

Inaugural dissertation
for
obtaining the doctoral degree
of the
Combined Faculty of Mathematics, Engineering and
Natural Sciences
of the
Ruprecht-Karls-University
Heidelberg

Presented by

M.Sc. Robin-André Koch

born in Herrenberg, Germany

Oral examination: 14.11.2023

**Time-resolved analysis
of cell colony growth in vitro
after irradiation**

Referees:
Prof. Dr. Ursula Kummer
Dr. Markus Alber

Abstract

Here I developed a novel method to investigate the growth of cell colonies in vitro. The method is inspired by and augments the standard in vitro clonogenic assay (IVCA). While the field of application is radiobiological research, the approach can be applied to any domain where colony growth of adherent cells is of interest. The method utilizes high numbers of time-resolved microscopy image series and hence requires largely automated image data acquisition, image processing, quantitative data extraction and single-colony growth characterization. I designed a multi-step analysis framework to implement these steps. This contrasts with traditional approaches relying on visual examination of cell culture containers and manual classification of cell colonies. This new approach allows yet unattained insights into growth behaviors and growth rates of large numbers of individual cell colonies.

In applying the new method to five different cell lines (H3122, H460, RENCA, SAT, UTSCC-5) in different experimental settings, the following main results were found:

- a) For some of the cell lines, the initial seeding density influences the growth dynamics of the resulting colonies in densities commonly used in standard experiments.
- b) Pre-experimental cell culture conditions influence the growth dynamics in two tested cell lines (SAT, UTSCC-5) without irradiation.
- c) Exponential growth rates of two tested cell lines (H3122, RENCA) are normally distributed independent of irradiation dose, but the average growth rate decreases linearly across commonly used doses.
- d) Some colonies growing from photon-irradiated cells exhibit a distinct *delayed abortive* growth behavior, as observed for the two analysed cell lines (H3122, RENCA). The frequency of this behavior increases with increasing dose.
- e) Survival rates, as traditionally determined via the standard IVCA, clearly depend on experimental readout choices, namely the time of readout and the size threshold used to score survival of colonies. My analysis indicates that this dependence emerges from observations c) and d).
- f) The observed influence of readout choices propagates into relative biological effectiveness quantification for carbon irradiation for three examined cell lines (H460, RENCA, UTSCC-5).

Hence, I demonstrate that the presented method can be used to inform experimental design decisions in standard IVCA experiments, to perform robustness analyses on these assays, and to find distinct types of growth behavior. Still, the application in its current form is limited to adherently growing cell lines forming contiguous colonies. In addition, due to the multi-step procedure, some underlying assumptions and methodological decisions need to be made which potentially influence the resulting findings. I discuss these aspects in a dedicated chapter.

In future work, potential extensions and combinations with quantitative single-cell analysis methods such as FACS, fluorescent live-cell imaging or single cell omics methods can make this method a cornerstone application to build on in order to understand not only how, but also why colonies grow the way they do.

In conclusion, the presented method elucidates colony growth in unprecedented detail. The presented results showcase the potential relevance of these details. However, to establish this method as a standard tool for applied research, a unified

analysis framework is necessary to standardize the methodological aspects, from image acquisition to colony growth type classification.

Zusammenfassung

In dieser Arbeit habe ich eine neue Methode entwickelt, um das Wachstum von Zellkolonien *in vitro* zu untersuchen. Die Methode ist inspiriert vom *in vitro* clonogenic assay (IVCA) und erweitert diesen. Der Anwendungsbereich ist die radiobiologische Forschung, doch der Ansatz kann in jedem Bereich angewandt werden, in dem das Koloniewachstum von adhären Zellen von Interesse ist. Die Methode verwendet eine große Anzahl von zeitaufgelösten Mikroskopie-Bildserien und erfordert daher eine weitgehend automatisierte Bilddatenerfassung, Bildverarbeitung, quantitative Datenextraktion und Charakterisierung des Einzelkoloniewachstums. Zur Umsetzung dieser Schritte habe ich eine mehrschrittige Analyse entwickelt. Diese steht im Gegensatz zum traditionellen Ansatz, der sich auf die visuelle Untersuchung von Zellkulturbehältern und die manuelle Klassifizierung von Zellkolonien stützt. Dieser neue Ansatz ermöglicht bisher unerreichte Einblicke in das Wachstumsverhalten und die Wachstumsraten einer großen Anzahl einzelner Zellkolonien.

