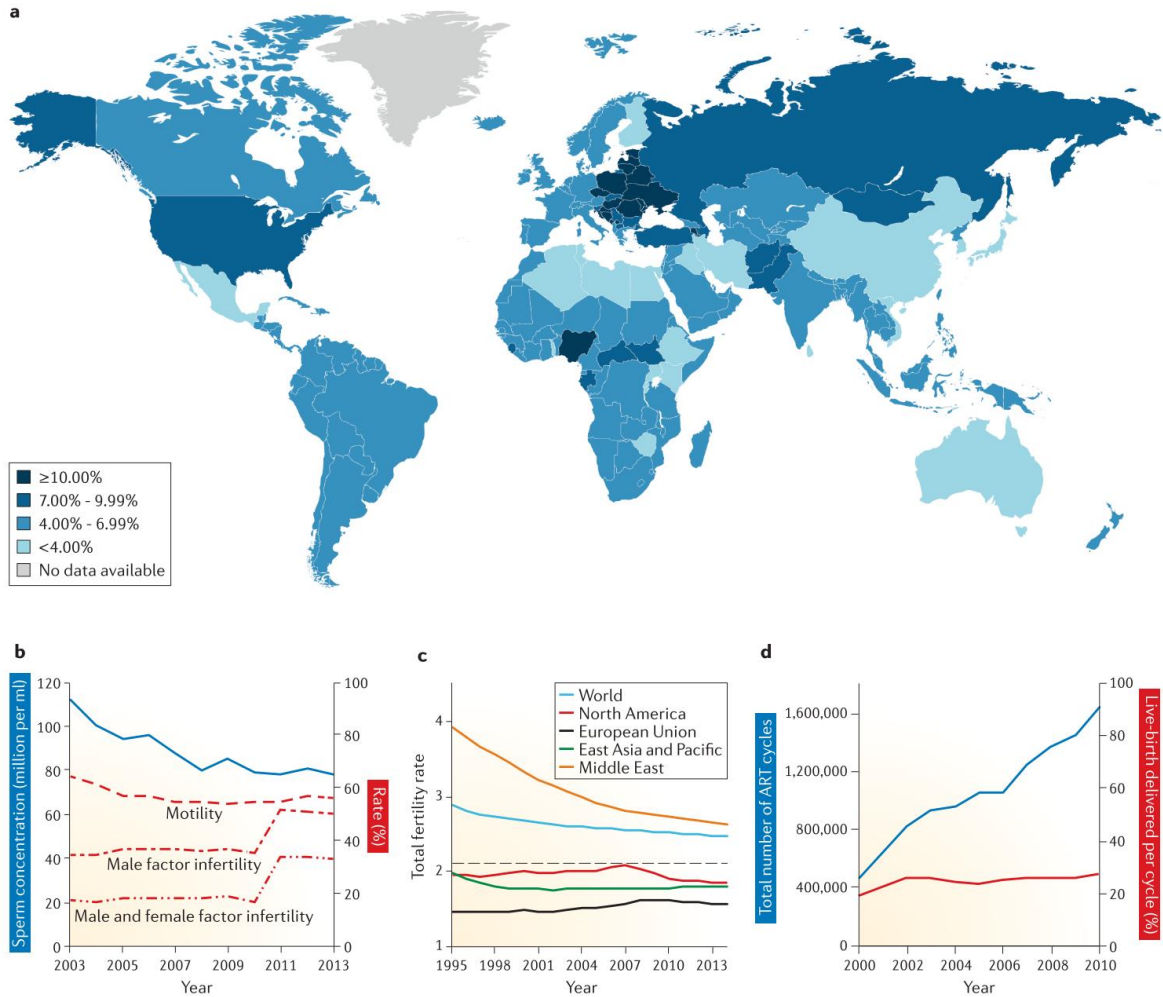
Figure 2.1. Male infertility and assisted reproduction globally. a) Map of the prevalence of male factor infertility in the world. b) Decreasing semen quality and increasing male factor infertility in the United States between 2003 and 2013. c) Total fertility rate between 1995 and 2014 for different regions and globally. d) Increasing number of ART cycles and plateaued success rate of ART per cycle in the United States between 2000 and 2010.[7] *Figure reproduced with permission of the rights holder, Springer Nature*

At the population level, recent geographical and demographic trends are emerging, exhibiting reduced semen quality, especially in sperm counts, viability and motility, for both fertile and infertile males.[8–10] More specifically, in USA, sperm concentration has declined 1.5% per year and a decreasing trend in motility in young adult men has been observed between 2003 and 2013.[9,11] In France, semen quality has also declined by ~1.5% each year from 1988 to 2017.[10] In Denmark, 20% of young men are estimated to have a severe decrease in semen quality and a higher portion of semen samples exhibit morphologically abnormal spermatozoa above 90%.[11]

This widespread decline is attributed to several biological, environmental and economic factors in modern lifestyle including obesity, smoking, exposure to pollutants and low relative income of young men relative to their parents[12–14]. Many male patients are not willing to do semen analysis because of embarrassment and anxiety during the semen sample collection in the laboratories and the clinics. The global trend of rising male infertility drives the demand and opens-up the market to transition semen diagnosis from research laboratories and clinics to home-based kits that make semen diagnosis affordable and accessible. Such kits could be used in developed regions as well as in resource limited environments where conventional assisted reproductive technologies (ART) are either unavailable or very costly.[3] By moving diagnostic processes more generally out of the clinic environment towards resource-limited settings, the World Health Organization (WHO) has expressed the need for these inexpensive, ease-of-use, and disposable medical devices.[15,16] WHO is rapidly expanding access to ART in developing countries with the goal of universal access.

The conventional laboratory-based semen analysis techniques, such as computer-assisted sperm analysis (CASA), counting chambers, and vitality assays, suffer from limitations that prevent their widespread application: long and expensive testing procedures requiring capital-intensive infrastructure.[17] As a result, male fertility tests are performed rarely, and only when recommended from within the healthcare system. A rapid, point-of-care and low-cost assay for semen analysis would be transformative for fertility patients. In addition, low-cost assays could enable monitoring of fertility potential over time, for instance, allowing one to track progress as a result of exercise, changes in diet, or smoking cessation. Mobile health technologies, further enhanced by ubiquitous smartphone technologies, are emerging for a wide variety of health screening applications, providing remarkable advantages of easy-to-use and portability.[18,19] However, male fertility diagnostics lag these other health arenas, leading to a vacancy for a mobile semen diagnosis technology. This opportunity is particularly significant for point-of-care diagnosis by patients at home, as well as rapid screening at fertility clinics.

2.1.2 Paper-based Microfluidic Testing Devices and Channel Designs

Low-cost diagnostic devices have recently gained attention, in part due to the rising cost of healthcare and the urgent and huge potential demand in resource-limited settings. Paper-based

microfluidic chips, especially with functionalized papers, have emerged recently combining the sensitivity and specificity with the low cost, readily available and ease of use of paper-based platforms.[20–23] The development of paper-based microfluidic analytical devices has realized many of the goals of point-of-care (POC) technologies. POC methods are attractive and have been developing quickly for home-based immediate medical diagnostics, for community-level disease monitoring, and for disease discovery at an early stage. Examples of asymptomatic diseases that can benefit from such POC technologies include a variety of cancers[24], and male infertility[25] among other conditions. Paper-based assays have gained their popularity especially in the biomedical area, for instance, paper ELISA[24,26], HIV chips[15,27], paper microfluidic immunoassays[27–29]and colourimetric diagnostic test strips.[18,30–32]

The Crooks Group has been exploring the design of three-dimensional (3D) paper-based microfluidic analytical devices and the wax patterning technology, and they have published a few papers regarding these subjects. Renault et al.[33] did fundamental studies on 3D wax patterning, involving 2 steps: (i) designing and printing wax patterns on the paper substrate and (ii) wax melting into the paper substrate. The traits of all three types were studied, including open channels, hemichannels and fully enclosed channels. Open channels and hemichannels can reduce the number of layers required for a 3D paper-based analytical device. Fully enclosed channels can prevent liquid evaporation and contamination. Renault et al.[34] also compared the flow rate in paper channels and in hollow channels. Their results showed that the fluid traveling speed in a hollow channel was 7 times faster than in a paper channel. This year, Channon et al.[35] developed this multilayer microfluidic paper-based analytical device with two hydrophilic layers on the top and bottom of a hollow channel to enhance the liquid flow rate. They studied the effect of channel height, tilting angle and the sample volume on the flow rate in a straight channel. Lastly, they applied their design in the stripping analysis of cadmium.

2.1.3 Conventional Techniques for Male Fertility Analysis

Quantitative semen analysis is critical for male infertility diagnosis. Conventional techniques for semen analysis include CASA [36], counting chambers, and vitality assays like hypotonic swelling and dye exclusion. The CASA system is currently the gold standard for semen diagnosis, using advanced optical microscopy to assess sperm concentration and motility via automatic tracking

of sperm in a digital image sequence[36,37]. The description and limitations of these techniques are organized in the table below.

Table 2-1: Comparison among conventional semen analysis techniques and their disadvantages

Semen Analysis Techniques	Advantages	Disadvantages
CASA System[36,37]	- Using optical microscopy - Digitizing sperm images	Subjective and/or equipment-specific measurements
Counting Chambers	- Manual microscopy inspection - Visually inspect sperm motility and concentration	Complicated testing procedure
Hypotonic Swelling[38]	- Manual microscopy inspection - Water influx induced pressure gradient	Expansive testing equipment
Dye Exclusion	- Use dye to selectively label live/dead sperms - Manual microscopy inspection	Limited to in-clinic use, not portable or suitable to home use.

2.1.4 Technologies for Home-based Semen Analysis

A variety of commercially available products for sperm concentration, viability and motility detection are currently on the market. SpermCheck developed lateral flow assay for semen analysis, however, the output is binary and based on total sperm concentration. SpermCheck and the related offering, SwimCount, can only detect live sperm concentration or motile sperm concentration within a range, using antibody reaction for colourimetric signals when sperm concentration is greater than a certain threshold. SpermCheck sets its sperm concentration threshold at 20 million sperms/mL with 98% accuracy and it retails for $39.99 USD.[39,40] Trak requires the uses of a small centrifuge to estimate sperm concentration on a microfluidic device. Trak categorizes sperm concentration into three subgroups, low, moderate and optimal with two dividing points as 15 and 55 million sperms/mL. Trak claims to have 97% accuracy (comparing to CASA as the gold standard). However, the Trak device cost $199.99 USD for the device and

four tests.[41] The semen diagnostic assay developed by the Harvard group provides a hardware microscope addition for smartphones to enable both quantitative sperm concentration and motility measurements. Their automated smartphone-based semen analyzer can provide a semen quality evaluation with 98% accuracy with regard to WHO guidelines for morphology.[42] The detailed comparison among these home-based semen diagnostic devices are demonstrated in the table blow. While all the devices have advantages and disadvantages, none are at a cost point that would make them accessible to patients in resource-poor environments, or make them an option for wide population based screening.

A former PhD student in our group designed paper-based chips with each test spot 4mm in diameter.[25] The chip required a minimum of 3 μL of solution at each test spot and a layer of laminate film was applied to prevent evaporation. These fertility-based chips are appealing and innovative due to its rapid testing and low cost. Most importantly, they can quantify three critical sperm parameters with one paper-based unit: live sperm concentration, motile sperm concentration, and sperm motility. Only 3 μL of sample is needed for each of the live sperm concentration and motile sperm concentration testing spots, and an additional of 3 μL of buffer solution is needed to pre-hydrate the motile sperm concentration test zone. The material costs are very low as well, estimated to be on the order of US$0.05 per device. The testing procedure is simple and the output is colourimetric, a dye colour change. A limitation of this device is the need for an optical scanner to accurately read out the results. While scanners are common in labs, the requirement of an optical scanner significantly reduces the commercial potential, and attractiveness, of this paper-based test.

Table 2-2: Competitive table of commercially-available semen diagnostic devices currently on the market.

Technology	Scientific basis	Quantitative results?	Only takes a few minutes or less?	Affordable regular tests at home (<$1 per test)?	Accurate App-based quantification?	Compatible with any smartphone?	Free of contamination on the phone?	Off the shelf	Cost per test (CAD)
SpermCheck®	Antibodies-based	✗	✓	✗	✗	N/A	N/A	✓	$50
SwimCount®	Colorimetric reaction-based MTT dye	✗	✗	✗	✗	N/A	N/A	✓	$70
FertilitySCORE®	Colorimetric reaction-based Resazurin dye	✗	✗	✗	✗	N/A	N/A	✓	$35
Trak®	Centrifuge-based	✓	✓	✗	✗	✓	✓	✓	$65
Yo Sperm®	Optical tracking-based	✓	✓	✗	✓	✗	✗	✓	$67
Harvard semen analysis	Optical tracking-based	✓	✓	✗	✓	✗	✗	✗	N/A Manufacturing cost less than $6

2.1.5 Colourimetric Cell Viability Assays

3-(4,5-dimethyl thiazol-2-yl)-2,5-diphenyl tetrazolium bromide (MTT) is a commonly used dye for cell proliferation assays. Its capability in evaluating the human sperm viability and its optimal conditions for MTT viability assay have been reported.[43] MTT is a positively charged

tetrazolium salt, so it is reduced primarily intracellularly by intracellular reductant NADH derived mainly from the live cells mitochondria.[44]

2,3-bis-(2-methoxy-4-nitro-5-sulfophenyl)-2H-tetrazolium-5-carboxanilide (XTT) is a relatively new dye that measures the colourimetric change of light yellow tetrazolium to orange formazan product, which in contrast to other tetrazolium salts like MTT, is water soluble. Electron coupling agents, like menadione (MEN) and phenazine methosulfate (PMS), have been added into XTT to promote XTT reduction for reagents such as human tumor cells[45], antimicrobial active plant species[46], and quantification of microbial growth/inhibition[47]. Unlike MTT, XTT is a negatively charged tetrazolium salt, thus it is most likely reduced at the cell surface.[44]

2.2   Mobile Image Capturing and Colour Calibration

2.2.1 Background and Motivation

The development of a smartphone-based colourimetric detection and a colour correction algorithm is another essential step to achieve a ready-to-go, point-of-care, commercial mobile health diagnosis. The combination of paper-based chips and a colourimetric assays enables rapid and inexpensive diagnostic testing. Colourimetric results can be viewed by naked eye, but it is difficult to precisely quantify the change in analyte concentration.[24,48] The conventional devices for precise colourimetric detection results are microscopes, spectrometers[18], scanners[49], digital cameras[50], and digital colour analyser[51], all of which require a connected computer for analysis, communication and transmission of results. To facilitate market adoption, a reliable, accessory-free and ease-of-use imaging and colour correction method are required for immediate POC sensing. The camera on a smartphone, without any other accessories, acts as the optical sensor of the colourimetric assay, and the analysis can be performed on the smartphone itself instead of a computer.

Realizing this potential, many studies have been completed on the adoption of cellphone camera as an imaging tool in colourimetric results detection, calibration, and representation. Cellphones are ubiquitous and they are becoming a viable alternative to traditional systems for fluorescent imaging[52,53], imaging analysis[31], and colourimetric diagnostics[24,54,55], because they are equipped

with digital colour cameras, internal memory, and central processing units.[56] However, the use of smartphones has not yet been fully commercialized for POC analysis for 4 reasons:

1. The changing lighting conditions are very hard to control. The colours can be detected differently under different lights. Recent results show greater measurement errors especially for outdoor illuminating conditions.[57]

2. Small colour changes can be another challenge for accurate colour measurement. Image analysis for this may require methods other than RGB intensity values.[31] The position and angle for image capture are also critical for precise colour recognition.[18]

3. Smartphone cameras often contain integrated automatic colour balancing functions which can optimize the quality of photos produced, yet these functions can also impair the accuracy for quantitative measurements by adjusting the RGB intensity values at different ratios.[31]

4. Most of the smartphone-based analyzers require an accessory for accurate measurements.[42,58] But these attachments have to change with phone manufactures, models, and camera technologies – an ever changing landscape. These attachments also increase the cost of the whole device.

More research is needed on smartphone viability as a photodetector for colourimetric detection and analysis.

2.2.2 Colourimetric Detection and Quantitative Analysis Methods

There are a range of colour measurement methods including direct RGB measurement using ΔRGB[26] or hue parameters[59], taking a ratio between red and green channels[55], or obtaining a colour ratio among 3 primary colours[57]. These methods have different adequacy and sensitivity under different conditions, and they are application specific. Colour systems can be classified by different types of transformations. RGB colour space uses intensity values for red, green and blue primary colours to describe the colour of interest. The International Commission on Illumination (CIE) 1931 XYZ colour space is a standard general space which uses x and y to quantify chromaticity of a colour and z to represent the brightness of a colour.[31] CIE 1931 XYZ

colour space is normally considered as a suitable method for two-colour shift (i.e. yellow to purple, clear to green to yellow, etc.) and particularly when the range of colour to be detected is large.[31,60] HSV and HSL colour space models are hue parameter oriented. A pixel in HSV can be defined by its hue(H) factor, saturation(S) factor and value coordinates/brightness(V).[18,59] Colours can be translated from one of these colour space descriptions to the other.

Jia et al.[19] brought attention to the effect of imaging distance and rotation/ tilted angle of cellphone camera on imaging results, while considering the variation in ambient light conditions. They tested "Four spots" method or "Two spots" method (Fig. 2.2) for image calibration with Equation (1). "Four spots" method was chosen for its simplicity and effectiveness, compensating the difference caused by cellphone models, lighting conditions and imaging positions.
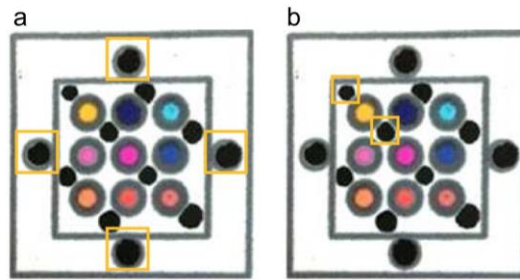
Figure 2.2. Left "Four Spots" Correction Method, Right "Two Spots" Correction Method[19] *Figure reproduced with permission of the rights holder, Elsevier*

To calibrate the lighting conditions, one first wants to know what is the real RGB value of Black (0, 0, 0) and White (255, 255, 255) being captured through the camera through measuring the RGB values of the black and white reference spots. After extracting this information, they use Equation 1 to correct the R, G and B value separately and compute the corresponding pixel value for the point of interest.

$$R_{corr} = (\frac{256}{R_w - R_B})(R_{measurement} - R_B) \tag{1}$$

As an imaging device which senses the light reflected from sensor surface, the smartphone camera captures RGB primaries only, with specific wavelength around 630nm, 550nm, and 450nm.[18,19] All of the other colours detected are a mixed ratio of these three primary colours. The RGB values from the sensing signals can be corrected with respect to the RGB measurements of

background/reference signals. Figure 2.3 shows that the correction equation adjusted the variant colours of the same sample before correction. The black reference bars at the extreme left and the extreme right of the urine strips is used in correction.
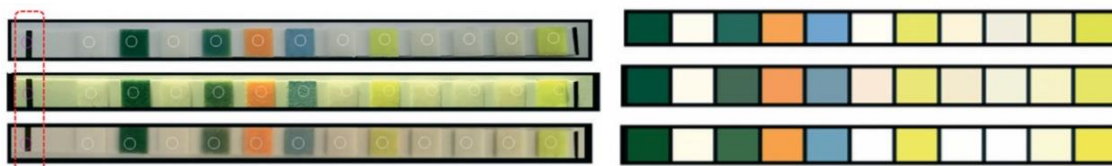
Figure 2.3. Images of a urine strip under indoor fluorescent light (top), outdoor sunlight(middle) and indoor low light intensity (bottom) conditions before(left) and after(right) adopting colour correction[18] *Figure reproduced with permission of the rights holder, Royal Society of Chemistry*

Hong and Chang[18] developed a smartphone-based POC urine diagnosis that converts the colour on the urine test paper to the concentration of the analyte. They measured their samples under different lighting: an indoor fluorescent light, outdoor sunlight and an indoor low light intensity. According to their paper, the intensity change was best characterized by the saturation (S) value in HSV colour space. They claim that the changes of wavelengths due to different illuminance do not correspond to the change of the relative ratio of the primary colours. However, a polynomial relationship can be found between the spectral wavelength and the hue value in HSV colour space, converted from RGB channels.[36,57] Oncescu, et al.[61] adopted the hue value in HSV colour space to quantify colour measurement without change in ambient light conditions. In their work they used only one colour coordinate (H or S out of HSV) for the calculation of concentration (while a colour is defined by three coordinates of either RGB or HSV). They indicate that adopting a complex algorithm for the combination of S and H could improve the precision, particularly with variant light conditions.

