Integrating Experimental and Computational Approaches to Optimize 3D Bioprinting of Cancer Cells

by

Dorsa Mohammadrezaei

A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Doctor of Philosophy

in

Applied Mathematics

Waterloo, Ontario, Canada, 2023

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner   Katrin Rohlf
             Professor
             Dep. of Mathematics
             Toronto Metropolitan University

Supervisor   Mohammad Kohandel
             Professor
             Dep. of Applied Mathematics
             University of Waterloo

Internal Member   Sivabal Sivaloganathan
             Professor
             Dep. of Applied Mathematics
             University of Waterloo

Internal Member   Brian Ingalls
             Professor
             Dep. of Applied Mathematics
             University of Waterloo

Internal-external Member   Ehsan Toyserkani
             Professor
             Dep. of Mechanical and Mechatronics Engineering
             University of Waterloo

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners. I understand that my thesis may be made electronically available to the public.

Abstract

A key feature distinguishing 3D bioprinting from other 3D cell culture techniques is its precise control over created structures. This property allows for the high-resolution fabrication of biomimetic structures with controlled structural and mechanical properties such as porosity, permeability, and stiffness. However, for bioprinting to be successful, a comprehensive understanding of cell behavior is essential, yet challenging. This includes the survivability of cells throughout the printing process, their interactions with the printed structures, and their responses to environmental cues after printing. There are numerous variables in bioprinting which influence the cell behavior, so bioprinting quality during and after the procedure. Thus, to achieve desirable results, it is necessary to consider and optimize these influential variables. So far, these optimizations have been accomplished primarily through trial and error and replicating several experiments, a procedure that is not only time-consuming but also costly. This issue motivated the development of computational techniques in the bioprinting process to more precisely predict and elucidate cells' function within 3D printed structures during and after printing.

During printing, we developed predictive machine learning models to determine the effect of different variables such as cell type, bioink formulation, printing settings parameters, and crosslinking condition on cell viability in extrusion-based bioprinting. To do this, we first created a dataset of these parameters for gelatin and alginate-based bioinks and the corresponding cell viability by integrating data obtained in our laboratory and those derived from the literature. Then, we developed regression and classification neural networks to predict cell viability based on these bioprinting variables. Compared to models that have been developed so far, the performance of our models was superior and showed great prediction results. The study further demonstrated

iv

that among the variables investigated in bioprinting, cell type, printing pressure, and crosslinker concentration, respectively, had the most significant impact on the survival of cells.

Additionally, we introduced a new optimization strategy that employs the Bayesian optimization model based on the developed regression neural network to determine the optimal combination of the selected bioprinting parameters for maximizing cell viability and eliminating trial-and-error experiments. In our study, this strategy enabled us to identify the optimal crosslinking parameters, within a specified range, including those not previously explored, resulting in optimum cell viability. Finally, we experimentally validated the optimization model's performance.

After printing, we developed a cellular automata model for the first time to predict and elucidate the post-printing cell behavior within the 3D bioprinted construct. To improve our model, we bioprinted a 3D construct using cell-laden hydrogel and evaluated cellular functions, including viability and proliferation in 11 days. The results showed that our model successfully simulated the 3D bioprinted structure and captured in-vitro observations. The proposed model is beneficial for demonstrating complex cellular systems, including cellular proliferation, movement, cell interactions with the environment (e.g., extracellular microenvironment and neighboring cells), and cell aggregation within the scaffold. We also demonstrated that this computational model could predict post-printing biological functions for different initial cell numbers in bioink and different bioink formulations with gelatin and alginate without replicating several in-vitro measurements.

Taken all together, this thesis introduces novel bioprinting process design strategies by presenting mathematical and computational frameworks for both during and after bioprinting. We believe such frameworks will substantially impact 3D bioprinting's future application and inspire

researchers to further realize how computational methods might be utilized to advance in-vitro 3D

bioprinting research.

Acknowledgments

I would like to express my sincere gratitude to my supervisor, Prof. Mohammad Kohandel, for his unwavering support, guidance, and care throughout my Ph.D. studies. His mentorship has pushed me to think outside the box and strive toward becoming a better scientist.

I extend my appreciation to my committee members, Prof. Sivabal Sivaloganathan and Prof. Brian Ingalls, for their valuable time, guidance, and suggestions from my initial seminar and comprehensive exam to my thesis defense. I am also grateful to Prof. Brian Ingalls for granting me access to his laboratory and facilities during my research.

I would like to thank my internal-external and external examiners, Prof. Ehsan Toyserkani and Professor Katrin Rohlf, for taking the time to read my thesis and attend my thesis defense.

I am grateful to Dr. Homeyra Pourmohammadali and Dr. Nafiseh Moghimi, previous post-doctoral fellows in the Mathematical Medicine lab, for their invaluable assistance and guidance in conducting lab experiments. I also express my gratitude to Dr. Altay Burak Dalan, a visiting scientist, whose experience proved to be valuable and helpful. I acknowledge Dr. Sara Hamis and Lena Podina for their valuable guidance and support throughout the computational studies. Additionally, I would like to thank Johanna De Silva for her help in conducting some experiments in the mathematical medicine lab.

Finally, I would like to offer my profound gratitude to my husband, Hossein Golzar, who has been my rock through all the challenges of my Ph.D. journey. His unwavering love, support, and encouragement have been instrumental in helping me navigate through tough times and preventing me from taking wrong turns.

Dedication

To my lovely family.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

2D Two-Dimensional

3D Three-Dimensional

AI Artificial Intelligence

BSA Bovine Serum Albumin

CAM Calcein AM

CA Cellular Automata

CaCl$_2$ Calcium Chloride

CPCs Cartilage Progenitor Cell

DMEM Dulbecco's Modified Eagle Medium

DMSO Dimethyl sulfoxide

ECM Extracellular matrix

ESCs Epidermal Stem Cells

FBS Fetal Bovine Serum

GelMA Gelatin Methacryloyl

GP-UCB Gaussian process upper confidence bound

ILP Inductive Logic Programming

ML Machine Learning

MLP Multi-Layer Perceptron

MSCs Mesenchymal Stem Cells

MTT 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide

ODEs Ordinary Differential Equations

PBS Phosphate Buffer Saline

PDEs Partial Differential Equations

PI Propidium Iodide

SVM Support Vector Machines

UV Ultraviolet

Chapter 1 Introduction

1.1   Cancer and 3D Cell Cultures

Based on the World Health Organization (WHO) reports, the number of cancers diagnoses worldwide has nearly doubled in the past two decades, from 10 million in 2000 to 19.3 million in 2020. Today, it is reported that 20% of people globally will experience cancer during their lifetime. It is projected that by 2040, there will be over 50% more people diagnosed with cancer than there were in 2020. Breast cancer is regarded as the most prevalent form of cancer in women, other than nonmelanoma skin cancer. In the US, 12% of all women are reported to develop breast cancer during their lives [1]. Breast cancer has proven to be challenging to study in an individual person, and therefore, much effort has been made to develop experimental systems to mimic this process in vitro.

Significant advances have been made in cancer biology during the past years, with the support of a wide range of cancer models, from traditional 2D cell cultures to in vivo models, both of which are essential for the successful discovery of novel therapies in laboratories around the world[2]. Recently, 3D cell culture has been introduced to these traditional cultural methods in order to investigate cancer cells in a natural 3D microenvironment more similar to that found in living organisms [3]. The emergence of 3D cell culture was critical for the advancement of the pharmaceutical industry because, even though many pharmaceuticals passed preclinical trials, the majority of them failed clinical trials. This could be because the 2D culture model of cancer cells does not simulate the 3D in-vivo tumor environment and cellular heterogeneity of a real tumor in humans and consequently have a poor ability to predict the clinical influence of therapies [4]. For instance, in 2D cell culture, most cancer cells are at a similar cell cycle phase, illustrating uniform

growth, while cancer cells in vivo are at various cell cycle phases [5,6]. As a result, 3D cell culture methods are being employed as a prerequisite to preclinical and clinical investigations to better predict the toxicity and efficacy of anti-cancer drugs [7]. In such cell culture, cancer cells get this chance to grow in a 3D microenvironment similar to the extracellular matrix (ECM) of tumors in terms of mechanical and biological features [7–9]. Though several 3D cancer models have made great advancements so far, like spheroids, which attempt to employ ECM or multiple cell lines that are distinctive to tumor microenvironments, they frequently lack spatial and temporal cell configuration and fail to simulate tumor heterogeneity. Therefore, drug studies on such models lack biological relevance or a high level of pharmacogenetic predictability [3]. As a result, there is a need to resolve the gap of mimicking tumor heterogeneity within a tumor model in a controlled manner. One of the latest 3D biofabrication methods is 3D bioprinting which provides this opportunity to customize the fabrication of a 3D in-vivo-mimetic system through predesigned computer-aided designs (CAD) [10]. This method facilitates accurate control over matrix structure resulting in optimal characteristics with modified cellular behavior. In the following sections, we'll go through the specifics of 3D bioprinting in greater depth.

1.2   3D Bioprinting

3D bioprinting is one of the burgeoning techniques widely applied in regenerative medicine and tissue engineering to fabricate complex tissue-mimetic structures simulating human tissues and organs. 3D bioprinting can be defined as a technique involving layer-by-layer deposition of (cells-laden) biomaterials referred to as bioink in a predesigned structural architecture to create human-like organs or tissues. This technology merges cells, biomaterials, and controlled motor systems to develop complicated 3D structures with greater accuracy than any other method used thus far. Therefore, the employment of 3D bioprinting in biomedical applications provides the opportunity

to fabricate bio-mimetic structures with controlled structural and mechanical properties such as porosity, permeability, and stiffness [11–13]. These fabricated 3D structures may be designed and created in CAD utilizing the complicated geometrical information gained from medical images like X-ray images, magnetic resonance imaging (MRI), and micro-computerized tomography scan (μ-CT-scan). Applying 3D bioprinting may be of benefit to the biomedical area, including the potential for increased accuracy, the formation of personalized patient designs, and the economic development of complicated 3D structures rapidly [14,15]. Additionally, 3D bioprinting is a potent technique which may bridge the gap between the inaccurate in-vitro therapeutic response and the in-vivo circumstances for biomedical research, chemotherapeutic responses, and the discovery of cancer treatments [16].

1.2.1. Different Types of 3D Printing Techniques

Among the available 3D printing techniques, inkjet, laser-assisted, and extrusion-based bioprinting are the leading technologies primarily used for the 3D bioprinting of live cells. As can be seen in Figure 1.1. a, in inkjet-based bioprinting, solutions with low viscosity, such as colloidal solutions or cell suspensions, are precisely deposited as droplets at high shear rates [17]. Then, the droplets can be solidified either using chemical or photocrosslinking processes. The volume of bio-ink droplets can be controlled through thermal or acoustic techniques. In thermal inkjet-based bioprinting, droplets are extruded from the printhead using air pressure made by the water vaporization when the bio-ink is heated up at the nozzle. The heater increases the temperature of the bio-ink slightly, which has been illustrated to have a low impact on the viability of cells. Acoustic inkjet printers print through either a piezoelectric actuator or an ultrasonic wave in order to form the droplets. This method does not put the cells at risk of high temperatures and pressures

related to the heated inkjets. However, acoustic damage to the cell membrane might happen if the frequencies are unsuitable [13].

Laser-assisted bioprinting works by using focused laser pulses directed at a laser-absorbing layer on a surface with another layer of bio-ink. The laser absorbing layer sends the laser energy to the bio-ink to produce a high-pressure bubble which leads the bioink onto a collector substrate. This laser provides high structural resolution regardless of biomaterial's viscosities. Also, since laser-assisted bioprinting is a nozzle-less method, no sign of a nozzle-clogging-related problem is associated with this technique. However, the issue with this kind of bioprinting is that each cell type or material should have its own ribbon that should be interchanged during the printing leading to slowing down the overall output [13,18].

An extrusion bioprinter is a robot-controlled machine that propels bio-ink onto a substrate via a nozzle to form a continuous filament. Extrusion printers include a printhead and a bed which can be moved in three dimensions. Most extrusion printers also consist of systems for adjusting temperature, humidity, and light sources used for photo-crosslinking. The primary components of this approach are a syringe, nozzle, and pressure system. The extrusion of bio-ink is frequently controlled using pneumatic or mechanical techniques. Pneumatic systems have a simple drive mechanism; however, one of the main issues with pneumatic systems is that there can be delays in the deposition process due to the time required for gas compression in the system. Mechanical systems such as pistons or screws have more control over the extrusion of the bio-ink but have smaller and more complicated components which are at the risk of breaking under high pressure[13].
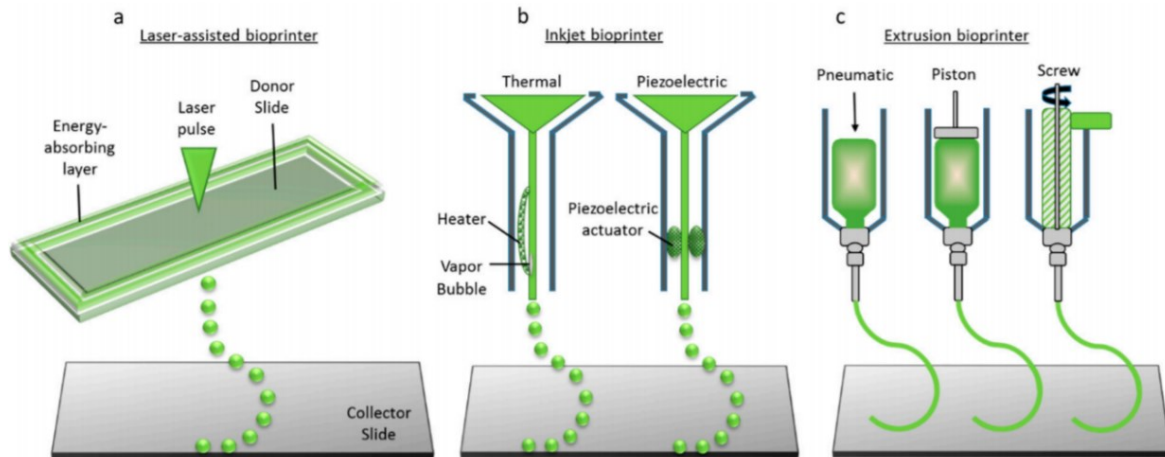
Figure 1.1 Three major bioprinting techniques: (a) Laser-assisted printing; (b) Inkjet-based printing; (c) Extrusion-based printing [19].

Table 1.1 Comparison between different categories of bioprinting technologies [18].

Performance	Extrusion bioprinting	Laser-assisted bioprinting	Inkjet bioprinting
throughput	Medium	Low to medium	High
Spatial resolution	Medium	Medium to high	Medium
Cell-encapsulation control	Medium	Medium to high	Low
Cell survival	40-80%	>95%	>85%
Cell density	High	Medium,10^8 cells/ml	Low, <10^6 cells/ml
Material\ viscosity	30 mPa.s to > 600 kPa.s	1-300 mPa.s	<10 mPa.s
Gelation method	Chemical, Ionic, Enzymatic, Photo Crosslinking, Thermal, PH	Ionic	Ionic, Enzymatic, Photo Crosslinking, Thermal
Gelation speed	Medium	High	High
Print/fabrication speed	High	Low	Medium
Printer cost	Medium	High	Low

Each 3D bioprinting method has its advantages and disadvantages, which are demonstrated in Table 1.1. Inkjet-based bioprinters are likely to be relatively low-priced as they can be produced with commercially available 2D ink-based printers by adding an electrically adjustable stage. However, this type of bioprinters tends to bring cells subject to mechanical and thermal stress, low bio-ink directionality, uncontrolled droplet size, high probability of clogging, and inaccurate cell encapsulation. Laser-assisted bioprinting has shown high resolution, which makes it a potential candidate for printing complicated structures, but they are likely to slow flow rates from the laser-energy absorbing layer. This deposition method is also expensive, making it unsuitable for basic tissue engineering studies. Compared to Inkjet-based and Laser-assisted bioprinting, extrusion-based bioprinting has become the preferred technique in biofabrication due to its simplicity, low cost, and scalability. Additionally, the capability of printing high cell density with high speed is another advantage of this method which provides this opportunity to reach required physiological cell densities in less time, allowing for more rapid cellular and mechanistic studies. Moreover, the multi-material printing capability of extrusion-based printers allows for the fabrication of heterogeneous structures, which is one of the essential aspects of bioprinting. This ability is particularly essential for applications such as tissue engineering and regenerative medicine, where the goal is to create functional tissue models that replicate the complex architecture and composition of native tissues [20–22]. However, this bioprinting method compared to the other two ones, has a lower cell viability. Extrusion-based bioprinting may affect the viability and functionality of cell lines due to the shear stress caused by the extrusion procedure. In the following section, we will discuss the impact of extrusion-based printing parameters on cellular viability and provide a more in-depth review of this issue.

1.2.2. Extrusion-based Bioprinting and Its Effect on Cell Viability

The extrusion-based bioprinting process generally consists of three main phases: pre-processing, processing, and post-processing [23]. In the pre-processing phase, 3D model geometry, number of layers, porosity, and material distribution are designed using CAD and Slicer software. The specified physical design is crucial in determining the scaffold's mechanical qualities. In addition, the bioink components, including appropriate cell types and biomaterials, are selected during this phase. To guarantee good cell viability and resolution, viscosity, biodegradation rate, biocompatibility, and printability should also be considered [24]. During the processing phase, the cell-laden bioink is loaded into the printer's cartridge. The structure is then printed by adjusting all bioprinter-related parameters, including extrusion pressure, cartridge temperature, bed temperature, nozzle size, and extrusion speed. Lastly, in the post-processing phase, before incubating the bioprinted structure, they usually need to be crosslinked. This stage is necessary to guarantee the structure's integrity and proper mechanical strength.

Several studies have examined the impact of various bioprinter-related parameters involved in the three phases on the printability of bioink, but for 3D printing to be successful, it is essential also to evaluate the effects of parameters on cell survival [24]. Cell viability is the main problem with extrusion-based bioprinters, as the extrusion of the bio-ink out of the nozzle can lead to high shear stress that might cause damage to cell membranes. Therefore, optimizing the extrusion-based bioprinting process concerning cellular viability is of great importance [25]. During the processing phase, cells are exposed to a variety of mechanical stresses that have been found to regulate cellular proliferation, differentiation, and migration [26–29]. This stress can be regulated by several variables, such as the pressure system, the size of the nozzle, and the bioink components, including cell types and specified biomaterials.

In the processing step, pressure is the most critical variable to adjust, as greater shear stress and tension is given to cells at higher pressures leading to cell injury. Nair et al.[30] have demonstrated that the impact of printing pressure on cell survival is significantly greater than that of nozzle size. This does not negate the nozzle diameter's influence on cell viability during and after the procedure. Nozzle size is mainly important to determine print resolution; however the smaller the size of the nozzle, the greater the velocity gradient, shear stress, and ultimately cell injury [31].

The choice of bioink components is another influential factor in cell viability. The viscosity of a bioink is the most vital measure of its resistance to extrusion. The greater the viscosity of the bioink, the greater the shear stress at the contact between the nozzle wall and the bioink, which can compromise cell membrane integrity and reduce cell survival [32]. To minimize shear stress-induced cellular viability loss, it is common practice in extrusion-based bioprinting to employ shear-thinning hydrogels, whose viscosities drop as shear strain increases, leading to higher cell survival. The sensitivity of a cell to pressure and shear stress is also dependent on the cell type. As an example: When subjected to the same level of shear stress, cells such as L929 fibroblasts, human mesenchymal stem cells, and HeLa cells experience less cell damage compared to mice embryonic stem cells [33–35]. Hence, optimizing cell viability in extrusion-based bioprinting also requires carefully evaluating the materials and cell types utilized as bioink.

1.2.3. Two-Step and One-Step Bioprinting

In addition to different 3D printing techniques, there are two distinct approaches to 3D bioprinting: one-step and two-step 3D bioprinting. One bioprinting approach is a 'two-step' biofabrication that involves cell seeding onto the prefabricated scaffold. In this method, 3D printing is employed to fabricate precise biocompatible scaffolds with interconnected pore networks for culturing cancer cells with adjustable structural properties and composition [16]. Using

this method results in keeping the shape and the mechanical features of the mimicked engineered tissue, which improve cell adhesion and provide the cell with a structure to proliferate into the 3D functioning tissues [36,37]. However, this method confronts some challenges, such as limited cell density, low accuracy in cell placement, repeatability, and inadequate interconnectivity [16]. Despite all of these challenges, there has been a growing interest in using this method of bioprinting in an effort to simulate the cancer microenvironment [38] due to some reasons: 1) being straightforward and easy to use; 2) providing researchers with geometrically well-controlled scaffolds in a fast and low-cost approach, using different materials such as ceramics, polymers and other cell-stimulating factors that make support for the seeded cells. This can provide high-throughput fabrication of copies of geometrically well-defined complex structures representing 3D complexity through multilayer printing [39–41]. In 2007, Fischbach et al.[42] constructed a 3D model containing oral squamous cell cancer, and they showed highly positive results similar to that of an in vivo model by comparing the results of in vitro and vivo 3D models. Since then, porous scaffolds consisting of different materials such as collagen, silk, and nanoparticles have been established to imitate angiogenesis, tumorigenicity, metastasis, and cancer therapy [43,44].

The one-step bioprinting technique prints a combination of hydrogel and cells as bioink, which is an effective method for producing 3D tumor models. This bioprinting method enables the 3D patterning of human cancer cells on hydrogel with 3D complexity and control across the microenvironment in terms of cell quantities and cell-cell distance [45]. The formation of cancer co-culture models for studying cell-cell interactions and tumor growth can be accelerated through this technique. Despite the fast growth of this bioprinting method, it has some challenges, including biological concerns involving cell viability and long-term functionality post printing. This technique might have apoptotic influences after and during bioprinting. In vivo, human cancer is

a 3D structure including neighboring cells, ECM, and arteries. Recent research on 3D culture has illustrated that cancer cells highly interact with their surrounding habitat during cell proliferation, angiogenesis, metastasis, and therapy [46,47]. Therefore, in-vitro research must simulate cancer cells and their adjacent environment. One-step bioprinting involving cancer cell encapsulation within hydrogel is one of the most common in vitro 3D culture methods, which can precisely imitate in-vivo conditions. Encapsulating cells can provide a valuable tool for studying cellular signal pathways, gene expression, angiogenesis, metastasis, as well as chemotherapy response [48,49]. Recently using multicellular tumor spheroid models has grabbed the researcher's attention to study gene expression and cancer treatment [50–52]. The spheroid culture was studied for the first time approximately 40 years ago, and progressively more refined models have been created to accommodate the rapidly expanding field of cancer treatment [53]. Jiang et al. [10] reported the successful creation of multicellular tumor spheroids (MCTS) of the MDA-MB-231 cell line using an extrusion bioprintable hydrogel bioink consisting of ionically crosslinked alginate and gelatin hydrogels without applying additional stresses. Cells grown within hydrogels not only interact with each other in their neighborhood but also show matrix adhesion [54].

1.3   Bioink

One of the significant elements of the 3D bioprinting technology is bioink which acts as the delivery and supportive materials for post-printing attachment, growth, and proliferation of the cells [55]. For the 3D bioprinting of cells and scaffolds in 3D cancer modeling to achieve the required native tumor microenvironment, including its biological components and microarchitecture, the selection of appropriate bioink materials is essential. Bioink features play a critical role in printing an integrated multilayer structure as well as the formation of 3D in-vivo-mimetic systems. Bioinks must represent gelation characteristics after extruding from the nozzle and show a rigid

construction that should have the sufficient mechanical strength to support the deposition of many layers [56]. Besides, bioinks must exhibit biocompatibility and be suitable microenvironments for live cell accommodation and activations such as cell growth, migration, and proliferation. An ideal bioink requires satisfying two critical physical and physiological properties; good printability to fabricate a successful multilayer structure with enough fidelity and integrity and representing a high cell viability rate post printing [33]. Therefore, it should be selected to satisfy these characteristics.

1.3.1   Materials Used in Bioinks

To choose an appropriate biomaterial for bioprinting a tumor, one must understand the specific tumor ECM contents associated with the target cancer type and its tissue. The biomaterial used in bioink is crucial to the bioprinting of an accurate tumor model because of its specific chemical and physical properties [16]. Biomaterials generally employed for 3D bioprinting applications can be separated into two classes: natural polymer-based and synthetic polymer-based. Natural biomaterials, such as Matrigel, collagen, gelatin, and alginate, consist of naturally derived polymers having ECM-like characteristics and a high level of biocompatibility with cells. In contrast, synthetic biomaterials are constructed of polymers manufactured in the laboratory, such as polylactic acid, polyglycolic acid, and hydroxyapatite. They include characteristics such as matrix rigidity, cell alignment, and other physical factors, which are required for 3D scaffolds or bioprinted structures to last longer than those made from natural biomaterials, which biodegrade at a faster rate [57].

Hydrogel-based bioinks, whether naturally or synthetically based, are a popular choice due to their biocompatibility as well as their ability to offer a well-hydrated microenvironment for cells in 3D structures by retaining a large amount of water [58,59]. Nevertheless, hydrogels by themselves

are frequently excessively soft and unstable to preserve the structure's ideal shape. Crosslinking is the method of forming covalent or noncovalent links between polymer chains in a hydrogel in order to strengthen its rigidity and endurance.

1.3.2 Crosslinking Methods of Hydrogels

Agents for crosslinking may be physical, such as ionic interaction, chemical, such as photo-crosslinking, or a combination of them. Crosslinking not only increases the hydrogel's rigidity but also offers physical stability for the cells, preventing their migration or separation during and after the bioprinting procedure [60].

Ionic interaction is among the most frequently used techniques for the physical crosslinking of 3D bioprinting hydrogels. Gelation is often induced by adding multivalent metal ions to the polymer network. This method of quick crosslinking enables the formation of hydrogels under moderate circumstances at room temperature and physiological pH, resulting in a desirable crosslinking technique. Nonetheless, it has several disadvantages, including mechanical fragility, poor stackability, and the likelihood of cation escape following implantation. The impact of ions on cell viability post printing is an additional problematic aspect. Recent research has demonstrated that the survival and proliferative potential of cells encapsulated in hydrogels vary with the amount of the cationic crosslinker and crosslinking duration. Indeed, more cations might be harmful to cells as the excessive concentration of ions damages the cell membrane, leading to an electrolyte imbalance [61–63]. Additionally, the level of sensitivity a cell exhibits in response to the crosslinking procedure is also dependent on its kind. For example, numerous cell types, such as fibroblasts [64], endothelial cells [65], and chondrocytes [66] were printed within hydrogel-based bioink, and each of them exhibited a unique cell sensitivity to the crosslinking procedure.

Along with the chemical crosslinking process, photocrosslinking occurs through light radiation that comes in the form of ultraviolet (UV), laser, or visible light. UV is the most widely employed of all the many types of photocrosslinking methods. Despite the prevalent application of UV crosslinking in 3D bioprinting, it poses biological concerns. The use of UV radiation for photocrosslinking has the potential to harm the cells of the printed structure. Hence, all crosslinking parameters must be tuned for the desired 3D bioprinted systems. Such variables as irradiation time, intensity, photoinitiator kind, and amount might affect the rigidity and biological activities. Gelatin methacryloyl (GelMA) is among the most broadly applied photocrosslinkable polymers in 3D bioprinting. Reduced crosslinking density results in significantly increased swellability and wider pores when making GelMA hydrogels, creating more favorable conditions for biomolecule incorporation. Unfortunately, the processability of bioinks with low crosslinking density is unsatisfactory [60]. Consequently, the optimal crosslinking conditions must be chosen to achieve a compromise between printability and biological efficiency.

1.4   Bioprinting Optimization and Computational Approaches

In general, for 3D bioprinting to be successful in creating functional and viable tissues or cancerous tumors, the different bioprinting variables need to be carefully optimized. Currently, researchers systematically vary the bioink composition and bioprinting parameters experimentally until they accomplish the desired result in order to optimize these factors. However, this strategy can be time-consuming, costly, and may not yield the best results. In addition, this approach may not completely account for the complex interactions between the various bioprinting factors, resulting in suboptimal outcomes. Therefore, more effective and efficient methods like computational techniques for optimizing the 3D bioprinting process are required.

Computational methods or in-silico approaches have been used to complement in vitro experiments and help resolve some of the limitations of this 3D technique [67–69]. Common in-silico methods consist of machine learning (ML) techniques and mechanistic modeling. ML models can be further classified into subgroups such as deep neural networks, random forests, support vector machines (SVM), and tree classifiers [70]. Mechanistic models include stochastic and agent-based models, discrete models, ordinary differential equations (ODEs), and partial differential equations (PDEs) [71]. These computational tools can help us develop more effective strategies for optimizing bioprinting systems.

In this research study, we provide a comprehensive study on optimizing bioprinting in two stages: during and after printing. At each stage, there are different variables affecting cell viability and cell functioning, which need optimization.

1- During printing, we applied ML algorithms to investigate cell viability based on the 3D bioprinting process parameters like bioink formulation and different adjustable printer parameters whose relationships and synergistic impacts on cell functions are not well-understood (Chapter 3).

2- After printing, we used Cellular Automata (CA) simulation to study the post-printing behavior of a population of cells embedded in a 3D bioprinted structure (Chapter 4).

In the following sections, these computational techniques and their application in bioprinting process are discussed in more detail.

1.4.1 Machine Learning

Currently, ML is one of the most quickly growing techniques which can play a significant role in optimizing a process through smarter and more effective use of required materials and services. ML is a subset of artificial intelligence (AI), which lets systems learn and predict according to

former experience, which is called a training dataset. ML algorithms require learning from training datasets during the training process to update parameters to develop a model. With modified parameters, this algorithm can make predictions for a new dataset that the model has not already observed [70]. Depending on the type of feedback provided to the learning model, ML methods can be mainly categorized into four groups: supervised learning [72], unsupervised learning [73], semi-supervised learning [74], and reinforcement learning [75]. Each ML type will be elaborately described in the next paragraphs.

In supervised learning, which is the most popular ML technique, the algorithm is trained using a labeled dataset. The labels of data points are representatives of a specific class that the data point belongs to. A supervised algorithm then attempts to determine the decision boundaries, which separate different classes of data in order to find the function between the inputs and the labeled outputs. So, it is capable of making a prediction of input features based on "desired" outcomes [76]. Based on whether the labels are discrete or continuous, supervised ML can be subdivided into the classification and regression disciplines, respectively. Whereas classification makes predictions relies on predetermined classes, regression typically makes predictions according to the characteristics of the data. Classification can either be "binary," where only two classes need to be predicted, or "multi-classes," where many more classes need to be predicted [77]. The most common supervised learning algorithms which have been applied in 3D printing technique include k-nearest neighbor [78], support vector machine [79], Naive Bayes [80], neural network, linear regression, radial basis function [81], convolutional neural network [76,82,83], and decision tree [78].