Bei der Anwendung der neuen Methode auf fünf verschiedene Zelllinien (H3122, H460, RENCA, SAT, UTSCC-5) in unterschiedlichen Versuchsumgebungen wurden die folgenden Ergebnisse festgestellt:

a) Bei einigen der Zelllinien beeinflusst die anfängliche Aussaatdichte die Wachstumsdynamik der entstehenden Kolonien für Dichten, die in Standardexperimenten üblicherweise verwendet werden.

b) Präexperimentelle Zellkulturbedingungen beeinflussen die Wachstumsdynamik in zwei getesteten Zelllinien (SAT, UTSCC-5) ohne Bestrahlung.

c) Die exponentiellen Wachstumsraten von zwei getesteten Zelllinien (H3122, RENCA) sind unabhängig von der Bestrahlungsdosis normalverteilt, wobei die durchschnittliche Wachstumsrate über üblicherweise verwendete Dosen linear abnimmt.

d) Einige Kolonien, die aus mit Photonen bestrahlten Zellen wachsen, zeigen ein ausgeprägtes verzögertes abortives Wachstumsverhalten, was für die beiden untersuchten Zelllinien H3122 und RENCA beobachtet wurde. Die Häufigkeit dieses Verhaltens nimmt mit zunehmender Dosis zu.

e) Die Überlebensraten, wie sie traditionell durch die Standard-IVCA bestimmt werden, hängen klar von der Wahl des experimentellen Auslesezeitpunkts und der für die Bewertung des Überlebens der Kolonien verwendeten Größenschwelle ab. Meine Analyse zeigt, dass sich diese Abhängigkeit aus den Beobachtungen c) und d) ergibt.

f) Der beobachtete Einfluss der Wahl der Ausleseparameter pflanzt sich in der Quantifizierung der relativen biologischen Wirksamkeit der Kohlenstoffionenbestrahlung für drei untersuchte Zelllinien (H460, RENCA, UTSCC-5) fort.

Folglich kann die vorgestellte Methode verwendet werden, um Entscheidungen über die Versuchsplanung in Standard-IVCA-Experimenten zu treffen, Robustheitsanalysen für diese Experimente durchzuführen und verschiedene Arten des Wachstumsverhaltens zu finden. Dennoch ist die Anwendung in ihrer derzeitigen Form auf adhärent wachsende Zelllinien beschränkt, die zusammenhängende Kolonien bilden.

Darüber hinaus müssen aufgrund des mehrstufigen Verfahrens einige zugrunde liegende Annahmen und methodische Entscheidungen getroffen werden, die die daraus resultierenden Ergebnisse potenziell beeinflussen können. Auf diese Aspekte gehe ich in einem eigenen Kapitel ein.

Potenzielle Erweiterungen und Kombinationen mit quantitativen Einzelzell-Analysemethoden wie FACS, Fluoreszenz-Lebendzell-Mikroskopie oder Einzelzell-omics-Methoden können diesen Ansatz in Zukunft zu einer Grundlagenmethode zu machen. Mit ihrer Hilfe kann man so nicht nur verstehen wie, sondern auch warum Kolonien so wachsen, wie sie es tun.