Shen et al.[31] used a smartphone to image colourimetric pH indicator strips and then the CIE chromaticity values for mixed colour quantification/pH concentration quantification. Figure 2.4a plot the colour change of the pH strip in a CIE 2D chromaticity diagram. Figure 2.4b related the analyte concentration to the chromaticity values x and y in a 3D view. Figure 2.4c demonstrated high accuracy in pH measurements with linear response ranges of 1-12.
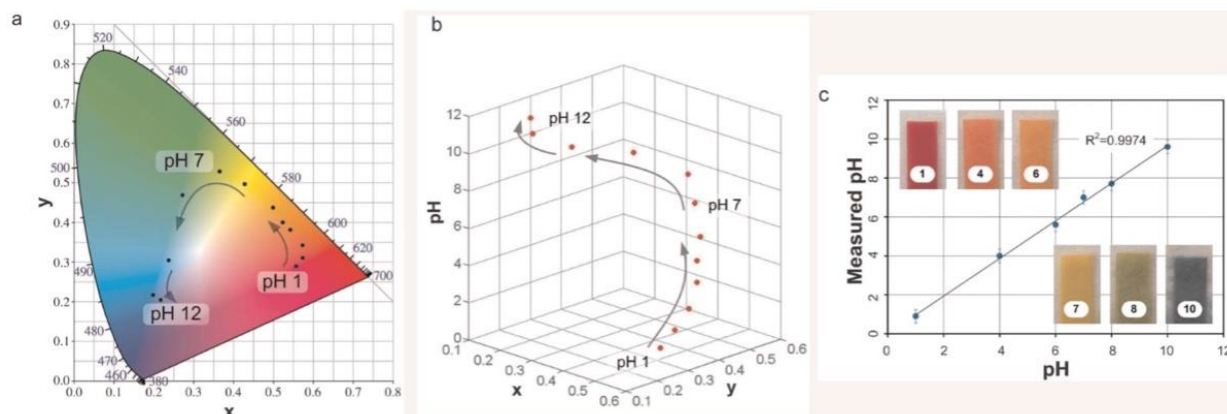
Figure 2.4. (a)(b)The relation of 2D CIE colour space and the 3D view with xy values with pH colour and pH values. (c) Best fit curve of the smartphone colourimetric pH VS pH buffer solutions.[31] *Figure reproduced with permission of the rights holder, Royal Society of Chemistry*

This group also developed a compensation equation for tests under various ambient light conditions, including light intensity, position of the light source, light temperature and outdoor lighting conditions. A reference chart was designed with 12 colour regions to compensate for ambient light conditions and to minimize the influence of Auto White Balance smartphone camera functions. They also found that by using their colour mapping approach, the calibration equation compensated for errors caused by both indoor and outdoor ambient lighting environment changes. The accuracy of this approach was higher than with the direct RGB measurement or ratios of red and green channels in pH concentration measurements because CIE 1931 has lower sensitivity to changes in brightness.[31,55,59] One way to improve upon these studies is to provide a reference colour chart directly on the test paper.

Cantrell et al.[59] first introduced the idea of a qualitative and quantitative colour property. Quantitative analysis is based on the absolute, relative, or reference measurements of intensity. For instance, both RGB and HSV are quantitative analytical parameters. The determination of the amounts of analytes require such quantitative parameters. Their study focused on the hue value in HSV colour space as a quantitative analysis method for bitonal optical sensors because they discovered that by quantifying H values using hue component only, the measurement of the concentration of the analyte achieved the best precision. Additionally, another advantage of this approach was direct calculation, better stability, insensitivity with respect to variations in illumination and indicator concentration, and the use of non-specialized, commercial cameras.

Murdock et al.[26] measured the average RGB intensity value for each tested spot and every intensity was converted into the corresponding CIE 1931 XYZ colour space to determine the image analysis method providing the best correlation. A Delta RGB (ΔRGB) value was calculated to quantify the difference between the RGB value of controlled spot and the RGB value of tested spot. R.C. Murdock reported an interesting finding that the analyte concentration correlated more consistently with the change in RGB colour space values since his paper-based enzyme-linked immunosorbent assays (P-ELISA) relies primarily on the intensity change of a single colour.

Sumriddetchkajorn et al.[57] proposed a mobile device-based self-referencing colourimeter for estimating the chlorine concentration in water. They specifically emphasized the importance of being able to fit both the reference scene and the tested sample in the same field of view of the camera. They analyzed the colours of the sample and the reference scene under the same illumination as a colour ratio (CR). They also investigated the performance of their cellphone colourimeter under various illuminating conditions, both indoor and outdoor. The experimental results were analyzed, and a fifth-order polynomial equation was developed to compensate for measurement errors. The effect of typical indoor illumination level on the cellphone-based colourimeter was acceptable with a less than 7% measurement error, whereas the outdoor results showed higher measurement errors, ranging from 23% to 33%, depending on the original chlorine concentration.

In this work, I chose to use RGB colour space because CMOS systems, most commonly used by smartphone cameras, sense red, green and blue colours and these have specific identifiable wavelengths, around 630, 550, and 450nm respectively. In the RGB colour model, each colour can be represented by assigning one numerical value in each of the red, green and blue additive primaries. Throughout this thesis, I call them red, green and blue channels. In each channel, the range of values is 8-bit, that is from 0 to 255. For example, (0, 0, 0) represents pure black and (255, 255, 255) represents pure white. The RGB colour space is the most commonly used colour space in colourimetric analysis. As an alternative, I also tried the HSV colour space but in general found more variability between phone manufacturers in that mode.

The previous work reviewed in this chapter provided the foundation for the work in the following two chapters. First the development of quantitative smartphone based imaging of our

paper-based assay applied to a simple yeast cell test (Chapter 3) and second the further development, testing and debugging of a colourimetric male fertility diagnostic.

Chapter 3

3   Reliable Smartphone-based Imaging and Colour Quantification Method

The content of this chapter constitutes a manuscript in preparation. The applicant is the lead and primary author for this work and played the primary role in experimental design, data collection and analysis, and write-up. The advisory and editorial contributions of Dr. Jae Bem You, Dr. Biao Zhang, Dr. Brian Nguyen, Farhang Tarlan and Professor David Sinton are also gratefully recognized.

3.1   Introduction

According to International Data Corporation (IDC), the worldwide smartphone market grew 13.0% in the year 2015. The explosive growth in smartphones usage has been a distinct trend in contemporary society. With more than 6 billion subscriptions worldwide, cellphones are ubiquitous.[31] Due to this growth, fine camera peripherals, and better mobile computing and central processing power are now more portable, easily-operated, and accessible with lower cost.[56] Owing to these technological advances, smartphones are becoming a good alternative for optical sensing[42,52,53] and imaging analysis[62], greatly enabling the application of point-of-care health diagnostics,[24,54,55] water quality monitoring,[57,63] chemical analysis,[64] and food nutrition tests.[58,65] The concurrent growth of paper-based diagnostics, detailed in the previous chapter, has enabled the powerful combination of cheap diagnostic testing with smartphone based readout. The convergence of convenient mobile technology with paper-based chips is particularly powerful in the context of paper-based assays with colourimetric detection.

However, current image quantification methods via smartphone have drawbacks that need to be overcome. First, many colour quantification methods have been proposed using the HSV colour space,[18,59] the CIE1932 colour system,[61] and RGB colour space.[26,64] (discussed earlier). Additionally, simple calculation of RGB values including grayscale intensity, red to blue ratio[63] and colour ratio[57], hierarchical cluster analysis[19,66] and the length of coffee-ring effect[67] have been leveraged as methods of colour quantification. However, smartphone cameras use CMOS imaging sensors that only detect the wavelengths of three primary colours: red, green and blue (RGB). CMOS detectors are inherently discriminating with respect to these three colours and the spectral

change observed by CMOS arrays results from the change in relative ratios of RGB colours, rather than the fundamental change in wavelength.[18,68] Furthermore, both HSV colour space and the CIE1932 colour system are conversions of RGB values. Therefore, working directly with RGB values is expected to be best when applying smartphones to colourimetric array readers.

Second, the variety of common lighting sources imposes great uncertainty and inconsistency in image digitization and colourimetric readings when imaging with a smartphone camera. Accessories have been developed for the purpose of eliminating the influence of changing lighting conditions.[30,42,55,58,65,66] However, these added components have obvious disadvantages in terms of cost, reusability and, in some cases, hygiene. Most importantly, these accessories must match the design and camera position of different types of phones, causing challenges for end users and/or requiring model-by-model matching of the diagnostic unit design to the phone (increasing costs for suppliers). Recently, some studies have been done on calibration methods to eliminate the effect of various lighting conditions and to distinguish among colours accurately, but have not been applied and tested in single-colour shift (e.g. from clear or white to blue, or from light red to dark red) quantification as of yet. For single-colour shift, relating RGB values directly to analytical values (without calibration) has been shown to perform better than converting each value into the corresponding HSV values or CIE1931 colour space coordinates.[26,57] This approach, however, is not expected to be as accurate as when both calibration and single-colour quantification are integrated.

Third, smartphone cameras often contain integrated automatic colour balancing functions intended to optimize the quality of photos produced. These functions act directly on the imaging data, and can introduce significant error, impairing quantitative measurements. All of these aspects are further complicated by the wide variety of phone brands and models worldwide, and the literature comparing various cellphones is sparse. To achieve the vision of widespread, low-cost paper-based diagnostics paired with ubiquitous smartphone technology, the field needs to overcome these limitations. Image acquisition by smartphones will need to be largely independent of environmental lighting conditions, independent of phone brand/model and applicable to single-colour shifts common in colourimetric assays.

In this chapter, I present a simple and reliable colour quantification method with the help of continuous smartphone LED light as a dominant light source and a distinctive combination of rescaling and colour quantification for a single-colour shift applications. The method was tested with three different phone models from the most representative smartphone brands: iPhone, Samsung and Huawei. The approach shown here (i) compensates for different lighting environments, (ii) minimizes the discrepancies in image quality from various phones, and (ii) provides uniform light to avoid any shadows during image acquisition. To demonstrate proof-of-concept the approach was applied to quantify yeast cell concentration in a paper-based device similar to that employed by our group previously for sperm testing. Sperm based testing is performed in the next chapter.

3.2   Fabrication of Three-trial Colourimetric Chip

3.2.1 Reagents and Chemicals

Malachite green solution and Rhodamine B solution were purchased from Sigma Aldrich as colourimetric dyes. Both the green dye and the red dye were diluted with distilled water into various concentrations to mimic lighter and darker colours. Phosphate buffered saline (PBS; pH 7.4, 1X) from Gibco® by Life Technologies was used for all experiments related to yeast. The 5 g/L 3-(4,5-dimethyl-2-thiazolyl)-2,5-diphenyl-tetrazolium bromide (MTT; Sigma-Aldrich) colourimetric reagent solution was prepared by dissolving MTT powders in PBS solution. Trypan blue stain 0.4% was purchased from Gibco® by Life Technologies as a dye for live-dead assay. Hydrogen peroxide 30% was purchased from Bio Basic Canada Inc.

3.2.2 Fabrication of Three-trial Colourimetric Chip

The device consists of 2 paper layers patterned with wax on top. The front layer is printed with both black reference spots boundaries and testing spots boundaries, while the back layer, a uniformly white reference scene, is printed with testing spots boundaries only. The back layer also prevents liquid penetration through the front paper layer. The device patterns were designed in AutoCAD, then printed on No. One Chromatography Paper from Sigma-Aldrich with a solid wax printer (ColourQube 8570N, Xerox Canada). Printed sheets of paper were placed in a preheated oven at 120°C for 3 minutes to allow the wax to fully melt and fill the pores through the thickness

of the paper. The size of the testing spots and background spot was 4-mm-diameter after heating. 1μL of black ink was added and dried to the designated 1-mm-diameter dots to prepare for black reference spots. Double-sided tape was used to bond 2 layers such that the testing spots boundaries aligned with each other. Lastly, the prepared dyes were pipetted and dried to the testing spots, T1, T2 and T3. Besides the green colour chips and the red colour chips, the paper-based sensor chips were also prepared for various concentrations of the green dye and the red dye.

3.2.3 Adoption of Dominant Light Source

The representative smartphones chosen were iPhone 7 Plus, Samsung Galaxy S7 edge and Huawei P8Lite ALE-L04. Photographs were obtained in four different lighting environments, including 50, 170, 515 and 1000 lux. These lighting conditions were measured with digital LUX meter (HP-881A, HoldPeak®). To acquire the images, the smartphone LED light was used as the dominant light source. To ensure the LED light is continuously on, camera was switched to video mode before imaging, such that the light could be used (Fig. 3.1a-b). This approach was effective in all three major manufacturer models tested. Images were acquired by taking snapshots while recording a video. During the imaging process the chip was placed at the left center of the imaging area since it is the brightest region and is intuitive for the user (Fig. 3.1b).

3.2.4 Colour Calibration and Representation

The snapshots captured by all three phones was opened using ImageJ for manual processing. The transition from the testing spots colour to wax boundary layer was excluded from the average RGB value measurement by selecting only 80% of the testing spot area. These measurements were then transferred into Excel for further colour calibration and quantification.

The digital RGB values of the coloured spots were detected by the smartphone camera and measured in ImageJ. Each testing spot was corrected individually by taking the black spots and white substrate around each spot as backgrounds. The black and white backgrounds were utilized as references to compensate the difference caused by uneven light distribution and light condition changes. After colour correction, two parameters were used to quantify the difference in colour change, (i) colour intensity difference between background(B) and testing spots $(I_B - I_T)$, (ii)

Delta RGB (ΔRGB)[26] the magnitudes of the vector in RGB colour space from background colour to testing colour.

3.2.5 Statistics

Statistical analysis was completed using the R studio. A test for significant differences in the rescaled ΔRGB values of testing spots imaged with continuous smartphone LED light search under different lighting conditions was completed using the one-way analysis of variance (ANOVA) tests (n=12, $\alpha < 0.05$). Statistically significant ANOVA results were then tested with Tukey HSD tests for pairwise comparisons.

3.2.6 Yeast Cell Sample Preparation and Detection Standards

White Labs WLP001 California Ale Yeast cells, from Toronto Brewing Homebrew Supplies website, were taken out of fridge, separated into two falcon tubes and warmed to room temperature for one hour. Then one tube was used for the preparation of a sample with lower live yeast cell concentration. This was achieved by mixing hydrogen peroxide with the yeast cells for another 30 minutes, after which the cells were washed and resuspended in PBS. These two yeast samples were mixed at various ratios to have various live cell concentrations. Aqueous solution of MTT dye (concentration = 5 g/L) was added to different yeast samples at volumetric ratio of 1:5. After 20 minutes, 5 µL of the mixture of yeast cells and colourimetric MTT dye was pipetted and dried to each testing spot.

To count the number of cells,10 µL yeast cell samples were first diluted to 2 mL of PBS (1:200 volumetric dilution ratio). Then 20 µL of diluted yeast sample was transferred to a centrifuge tube and mixed with 20 µL of trypan blue solution. 10 µL of dyed yeast cells were then added to each chamber of counting slides (Bio-Rad). Lastly, the counting slides were inserted into the automated cell counter (Bio-Rad T20) and readings were recorded. The actual concentration of the yeast cells equals the reading directly from the cell counter times 200.

3.3 Results and Discussion

3.3.1 Colour Rescaling and Quantification

Given the demands of paper-based diagnostic testing applications, smartphone-based measurement should be insensitive to surrounding light conditions, accessory-free, and easy-to-use so that even untrained end-users can readily learn the operating steps quickly. The first approach here is to combine an efficient colour calibration method and a suitable colour quantification method, to measure, correct, and convert sensed colours to analytical values.
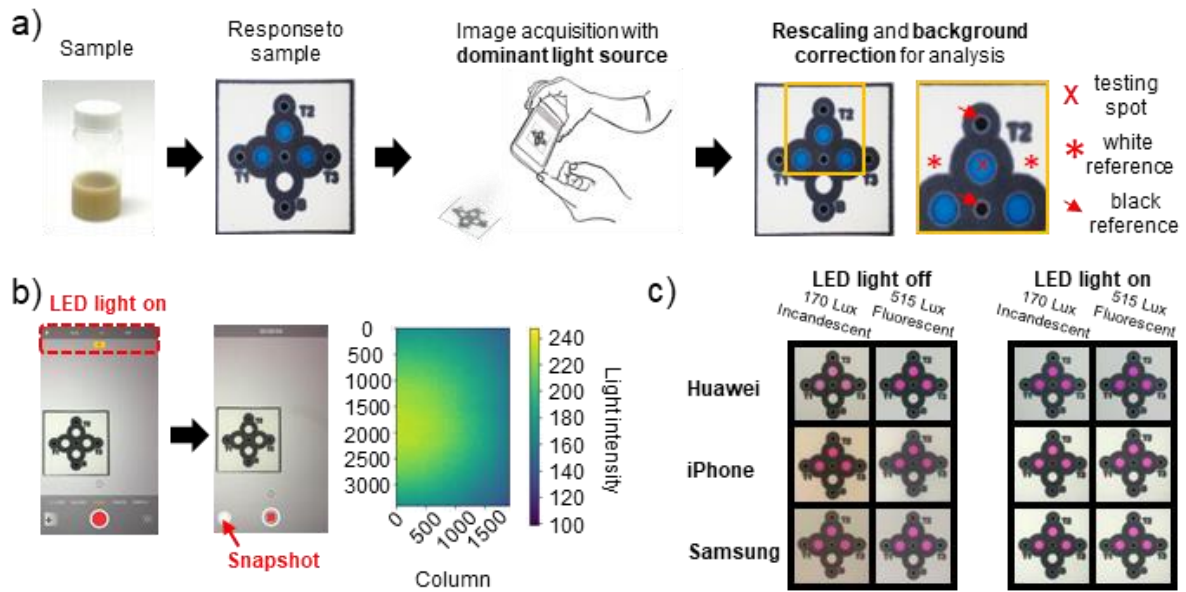
Figure 3.1. (a) Schematic for the determination of the concentration of a sample using a smartphone-based colourimetric chip. The image acquisition and rescaling method are illustrated. (b) includes the screenshots of the "LED light on" setting and an indication of the position to take snapshots under video mode with iPhone 7 Plus. Right: A light intensity distribution map of the camera field when LED light of iPhone 7 Plus is continuously on.

The correction process uses RGB values of the black and white backgrounds as references because black and white are two extremes in RGB colour space. The colour of the selected background spot or testing spot ($R_{meas}, G_{meas}, B_{meas}$) was detected by the smartphone and digitalized to the range of pixel values (0, 0, 0) to (255, 255, 255), assuming 8-bit RGB colour space. The black and white reference values were measured at the black spots and white background close to the measured spot, shown in Figure 3.1a. Taking the red channel as an

example, the rescaling of digital RGB values of the testing spots on the chip was demonstrated by equation (1):

$$R_{corr} = \left(\frac{256}{R_w - R_B}\right)(R_{meas} - R_B) \tag{1}$$

where R_{corr} is the corrected value after rescaling, R_w is the white reference value, R_b is the black reference value and R_{meas} is the measured value from the smartphone pictures.

After the colour correction, Delta RGB (ΔRGB)[26] calculation, shown below as equation (2) was used for colour quantification. The value for ΔRGB can be understood as the magnitude of a vector in a 3D colour space, from original colour with (R_0, G_0, B_0) to testing colour with (R_t, G_t, B_t).[26] Since the maximum possible range for each R, G and B channel is 0 to 255, the maximum possible range for ΔRGB can be easily calculated as 0 to 442.

$$\Delta\text{RGB} = \sqrt{(R_t - R_0)^2 + (G_t - G_0)^2 + (B_t - B_0)^2} \tag{2}$$

$$\text{Intensity Difference} = I_t - I_b = \frac{R_t + G_t + B_t}{3} - \frac{R_b + G_b + B_b}{3} \tag{3}$$

In this work, I chose to use RGB colour space for two reasons: (i) When a smartphone is utilized as an image sensor, only red, green, and blue light reflected from the paper-based colourimetric devices can be detected. Thus, calculating the right values and ratios for these three colours is crucial before any further quantification and conversion and this is also the reason why colour calibration was done prior to quantification. (ii) Since our assays are based a single-colour shift (like many colourimetric assays), relating corrected RGB values of digitalized images directly to analytical values was a better fit than converting each value into the corresponding HSV values or CIE1931 colour space coordinates.