Unsupervised ML is implemented with indistinct and unstructured datasets [84]. Unsupervised ML is more efficient when it has access to more data. Due to the lack of a training and labeled dataset, unsupervised learning algorithms must first, using the self-taught rule, learn and extract

the features of the inputs to identify naturally existing trends within the training data [76]. Unsupervised ML can serve as an initial step in identifying significant characteristics of a dataset before utilizing advanced supervised ML. Indeed, unsupervised ML can be a powerful tool for gaining insights into complex datasets and identifying important patterns or structures that may not be immediately obvious [85]. These unsupervised ML features have several potential applications, such as clustering customers having the same preferences in a business, classifying closely related flower or animal species, and recognizing unusual network connections. K-means clustering [86] and self-organizing maps [78,87] are examples of this learning method applied in 3D printing.

Semi-supervised ML is a mixture of both supervised and unsupervised ML techniques. During training, semi-supervised ML algorithms use a small quantity of labeled data in addition to a great deal of unlabeled data. Semi-supervised machine learning is frequently employed to improve classification results by incorporating unlabeled data when the available labeled data is limited. As the capabilities of ML increase, labeling has moved from being a straightforward procedure to a complex and highly professional task. Hence, semi-supervised ML has attracted attention in numerous applications by processing quite unlabeled data [74].

In reinforcement learning, the input data is between unsupervised and supervised learning, which can just provide an indication as to whether the algorithm is correct or not [70]. This model attempts to understand the relationship between input and output by interacting with their environment. To boost performance, it does not rely on training data like supervised/unsupervised ML does but instead rewards the outcome of actions related to a particular state. In reinforcement ML, training steps are repeated to gain better incentives. It needs information produced from environments, carefully specified reward and punishment criteria, and adequate computing power.

Reinforcement ML can be performed using a Markov decision process that carries out consecutive actions in a predefined setting [74]. These are distinct from traditional supervised ML, which is characterized by a concentration on online performance and does not necessitate familiarity with the Markov decision process. Alternatively, reinforcing ML achieves an equilibrium between investigating previously unexplored areas and utilizing existing knowledge. All that's needed is the present state and action; from there, the reward or penalty can be calculated. This approach is influenced by behavioral psychology, and Q-learning and deep adversarial networks are examples of this learning method [76,88].

1.4.2 Application of ML in 3D Bioprinting

Recently, the application of ML to bioprinting-relevant studies has attracted the researchers' attention, and there is a rise in the number of papers published in this area. The application of ML in 3D bioprinting can decrease fabrication time and cost and increase the quality of bioprinting to satisfy desired purposes. In recent publications, ML methods have been used in the optimization of system and bioink [89,90], dimensional accuracy control [91], defect detection in the process of fabrication [92], and material features prediction [85].

As mentioned above, an ideal bioink, in 3D bioprinting, requires satisfying two critical physical and physiological properties; good printability to fabricate a successful multilayer structure with enough shape integrity and fidelity and a high cell viability rate post printing [33]. Based on the literature, these two criteria depend on the physicochemical properties of bioink, 3D printer parameters, and post-printing treatments. Although there are many studies that have investigated the optimization of bioink printability and viability by doing rheological measurements [93–95], mathematical simulation on bioink flow [95], and image analysis [96], researchers still require to do

17

trial-and-error experiments in order to bioprint a 3D structure with good characteristics. Hence, there is an increasing need to find a smart and fast technique for optimizing an ideal bioink.

Several studies have tried to develop different ML-based techniques to optimize the printability of bioinks. For instance, Lee et al. [97] utilized inductive logic programming (ILP) and regression analysis to create a printable bioink. They used different forms of natural collagen and atelocollagen in combination with hyaluronic acid and fibrin to create a variety of bioink formulations. Using ML, they could find a correlation between the mechanical characteristics of bioink and its printability; more specifically, a higher elastic modulus resulted in better shape integrity and that extrusion could happen below the threshold yield stress. From this correlation, they used multiple regression analysis to develop many formulations of bioinks that maintained a high degree of shape accuracy while printing. Another study used a Bayesian optimization algorithm to fine-tune the printability of bioinks consisting of GelMA and hyaluronic acid methacrylate for extrusion-based bioprinting. This study optimized bioink in terms of layer stacking and filament morphology by changing GelMA composition, ink cartridge temperature, printing pressure, speed, and platform temperature [55]. Similarly, using ML methods, Bone et al. [98] attempted to systematically maximize the shape fidelity of printed alginate-based scaffold based on print speed, flow rate, alginate concentration, and nozzle size.

Cell viability is one of the most critical factors in 3D bioprinting. The main objective of 3D bioprinting is to create functional tissues, organs or tumours, which require living cells to grow, proliferate or differentiate into specific tissue types. The viability of cells in 3D bioprinting determines the success of the printing process and the functionality of the printed construct. Hence, it is important to optimize the printing conditions to maintain high cell viability and ensure the success of the 3D bioprinting process. However, despite the importance of cell viability, the effect

of this factor on bioprinting success is often undermined. To the best of our knowledge, only a few studies have been working on applying ML algorithms to study the effect of 3D bio-printing parameters on cell viability. Therefore, the relationship between these parameters with cell survival rate is not yet well-defined.

For instance, Xu et al. [99], created an ensemble learning model to predict cell survival in stereolithography-based bioprinting. The model achieved a high level of accuracy, as evidenced by R^2 of 0.953 when tested on 10% of the dataset. This score indicates that the model is well-suited for predicting the viability of cells when given new parameter combinations. In a recent study, Tian et al. [100] also studied the capability of ML regression and classification approaches in predicting the survival of cells and the printability of cell-containing alginate and gelatin-based bioinks. To do this, they collected a dataset on bioink material content, solvent utilized, crosslinking data, printing settings, survival rate, and printability results from various bioprinting laboratories published in the last 13 years. This research compared three regression models (support vector regression, linear regression, and random forest regression) and three classification models (random forest classification, logistic regression classification, and support vector machines). The random forest models yielded the highest R^2 in regression and prediction accuracy in classification, compared to other models. Additionally, based on feature importance evaluation through decision, Printing pressure, particular material concentration, solvent selection, nozzle size, and printing temperatures were found to have the most influence on cell survival predictions. Results taken from this study illustrate how ML models could be useful for experimental design in bioprinting.

Despite these studies, research analyzing the characteristics associated with cell viability and proliferation in bioprinting is inadequate due to the lack of a comprehensive understanding of how

cell viability and proliferation are affected in the bioprinting process. Hence, the next stage in the use of ML in bioprinting appears to be integrating cellular parameters and predicting cellular performance in bioprinted structures.

Despite the availability of a few studies on the subject, our understanding of the bioprinting factors that affect cell viability and our ability to optimize the bioprinting process remain limited. To leverage the full potential of ML in bioprinting and advance the field towards more complex and functional bioprinted tumor models, tissues, and organs, more efforts should focus on integrating bioprinting process with ML approaches and developing models to predict cellular performance.

1.4.3 Cellular Automata Modeling

CA is a type of mathematical model, under the larger framework of agent-based modeling, used to study the behavior of complex systems, particularly those that evolve over time. In CA, the system is spatially and temporally discretized. In addition, the system is represented as a grid of cells or elements, each of which has a state and a neighborhood. The state of each cell can change over time and be updated at discrete time intervals in accordance with a set of rules that specify how the state of a cell is dependent on its present state and the state of its surrounding elements. This allows the system to evolve in a dynamic and frequently unexpected manner over time [101]. CAs have been used to simulate a variety of phenomena, including physical systems, biological systems, and social systems, and they have applications in domains including physics, biology, computer science, and economics. Neighborhoods refer to the set of elements that interact with a certain element. Von Neumann and Moore neighborhoods are the two most common neighborhoods in CA. A cell's first-order von Neumann neighborhood is composed of the four elements directly adjacent to it, the cells sharing an edge with the cell of interest [102]. A cell's first-

order Moore neighborhood includes the eight cells immediately adjacent to it. First-order and second-order Von Neumann and Moore neighborhoods are illustrated in Figure 1.2. More detailed aspects of CA modeling are described in Chapter 4.
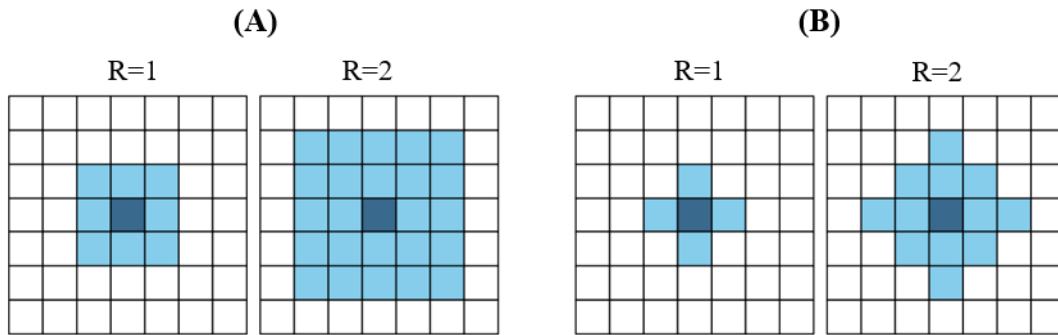
Figure 3 (A): First (R=1) and second (R=2) order Moore neighborhood; (B): First and second order Von Neumann neighborhood [103].

1.4.4 Application of Cellular Automaton Model in 3D Bioprinting

As discussed previously, cell culture experiments are experiencing a switch from a traditional 2D system to a 3D culture, which provides researchers with more realistic representations of in vivo cell behavior. Indeed, these 3D methods incorporate important aspects of cellular habitat, including a non-uniform microenvironment, complex interactions of cells with their neighbor cells and with local ECM, as well as more complicated diffusion processes of nutrients and oxygen. Thus, these methods can be used to better represent insights into cell growth mechanisms and provide a closer prediction of the real system. However, these complexities make the experimental design as well as data collection more complicated, so there are several issues that require to be addressed in these 3D culture methods. Computational modeling can be considered a good complement to the in-vitro experiments and can assist in addressing some of the limitations of 3D method settings [68,104]. For example, in 3D bioprinting method, simulation of cellular behavior cultured within the hydrogel helps by providing researchers with more insights into cell activities during or after the bioprinting. In addition, it allows them to predict cellular behavior in this process

21

and optimize the cellular microenvironment to better simulate the in-vivo ECM, saving more money and time.

CA modeling has been widely used in recent years to investigate different cancer cell mechanisms based on different static automaton rules [105]. For instance, Cui et al. [54] used CA modeling to simulate and compare cancer cell spheroids' formation in two environments; microgel scaffold and suspension. In their model, cells behave based on the rules for movement, proliferation, and death and could simulate the experimental results of size distribution and spheroid formation rate in both environments. The CA model suggested an important effect of initial cell density in the formation of spheroids and their features, such that more initial cells caused larger spheroids and more variability in their size, which was shown to be more pronounced for cells within the microgel scaffold compared to the suspension cultures. Additionally, in another recent study, Cortesi et al. [106] introduced a ScAffoLd SimulAtor (SALSA) as a computational tool to study the cellular behaviors of two types of breast cancer cells (MCF-7 and MDA-MB-231) cultured in a 3D system. SALSA was used to investigate both the cellular (such as position, status, and behavior) and the local microenvironmental (such as concentration of glucose and oxygen, as well as stiffness of the scaffold) of relevant properties. The results illustrated that combining experiments and computational modeling could modify the study of complicated cell culture within the 3D scaffold, representing more complete data and tools for researchers to do experimental screening and optimization. The promising results of these studies motivated us to apply the CA model in 3D bioprinting method for the first time to study the behaviors of a population of arbitrary cancer cells encapsulated and grown in a 3D structure.

1.5   Motivation and Scope of the Thesis

As discussed above, extrusion-based bioprinting is a 3D bioprinting technique that has made significant progress in recent years; however, it has its particular challenges. Various variables are involved in the bioprinting process that can influence printing quality and cell behavior during and post printing. Hence, depending on the intended purpose of bioprinting, it is essential to carefully consider and optimize different variables to ensure successful outcomes. This optimization can be categorized into two stages: during and after printing.

During the printing, variables such as material choice, cell type, printing settings, and post-printing treatments must be optimized for optimal cell viability and functionality. Optimization in this stage is typically accomplished by replicating numerous trial-and-error experiments in research laboratories, a tedious and costly procedure. To overcome this challenge, in this study, we developed Neural Networks to predict cell viability based on a variety of effective characteristics for both classification and regression purposes. In addition, we utilized a Bayesian optimization model to inversely determine the optimum global value of the selected bioprinting variables that would result in the highest cell viability during the printing process. To achieve this, we combined the data collected in our laboratory under various experimental conditions with acquired data from previously published bioprinting research to create a dataset to train and test our neural network models.

After the printing, cells' behavior, including proliferation and viability, is influenced by a variety of factors, such as initial cellular density in bioink, cell-cell interactions, cell-microenvironment interactions, and source diffusion. As the cells are embedded in a 3D structure, it is challenging to measure these factors experimentally, particularly over an extended period of time. Therefore, it is difficult to obtain insight into the cell dynamics within a 3D bioprinted

23

structure without costly and time-consuming experiments. To address this issue, we developed a CA model for the first time to simulate post-printing cell behavior within the 3D bioprinted construct. To improve our model, we bioprinted a 3D construct using cell-laden hydrogel bioink and evaluated cellular functions, including viability and proliferation in 11 days. Using this model, researchers can simulate the dynamics of cells within a 3D bioprinted structure, study the effect of various variables on cellular functions, and optimize these variables to accomplish the desired cellular behavior more efficiently and cost-effectively. This thesis is structured as follows:

Chapter 2 introduces and discusses some of the experimental research methods and computational and mathematical methodologies applied throughout this study. In **Chapter 3**, we present ML algorithms for optimizing the bioprinting process during printing. In **Chapter 4**, we introduce the integrated CA model with the bioprinted tumor model to predict and elucidate the post-printing behavior of 3D-printed cancer cells in hydrogel structures. In **Chapter 5,** we present the conclusion of the findings drawn from this study. This chapter provides a comprehensive overview of this study's contributions to the field of bioprinting and recommends prospects for future research in this area.

This thesis partly adapts the publication "Mohammadrezaei, D. et al. Predicting and elucidating the post-printing behavior of 3D printed cancer cells in hydrogel structures by integrating in-vitro and in-silico experiments. Sci. Rep. 13, 1–13 (2023)" [107].

Chapter 2 Method Background

2.1 Introduction

This chapter outlines the methodology used for this thesis, which is divided into two discrete sections: in-vitro studies and computational studies. The in-vitro studies consist of cellular assays and bioink characterization, which provide a good understanding of the cellular behavior and material properties in the bioprinting procedure. The computational techniques utilized in this study include the development of a neural network to predict cell survival during the bioprinting based on various variables in this process. In order to prevent overfitting and assure the reliability of our models, we also employed regularization and cross-validation techniques. We also used Bayesian optimization to optimize the system and determine the global optimal parameter combinations for the highest viability. A comprehensive explanation of these techniques is provided in this chapter. In addition, we elucidated and predicted the cellular behavior within the 3D-printed structure after printing using CA method, whose methodology and specifics are described in Chapter 4 and Appendix 2.

2.2 In-vitro Studies

2.2.1 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide assay

The colorimetric measurement of 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) was initially invented as a fast technique for measuring the growth and viability of mammalian lymphoma cells [108]. The respiratory chain and other electron transportation networks decrease MTT as well as other tetrazolium compounds inside the cell, producing non-water-soluble violet formazan crystals [109]. A spectrophotometric assessment of the concentration of these crystals provides an approximation of the sample's mitochondrial content and, therefore, its cell

viability. These properties can be used in cytotoxicity and cell proliferation experiments, which are common methodologies in the fields of immunology, toxicology, and cell biology [110].

2.2.2 Ki67 Immunohistochemical Assay

Ki67 is a type of protein that is produced in every mammalian cell throughout the growth cell cycle but is greatly suppressed in the G0 phase. Because of this property, Ki67 has become an important biomarker for evaluating proliferation in clinical samples. The Ki-67 proliferative index is a well-characterized indicator of breast cancer progression and is the proportion of proliferating cells relative to the overall number of cells. The measurement of the Ki-67 proliferative index involves the use of immunohistochemistry, a technique that uses specific antibodies to detect and visualize the Ki-67 protein. In this technique, the sample is treated with the primary Ki-67 antibody, which binds specifically to the Ki-67 protein present in the proliferating cells. Afterward, a secondary antibody, labeled with a visible marker, is introduced to bind to the primary antibody which highlights the locations of the Ki-67 protein in cells, allowing for visualization under a microscope [111]. To calculate the Ki-67 proliferative index, the number of ki-67 positive cells is counted, and this count is divided by the total number of cells present in the same sample. The result is then expressed as a percentage, representing the proportion of proliferating cells.

2.2.3 Calcein AM and Propidium Iodide

Calcein AM (CAM) is a non-fluorescent dye that, inside living cells, is rapidly degraded by esterases, converting into the bright green fluorescing calcein [112]. Therefore, cells lacking these esterases will not be able to convert CAM to calcein, preventing them from producing fluorescing green. This facilitates the rapid and straightforward analysis of samples for the presence of metabolically active cells.

Propidium iodide (PI) is a red fluorescent dye that cannot enter living cells due to their intact membranes but can enter non-living cells with damaged membranes, such as dead or dying cells. PI interacts with DNA and emits red fluorescence when excited by specific light wavelengths [113].

Therefore, CAM and PI can be used together to assess a sample's viability by counting the proportion of metabolically active (green fluorescent) cells to those that are not viable (red fluorescent) when observed by fluorescence microscopy or flow cytometry.

2.3   Rheological Properties of Bioinks

In bioprinting, the rheological properties of bioinks determine their printability and ability to retain their morphology after printing. In general, there are two categories of fluid rheology: Newtonian and non-Newtonian fluids (Figure 2.1). In contrast to non-Newtonian fluids, Newtonian fluids' viscosity does not correlate with shear rate. Non-Newtonian fluids do not follow Newton's principle of viscosity and can behave differently when subjected to shear strain. The features of non-Newtonian fluids allow for three distinct categories: Bingham, shear-thinning, and shear-thickening. For instance, in Bingham polymers, like toothpaste, shear stress and shear rate are linearly related, however, yield stress is required to initiate its flow. In contrast, shear-thinning fluids exhibit a decrease in viscosity as the shear rate increases [114].

Typically, bioinks are non-Newtonian fluids, with many displaying shear-thinning behavior. This enables them to be extruded through small nozzles during the 3D printing process with minimal impact on the cells while retaining their structure without collapsing after extrusion.
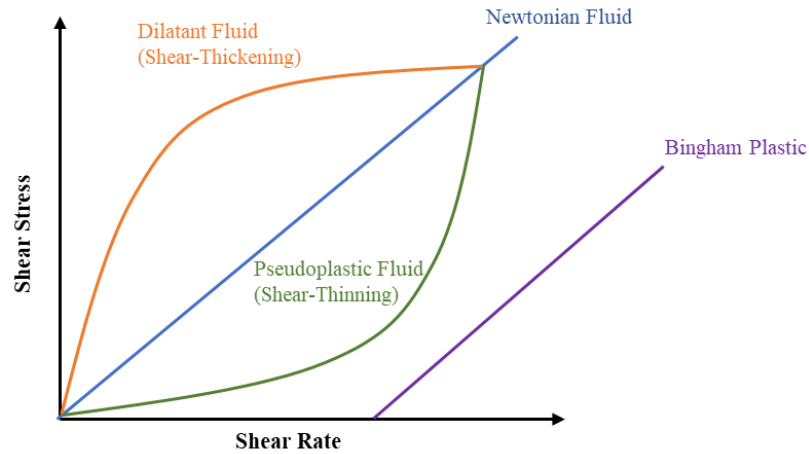
Figure 2.1 Relationship between shear stress and shear rate for various types of fluids.

2.4   Computational Studies

2.4.1   Multi-layer Perceptron Neural Network

One of the purposes of the proposed work is the classification and regression application of multi-layer perceptron (MLP) neural networks. An MLP neural network is a feedforward neural network that employs the backpropagation learning approach utilizing gradient descent, in order to optimize its weights. In MLP neural network, the input layer of nodes is followed by one or more layers of hidden layers of neurons that process the data and perform iterations and an output layer that estimates the output (Figure 2.2) [115].
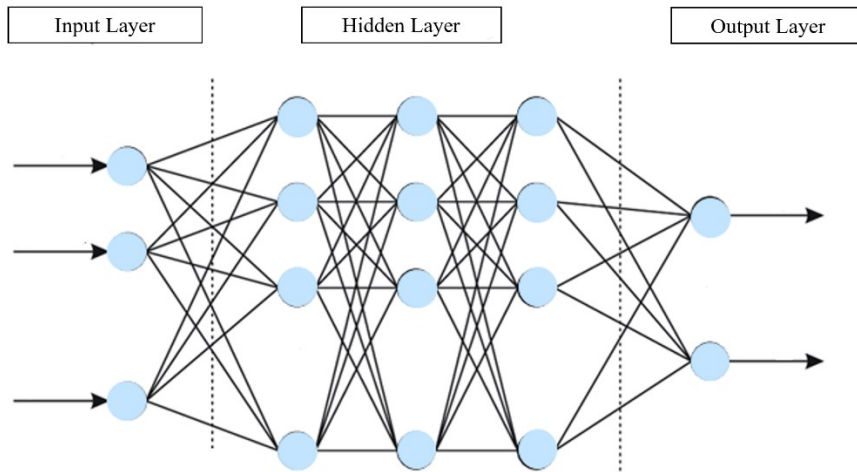
Figure 2.2 Artificial neural network structure.

Each node in a layer is linked to every node in the preceding layer, and the links are associated with weights that are adapted during training [116]. The weights in input neurons are used to generate the output by applying a nonlinear activation function. The perceptron computes a linear combination of multiple real-valued inputs to produce an output neuron. This computation can be expressed using the following equation [117]:

$$y = \varphi(\sum_{i=1}^{n} w_i x_i + b) = \varphi(w^T x + b)$$

Here x, w, φ, n and b denote the input vector, weight, activation function, number of input neurones and bias, respectively. Numerous nonlinear activation functions include ReLU, Sigmoid, Tanh, softmax, and others. In most cases, the activation function used on the output layers will differ from those used on the hidden layers. The selection of the activation function relies on the expected output value range and the problem kind. For example, in a binary classifier, typically, there is only one neuron in the output layer, and its value is between 0 and 1 to indicate the level of certainty that the input vector corresponds to one of two classes. Likewise, in the multi-class classifier, the number of output neurons is equal to the number of classes, and the output neurons reflect the probabilities of an input vector belonging to a specific class. As a result, a distinct

activation function is employed, as the output neurons must add up to 1. Therefore, the "softmax" function is the most frequently employed activation function for the output layer neurons. For this reason, "softmax" is the most popular activation function for neurons in the output layer.

Notationally, if $\vec{a}^{(0)}$ and $\vec{a}^{(l)}$ denote the input vector, and lth layer of $\vec{x}$, and $\varphi^{(l)}$ represents the lth layer's activation function, then the lth layer's value of the network is calculated as

$$a(l) = \varphi^{(l)}\left(W^{(l)} a^{(l-1)} + b^{(l)} \right)$$

Meaning that each layer's value is calculated by using the layer's weights and biases in conjunction with the output of the preceding layer and then putting this information into the activation function. So, finding the neural network's output for a given input vector is quite straightforward, given a set of weights and biases. $W^{(l)}$ is a weight matrix in lth layer, with the shape (output neurons, input neurons), where output neurons represent the number of hidden units or number of neurons in the following layer, and input neurons represent the number of neurons in the input layer or previous layer that fits $a^{(l-1)}$. The challenge of ML, thus, is to establish appropriate values for these weights and biases. Backpropagation comes into action at this point. Backpropagation is a frequently employed training strategy for neural networks that relies on a loss function and optimization algorithm. Backpropagation commonly selects starting weights and biases from Gaussian data [118] and trains them using stochastic gradient descent optimization.

There are different loss functions that can be implemented in neural network training. The mean squared error (MSE) function is among the most often employed loss functions for regression tasks. It determines the average squared error between predicted and actual output. The MSE loss function is calculated as:

$$L(y, \hat{y}) = \frac{1}{n} \sum_{i=1}^{n} \left(y_i - \hat{y}_i \right)^2$$

Where n represents the sample size, y_i represents the actual output, and $\hat{y}_i$ represents the predicted output. Typically, the categorical cross-entropy function is applied to classification tasks. It quantifies the difference between predicted and actual class probabilities. The definition of the categorical cross-entropy loss function is:

$$L(y,\hat{y}) = -\sum_{i=1}^{n} y_i \log(\hat{y}_i)$$

It is important to note that the categorical-cross entropy loss function is only used here if the activation function of the last layer is the "softmax" function that assures a loss function with a positive, real-valued sign.

To reduce this loss function, numerous optimization procedures can be utilized, although stochastic gradient descent variations are the most common. The purpose of the training procedure is to calculate the gradient of the loss function in regard to the neural network's weights and biases. Automatic differentiation is utilized to calculate a numerical gradient for each training sample. Then for a specific learning rate (η), the weights in the neural network are adjusted using the backpropagation algorithm. The algorithm consists of the following steps:

1- Pass the input ($\vec{x}$) through the neural network.
2- Calculate the error δ for the output layer using the following equation:

$\delta = \nabla L \odot \varphi^{(l)'}(z^{(l)})$

Where $\odot$ denotes element-wise multiplication, L is cost function, $z^{(l)} = w^{(l)} a^{(l-1)} + b^{(l)}, a^{(l)} = \varphi(z^{(l)})$, φ is the activation function, and ∇L is the gradient of the loss function with regard to the output.

3- Backpropagate the error δ from the output to the input layer by computing the error for each layer using the equation below:

$\delta^{(l)} = (w^{(l+1)\,T} \delta^{(l+1)}) \odot \varphi^{(l)'}(z^{(l)})$

Where $w^{(l+1)}$ is the weight matrix of the next layer, $\delta^{(l+1)}$ is the error of the next layer, and $z^{(l)}$ is the weighted sum of inputs for the current layer.

4- Compute the cost function's gradient with regard to the weights and biases using the following equations:

$$\frac{\partial L}{\partial w_{ij}^{(l)}} = a_j^{(l-1)} \delta_i^{(l)}$$

$$\frac{\partial L}{\partial b_i^{(l)}} = \delta_i^{(l)}$$

Where L is the cost function, $a^{(l-1)}$ is the activation of the previous layer, and $\delta_i^{(l)}$ is the error for the current layer.

5- Update the weights and biases using the following equations:

$$w^{(l)} = w^{(l)} - \eta \left(\frac{\partial L}{\partial w^{(l)}} \right)$$

$$b^{(l)} = b^{(l)} - \eta \left(\frac{\partial L}{\partial b^{(l)}} \right)$$

Where η is the learning rate.

These steps are repeated for each training example in the dataset for a certain number of epochs until the weights and biases converge to a minimum of the cost function.

The stochastic gradient descent technique calculates gradients by randomly dividing training data into batches of a predetermined size. Using backpropagation with fixed weights, gradients are computed for each batch. Following this, the gradient technique is used to fine-tune the weights, but this time the learning rate is $\frac{\eta}{b_s}$. This procedure continues until all training data is processed, which represents one learning epoch. Generally, training continues for a number of epochs until the loss function stops diminishing or reaches a minimum amount. The Adam optimizer is a popular form of stochastic gradient descent that adjusts the learning rate for each parameter based on the mean and standard deviation of recent gradients for the weights [119].

2.4.2 Regularization

Overfitting is a common problem in ML where a model becomes overly complex and starts to memorization of the training data, leading to poor generalization accuracy on new, unseen data.

Overfitting has three major causes: First: noise in the training data, which can arise when the training data is too small or contains too many irrelevant features. This circumstance makes it likely that the noises will be learned and subsequently act as the foundation for predictions. Thus, a properly operating algorithm must be capable of distinguishing between representative data and noise [120]. Second: the model's complexity can result in a bias–variance trade-off. When there are too many inputs into a model, it can become more accurate on average but less consistent across various datasets. Third: Multiple comparison techniques, which are prevalent in induction and AI algorithms [121]. Throughout these procedures, we regularly analyze various evaluation metric scores and pick the one with the highest score. Nevertheless, this can lead to the selection of items that do not increase prediction accuracy.

Regularization is one of the techniques used to prevent overfitting by incorporating a penalty factor to the model's objective function that is attempting to optimize. This factor motivates the model to learn simpler and more generalizable patterns. There are several common regularization techniques used in ML, including L1 and L2 regularization, as well as dropout regularization [121].

L1 regularization, which is also referred to as Lasso regularization, introduces a penalty to the objective function that is proportional to the absolute value of the model's weights. This penalty term encourages the model to learn sparse feature representations, which can be useful in situations where feature selection is important.

$$L1 : \lambda \sum_{i=1}^{m} \left| w_i \right|$$

L2 regularization, which is also referred to as Ridge regularization, introduces a penalty factor that is proportional to the square of the model's weights. This penalty term motivates the model to

learn small weight values, which aids in eliminating overfitting by reducing the model's sensitivity to small fluctuations in the training data.

$$L2 : \frac{\lambda}{2} \sum_{i=1}^{m} \omega_i^2$$

The factor of 1/2 is multiplied to simplify the gradient term [122].

Regularization strength (also known as the regularization parameter or lambda value) controls the impact of the penalty term on the model's loss function. The effect of the regularization strength on the model's performance and the sparsity of the generated coefficients depends on the specific problem and dataset employed. Increasing the regularization strength can assist in enhancing the model's generalization performance by reducing overfitting, but it may also raise the model's bias. Increasing the regularization strength can also lead to sparser coefficient values, as the penalty term encourages parameter values to be closer to zero. However, if the regularization strength is too high, the model may become overly simplistic and underfit the data. Typically, the best regularization strength for a particular problem and dataset is determined by trial and error and model selection techniques, such as cross-validation.

Dropout is another popular technique commonly employed in deep neural networks to prevent overfitting. This technique involves temporarily "dropping out" or setting a random subset of the neurons in a layer to zero during training. This approach is aimed at reducing the network's dependence on specific neurons and promoting the learning of more robust features that are useful for making predictions, even when some of the neurons are missing. Dropout can be applied to any layer in a neural network, but it is usually applied to fully connected layers. The probability of dropout is typically set between 0.2 and 0.5 during training. By randomly dropping out neurons

during training, the neural network is forced to learn more generalized features and reduce the risk of overfitting the training data [123].

The choice of regularization technique will depend on the specific requirements of the model and the dataset. While both L1 and L2 regularization can help to prevent overfitting, L2 regularization has several advantages over L1 regularization [122]. One advantage of L2 regularization is that it tends to produce smoother and more stable models than L1 regularization. This is because the penalty term in L2 regularization is proportional to the square of the weights, which means that small changes to the weights impact the objective function less than they do in L1 regularization. However, if feature selection is important and the dimensions of the data are high, L1 regularization may be preferred. It is also common to use a combination of both techniques, known as Elastic Net regularization, to balance the advantages of both L1 and L2 regularization [124].

Overall, regularization is a fundamental aspect of building effective and reliable machine learning models. By choosing the appropriate regularization technique, developers can produce models that are better suited for their specific applications and that can generalize well to new, unseen data.

2.4.3 Cross Validation

Cross-validation is a widely used statistical technique in machine learning and data analysis to measure the performance of a predictive model. The purpose of cross-validation is to estimate how well a given model can generalize to new, unseen data. The primary idea is to partition the available data into two separate subsets: a training set that is utilized to train the model, and a validation set, that is utilized to evaluate its performance. However, a single split of the data into

training and validation sets may result in a biased estimate of the model's performance, as the choice of the split may impact the model's capability to generalize to new data.

K-fold cross-validation is a popular method to address this issue. It involves splitting the data into K equal-sized folds, where K is a user-defined parameter typically set to 5 or 10. The model undergoes training on K-1 folds and is validated on the rest of the fold; this procedure is carried out K times, with each fold functioning as the set for validation once. The results are then averaged over the K folds to obtain a further reliable measure of the model's performance.

Optimizing the value of K is important to balance the bias-variance trade-off. A small value of K (e.g., K=2) may result in a high bias estimate, as the model is trained on a small subset of the data. On the other hand, a large value of K (e.g., K=10) may result in a high variance estimate, as each fold may contain a small number of samples, leading to instability in the estimated performance. In practice, values of K between 5 and 10 are commonly used.

Using k-fold cross-validation can aid in eliminating overfitting and enhance the model's generalization performance. By estimating the model's performance on unseen data, cross-validation can help identify whether the model is overfitting to the training data or can be generalized to new, unseen data. It can also be used to tune the model's hyperparameters, such as the regularization parameter in L2 regularization, to improve its performance on new data [124].

2.4.4   Bayesian Optimization

Bayesian optimization is a technique of optimization that employs a probabilistic model to optimize objective functions like black boxes that might be unknown, noisy, or expensive. Bayesian optimization tries to determine the objective function's global optimum in a few steps by incorporating prior beliefs about the function and updating this prior belief with new samples derived from the function to produce a posterior that more closely approximates the

objective function. It is also beneficial when evaluations are costly or when the questioned problem is non-convex [125].

Bayesian optimization uses a probabilistic surrogate model to approximate the objective function and a Bayesian technique to update the surrogate model according to observed function evaluations. A surrogate model is used for predicting the objective function at unexplored points, and an acquisition function is utilized to determine the next point to be evaluated [125]. Bayesian optimization has been effectively implemented in numerous applications, such as hyperparameter tuning of ML models, optimization of costly simulations, and drug discovery [126].

Bayesian optimization is comprised of two primary elements: a statistical model for conveying hypotheses about the back-box function and an acquisition function for determining next promising points to evaluate.

The Bayesian optimization model algorithm is as following [127]:

1- Apply a prior distribution to the black-box function f(x), typically a prior based on the Gaussian process.

2- Select typically randomly n=n_0 initial points x_1, x_2,..., x_{n0} from the domain of interest. At these points, evaluate the function f(x) and derive the outputs y_1, y_2,..., y_{n0}.

3- Repeat steps 4 to 7 for n = n_0 + 1, n_0+ 2,…,N until the budget for N function evaluations is achieved.

4- Update the Gaussian process model of the function f(x) utilizing the collected data so far, which generates a posterior distribution on each f(x).

5- Use the determined acquisition function, which optimizes exploration and exploitation of the function space. The selected acquisition function maximizes the trade-off between

exploring unexplored regions of the function space and exploiting promising areas. This optimization method attempts to reduce the number of objective function evaluations.

6- Optimize the acquisition function to identify the subsequent point x_n at which to compute the function f(x).

7- Estimate the function f(x) at the selected point x_n and calculate the output y_n. Incorporate this point into the list of assessed points and update the Gaussian process model.

In the end, the point with the highest observed or predicted output value can be selected as the solution. Standard Bayesian optimization for a 1D problem is depicted in Figure 2.3. During each iteration, the maximum of the acquisition function shows the next most viable point, based on the mean and variance of predictions across the space, to evaluate using the objective function. After updating the Gaussian process model using the new data, the procedure is repeated for the subsequent iteration [125].
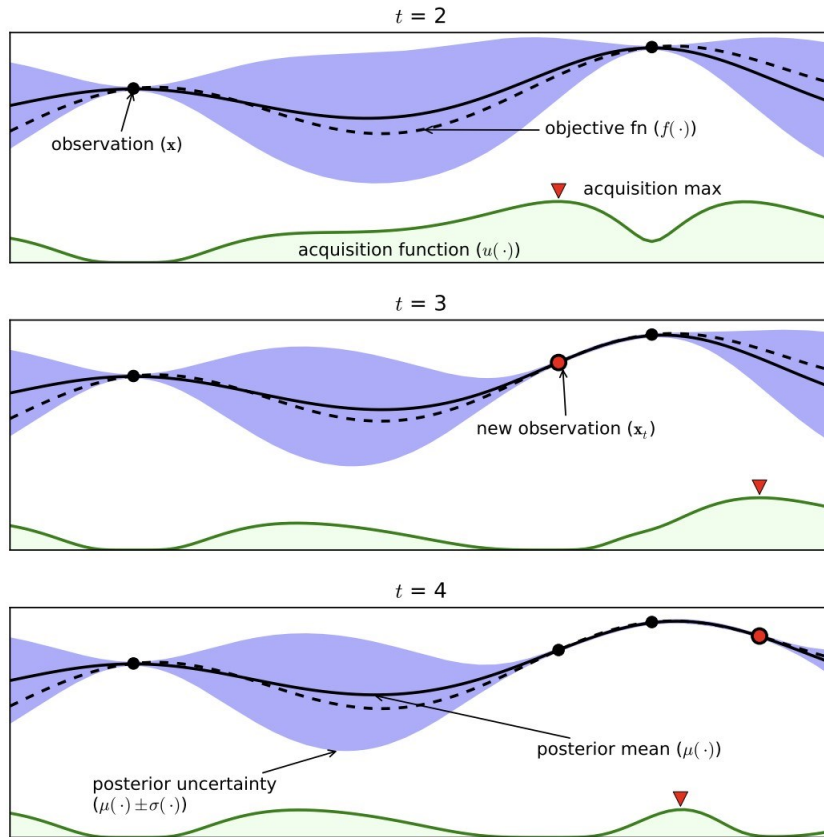
Figure 2.3 A Bayesian optimization procedure in four iterations (t) for a 1D problem. The figures depict true values of observations of the objective function f(x) at two sites (black dots) at t=2, and the dashed line is indicative of the true f(x). The solid black line represents the approximates of the f(x), the posterior mean of the Gaussian process, while the shaded area represents the posterior mean +/-standard deviation. The acquisition function is also depicted in the shaded graphs located below. Prioritizes regions with high predicted objective function values (exploitation) and high prediction uncertainty (exploration). Therefore, locations possessing both characteristics are sampled first. However, the area on the far left is not sampled, despite its high degree of uncertainty, because the algorithm predicts it will not increase f(x) to reach the highest observation [128].

2.4.4.1 Gaussian Processes Regression

A Gaussian process is a stochastic process with an unlimited number of dimensions that is defined by expanding the concept of a multivariate Gaussian distribution. In this process, any finite combination of dimensions is itself a Gaussian distribution. The Gaussian process regression methodology focuses on the values of a function f(x) at a restricted set of points $x_1, x_2, \ldots, x_k \in R^d$ These function values can be collected in the vector $\left[f(x_1), f(x_2), \ldots, f(x_k) \right]$ Similar to how the mean and standard deviation of a random variable completely define the Gaussian distribution, the mean function, m, and the covariance function, Σ of a Gaussian process, define the entire

distribution of functions. The covariance function, or kernel Σ_0 is calculated at each pair of points x_i and x_j to provide the covariance matrix for the Gaussian process. As a result, the prior distribution on the interval $\left[f(x_1), f(x_2), \ldots, f(x_k) \right]$ is [129]:

$$f(x_{1:t}) \sim Normal\left(m_0(x_{1:k}), \Sigma_0(x_{1:k}, x_{1:k}) \right)$$

Where $f(x_{1:k}) = \left[f(x_1), \ldots, f(x_k) \right]$, $m_0(x_{1:k}) = \left[m_0(x_1), \ldots, m_0(x_k) \right]$, and $\Sigma_0(x_{1:k}, x_{1:k}) =$

$[\Sigma_0(x_1, x_1), \ldots, \Sigma_0(x_1, x_k); \ldots, \Sigma_0(x_k, x_1), \ldots, \Sigma_0(x_k, x_k)]$.

The kernel is chosen so that there is a strong positive correlation between the points x_i and x_j that are nearer together in the feature space [127]. Now, assume we had noise-free observations $\{x_{1:n}, f_{1:n}\}$, for some n, and aim to choose the next point in the Bayesian optimization process. Thus, we would want to derive f(x) at the new point x. Let k = n+1 and $x_k = x$; we can then determine the conditional distribution of f(x) given the previous observations [129].

$$f(x) | f(x_{1:n}) \sim Normal(m_n(x), \sigma_n^2(x))$$

$$m_n(x) = \Sigma_0(x, x_{1:n}) \Sigma_0(x_{1:n}, x_{1:n})^{-1} (f(x_{1:n}) - m_0(x_{1:n})) + m_0(x)$$

$$\sigma_n^2(x) = \Sigma_0(x, x) - \Sigma_0(x, x_{1:n}) \Sigma_0(x_{1:n}, x_{1:n})^{-1} \Sigma_0(x_{1:n}, x)$$

The conditional distribution obtained by using Bayes' theorem is known as the posterior probability distribution. The posterior mean $m_n(x)$ is calculated via a weighted average of the prior mean $m_0(x)$ and a data-driven estimate based on $f(x_{1:n})$. The weight assigned to each term is calculated using the kernel used. The posterior variance $\sigma_n^2(x)$ can be calculated by subtracting the variance that is omitted by observing the data $f(x_{1:n})$ from the prior covariance $\Sigma_0(x, x)$.

Moreover, in the training of the Gaussian Process model, a small positive number, such as $\alpha = 10^{-6}$, can be added to each element on the diagonal of the covariance matrix $\Sigma_0\left(x_{1:n}, x_{1:n} \right)$ to modify the numerical stability of this method, particularly when two or more nearby points are

present in the set of points $x_{1:n}$. This ensures that the eigenvalues of the covariance matrix do not approach zero, thereby minimizing the impact of numerical errors on the predictions [127].

2.4.4.2 Covariance Function Selection

We will now discuss the selection of covariance functions or kernels. As mentioned before, points nearer together in the input space are generally further correlated, for example, $||x - x'|| < ||x - x''||$ for a norm $||.||$, then $\Sigma_0(x, x') > \Sigma_0(x, x'')$. Thus, when the values are very close, the covariance function approaches 1, and when they are pretty distant, it approaches 0. Indeed, two points that are near together can be anticipated to significantly impact one another, while points far apart exert minimal impact. This section describes two common kernels in the Gaussian process.

One frequently employed and straightforward covariance function is the power exponential or Gaussian covariance function,

$$\Sigma_0(x, x') = \alpha_0 \, exp\left(- ||x - x'||^2 \right)$$

Where $||x - x'||^2 = \sum_{i=1}^{n} \alpha_i \left(x_i - x'_i \right)^2$. α_i is a kernel parameter whose different values lead to various expectations for how quickly f(x) varies with x.

The second frequently used covariance function in Bayesian optimization is the Matern kernel, which includes a smoothness parameter (ν) to allow increased flexibility in modeling functions [125]:

$$\Sigma_0(x, x') = \alpha_0 \frac{1}{2^{\nu-1} \Gamma(\nu)} \left(2\sqrt{\nu} ||x - x'|| \right)^\nu H_\nu \left(2\sqrt{\nu} ||x - x'|| \right)$$

Where $\Gamma(.)$ and $H_\nu(.)$ are the Gamma function and the Bessel function of order ν. Three approaches are typically used to select the hyperparameters for these kernels: maximum likelihood

estimate (MLE), fully Bayesian approach, and maximum a posteriori. This method's specifics can be found in [127].

2.4.4.3 Acquisition Functions

Acquisition functions are another important component in the Bayesian optimization process. A well-designed acquisition function should balance the trade-off between exploring new areas within the search space and exploiting the promising areas. If an acquisition function does not sufficiently explore, it may become stuck in exploiting suboptimal areas and never discover the global maximum solution. Nevertheless, it may lose expensive evaluations modifying an already developed Gaussian process if it explores excessively. Different acquisition functions exist, such as the Gaussian process upper confidence bound, entropy search, Expected Improvement, Probability of Improvement, and knowledge gradient. One of the most commonly used functions is the Gaussian process upper confidence bound (GP-UCB), which is among the most intuitive acquisition functions, balancing exploration and exploitation with a single hyperparameter κ [130].

$$GP - UCB_t(x) \ = \ m_t(x) \ + \ \kappa\sigma_t(x)$$

Where $m_t(x)$ and $\sigma_t(x)$ are the mean and the standard deviation of the Gaussian process at iteration t and point x, respectively. Increasing κ favors points with a high variance promotes exploration, whereas decreasing κ favors points with a high mean, which promotes exploitation.

By identifying the x value that maximizes the GP-UCB (x), the next best point to evaluate for the function f(x) is determined.

Chapter 3 Predicting and Optimizing Cell Viability in Bioprinting through Regression, Classification Neural Networks, and Bayesian Optimization

3.1   Introduction

In the domains of regenerative medicine and cancer modeling, 3D bioprinting has witnessed rapid growth and is rising in popularity [131]. 3D bioprinting is a biofabrication process that refers to computer-guided additive manufacturing that allows the development of extremely precise and complex 3D structures utilizing biologic materials in a pre-designed geometry [132,133]. There are two approaches for including cells in 3D printing: two-step bioprinting entails seeding the cells on the surface of pre-printed constructions, and one-step bioprinting involves incorporating the cells directly into the bioink before making cell-laden 3D structures [134].

In one-step bioprinting, bioink is a key component of the process, as it serves as the building block for the final construct. Bioinks can be made from a variety of materials, including hydrogels, polymers, and decellularized extracellular matrices [135]. Hydrogels derived from natural or synthetic materials are widely used as bioinks in one-step bioprinting. They can provide the high-water content of the natural ECM and let tunability of the mechanical strength and biochemical properties. Most of these materials need optimization before they can be effectively applied in bioprinting. This optimization involves modifying the hydrogel's viscosity and mechanical strength which can be achieved by the addition of thickening materials or by crosslinking to create a scaffold that closely resembles the ECM [136].

Alginate is one of the commonly used hydrogels in 3D bioprinting, which is biocompatible and biodegradable and can be crosslinked by multivalent cations such as Ca^{2+} [137]. After crosslinking, the alginate-containing scaffolds are illustrated to keep their morphology and enhance their mechanical properties [138]. One disadvantage of alginate is its poor capability to provide enough binding sites for cells to attach and proliferate [139]. However, this problem can be solved by introducing cell attachment peptides such as the RGD sequence, a typical cell binding site, to modify the properties of alginate [140]. Of all the different polymers that can provide cell attachment peptides, gelatin attracted great interest during the last decade. Gelatin is a water-soluble polymer that is the product of partial hydrolysis of natural polymer collagen and has a similar structure to that of collagen, the main ECM component. Gelatin contains several biological signals that improve cellular functions, including cell attachment, growth, and proliferation [138, 141]. However, the disadvantages of applying gelatin as bioink are its relatively low mechanical strength and fast biodegradation [141]. Hence, the composition of the gelatin and alginate can be considered as an effective way to develop a bioink with modified properties to be used for bioprinting 3D cell-laden structure in this study.

Although the introduction of one-step bioprinting is significant progress in 3D printing, cells may be impacted or even lost during the printing procedure. In extrusion-based bioprinting, the extrusion of bio-ink from the nozzle can result in severe shear stress that can lead to cell membrane damage. Hence cell viability, defined as the proportion of live cells at the end of the bioprinting, is an essential factor in bioprinting in different applications like cancer research, as it can affect the accuracy of the 3D bioprinted model, the reliability of the results, and the development of new therapies. Therefore, it is crucial to optimize the bioprinting conditions to maintain high cell viability and ensure the success of bioprinting.

Several efforts have been made to produce biomimetic scaffolds with high cell viability using an extrusion-based bioprinting technique [31]. Cell survival, depending on the type of cell line, in the 3D bioprinting process can be affected by different variables, including bioink formulation, 3D printer parameters, and post-printing treatments. Each of these categories may be subdivided into distinct parameters: Bioink formulation consisting of the type and concentration of biomaterials; 3D printing parameters including cartridge temperature, bed temperature, nozzle size, printing pressure, printing speed, etc.; and post-printing treatment involving crosslinking process that can be adjusted by choosing a suitable crosslinking method and fine-tuning the amount of crosslinker and crosslinking time.

Until now, optimization of the extrusion-based bioprinting settings has been conducted using time-consuming and expensive trial-and-error experimentation. This optimization process can become even more challenging when it comes to incorporating various biomaterials- and printer-related factors into the process. [100]. As a solution, computational techniques have been introduced in 3D bioprinting. These days, with the emergence of supercomputers, ML has provided a new perspective for numerous sectors of science and engineering, such as biofabrication. It is anticipated that ML can speed up and improve bioprinting process by optimizing research design and findings, diminishing trial-and-error experiments, and decreasing fabrication time and expense[31,142].

ML as one the recent computational techniques can find connections between input parameters and forecast expected results based on these connections The majority of ML applications have focused on improving printability [97], optimizing material utilization [143], and architectural design factors to enhance material characteristics [144,145] . However, only a few studies have been working on applying ML algorithms to study the effect of 3D bio-printing parameters on cell viability.

Therefore, the relationship between the effective aforementioned-bioprinting parameters with cell survival rate is not yet well-studied. In one of these studies, Xu et al. [99] created an ensemble learning model to predict cell survival in stereolithography-based bioprinting by considering the impact of four crucial parameters, including UV intensity, length of exposure, GelMA content, and layer thickness. The model achieved a high level of accuracy, as evidenced by R^2 of 0.953 when tested on 10% of the dataset. In another recent study, Tian et al. [100] studied the capability of ML regression and classification approaches in predicting the survival of cells and the printability of cell-containing alginate and gelatin-based bioinks in extrusion-based bioprinting. To do this, they collected a dataset on bioink material content, solvent utilized, crosslinking data, printing settings, survival rate, and printability results from various bioprinting laboratories. Their results showed that the random forest models yielded the highest accuracy in both regression and classification.

So far, the existing studies have focused on different optimizing bioprinting parameters without taking into account the cell type. However, it is essential to note that different cell types have different optimal bioprinting parameters and have different characteristics that affect their viability during the bioprinting process, such as their size, shape, and susceptibility to mechanical stress. Therefore, the importance of cell type in predicting cell viability and optimizing bioprinting parameters is crucial for the development of effective bioprinting strategies.

In this study, we aimed to achieve two objectives:

First, to develop predictive regression and classification neural networks in bioprinting process. To accomplish this, we gathered 92 data points in our laboratory under a variety of experimental conditions, including bioink formulations incorporating gelatin and alginate, 3D printing parameters of extrusion-based bioprinting, and crosslinking conditions. To train and test our ML

46

model, we also acquired data from previously published bioprinting research involving and combined it with our own data to form a dataset of 713 instances. We evaluated the capability of ML regression and classification algorithms to effectively and precisely predict cell viability result depending on different bioprinting conditions. We also identified the most important parameters influencing cell viability prediction through network training in regression model.

Second, we integrated a Bayesian optimization model with the regression neural network and subset of bioprinting experimental data for the first time in bioprinting field to inversely predict the optimum values of selected bioprinting parameters that yields the highest cell viability. Bayesian optimization is a powerful tool used in ML to optimize complex systems [127]. The method uses a probabilistic model to predict the performance of multiple input values to discover the best possible input for a given function by iteratively selecting new input values based on the outcomes of prior evaluations. The key idea about this optimization is to use prior knowledge about the function by building a probabilistic model using a Gaussian process prior distribution, which is updated as new function evaluations become available [125].

This study used the developed regression neural network as the objective function of the system in the Bayesian optimization model, focusing on fine-tuning the crosslinking condition including the crosslinker concentration and crosslinking time for specific sets of parameters while holding all other parameters constant. The Bayesian model, with its probabilistic approach and Gaussian process prior distribution, was able to provide a robust framework for estimating uncertainty in viability prediction. Additionally, using the trained neural network to make predictions in the Bayesian model could also improve the accuracy of the predictions by guiding the optimization search more effectively using the information acquired from the available data. This integration

could lead to faster convergence and improved system performance, making it an efficient method for obtaining more accurate predictions with limited data.

3.2 Materials and Methods

3.2.1 In-vitro Studies

3.2.1.1. Materials

The alginic acid Sodium salt, dimethyl sulfoxide (DMSO), Sodium chloride, Calcium chloride (CaCl$_2$) and gelatin from bovine skin (type B) were obtained from Sigma-Aldrich (Canada). For cell culture studies, MDA-MB-231 were purchased from ATCC, Dulbecco's Modified Eagle Medium (DMEM), fetal bovine serum (FBS), penicillin/streptomycin, Trypsin/EDTA solution at 0.25% (w./v.) and phosphate buffer saline (PBS) tablets were bought from Wisent Bioproducts. Furthermore, a Live-Dead Cell Viability assay kit (CBA415), Hoechst 33342Nuclei Dye were provided from Sigma-Aldrich (Canada).

3.2.1.2. Cell Culture

MDA-MB-231 cells were cultured in DMEM supplemented with 10% FBS and 1% 100 Uml−1penicillin/streptomycin in T-75 flasks. Cells were incubated at 5% CO2 and 37 °C, and the media was exchanged every other day. When cells reached 80% confluency, the cells were rinsed twice with DPBS and then suspended with trypsin/EDTA (0.25%-1X).

3.2.1.3. Bioprinting

To prepare bioink, first 0.5% NaCl solution using DI water was prepared. Then, gelatin and alginic acid with different concentrations from 1% (w/v) to 8% (w/v) were added, and the solution was stirred vigorously for 1 hr. The prepared solution was sterilized by heating at 70 °C

(30 min, three times) and then kept at 4 °C before usage. Prior to bioprinting, MDA-MB-231 suspension was gently and evenly mixed with prepared gelatin/alginate solution using a mixing syringe to make a bioink with final cell number of 1-5-2 $\times$ 10^6 MDA-MB-231 cells mL^{-1}. Bioprinting was carried out using CELLINK INCREDIBLE+ extrusion-based bioprinter. The prepared bio-ink was extruded using a needle to fabricate the 3D cell-hydrogel layered grid constructs. More specifically, printing nozzles applied for these experiments were standard conical nozzles (25G and 22G), and the moving speed of the needle was adjusted at 20 mms^{-1}. The printings were done at room temperature by applying appropriate various pressure. The structures were designed using Solidwork software and sliced in slicer software to 10 layers with the size of $15 \times 15 \times 3 \, mm^3$ with a rectilinear filling pattern. During the printing, we made an effort to maintain the pressure at the least level to have the minimum damage to cell viability. Subsequently, the cell-laden constructs were immersed in sterile calcium chloride solutions with two different concentrations (3% and 5% (w/v)) for different periods of time (5, 10, 15, 20 min) for crosslinking sodium alginate with calcium ions [10]. Then, after being washed with PBS three times, each structure was cultured in DMEM medium containing 10% FBS and 1% penicillin/streptomycin in a 12-well plate and incubated at 5% CO2 and 37 °C.

3.2.1.4. Live-Dead Assay

The 3D bioprinted cell-laden constructs were stained after printing using Live-Dead staining viability kit based on the manufacturer's instructions to determine cell viability. Briefly, each construct was washed in PBS three times. Then, for staining each three repeats, we first prepared a 12 ml dye solution by mixing 6 ml of cell complete medium and 6 ml of PBS (1:1), followed by adding 5 µl of CAM and 20 µl of PI (present in Live-Dead Cell Viability assay kit (CBA415)). Then 2 ml dye mixture was applied to each well to stain cells while incubated for 1 hr in the

incubator. A laser scanning confocal microscope (Zeiss LSM 700) was used for imaging the live and dead cells within the 3D constructs at multiple spots. Then, to quantify cell viability, images were analyzed using ImageJ software. First, a pre-processing step was performed to improve quality of images and eliminate any background noise. This step involved adjusting the brightness and contrast of the images to ensure optimal visibility of the fluorescent signals. the analysis proceeded by segmenting the green and red channels in the images. Then, through careful manual counting, the number of cells exhibiting green fluorescence was determined from the green channel, while the number of cells showing red fluorescence was counted from the red channel. These counts represented the population of live cells and dead cells, respectively. At the end, Cell viability was counted by dividing the number of green (living) cells by the total number of cells in each image.

3.2.2   Computational Studies

3.2.2.1. Dataset Preparation

The dataset created for this thesis consists of two sections. The first section provides 92 data points collected in our laboratory under a range of experimental settings, including several bioink formulations, particular 3D printing parameters, and particular crosslinking conditions. The dataset's second section comprises 63 previously reported bioprinting studies, including 14 distinct cell types. This section was integrated with our own data to create a comprehensive dataset of 713 instances, each with a distinct cell viability value.

Parameters related to bioink formulations mainly include "Gelatin_Concentration(%w/v) " and "Alginate_Concentration(%w/v)". In some research studies, gelatin and/or alginate were mixed with certain concentrations of other hydrogels, such as GelMA, hyaluronic acid, and

nanocellulose, which are also added to the dataset. The 3D printing parameters include "Printing_Pressure(kPa)," "Movement_Speed(mm/s)", "nozzle_size(µm)", "Cartridge_Temperature(°C)", and "Bed_Temperature(°C)." Finally, post-printing treatments, including "crosslinker(CaCl2)_concentration(mM)" which in this study is calcium chloride (CaCl$_2$), and "Physical_Crosslinking_time_(s)" are added to the dataset. In addition, the duration of photocrosslinking in some experiments, such as the ones using GelMA in bioink formulations, is incorporated into the dataset.

Moreover, "Cell type" and "cell number" are other important characteristics of our dataset collection. "Cell type" specifies the type of bioprinted cells, such as "MDA-MB-231," the cell type employed in our experiments, and "Cell_Number(cells/mL)" represents the starting cell number per mL of bioink with which bioprinting was initiated. In research studies, thirteen other cell types were utilized, including "Mesenchymal Stem Cells (MSCs)", "Endothelial Cells", "Schwann Cells", "Fibroblast Cells", "Glioma Cells", "Femur-derived Cells", "Chondrocytes", "HepG2 Cells", "Myoblasts", "Epidermal Stem Cells (ESCs)", "Cartilage Progenitor Cell (CPCs)", "Embryonic Stem Cells ", and "Hela Cells."

"Time of measurement(day)" is another characteristic of the dataset that indicates how long after bioprinting the cell viability was evaluated. Cell survival at the measurement time is represented by the target variable "Cell Viability (%)".

3.2.2.2. Data Preprocessing

Preprocessing is a critical step in ML that involves transforming raw data into a format that can be effectively used by ML models. It typically involves cleaning the data, scaling the data, encoding categorical variables, and handling missing values. During the preprocessing stage of

this research, first, we removed all the rows where the "Alginate_Concentration(%w/v)" column and the "Gelatin_Concentration(%w/v)" column were both 0 because we wanted to focus on gelatin- and alginate-based bioinks containing at least one of these materials. Then, numerical data were normalized using MinMaxScaler function for better model evaluation. This function scales the values of each feature between 0 and 1 to ensure that each feature contributes equally to the model. One-hot-encoding was utilized to convert categorical data ("Cell type") into numerical data. This technique involves creating a binary column for each category of a categorical variable. Additionally, the instances in which cartridge and substrate temperatures were null were assigned a value of 22.5 °C, as it was assumed that the experiments were performed at room temperature. In addition, k-nearest-neighbors imputing was applied to impute the missing values of the other variables, with a neighbor range of 10. In this method, by considering the known values of the 10 neighboring data points, the algorithm estimated the missing values for the variables of interest, resulting in a complete dataset suitable for further analysis. Then, the dataset is split into a training set and a test set using the train_test_split() function from the scikit-learn library, with a test size of 0.15.

3.2.2.3. Regression Neural Networks

For the cell viability regression model, the neural network consists of three fully-connected layers: a fully-connected layer with 44 neurons and a hyperbolic tangent activation function (tanh), a fully-connected layer with 120 neurons, a hyperbolic tangent activation function, and L2 regularization with a weight of 0.01, and a fully-connected layer with a single neuron and a linear activation function. The model is compiled with the MSE loss function and the Adam optimizer with a learning rate of 0.01. The neural network used in this study is based on the Keras library, which is a high-level API that allows the creation of neural network models with minimal code.

The dataset is split into a training set and a testing set using the train_test_split function with a test size of 15%. The K-Fold cross-validation technique is applied with 10 folds to evaluate the model performance. For each fold, the dataset was scaled using the MinMaxScaler, and the model was trained on the training set for 1000 epochs with a batch size of 5. The training and validation set MSE and R^2 cores are calculated for each fold, and the average scores are computed across all folds. Finally, the trained model is evaluated on the test set, and the predicted values for cell viability are compared with the true values.

3.2.2.4. Classification Neural Network

For cell viability classification models, a binary class was created from the numerical cell viability data by converting the target variable to binary by assigning 1 to values greater than 70 and 0 to values less than or equal to 70. The neural network model has three dense layers. The first dense layer has 44 nodes with a ReLU activation function. The second dense layer has 200 nodes with a ReLU activation function and L2 regularization with a lambda value of 0.01. A dropout layer with a rate of 0.7 follows the second dense layer. Finally, a dense layer with a single node and a sigmoid activation function is used to produce the binary classification output. The model is compiled with a binary cross-entropy loss function, accuracy metrics, and the Adam optimizer with a learning rate of 0.001. The dataset is split into a training set and a testing set using the train_test_split function with a test size of 15%. The model is trained using K-fold cross-validation with 10 folds on the training set. The data is scaled using MinMaxScaler, and the training and validation set accuracy, precision, and recall are recorded for each fold, and the averages across all folds are computed. Finally, the trained model is evaluated on the test set, and the predicted values are compared with the actual values.

3.2.2.5. Hyperparameter Fine-tuning

We utilized the GridSearchCV module from the scikit-learn library to conduct hyperparameter tuning on neural networks. This module allowed us to systematically search for the optimal hyperparameter combination within a given range by conducting a grid search over a specific hyperparameter space. These hyperparameters include activation function, dropout rate, L2 regularization strength, learning rate, number of layers, and number of units in the hidden layers. The model is then trained and evaluated for each possible combination by means of cross-validation, and let us find the optimal hyperparameters that produced the greatest performance on the validation dataset and reduce the possibility of overfitting.

3.2.2.6. Permutation Importance

We used the Permutation Importance method to evaluate the significance of each feature in the built regression model for predicting the target variable. To do this, we utilized the "permutation_importance" utility from the scikit-learn library. This method involves arbitrarily permuting the values of a single feature and measuring the resulting decrease in the model's performance or score. This procedure is repeated K times to estimate feature significance accurately. The importance (i_j) for each feature j for K repetition is calculated as:

$$i_j = s - \frac{1}{K} \sum_{K=1}^{K} s_{K,j}$$

Where s is the reference score of the model, and $s_{K,j}$ is score of the model for permuted data which feature j and on K^{th} repeat. In this study, $K = 10$ and, since the neural network's loss function is MSE, the permutation importance score is also determined by the decrease in MSE when each feature is permuted.