To demonstrate the effect of this colour correction and colour evaluation method on various ambient light conditions and smartphone brands, the paper-based chip was first imaged under different illuminance conditions of four lighting intensities: 50, 170, 515 and 1000 Lux and two types of light sources: incandescent (170 Lux) and fluorescent (50, 515, and 1000 Lux). The three most popular phone brands on the market, namely iPhone (7 Plus), Samsung (Galaxy S7 edge) and

Huawei (ALE-L04), were tested with all illuminance conditions. I compared the colour intensity difference between the testing spots and the background spot $(I_t - I_b)$ and ΔRGB values before and after rescaling. These testing spots represented three trials for same colour. The values of all conditions mentioned above were normalized to their corresponding values at 50 Lux (Fig. 3.2a-b). The ΔRGB value narrows the differences significantly for Huawei red and slightly for Huawei green. This same trend was also observed with Samsung red, Samsung green and iPhone green. The detailed results of Samsung and iPhone are shown in Appendix A. The rescaling method provides an additional compensation for the colour variation caused by changing lighting conditions with both intensity difference and ΔRGB. This correction effect is significant for both red and green colours and all three types of phones. For all Huawei green, iPhone red and green and Samsung red and green, rescaling the ΔRGB values correct the colours most effectively. As mentioned above, ΔRGB has a larger range than intensity difference, and thus a similar absolute variation results in less relative error. The result is that the ΔRGB value with rescaling corrects for light intensity most effectively in this green spot case. For Huawei with a red spot case (Fig. 3.2b), the standard deviation of ΔRGB and rescale measurements were greater than intensity difference and rescale; however, the percentage value of standard deviation divided by the average of measurements for ΔRGB and rescale is still smaller than that of intensity and rescale. Across all three brands and both colours, ΔRGB with rescale corrects best (Fig. 3.2c-d). This performance is attributed to ΔRGB including contributions from each R, G, and B channel separately. In contrast, in the calculation of intensity difference, values for different channels are averaged for each colour before comparison, thus blurring the contributions of different channels and reducing the inherent resolution of the measurement.
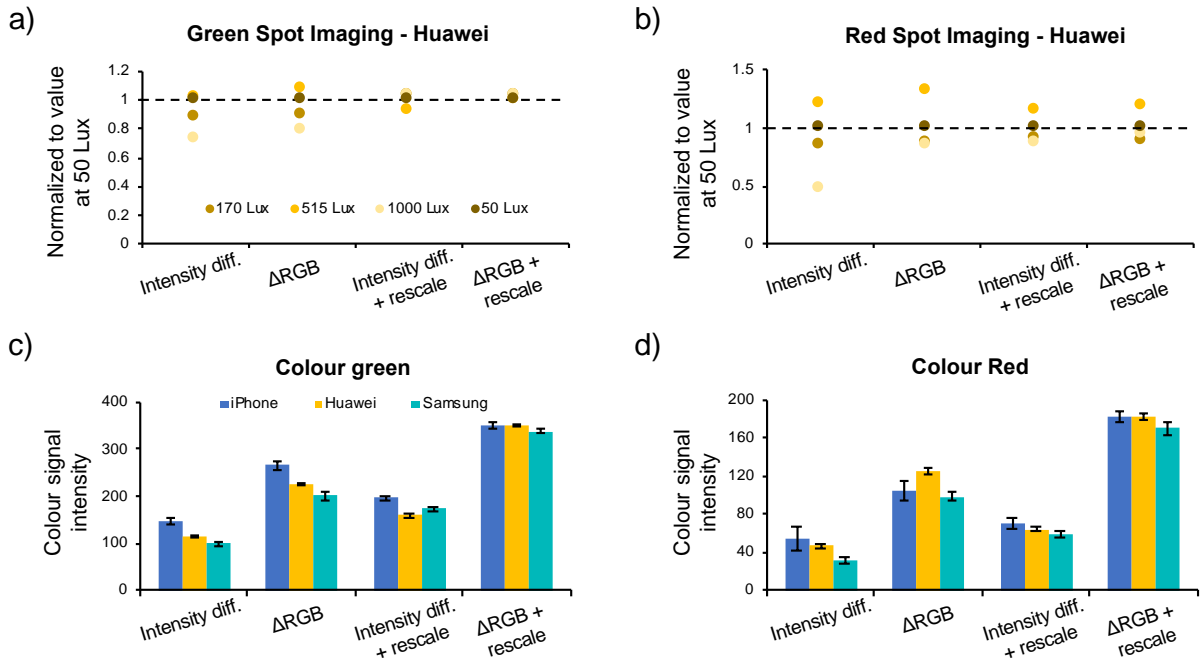
Figure 3.2. Comparison of two RGB colour quantification methods (**ΔRGB** and intensity difference) before and after rescaling, without dominant light source. Images were obtained under various ambient lighting conditions. (a-b) The RGB values of 170 lux, 515 lux, 1000 lux were normalized to the value at 50 lux in their own category. (a) Green colour chip images imaged with Huawei P8Lite ALE-L04. (b) Red colour chip images imaged with Huawei P8Lite ALE-L04. (c-d) The average value of colour signal intensity were compared among three representative phones in each colour quantification category. (c) Green (d) Red

3.3.2 The Impact of Lighting Conditions

The colour response testing as a function of light intensity (Fig. 3.2) was conducted under highly controlled conditions. In reality, paper-based diagnostics will be used in a wide variety of non-ideal lighting conditions. Preliminary testing with the three common smartphone makes across an array of typical lighting conditions (interior room, hallway, elevator, outdoor bright, outdoor dark). These tests showed wide variability. I hypothesized that one way to reduce the influence of external lighting conditions on the imaging would be to dominate the lighting conditions using the LED light integrated in each smartphone – used for LED light photography as well as LED light apps etc. Simply using the LED light did not, however, provide consistent results, because of variabilities in the camera settings under LED light conditions. In addition, it was difficult for users to align the LED light-illuminated region and the area to be imaged at close range, and alignment was critical. Having the light on continuously for the user, and the camera, was

preferable. It was possible to achieve this imaging condition in all three manufacturers by switching to video mode and adding the LED light (Fig. 3.1b left). The built-in LED light provided a dominant light source, reducing the influence of the surroundings (bright or dark) and allowed alignment of the device. Pictures of the devices could then be taken separately, while in video mode, in all three phones with continuous LED illumination (Fig. 3.1b middle). The light intensity heat map in Fig. 3.1b illustrates the light intensity distribution on a picture taken by smartphones and that the center of the halo created by all three representative phones are at the left center region (in keeping with the position of the LED on the back of the device). To maximize the efficiency of the dominant light source and minimize the light intensity gradient on the testing paper, the chip was placed at the brightest region in the field of view. This specific requirement could become a source of error which can be rectified by using an instructional app to visually guide the end user to place the chip at the right spot. Since the smartphones were placed closer to the paper chip and smartphone LED light is a point light source, the light distribution in the picture is less uniform, in general, compared to typical lighting conditions. To compensate for expected non-uniformity a two-spot correction method[19] (described in the previous chapter), was employed for calibration using two very proximate reference spots as shown in Figure 3.1a right. While in this work I used the video recording mode to provide the dominant light source continuously, in ultimate application it is likely that the user would be using an instructional app for the paper-based diagnostic, that would also control the LED and imaging hardware automatically. Such an app could also provide instructions for the user in preparation of the assay and a visual alignment/sizing guide when imaging. Our group had the basics of such an app developed, however, the imaging variability was the main issue, and the motivation for the work in this chapter. The key finding here is that with this use of the integrated LED and appropriate calibration/rescaling, variations in lighting levels and conditions can be readily removed for more quantitative imaging and ultimately diagnosis.

As shown in Figure 3.1c, without the adoption of continuous LED light, images captured under various light conditions using iPhone 7 Plus, Samsung Galaxy S7 edge and Huawei ALE-L04 present large discrepancies. After I applied the dominant light source, the variance caused by changing illuminance environments are significantly minimized, especially in the photos obtained by iPhone (7 Plus) and Samsung (Galaxy S7 edge).

Here the benefit of the continuous lighting strategy was quantified. First, the pictures were taken under the illuminance of different environments with both LED light on and LED light off, and colourimetric results were compared afterwards (Fig. 3.3a-b). ANOVA and Turkey HSD comparison test were used as the statistical analysis method. $P < 0.05$ was considered as a statistically significant difference. After the adoption of LED light on top of calibration, the ΔRGB values in various incident light illuminance failed to reject the null hypothesis for the ANOVA test with alpha level of 0.05 for all tests, since its p value was greater than 0.05. Detailed ANOVA results for various phones with dominant light source are listed in Table 3-1 below. The null hypothesis for the test implies that the means for four lighting conditions are equal with 12 replicates used for each lighting condition. The conclusion for ANOVA tests for both green and red colour indicate that there is no significant difference among the corrected ΔRGB values in these different lighting conditions – an affirming result. The ANOVA test results for experiments with LED off showed otherwise. The p value from ANOVA test for samples with no LED were smaller than 0.05, meaning there was indeed significant differences introduced by the lighting conditions. Detailed ANOVA results for various phones without dominant light source are listed in Table 3-2 below. For all three phones the ANOVA test showed the same results as mentioned above. The standard deviation for each lighting condition also decreased after continuous LED light was used (Fig. 3.3a-b). Thus, within each cellphone brand, the use of continuous LED light as a dominant light source largely compensated for the different intensity of the ambient light illuminating environments.

Table 3-1: ANOVA Summary Table for Various Phone Models with Dominant Light Source

Colour	Phone Model	P value
Red	iPhone 7 Plus	0.3750
	Huawei ALE-L04	0.9135
	Samsung Galaxy S7 edge	0.0975
Green	iPhone 7 Plus	0.5268
	Huawei ALE-L04	0.5877
	Samsung Galaxy S7 edge	0.3856

Table 3-2: ANOVA Summary Table for Various Phone Models without Dominant Light Source

Colour	Phone Model	P value
Red	iPhone 7 Plus	0.0011
	Huawei ALE-L04	3.289e-13
	Samsung Galaxy S7 edge	1.996e-05
Green	iPhone 7 Plus	0.0073
	Huawei ALE-L04	7.854e-05
	Samsung Galaxy S7 edge	5.892e-09

To further investigate the statistical difference between each lighting conditions, I performed Tukey HSD tests after ANOVA with and without LED light for all three smartphones. Significant differences were observed between some of the lighting conditions when LED light was not adopted. The iPhone camera works better in brighter lighting conditions, whereas, it is not as stable in dimmer lighting conditions without the use of its LED light. For both green and red colour, there is a significant difference between 50 Lux and the rest of lighting conditions. Huawei and Samsung smartphone camera show better performance in the dark light environment, but their sensors have slightly weaker sensitivity in strong ambient light illuminating conditions. For instance, when the dominant LED light source off, Tukey HSD results show significant statistical difference between 1000 Lux and the other illuminance intensities for both Huawei green and Samsung red. For Huawei red, there is a significant difference between 515 Lux and the other lighting conditions, as well as between 50 Lux and 170 Lux. When the LED is used, though, I do not find significant differences (p value ≥ 0.05) between each lighting conditions for either Huawei, Samsung or iPhone. The dominant LED light at the back of the smartphone is expected to compensate for the ambient illuminance intensities up to at least 1000 lux since the light intensity of all three LED lights, measured at a distance of 10 cm in a dark environment by the light meter are greater than 1450 lux. The intensity of the LED light outcompetes ambient lighting environments typical of interior spaces (150 ~ 550 lux) or overcast outdoor spaces (~ 1000 lux), minimizing the influence of lighting environment.

To evaluate the improvement in accuracy of this imaging method among various phone brands, the colourimetric results of Huawei, iPhone and Samsung, with the LED light on and the LED light off, were measured and compared in Figure 3.3c. All of the results were rescaled and represented by ΔRGB. For both green and red, the trendlines of three phones are brand-specific

without LED light, the trendline of the Huawei phone is particularly out. After applying the dominant light from smartphones, the range of rescaled ΔRGB values from all three phone models reduces from 249.94 to 22.96 for colour green and from 93.07 to 26.57 for colour red, and the trendlines show greater overlap, demonstrating that the colourimetric results are more consistent in diverse ambient lighting conditions and among different phone brands. Furthermore, the rescaled ΔRGB values from all phones are more clustered with the LED light, meaning that LED light also reduces the standard deviation among different trials within each lighting intensity, for instance, in 1000 Lux lighting condition, the standard deviation of Huawei colour green decreases from 21.49 to 8.63. The adoption of the LED light also reduces the size of the error bars from multiple tests, especially for 1000 lux iPhone green and 1000 lux Samsung red. The corresponding detailed results of Samsung and iPhone are shown in Appendix B. As the surrounding light intensity increases, it becomes harder to avoid shadows casted on the paper-based chip due to the placement of cellphone. Figure 3.3c reveals that ΔRGB under 1000 Lux without the LED light are distributed in a wider error range, resulting in a slanted trendline for Huawei Green. In addition to shadows that occur naturally in the environment, the phone itself is a major, and largely unavoidable, source of shadows in images taken without the LED light. Lastly, I tested the integration of imaging and calibration methods and the three-spot paper-based chip with diluted green dye (Fig. 3.3d). The ΔRGB values were plotted against the concentration of green dye, obtaining a logarithmic best-fit equation since the colour changes become smaller when the colour reaches saturation. Zooming into the low concentration part of the diagram, the direct relationship between ΔRGB values and the dye concentration is linear in the range where ΔRGB is greater than 50 and smaller than 300. With the LED, ΔRGB and colour calibration methods, a set of new calibration curves can be made for various applications. Then, the concentrations of each analyte can be described as a function of a simple variable, ΔRGB of the corrected colours. These calibrated curves can be loaded onto the smartphone ready to be used for various applications. In the next section this approach is applied to a paper-based viability test for yeast cell samples.
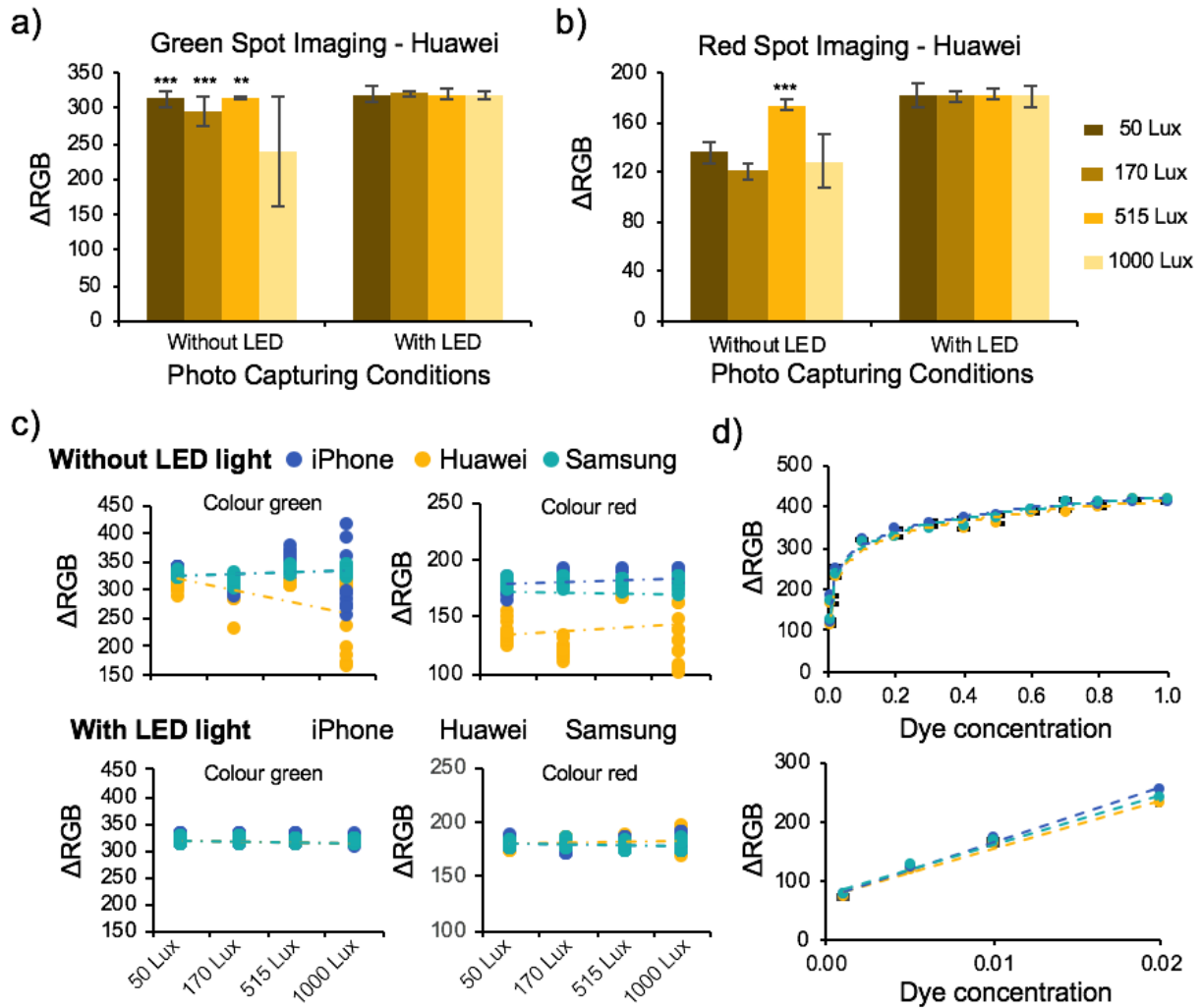
Figure 3.3. (a-b) The comparison between two imaging conditions, with LED light and without LED light. ANOVA and Turkey HSD comparison test were used as statistical analysis method. P < 0.05 was considered as a statistically significant difference. **p<0.01, ***p<0.001 (a) Green colour chip images obtained with Huawei P8Lite ALE-L04 under various ambient lighting conditions. (b) Red colour chip images obtained with Huawei P8Lite ALE-L04 under various ambient lighting conditions. (c)The effect of using the dominant light source on calibration of various phone models. (d) The Δ**RGB** values agree with a single-colour shift of green colour dye resulted by various dye concentrations. Dye concentration was varied through serial dilution.

3.3.3 Yeast Cells Concentration Analysis

While the ultimate goal of the imaging method developed in this chapter is application to male fertility diagnostics, a simple yeast based cell viability test was chosen here to prove the imaging

approach, separate from other variables associated with sperm testing (to be discussed in Chapter 4).

While abundant studies have been done on distinguishing among various colours for commercially-available paper microfluidic diagnostic devices such as pH paper test and urine test strips,[18,19] the application of yeast cell viability analysis and concentration monitoring has yet to be explored. Yeast cells are most commonly found in the process of alcoholic beverage making (i.e. beer brewing and wine making) and bread baking.[69,70] In recent years, yeast has also been considered for other applications such as in producing biofuels, providing a renewable source of bioethanol and butanol, as well as in making feedstock chemicals, such as acids and resveratrol, at industrial scale.[71–74] Monitoring yeast cells in all these applications currently requires a microscope and a technician. Feizi, et. al. developed a rapid, portable and cost-effective yeast cell viability and concentration analysis method, but it also involves a lens-free microscope.[75]

The approach here was a paper-based colourimetric test for yeast live cell concentration and viability quantification. The viability of the yeast cells was varied by mixing different ratios of hydrogen peroxide-killed yeast with the original yeast solution. Colourimetric reagent MTT was used for the cell viability assay since it reacts with live cells and forms purple formazan. After 20 minutes, the purple solution was transferred to the testing spots on the paper device (Fig. 3.4a) and this colourimetric change was captured using the smartphone's camera with dominant LED light and colour correction. The accuracy of the paper-based quantification method was characterized by directly relating the rescaled ΔRGB values to the live cell concentrations and viability obtained from the automated cell counter and linear regression analysis was performed (Fig. 3.4b and c). The calibration curves for Samsung and iPhone are shown in Appendix C. For all three phone models, both live cell concentration and viability show very linear correlation with the ΔRGB values, with $R^2 > 0.95$. This linear correlation is expected based on previous testing because the ΔRGB values for the concentrations tested were below 300, corresponding to the linear region in Figure 3.3d. For the control testing in the automated cell counter, all tested samples were pre-diluted 1:200 for compatibility with the counter, and results were converted back for comparison with the paper-based test. The test concentrations here (7.740×10^7 ~ 9.880×10^8 live cells/mL), reflect typical concentrations used in wine fermentation and beer brewing industries. Even higher concentrations are possible in biofuel production.[74] The calibration curves for three phone brands

agree well (Fig. 3.4d), indicating that the combination of LED-assisted imaging and rescaling for single-colour shift quantification was effective in this application, and could be applied directly to any of the three smartphone manufacturers without issue.
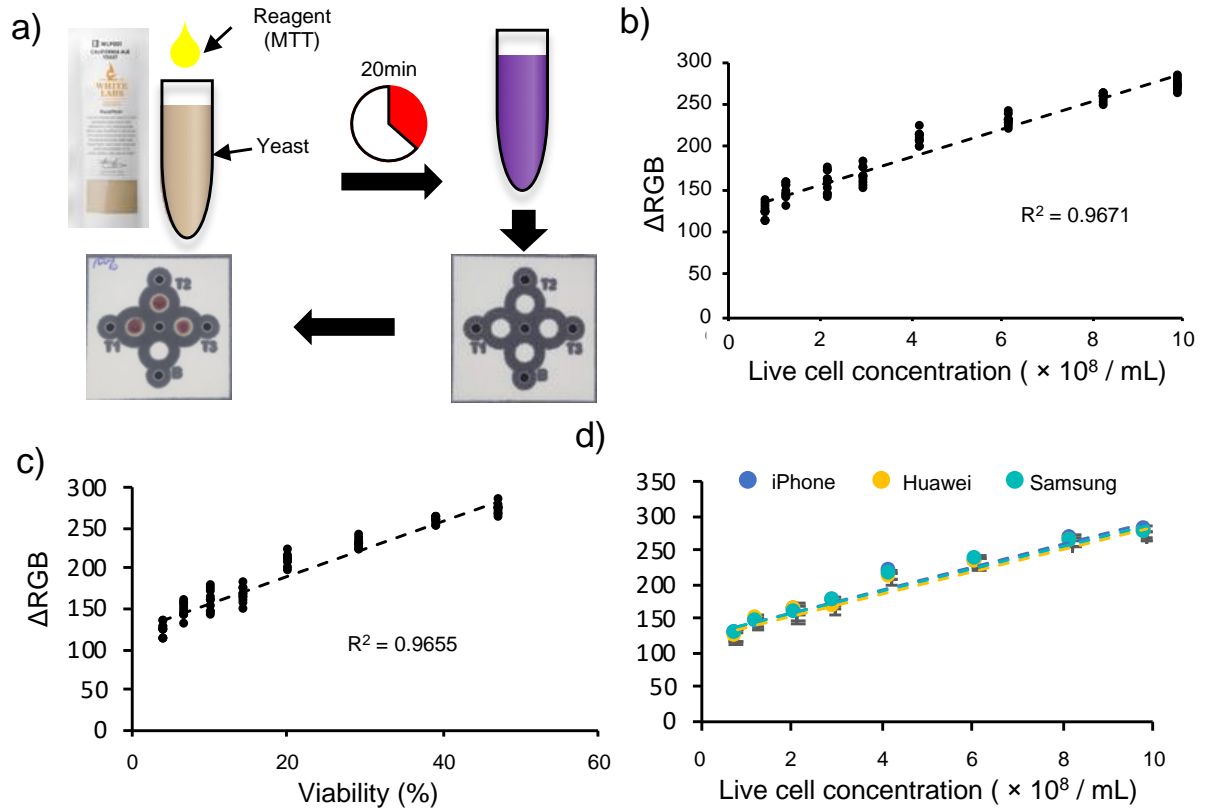
Figure 3.4. (a) Schematic for the determination of yeast live cell concentration and viability using 3 trial paper-based chips. (b-c) Huawei P8Lite ALE-L04 calibration curves for colour change on paper-based chips (Δ**RGB**) with respect to live cell concentration and viability measured by automated cell counter counting. (d) The live cell concentration calibration curve for iPhone (dotted orange line), the calibration curve for Huawei (dotted blue line) and the calibration curve for Samsung (dotted green line) agrees with one another.

3.4   Conclusion

In this chapter, a fast and reliable imaging, rescaling and colour quantification method was presented for a single-colour shift analysis using smartphone imaging. This process involves a using the continuous LED light built-in to standard phones as the dominant light source, in combination with rescaled ΔRGB. I demonstrate that the integrated method of calibration, ΔRGB and smartphone LED light can compensate for ambient lighting environments and mitigate

variances among three phone manufacturers iPhone (7 Plus), Samsung (Galaxy S7 edge) and Huawei (P8Lite ALE-L04), for single-colour quantification. The ANOVA analysis shows no significant difference (p value >0.05) between measurements at different lighting conditions with the adoption of the above method.