3.2.2.7. Bayesian Optimization Model

We employed a Bayesian optimization model based on a constructed regression neural network in order to optimize the bioprinting variables for maximizing cell viability. We aimed to determine the optimal "physical-crosslinking time" and "crosslinker ($CaCl_2$) concentration" for a given set of bioprinting parameters to achieve the highest level of cell viability possible.

The algorithm of this neural network-based Bayesian optimization model is depicted in Figure 3.1. Here, the based Bayesian optimization framework started by constructing the Gaussian Process model using the created dataset of variables in the bioprinting process and their associated cell viability. The Gaussian Process model consists of a mean function and a covariance function. The covariance function, which in this study is the Matern kernel, with a smoothness parameter ν equal to 2.5, conveys the smoothness of the function and determines how the cell viability percentage at one location affects the estimate of the values surrounding it [128]. Moreover, in the training of the Gaussian Process model, a small positive number, $\alpha = 10^{-3}$, was added to each element on the diagonal of the covariance matrix to modify the numerical stability of this method.

The mean and covariance from the Gaussian Process model were utilized to develop an acquisition function that suggests the next questioned bioprinting variables to probe. The suggested parameters were chosen to maximize the probability of achieving the highest cell viability while maintaining a balance between exploration and exploitation. In this study, the GP-UCB acquisition function, with κ=2.5, was applied [130]. New suggestions for questioned parameters were chosen within the range specified for them. Some of these ranges in our bioprinting dataset can be found in Table 3.1. Then, the newly recommended crosslinking time, amount, and other constant variables were fed into the developed regression neural network in each iteration. Utilizing the developed neural network, we were able to predict cell viability without having to conduct time-

consuming and laborious laboratory experiments. Then, using the collected data so far, the Gaussian process model was updated and generated a posterior distribution. This process continued for 50 iterations, then the best combination of parameters with the highest observed or predicted cell viability was chosen as the solution.
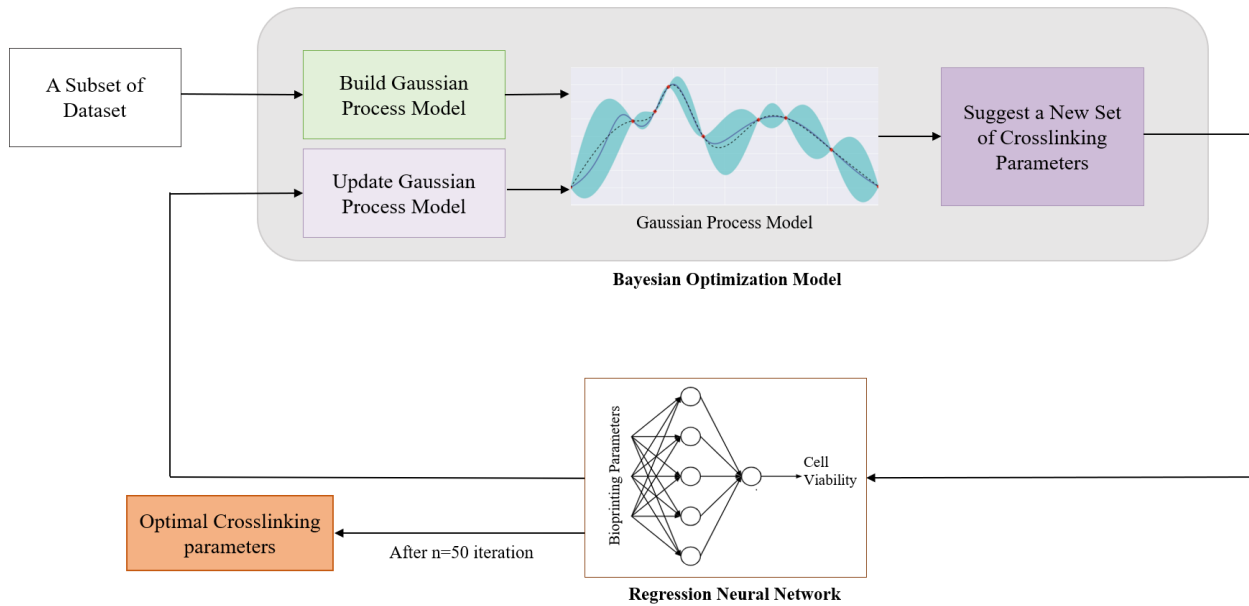
Figure 3.1 Algorithm of Bayesian Optimisation Model based on regression neural network for bioprinting process optimization.

3.3   Result and Discussion

3.3.1   In-vitro Studies

In this study, we conducted in-vitro experiments and bioprint gelatin and alginate-based bioinks under various printing and post-printing conditions. We selected gelatin and alginate for formulating the bioink because they are the most common hydrogels used in bioprinting studies due to their suitable properties in terms of biocompatibility and rheology [146]. Then, we utilized live-dead assay and confocal microscopy to measure cell survival following printing for each set of parameters and collected a dataset of effective parameters and corresponding cell viability. To create a more comprehensive dataset, we combined our results with those from other laboratories.

This resulted in a dataset with a categorical parameter, "Cell type", encompassing 14 distinct cell types listed in the previous section. Some of the numerical variables in the final dataset are also summarized in Table 3.1, including their maximum and minimum values. Based on this table, for instance, data points are generated with various bioink formulations containing varying concentrations of gelatin and alginate ranging from 0 to 20% (w/v), and with varying printing setting parameters such as extrusion pressures ranging from 5 kPa for low-viscosity bioinks to 300 kPa for highly viscous ones.

Table 3.1 A subset of numeric variables in the bioprinting dataset and their range.

Feature	Unit	Min	Max
Alginate Concentration	%w/v	0	20
Gelatin Concentration	%w/v	0	20
GelMA Concentration	%w/v	0	7
Hyaluronic acid Concentration	%w/v	0	2
NaCl2 Concentration	mM	0	272
Agarose Concentration	%w/v	0	1
Crosslinker (CaCl2) Concentration	mM	0	901
Physical Crosslinking time	s	0	1800
Photo-crosslinking time	s	0	2400
Printing Pressure	kPa	5	300
Movement Speed	mm/s	0.666	235
Nozzle Size	µm	101.6	1600
Cell Number	cells/mL	100,000	20,000,000
Cartridge Temperature	°C	4	40
Bed Temperature	°C	3.5	25
Time of Measurement	day	0	30

Each of these parameters has the potential to affect the viability of cells in the bioprinting procedure. For example, cells printed within the same bioink (4%(w/v) gelatin - 3%(w/v) alginate), crosslinking conditions, and printing settings but with different nozzle sizes showed different cell viabilities. Cells extruded using a nozzle with a diameter of 250 micrometers exhibited approximately 61% cell viability, while those printed using a nozzle with a diameter of 410

micrometers showed 82% cell viability. This disparity in cell viability is due to the difference in nozzle size; the smaller the nozzle, the lower the cell survival rate. Another example is for cells printed with different bioink formulations. Figure 3.2 displays images of stained cells in 4%(w/v) gelatin-8%(w/v) alginate and 4 (%w/v) gelatin- 4(%w/v) alginate using the same other parameters except for the printing pressure, which varied based on the bioink viscosity. Cells in 4%(w/v) gelatin-8%(w/v) alginate exhibited approximately 57% viability, whereas those in 4%(w/v) gelatin-4%(w/v) alginate exhibited about 68% viability. The lower cell viability in the 4%(w/v) gelatin-8%(w/v) alginate hydrogel may be attributable to the higher amount of alginate, resulting in a more viscous and stiffer bioink necessitating a greater printing pressure.
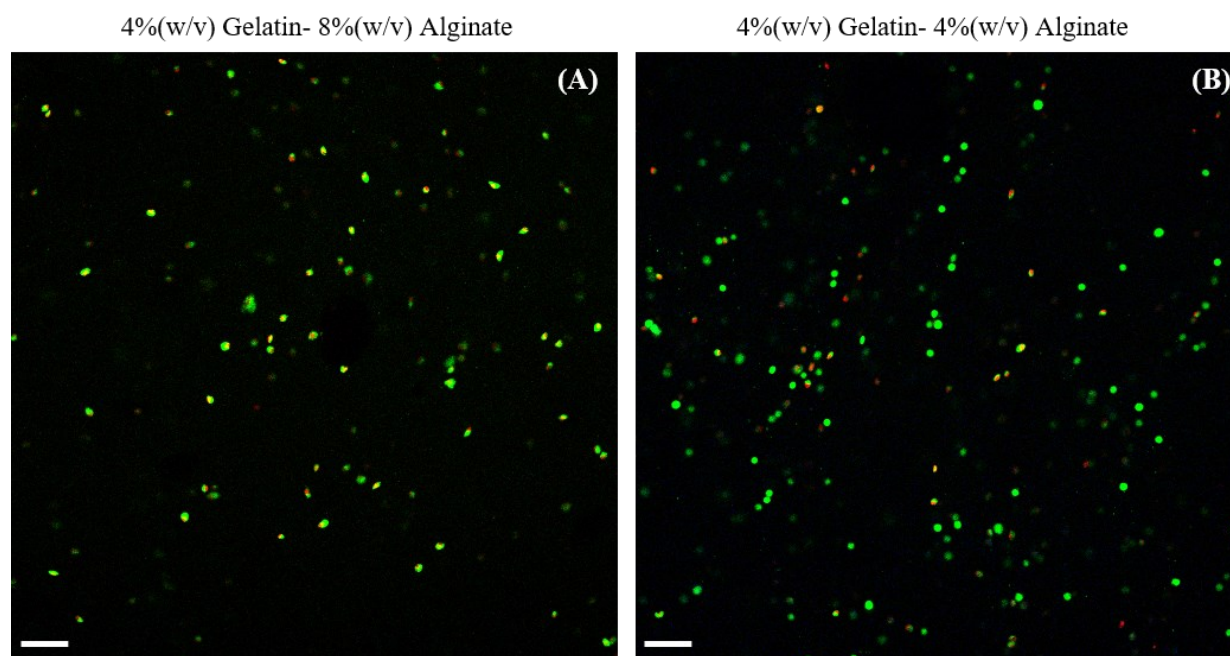
Figure 3.2 Microscopic images illustrating viability of MDA-MB-231 cells within 3D hydrogel constructs using the fluorescent live-dead assay kit on day 0. Bioink is consist of: (A) 4 (%w/v) Gelatin-8 (%w/v)Alginate, (B) 4 (%w/v)Gelatin- 4 (%w/v) Alginate. Live and dead cells were stained using calcein-AM and PI (green color represents live cells; red color represent dead cells). Cells were imaged using the laser scanning confocal microscope. Scale bar, 100 μm.

3.3.2 Computational Studies

3.3.2.1 Regression and Classification Neural Networks

To improve the bioprinting optimization process, we developed neural networks for regression and classification purposes to predict cell viability based on key variables. To construct these networks, we randomly divided the dataset into a training set containing 85% of the data points and a test set containing 15% of the data points. We utilized feature scaling techniques to ensure that all numeric features were normalized to the same range of distribution. Cross-validations were performed on the training datasets to determine the optimal configuration of ML algorithms. In order to prevent data leakage, the encoding and scaling methods were performed exclusively on the subtraining dataset and then applied to the subvalidation dataset during cross-validation.

The performance of the regression neural network model on the train set and test set is shown in Table 3.2. This graph depicts the mean and standard deviation of performance metrics, including MSE, MAE, and R^2 across 10 folds for the train set. This figure also showcases the corresponding values of MSE, MAE and R^2 for the test set. The mean MSE, MAE, and R^2 values for the train set were found to be 0.010, 0.075, and 0.81, respectively. Similarly, the corresponding metrics for the test set yielded values of 0.016, 0.089, and 0.71, respectively.

These metrics provide insight into the model's success in prediction. Lower MSE and MAE values indicate that the model's predictions are closer to the actual values, whereas a higher R^2 value indicates that the model explains a substantial portion of the data's variance. Thus, the relatively low MSE and MAE values for both the train and test sets and the relatively high R^2 value suggests that our model was trained effectively on the dataset and can predict cell viability with good performance.

Table 3.2 Performance of the regression neural network model evaluated using MSE, MAE, and R^2 for the train set and test set. ± Std indicates the standard deviation for each metric for the train set across 10 folds.

Metric	Train Set (Mean ± Std)	Test Set
MSE	0.01 ± 0.0024	0.016
MAE	0.075 ± 0.0089	0.089
R^2	0.81 ± 0.03	0.71

Table 3.3 compares the cell viability of random six laboratory-collected samples, available in the test set, with the predicted values derived from the regression model. As shown in this table, the predicted viability and actual one are in good agreement confirming the model's good performance in predicting cell viability and its ability to replicate laboratory measurements.

Table 3.3 A comparison between the predicted viability values using the regression neural network model and the actual value produced in the laboratory.

Sample	Predicted Viability (%)	Actual viability (%)	Error%
1	81.636	84.846	3.783
2	82.504	95.986	14.046
3	95.917	94.49	1.510
4	97.891	94.9821	3.063
5	80.656	73.204	10.179
6	42.099	41.792	0.734

For the purpose of the classification, we defined two classes for cell viabilities: 0 for viabilities below 70% and 1 for viabilities greater than or equal to 70%, the threshold accepted in the literature as indicative of acceptable cell survival in 3D bioprinting [147]. To evaluate the neural network's performance in the training and validation process, we measured the average accuracy, precision, and recall over 10 folds in the training. Similar to regression model evaluation, a test set, 15% of the whole dataset, was used to assess its ability to predict cell viability on unseen data. The results of the model performance on the train set and test set are presented in Figure 3.3. This figure shows

that the model achieved a mean accuracy of 0.91, a mean precision of 0.96, and a mean recall value of 0.90 for the train set. Additionally, this figure shows these metrics for the test set, which are 0.86, 0.85, and 0.97, respectively. Thus, with an accuracy score of 0.91 in seen data and 0.86 in unseen data, it can be concluded that the model correctly classified 91% and 86% of instances in the train set and test set, respectively, representative of how well the model could be trained and predict the objective. Moreover, with a recall score of 0.97, the model accurately captured 97% of the actual positive or high-viability (>70%) instances in the test set.

Thus overall, these findings suggest that both regression and classification neural networks are effectively trained on the created dataset and can predict cell survival with excellent performance, making it a reliable tool for predicting cell viability in bioprinting processes with high generalization.

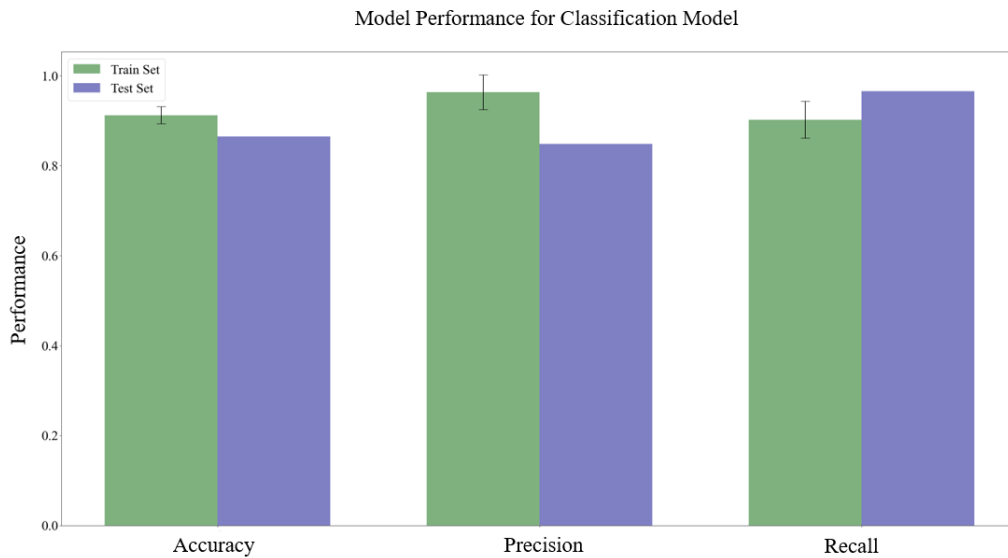**Figure 3.3** Performance of a neural network Classification model evaluated using Accuracy, Precision, and Recall for train set and test set. Error bars indicate the standard deviation for each metric for train set across 10 folds.

Our developed neural networks for both regression and classification also outperformed the only study that predicted cell viability so far in extrusion-based bioprinting. This study, conducted

by Tian et al. [100], which was also used as one of the sources for creating our dataset, compared three regression models (support vector regression, linear regression, and random forest regression) and three classification models (random forest classification, logistic regression classification, and support vector classification). Compared to other models, the random forest models produced the highest R^2 (approximately 0.38) and prediction accuracy (approximately 65%) in regression and classification. However, the trained neural networks in our study yielded a regression R^2 value of 0.71 and classification accuracy of 0.86, representing a significant improvement over previous research. These results demonstrate the efficacy of neural networks for predicting cell viability in bioprinting and imply that they may offer significant advantages over conventional regression and classification models. Additionally, we believe the models' performance can be enhanced even further by expanding the dataset to capture a broader range of patterns and relationships, reducing overfitting, and enhancing the model's generalization ability. Increasing the number of instances with diverse combinations of bioprinting parameters and corresponding cell viability measurements can be accomplished by collecting additional in-vitro data in the laboratory or incorporating data from other sources.

3.3.2.2 Permutation Importance of Bioprinting Parameters

Permutation Importance was applied to determine the most significant features of the bioprinting for predicting cell viability using a neural network for regression. Figure 3.4 shows 10 of the most important features; the x-axis represents the significance score, while the y-axis represents the evaluated bioprinting parameters.
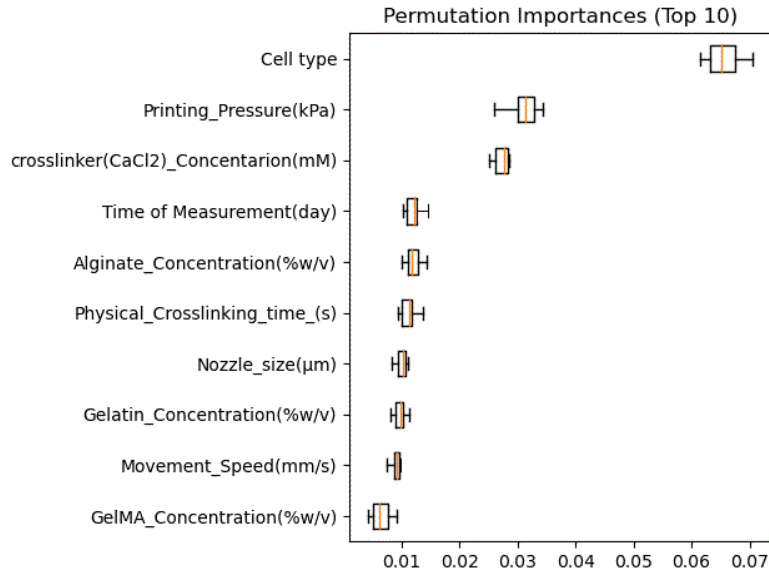
Figure 3.4 Permutation Importance of bioprinting parameters on cell viability, based on regression neural network. The boxplot illustrates the distribution of each parameter's importance scores.

As demonstrated in this figure, the "Cell type" has the most significant impact on optimizing the bioprinting process compared to other variables. In the dataset, we included cancer cells for cancer research and normal cells for tissue engineering purposes. In the bioprinting procedure, the survival and viability of these two cell types can vary significantly. Typically, cancer cells with a higher proliferation rate may be better able to recover from any printing-related damage, while normal cells may be more fragile and susceptible to injury. Different cell types within cancerous and normal cell populations may demonstrate varying bioprinting sensitivity. The degree to which cells are affected by pressure and shear stress depends on the cell type. This variation can be attributed to the unique physiological properties of each cell type. For instance, cells with a more robust cytoskeleton may be more deformation-resistant during printing. Studies showed that, for example, immortalized cells like L929 fibroblasts, human mesenchymal stem cells, and HeLa cells experience less cell damage compared to mice embryonic stem cells when exposed to the same level of shear stress [34,35,33]. In other studies, numerous cell types, including fibroblasts [64], myoblasts [148], endothelial cells [65], chondrocytes [66], and Schwann cells [149], were printed in alginate-based

structures, and each exhibited unique cell vulnerability to the post-printing process. Hence, these observations highlight the significance of considering cell types as a critical element for optimizing cell viability in the bioprinting process.

The second most significant bioprinting variable is "Printing_Pressure", influencing cell viability prediction. The extrusion involves using pressure to force the bioink through the nozzle and deposit it onto the substrate. The amount of pressure applied determines the flow rate and deposition speed of the bioink, affecting the tension experienced by the cells during printing. Previous research has indicated that increasing dispensing pressure during bioprinting has a damaging effect on cell viability [150]. This is primarily due to the mechanical stress and shear forces imparted on the cell membrane by the pressure, causing damage and diminishing viability. Additionally, high pressure may cause the formation of air bubbles or the fragmentation of the bioink, which can further damage cells. Consequently, optimizing printing pressure is essential for achieving high cell viability during the bioprinting procedure. It is important to note that pressure can be controlled by considering factors such as nozzle size and material properties within the bioink, which will be further explained in the subsequent explanations.

Among other 3D printer parameters, Permutation Importance in our study shows that after "Printing_Pressure", "Nozzle_size" and "Movement_speed" are among the 10 most important variables affecting cell viability prediction. This result aligns with previously conducted research that concluded that the dispensing pressure has a more substantial impact when compared to the nozzle size [30]. Smaller nozzle sizes typically result in higher printing pressures due to the restricted flow area, whereas wider nozzles allow for reduced extrusion pressures resulting in decreased shear stress. Using wider nozzles, however, diminishes the filament's resolution. Hence, selecting the proper nozzle size is one of the parameters that must be tuned to produce the best outcomes in

terms of printability and viability [151]. High printing speeds, as the 9th important parameter, can also lead to mechanical stress on cells, which can affect their viability and functionality. In addition, high extrusion speeds can reduce the precision and resolution of 3D printed structures, resulting in regions of high stress or low nutrient supply, which affects cell survivability post printing. On the other hand, low printing speeds can result in prolonged cell exposure to the printing environment and excessive residence time in the nozzle, which is harmful to cells [31]. Therefore, the optimal printing speed may depend on several variables, including the required printing resolution, as well as the sensitivity of the cells to mechanical stress.

The third significant parameter in predicting cell viability, identified in our study, is "crosslinker(CaCl2)_concentration". The crosslinking method, including physical and chemical, can be selected based on the bioink formulation [60]. For example, in this study, Ca^{2+} in $CaCl_2$ solution is used as a crosslinker for an alginate-containing scaffold, which forms a flexible and solid structure. Increasing the amount of crosslinker means more calcium ions can move into the 3D network, leading to more crosslinked and a firmer construct, while a higher amount of Ca^{2+} can be toxic to cells and decrease cell viability [138]. Another important parameter in controlling the crosslinking process is "physical_Crosslinking_time" which is the length of time the printed structure is immersed in the $CaCl_2$ solution, and is among the 10 most important factors in optimizing cell viability during bioprinting. High "physical_Crosslinking_time" and long-term exposure to Ca^{2+} ions can further reduce cell viability, highlighting the significance of reducing the physical crosslinking duration as well as its amount to minimize cell damage.

However, it is important to note that in alginate-based bioprinting, the absence or improper execution of crosslinking can also lead to reduced cell viability. Proper crosslinking is critical for preserving structural integrity, preventing cell leakage and relocation, and supplying nutrients and

oxygen. Without sufficient crosslinking, the cell-laden structure may lack mechanical integrity, resulting in ineffective cell attachment and arrangement. Moreover, inadequate crosslinking can result in cell leakage, uneven distribution, and restricted nutrient exchange, all contributing to a decline in cell viability. This creates a trade-off, emphasizing the significance of optimizing the crosslinking condition to balance structural integrity and cell viability. Therefore, in order to achieve optimal cell viability and overall success in alginate-based bioprinting, it is essential to ensure appropriate crosslinking conditions.

The "Time of Measurement" parameter is demonstrated to be the 4th important parameter in expected cell viability. Over time, after the printing, cells within the bioprinted structure, depend on their proliferation rate, are likely to grow and repair any damage caused by the bioprinting process. However, after some time, there is a growth capacity threshold beyond which cell viability may plateau or decline. This phenomenon was illustrated in one of our studies described in Chapter 4, where we assessed the viability and proliferation of MDA-MB-231 cells over an 11-day period after bioprinting using live-dead assay and MTT assay. Our results, illustrated in Figure 4.5, indicated that the viability rate of cells was 76±2% on the day of printing (day 0), and the viability rate progressively increased during the first week, reaching 98±1% on day 4 and 99±1% on day 7. Then, from say 7 to day 11, the rate of viability decreased a little and reached 96±2% on day 11[107]. Therefore, the time at which measurements are taken has a significant impact on the observed cell viability and should be carefully considered in accordance with the specific objectives of the study.

From the material perspective, "Alginate_Concentration" and "Gelatin_Concentration" are shown to be the 5th and 8th significant variables in predicting cell viability based on our built neural network. Indeed, the type of hydrogel and its concentration define the bioink property, which is one of the critical factors determining bioprinting success in terms of cell survivability and

printability. Depending on the bioink component, the rheology behavior, such as viscosity and mechanical properties of bioink can change [131]. When bioink viscosity is high, shear stress increases at the point of contact between the bioink and the nozzle wall, leading to cellular damage, such as membrane rupture, cytoskeletal damage, and DNA damage [32]. In addition, bioinks with a greater viscosity or stiffness require higher extrusion pressures to overcome resistance and flow through the nozzle to accomplish proper flow and deposition. For instance, increasing alginate concentration can increase the hydrogel's rigidity and mechanical strength, exposing the cells to greater mechanical stress during extrusion and printing [100]. This increased rigidity can also reduce the diffusion of nutrients and oxygen to the cells, resulting in reduced cell viability and growth for more extended studies after printing. Hence, this result shows the significance of bioink formulation in achieving high viability during and after printing.

3.3.2.3 Bayesian Optimization Model Based on the Built Regression Neural Network

This study utilized a neural network-based Bayesian optimization model to optimize the bioprinting process to obtain the highest possible cell viability. Bayesian optimization is a powerful ML optimization tool that is used to find the optimal set of parameters for a given objective function [127]. This method in extrusion-based bioprinting was used in the study by Ruberu et al. [131] for the first time to optimize the printability of GelMA and GelMA/hyaluronic acid methacrylate bioinks. The researchers utilized Bayesian optimization as a novel method for determining the optimal printing parameters while minimizing the number of necessary experiments. The study demonstrated the promising outcome of this method in speeding up the extrusion bioprinting experimentation process compared to conventional optimization.

Our current work aims to eliminate laborious trial-and-error steps in optimizing cell viability in extrusion-based bioprinting by integrating our built neural network with the Bayesian optimization

67

technique. Inspired by the pioneering efforts of Gau et al. [152], who successfully implemented the ML-Based Bayesian optimization model to revolutionize membrane design, we aimed to significantly accelerate the optimization process in bioprinting process.

Integrating the Bayesian optimization model with a neural network not only overcomes the noise and uncertainty of the viability prediction but also provides us with a model that can predict the optimal bioprinting experimental features for maximizing cell viability. Using the trained neural network to make predictions in the Bayesian model can also improve the accuracy of the predictions by guiding the optimization search with the available data more effectively. This combination can expedite convergence and boost overall optimization system performance.

Analysis of Permutation Importance revealed the significance of the crosslinking conditions on cell viability in the bioprinting process for alginate-and gelatin-based bioinks, which encouraged us to optimize these parameters in this study. Hence, our optimization model focused on finding the optimal "Crosslinking (CaCl2)_Concentration" and "Physical_crosslinking_time" parameters, presuming that all other bioprinting parameters were already predefined.

We did not include the first two important parameters, cell type, and printing pressure, in our optimization model for the following reasons: Firstly, the choice of cell type is typically decided by researchers based on the specific objectives of their project, and it is not a parameter that needs to be optimized using our model. Secondly, in extrusion-based bioprinting, the printing pressure normally cannot be predetermined as it depends on various factors such as the bioink formulation, temperature, and printing conditions. Consequently, the pressure should be adjusted during the printing process, and determining it in advance would not be beneficial. Hence, in this study, we focused on determining optimal crosslinking parameters to optimize cell viability. This method allowed us to determine the optimal combinations for crosslinking parameters along with the

desired bioink formulation and printing settings. Indeed, using this method, we are able to print cell-laden scaffolds with superior cell viability and minimal crosslinking-induced injury.

To execute Bayesian optimization, we first defined the range space for "Crosslinking (CaCl2)_Concentration" and "Physical_crosslinking_time" by determining minimum and maximum values. We set the minimum value to be greater than zero because we intended to evaluate optical crosslinking parameters rather than zero for our particular bioprinting parameters. In addition, we defined the range's maximal value as the highest value observed in our dataset. The Bayesian optimization procedure relied on the neural network we developed for regression. During each iteration of the optimizer, the suggested "Crosslinking (CaCl2)_Concentration" and "Physical_crosslinking_time" with other predefined and desired variables were fed into the trained neural network to predict cell viability with good accuracy. This approach enabled us to rapidly investigate numerous unexplored parameter possibilities without the need for time-taking and costly experiments, ultimately identifying the most efficient parameter combination during printing.

The Gaussian process in this process model the underlying patterns and correlations in the data, providing a probabilistic estimate of the behavior of the cell viability. The bioprinting dataset used for developing the Gaussian process in Bayesian optimization was limited to a smaller subset compared to the data used for training the neural networks. We specifically selected a portion of the dataset that met certain constraints, including: having non-zero concentration values for both gelatin and alginate, using MDA-MB-231 as the cell type, and ensuring non-zero values for the crosslinking parameters. We mainly focused on the most essential data necessary for our specific experiments in the laboratory.

To demonstrate the practical implementation of our model, we designed a process of printing cell-laden structure using MDA-MB-231 using alginate and gelatin hydrogels in the bioink. Initially, we determined the desired concentrations of hydrogels and specific printing settings for two different structures: 1) 4%(w/v) gelatin-3%(w/v) alginate and 2) 3%(w/v) gelatin-7%(w/v) alginate containing 2×10^6 and 1.5×10^6 cells/ml, as outlined in Table 3.4, respectively.

Table 3.4 Predefined values of bioprinting parameters for Experiment 1 and Experiment 2

Bioprinting Parameter	Experiment 1	Experiment 2
Cell type	MDA_MB-231	MDA_MB-231
Gelatin_Concentration(%w/v)	3	5
Alginate_Concentration(%w/v)	2	6
Printing_Pressure(kPa)	23	67
Movement_Speed(mm/s)	15	22
Nozzle_size(µm)	250	410
Cell_number(cells/mL)	2,000,000	1,500,000
Cartridge_Temperature(°C)	22.5	22.5
Bed_Temperature(°C)	22.5	22.5
Time of Measurement(day)	0	0

Using the Bayesian optimization model in combination with our regression neural network, we could identify the optimal "Crosslinking (CaCl2)_Concentration" and "Physical_crosslinking_time" for each structure, represented in Table 3.5, thereby predicting the associated optimal cell viability. To assess the reliability and accuracy of our model, we printed the cell-laden structures and crosslinked the bioprinted construct with the optimal crosslinking conditions identified by the Bayesian optimization in our laboratory. Finally, the actual cell viability was evaluated using live-dead assay. The results of prediction values of cell viability and actual viability for both structures are represented in Table 3.5.

Table 3.5 Predicted optimal crosslinking parameters and cell viability using neural network-based Bayesian optimization model and the actual cell viability of Experiment 1 and Experiment 2.

	Experiment 1	Experiment 2
Optimal Crosslinker(CaCl2)_Concentarion(mM)	376	568
Optimal Physical_Crosslinking_time_(s)	1796	1800
Optimal Predicted Cell Viability (%)	85	79
Actual Cell Viability (%)	80±1.6%	64±2.3%

The findings demonstrate that this approach can successfully improve the bioprinting process by determining the optimum crosslinking condition. There is a good agreement between the actual and predicted value of cell viability for both experiments, showing the promising application of this optimization method for bioprinting. Using this neural network-based Bayesian optimization model, researchers can determine optimal crosslinking conditions for any sets of parameters for alginate-based bioink and optimize the bioprinting without any trial-and-error experiments. Additionally, this optimization method can be extended to predict other bioprinting parameters, such as concentration of materials, printing setting parameters, or combinations of all of them, to optimize cell survival in future studies.

Although there is agreement between the predicted and actual viabilities, there are still differences. This discrepancy can be attributed to the performance of neural network. Indeed, our Bayesian optimization model performance is based on the built regression neural network, whose prediction ability is highly dependent on the dataset's precision, and size. Our neural network is developed based on the limited dataset mainly collected from the literature. Hence, the performance of the optimization model can be further improved by including more precise data collected from different bioprinting laboratories allowing for more accurate prediction of optimal parameter combinations.

Finally, it can be concluded that the application of neural network-based Bayesian optimization provides a novel and efficient method for the reverse design of the bioprinting process for gelatin and alginate-based bioinks. This technique eliminates the need for the typical trial-and-error optimization, paving the way for a more straightforward and efficient optimization technique of the bioprinting process in the future.

Chapter 4 Predicting and Elucidating the Post-printing Behavior of 3D Printed Cancer Cells in Hydrogel Structures by Integrating In-vitro and In-Silico Experiments

4.1   Introduction

One of the burgeoning 3D biofabrication methods is 3D bioprinting, which is widely applied in regenerative medicine and tissue engineering to fabricate complex tissue-mimetic structures.[153] The application of this technique has great potential in personalized therapy with more concentration on controlling drug release, drug screening for cancer treatment, studying possible side effects, and analyzing the metastasis and invasion of tumour cells.[154] 3D bioprinting technique combines cells, biomaterials and controlled motor systems to develop complex 3D structures and has accurate control over the structure's features such as mechanical properties, porosity, permeability and stiffness.[11,12,155] This technique can overcome multiple limitations of the traditional 3D methods by incorporating important aspects of cellular habitat. These aspects include a nonuniform 3D microenvironment similar to the natural extracellular matrix (ECM) of tumours, complex interactions of cells with their neighbouring cells and with local ECM, and complicated diffusion processes of nutrients and oxygen.[7,9,156] Thus, this method can be used to better represent insights into cell growth mechanisms and provide a closer prediction of in vivo tumour dynamics and cancer cells' response to therapies.[156]

Despite the fast advancement of the 3D bioprinting method, there are some challenges that need to be addressed. Currently, the bioprinting technique is mainly on a trial-and-error basis to achieve the desired output, increasing the need for experimental techniques. This trial-and-error basis includes optimizing bioink properties and its printability, structure mechanical strength, and cell viability during and post-printing. Therefore, it is very costly to optimize bioprinting-related

experiments.[157] These challenges make the experimental design and data gathering more complicated in this process.

In silico methods can be used to complement in vitro experiments and assist in addressing some of the limitations of this 3D method.[158–160] Common in-silico approaches include ML methods and mechanistic modelling. We developed ML models for optimizing cell survival during bioprinting procedure (Chapter 3); in this chapter, we would like to apply mechanistic modelling for simulating embedded cells' activity in 3D printed structure post printing.

Mechanistic modelling concerns describing phenomena via hypotheses for underlying mechanisms. Mechanistic models include stochastic and agent-based models, discrete models, ODEs, and PDEs. Such mechanistic models may provide insight into the mechanistic aspects of cell behaviors in bioprinted structures and are our preferred mathematical technique in this study.

There is increasing research on mechanistic modelling and simulation as part of the 3D bioprinting process. For example, there are multiple mathematical studies predicting the shear stress applied on the bioink and cells in the nozzle;[161,162] predicting the mechanical properties of the final printed structure based on bioink material properties;[163] and simulating the bioink deposition and the ultimate shape of the created 3D structure.[157] These mathematical techniques have tried to simulate the bioprinting process - during and after bioprinting - to investigate the effect of different parameters on various aspects of bioprinting, including cell viability and structural stability.

However, a comprehensive computational study of the post-printing behaviour of cells embedded in 3D bioprinted structures is yet to be reported. Such mathematical studies can provide researchers with more insights into post-printing cell functioning and allow them to predict complex cellular activities and optimize the cellular microenvironment to save money and time by

keeping experimental evaluations to a minimum. This is the first study that aims to develop an integration of in-vitro and in-silico 3D bioprinted breast cancer models to study the post-printing behaviour of a population of breast cancer cells embedded in a 3D bioprinted structure (Figure 4.1). In the in-vitro study, we used the MDA-MB-231 cell line, one of the most aggressive breast cancer cell lines most frequently used in cancer-related studies. To develop the bioink, a mixture of gelatin and alginate is selected as the most common bioink material in 3D bioprinting due to its similar characteristic to native ECM and appropriate physiological and biological properties for bioprinting application [10]. For our in-silico study, we selected cellular automata modeling, in which most of the parameters were picked from experimental data. Comprehensive in-vitro assays were also done to improve the reliability of the model.
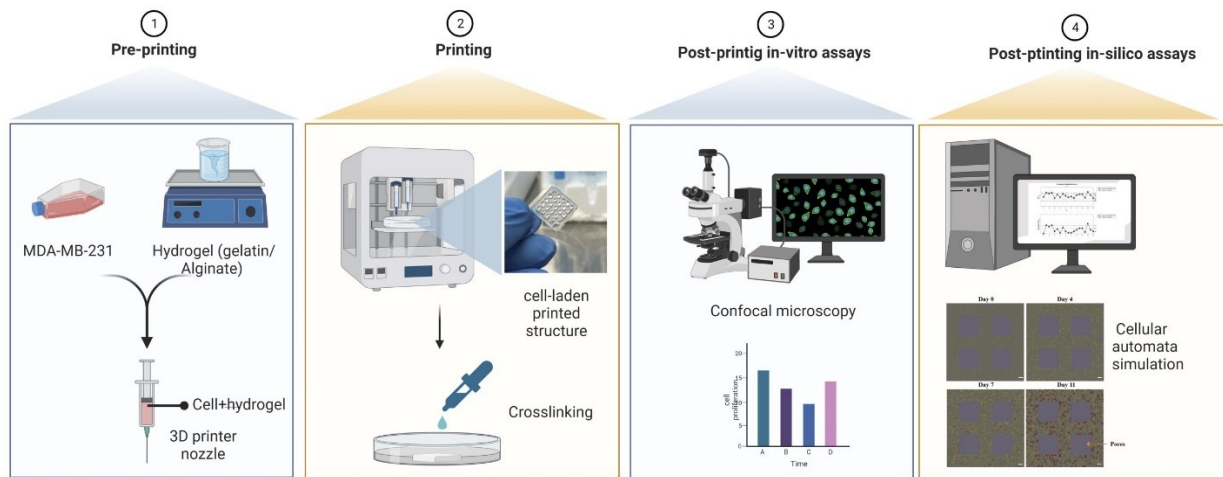
Figure 4.1 Schematic illustration of the main steps involved in this study; step 1: preparing bioink consisting of gelatine, alginate, and MDA-MB-231 cell line; step 2: printing and crosslinking the cell-laden 3D structure; step 3: performing post-printing in-vitro assays; step 4: developing and calibrating an in-silico model. Created with BioRender.com.

The proposed agent-based model in this study is beneficial for demonstrating complex cellular systems, including cellular proliferation, movement, cell interactions with the environment (e.g., ECM, neighbouring cells, and resource consumption) and cell aggregation within the scaffold. In this study, we demonstrate that mathematical models and in-silico simulations can be used to

capture in-vitro dynamics. Such a combination of in-vitro and in-silico studies can improve researchers' understanding of cell activities within 3D bioprinted constructs with the purpose of accelerating and increasing the accuracy of optimization and experimental settings. The proposed in-silico model is very promising and is able to be further developed to be applied in 3D cell culturing using bioprinting techniques in different biomedical applications.

4.2   Materials and Methods

4.2.1   Experimental Methodology

4.2.1.1 Materials

The alginic acid Sodium salt, dimethyl sulfoxide (DMSO), Sodium chloride, $CaCl_2$ and gelatin from bovine skin (type B) were obtained from Sigma-Aldrich (Canada). For cell culture studies, MDA-MB-231 were purchased from ATCC, DMEM, FBS, penicillin/streptomycin, Trypsin/EDTA solution at 0.25% (w./v.) and PBS tablets were bought from Wisent Bioproducts. MTT powder, Triton X-100, BSA and paraformaldehyde were bought from Sigma-Aldrich (Canada). Furthermore, a Live-Dead Cell Viability assay kit (CBA415), Hoechst 33342Nuclei Dye were provided from Sigma-Aldrich (Canada). Additionally, Anti-Ki67 antibody (ab15580), Alexa Fluor 546 Goat anti-Rabbit IgG (H+L) were purchased from Abcam and Invitrogen (Canada), respectively.

4.2.1.2 Rheological Measurement

Rheological measurements are carried out using a Bohlin-CS Rheometer with a parallel plate geometry. The complex viscosity was measured at different frequencies ranging from 0.1 to 100

(Hz). The strain sweep test is performed at room temperature from 0.1 to 100% to investigate the viscoelastic properties of the bioink. All rheological tests are carried out three times.

4.2.1.3 Cell Culture

MDA-MB-231 cells were cultured in DMEM supplemented with 10% FBS and 1% 100 Uml−1penicillin/streptomycin in T-75 flasks. Cells were incubated at 5% CO_2 and 37 °C, and the media was exchanged every other day. When cells reached 80% confluency, the cells were rinsed twice with DPBS and then suspended with trypsin/EDTA (0.25%-1X).

4.2.1.4 Bio-printing

To prepare bioink, according to the protocol provided by Ouyang et al.[164] first 0.5% NaCl solution using DI water was prepared. Then, gelatin and alginic acid sodium salt powder were added, and the solution was stirred vigorously for 1 hr. The prepared solution was sterilized by heating at 70 °C (30 min, three times) and then kept at 4 °C before usage. Prior to bioprinting, MDA-MB-231 suspension was gently and evenly mixed with prepared gelatin/alginate solution using a mixing syringe to make a bioink with final concentrations of 4% (w/v) gelatin, 4% (w/v) alginate and 2-2.5 $\times$ 10^6 MDA-MB-231 cells mL^{-1}.

Bioprinting was carried out using CELLINK INCREDIBLE+ extrusion-based bioprinter. The prepared bio-ink was extruded using a needle to fabricate the 3D cell-hydrogel layered grid constructs. More specifically, the printing nozzle applied for this experiment was a standard conical nozzle (22G), and the moving speed of the needle was adjusted at 5 mms^{-1}. The printing was done at room temperature by applying appropriate pressure. The structure was designed using Solidwork software and sliced in sli3r software to 10 layers with the size of $10 \times 10 \times 3\ mm^3$ with a rectilinear filling pattern. During the printing, we made an effort to maintain the pressure at

the least level to have the minimum damage to cell viability. Subsequently, the cell-laden constructs were immersed in 3% (w/v) sterile calcium chloride solution (20 min) for crosslinking sodium alginate with calcium ions.[10] Then, after being washed with PBS three times, each structure was cultured in DMEM medium containing 10% FBS and 1% penicillin/streptomycin in a 12-well plate and incubated at 5% CO2 and 37 °C for a predetermined period of time. Culture medium was exchanged every other day.

4.2.1.5 MTT Assay

Cell proliferation within the 3D network was assayed using MTT assay. After predetermined days of incubation, the medium in each well of a 12-well plate containing a 3D structure was removed, and 900 μL fresh media and 100 μL of MTT solution (0.5 mg mL^{-1} in PBS) was added into each well, and the plate was incubated for 4 hr at 37°C in a CO_2 incubator. The control was a 3D structure without any cells. Then, after aspirating the medium, 1 mL of DMSO was added into each well to dissolve formed formazan crystals for 30 min at 37°C in a CO_2 incubator. In the end, the intensity of the solubilized formazan crystals was recorded using the Absorbance Microplate Reader at 540 nm.

4.2.1.6 Live-Dead Assay

The 3D bioprinted cell-laden constructs were stained at specific time points using Live-Dead staining viability kit based on the manufacturer's instructions to determine cell viability. Briefly, each construct was washed in PBS three times. Then, 1 μM CAM and 2 μM PI were applied to stain cells while incubated in darkness. A laser scanning confocal microscope (Zeiss LSM 700) was used for imaging the live and dead cells within the 3D constructs at multiple spots. Then,

images were analyzed using ImageJ software. Cell viability was counted by dividing the number of green (living) cells by the total number of cells in each image.

4.2.1.7 Cellular Proliferation

Ki-67 Cell Proliferation Assay is a quantitative technique to evaluate cell proliferation in vitro. Using this method, the Ki-67 protein is applied as a marker for cell proliferation as it is expressed during the active cell cycle (G1, S, G2, and M) but not the stationary phase (G0). For this essay, first, the medium of each well at predetermined timesteps was removed entirely, and each scaffold was washed two times with PBS. Then, cells were become fixed by incubating scaffolds in 4% paraformaldehyde in PBS for 1hr at room temperature. Fixed structures were permeabilized in PBS containing 0.1-0.25% Triton X-100 for 15 min after being washed with PBS two times. Afterward, cells were blocked with 3% BSA/0.1% Triton X-100 in PBS for 1hr before the addition of Anti-Ki67 antibody with a concentration of $5 \, \mu gmL^{-1}$ in PBS/1% BSA and incubation overnight at 4°C. Then, the 3D structures were washed three times for 5 min with PBS/1% BSA prior to the addition of the Goat anti-Rabbit IgG (H+L) Secondary Antibody and incubation for 2 hr. Next, the structures were rewashed (two times, 5 min) with PBS/BSA before adding Hoechst 33342 (incubation of 1 hr). Ultimately, the cells were imaged using the laser scanning confocal microscope.

4.2.2 Computational Methodology

4.2.2.1 Cellular Automata Model Description

This model simulates time as discrete, uniform time steps; each time step is 1 hr, and each simulation lasts for 11 days. A subdomain of the porous cell-laden scaffold fabricated using the 3D bioprinting method is simulated in this model, which comprises a rectangular lattice of 190 ×

190×30 lattice points, symmetrically consisting of four pores with widths of $50 \times 50 \times 30$ lattice points. Each lattice point within the hydrogel can be occupied by a cell or remain vacant, while the grid points in the pores should remain unoccupied, as the cells in the corresponding in-vitro experiments do not move into the pores (Appendix A, Figure A2). A simulation is instantiated by placing a specified initial number of cells at random locations on the lattice within the hydrogel. At each time step, cells behave according to a set of stochastic rules that describe cellular processes such as proliferation, movement, and death. The main algorithm and schedule of processes within each time step in bioprinting are depicted in Figure 4.2. The computational framework builds upon previous theoretical cell population work [165,166].
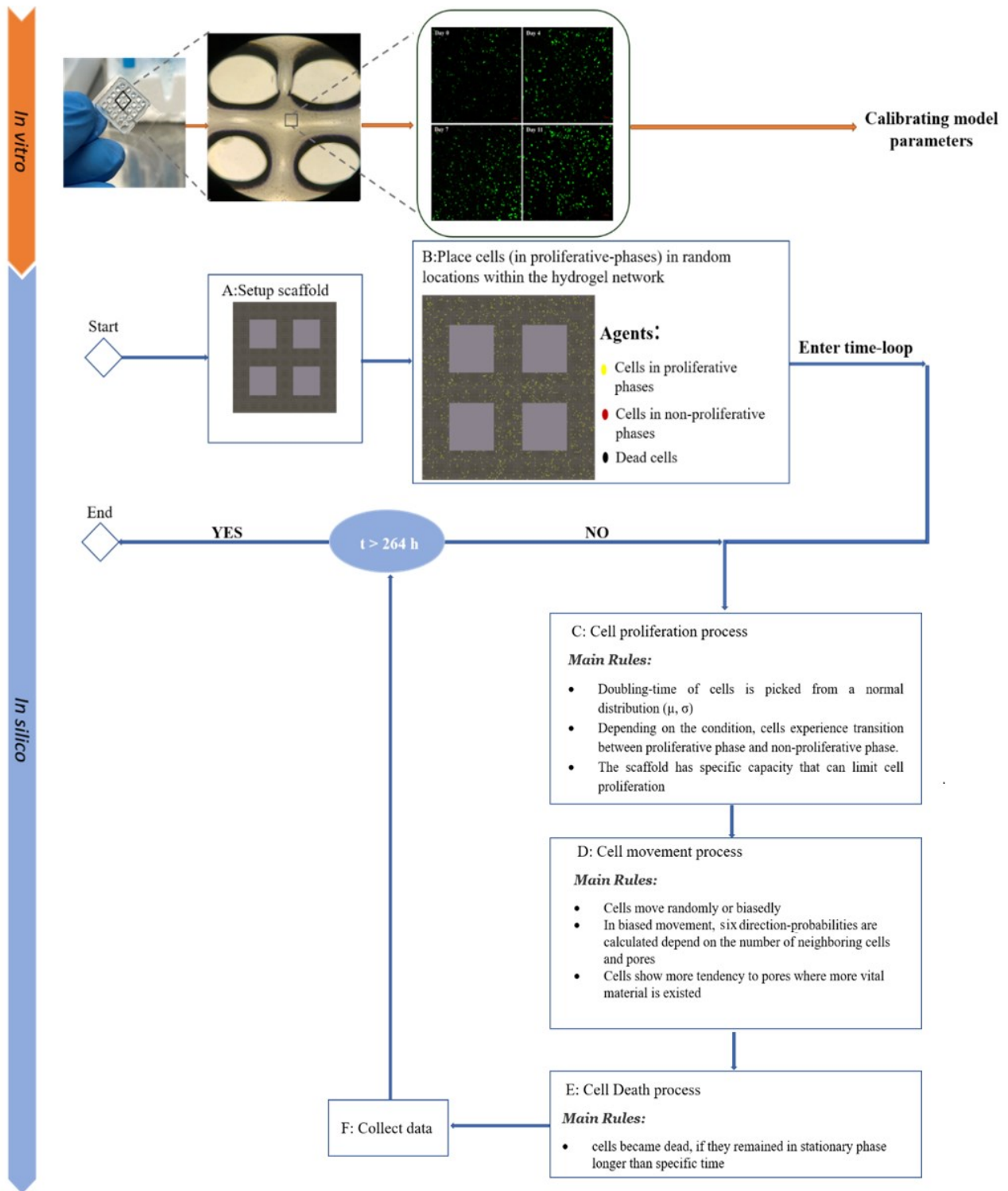
Figure 4.2 Main steps in in-vitro and in-silico bioprinting.

4.2.2.2 Cell Proliferation Process

This process is simulated for each cell individually to consider variations among cells. Individual cells are characterized by a specific, stochastic doubling-time, which attributes to the time it takes for each cell to complete one cell cycle to divide. Based on our experimental data, the doubling-time for each cell is picked from a normal distribution with a value μ=96 hrs and a standard deviation σ=6 hrs. The modeled cell cycle process in proliferation consists of transitions between proliferative and non-proliferative (G0) phases. After dividing, parent cells remain at their current position, and daughter cells can be placed in an unoccupied lattice point in the neighbourhood of the parent. If every neighbouring site is already occupied, parent cells enter the G0-phase, known as the resting or quiescence phase, where cells become inactive.[167] First, second and third-order Moore neighbourhoods are used for placing daughter cells. The scaffold has a specified maximum carrying capacity C to accommodate cells, and upon reaching the total number of cells to this capacity, the proliferation is aborted with a specified probability (P_0). The carrying capacity depends on spatial and nutrient limitations in the in-vitro system. The values of C and P_0 are calibrated according to the in-vitro observations. Due to the unequal distribution of nutrients and oxygen within the scaffold, some cells can still proliferate after reaching the scaffold's carrying capacity, while others enter the G0 phase. This means that cells can continue to proliferate at sites with sufficient nutrients and free neighbouring lattice points. Table S1 contains the values for all in-silico parameters.

4.2.2.3 Movement Process

In the movement process, cells move every specific time known as m_c, however as all the cells might not be synchronized in their movement, a random number was introduced to be added to

m_c. The value of the m_c parameter was also calibrated using in-vitro data. In this process, individual cells can move in a random manner with random-probability, or biased-random manner with biased-probability, and change their positions if there is a free lattice point in their neighbourhood [168]. Cells can move toward one of six directions (right, left, up, down, forward, and backward) of their neighbourhood with equal probability, which is known as random movement, or move in a biased-random manner with weighted probability, where cells are attracted to other cells and the pores in their vicinity, as motivated by empirical observations from the in-vitro experiments. In the biased-random movement, we compute the probability of movement in each direction (direction-p) depending on the number of neighboring cells at that direction of a cell within its range of attraction, defined as (L_c). Also, the probability in each direction can be increased depending on the Euclidean distance between the individual and pores within a specific attraction range known as (L_p). In the random movement case, $P_{\text{up}} = P_{\text{down}} = P_{\text{left}} = P_{\text{right}} = P_{\text{forward}} = P_{\text{backward}} = 1/6$, where 6 is the number of directions, and P_{up}, P_{down}, P_{left} P_{right}, $P_{forward}$, and P_{backward} are the probability of cell movement in the up, down, left and right direction, respectively. In the random-biased case, $P_{\text{up}} = P_1$, $P_{\text{down}} = P_2$, $P_{\text{right}} = P_3$, $P_{\text{left}} = P_4$, $P_{forward} = P_5$, and $P_{backward} = P_6$, with $\sum_{i=1}^{6} P_i = 1$. The probabilities P_i are computed by scanning and summing the neighboring cells and pore lattice points. More details are described in the Appendix B. Finally, each cell moves with more tendency in the direction where higher probability is computed.

4.2.2.4 Death

In the simulations, we assumed that the porous structure of the 3D bioprinted scaffold allowed all cells to access nutrients and oxygen, which prevented death due to the lack of such vital

materials. However, cells lose their viability if they remained in the stationary phase for more than a specific time defined as C_d, calibrated within a range of values, with the probability of P_d. Table S1 contains the values for all in-silico parameters.

4.2.3   Statistical Analysis

Image J software was utilized to analyze the images. All error bars in the figures and reported data indicated ±SD from at least three repeats (n ≥ 3). The results were reported in the format of mean± SD. The reported statistics from confocal microscopy images were obtained by averaging the number of cells in at least five distinct points of each structure.

4.3   Result and Discussion

4.3.1   Rheological Measurements

As previously stated, the ideal bioink for extrusion-based printing should possess suitable rheological characteristics, such as shear-thinning properties. As demonstrated in Figure 4.3, the complex viscosity of the 4%gelatin-4%alginate bioink was assessed as a function of frequency for this purpose, revealing that as the frequency increased from 0.1 to 10 Hz, the complex viscosity dropped, demonstrating its shear-thinning behavior. Hence, during extrusion, when cell-laden bioink is exposed to high shear stress at its interface with the nozzle, it can be easily extruded through the nozzle. This characteristic not only improves the printability of bioink, but also reduces the shear stress-induced loss of cell viability, resulting in greater cell survival.
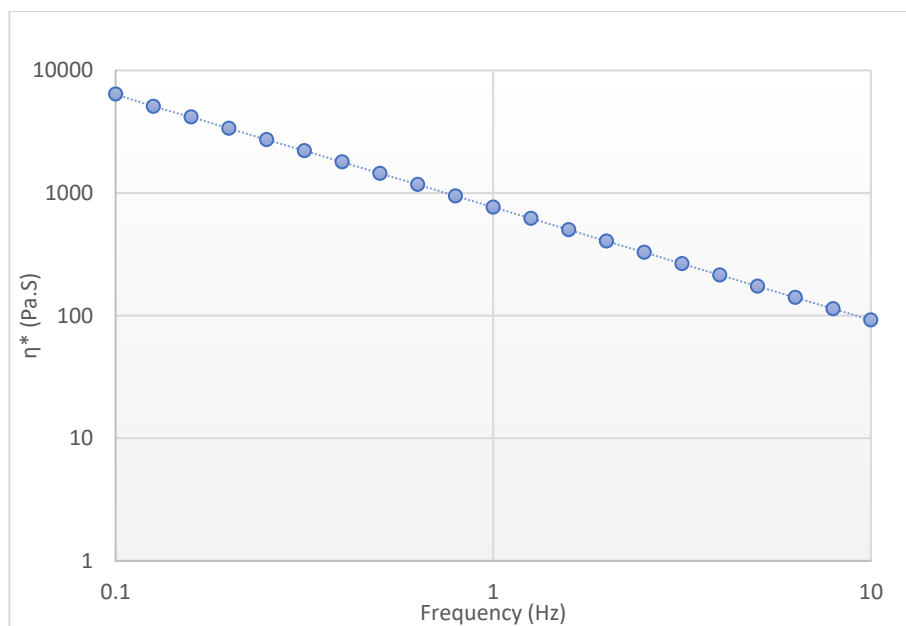
Figure 4.3 The complex viscosity of the 4%gelatin-4%alginate formulation vs. frequency.

4.3.2 In-vitro Cell Studies

The tumour-like hydrogel network model was successfully printed (Figure 4.4 (A)). To determine the proliferation of MDA-MB-231 cells embedded within the hydrogel, MTT assay was used to monitor cellular metabolic activity on days 0, 4, 7, 10, and 11. As illustrated in Figure 4.4 (B), compared to day 0, MDA-MB-231 cells in the 3D cell/hydrogel construct demonstrated 1.86-, 2.7- and 2.78- and 2.8- fold proliferation on days 4, 7, 10, and 11, respectively. Indeed, cells showed rapid proliferation in the first 7 days and almost maintained almost a plateaued cell proliferation from day 7 to 11. Interestingly, the results showed that the doubling time of the MDA-MB-231 encapsulated in the 3D microenvironment was three times higher than that of cells grown in 2D cultures, which can be attributed to the reduced cell activities within the 3D matrix [169]. From day 7 to day 11, the number of proliferating cells remained approximately constant and left a question behind: whether cells died or entered to non-proliferating phase (quiescent) due to

85

undesired conditions. To answer this question, live-dead assay, as well as Ki-67 immunostaining, were further employed during 11-days period to get a better understanding of cell behaviour growth encapsulated in the 3D scaffold.
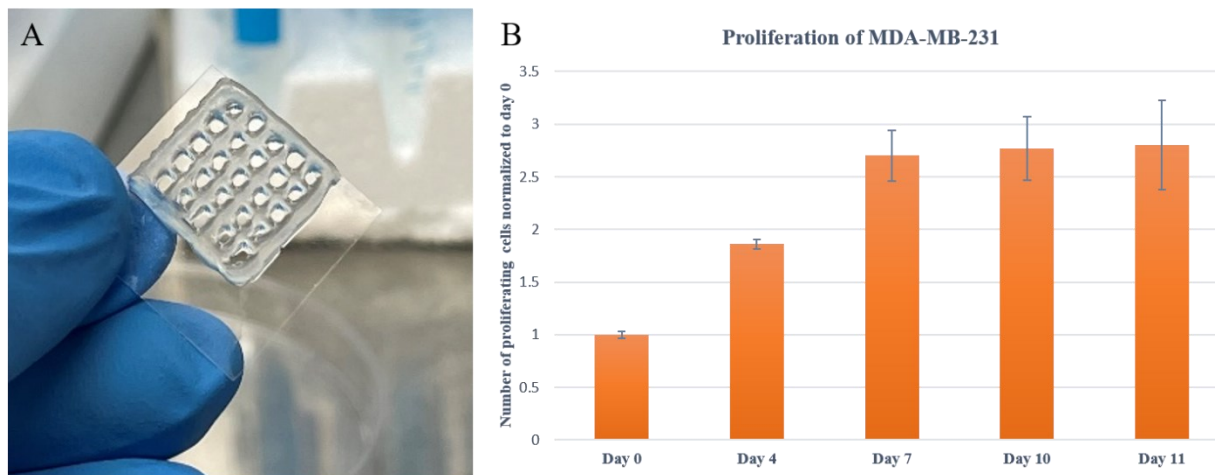

Figure 4.4 (A): Fabricated cell/hydrogel structure using 3D bioprinting technique; (B): MDA-MB-231 proliferation embedded in hydrogel network post printing from day 0 to day 11. Error bars represent ± SD, n ≥ 3.