Our integrated smartphone and paper-based yeast viability detection technology is an attractive alternative to conventional laboratory and/or industrial cell counting. Moreover, I have demonstrated that the accuracy of the quantified results obtained using our method are comparable with those of a commercial yeast kit. With the method established and proven here, a smartphone application can be developed to control the LED light in camera mode and an algorithm can be developed to digitalize, correct, and analyze the image to provide the end user with live yeast cell concentrations in a user-friendly manner.

One limitation of this imaging method is when the surrounding light noise equals or exceeds the LED light intensity, such as full daylight and direct sunlight with a light intensity up to 30,000 lux. In the context of both paper-based yeast sensing, and paper-based male fertility diagnostics, this high-daylight issue is unlikely to be an issue (all testing expected to be indoors). It would also be straightforward to implement a safety in this regard, that is, having the app warn the user if the overall light levels in the image are too high, or reject and image if the brightness is too varied across the image.

3.5   Supporting Results Section

3.5.1 Calibration of Imaging Angle

All iPhone 7 Plus, Samsung Galaxy S7 edge and Huawei P8Lite ALE-L04 were employed for the imaging of the red or green dye-contained sensor array at difference imaging angle. The LED light of the phone was continuously on during the imaging process. The phones were placed on top of the chip and with a 30° tilting angles to study the effect of imaging angle of smartphone camera on the results. The vertical imaging distance from the camera and the paper-based chip was controlled during the experiment. The raw colours of testing spots were corrected utilizing the two spots, black and white rescaling method and then represented by ΔRGB, for both flat and angled imaging.

The raw RGB values of the colour are significantly different and distinguishable before correction and utilization of LED light. After the dominant light source and the calibration are adopted, the ΔRGB values becomes similar no matter the smartphone is placed flat or angled. According to the results shown in Figure 3.5a, this integrated method can compensate for various tilting angles for both red and green and it should be applicable with other colours. As illustrated in Figure 3.5b, our integrated method's ability to compensate tilting angles also works for all three representative phones.
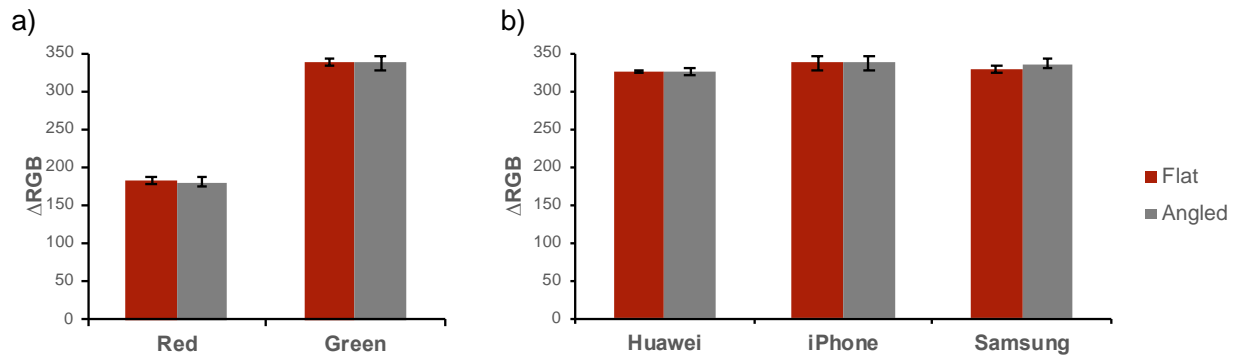
Figure 3.5. (a) **ΔRGB** representation of the corrected RGB values of red and green chips at different tilting angels. Images were obtained from iPhone 7 Plus. (b) **ΔRGB** representation of the corrected RGB values of green chips imaged by three representative smartphone at different tilting angels.

3.5.2 Imaging with Camera Mode vs. with Video Mode

Among the representative cellphones, only Huawei P8Lite ALE-L04 can have the LED light continuously on in both camera mode and video mode, so the following experiment was conducted with Huawei smartphone only. Colour information of images from both camera mode and video mode were extracted, corrected and represented in ΔRGB. The ΔRGB values measured with camera mode and video mode are within 96% similarity. Thus, the camera mode can also be used if an application can help control the LED light of the phones.
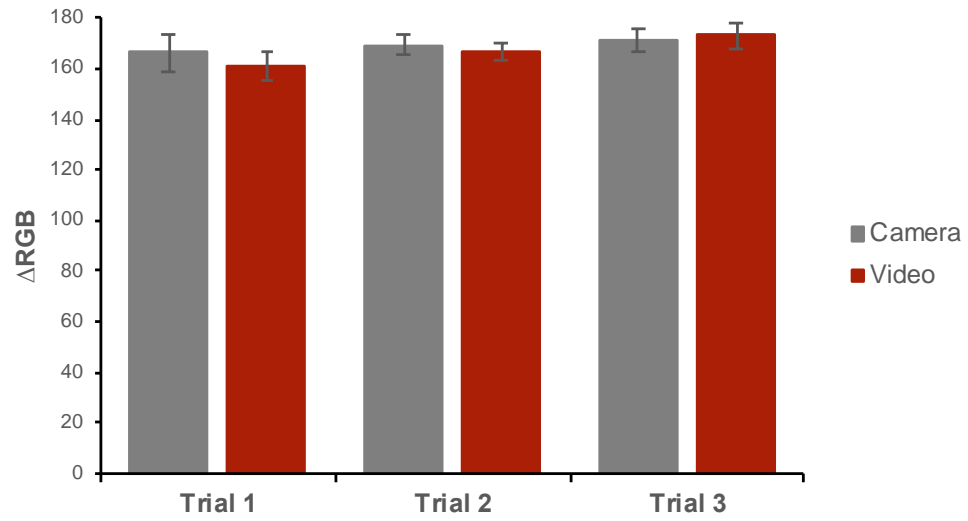
Figure 3.6. Comparison on **ΔRGB** representation of the corrected RGB values of red chips from Huawei camera mode and video mode, with 3 repetitions.

In the next chapter, I will apply this proven smartphone imaging and correction platform to male infertility to determine sperm concentration and motility.

Chapter 4

4 Paper-based Microfluidic Chip Development and Optimization for Male Infertility Diagnostics

4.1 Introduction

The paper-based male infertility diagnostic concept (Fig. 4-1) was developed by Dr. Reza Nosrati, a former member of the Sinton lab, using microfluidics and sperm behavior expertise[21,25]. The simple design and fabrication process required minimal reagents and equipment, with a total material cost of ~US$0.05 (5 cents) per device. The device operates using the enzymatic colourimetric MTT assay. Briefly, the yellow MTT tetrazolium converts to purple formazan upon removal of the bromide by the diaphorase flavoprotein enzyme present in metabolically active human sperm, and the motile sperm are distinguished based on their ability to swim through the pore size membrane prior to reaching the dye. This approach allows simultaneous measurement of three critical semen parameters in 10 min: live and motile sperm concentrations and sperm motility (derived from the other two). The majority of competing male fertility diagnostics provide only total sperm cell count[39–41], whereas motile count is generally seen as more predictive of male fertility potential.
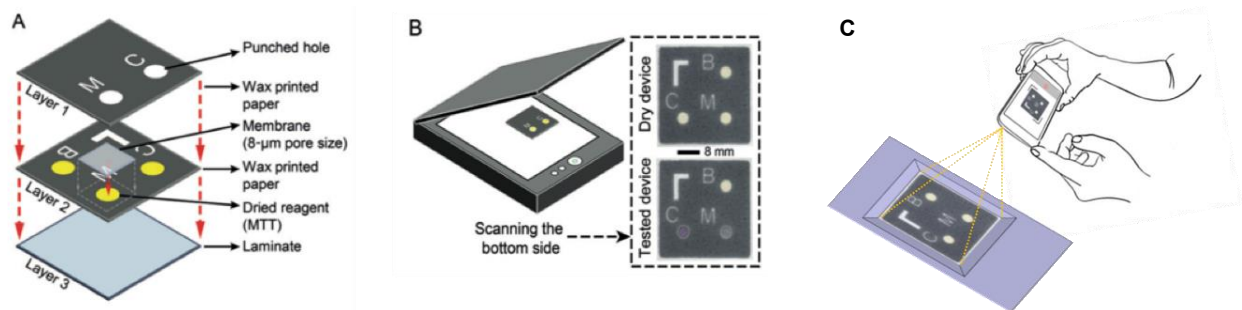
Figure 4.1. Paper-based device for mobile male fertility screening. a) Exploded view of the paper-based chip. b) Schematic view of the assembled device and scanned images of devices before and 10 min after applying a semen sample. c) Schematic of smartphone-based imaging.[25] *Figure reproduced with permission of the rights holder, Clinical Chemistry*
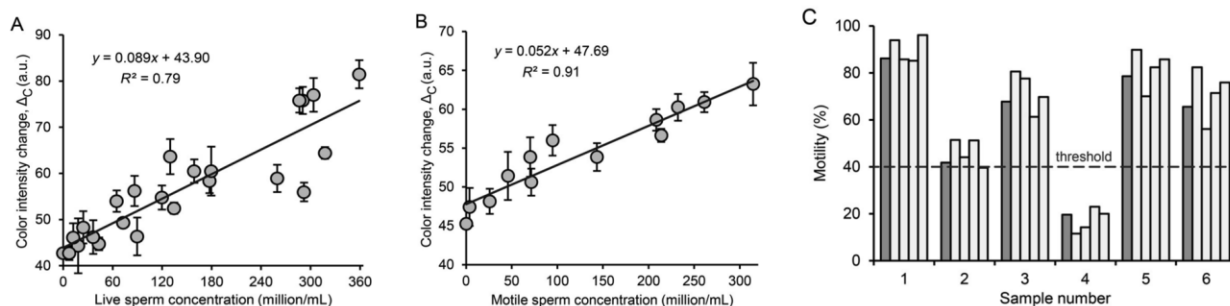
Figure 4.2. Previous clinical assessment of the paper-based male fertility device. A-B) Calibration curves for live and motile sperm concentrations. C) Comparison of motility values measured with the device vs. those measured with standard clinical approaches.[25] *Figure reproduced with permission of the rights holder, Clinical Chemistry*

Previously the colourimetric sensor was read by a scanner (Fig. 4.1b). The scanner was essential to obtain quantitative results from the colorimetric assay. In this previous clinical testing, proof of concept devices were tested using human semen samples from patients ($n=5$) and healthy donors ($n=12$), using a scanner. Results from this paper-based device strongly correlated with clinical data from CASA and dye exclusion vitality assays run in parallel (Fig. 4.2a,b). In addition, device robustness was also tested, demonstrating resistance to high absolute humidity conditions (22.8 g/m^3) for >16 weeks when packed with desiccant. Comparative side-by-side testing of the paper-based device with current clinical approaches (CASA and dye exclusion vitality assays) was conducted along with testing using semen obtained from patients ($n=5$). In terms of clinical outcome, this device provided 100% agreement with the established test for patient samples tested (Fig. 4.2c).

This original device had two main shortcomings. First, the requirement for a scanner was not desireable. A prototype 'stand-along' mobile application has been developed to measure the colour change and perform data analysis. Preliminary results showed that, while the application was effective, variations in ambient light conditions reduced the accuracy of measurements. For these reasons, one focus of this thesis is to implementing and validate a smartphone imaging approach in this application, informed by the work of the previous chapter. Secondly, during the original testing a pipette was used to add small volumes (3-5 µL) of buffer. However, use of such tools is generally antithetical to point-of-care diagnostics which need to be very simple and single-component. Thus, a user-friendly, robust fluid delivery interface is needed to be integrated into the

device. These two shortcomings inform the objectives of this chapter: i) optimize the device for smartphone/mobile health diagnosis; and ii) develop a robust user-friendly fluid delivery interface.

4.2   Optimization of the Paper-based Device for Mobile Health Diagnosis

The first objective was to integrate the device with smartphone imaging and quantification technology. After colourimetric reaction, the paper-based device was imaged at various light conditions and tested for reproducibility and accuracy. The device was modified and optimized by changing the laminate film and spot diameter, and, most importantly, integrating local reference spots that enable the smartphone-based imaging app to make local adjustments for lighting. A feasible, robust light compensation method was developed.

4.2.1 Standardized Chip Fabrication Process with XTT as the Reagent

This standardization can help reduce the error caused by the fabrication process of the chips, thus contributing to more uniform results. I redesigned the chip to become more compact and symmetrical, and the background spot (B) is placed on the centerline between live concentration spot (C) and motile concentration spot (M) to better compensate for of uneven light distribution. Then, chromatography paper sheets were pretreated at 120°C for 1min to remove moisture. Painted chromatography paper sheets were heated in oven at 120°C for 3 minutes to allow wax to penetrate through the thickness forming hydrophobic layers while it does not penetrate too much in the horizontal direction. The size of the testing spots was controlled at a diameter of 4mm after heating. Reference bars and spots were printed on the same chromatography paper as the testing spots to ensure same reflectivity.

In addition to these imaging changes, I substituted purple MTT assay with orange XTT colourimetric assay. The reasoning was that the CMOS imaging array has higher sensitivity at longer wavelengths, owing to the inherent material response of Si. Both MTT and XTT colourimetric assays can quantify cellular viability, but XTT has better water solubility which is another advantage. I also determined the protocol for activated-XTT solution (XTT and PMS),

after I studied the protocol for quantification of cell number based on metabolic activity in microplates.

The dye {2,3-bis (2-methoxy-4-nitro-5-sulfophenyl)-5-[(phe- nylamino) carbonyl]-2H-tetrazolium hydroxide} (XTT) and Phenazine methosulfate (PMS) was purchased from Sigma Aldrich. A 50mM PMS solution was prepared in phosphate-buffered saline. Then, 5mg of XTT was dissolved in 1mL of phosphate-buffered saline. The XTT and PMS solutions were sealed and stored separately when not used. A 10μL of the PMS solution was added to 4mL of XTT solution created in previous step immediately before labelling cells, achieving a final concentration of 125μM. The life of paper-based chips with XTT agent were tested till 3 months with drying agent and avoiding sunlight.

4.2.2 Paper-based Device and XTT Characterization

After the new XTT colourimetric assay was adopted, characterization of the operating parameters for the paper-based semen analysis devices was performed with human semen sample, shown in Figure 4-3. The XTT density was characterized per millimeter square of testing spots, and this was controlled by the volume of 5mg/mL XTT solution dried to each testing spot. During dye density characterization, $3\mu L$ of semen was added to each testing spot for a 10 minute reaction, similar to the optimization and testing process applied previously for MTT.

The differences among these three dye densities are small but distinguishable. The 2 $\mu g/mm^2$ case shows best sensitivity and agreement in the calibration curve (Fig. 4.3a) in devices with 4mm diameter reaction spots. The limits of detection (LODs) for sperm concentration assays can be calculated as 3 times the standard deviation of the signal from a blank device with zero sperm concentration (applying only buffer on the spots and scanning after 10min). The LOD value for live sperm concentration is lowest at 2 $\mu g/mm^2$ (Table 4-1). Therefore, I selected 2 $\mu g/mm^2$ as the XTT dye density for our paper-based device, although both XTT and MTT were used in debugging the paper-based diagnostic as described in the next section.

Table 4-1: The limits of detection for sperm concentration assays for various dye concentrations

Dye Density	3σ (ΔRGB)
1 μg/mm^2	10.0
2 μg/mm^2	3.1
3 μg/mm^2	9.6

With regard to sample volume, the colour change resulting from $5\mu L \sim 20\mu L$ semen volume at the testing spot showed no significant differences in signal (Fig. 4.3b). This indicates that small variances on the semen volume delivered to the testing spot will have minor effect on the colour change and the final result. A volume of $3\mu L$ of semen resulted in a significant smaller colour change (Fig. 4.3b) due to the evaporation and drying effect on such a small volume, indicating a lower bound on the volume. With respect to reaction time, the colour change represented by ΔRGB increased from 5 minutes to 15 minutes and plateaued thereafter for all sufficient semen volumes (over $3\mu L$) (Fig. 4.3b). Thus, the recommended conditions for semen volume and reaction time are 10 μL & 15 min.
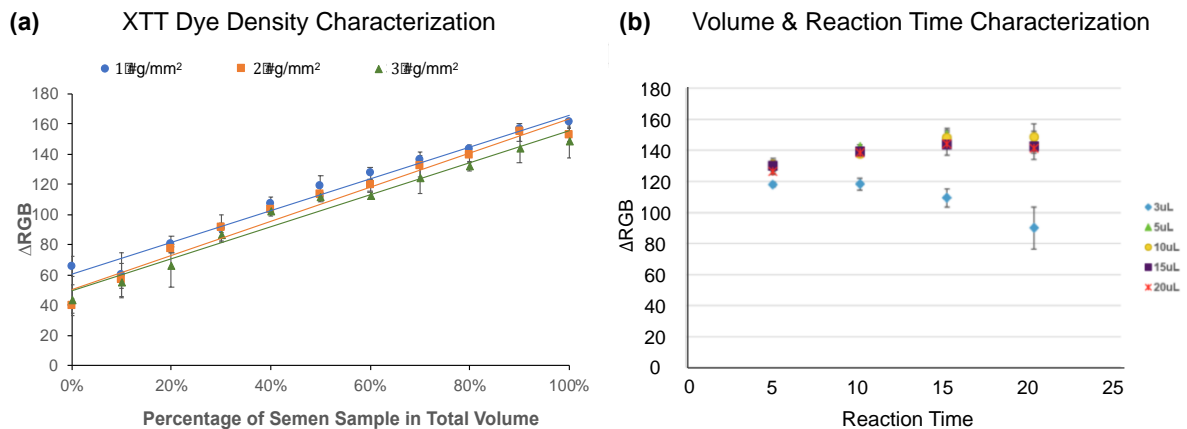
Figure 4.3. Characterization of the operating parameters of the paper-based semen analysis device. (a) **ΔRGB** as a function of percentage of semen sample in total volume for three difference XTT dye density per millimeter square. Sample concentration was varied by serial dilution. (b) **ΔRGB** as a function of sample volume and reaction time in minutes.

4.3 Bacteria Cross-sensitivity Issue and Solution

4.3.1 Cross-sensitivity Issue Realized

During clinical validation of the paper-based semen diagnostic device with smartphone imaging method and fluid delivery systems (illustrated in Section 4.4), a limitation of the original paper-based device was discovered. Specifically, the colourimetric reaction occurred even with azoospermia and post-vasectomy semen samples (i.e. false positives indicating sperm when negligible sperm present) in cases where those samples contained high numbers of bacteria. Azoospermia is the condition of a man whose semen contains no sperm and vasectomy is a surgical process to prevent fertilization. These samples were proven to have little or no sperms. For some azoospermia patient samples and post vasectomy patient samples, the device gave evident colour change (Fig. 4.4), corresponding to sperm concentration ≥20 million/mL. As specialists in the most challenging cases of male infertility, the Jarvi lab has provided us access to these important test cases – through which we discovered the cross-reactivity issue.
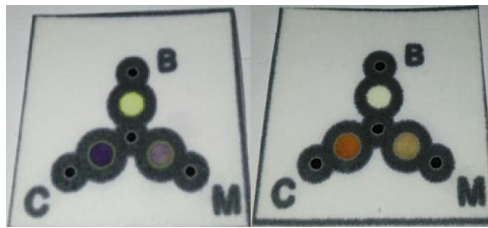
Figure 4.4. The colourimetric reaction of an azoospermia sample with MTT (left) and XTT (right).

To investigate the cause of the colour change, I centrifuged the semen sample and separated the supernatant and the participated pellet into 2 centrifuge tubes. Since sperm cells have greater mass than other cells in the semen, the pellet consisted of concentrated sperms, and the supernatant was seminal plasma containing bacteria and other cells. After separation, I resuspended the concentrated sperm cells with PBS to its original concentration. This process is illustrated in Figure 4.5a. Later, MTT was added to both tubes. The colourimetric reaction in the seminal plasma tube occurred right away with significant colour change observed in 10 minutes. However, noticeable colour change took 30 minutes to occur in the resuspended sperms tube. The colour change of MTT reaction with seminal plasma and with raw semen were compared for both concentration (C) spot and motility (M) spot in 10 minutes (Fig. 4.5b). Seminal plasma contributes to 92% of the colour change at C spot, and 84% of the colour change at M spot. If

error bars are considered, the results from raw semen show no difference to which of seminal plasma. The resuspended sperms were incubated for 3 hrs, then added to the paper-based device with MTT on the testing spots to investigate whether the product of sperm cells can cause colourimetric reaction (Fig. 4.5c). Results show that longer incubation time does not lead to detectable colour change at either C or M spot.