The viability of MDA-MB-231 over 11 days was visualized using live-dead staining, which was in line with MTT assay results. As shown in Figure 4.5 (A), most cells were viable after printing, and the viability rate of cells was 76±2% on day 0, which clearly demonstrated minor damage of the bioprinting process on cell viability. The rate of viability increased over the first week and reached 98±1% and 99±1% on day 4 and day 7, respectively. Therefore, the structure was porous enough for oxygen and glucose to diffuse and distribute through the hydrogel scaffold, which could provide a proper living environment for cells. From day 7 to day 11, although some cells died, the majority of cells survived, and the rate of viability reached 96±2% on day 11. By comparing the results of live-dead assay and MTT assay, it can be concluded that a significant portion of cells entered the resting phase after seven days since the maximum capacity of the

scaffold had been achieved, and some of the cells began to die during the long-term stationary phase, or due to the lack of resources. Cell death in this experiment is almost negligible, which illustrates the promising potential of gelatin/alginate scaffold for bioprinting applications for a long time. Furthermore, for better visual representation of the printed cells within the three-dimensional structure, the distribution of cells within the hydrogel is depicted in Figure 4.5 (B).
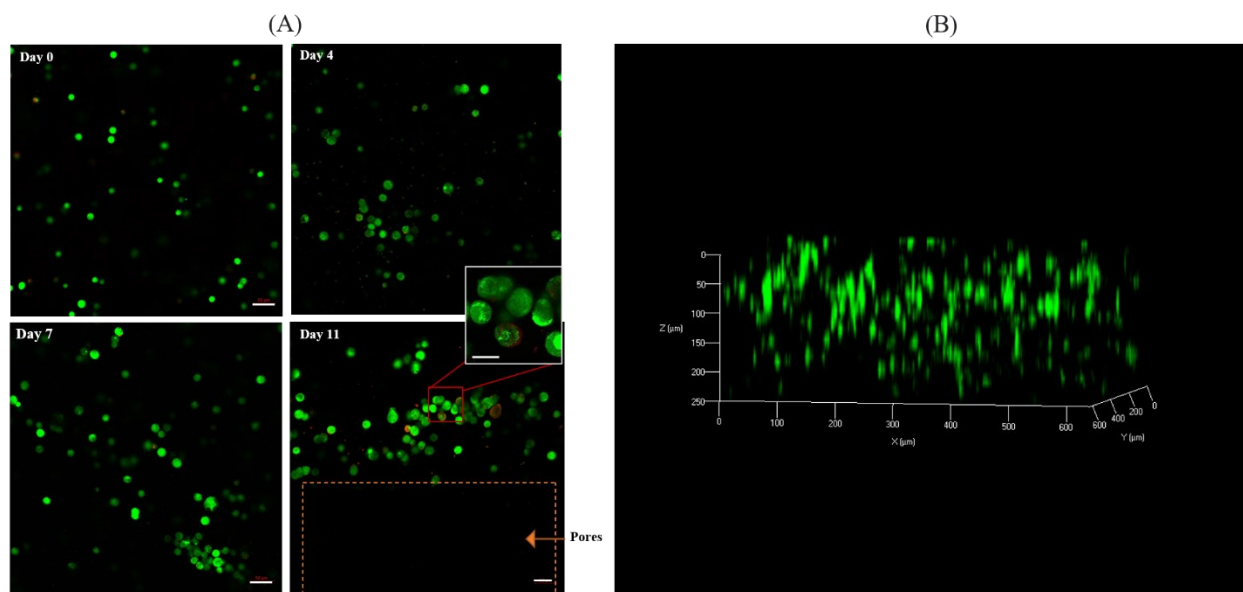

Figure 4.5 (A): Microscopic images demonstrating viability of MDA-MB-231 cells within 3D hydrogel constructs using the fluorescent live-dead assay kit from day 0 to day 11. Live and dead cells were stained using CAM and PI, respectively; (B): Distribution of cells whithin 3D hydogel-based construct on day 0. Green color represents live cells; red color represent dead cells. Cells were imaged using the laser scanning confocal microscope. Scale bar, 50 μm (enlarged images, scale bar, 30 μm).

To visualize the proliferative capacity of MDA-MB-231, cells were fixed, and an anti-Ki-67 antibody was used to image proliferating cells using a confocal microscope (Figure 4.6). Ki-67 is a common-used marker that is present for all active phases of the cell cycle but absent in cells at the stationary phase.[170] The results demonstrated that on days 0 and 4, almost 98±1% and 95±2% of the cells were positive for ki-67, respectively, while this number decreased to 86±2% on day 7, followed by a dramatic drop to about 48.2±2.4 on day 11. This result is in agreement with the data in the previous Figure (Figure 4. 5 (A)), illustrating that within seven days, cells not only survive

but also maintain their proliferating ability. From day 7 to day 11, although a high proportion of cells demonstrated to be alive, they were quiescent and were not able to proliferate anymore because of the lack of enough space for cells to proliferate. Additionally, on days 7 and 11, cells became more aggregated, particularly close to the pores, and cells at the center of the cell aggregates were shown to be non-proliferating due to being surrounded by other cells. Therefore, the proliferation process was aborted as there was not enough space for daughter cells to be placed and due to the lack of nutrients and oxygen at the center of aggregates.

The results of the in-vitro experiments were applied to parameterize the developed mathematical model.
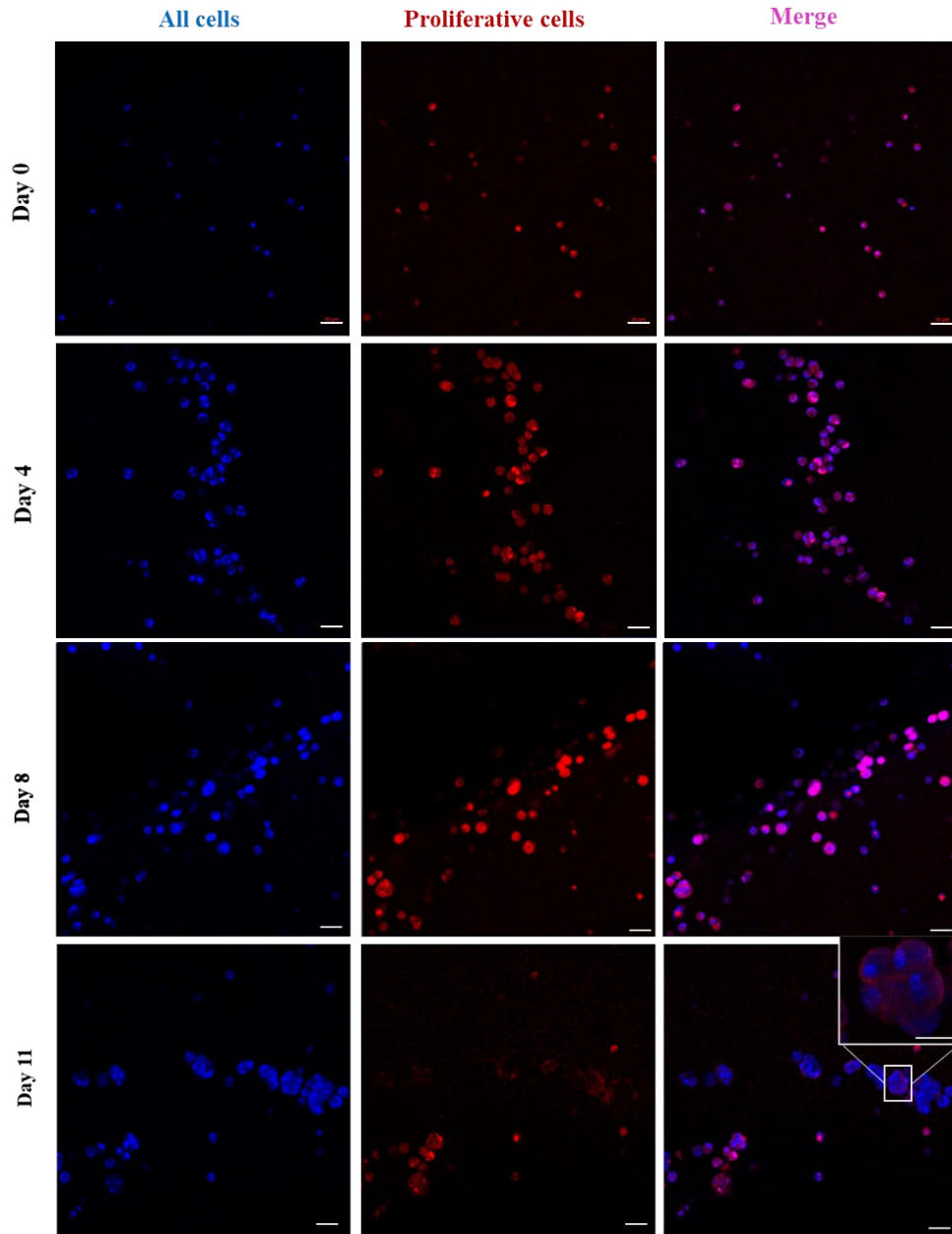
Figure 4.6 Ki-67 staining of encapsulated MDA-MB-231 cells within 3D bioprinted constructs. Cells were stained using anti-Ki-67 antibodies visualized with Alexa Fluor 546, and Hoechst 33342 (red color represents cells positive to ki-67; blue color represents all cells). Cells were imaged using the laser scanning confocal microscope. Scale bar, 50 μm (enlarged image, scale bar, 30 μm).

4.3.3   In-silico cell studies

Individual-based modelling using a cellular automaton is one method to simulate the spatial and temporal mechanism of cell growth at the cellular level [171,172]. This approach is a dynamic system that includes grids of cells, and each cell has sets of discrete states.[101] CA modelling has been widely used in recent years to investigate different cancer cell mechanisms based on a variety of static automaton rules [105]. However, to date, no study has applied CA modelling in 3D culturing cancer cells using 3D bioprinting. We selected this mathematical method for this study on simulating cell growth encapsulated in a 3D bioprinted structure due to its ability to capture spatial properties of the 3D printing structure and its flexibility to explore different hypotheses. Additionally, since the data obtained from our in-vitro experiments contained cells in discrete form, this discrete mathematical technique would have a more accurate simulation of this process. The framework developed in this study represents rules in cellular proliferation, viability, movement, and interactions with the environment, which includes hydrogel and neighbouring cells, as well as cluster formations within the hydrogel network. The in-silico results demonstrated in this article are based on mean values, and standard deviations from n=100 simulation runs, where n is motivated by consistency analysis (Appendix C, Consistency Analysis).

Figure 4.7, showing the cell proliferation patterns for 11 days within the scaffold for both in-vitro and in-silico, illustrates them in agreement with each other. In cell proliferation, the initial cell density and scaffold capacity are two key parameters that we specified as $C_{initial}$ and C variables, respectively. The corresponding values of these parameters were determined using calibration to produce the best fit to the in-vitro studies. Note that, the initial cell density in the developed model represents the effective initial population of cells that can interact with one another within a thin layer, not the total number of cells in the scaffold. Therefore, the simulated

cell density was reduced by a scale factor compared to the cell density in the experimental settings. The results showed that cells reached the maximum cell density of cells in both simulation and experiments after seven days. The time it took for cells to achieve maximum density in the scaffold was strongly dependent on the number of initial cells and the capacity of the printed scaffold. The greater the number of initial cells and the less the scaffold capacity, the sooner cells reached maximum cell density and stopped proliferation. Therefore, by fine-tuning these parameters and running simulations in different scenarios, researchers can design the experiments to achieve desired results without repeating in-vitro assays.
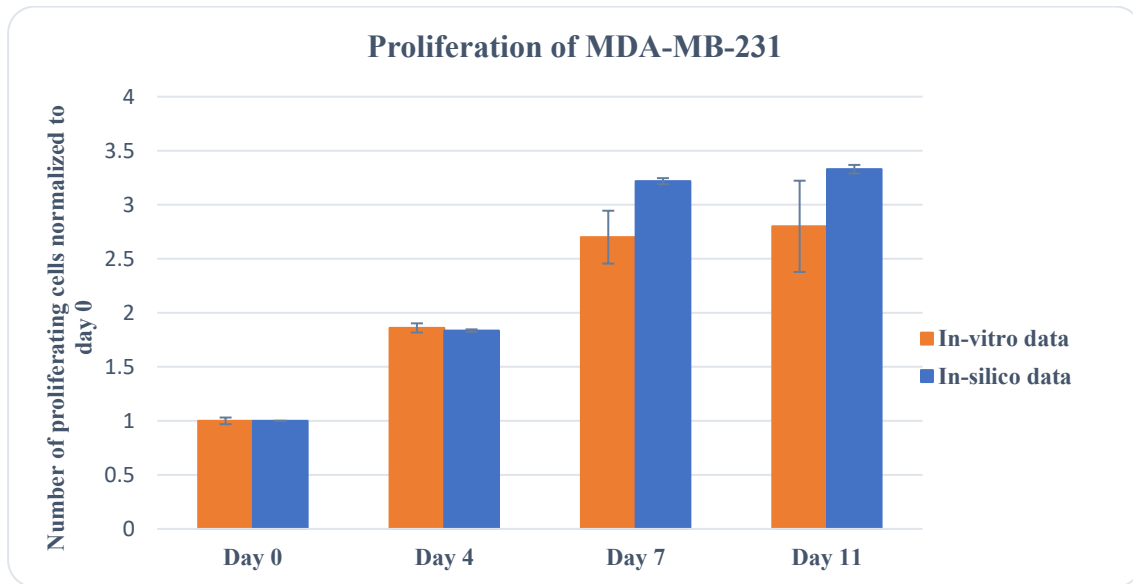
Figure 4.7 Comparison between in-vitro and in-silico results of MDA-MB-231 cell proliferation within the 3D hydrogel network. Error bars represent ± SD, n ≥ 3.

Simulated data were also able to consistently replicate the viability and proliferation experimental results. As explained in the previous section, although cells showed around 99% viability within seven days of printing, the number of dead cells slightly increased from day 7 to day 11 when the significant portion of cells was in the resting phase. Hence, to precisely simulate the in-vitro condition, it was assumed that cells that remained in a prolonged stationary phase for

more than the specified hours, defined by a stochastic number (C_d), started to die with a specified probability (P_d). This observation can be explained biologically by the cells' inability to re-enter the cell cycle after entering the cell stationary phase.

Figure 4.8 shows the snapshots of in-silico MDA-MB-231 cells growing within the hydrogel network; as well as in-vitro microscopic images of cells in the 3D structure. In this figure, you can see cells' distribution and progression of cell cluster formation on days 0, 4, 7 and 11 for both in vitro and in silico. In the in-silico images, yellow, red and black coloured cells are representative of proliferating, non-proliferating and dead cells, respectively.

On day 0, a few cells were distributed inside the hydrogel network. Over time, cells proliferated and created the first two-cell clusters and then bigger ones. Similar to in-vitro observations, the percentage of viability remained around 100% until day 11, when the viability slightly decreased and reached 93.74±0.5%. Furthermore, simulated cell proliferation decreased over time and after seven days experienced a significant drop due to achieving the maximum capacity of the scaffold; and finally dwindled to 54.14±0.25% on day 11.

Another important factor apart from cell viability and proliferation is the ability of the cells to move in their surrounding matrix. This simulation can also be applied to analyze the cell movement as well as the structure and distribution of formed tumour clusters in the hydrogel network without experimental assessment. Tumour clusters might be created due to interactions between neighboring cells or between parent and daughter cells, depending on their position and the microenvironment [173]. Indeed, cells coordinate through cell-cell physical and signalling interactions and create clusters.

In Figure A1 (Appendix A), cells show a trend of crawling toward scaffold pores followed by forming clusters around those pores, as essential resources are in greater concentration there,

particularly after seven days. This fact suggests that the hydrogel networks had limited resources transport capacity. Thus, we have defined particular ranges of attraction for both cell-cell signalling (L_C) and cell-pore attraction (L_p) to consider different cell migration directions. Speed of movement was another important parameter affecting cluster formation. Fallica et al.[169] illustrated that the movement of cancer cells is inhibited in 3D microenvironments and shows extremely low speed due to the combination of material rigidity and the anchoring of cell receptors. Therefore, based on the observations in this and previous studies, we have calibrated cell movement speed. In the movement processes, each individual moved at a specific time defined as m_C in a direction determined based on cell attraction. During the process of calibration, by comparing in-vitro and in-silico results, it was concluded that cells had less tendency to move toward neighbouring cells within a smaller attraction range (L_C) compared to the surface/pores of the scaffold (L_p). Applying these rules in the model, cells mimic the in-vitro cellular behaviour in terms of movement and cluster formation. The results of the proposed model are consistent with those of previous studies [169,174].
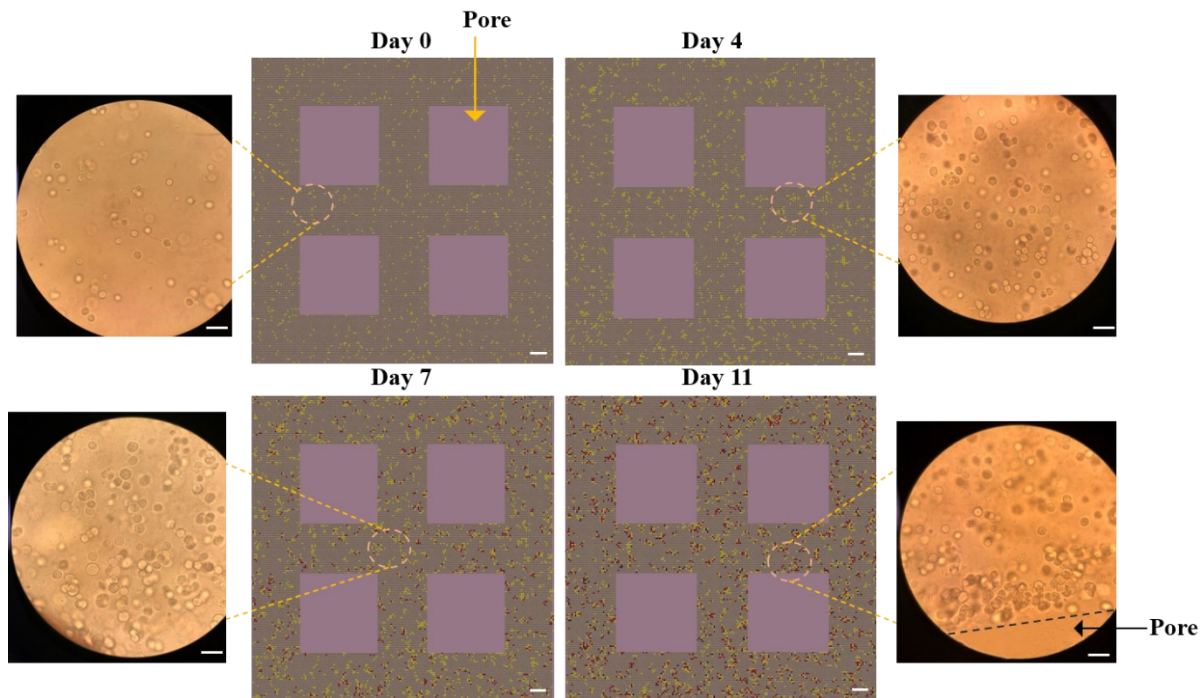
Figure 4.8 Middle panels visualize the MDA-MB-321 growth within the 3D hydrogel construct in silico; yellow represents proliferating cells; red represents non-proliferating cells; black represents dead cells. The right and left panels represent MDA-MB-231 cells encapsulated in 3D hydrogel constructs observed by a phase-contrast microscope on day 0, day 4, day 7, and day 11: scale bar, 50 μm.

In general, this model is developed to combine with in-vitro 3D-bioprinting evaluations, leading to a comprehensive analysis of the whole 3D fabricated structures. One of the main applications of this simulation is to predict the post-printing cellular behaviour in an unpracticed microenvironment which improves its capability to replicate desired biological settings. For example, this model provides the opportunity to evaluate the impact of different important parameters such as various initial cell densities on cellular behaviour in a long-term period. This can be of benefit to researchers to generate a more suitable microenvironment for cell growth without the need to repeat experiments. For instance, they can design the scaffold in terms of size or structural shape with the purpose of modifying the scaffold capacity to improve cell proliferation and decrease cell death.

4.3.4   In-silico Model Validation

To further validate the in-silico model, we performed the bioprinting procedures with different experimental variables in two different situations: case 1: varying initial cell densities; case 2: varying bioink formulation. In case 1, we did bioprinting with a bioink with 4% (w/v) gelatin, 4% (w/v) alginate, and 1.5×10^6 MDA-MB-231 cells mL^{-1}. In case 2: we performed bioprinting experiments with a bioink with final concentrations of 4% (w/v) gelatin, 5% (w/v) alginate, and 2×10^6 MDA-MB-231 cells mL^{-1}. Using the calibrated in-silico model, we would like to predict the proliferation pattern of cells in these two new conditions.

Figure 4.9, showing the cell proliferation patterns for 10 days for case 1 for both in-vitro and in-silico, illustrates them in agreement with each other. In the in-silico model, all parameters except C_{initial} (the initial cell density) have the same values as in the calibrated model. The simulated cell density is reduced by the same scale factor compared to the cell density in the experimental settings and set to $C_{\text{initial}} = 2000$. Both simulations and experiments demonstrated that cells did not achieve the maximal cell density after 7 days and kept growing. Hence, when the number of beginning cells reduced, the later cells attained their scaffold capacity and consequently ceased proliferating.
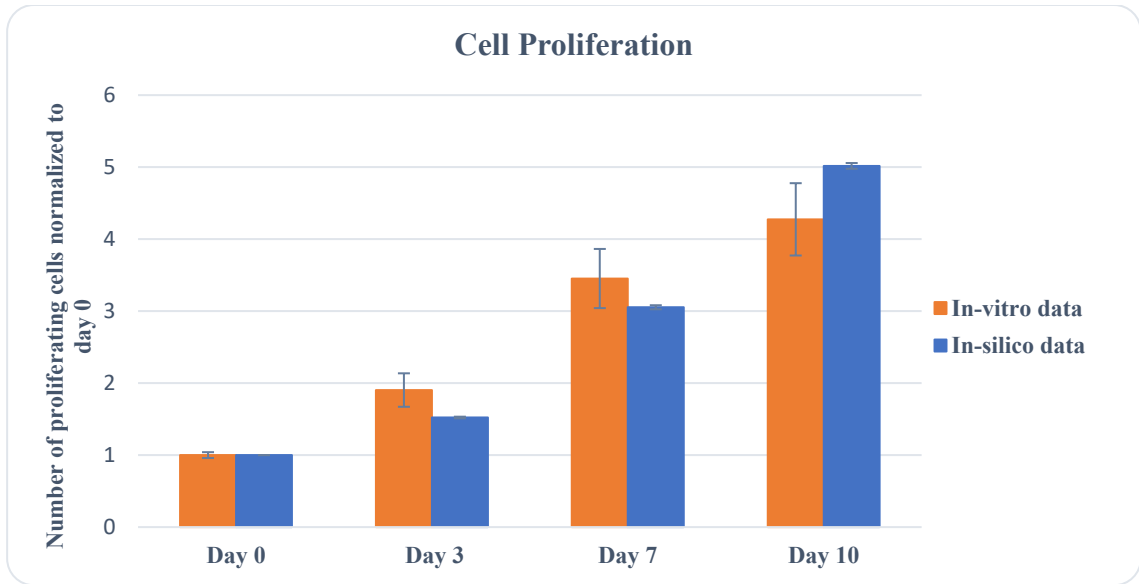
Figure 4.9 Prediction of cell proliferation using in-silico model for case 1: 1.5×10^6 MDA-MB-231 cells mL^{-1} in 4% gelatin/4% alginate bioink; on day 0, 3, 7, and 10. In vitro study was performed using the MTT assay. Error bars represent $\pm$ SD, n $\geq$ 3

Figure 4.10, comparing in-vitro and in-silico cell proliferation patterns for case 2, also shows agreement. In this case, experimentally, we altered the formulation of bioink. Increasing the alginate concentration can increase the rigidity of a hydrogel-based construction, as demonstrated earlier [146]. It has also been found that the stiffness of the microenvironment would also affect cell movement and spheroid formation within the scaffold [168]. Although parameters directly related to stiffness have not yet been integrated into our model, we may regulate cellular behavior and investigate the impacts of bioink formulation and stiffness on proliferation and migration by varying some rules of cellular movement. In the calibrated model for bioink containing 4% (w/v) gelatin and 4% (w/v) alginate, cells move every 15 hours, denoted by m_c. Thus, with 4% (w/v) gelatin, 5% (w/v) alginate-based bioink, we reduce the movement speed of cells encapsulated in a stiffer microenvironment and change m_c to 20 hours while keeping other parameters unchanged. Comparing the results of an in-silico model to in-vitro data, we conclude that for 4% (w/v) gelatin and 5% (w/v) alginate-based bioink, m_c =20 hours closely match the in vitro proliferation trend within 11 days.

The in vitro observations revealed that on day 11, cell proliferation decreased slightly, which can be explained by the stiffer microenvironment. Indeed, the rigidity might reduce cell movement and proliferation; and hinder the transport of nutrients, resulting in cell death over time. For more significant changes in bioink formulation, it is necessary to incorporate microenvironment stiffness or bioink-related parameters into the model to anticipate cell behaviour accurately. However, with 4% (w/v) gelatin, 5% (w/v) alginate, and minor modifications to the bioink formulation, our developed model can be used successfully.
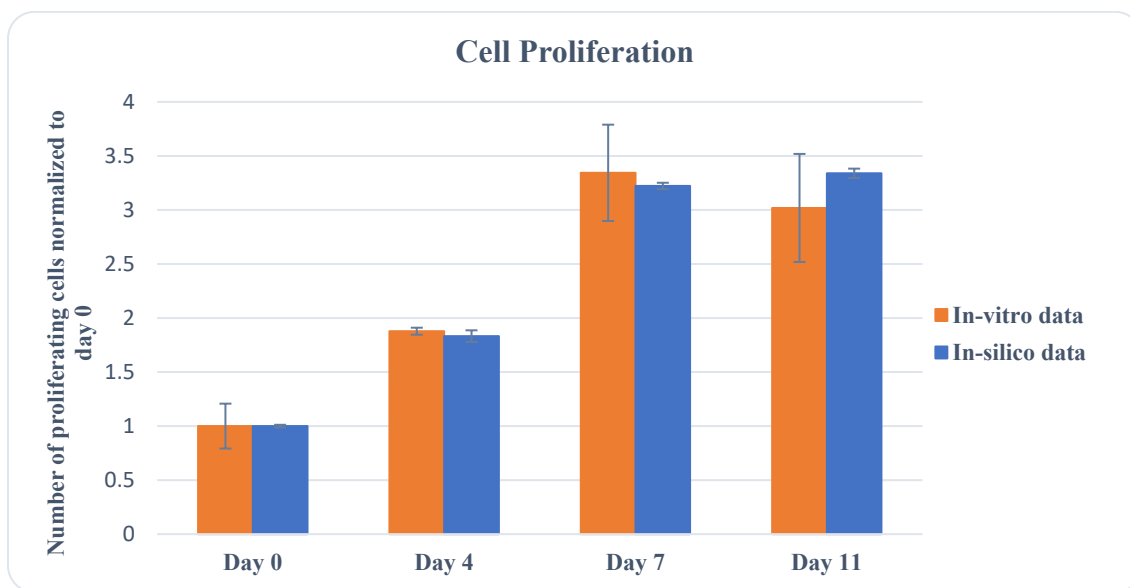
Figure 4.10 Prediction of cell proliferation using in-silico model for case 2: 2×10^6 cells mL^{-1} in 4% gelatin/5% alginate bioink. In vitro study was performed using the MTT assay. Error bars represent ± SD, n ≥ 3

Taken together, we could validate our model by creating variations in the in vitro data and successfully simulating the varied situation. Hence, we can confidently claim that this model can help researchers plan experiments more accurately by predicting the outcome. In fact, researchers executing simulations under various situations and fine-tuning related parameters can design experiments to reach the desired results without repeating in-vitro procedures.

Chapter 5 Conclusions and Future Work

5.1 Conclusion

Studying and optimizing cell behavior, such as cell viability and growth, and understanding how different variables in the bioprinting process affect cell functioning during and after printing are essential for the success of the bioprinting process. In this study, we attempted to understand cell behaviour in this process better and optimize the bioprinting procedure by integrating computational techniques and in-vitro bioprinting studies to eliminate the need for costly and time-consuming trial-and-error methods.

In Chapter 3, we developed a novel cell viability optimization process during the bioprinting process through the application of ML techniques. To do this, we combined the 92 datapoints generated in our laboratory with those collected from other laboratories to create a larger dataset with 713 data points. Using the created dataset, we successfully created regression and classification neural network models to predict cell viability for gelatin and alginate-based bioinks. The trained neural network for regression in our study yielded regression R^2 value of 0.71 and classification accuracy of 0.86. In conclusion, the neural networks represent great performance and a significant improvement over the previous studies using other ML techniques conducted so far.

Through the neural network for regression and Permutation Importance, we identified the bioprinting parameters significantly impacting cell viability prediction. Among different parameters, "cell type" emerges as the most critical variable, highlighting different sensitivities of various cell types to the bioprinting procedure. In addition, "extrusion pressure" is identified as the second most significant parameter, demonstrating the detrimental impact of

excessive pressure on cell viability due to mechanical stress and shear forces on the cell membrane. After the bioprinting procedure, the "crosslinker (CaCl2) concentration" and "physical crosslinking time" are identified as the third and fifth significant features, respectively, which balance the structural integrity and cell viability of bioprinted structures. Therefor, we can conclude that tuning these effective parameters can highly impact the survival of cells during the bioprinting procedure.

In this study, we are pioneers in conducting a computational investigation to explore the profound influence of cell type on cell viability in bioprinting. Our dataset not only includes cancer cell lines but also encompasses diverse cell types, including mesenchymal stem cells, allowing this model to be utilized not only in cancer research but also in various tissue engineering applications.

We finally developed a novel Bayesian optimization model based on the created trained neural network to inversely predict optimal bioprinting crosslinking parameters achieving the highest cell viability without any trial-and-error experiments. By integrating the synergistic capabilities of our neural network and Bayesian optimization model, we achieved quicker and more efficient optimization that surpasses the limitation of traditional optimization approaches. In addition, we validated the performance of our Bayesian optimization model by conducting two distinct laboratory experiments. These experiments showed good agreement between the final predicted and actual cell viability, demonstrating the promising potential of this optimization technique for determining optimal bioprinting parameters.

In Chapter 4, we developed a CA model to simulate post-printing cells' dynamics within the 3D bioprinted construct. To this end, we first successfully printed MDA-MB-231/gelatin/alginate bioink and evaluated cellular behavior in 11 days using MTT, Live-dead, and

Ki-67 cell proliferation assays. Using in-vitro results, we defined rules in the CA model for cell proliferation, viability, movement and cluster formation within the 3D hydrogel network and calibrated model parameters such as doubling-time, movement speed, and probability of death in 11 days.

Our model could quantitatively capture the post-printing in-vitro behavior of cells in the 3D scaffold and predict and elucidate the cell behavior for different bioprinting conditions. For example, it replicated the cellular movement and cell crawling toward the pores, followed by forming clusters after seven days due to uneven distribution of nutrients and oxygen. Furthermore, the in-silico data elucidated the dependence of cell proliferation on the initial number of cells and the capacity of the printed hydrogel network and could predict post-printing cellular proliferation based on the initial quantity of cells in bioink. This in-silico model could also represent cell activity using various network formulations containing gelatin and alginate. The proposed mathematical framework can be of benefit to researchers to generate a more suitable microenvironment for cell growth without the need to repeat experiments.

In the end, we believe that the encouraging outcomes of our study and our developed mathematical and computational frameworks provide a solid foundation for advancing bioprinting applications.

5.2 Future Work

In addition to the accomplishments and contributions outlined in this study, several promising directions for the future can enhance the performance and broaden the applicability of our research:

1- By expanding the available dataset with additional literature and experimental data, with diverse values for each parameter, such as varying hydrogels, cell types, printing

settings, and crosslinking conditions, we can further improve the accuracy of our predictions of both neural networks and the Bayesian optimization model.

2- The inclusion of additional parameters, such as the geography of the printed structure and the thickness, in the dataset for each set of data is likely to improve the accuracy of cell viability prediction.

3- After modifying the neural network, backpropagation optimization is another powerful ML technique that can maximize cell viability by tuning bioprinting parameters. Backpropagation can be used in this method to efficiently compute gradients that indicate the sensitivity of cell viability to parameter changes. These gradients lead the optimization process to search for parameter values that result in the highest cell viability.