In the centrifuge tests the MTT reported purple reaction product in both the supernatant and the pellet (with sperm). The implication was that the bacteria in the supernatant were producing significant MTT signal, in the absence of sperm. Both MTT and XTT can give colourimetric signal with azoospermia samples due to cross sensitivity with bacteria cells in semen sample. Bacteriospermia is a common phenomenon in semen samples and the concentration of bacteria in semen varies from person to person.
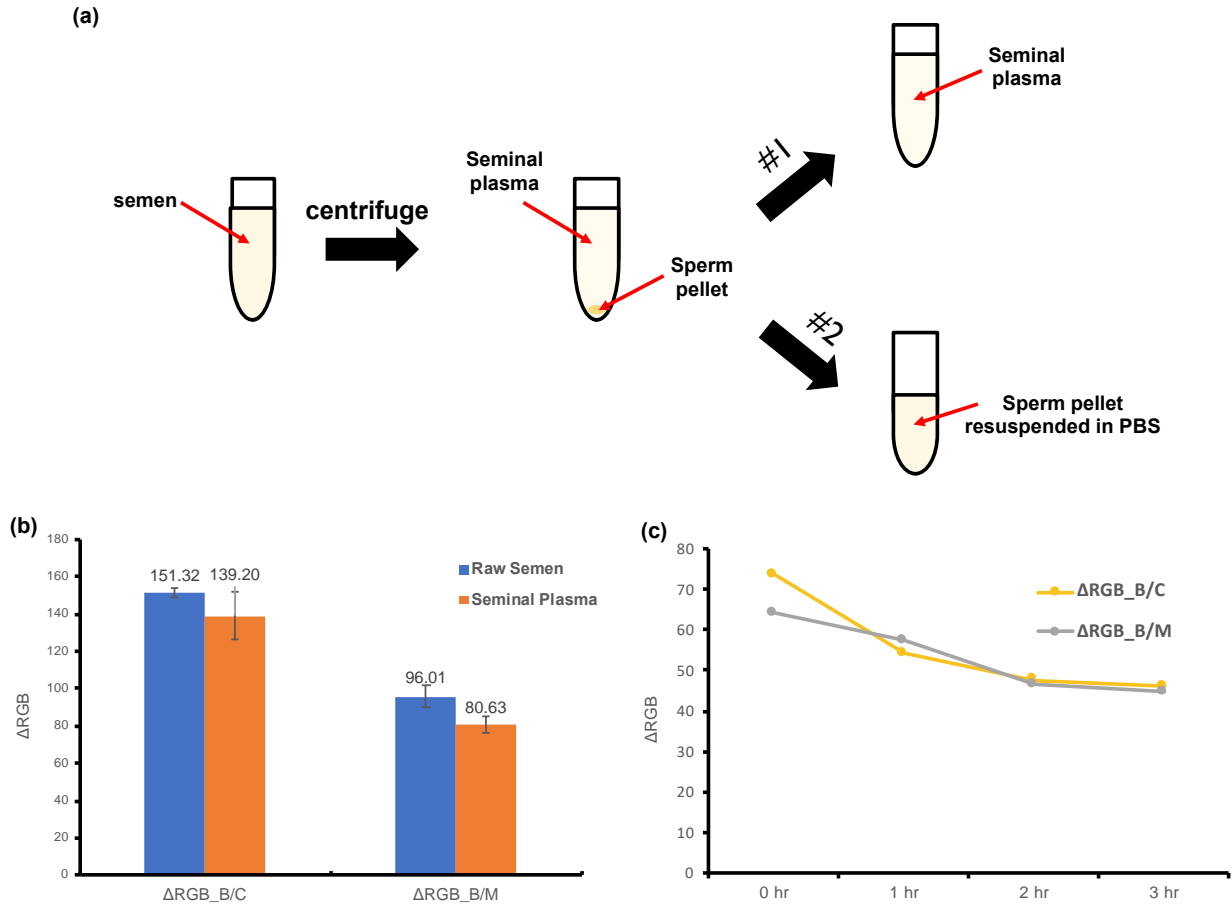
Figure 4.5. (a) Schematic of the centrifugation process to separate seminal plasma and sperm cells. (b) **ΔRGB** representation of MTT reaction with seminal plasma and with raw semen were compared for both concentration spot and motility spot in 10 minutes. (c) The effect of incubation time on colour change represented by **ΔRGB**.

4.3.2 Bacteriospemia Issue Solution

In order to solve the cross sensitivity problem explained in the previous section, I developed a cost-effective and rapid method to remove the centrifugation process as shown in Figure 4.6. The method uses a syringe and a syringe filter with pore size of 5 μm. This size is smaller than the sperm head diameter but larger than the size of most bacteria in semen samples. The purifying process includes four steps, shown in Figure 4.8: (i) withdraw the raw semen into the syringe to position A, (ii) attach the filter to the syringe, empty the syringe to position B and filter out other bacteria and other cells with smaller sizes through the syringe filter, (iii) withdraw fresh PBS into the syringe to position A to resuspend the sperm, (iv) remove the syringe filter, empty the resuspended sperms into a syringe tube, add MTT to resuspended sperm, wait 30 minutes and add it to the spots on paper-based chips.
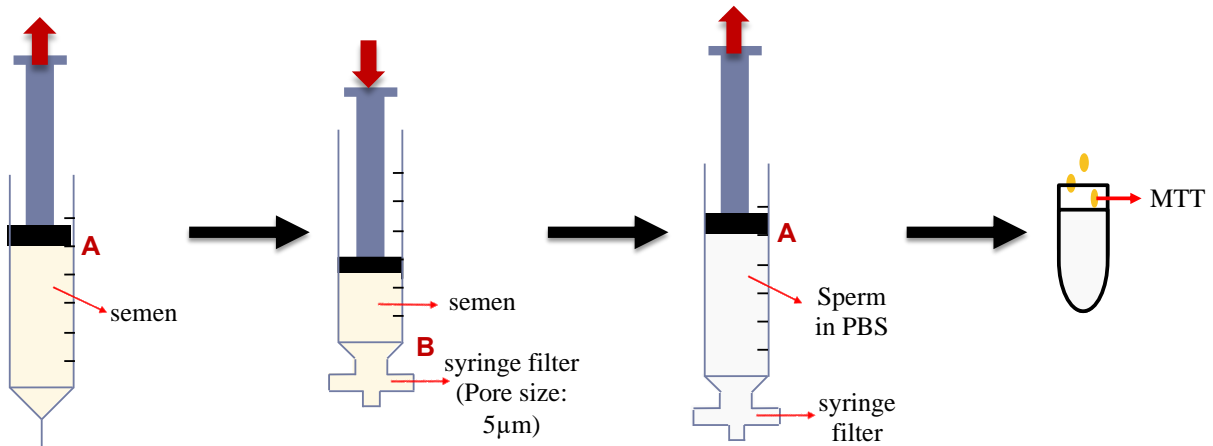
Figure 4.6. Schematics of a cost-effective, rapid and portable semen purification method with a syringe and a syringe filter.

Resuspended sperm, using the approach shown in Figure 4.6, was applied to the original paper-based test (instead of raw semen). Neither the motility test (swim through membrane to dye) nor the count spot, were effective with resuspended sperm. MTT, positively charged tetrazolium salt, is reduced primarily intracellularly. By dying the cells, MTT also slow kills them. Based on the literature and our experiments, MTT has minor impact on sperm cell motility within 15 minutes. However, after 15 minutes, the number of motile sperms reduces and the remaining motile cells also progress slower. The colourimetric reaction of MTT and resuspended sperm cells within 15 minutes remained undetectable. The only viable way to obtain signal from the resuspended sperm was to dye the fluid directly as shown in Figure 4.6. This altered method can be used for measuring only the live cell concentrations, not the motile sperm concentrations or the motility. Perhaps with future optimization of the dye, filtration and swimming conditions this method could be fully applied in the paper-based mode originally intended.

The relationship between live sperm cell concentration and colour change obtained using syringe filter method with MTT is demonstrated in Figure 4.7 with the y axis in logarithmic scale. It is common to plot ΔRGB values in a semilog or log-log figure, for instance, Murdock et al.[26] plotted the ΔRGB values against antigen concentration with a semilog best fit equation. The linear correlation between live cell concentration and ΔRGB values by iPhone has $R^2=0.969$, followed by Huawei ($R^2=0.99355$). The correlation by Samsung results are slightly weaker than

the other results from the other two phones, with $R^2 = 0.8998$. The calibration curves obtained from these three representative phones agree reasonably well.
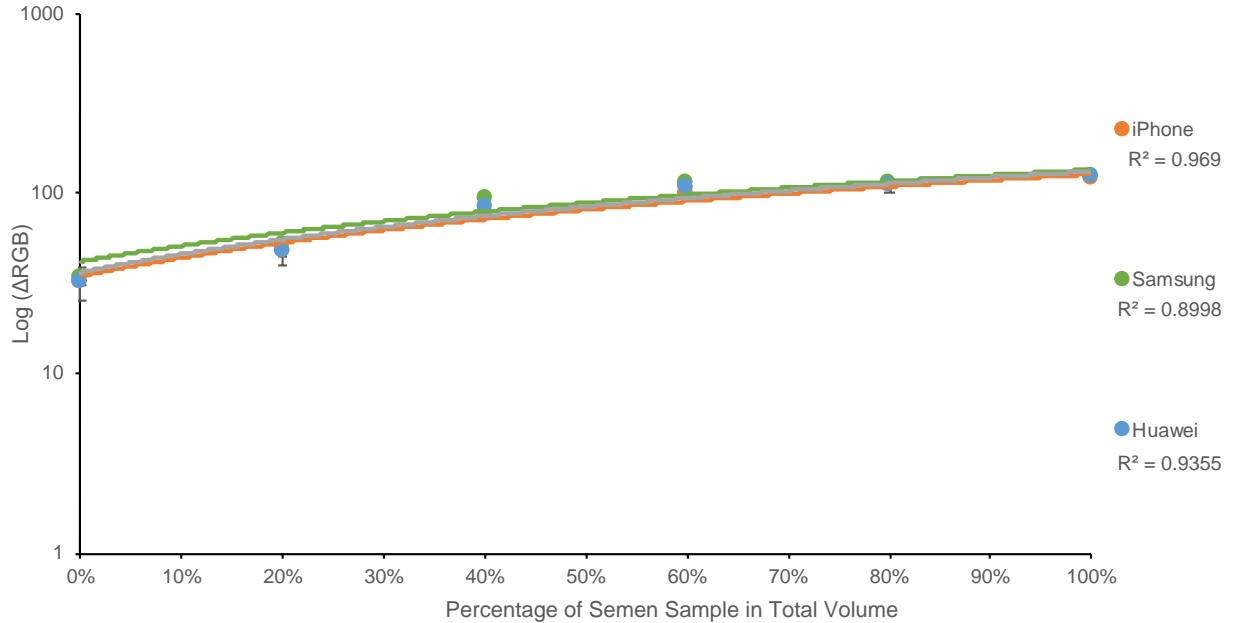
Figure 4.7. Calibration curves for colour change on paper-based chips (**ΔRGB**) with respect to live cell concentration of a semen sample with y axis in log scale for three representative phones. Sample concentration was varied by serial dilution.

4.4 Development of a Robust User-friendly Fluid Delivery Interface

As noted in the introduction, a second shortcoming of the original device was the requirement of external user fluid handling. Here I developed and integrated a robust user-friendly fluid delivery interface, integrated within the device. The previous protocol was to use a pipette to add small volumes (3 - 5μL) of buffer and semen to individual spots on the device. While very familiar to all those in the life sciences, this step represents a technical hurdle to untrained end users. For widespread implementation, the pipette must be replaced with an easy-to-use and robust fluid delivery interface.

4.4.1 Alternative Design 1: Slide Chip

To meet these requirements, a novel fluid delivery interface was proposed and shown in Figure 4.8 A sliding plate was placed onto the back side of paper-based device in hard contact. Large droplets (1~2 mL) of buffer and semen was added to the rectangular area using a dropper, which is familiar to most end-users. The plate was slid step-by-step to align the arrow to the markers, allowing for robust and accurate control of buffer and semen addition. In short, the sliding device ensured the sensor exposure and fluid volumes needed, so the user did not need to measure or apply precisely. To ensure biocompatibility, the plate and the separator layer was fabricated through 3D printing and laser cutting of polycarbonate which has regulatory approval for medical devices. Although the approach is necessarily simple, engineering this interface so that it is both simple to use *and* effective in untrained hands, is a real challenge. The design and testing took into full account the effects of cover geometry, wetting, expected user errors, variations in volumes applied, and variations in the viscosity of the semen applied.
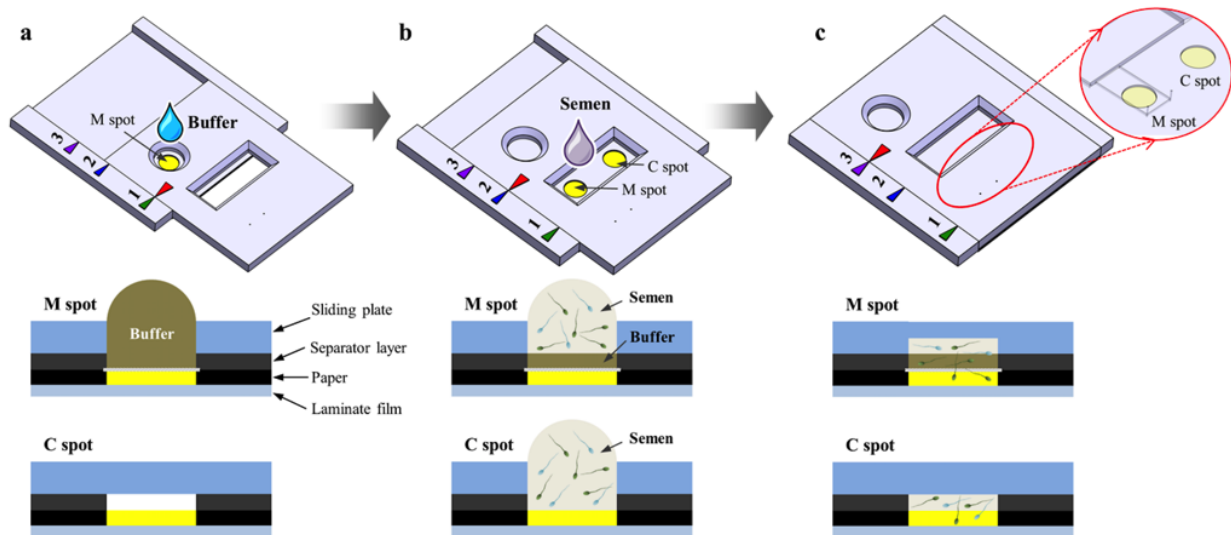
Figure 4.8. Slip chip strategy for the fluid delivery interface, alleviating the requirement of precise user fluid volume measurement. a) At position 1, add ~1 mL buffer to the circular opening, and fill the reservoir in the separator layer. b) Move the sliding plate to position 2, removing extra buffer on the M spot. Add semen sample (1~2 mL) to the rectangular opening. c) Move the sliding plate to position 3, removing extra semen on the M and C spots. Fixed small volume (3 μL) of semen sample will be left in the reservoir on the C spot, and in the gap on the M spot.

4.4.2 Alternative Design 2: Integrated Paper-based Hollow Channel with Pressure Pump Fluid Delivery Device

This second design of the fluid delivery system includes a buffer encapsulated pressure pump and a paper-based hollow channel for automatic semen delivery (Fig. 4.9a). A pressure pump was made out of two PET sheets and a laser-cut PMMA middle layer. All layers were bonded by double-sided tape with a narrow weak bonding spot (Fig. 4.9d). This weak bonding spot kept buffer inside the device without external force and also allowed buffer to be guided to the motility spot when pressing. The buffer volume was controlled by the volume prestored in devices. Although the pressure pump has a weakly bonded spot, fluid can be trapped for more than a week at the temperature of 37 ℃ and at 17% humidity.

After bonding, the bottom tip of the paper-based hollow channel chip was put into semen sample which was automatically introduced into the channel and delivered to both concentration and motility spot. I used the laser cutter to create the pattern for the hollow channel and then bonded it to the two outside layers with double-sided tape. The semen volume at the testing spots was controlled by the thickness of hollow layers at these spots. The paper-based hollow channel device not only allows fluid to travel upward through the channel, it also increases the fluid traveling speed, especially at the beginning of fluid delivery. It took about 30s for semen to travel to concentration spot and it took about 60s for semen to travel to motility spot because motility spot was prewetted.
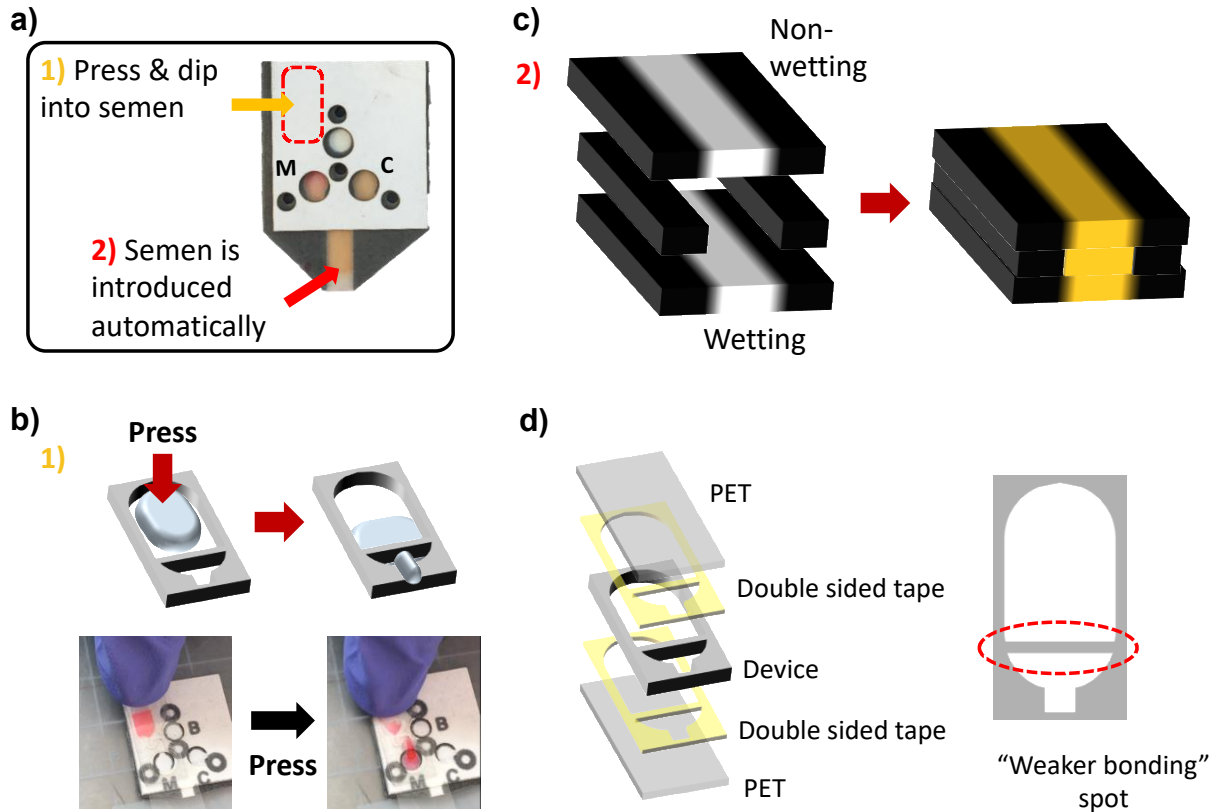
Figure 4.9. Illustration of the paper-based hollow channel 3D device. (a) demonstrates the 2 steps working principle of the device. (b) illustrates how to use buffer encapsulated pressure pump. (c) is the cross section view of the 3D paper-based hollow channel for automatic semen introduction. (d) is the detailed illustration of the pressure pump for buffer.

4.5   Conclusion

Recalling the two main objectives for this development of this device toward commercialization. The optimization of the device for smartphone health diagnosis, in applying the imaging method developed in Chapter 3, was successful. XTT was a better choice for smartphone applications than MTT, because CMOS imaging array in smartphone camera has higher sensitivity at longer wavelengths. The colourimetric reaction for XTT is from light yellow to orange, the reaction for MTT is from yellow to purple and orange has a longer wavelength than purple. The XTT performance was further optimized for dye density, semen volume, and reaction time. For cellphone results, logarithmic best fit curves show better approximation of data points and better agreement among the best fit curves of different smartphones.

In clinical testing with challenging samples, however, I discovered a cross-reactivity issue with the device and protocol. Namely, bacteria in the semen sample provided substantial signal in some cases – in some extreme cases, false-positives were recorded for samples with no sperm. To mitigate this issue I developed a syringe filter approach (filtering out the bacteria in the semen and resuspending the sperms) to solve bacteriospermia problem. While successful in tube-based testing this method has not – as yet – been successful in the paper based platform originally envisioned. The addition of this user processing step is also not in the spirit of POC diagnostics, and presents some barriers to commercialization and widespread use: (i) the procedures can be too complex for untrained users to operate, (ii) the addition of syringe and syringe filters increases the material cost of the diagnostic device, and (iii) this device can effectively measure live sperm concentration, but it is not capable of motility-spot testing which was a differentiating feature in the marketplace.

The second shortcoming of the original device was a lack of integration of the buffer addition processing step. Here, two integrated fluid delivery interfaces were developed and fabricated for clinical validation. The methods were successful in preliminary tests, and show promise for future paper based diagnostic testing.

Chapter 5

5 Conclusion

5.1 Summary

Male infertility is a growing global concern, which drives the development of affordable, accessible, accurate and point-of-care male fertility diagnosis. The thesis describes the development of mobile health screening for male fertility. This development can be divided into three stages: i) develop an image sensing platform and optimize the device for mobile health diagnosis; ii) develop a robust user-friendly fluid delivery interface; iii) perform necessary clinical validation and testing essential to engage early-stage investment.

At the first stage, a smartphone colour imaging, calibration and representation platform that can provide quantitative data for precise single-colour shift analysis was developed. This comprehensive method is reliable, accessory-free and it can compensate for different indoor lighting sources, various lighting intensities and different brands and models of smartphones partially due to using the continuous LED light at the back of the phone at dominant light source when imaging. After imaging, the RGB values were calibrated with the black and white backgrounds and then represented by ΔRGB. The combination of the calibration method and representation further reduces the effect of changing ambient lights and different cellphone models. After utilizing the method developed, the ANOVA test result shows no significant difference (p value >0.05) among various lighting conditions by different phone models. this imaging method is combined with paper-based technology for a low cost and ease-of-use yeast viability quantification and it becomes an attractive alternative to conventional laboratory cell counting, with future potential for home brewing application. The optimization of the paper-based chip for mobile screening diagnostics is achieved by choosing the colourimetric dye that works better with CMOS array and choosing the optimal dye density, semen volume and reaction time.

At the second stage, two fluid delivery systems were designed, fabricated and tested to achieve ease of use. The difficulties in designing the fluid delivery systems are (i) ensuring small amount of fluid delivered, (ii) buffer addition prior to semen, (iii) robust assembly and (iv) design for ultimate mass manufacturing. For slip chip, accurate addition of fluid to the spot are controlled

by simply moving of the top layer of the slip device. By moving from position 1 to position 2, controlled amount of buffer was added to the motility spot only. Sliding the top layer from position 2 to position 3 adds a small volume of semen sample to both live sperm concentration spot and motility spot. The usage of the integrated paper-based hollow channel and pressure pump device is even simpler. A simple press is all the user need to do to add the buffer. Then the tip of the chip can be left in semen and fluid will travel to both C and M spots automatically.