4- To expand the bioprinting dataset and improve model performance, considering data from other types of printing techniques like stereolithography and inkjet-based bioprinting, in addition to extrusion-based bioprinting, can be beneficial. By incorporating diverse printing data, the dataset becomes more comprehensive and representative of various bioprinting scenarios. Subsequently, pretraining the model on this augmented dataset and applying transfer learning to fine-tune the model specifically for extrusion-based bioprinting can enhance its performance and predictive capabilities. This approach leverages the knowledge gained from different printing techniques and adapts it to improve cell viability prediction in the context of extrusion-based bioprinting.

5- Optimizing the printability of structures is a crucial aspect of bioprinting that has traditionally been accomplished through tedious trial-and-error experiments. In order

to improve the predictive capabilities of neural networks in determining successful bioprinting of a cell-laden structures, it is advantageous to expand the dataset with data regarding printability in addition to cell viability. By incorporating printability information, the model can predict cell viability as well as printability based on particular bioprinting parameters. This strategy provides bioprinting laboratories with a broader understanding of bioprinting, allowing them to optimize bioprinting outcomes with more informed decisions.

6- The structural integrity of the bioprinted constructs is another factor playing a vital role in the overall success of the bioprinting process. Additionally, optimal crosslinking parameters can be accomplished more precisely by creating a balance between preserving cell viability and assuring the desired structural integrity of printed constructs. Hence, to enhance the predictive ability of the Bayesian optimization model for predicting the crosslinking parameters, we can include both structural integrity and cell viability in the objective function. By incorporating these variables, we can create a more comprehensive optimization model allowing more effective prediction of crosslinking parameters to preserve cell viability while strengthening the structure.

7- Our neural network-based Bayesian optimization technique has the potential to be expanded to address demands of different bioprinting projects. Using this method, we may explore the inverse prediction of other bioprinting parameters, such as initial cell density, bioink formulation, printer setting parameters, or a combination of all of them to optimize cell viability. This expansion of our Bayesian optimization

model enables a comprehensive optimization framework that accommodates the diverse needs and goals of bioprinting projects.

8- The developed CA framework in this study can be extensively applied in different bioprinting-related studies for various applications such as tissue engineering, oncology, and the pharmaceutical industry by extending its rules and improving its ability to provide an accurate prediction of biological systems. Our CA model can be expanded by incorporating bioink-related parameters such as stiffness and structural integrity, which regulate cellular behavior, including proliferation, migration, and oxygen/nutrient diffusion to the 3D network.[131,175–178]

9- The proposed CA model can be integrated with the machine learning algorithms and provide researchers with this opportunity to predict the temporal or structural effect of the hydrogel network on any desired objectives in the biological system. Furthermore, we can use CA simulation to pre-train the ML algorithm, and then a transfer learning approach can be applied to train the experimental data.

10- Another prospect of our CA model is its application in a heterogeneous environment with multiple cell lines for studying cell-cell interaction and cell-ECM interactions. Besides, by adjusting the rules, this model can be integrated with pharmacokinetic modelling techniques to simulate drug treatment responses in 3D cell culture using 3D bioprinting to help study tumor development and metastasis, drug screening, and other aspects of cancer research.

11- A 3D bioprinted tumor model has the potential to advance cancer therapy research significantly. By successfully printing a 3D model with the MDAMB-231 cell line using gelatin and alginate-based bioink in our laboratory, we can further study the

effect of various cancer treatments on cells in an environment that closely resembles

the ECM. Compared to traditional cell cultures, this model provides a more biologically

accurate representation of actual cancer cells in the human body

References

1. Waks, A. G. & Winer, E. P. Breast Cancer Treatment: A Review. *JAMA - Journal of the American Medical Association* **321**, 288–300 (2019).

2. Padrón, J. M. *et al.* The multilayered postconfluent cell culture as a model for drug screening. *Critical Reviews in Oncology/Hematology* **36**, 141–157 (2000).

3. Jong, B. K. Three-dimensional tissue culture models in cancer biology. *Seminars in Cancer Biology* **15**, 365–377 (2005).

4. Hosoya, H. *et al.* Engineering fibrotic tissue in pancreatic cancer: A novel three-dimensional model to investigate nanoparticle delivery. *Biochemical and Biophysical Research Communications* **419**, 32–37 (2012).

5. Meacham, C. E. & Morrison, S. J. Tumour heterogeneity and cancer cell plasticity. *Nature* vol. 501 328–337 Preprint at https://doi.org/10.1038/nature12624 (2013).

6. Gupta, P. B., Pastushenko, I., Skibinski, A., Blanpain, C. & Kuperwasser, C. Phenotypic Plasticity: Driver of Cancer Initiation, Progression, and Therapy Resistance. *Cell Stem Cell* vol. 24 65–78 Preprint at https://doi.org/10.1016/j.stem.2018.11.011 (2019).

7. Rønnov-Jessen, L., Petersen, O. W. & Bissell, M. J. Cellular changes involved in conversion of normal to malignant breast: Importance of the stromal reaction. *Physiological Reviews* **76**, 69–125 (1996).

8. Ingber, D. E. Can cancer be reversed by engineering the tumor microenvironment? *Seminars in Cancer Biology* vol. 18 356–364 Preprint at https://doi.org/10.1016/j.semcancer.2008.03.016 (2008).

9.   Zhang, S. Beyond the Petri dish. *Nature Biotechnology* **22**, 151–152 (2004).

10.   Jiang, T. *et al.* Directing the Self-Assembly of Tumour Spheroids by Bioprinting Cellular Heterogeneous Models within Alginate/Gelatin Hydrogels. *Sci Rep* **7**, 1–9 (2017).

11.   Ahn, H. J. *et al.* Serial Analysis of Tracheal Restenosis After 3D-Printed Scaffold Implantation: Recruited Inflammatory Cells and Associated Tissue Changes. *Tissue Engineering and Regenerative Medicine* **14**, 631–639 (2017).

12.   Gu, B. K. *et al.* 3-Dimensional Bioprinting for Tissue Engineering Applications. *Biomaterials Research* **20**, 1–8 (2016).

13.   Murphy, S. V. & Atala, A. 3D bioprinting of tissues and organs. *Nature Biotechnology* vol. 32 773–785 Preprint at https://doi.org/10.1038/nbt.2958 (2014).

14.   Guvendiren, M., Molde, J., Soares, R. M. D. & Kohn, J. Designing Biomaterials for 3D Printing. *ACS Biomaterials Science and Engineering* vol. 2 1679–1693 Preprint at https://doi.org/10.1021/acsbiomaterials.6b00121 (2016).

15.   Guillotin, B. *et al.* Laser assisted bioprinting of engineered tissue with high cell density and microscale organization. *Biomaterials* **31**, 7250–7256 (2010).

16.   Knowlton, S., Onal, S., Yu, C. H., Zhao, J. J. & Tasoglu, S. Bioprinting for cancer research. *Trends in Biotechnology* **33**, 504–513 (2015).

17.   Ozbolat, I. T., Peng, W. & Ozbolat, V. Application areas of 3D bioprinting. *Drug Discovery Today* vol. 21 1257–1271 Preprint at https://doi.org/10.1016/j.drudis.2016.04.006 (2016).

18.   Vijayavenkataraman, S., Fuh, J. Y. H. & Lu, W. F. 3D printing and 3D bioprinting in pediatrics. *Bioengineering* **4**, 1–11 (2017).

19.   Vijayavenkataraman, S., Lu, W. F. & Fuh, J. Y. H. 3D bioprinting of skin: A state-of-the-art review on modelling, materials, and processes. *Biofabrication* **8**, (2016).

20.   Pati, F., Gantelius, J. & Svahn, H. A. 3D Bioprinting of Tissue/Organ Models. *Angewandte Chemie International Edition* **55**, 4650–4665 (2016).

21.   Derakhshanfar, S. *et al.* 3D bioprinting for biomedical devices and tissue engineering: A review of recent trends and advances. *Bioactive Materials* **3**, 144–156 (2018).

22.   Moroni, L. *et al.* Biofabrication strategies for 3D in vitro models and regenerative medicine. *Nature Reviews Materials 2018 3:5* **3**, 21–37 (2018).

23.   Clinical Perspectives on 3D Bioprinting Paradigms for Regenerative Medicine. *Regenerative Medicine Frontiers* (2019) doi:10.20900/RMF20190004.

24.   Boularaoui, S., Al Hussein, G., Khan, K. A., Christoforou, N. & Stefanini, C. An overview of extrusion-based bioprinting with a focus on induced shear stress and its effect on cell viability. *Bioprinting* **20**, e00093 (2020).

25.   Murphy, S. V. & Atala, A. 3D bioprinting of tissues and organs. *Nature Biotechnology 2014 32:8* **32**, 773–785 (2014).

26.   White, C. R. & Frangos, J. A. The shear stress of it all: the cell membrane and mechanochemical transduction. *Philosophical Transactions of the Royal Society B: Biological Sciences* **362**, 1459–1467 (2007).

27.   Moehlenbrock, M. J., Price, A. K. & Martin, R. S. Use of microchip-based hydrodynamic focusing to measure the deformation-induced release of ATP from erythrocytes. *Analyst* **131**, 930–937 (2006).

28. Wong, K. H. K., Chan, J. M., Kamm, R. D. & Tien, J. Microfluidic Models of Vascular Functions. *https://doi.org/10.1146/annurev-bioeng-071811-150052* **14**, 205–230 (2012).

29. Zhao, F., Chella, R. & Ma, T. Effects of shear stress on 3-D human mesenchymal stem cell construct development in a perfusion bioreactor system: Experiments and hydrodynamic modeling. *Biotechnology and Bioengineering* **96**, 584–595 (2007).

30. Nair, K. *et al.* Characterization of cell viability during bioprinting processes. *Biotechnol J* **4**, 1168–1177 (2009).

31. Malekpour, A. & Chen, X. Printability and Cell Viability in Extrusion-Based Bioprinting from Experimental, Computational, and Machine Learning Views. *J Funct Biomater* **13**, (2022).

32. Billiet, T., Gevaert, E., De Schryver, T., Cornelissen, M. & Dubruel, P. The 3D printing of gelatin methacrylamide cell-laden tissue-engineered constructs with high cell viability. *Biomaterials* **35**, 49–62 (2014).

33. Ouyang, L., Yao, R., Zhao, Y. & Sun, W. Effect of bioink properties on printability and cell viability for 3D bioplotting of embryonic stem cells. *Biofabrication* **8**, 035020 (2016).

34. Blaeser, A. *et al.* Controlling Shear Stress in 3D Bioprinting is a Key Factor to Balance Printing Resolution and Stem Cell Integrity. *Adv Healthc Mater* **5**, 326–333 (2016).

35. Shi, J. *et al.* Shear stress analysis and its effects on cell viability and cell proliferation in drop-on-demand bioprinting. *Biomed Phys Eng Express* **4**, 045028 (2018).

36. Dababneh, A. B. & Ozbolat, I. T. Bioprinting Technology: A Current State-of-the-Art Review. *Journal of Manufacturing Science and Engineering, Transactions of the ASME* **136**, 1–11 (2014).

37. Smith, I. O., Liu, X. H., Smith, L. A. & Ma, P. X. Nanostructured polymer scaffolds for tissue engineering and regenerative medicine. *Wiley Interdisciplinary Reviews: Nanomedicine and Nanobiotechnology* **1**, 226–236 (2009).

38. Mironov, V., Kasyanov, V., Drake, C. & Markwald, R. R. Organ printing: Promises and challenges. *Regenerative Medicine* **3**, 93–103 (2008).

39. Moon, S. *et al.* Layer by layer three-dimensional tissue epitaxy by cell-laden hydrogel droplets. *Tissue Engineering - Part C: Methods* **16**, 157–166 (2010).

40. Mironov, V., Boland, T., Trusk, T., Forgacs, G. & Markwald, R. R. Organ printing: Computer-aided jet-based 3D tissue engineering. *Trends in Biotechnology* **21**, 157–161 (2003).

41. Hamid, Q. *et al.* Fabrication of three-dimensional scaffolds using precision extrusion deposition with an assisted cooling device. *Biofabrication* **3**, 034109 (2011).

42. Fischbach, C. *et al.* Engineering tumors with 3D scaffolds. *Nature Methods* **4**, 855–860 (2007).

43. Horning, J. L. *et al.* 3-D tumor model for in vitro evaluation of anticancer drugs. *Molecular Pharmaceutics* **5**, 849–862 (2008).

44. Martin, M. D. *et al.* Establishment and quantitative imaging of a 3D lung organotypic model of mammary tumor outgrowth. *Clinical and Experimental Metastasis* **25**, 877–885 (2008).

45. Dababneh, A. B. & Ozbolat, I. T. Bioprinting Technology: A Current State-of-the-Art Review. *Journal of Manufacturing Science and Engineering, Transactions of the ASME* **136**, 1–11 (2014).

46. Fischbach, C. *et al.* Engineering tumors with 3D scaffolds. *Nature Methods* **4**, 855–860 (2007).

47. Han, J. *et al.* Molecular predictors of 3D morphogenesis by breast cancer cell lines in 3D culture. *PLoS Computational Biology* **6**, e1000684 (2010).

48. Kenny, P. A. *et al.* The morphologies of breast cancer cell lines in three-dimensional assays correlate with their profiles of gene expression. *Molecular Oncology* **1**, 84–96 (2007).

49. Weiss, M. S. *et al.* The impact of adhesion peptides within hydrogels on the phenotype and signaling of normal and cancerous mammary epithelial cells. *Biomaterials* **33**, 3548–3559 (2012).

50. Friedrich, J., Ebner, R. & Kunz-Schughart, L. A. Experimental anti-tumor therapy in 3-D: Spheroids - Old hat or new challenge? *International Journal of Radiation Biology* vol. 83 849–871 Preprint at https://doi.org/10.1080/09553000701727531 (2007).

51. Eke, I. & Cordes, N. Dual targeting of EGFR and focal adhesion kinase in 3D grown HNSCC cell cultures. *Radiotherapy and Oncology* **99**, 279–286 (2011).

52. Martin, M. D. *et al.* Establishment and quantitative imaging of a 3D lung organotypic model of mammary tumor outgrowth. *Clinical and Experimental Metastasis* **25**, 877–885 (2008).

53. Cecchini, M. G., Wetterwald, A., van der Pluijm, G. & Thalmann, G. N. Molecular and biological mechanisms of bone metastasis. *EAU Update Series* **3**, 214–226 (2005).

54. Cui, X. *et al.* A mechanistic study on tumour spheroid formation in thermosensitive hydrogels: experiments and mathematical modelling. *RSC Advances* **6**, 73282–73291 (2016).

55. Ruberu, K. *et al.* Coupling machine learning with 3D bioprinting to fast track optimisation of extrusion printing. *Appl Mater Today* **22**, 100914 (2021).

56. Gopinathan, J. & Noh, I. Recent trends in bioinks for 3D printing. *Biomaterials Research 2018 22:1* **22**, 1–15 (2018).

57. Babilotte, J. *et al.* 3D printed polymer–mineral composite biomaterials for bone tissue engineering: Fabrication and characterization. *Journal of Biomedical Materials Research Part B: Applied Biomaterials* **107**, 2579–2595 (2019).

58. O'Connell, C. *et al.* Characterizing Bioinks for Extrusion Bioprinting: Printability and Rheology. *Methods in Molecular Biology* **2140**, 111–133 (2020).

59. Stanton, M. M., Samitier, J. & Sánchez, S. Bioprinting of 3D hydrogels. *Lab Chip* **15**, 3111–3115 (2015).

60. Ghavaminejad, A. *et al.* Crosslinking Strategies for 3D Bioprinting of Polymeric Hydrogels. *Small* **16**, 2002931 (2020).

61. Fonseca, K. B. *et al.* Enzymatic, physicochemical and biological properties of MMP-sensitive alginate hydrogels. *Soft Matter* **9**, 3283–3292 (2013).

62. Sun, J. & Tan, H. Alginate-Based Biomaterials for Regenerative Medicine Applications. *Materials 2013, Vol. 6, Pages 1285-1309* **6**, 1285–1309 (2013).

63. Demirtaş, T. T., Irmak, G. & Gümüşderelioğlu, M. A bioprintable form of chitosan hydrogel for bone tissue engineering. *Biofabrication* **9**, 035003 (2017).

64. Xiong, R., Zhang, Z., Chai, W., Huang, Y. & Chrisey, D. B. Freeform drop-on-demand laser printing of 3D alginate and cellular constructs. *Biofabrication* **7**, 045011 (2015).

65. Izadifar, M., Babyn, P., Kelly, M. E., Chapman, D. & Chen, X. Bioprinting Pattern-Dependent Electrical/Mechanical Behavior of Cardiac Alginate Implants: Characterization and Ex Vivo Phase-Contrast Microtomography Assessment. *Tissue Eng Part C Methods* **23**, 548–564 (2017).

66. Kundu, J., Shim, J. H., Jang, J., Kim, S. W. & Cho, D. W. An additive manufacturing-based PCL–alginate–chondrocyte bioprinted scaffold for cartilage tissue engineering. *J Tissue Eng Regen Med* **9**, 1286–1297 (2015).

67. Reina-Romo, E., Papantoniou, I., Bloemen, V. & Geris, L. Computational design of tissue engineering scaffolds. *Handbook of Tissue Engineering Scaffolds: Volume One* 73–92 (2019) doi:10.1016/B978-0-08-102563-5.00004-6.

68. Bersini, S. *et al.* Human in vitro 3D co-culture model to engineer vascularized bone-mimicking tissues combining computational tools and statistical experimental approach. *Biomaterials* **76**, 157–172 (2016).

69. Grant, M. R., Mostov, K. E., Tlsty, T. D. & Hunt, C. A. Simulating Properties of In Vitro Epithelial Cell Morphogenesis. *PLOS Computational Biology* **2**, e129 (2006).

70. Yu, C. & Jiang, J. A Perspective on Using Machine Learning in 3D Bioprinting. *International Journal of Bioprinting* **6**, 4–11 (2020).

71. Ng, W. L., Chan, A., Ong, Y. S. & Chua, C. K. Deep learning for fabrication and maturation of 3D bioprinted tissues and organs. *https://doi.org/10.1080/17452759.2020.1771741* **15**, 340–358 (2020).

72. Caruana, R. & Niculescu-Mizil, A. An empirical comparison of supervised learning algorithms. *ICML 2006 - Proceedings of the 23rd International Conference on Machine Learning* **2006**, 161–168 (2006).

73. Francis, L. Unsupervised Learning. *Predictive Modeling Applications in Actuarial Science: Volume I: Predictive Modeling Techniques* 280–312 (2014) doi:10.1017/CBO9781139342674.012.

74. Shin, J. *et al.* Optimized 3D Bioprinting Technology Based on Machine Learning: A Review of Recent Trends and Advances. *Micromachines 2022, Vol. 13, Page 363* **13**, 363 (2022).

75. Li, C. Deep Reinforcement Learning. *Reinforcement Learning for Cyber-Physical Systems* 125–154 (2019) doi:10.1201/9781351006620-6.

76. Goh, G. D., Sing, S. L. & Yeong, W. Y. *A review on machine learning in 3D printing: applications, potential, and challenges*. *Artificial Intelligence Review* vol. 54 (Springer Netherlands, 2021).

77. Nasteski, V. An overview of the supervised machine learning methods. (2017) doi:10.20544/HORIZONS.B.04.1.17.P05.

78. Wu, M., Phoha, V. V., Moon, Y. B. & Belman, A. K. Detecting malicious defects in 3D printing process using machine learning and image classification. *ASME International Mechanical Engineering Congress and Exposition, Proceedings (IMECE)* **14**, 4–9 (2016).

79. Gobert, C., Reutzel, E. W., Petrich, J., Nassar, A. R. & Phoha, S. Application of supervised machine learning for defect detection during metallic powder bed fusion

additive manufacturing using high resolution imaging. *Additive Manufacturing* **21**, 517–528 (2018).

80.   Bacha, A., Sabry, A. H. & Benhra, J. Fault diagnosis in the field of additive manufacturing (3D printing) using Bayesian Networks. *International journal of online and biomedical engineering* **15**, 110–123 (2019).

81.   Vahabli, E. & Rahmati, S. Application of an RBF neural network for FDM parts' surface roughness prediction for enhancing surface quality. *International Journal of Precision Engineering and Manufacturing* **17**, 1589–1603 (2016).

82.   Pham, G. N., Lee, S. H., Kwon, O. H. & Kwon, K. R. Anti-3DWeapon model detection for Safe 3D printing based on convolutional neural networks and D2 shape distribution. *Symmetry* **10**, (2018).

83.   Ludwig, M., Meyer, G., Tastl, I., Moroney, N. & Gottwals, M. An appearance uniformity metric for 3D printing. *Proceedings - SAP 2018: ACM Symposium on Applied Perception* (2018) doi:10.1145/3225153.3225169.

84.   Olaode, A., Naghdy, G. A., Todd, C. & Naghdy, G. Unsupervised Classification of Images: A Review Biometric gait, motion and fall risk analysis in older people View project Image Retrieval View project Unsupervised Classification of Images: A Review. *International Journal of Image Processing (IJIP)* 325 (2014).

85.   Gu, G. X., Chen, C. T., Richmond, D. J. & Buehler, M. J. Bioinspired hierarchical composite design using machine learning: Simulation, additive manufacturing, and experiment. *Materials Horizons* **5**, 939–945 (2018).

86.	Snell, R. *et al.* Methods for Rapid Pore Classification in Metal Additive Manufacturing. *Jom* **72**, 101–109 (2020).

87.	Jafari-Marandi, R., Khanzadeh, M., Tian, W., Smith, B. & Bian, L. From in-situ monitoring toward high-throughput process control: cost-driven decision-making framework for laser-based additive manufacturing. *Journal of Manufacturing Systems* **51**, 29–41 (2019).

88.	Rana, H., Benoit, N. & Valamanesh, A. Applying Machine Learning for Real Time Optimization of Powder Bed Manufacturing. (2019).

89.	Guvendiren, M., Molde, J., Soares, R. M. D. & Kohn, J. Designing Biomaterials for 3D Printing. *ACS Biomaterials Science and Engineering* vol. 2 1679–1693 Preprint at https://doi.org/10.1021/acsbiomaterials.6b00121 (2016).

90.	Aoyagi, K., Wang, H., Sudo, H. & Chiba, A. Simple method to construct process maps for additive manufacturing using a support vector machine. *Additive Manufacturing* **27**, 353–362 (2019).

91.	Francis, J. & Bian, L. Deep Learning for Distortion Prediction in Laser-Based Additive Manufacturing using Big Data. *Manufacturing Letters* **20**, 10–14 (2019).

92.	Scime, L. & Beuth, J. Using machine learning to identify in-situ melt pool signatures indicative of flaw formation in a laser powder bed fusion additive manufacturing process. *Additive Manufacturing* **25**, 151–165 (2019).

93.	Chung, J. H. Y. *et al.* Bio-ink properties and printability for extrusion printing living cells. *Biomater Sci* **1**, 763–773 (2013).

94. Zhao, Y., Li, Y., Mao, S., Sun, W. & Yao, R. The influence of printing parameters on cell survival rate and printability in microextrusion-based 3D cell printing technology. *Biofabrication* **7**, (2015).

95. Paxton, N. *et al.* Proposal to assess printability of bioinks for extrusion-based bioprinting and evaluation of rheological properties governing bioprintability. *Biofabrication* **9**, (2017).

96. Strauß, S., Meutelet, R., Radosevic, L., Gretzinger, S. & Hubbuch, J. Image analysis as PAT-Tool for use in extrusion-based bioprinting. *Bioprinting* **21**, (2021).

97. Lee, J. *et al.* Machine learning-based design strategy for 3D printable bioink: Elastic modulus and yield stress determine printability. *Biofabrication* **12**, (2020).

98. Bone, J. M. *et al.* Hierarchical Machine Learning for High-Fidelity 3D Printed Biopolymers. *ACS Biomaterials Science and Engineering* **6**, 7021–7031 (2020).

99. Xu, H. *et al.* Prediction of cell viability in dynamic optical projection stereolithography-based bioprinting using machine learning. *J Intell Manuf* **33**, 995–1005 (2022).

100. Tian, S., Stevens, R., McInnes, B. T. & Lewinski, N. A. Machine assisted experimentation of extrusion-based bioprinting systems. *Micromachines (Basel)* **12**, (2021).

101. Hewison, D. & Kuras, M. A New Kind of Science. *Appl Mech Rev* **56**, B17–B33 (2003).

102. Childress, W. M., Rykiel, E. J., Forsythe, W., Li, B. L. & Wu, H. I. Transition rule complexity in grid-based automata models. *Landscape Ecology* **11**, 257–266 (1996).

103. Evsutin, O., Shelupanov, A., Meshcheryakov, R., Bondarenko, D. & Rashchupkina, A. The algorithm of continuous optimization based on the modified cellular automaton. *Symmetry* **8**, (2016).

104. Guyot, Y. *et al.* A computational model for cell/ECM growth on 3D surfaces using the level set method: a bone tissue engineering case study. *Biomech Model Mechanobiol* **13**, 1361–1371 (2014).

105. Patel, A. A., Gawlinski, E. T., Lemieux, S. K. & Gatenby, R. A. A cellular automaton model of early tumor growth and invasion: The effects of native tissue vascularity and increased anaerobic tumor metabolism. *Journal of Theoretical Biology* **213**, 315–331 (2001).

106. Cortesi, M., Liverani, C., Mercatali, L., Ibrahim, T. & Giordano, E. An in-silico study of cancer cell survival and spatial distribution within a 3D microenvironment. *Scientific Reports* **10**, 1–14 (2020).

107. Mohammadrezaei, D. *et al.* Predicting and elucidating the post-printing behavior of 3D printed cancer cells in hydrogel structures by integrating in-vitro and in-silico experiments. *Scientific Reports 2023 13:1* **13**, 1–13 (2023).

108. Ferrari, M., Fornasiero, M. C. & Isetta, A. M. MTT colorimetric assay for testing macrophage cytotoxic activity in vitro. *Journal of Immunological Methods* **131**, 165–172 (1990).

109. Altman, F. P. Tetrazolium Salts and Formazans. *Progress in Histochemistry and Cytochemistry* **9**, III–51 (1976).

110. Freimoser, F. M., Jakob, C. A., Aebi, M. & Tuor, U. The MTT [3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide] assay is a fast and reliable method for colorimetric determination of fungal cell densities. *Applied and Environmental Microbiology* **65**, 3727–3729 (1999).

111. Aung, T. N. *et al.* A new tool for technical standardization of the Ki67 immunohistochemical assay. *Modern Pathology* **34**, 1261–1270 (2021).

112. Braut-Boucher, F. *et al.* A non-isotopic, highly sensitive, fluorimetric, cell-cell adhesion microplate assay using calcein AM-labeled lymphocytes. *Journal of Immunological Methods* **178**, 41–51 (1995).

113. Bulin, A. L., Broekgaarden, M. & Hasan, T. Comprehensive high-throughput image analysis for therapeutic efficacy of architecturally complex heterotypic organoids. *Scientific Reports* **7**, (2017).

114. Cheneler, D. *Viscoelasticity of Polymers: Theory and Numerical Algorithms*. *Applied Rheology* vol. 26 (2016).

115. Desai, M. & Shah, M. An anatomization on breast cancer detection and diagnosis employing multi-layer perceptron neural network (MLP) and Convolutional neural network (CNN). *Clinical eHealth* **4**, 1–11 (2021).

116. Bre, F., Gimenez, J. M. & Fachinotti, V. D. Prediction of wind pressure coefficients on building surfaces using artificial neural networks. *Energy and Buildings* **158**, 1429–1441 (2018).

117. Ting, F. F. & Sim, K. S. Self-regulated multilayer perceptron neural network for breast cancer classification. *Proceeding of 2017 International Conference on Robotics, Automation and Sciences, ICORAS 2017* **2018-March**, 1–5 (2018).

118. Haykin, S. S. *Neural networks and learning machines*. *Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics* vol. 3 (Pearson education Upper Saddle River, 2009).

119. Kingma, D. P. & Welling, M. An introduction to variational autoencoders. *Foundations and Trends in Machine Learning* **12**, 307–392 (2019).

120. Paris, G., Robilliard, D. & Fonlupt, C. Exploring Overfitting in Genetic Programming. 267–277 (2004) doi:10.1007/978-3-540-24621-3_22.

121. Ying, X. An Overview of Overfitting and its Solutions. *Journal of Physics: Conference Series* **1168**, (2019).

122. Demir-Kavuk, O., Kamada, M., Akutsu, T. & Knapp, E. W. Prediction using step-wise L1, L2 regularization and feature selection for small data sets with large number of features. *BMC Bioinformatics* **12**, 1–10 (2011).

123. B, K. H., Johannsen, O., Kondermann, D. & Goldluecke, B. *Computer Vision ACCV 2016*. vol. 10111 (2017).

124. Kumari, S., Tripathy, K. K. & Kumbhar, V. *Data Science and Analytics*. *Data Science and Analytics* (Springer Singapore, 2020). doi:10.1108/9781800438767.

125. Jasper Snoek, Hugo Larochelle, R. P. A. Practical Bayesian Optimization of Machine Learning Algorithms. *Advances in Neural Information Processing Systems 25 (NIPS 2012)* **25**, (2012).

126. Zanjani Foumani, Z., Shishehbor, M., Yousefpour, A. & Bostanabad, R. Multi-fidelity cost-aware Bayesian optimization. *Computer Methods in Applied Mechanics and Engineering* **407**, 1–37 (2023).

127. Frazier, P. I. A Tutorial on Bayesian Optimization. 1–22 (2018).

128. Brochu, E., Cora, V. M. & de Freitas, N. A Tutorial on Bayesian Optimization of Expensive Cost Functions, with Application to Active User Modeling and Hierarchical Reinforcement Learning. (2010).

129. Heimann, P. & Isaacs, S. Regression. *Developments in Psychoanalysis* 169–197 (2018) doi:10.4324/9780429473661.

130. Berk, J., Gupta, S., Rana, S. & Venkatesh, S. Randomised Gaussian process upper confidence bound for Bayesian optimisation. *IJCAI International Joint Conference on Artificial Intelligence* **2021-Janua**, 2284–2290 (2020).

131. Ruberu, K. *et al.* Coupling machine learning with 3D bioprinting to fast track optimisation of extrusion printing. *Applied Materials Today* **22**, 100914 (2021).

132. Mironov, V., Boland, T., Trusk, T., Forgacs, G. & Markwald, R. R. Organ printing: computer-aided jet-based 3D tissue engineering. *Trends Biotechnol* **21**, 157–161 (2003).

133. Mironov, V., Reis, N. & Derby, B. Review: Bioprinting: A Beginning. *https://home.liebertpub.com/ten* **12**, 631–634 (2006).

134. Landers, R., Hübner, U., Schmelzeisen, R. & Mülhaupt, R. Rapid prototyping of scaffolds derived from thermoreversible hydrogels and tailored for applications in tissue engineering. *Biomaterials* **23**, 4437–4447 (2002).

135. Gopinathan, J. & Noh, I. review 5 2018 Recent trends in bioinks for 3D printing.pdf. *Biomater Res* 1–15 (2018).

136. Williams, D., Thayer, P., Martinez, H., Gatenholm, E. & Khademhosseini, A. A perspective on the physical, mechanical and biological specifications of bioinks and the development of functional tissues in 3D bioprinting. *Bioprinting* **9**, 19–36 (2018).

137. Schloßmacher, U. *et al.* Alginate/ silica composite hydrogel as a potential morphogenetically active scaffold for three-dimensional tissue engineering. *RSC Adv* **3**, 11185–11194 (2013).

138. Luo, Y., Lode, A., Akkineni, A. R. & Gelinsky, M. Concentrated gelatin/alginate composites for fabrication of predesigned scaffolds with a favorable cell response by 3D plotting. *RSC Adv* **5**, 43480–43488 (2015).

139. Patel, A. S. *et al.* Encapsulation of angiogenic monocytes using bio-spraying technology. *Integrative Biology* **4**, 628–632 (2012).

140. Gomez-Guillen, M. C., Gimenez, B., Lopez-Caballero, M. E. & Montero, M. P. Functional and bioactive properties of collagen and gelatin from alternative sources: A review. *Food Hydrocoll* **25**, 1813–1827 (2011).

141. Lee, D. H. *et al.* Cellular Orientation on Repeatedly Stretching Gelatin Hydrogels with Supramolecular Cross-Linkers. *Polymers 2019, Vol. 11, Page 2095* **11**, 2095 (2019).

142. An, J., Chua, C. K. & Mironov, V. Application of Machine Learning in 3D Bioprinting: Focus on Development of Big Data and Digital Twin. *Int J Bioprint* **7**, 1–6 (2021).

143. Jiang, J. *et al.* Analysis and prediction of printable bridge length in fused deposition modelling based on back propagation neural network. *https://doi.org/10.1080/17452759.2019.1576010* **14**, 253–266 (2019).

144. Bayraktar, Ö., Uzun, G., Çakiroğlu, R. & Guldas, A. Experimental study on the 3D-printed plastic parts and predicting the mechanical properties using artificial neural networks. *Polym Adv Technol* **28**, 1044–1051 (2017).

145. Mohamed, O. A., Masood, S. H. & Bhowmik, J. L. Influence of processing parameters on creep and recovery behavior of FDM manufactured part using definitive screening design and ANN. *Rapid Prototyp J* **23**, 998–1010 (2017).

146. Gao, T. *et al.* Optimization of gelatin-alginate composite bioink printability using rheological parameters: A systematic approach. *Biofabrication* **10**, 34106 (2018).

147. Markstedt, K. *et al.* 3D bioprinting human chondrocytes with nanocellulose-alginate bioink for cartilage tissue engineering applications. *Biomacromolecules* **16**, 1489–1496 (2015).

148. Zhang, T., Yan, K. C., Ouyang, L. & Sun, W. Mechanical characterization of bioprinted in vitro soft tissue models. *Biofabrication* **5**, 045010 (2013).

149. Ning, L. *et al.* 3D bioprinting of scaffolds with living Schwann cells for potential nerve tissue engineering applications. *Biofabrication* **10**, 035014 (2018).

150. Xu, H. Q., Liu, J. C., Zhang, Z. Y. & Xu, C. X. A review on cell damage, viability, and functionality during 3D bioprinting. *Mil Med Res* **9**, (2022).

151. Li, M., Tian, X., Kozinski, J. A., Chen, X. & Hwang, D. K. MODELING MECHANICAL CELL DAMAGE IN THE BIOPRINTING PROCESS EMPLOYING A CONICAL NEEDLE. *https://doi.org/10.1142/S0219519415500736* **15**, (2015).

152. Gao, H. *et al.* Revolutionizing Membrane Design Using Machine Learning-Bayesian Optimization. *Environ Sci Technol* **56**, 2572–2581 (2022).

153. Yi, H. G. *et al.* Application of 3D bioprinting in the prevention and the therapy for human diseases. *Signal Transduction and Targeted Therapy* **6**, (2021).

154. Sharifi, M. *et al.* 3D bioprinting of engineered breast cancer constructs for personalized and targeted cancer therapy. *Journal of Controlled Release* **333**, 91–106 (2021).

155. Murphy, S. V. & Atala, A. 3D bioprinting of tissues and organs. *Nature Biotechnology* **32**, 773–785 (2014).

156. Ingber, D. E. Can cancer be reversed by engineering the tumor microenvironment? *Semin Cancer Biol.* **18**, 356–364 (2009).

157. Göhl, J. *et al.* Simulations of 3D bioprinting: Predicting bioprintability of nanofibrillar inks. *Biofabrication* **10**, (2018).

158. Reina-Romo, E., Papantoniou, I., Bloemen, V. & Geris, L. *Computational design of tissue engineering scaffolds*. *Handbook of Tissue Engineering Scaffolds: Volume One* (Elsevier Ltd, 2019). doi:10.1016/B978-0-08-102563-5.00004-6.

159. Bersini, S. *et al.* Human in vitro 3D co-culture model to engineer vascularized bone-mimicking tissues combining computational tools and statistical experimental approach. *Biomaterials* **76**, 157–172 (2016).

160. Grant, M. R., Mostov, K. E., Tlsty, T. D. & Hunt, C. A. Simulating properties of in vitro epithelial cell morphogenesis. *PLoS Computational Biology* **2**, 1193–1209 (2006).

161. Müller, M., Öztürk, E., Arlov, Ø., Gatenholm, P. & Zenobi-Wong, M. Alginate Sulfate–Nanocellulose Bioinks for Cartilage Bioprinting Applications. *Annals of Biomedical Engineering* **45**, 210–223 (2017).

162. Leppiniemi, J. *et al.* 3D-Printable Bioactivated Nanocellulose-Alginate Hydrogels. *ACS Applied Materials and Interfaces* **9**, 21959–21970 (2017).

163. Hölzl, K. *et al.* Bioink properties before, during and after 3D bioprinting. *Biofabrication* **8**, (2016).

164. Ouyang, L., Yao, R., Zhao, Y. & Sun, W. Effect of bioink properties on printability and cell viability for 3D bioplotting of embryonic stem cells. *Biofabrication* **8**, (2016).

165. Hamis, S., Stratiev, S. & Powathil, G. G. Uncertainty and sensitivity analyses methods for agent-based mathematical models: An introductory review. *Physics Of Cancer, The: Research Advances* 1–37 (2020) doi:10.1142/9789811223495_0001.

166. Hamis, S., Kohandel, M., Dubois, L. J., Yaromina, A. & Lambin, P. Combining Hypoxia-Activated Prodrugs and Radiotherapy in silico : Impact of Treatment Scheduling and the Intra-Tumoural Oxygen Landscape Keywords. 1–36.

167. Sachlos, E., Czernuszka, J. T., Gogolewski, S. & Dalby, M. Making tissue engineering scaffolds work. Review on the application ofsolid freeform fabrication technology to the production of tissue engineeringscaffolds. *European Cells and Materials* **5**, 29–40 (2003).

168. Cui, X. *et al.* A mechanistic study on tumour spheroid formation in thermosensitive hydrogels: Experiments and mathematical modelling. *RSC Advances* **6**, 73282–73291 (2016).

169. Fallica, B., Maffei, J. S., Villa, S., Makin, G. & Zaman, M. Alteration of Cellular Behavior and Response to PI3K Pathway Inhibition by Culture in 3D Collagen Gels. *PLoS ONE* **7**, 1–11 (2012).

170. Dobos, A. *et al.* Thiol–Gelatin–Norbornene Bioink for Laser-Based High-Definition Bioprinting. *Advanced Healthcare Materials* **9**, 1–9 (2020).

171. Powathil, G. G., Gordon, K. E., Hill, L. A. & Chaplain, M. A. J. Modelling the effects of cell-cycle heterogeneity on the response of a solid tumour to chemotherapy: Biological insights from a hybrid multiscale cellular automaton model. *Journal of Theoretical Biology* **308**, 1–19 (2012).

172. Hamis, S., Powathil, G. G. & Chaplain, M. A. J. Blackboard to Bedside: A Mathematical Modeling Bottom-Up Approach Toward Personalized Cancer Treatments. *JCO Clinical Cancer Informatics* 1–11 (2019) doi:10.1200/cci.18.00068.

173. Wu, Y., Zhao, Z., Guan, Y. & Zhang, Y. Galactosylated reversible hydrogels as scaffold for HepG2 spheroid generation. *Acta Biomaterialia* **10**, 1965–1974 (2014).

174. Wang, X. *et al.* Tumor-like lung cancer model based on 3D bioprinting. *3 Biotech* **8**, 1–9 (2018).

175. Freeman, F. E. & Kelly, D. J. Tuning alginate bioink stiffness and composition for controlled growth factor delivery and to spatially direct MSC Fate within bioprinted tissues. *Scientific Reports* **7**, 1–12 (2017).

176. Zaman, M. H., Kamm, R. D., Matsudaira, P. & Lauffenburger, D. A. Computational model for cell migration in three-dimensional matrices. *Biophysical Journal* **89**, 1389–1397 (2005).

177. Ouyang, L., Yao, R., Zhao, Y. & Sun, W. Effect of bioink properties on printability and cell viability for 3D bioplotting of embryonic stem cells. *Biofabrication* **8**, (2016).

178. Ouyang, L., Highley, C. B., Rodell, C. B., Sun, W. & Burdick, J. A. 3D Printing of Shear-Thinning Hyaluronic Acid Hydrogels with Secondary Cross-Linking. *ACS Biomaterials Science and Engineering* **2**, 1743–1751 (2016).

179. Grimm, V. *et al.* A standard protocol for describing individual-based and agent-based models. *Ecological Modelling* **198**, 115–126 (2006).

180. Hamis, S., Yates, J., Chaplain, M. A. J. & Powathil, G. G. Targeting Cellular DNA Damage Responses in Cancer: An In Vitro-Calibrated Agent-Based Model Simulating Monolayer and Spheroid Treatment Responses to ATR-Inhibiting Drugs. *Bulletin of Mathematical Biology* **83**, 1–21 (2021).

Appendix A

Microscopic Images of Cells Embedded in 3D Constructs

Cells encapsulated in 3D bioprinted structures were monitored by a phase-contrast microscope on days 0, 4, 7 and 11 (Figure A1). Cells show a trend of crawling toward scaffold pores over time, followed by forming clusters.
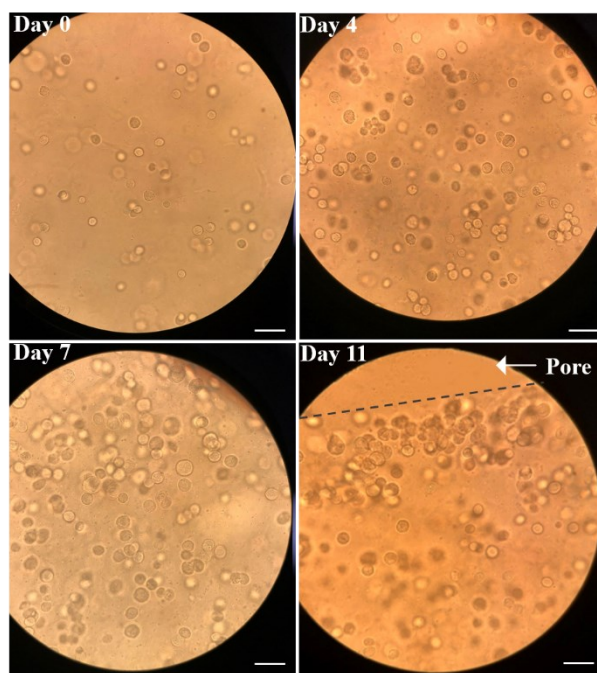

Figure A1. MDA-MB-231 cells encapsulated in 3D bioprinted structures observed by a phase-contrast microscope on days 0, 4, 7 and 11. Scale bar, 50 μm.
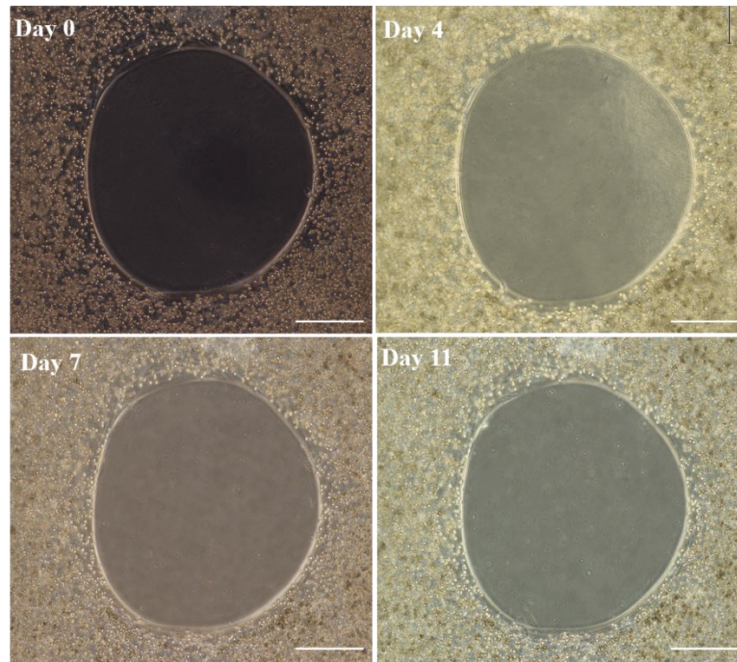
Figure A2. 3D bioprinted constructs observed by a phase-contrast microscope on days 0, 4, 7 and 11. Scale bar, 500 μm.

In addition, as shown in Figure A2, the 3D printed construct could hold its structure and maintain a good form within 11 days. During this period, the pore size remained unchanged, and cells were unable to exit the hydrogel network.

Appendix B

ODD (Overview, Design Concepts, and Details) Protocol of CA Model

To improve our model description with the purpose of ensuring that enough information is communicated to help model understanding and replication, we used the ODD protocol proposed by Grimm et al.[179].

Model description

Overview

Purpose

The purpose of this model is to study the post-printing cellular behaviour of MD-MB-231 cells encapsulated and grown in a 3D hydrogel-based structure fabricated using bioprinting, and investigate the impact of this culture method on cellular proliferation as well as cluster formation over time.

State Variables and scales

This model simulates time as discrete, uniform time steps; each time step is 1 hour, and each simulation lasts for 11 days (264 time steps). This model applies a two-dimensional lattice for the purpose of developing a better simulation of the obtained 2D in-vitro data. A subdomain of the porous cell-laden scaffold fabricated using the 3D bioprinting method is simulated in this model, which comprises a rectangular lattice of $190 \times 190 \times 30$ lattice points, symmetrically consisting of four pores with widths of $50 \times 50 \times 30$ lattice points. Each lattice point spans a volume of $1000 \mu m^3$. Each grid point within the hydrogel can be occupied by a cell or remain vacant, while the grid points in the pores should remain unoccupied. A simulation is initiated by placing a specified initial number of cells (init_c) at random locations on the lattice within the

hydrogel. At each time step, cells behave according to a set of stochastic rules that describe cellular processes such as proliferation, movement, and death. Each cell can be characterized by cell-cycle phase and cell-cycle length. Cell-cycle phases include the proliferative phase (G) and non-proliferative or stationary phase (G0), and cell-cycle length is chosen from a normal distribution with a mean of 96 hours and a standard deviation of 6 hours. The mean and standard deviation of this distribution are selected to match the in-vitro results of this study for MDA-MB-231 cells. An overview of state variables, their default values, and experimentally observed data used for their calibration is illustrated in Table 1.

Table B1. Overview of state variables, default values and experimentally observed data used for their calibration.

Parameter	Value	experimentally observed data used for calibration
The total size of the scaffold	190× 190 × 30 lattice points	Size of 3D bioprinted scaffold
Pore size	50×50× 30 lattice points	Size of 3D bioprinted scaffold
Time step	1 h	-
Initial distribution of cells	Random	-
Cell phases	Proliferative phase (G), non-proliferative phase (G0)	Ki-67 test
Cell-cycle length	Picked from a normal distribution	MTT assay
Mean of cell cycle length distribution (μ)	96 h	MTT assay
The standard deviation of cell cycle length distribution (σ)	6 h	MTT assay
Probability of random movement	0.25	Microscopic images

Probability of biased-random movement	0.75	Microscopic images
m_C	15 h	Microscopic images
s	0 or 1 h	-
C	10,000 cells or lattice points	Proliferation assays
L_C	5 lattice points	Microscopic images
L_p	10 lattice points	Microscopic images
$C_{initial}$	3000 cells	Proliferation assays
P_0	0.5	Proliferation assays
C_d	48 h	Proliferation assays
P_d	0.03	Proliferation assays

For example, as part of the process of calibrating μ of cell cycle, we defined a parameter range based on the MTT assay results obtained from multiple repetitions and samples. The range extended from 85 hours to 130 hours, with a 5-hour increment. Subsequently, we executed the model multiple times, utilising distinct parameter values within the predetermined range. By conducting a thorough analysis of the results, we determined the parameter values that produced the most accurate replication of the in-vitro proliferation pattern. We repeated the same procedure for all calibrated parameters.

Process Overview and Scheduling

Process overview and scheduling are illustrated in Figure 3.2. At each time step, cells proliferate and divide to create a daughter cell. Also, individual cells move randomly or biasedly and change their positions if there is a free lattice point in their neighbourhood. In our study, cell death occurs under specific conditions.

Design Concepts

Emergence

The emergent behaviour in this model exists in the cellular proliferation of cancer cells. Indeed, cancer cells encapsulated in the hydrogel network can show different proliferation rates post-bioprinting depending on the initial number of cultured cells; cell movement (biased-random or random) and the direction they move; the position they plant their daughter cells in their neighbourhood; carrying capacity of the scaffold; and accumulation of vital materials.

Sensing

In the proliferation process, the parent cells sense the presence of other cells in their neighbourhood, and daughter cells can only be placed in an unoccupied lattice point in the neighbourhood of the parental cell. If all lattice points within a third-order Moore neighbourhood of the parental cell is occupied, no daughter cell will be placed, and the parental cell will enter the G0 phase. Additionally, in the biased-random movement toward neighbouring cells, each cell can access the number of cells within its range of attraction and attempts to move toward the direction in which there are more cells.

Interaction

Cells in this model have indirect interactions as the proliferation is assumed to be aborted when the total number of cells reaches the maximum capacity of the scaffold or all neighbouring positions are occupied by other cells. Cells also move towards each other in biased-random movement, and daughter cells are placed in the neighbourhood of parent cells.

Stochasticity

Behavioural stochasticity is represented in this simulation. In the initialization step of the simulation, cells in the active phase (G1) are seeded at random positions. Additionally, for proliferation and movement and death processes, there are associated probabilities that cells make

certain decisions. For example, in cell movement, cells do not always move randomly and sometimes move in a biased-random manner instead. Whether a cell decides to do a biased-random movement or random movement at any time step is simulated stochastically (using parameters called biased probability and random probability). Cells move every $m_c + s$ hours, where s is a random number (0 or 1) in order to avoid the synchronized movement of all cells. The direction of movement is finally selected stochastically depending on the probability computed for each direction. Also, in cell proliferation, each cell is given an individual stochastic doubling-time that is derived from a normal distribution with a determined mean and standard deviation.

Details

Initialization

The values of parameters utilized in initialization are based on the carried out experiments. Random lattice points are initially occupied by the specific initial density of cells in a two-dimensional lattice. For initial values of parameters, look at table 1. Initialization consists of two steps:

Setup scaffold

A scaffold is set in lattice points, such that scaffold[z][n]=1, where there are printed layers of hydrogel, and scaffold[z][n]=0 where there are pores (z and n are representatives of the lattice point coordinates).

Place Cells in Random Locations

Each initial cell is planted randomly in the position (z,n), if and only if scaffold[z][n]=1.

Check if the new location is free or occupied by a cell. If it is free, a cell is planted in a new random location, and the counter of cell density "no_planted_cells" is incremented by one;

otherwise, a different location is tried by redoing the while loop without incrementing the number of planted cells.

After planting each cell at any position, the status of positions is updated by getting changed from 0 to 1. On day 0, the in-vitro results showed that 24±2% of the cells in bioink were damaged and dead due to being exposed to the high pressure during the bioprinting process. However, we did not consider the initial dead cells in the model, and cultured only viable cells on day 0.

Input

A subdomain of the porous cell-laden scaffold fabricated using 3D bioprinting method is simulated in this model, which comprises a square lattice of $190 \times 190 \times 30$ lattice points, symmetrically consisting of four pores with $50 \times 50 \times 30$ lattice points.

Submodels

Within each time step, the following processes are defined in this model.

Cell movement

In this process, each cell can move in a random way or biased-random manner and change its positions if there is a free lattice point in its neighbourhood. The probability of biased-random movement is considered to be 0.75, and the probability of random movement is equal to 0.25. Cells move every 15 hours known as m_c, however as all the cells might not be synchronized in their movement, a random number (0 or 1) was introduced to be added to m_c. The value of the m_c parameter was also calibrated using in-vitro data. In the biased movement, the probability in each direction is computed using the number of neighbouring cells at that direction of a cell within $L_c = 5$ (lattice points) and the Euclidean distance between the individual and pores within $L_p = 10$.

For example, the probability of movement in up direction is computed similar to the following:

For the range of L_p lattice points around each focal cell:

$$X_{up} = \sum_i \frac{h_i}{l_i} W$$

For the range of L_c lattice points around each focal cell:

$$p_{up} = Y$$

Here W is a constant weight, i is sum over L_p lattice points that are above the focal cell, $h_i = \begin{cases} 1, & \text{if lattice point is a pore} \\ 0, & \text{otherwise} \end{cases}$, L_i is the Euclidean distance between the focal cell and lattice point

i. Y is the number of neighboring cells in L_c lattice The value $W = 100$, $L_c = 5$ (lattice points) and $L_p = 10$ (lattice points) are calibrated to consider the empirical observation that cells were more desired to move toward pores where more vital materials were found, in comparison to the movement toward other cells. Finally,

$$P_{up} = (X_{up} + p_{up}) N$$

Indeed, X_{up} and p_{up} are the contribution from pores and cells, respectively, and N is normalization constant. Therefore, we end with $P_{up} + P_{down} + P_{right} + P_{left} + P_{forward} + P_{backward} = 1$, and each focal cell is more likely to move in the direction where a higher directional probability is computed.

Other submodels (Cell Proliferation and Cell Death) are described in the Method section. All the parameters in these submodels are calibrated to best fit the experimental data.

Appendix C

Consistency Analysis

Consistency analysis is a technique that is applied to find how many times the simulation should be run before describing our results in order to minimize the uncertainty originating from inherent model stochasticity. This simulation experiment was conducted based on the study performed by Hamis et al. [180] To do this, we produced five different groups that each contains $k = 20$ distributions of size $n = \{1, 5, 50, 100, 200\}$. We also specified two outputs of interest: X1: the number of living cells on day 5, and X2: the number of proliferating cells on day 5. The purpose of consistency analysis is to determine a distribution size that leads to small statistical significance for these results. In each group, we computed and plotted the maximal Â-measure [165] for $k = 1, 2, ..., 20$ distributions and for both X1 and X2 separately, as is shown in Figures C1, C2, C3, and C4.
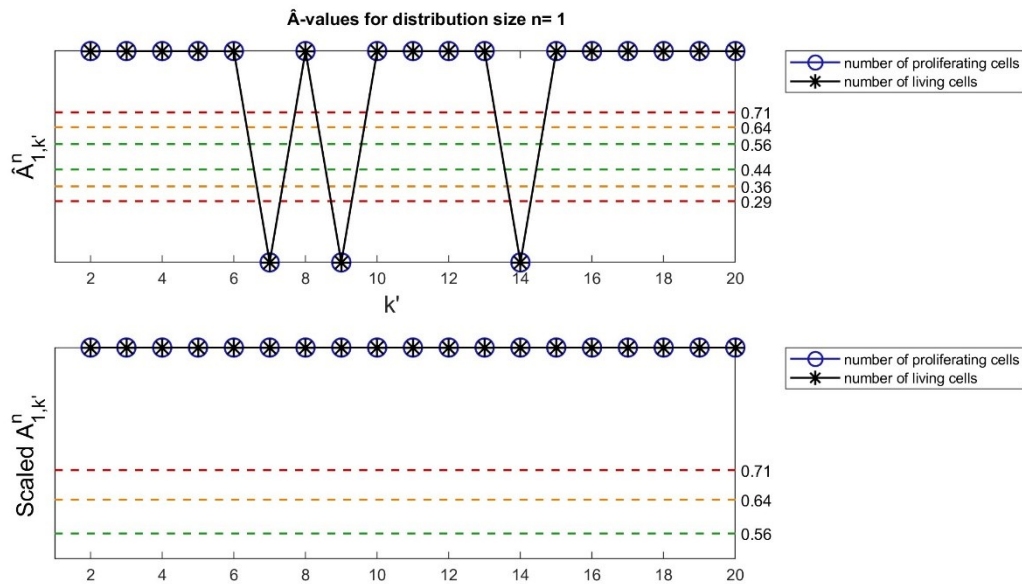

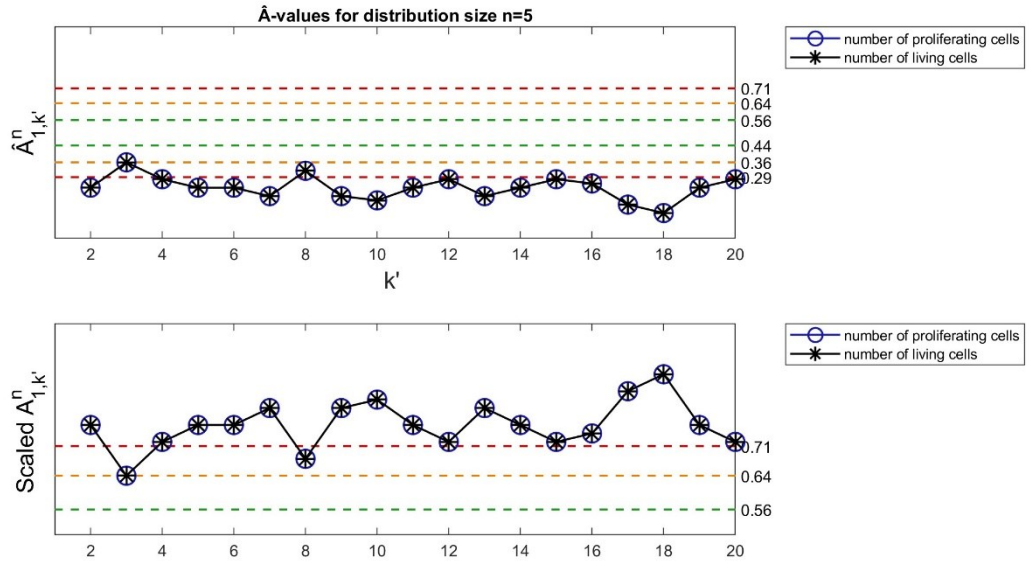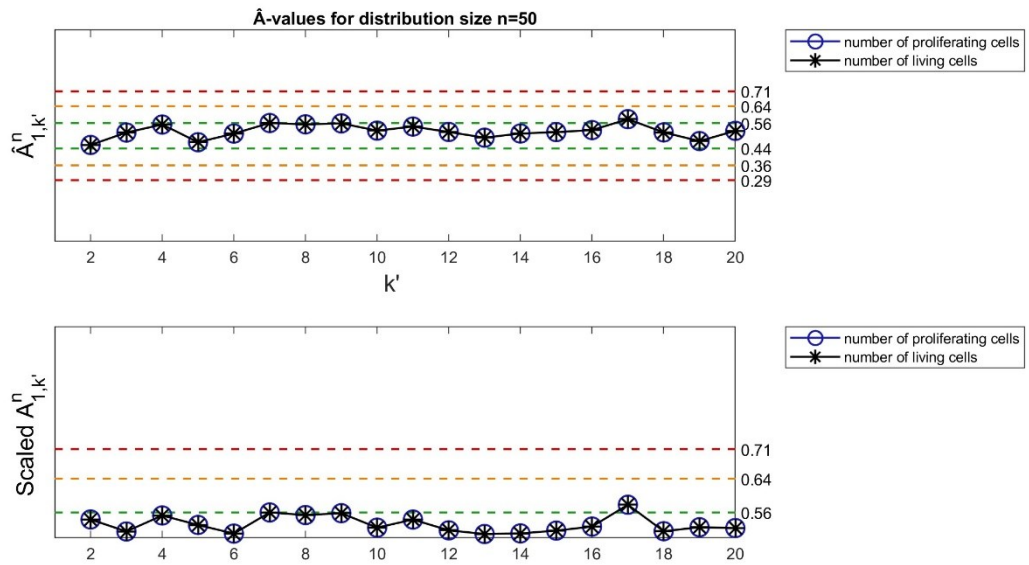**Figure C1**. Consistency Analysis, Â-values and scaled Â-values for $n = 1$.

Figure C2. Consistency Analysis, Â-values and scaled Â-values for $n = 5$.

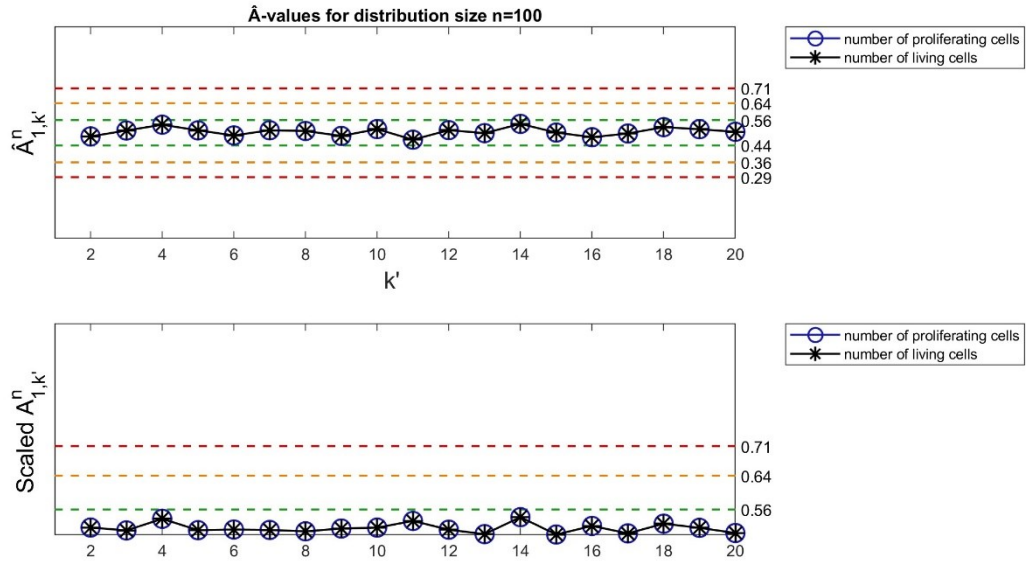**Figure C3**. Consistency Analysis, Â-values and scaled Â-values for $n = 50$.

Figure C4. Consistency Analysis, Â-values and scaled Â-values for $n = 100$.

The figures above demonstrated that the minimum distribution size for which the statistical significance is considered small (< 0.56) for both outputs of interest is $n = 100$. Therefore, all in-silico results are based on an average obtained from 100 simulation run1.