At the third stage, the paper-based chip was found giving false positive results due to its cross sensitivity with bacteria in the semen. In order to solve this problem, I developed a syringe filter method which filters out bacterial and resuspends the sperm cells to remove bacteria from semen effectively. This approach advances the accuracy of the portable male fertility diagnostics technology. However, the added steps are too complicated for untrained users to learn. More importantly, only live sperm concentration can be accurately measured after modification. This approach is not capable of providing sufficient signal for motility-spot testing, but motility is a crucial indicator of male fertility potential. This huge trade-off of accuracy vs. complexity and cost makes the device less commercially viable and it stops us from pursuing commercialization in this project. A few lessons learned are that: (i) bacteriospermia is common in male semen samples and it has always been overlooked in home-based semen diagnostics, (ii) the gap between laboratory-based testing and low-cost diagnostics is huge, normally due to the trade-offs of performance vs cost and complexity, (iii) the collaboration with labs/clinics with male fertility background is important and valuable.

As an additional contribution to work in this area, I partnered with lab members in pursuing a machine learning approach to sperm analysis at single cell level using a non-linear regression modeling. As these contributions are not directly in line with the central theme of the thesis they are including only in Appendix D. I note here, however, my interest, enthusiasm, and some early success in pursuing this direction. With the use of a five-layer deep learning neural network, the accuracy of the approximation doubled, compared to the accuracy of linear fitting. This deep learning neural network approach is promising and has the potential of better performance with further optimizations on both accuracy and stability – all of which would benefit the larger challenge of male fertility.

5.2 Future Outlook

The current methods for both smartphone imaging and fluid delivery were successful to some extent. The cross-reactivity shortcoming was a surprise and necessitated some changes in the workflow and emphasis in this work. All contributions here have applicability both in male fertility diagnostics as well as more broadly in paper-based diagnostics and sperm analysis. These methods could also be further improved in the future. There is significant motivation to do so both from the perspective of human health and the market demand for diagnostics.

For the smartphone imaging, calibration and representation method discussed in Chapter 3, images are analyzed with ImageJ. The imaging process is currently under video mode due to the fact that most of smartphones could not keep the LED light continuously on under camera mode. The next step of this work can be developing a smartphone app to control the LED light in camera mode and writing an algorithm to detect, digitalize, correct and analyze images. This app can also provide guidance to end users in preparation of the assay and alignment of the chip. Furthermore, the calibration curve of yeast cell viability can be preinstalled into the software, then this app can be used towards, for instance, by individuals for home brewing or a wide range of industrial processes involving yeast cultures. Also the variance among the LED light for different phones can be further investigated. While I make the case that the general method and test hardware should be phone-independent (as achieved here), a smartphone app could slightly tune, for instance the LED brightness, to correct for model variations. Another noted limitation of the imaging method is when the surrounding light noise equals or exceeds the LED light intensity, such as full daylight and direct sunlight with a light intensity up to 30,000 lux. If not solution is found for this high-light issue, the smartphone app could simply alert the user when environmental light intensity is too high and reject the taken image.

The future improvements on the fluid delivery system and male fertility diagnostics discussed in Chapter 4 could involve other substrate materials, like PDMS, instead of paper. Both Trak and YoSperm competing product offerings are simplified versions of laboratory-based experiments, and the ubiquitous classical swim-up test has not been simplified for diagnostic purposes yet. The possible next step for male fertility diagnostics can be to develop an ease-of-use and low cost version of swim-up test.

Lastly, I include here some critical reflection on the additional contributions in Appendix D - non-linear regression model for sperm quality, with limitations of the study and potential improvements identified. The model is inspired by the neural network toolbox developed by MATLAB. Therefore, some of the functions and commands used in the current model are built-in functions with slight modification. Thus, future work on the model should involve further customization and optimization to improve the model's accuracy and stability. These possible steps are listed as follows: (i) explore the effect of number of nodes on accuracy, (ii) explore the effect of number of hidden layers on accuracy, (iii) test the model with other training functions, such as scaled conjugate gradient, Bayesian regularization backpropagation, and resilient backpropagation, (iv) adjust the learning rate for these training functions. Bayesian regularization backpropagation works better for some noisy and small problems, but it takes longer to train and to obtain solution. Scaled conjugate gradient is usually applied to large dataset and neural networks problems since it uses gradient calculations which are more memory efficient than the Jacobian calculations, used by LM backpropagation. Furthermore, the R value for training is currently less than 0.7. It is possible that the model is still undertrained. Possible solutions for this specific problem can be similar to the first three steps listed above, increasing the number of nodes in each hidden layer, increasing the number of hidden layers, and testing alternative training functions. An important objective is to avoid overtraining, which can be prevented by the following actions: (i) monitoring the R value for testing and ensuring that it does not drop while more layers and more nodes are added, (ii) keeping track of the mean squared error of all training, validation, and testing data and stopping the training process at the repetition of the best validation performance.

Collectively this work advanced the field of male fertility, with emphasis on paper-based male fertility diagnostics. While challenges were encountered, several solutions were presented that bring these technologies closer to market. This work also highlighted the challenges of developing low cost, simple diagnostics for complex analysis of complex fluids! The additional machine learning element provided an excellent introduction for me in this growing field – one that shows great potential to inform on the mechanisms of male infertility.

References or Bibliography

1. Bushnik, T., Cook, J. L., Yuzpe, A. A., Tough, S. & Collins, J. Estimating the prevalence of infertility in Canada. *Hum. Reprod.* **27,** 738–746 (2012).

2. Boivin, J., Bunting, L., Collins, J. A. & Nygren, K. G. International estimates of infertility prevalence and treatment-seeking: Potential need and demand for infertility medical care. *Hum. Reprod.* **22,** 1506–1512 (2007).

3. Ombelet, W., Cooke, I., Dyer, S., Serour, G. & Devroey, P. Infertility and the provision of infertility medical services in developing countries. Hum Reprod Update. **14,** 605–21 (2008).

4. Schultz, R. M. The Science of ART. *Science (80-.).* **296,** 2188–2190 (2002).

5. Agarwal, A. & Said, T. M. Role of sperm chromatin abnormalities and DNA damage in male infertility. *Hum. Reprod. Update* **9,** 331–345 (2003).

6. Agarwal, A., Mulgund, A., Hamada, A. & Chyatte, M. R. A unique view on male infertility around the globe. *Reprod. Biol. Endocrinol.* **13,** 1–9 (2015).

7. Nosrati, R. *et al.* Microfluidics for sperm analysis and selection. *Nat. Rev. Urol.* **14,** 707–730 (2017).

8. Lackner, J. *et al.* Constant decline in sperm concentration in infertile males in an urban population: Experience over 18 years. *Fertil. Steril.* **84,** 1657–1661 (2005).

9. Centola, G. M., Blanchard, A., Demick, J., Li, S. & Eisenberg, M. L. Decline in sperm count and motility in young adult men from 2003 to 2013: Observations from a U.S. sperm bank. *Andrology* **4,** 270–276 (2016).

10. Geoffroy-Siraudin, C. *et al.* Decline of semen quality among 10 932 males consulting for couple infertility over a 20-year period in Marseille, France. *Asian J. Androl.* **14,** 584–590 (2012).

11. Joensen, U. N., Skakkebæk, N. E. & Jørgensen, N. Is there a problem with male

reproduction? *Nat. Clin. Pract. Endocrinol. Metab.* **5,** 144–145 (2009).

12. Storgaard, L., Bonde, J. P. & Olsen, J. Male reproductive disorders in humans and prenatal indicators of estrogen exposure: A review of published epidemiological studies. *Reprod. Toxicol.* **21,** 4–15 (2006).

13. Skakkebaek, N. E. *et al.* Male Reproductive Disorders and Fertility Trends: Influences of Environment and Genetic Susceptibility. *Physiol. Rev.* **96,** 55–97 (2015).

14. Bieniek, J. M. *et al.* Influence of increasing body mass index on semen and reproductive hormonal parameters in a multi-institutional cohort of subfertile men. *Fertil. Steril.* **106,** 1070–1075 (2016).

15. Wang, S. Q., Xu, F. & Demirci, U. Advances in developing HIV-1 viral load assays for resource-limited settings. *Biotechnol. Adv.* **28,** 770–781 (2010).

16. Lee, W. G., Kim, Y. G., Chung, B. G., Demirci, U. & Khademhosseini, A. Nano/Microfluidics for diagnosis of infectious diseases in developing countries. *Adv. Drug Deliv. Rev.* **62,** 449–457 (2010).

17. Maatman, T. J., Aldrin, L. & Carothers, G. G. Patient noncompliance after vasectomy. *Fertil. Steril.* **68,** 552–555 (1997).

18. Hong, J. Il & Chang, B. Y. Development of the smartphone-based colorimetry for multi-analyte sensing arrays. *Lab Chip* **14,** 1725–1732 (2014).

19. Jia, M. Y. *et al.* The calibration of cellphone camera-based colorimetric sensor array and its application in the determination of glucose in urine. *Biosens. Bioelectron.* **74,** 1029–1037 (2015).

20. Cheng, C. M. *et al.* Paper-based elisa. *Angew. Chemie - Int. Ed.* **49,** 4771–4774 (2010).

21. Gong, M. M., Nosrati, R., San Gabriel, M. C., Zini, A. & Sinton, D. Direct DNA Analysis with Paper-Based Ion Concentration Polarization. *J. Am. Chem. Soc.* **137,** 13913–13919 (2015).

22. Gong, M. M., Macdonald, B. D., Nguyen, T. V., Van Nguyen, K. & Sinton, D. Lab-in-a-pen: A diagnostics format familiar to patients for low-resource settings. *Lab Chip* **14,** 957–963 (2014).

23. Gong, M. M., Zhang, P., Macdonald, B. D. & Sinton, D. Nanoporous membranes enable concentration and transport in fully wet paper-based assays. *Anal. Chem.* **86,** 8090–8097 (2014).

24. Wang, S. *et al.* Integration of cell phone imaging with microchip ELISA to detect ovarian cancer HE4 biomarker in urine at the point-of-care. *Lab Chip* **11,** 3411–3418 (2011).

25. Nosrati, R. *et al.* Paper-based quantification of male fertility potential. *Clin. Chem.* **62,** 458–465 (2016).

26. Murdock, R. C. *et al.* Optimization of a paper-based ELISA for a human performance biomarker. *Anal. Chem.* **85,** 11634–11642 (2013).

27. Zhao, C. & Liu, X. A portable paper-based microfluidic platform for multiplexed electrochemical detection of human immunodeficiency virus and hepatitis C virus antibodies in serum. *Biomicrofluidics* **10,** (2016).

28. Song, S. *et al.* Biosensors and Bioelectronics Highly sensitive paper-based immunoassay using photothermal laser speckle imaging. *Biosens. Bioelectron.* **117,** 385–391 (2018).

29. Manuscript, A. & Structures, T. NIH Public Access. **6,** 247–253 (2009).

30. Oncescu, V., O'Dell, D. & Erickson, D. Smartphone based health accessory for colorimetric detection of biomarkers in sweat and saliva. *Lab Chip* **13,** 3232–3238 (2013).

31. Shen, L., Hagen, J. A. & Papautsky, I. Point-of-care colorimetric detection with a smartphone. *Lab Chip* **12,** 4240–4243 (2012).

32. Fabrication of low-cost paper-based diagnostic .

33. Renault, C., Koehne, J., Ricco, A. J. & Crooks, R. M. Three-dimensional wax patterning of paper fluidic devices. *Langmuir* **30,** 7030–7036 (2014).

34. Renault, C., Anderson, M. J. & Crooks, R. M. Electrochemistry in hollow-channel paper analytical devices. *J. Am. Chem. Soc.* **136,** 4616–4623 (2014).

35. Channon, R. B. *et al.* Rapid flow in multilayer microfluidic paper-based analytical devices. *Lab Chip* **18,** 793–802 (2018).

36. Amann, R. P. & Waberski, D. Computer-assisted sperm analysis (CASA): Capabilities and potential developments. *Theriogenology* **81,** 5–17 (2014).

37. Amann, R. P. & Katz, D. F. Andrology Lab Corner*: Reflections on CASA After 25 Years. *J Androl* **25,** 317–325 (2004).

38. Jeyendran, R. S., Van der Ven, H. H., Perez-Pelaez, M., Crabo, B. G. & Zaneveld, L. J. D. Development of an assay to assess the functional integrity of the human sperm membrane and its relationship to other semen characteristics. *Reproduction* **70,** 219–228 (1984).

39. Coppola, M. A. *et al.* SpermCheck®Fertility, an immunodiagnostic home test that detects normozoospermia and severe oligozoospermia. *Hum. Reprod.* **25,** 853–861 (2010).

40. Yu, S. *et al.* Emerging technologies for home-based semen analysis. *Andrology* **6,** 10–19 (2018).

41. Chen, Y. A. *et al.* Analysis of sperm concentration and motility in a microfluidic device. *Microfluid. Nanofluidics* **10,** 59–67 (2011).

42. Kanakasabapathy, M. K. *et al.* An automated smartphone-based diagnostic assay for point-of-care semen analysis. *Sci. Transl. Med.* **9,** 1–14 (2017).

43. Nasr-Esfahani, M. H., Aboutorabi, R., Esfandiari, E. & Mardani, M. Sperm MTT viability assay: A new method for evaluation of human sperm viability. *J. Assist. Reprod. Genet.* **19,** 477–482 (2002).

44. Trendowski, M. *et al.* The real deal: Using cytochalasin B in sonodynamic therapy to preferentially damage leukemia cells. *Anticancer Res.* **34,** 2195–2202 (2014).

45. Roehm, N. W., Rodgers, G. H., Hatfield, S. M. & Glasebrook, A. L. An improved

colorimetric assay for cell proliferation and viability utilizing the tetrazolium salt XTT. *J. Immunol. Methods* **142,** 257–265 (1991).

46. Al-Bakri, A. G. & Afifi, F. U. Evaluation of antimicrobial activity of selected plant extracts by rapid XTT colorimetry and bacterial enumeration. *J. Microbiol. Methods* **68,** 19–25 (2007).

47. Gabrielson, J. *et al.* Evaluation of redox indicators and the use of digital scanners and spectrophotometer for quantification of microbial growth in microplates. *J. Microbiol. Methods* **50,** 63–73 (2002).

48. Yu, L. *et al.* Flow-through functionalized PDMS microfluidic channels with dextran derivative for ELISAs. *Lab Chip* **9,** 1243–1247 (2009).

49. Soldat, D. J., Barak, P. & Lepore, B. J. Microscale colorimetric analysis using a desktop scanner and automated digital image analysis. *J. Chem. Educ.* **86,** 617–620 (2009).

50. Tohda, K. & Gratzl, M. Micro-miniature autonomous optical sensor array for monitoring ions and metabolites 1: design, fabrication, and data analysis. *Anal. Sci.* **22,** 383–388 (2006).

51. Suzuki, K. *et al.* Ionophore-based lithium ion film optode realizing multiple color variations utilizing digital color analysis. *Anal. Chem.* **74,** 5766–5773 (2002).

52. Lee, D., Chou, W. P., Yeh, S. H., Chen, P. J. & Chen, P. H. DNA detection using commercial mobile phones. *Biosens. Bioelectron.* **26,** 4349–4354 (2011).

53. Zhu, H., Mavandadi, S., Coskun, A. F., Yaglidere, O. & Ozcan, A. Optofluidic fluorescent imaging cytometry on a cell phone. *Anal. Chem.* **83,** 6641–6647 (2011).

54. Martinez, A. W. *et al.* Simple Telemedicine for Developing Regions : Camera Phones and Paper-Based Microfluidic Devices for Real-Time , Off-Site Diagnosis. *Anal. Chem.* **80,** 3699–3707 (2008).

55. García, A. *et al.* Mobile phone platform as portable chemical analyzer. *Sensors Actuators, B Chem.* **156,** 350–359 (2011).

56. Ozcan, A. Mobile phones democratize and cultivate next-generation imaging, diagnostics and measurement tools. *Lab a Chip - Miniaturisation Chem. Biol.* **14,** 3187–3194 (2014).

57. Sumriddetchkajorn, S., Chaitavon, K. & Intaravanne, Y. Mobile-platform based colorimeter for monitoring chlorine concentration in water. *Sensors Actuators, B Chem.* **191,** 561–566 (2014).

58. Lee, S., Oncescu, V., Mancuso, M., Mehta, S. & Erickson, D. A smartphone platform for the quantification of vitamin D levels. *Lab a Chip - Miniaturisation Chem. Biol.* **14,** 1437–1442 (2014).

59. Cantrell, K., Erenas, M. M., De Orbe-Payá, I. & Capitán-Vallvey, L. F. Use of the hue parameter of the hue, saturation, value color space as a quantitative analytical parameter for bitonal optical sensors. *Anal. Chem.* **82,** 531–542 (2010).

60. Dawidson, I., Blom, M., Lundeberg, T., Theodorsson, E. & Angmar-m, B. Neuropeptides in Tfie Saliva of Healthy Subjects. **60,** 269–278 (1997).

61. Oncescu, V. *et al.* Point-of-care colorimetric detection with a smartphone. *Lab Chip* **85,** 4240–4243 (2014).

62. Yetisen, A. K., Martinez-Hurtado, J. L., Garcia-Melendrez, A., Da Cruz Vasconcellos, F. & Lowe, C. R. A smartphone algorithm with inter-phone repeatability for the analysis of colorimetric tests. *Sensors Actuators, B Chem.* **196,** 156–160 (2014).

63. Sicard, C. *et al.* Tools for water quality monitoring and mapping using paper-based sensors and cell phones. *Water Res.* **70,** 360–369 (2015).

64. Zhang, Y. *et al.* Postage stamp-sized array sensor for the sensitive screening test of heavy-metal ions. *Analyst* **139,** 4887–4893 (2014).

65. Coskun, A. F. *et al.* A personalized food allergen testing platform on a cellphone. *Lab Chip* **13,** 636–640 (2013).

66. Salles, M. O., Meloni, G. N., De Araujo, W. R. & Paixão, T. R. L. C. Explosive colorimetric discrimination using a smartphone, paper device and chemometrical

approach. *Anal. Methods* **6,** 2047–2052 (2014).

67. Zhang, D., Gao, B., Chen, Y. & Liu, H. Converting colour to length based on the coffee-ring effect for quantitative immunoassays using a ruler as readout. *Lab Chip* **18,** 271–275 (2018).

68. Shen, L., Ratterman, M., Klotzkin, D. & Papautsky, I. A CMOS optical detection system for point-of-use luminescent oxygen sensing. *Sensors Actuators, B Chem.* **155,** 430–435 (2011).

69. Pretorius, I. S. Tailoring wine yeast for the new millennium: Novel approaches to the ancient art of winemaking. *Yeast* **16,** 675–729 (2000).

70. Lodolo, E. J., Kock, J. L. F., Axcell, B. C. & Brooks, M. The yeast Saccharomyces cerevisiae - The main character in beer brewing. *FEMS Yeast Res.* **8,** 1018–1036 (2008).

71. Nielsen, J. Yeast cell factories on the horizon. *Science (80-.).* **349,** 1050–1051 (2015).

72. Basso, L. C., De Amorim, H. V., De Oliveira, A. J. & Lopes, M. L. Yeast selection for fuel ethanol production in Brazil. *FEMS Yeast Res.* **8,** 1155–1163 (2008).

73. Caspeta, L., Buijs, N. A. A. & Nielsen, J. The role of biofuels in the future energy supply. *Energy Environ. Sci.* **6,** 1077–1082 (2013).

74. Lopes, M. L. *et al.* Ethanol production in Brazil: a bridge between science and industry. *Brazilian J. Microbiol.* **47,** 64–76 (2016).

75. Feizi, A. *et al.* Rapid, portable and cost-effective yeast cell viability and concentration analysis using lensfree on-chip microscopy and machine learning. *Lab Chip* **16,** 4350–4358 (2016).

76. Salomone, J. Women as wombs: Reproductive Technologies and the Battle over Women's Freedom. *Womens. Stud. Int. Forum* **19,** 346–347 (1996).

77. Ombelet, W., Cooke, I., Dyer, S., Serour, G. & Devroey, P. Infertility and the provision of infertility medical services in developing countries. *Hum. Reprod. Update* **14,** 605–621

(2008).

78. Edition, F. WHO laboratory manual for the Examination and processing of human semen. *World Health* **Edition, V,** 286 (2010).

79. Hamilton, J. A. M. *et al.* Total motile sperm count: A better indicator for the severity of male factor infertility than the WHO sperm classification system. *Hum. Reprod.* **30,** 1110–1121 (2015).

80. Nagata, M. P. B. *et al.* Live births from artificial insemination of microfluidic-sorted bovine spermatozoa characterized by trajectories correlated with fertility. *Proc. Natl. Acad. Sci.* 201717974 (2018). doi:10.1073/pnas.1717974115

81. R.G.Edwardsb, P. C. S. Birth after the reimplantation of a human embroyo. *Lancet* **381,** 1620 (1978).

82. Palermo, G., Joris, H., Devroey, P. & Van Steirteghem, A. C. Pregnancies after intracytoplasmic injection of single spermatozoon into an oocyte. *Lancet* **340,** 17–18 (1992).

83. Lundin, K., Söderlund, B. & Hamberger, L. The relationship between sperm morphology and rates of fertilization, pregnancy and spontaneous abortion in an in-vitro fertilization/intracytoplasmic sperm injection programme. *Hum. Reprod.* **12,** 2676–2681 (1997).

84. Nagy, Z. P., Verheyen, G., Tournaye, H. & Van Steirteghem, A. C. Special applications of intracytoplasmic sperm injection: The influence of sperm count, motility, morphology, source and sperm antibody on the outcome of ICSI. *Hum. Reprod.* **13,** 143–154 (1998).

85. Lo Monte, G., Murisier, F., Piva, I., Germond, M. & Marci, R. Focus on intracytoplasmic morphologically selected sperm injection (IMSI): A mini-review. *Asian J. Androl.* **15,** 608–615 (2013).

86. Sakkas, D., Ramalingam, M., Garrido, N. & Barratt, C. L. R. Sperm selection in natural conception: What can we learn from Mother Nature to improve assisted reproduction outcomes? *Hum. Reprod. Update* **21,** 711–726 (2015).

87. Simon, L., Zini, A., Dyachenko, A., Ciampi, A. & Carrell, D. A systematic review and meta-analysis to determine the effect of sperm DNA damage on IVF and ICSI outcome. *Asian J. Androl.* **0,** 0 (2016).

88. Zini, A., & Libman, J. Clinical Significance in the Era of Assisted Reproduction. *Can. Med. Assoc. J.* **175,** 495–500 (2006).

89. Zini, A. & Libman, J. Sperm DNA damage: clinical significance in the era of assisted reproduction. *CMAJ* **175,** 495–500 (2006).

90. Sakkas, D. *et al.* Sperm nuclear DNA damage and altered chromatin structure: effect on fertilization and embryo development. *Hum. Reprod.* **13 Suppl 4,** 11–19 (1998).

91. Concepts, N., Agarwal, A. & Setti, A. S. *Non-Invasive Sperm Selection for In Vitro Fertilization*. (2015). doi:10.1007/978-1-4939-1411-1

92. Cassuto, N. G. *et al.* A new real-time morphology classification for human spermatozoa: a link for fertilization and improved embryo quality. *Fertil. Steril.* **92,** 1616–1625 (2009).

93. Mirroshandel, S. A., Ghasemian, F. & Monji-Azad, S. Applying data mining techniques for increasing implantation rate by selecting best sperms for intra-cytoplasmic sperm injection treatment. *Comput. Methods Programs Biomed.* **137,** 215–229 (2016).

94. Dariš, B., Goropevnek, A., Hojnik, N. & Vlaisavljević, V. Sperm morphological abnormalities as indicators of DNA fragmentation and fertilization in ICSI. *Arch. Gynecol. Obstet.* **281,** 363–367 (2010).

95. Aydos, O. S., Yukselten, Y., Kaplan, F., Sunguroglu, A. & Aydos, K. Analysis of the correlation between sperm DNA integrity and conventional semen parameters in infertile men. *Türk Üroloji Dergisi/Turkish J. Urol.* **41,** 191–197 (2015).

96. Cassuto, N. G. *et al.* Correlation between DNA defect and sperm-head morphology. *Reprod. Biomed. Online* **24,** 211–218 (2012).

97. Garolla, A. *et al.* High-power microscopy for selecting spermatozoa for ICSI by physiological status. *Reprod. Biomed. Online* **17,** 610–616 (2008).

98. Wilding, M. *et al.* Intracytoplasmic injection of morphologically selected spermatozoa (IMSI) improves outcome after assisted reproduction by deselecting physiologically poor quality spermatozoa. *J. Assist. Reprod. Genet.* **28,** 253–262 (2011).

99. D. P. Evenson, Z. D. and M. R. M. Relation of mammalian sperm chromatin heterogeneity to fertility. *Science (80-.).* **210,** 1131–1133 (1980).

Appendices

A. Plots of the Comparison on Colour Quantification Methods and Rescaling from iPhone and Samsung
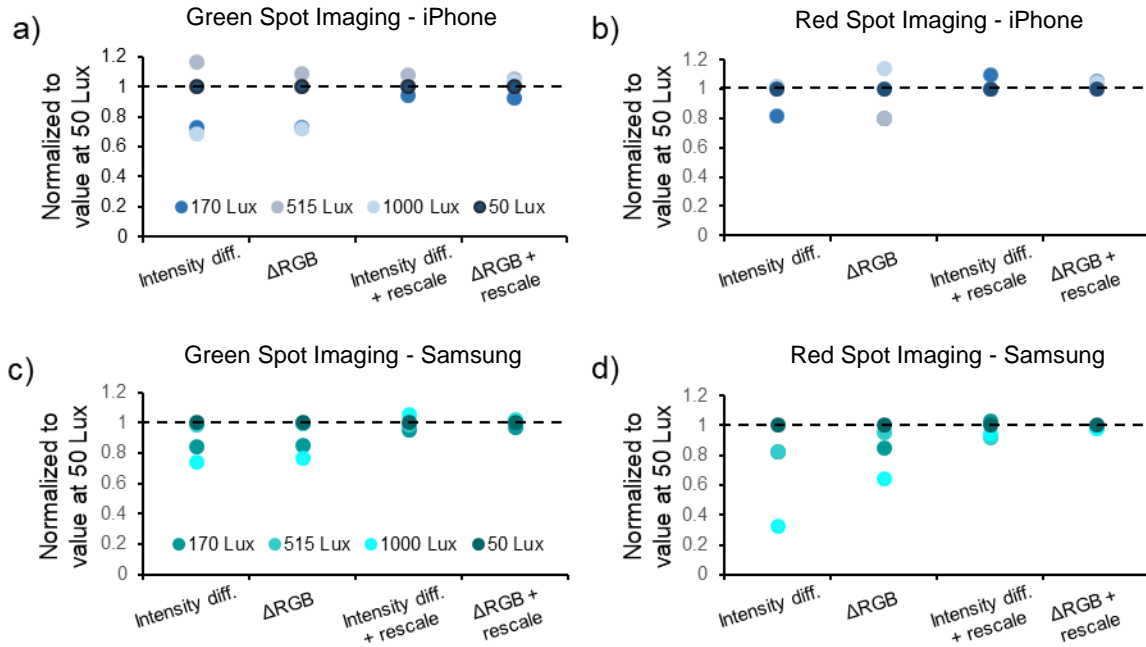
Figure A.1. Comparison of two RGB colour quantification methods(**ΔRGB** and intensity difference) before and after rescaling, without dominant light source. Images were obtained under various lighting conditions. The RGB values of 170 lux, 515 lux, 1000 lux were normalized to the value at 50 lux in their own category. (a) Green colour chip imaged with iPhone 7 Plus. (b) Red colour chip imaged with iPhone 7 Plus. (c) Green colour chip imaged with Samsung Galaxy S7 edge. (b) Red colour chip imaged with Samsung Galaxy S7 edge.

B. Plots of the Comparison between Two Imaging Conditions from iPhone and Samsung
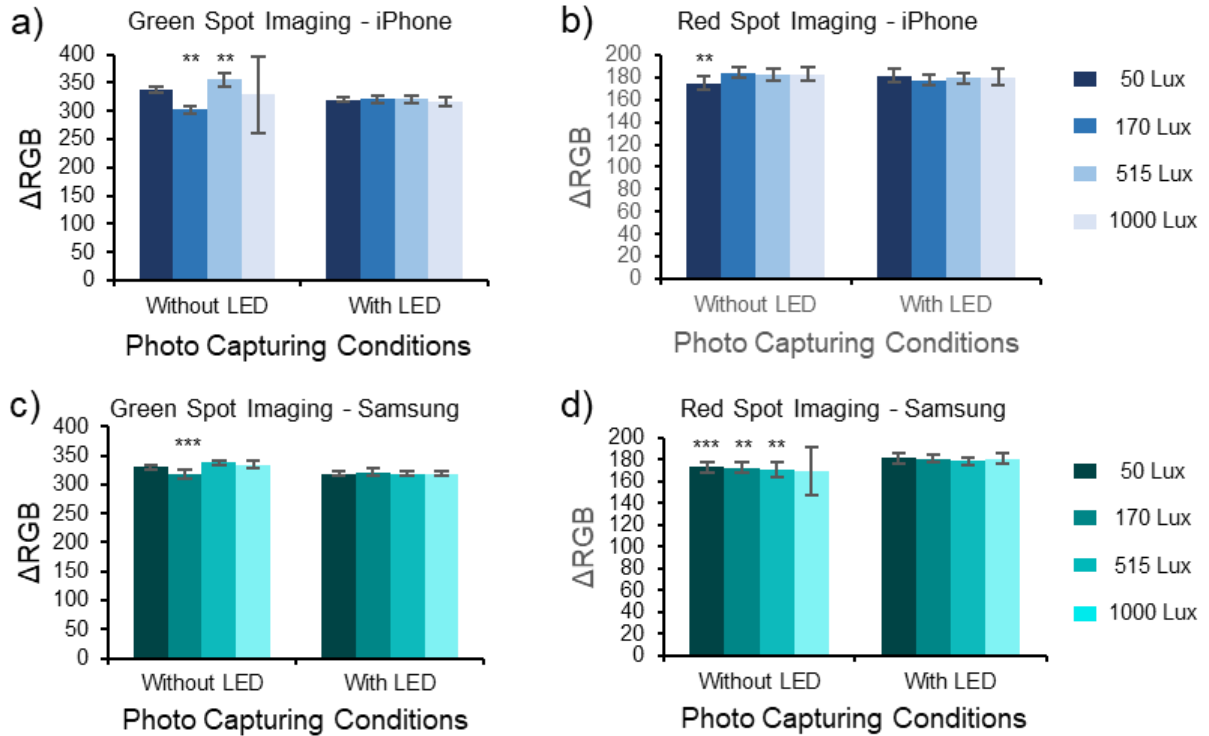
Figure B.2. The comparison between two imaging conditions, with LED light and without LED light. ANOVA and Turkey HSD comparison test were used as statistical analysis method. P < 0.05 was considered as a statistically significant difference. **p<0.01, ***p<0.001 (a-b) Images obtained with iPhone 7 Plus under various ambient lighting conditions (a) Green colour (b) Red colour (c-d) Images obtained with Samsung Galaxy S7 edge under various ambient lighting conditions (c) Green colour (d) Red colour chip

C. Plots of the Calibration Curves of Yeast Live Cell Concentration, Its Viability and ΔRGB from iPhone and Samsung
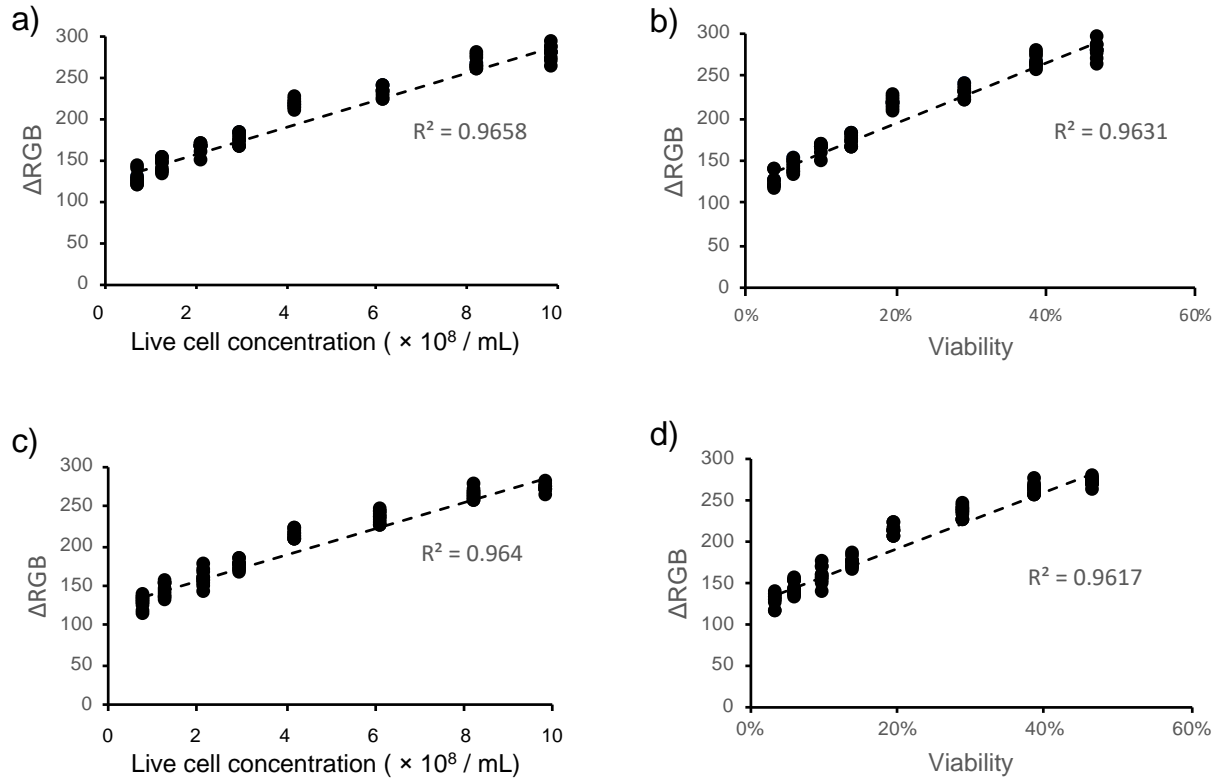
Figure C.3. Calibration curves for colour change on paper-based chips (**ΔRGB**) with respect to live cell concentration and viability measured by automated cell counter counting. (a-b) iPhone 7 Plus (c-d) Samsung Galaxy S7 edge

D. Correlation between Sperm Morphology and DNA Integrity at Single Cell Level

This is an ongoing collaborative project in which I have made significant contributions to cell imaging, image analysis, non-linear regression modeling and deep learning. These aspects are detailed in this chapter.

D.1 Introduction

Male-factor infertility is a global health concern, affecting 25-50 million couples worldwide[2,76,77]. Several semen parameters, including sperm count, morphology, motility, and DNA integrity are utilized to evaluate of sperm fertility potential in male fertility diagnositics[78–80]. Intracytoplasmic sperm injection (ICSI) is an assisted reproductive technology (ART) allowing single sperm deposition into the cytoplasm of the oocyte and making it possible to fertilize an egg with a sperm that may not penetrate the oocyte naturally. Therefore, ICSI enhances fertilization rate more effectively compared to other ARTs, such as traditional *in vitro* fertilization (IVF).[81,82] In ICSI, sperm morphology and motility are two main criteria for the selection of sperms for injection, as evidences show the correlations between these parameters with fertility outcome.[83–85] Nevertheless, the sperms selected by these external traits cannot directly reveal the inherent sperm properties, such as DNA integrity, chromosome abnormality, and chromatin structure. Moreover, other factors, such as female factors, can also alter the fertility and pregnancy outcome and thus make the roles of these parameters inconclusive. DNA integrity is an important predictor of the fertility potential as evidenced by its correlation with fertility and embryonic development.[86–88] There is more evidence to show that to measure male fertility potential, DNA integrity may be a better predictor than the conventional measurements.[89,90] However, the destructive nature of DNA integrity analysis makes it impossible to utilize the sperms after DNA analysis for fertilization.[91] Therefore, noninvasive predictors of sperm DNA integrity are imperative for the sperm selection in ICSI.

Whole sample morphology parameters are correlated with DNA integrity.[92–96] Studies attempting to determine the relationship between sperm morphology and DNA integrity are mainly estimated from groups of cells, while the correlations at the single-cell level remain unclear. Moreover, these correlations are established between single independent variables with DNA integrity.[96–98] Therefore, these studies cannot fully reveal the roles of each parameter and their

interactions that predict sperm DNA integrity. A statistical model of the correlation between sperm morphological parameters and DNA integrity established at a single cell level can provide important criteria for the selection of sperms with low fragmentation and potentially high fertility outcome in ICSI.

In this work, to establish the correlation between single sperm morphology and DNA integrity, I developed a non-fixation method which enables simultaneous analysis sperm morphology and DNA integrity. Several sperm morphological parameters, such as sperm head area size, head width, head length, head aspect ratio, head circularity, and middle piece width, are evaluated from high magnification bright field images. The DNA integrity, quantified as DNA fragmentation index (DFI), of the sperm is determined by the fluorescence intensity of acridine orange bound to DNA.[99] A linear regression model is used to correlate sperm morphological parameters with sperm DFI for individual sperms. Then a deep learning neural network is utilized to generate non-linear regression model using unstructured data. Meanwhile, the method for sperm sample preparation by hyaluronic acid (HA) binding offers a noninvasive way for single sperm selections. Overall, our discoveries bring new focus and guidance for the selection of sperm with high DNA integrity for ICSI.

D.2 Experiments and Methods

D.2.1 HA Functionalization of Glass

A glass cover slide (50mm x 24mm, 0.16mm to 0.19mm thickness) was treated with piranha solution (sulfuric acid: H_2O_2 = 3:1) for 30 min. After treatment, the glass slide was merged into a solution of 10% v/v (3-Aminopropyl)triethoxysilane (APTES) in acetone at 60°C for 30 min, rinsed with acetone, and dried in air. The glass slide was then heated up to 110°C in an oven for 60 min. Hyaluronic acid (HA), N-(3-Dimethylaminopropyl)-N′-ethylcarbodiimide hydrochloride (EDC·HCl), and N-Hydroxysuccinimide (NHS) was dissolved in MES buffer (50 mM, pH= 5) to a final concentration of 5 mg/mL and the solution was stirred for 1hr till HA was completely dissolved. The silanized glass slide was treated with the above solution for 30 min and rinsed with Hanks' Balanced Salt Solution (HBSS) before cell suspension was loaded on to the glass.

## D.2.2 Sperm Sample Preparation and Acridine Orange Stain

The raw semen sample was purified by dispersing 200 µL raw semen into 2mL pure sperm wash medium (Nidacon, PSW-100) and centrifuging at 300g for 5min. The cells were washed with pure sperm wash medium at the same conditions twice. A reservoir made from binding a polydimethylsiloxane (PDMS, Dow Corning) slab cut with a hole (8mm in diameter) to the HA-functionalized glass cover slide was used to contain cells and solutions. Before binding, the surfaces of HA functionalized glass slide and PDMS slab were treated with plasma by a Surface Energy Treater for 30s. A drop of 5µl purified sperm suspension was loaded into the reservoir on the HA functionalized glass cover slide. After the sperms bound to the glass, they were dried in air for 10 min till the water on glass was completely evaporated. Then 20 µL TNE buffer (1mM Tris–HCl, 15mM NaCl, 0.1 mM EDTA, pH= 7.4) was added to the reservoir. The sperms were treated with 40 µL acid-detergent solution (80 mM HCl, 150mM NaCl, 1 mg/mL Triton X-100, pH= 1.2) for 30s before 120 µL 6µg/mL acridine orange (AO) in buffer (37mM citric acid, 126mM $Na_2HPO_4 \cdot 2H_2O$, 150mM NaCl, 1.10mM EDTA-2Na$\cdot 2H_2O$, pH= 6) was added to stain the cells.

D.2.3 Sperm Acrosome Stain

The acrosomes of the air-dried sperms on the HA modified glass were stained with a solution of 40 µg/mL lectin-PNA in HBSS and imaged immediately under the spinning disk confocal microscopy.

D.2.4 Cell Imaging by Spinning Disk Confocal Microscopy

After stained, the sperms were immediately imaged under a spinning disk confocal microscopy (Zeiss, AxioObserverZ1 inverted) under a total magnification of 1000 times. The excitation wavelength of the laser source was 488nm and the bands for the emission filters were set at 500-550nm and 665-715nm for green and red colour, respectively.
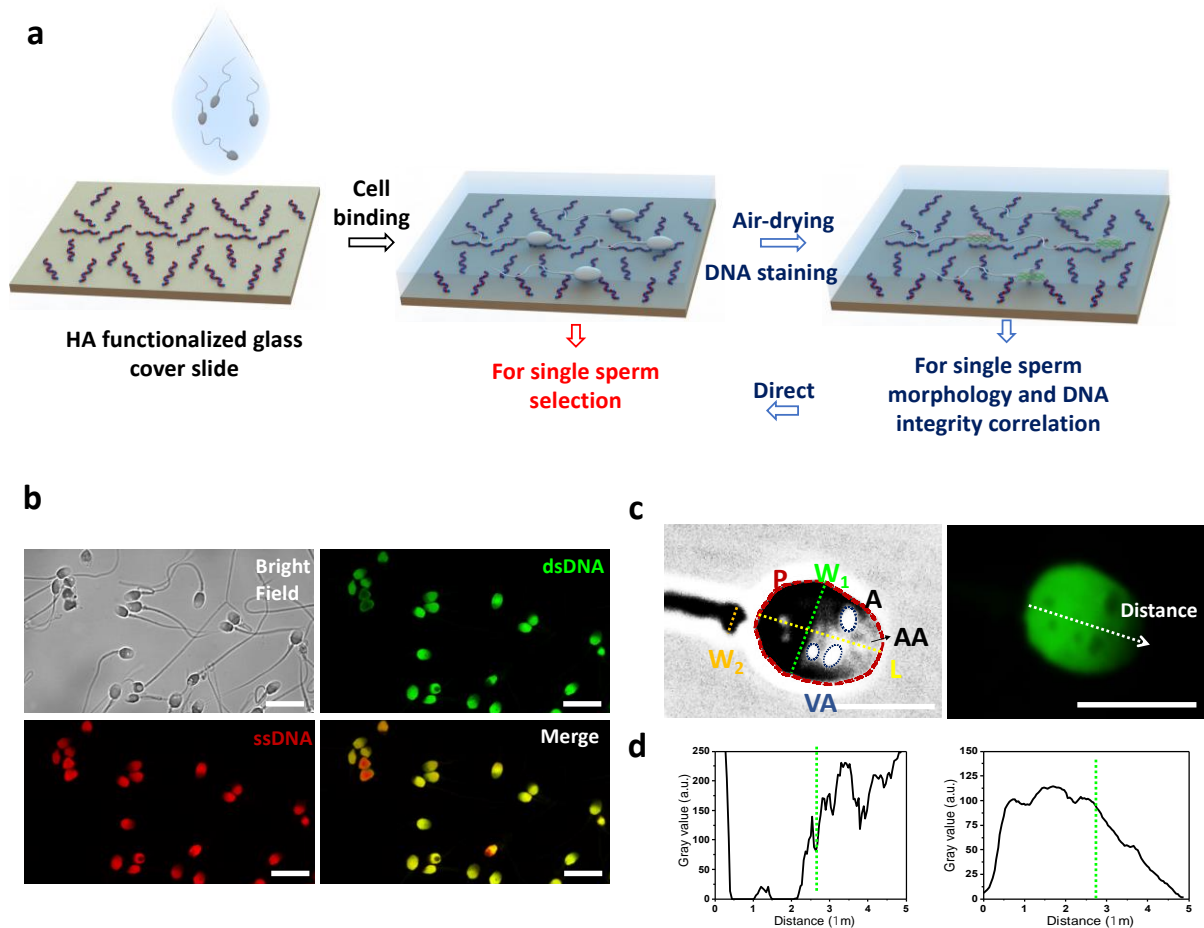
Figure D.4. a) Schematic for a sperm-binding and air-drying method to prepare sperms for morphological and DNA integrity analysis at single cell level. b) Images were obtained from the confocal microscope with 100x magnification in objective lens and 10x magnification in eye lens. Various filters were used to observe stained cells and to calculate DFI for each sperm cell. c) The high contrast image of a single sperm are labeled with independent measurements to identify the sperm morphology. The brighter part within the sperm head is acrosome. d) These line charts, presenting the relationship of grey scale intensity with respect to the distance of along the length of sperm head, are used to find the exact location dividing the acrosome and the rest of sperm head.

D.2.5 Neural Network for Non-linear Regression Approximation

Non-linear regression analysis was implemented in MATLAB software from Mathworks with Neural Network Toolbox. With the data collected from above experiments, the neural network model created to fit our data is a five-layer feed-forward network, with a tan-sigmoid transfer function in each hidden layer and a linear transfer function in the output layer. 20 hidden neurons were used in the first three hidden layers, 10 neurons were used in the fourth hidden layer. The

input dataset was imported as a 6×1004 numerical matrix and the target dataset was imported as a 1×1004 column vector. Then, the input vectors and target vector were randomly divided into three sets of samples as follows: 65% of data was used for training the algorithm, 10% was used to validate the network performance and 25% was presented to independently test the network. The neural network was trained with Levenberg-Marquardt (LM) backpropagation function. The training continued until the validation error failed to decrease for six iterations.

D.3   Results and Discussion

D.3.1   Deep Learning Algorithm for Non-linear Regression Model

Our non-linear regression model uses feedforward neural network in MATLAB neural network toolbox. The default settings of this neural network have one hidden layer of ten tan-sigmoid neurons followed by an output layer with linear neurons (Fig. D.5a). A linear transfer function was chosen for the output layer for fitting the continuous output values. In order to increase the accuracy of prediction and the stability of the model, *newff* network creation function was chosen for the creation and customization of the deep-learning neural network with four hidden layers and one output layer. Multilayer neural network architecture contains 20 nodes for the first three hidden layers and 10 nodes for the fourth hidden layer, shown in Fig. D.5b.

The "Import data" function was used to select and load input matrix and target vectors to the algorithm, shown in Figure D.5c below. The inputs for the algorithm involve six sperm morphology parameters: sample number, head width, head length, head circularity, midpiece width, and ratio. Among these parameters, head circularity and ratio are calculated values based on measurements, in Figure D.5c above, through the following equations:

$$\text{head circularity} = \frac{4\pi * head\ area}{head\ perimeter^2} \tag{4}$$

$$\text{ratio} = \frac{(acrosome\ size - vacuoles\ size)}{head\ area} \tag{5}$$

Then, these input vectors and target vectors were randomly divided into three sets. 65% of them were presented to the network during training stage, when the network was adjusted and trained according to its error. 10% was used to validate network performance and to stop training when performance stopped improving, in the validation stage. The last 25% had no effect on the training and these were provided as a completely independent test of network generalization, in testing stage. This division of data was optimized through various trials. The division of data after optimization was 65% - 10% - 25% for training, validation and testing.

Then, the network was trained and optimized with LM backpropagation function. LM network training function updates weight and bias values according to neurons in the hidden layers and the results from previous iteration. This training function was chosen because it was the fastest backpropagation algorithm in MATLAB toolbox and it was recommended for most of the problems with moderate-sized neural networks. One drawback of this algorithm is that it requires more memory than other algorithms, but this does not affect regular computer performance with the size of our dataset.
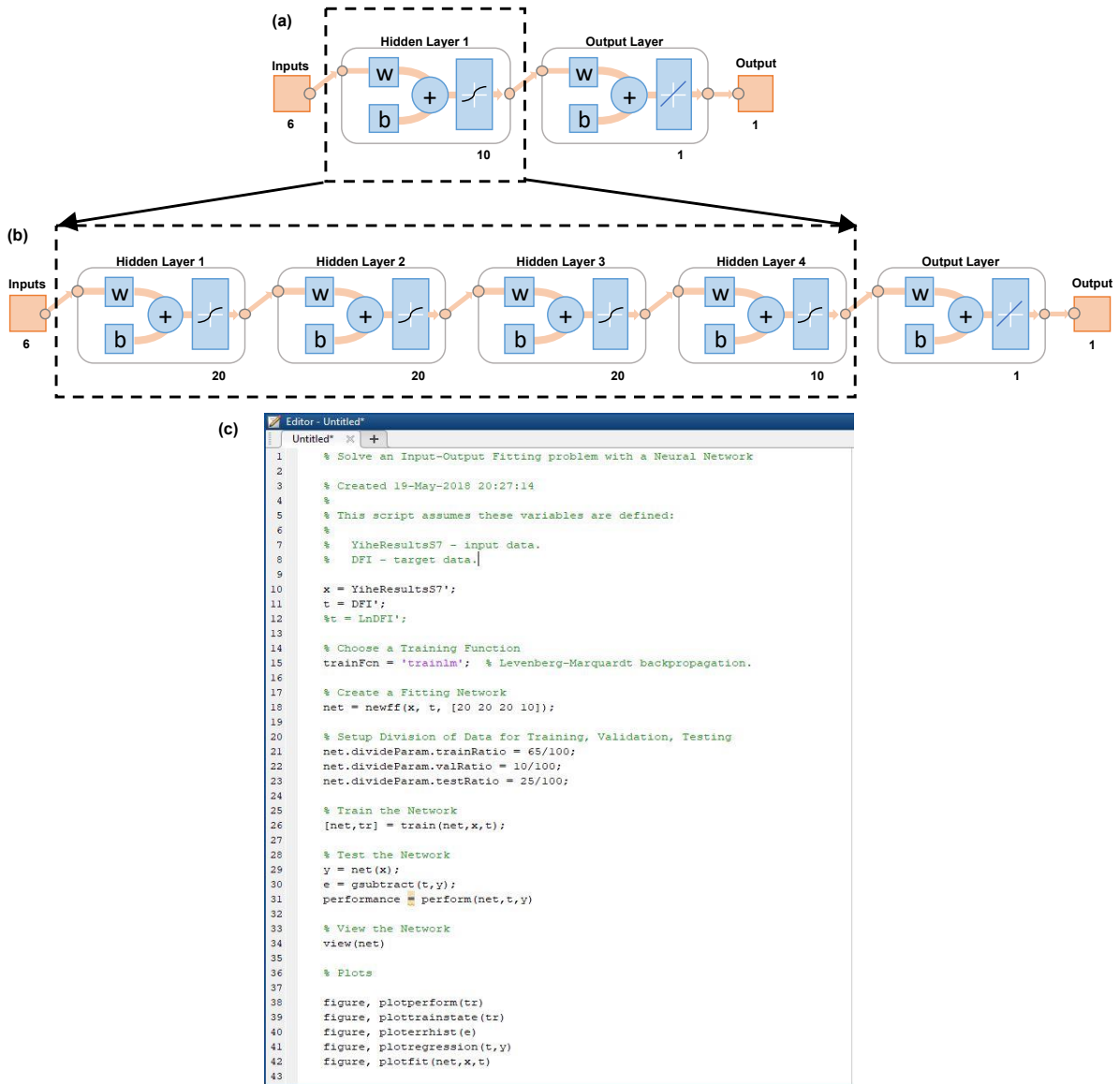
Figure D.5. (a-b) Neural network architecture for input-output fitting. This neural network uses tan-sigmoid function in the hidden layers and a linear function output function in the output layer. W is the weight matrix and b is the bias. (a) This architecture is the default two-layer feed-forward network in MATLAB. (b) This deep-learning neural network is customized to contain five-layer with 20 nodes in each of the first three hidden layers and 10 nodes in the fourth hidden layer. (c) The algorithm for this deep-learning neural network.

D.3.2 Mean Squared Error and Validation

The network trained with LM backpropagation function must use either mean squared error performance function or sum squared error performance function because LM function uses the

Jacobian for calculation. In this work, I chose the mean squared error performance function for validating purposes. The mean squared error is the difference between the predicted value from the neural network and the actual value from loaded target vectors. The lower the value is, the better the network performs, and zero means no error. The validation dataset is used to stop the training of algorithm automatically when the network performance on validation stops improving or remains the same, as indicated by an increase in the means square error of the validation dataset. As shown in Figure D.6, the best validation performance is with mean squared error of 0.001464 at epoch 6, meaning that the validation error decreased until iteration 6 and started to rise thereafter.

During the optimization of the algorithm, both the number of neurons in the hidden layers and the number of the hidden layers were continuously increased until the mean square error of the network performs well. I also tried to avoid overtraining when increasing the number of hidden layers and the number of neurons in the hidden layers. According to Figure D.6, none of training, validation or testing error increased before iteration 6, therefore it proved that overtraining did not occur at iteration 6.
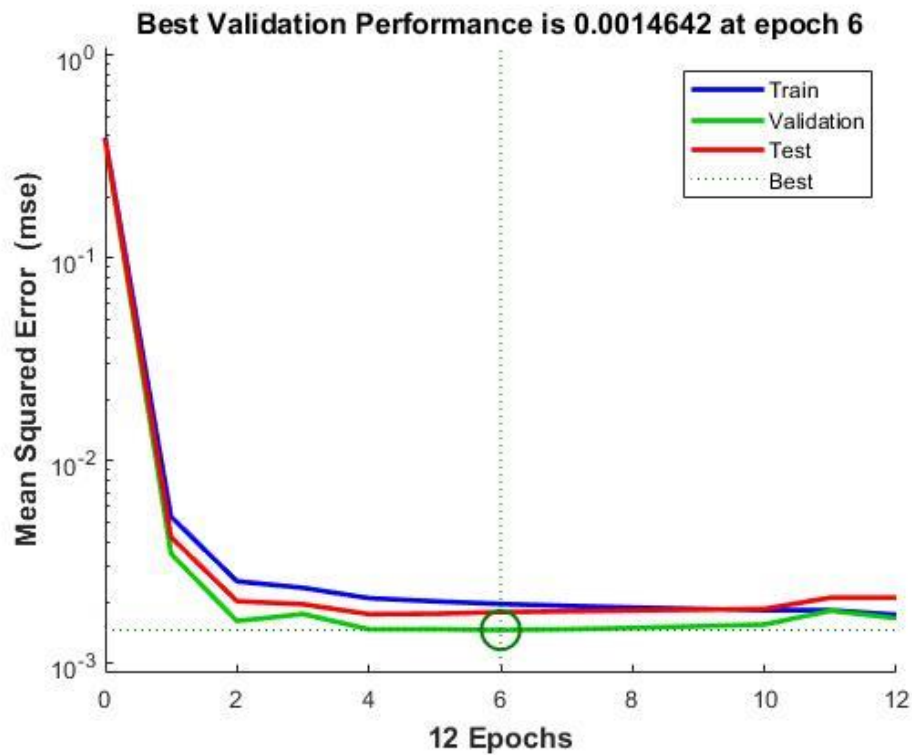

Figure D.6. The best validation performance is with mean squared error of 0.001464 at epoch 6.

D.3.3 Non-linear Regression Analysis

The regression R value measures the correlation between the predicted values and the actual values. When R value equals to 1, it means a close relationship and the predicted values are equal to their corresponding actual values. When R value equals to 0, it indicates a random relationship and there is no correlation between the predicted values and the actual values. The larger the R value is, the better the model predicts. The regression plots illustrate the network predicted values with respect to the actual values for training, validation and testing datasets. When the data fits perfectly, it should fall along the dashed 45 degree line (R=1), where the network predicted values are equal to the actual values. The solid line in each plot represents the best fit linear regression line between outputs and targets. For our problem, the fit is reasonably moderate for all data sets, with R values in both training and testing case of 0.6 above (Fig. D.7). The result shows R values for the training and testing subsets are similar, proving that the algorithm is not overtrained.

Increasing the number of hidden layers and changing the number of nodes in each hidden layer results in significantly larger R values in all training, validating and testing. Multiple hidden layers of neurons with tan-sigmoid transfer function allow the network to learn linear relationships between the six input parameters and the output DFI more accurately and with better stability than the network with single hidden layer. Decreasing the number of data used towards training and validating by 5% each did not affect the accuracy of the model significantly. However, increasing the size of testing data by 10% enhanced its accuracy on network generalization. Overall, optimization on the number of hidden layers and the nodes in each layer has a greater impact on the accuracy of the program than altering the division of data.
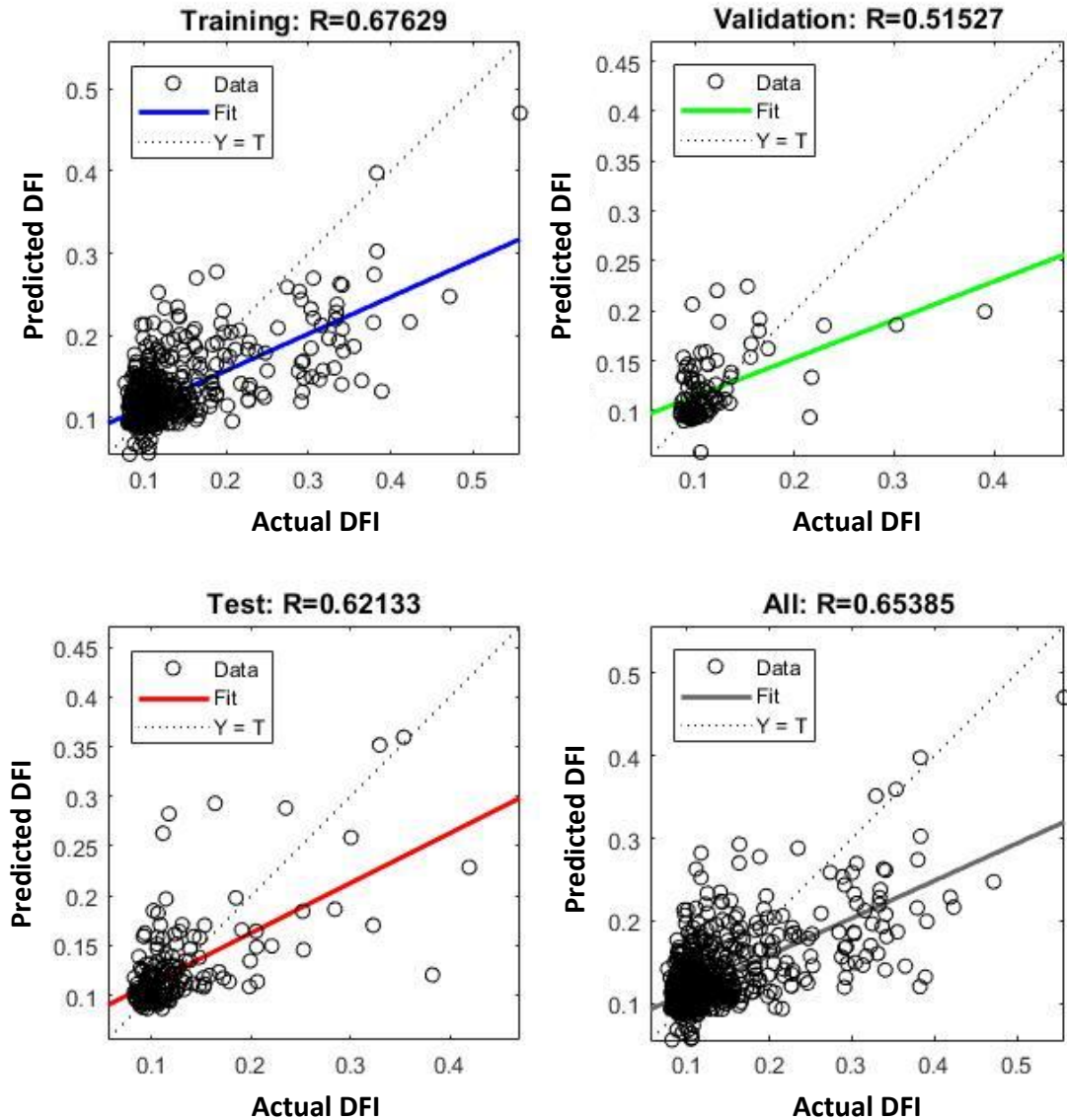
Figure D.7. Non-linear regression approximation of the network predicted values with respect to actual values for training, validation, test and overall datasets.

D.3.4 Histogram Analysis

The error histogram provides additional verification of network performance. The blue bars represent training data, the green bars represent validation data, and the red bars represent testing data. Error here is defined by subtracting the predicted values from the actual values, and these

errors are divided into small subgroups as categories on the horizontal axis. The number/frequency of occurrence is set as the vertical axis. The histogram provides a clear indication of the data points where the fit is significantly worse than the majority of data. As shown in Figure D.8, most errors fall between -0.02634 and 0.01638 and zero error line is within this range. There is a training point with an error of 0.1755 and a validation point with an error of -0.1235 and a testing point with an error of 0.2465. These outliers are also visible on the regression plot. The histogram for this non-linear regression model is relatively symmetrical with thin tails. The error range with the highest frequency of occurrence includes zero error. Thus, the network performance from histogram analysis is positive.
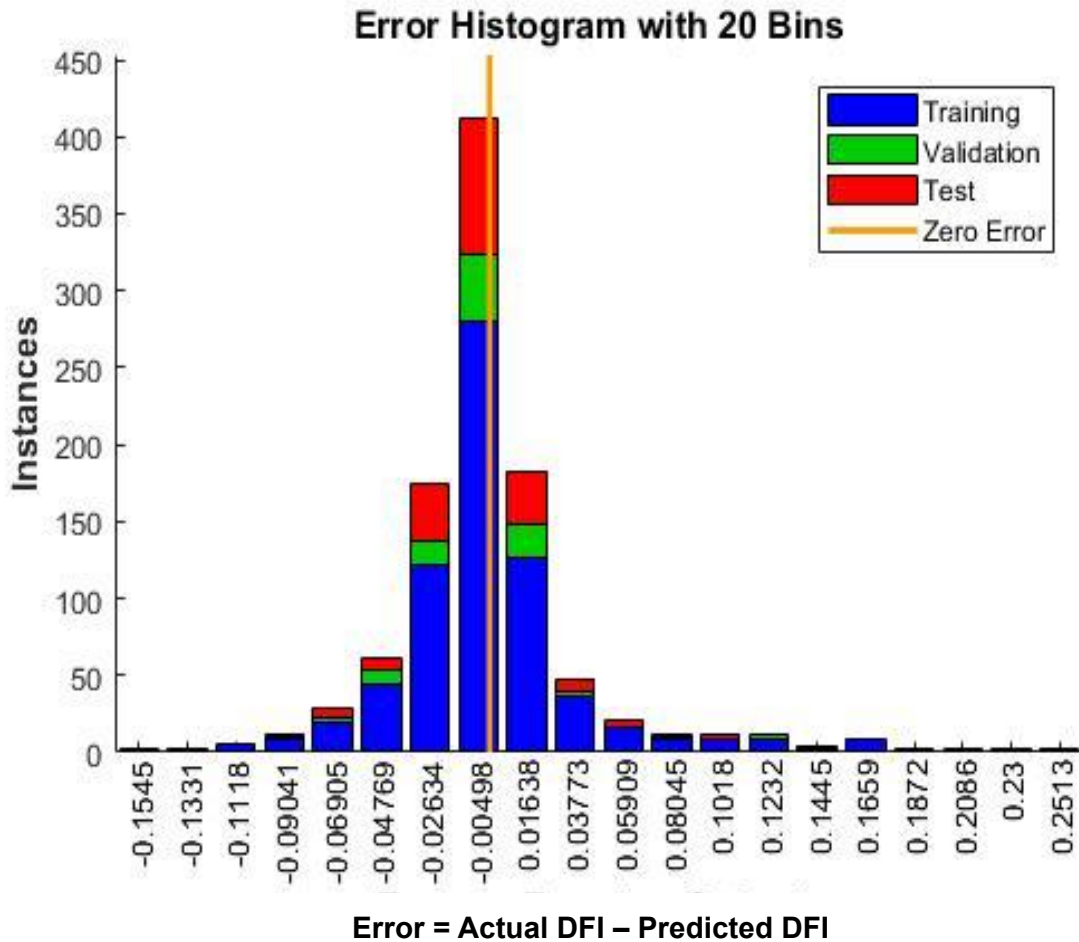
Figure D.8. Histogram analysis of number of instances with respect to the error between actual values and the network predicted values for training, validation and test datasets.

D.4   Conclusion

The deep learning algorithm after the first-step optimization performs moderately well for non-linear regression purpose. The training stops when the mean squared error of the validating dataset increases for six consecutive times to avoid overtraining. Increasing the number of neurons and the number of hidden layers requires more computation, and this has a tendency to overfit the data when the numbers are set too high, but it allows the network to solve more complicated problems with better accuracy, stability and efficiency. I tried to increase the number of neurons to 100 in the one hidden layer case, the performance on the training set was good, but the test set performance was significantly worse. This phenomenon is a result of overfitting, thus adjusting the number of neurons improved our test set results. Comparing to linear regression model, the non-linear regression model is presenting a much stronger correlation. Besides, the error histogram is also symmetrical and thin-tailed, proving that errors between the actual values and the predicted values occurs most frequently within the range with zero error. Overall, this deep learning neural network has shown positive results with the potential of better performance with further optimizations on both accuracy and stability. To ensure the stability of the algorithm, the difference in results from multiple trainings should be obtained and analyzed. This variance is normally due to the change in initial conditions of the network and different random sample division.