Zusammenfassend lässt sich sagen, dass die vorgestellte Methode das Wachstum von Kolonien in noch nie erreichter Detailliertheit aufzeigt. Die vorgestellten Ergebnisse zeigen die potenzielle Relevanz dieser Details. Um diese Methode als Standardwerkzeug für die angewandte Forschung zu etablieren, ist jedoch ein einheitliches Analysesystem erforderlich um die methodischen Aspekte von der Bildaufnahme bis zur Klassifizierung des Koloniewachstumstyps zu standardisieren.

Contents

I	Introduction	
1	Radiobiology reveals and explains radiation effects	3
2	The importance of colony formation assays	5
3	A new concept: studying cell colony growth in vitro	7
4	Technical background	13
4.1	Irradiation and biological effects	13
4.2	Image processing	16
4.3	Analysis of time series	18
II	Material & Methods	
5	Cell Culture	23
6	Irradiation Treatment	25
6.1	Photon irradiation	25
6.2	Carbon Ion irradiation	25
7	Imaging	27
7.1	IncuCyte	27
7.2	Zeiss Axio CellObserver	27
8	Image Analysis	29
8.1	Stitching	29
8.2	Registration	30
8.3	Segmentation	31
8.4	Quantification	33

9	Postprocessing	37
9.1	Tracking	37
9.2	Growth rate quantification	38
9.3	Growth curve classification	38
9.4	Survival rate calculation and survival curve fitting	40
9.5	RBE calculation	41

III

Results

10	Registration	45
11	Segmentation	47
12	Tracking	51
13	Growth at different seeding densities	53
14	Growth of different culture passages	57
15	Photon irradiation influences growth rates	61
16	Photon irradiation influences growth behaviors	63
17	Readout choices influence survival curves	67
18	Readout choices influence RBE results	69

IV

Discussion & Conclusion

19	New perspectives from the time-resolved IVCA method	73
20	Implications of biological results	75
20.1	Variability between cell lines	75
20.2	Seeding densities	75
20.3	Passage numbers	76
20.4	Growth rates	76
20.5	Growth behaviors	77
21	Potential, challenges & limitations of the method	79
21.1	Cell culture	79
21.2	Imaging	80

21.3	Image analysis	81
21.4	Threshold-based survival rate and cell counts	81
21.5	Growth curve analysis	82

22 Subsequent / additional analyses 83

22.1	Examining the influence of cell cycle status at irradiation	83
22.2	Discovery of predictive proteomic profiles	83

V Appendix and Bibliography

Bibliography	87
---------------------	-----------

List of Figures	95
------------------------	-----------

List of Tables	97
-----------------------	-----------

Acronyms	99
-----------------	-----------

Part I

Introduction

Radiobiology reveals and explains radiation effects 3

The importance of colony formation assays 5

A new concept: studying cell colony growth in vitro 7

Technical background .. 13

Radiobiology reveals and explains radiation effects

Radiobiological research aims at quantifying, understanding and predicting the influence of radiation on biological systems, from cells and organs to organisms and populations. This knowledge is essential for reasonable risk management in industries concerned with radiation, such as space travel, the nuclear power industry and medicine [1].

Especially in the medical industries, many different types of radiation are relevant in diagnostic as well as therapeutic applications. The pivotal goals of minimizing risks to the patient and maximizing therapeutic effects can only be accomplished based on sufficiently accurate models which capture the effects of the applied radiation type on the tissue in question. Several different strategies aim at creating and informing these models, ranging from empirical, descriptive approaches [2] to mechanistic approaches [3]. This spectrum mostly coincides with the scale of the investigated system: For populations, patients, and multicellular systems such as organs, organoids and spheroids descriptive models prevail, while mechanistic, causal models are restricted to cellular and subcellular systems due to the high complexity of larger systems [4]. Considering the stochastic nature of radiation, these mechanistic models are usually stochastic in nature and use track structures and spatial energy deposition distributions as modeled by Monte Carlo Simulations [5].

The concept of informing radiation effect models based on patient data faces several obstacles: Practical limitations include a lack of observability of processes within patients, the presence of many unknown, potentially confounding variables, which in turn can lead to large inter-patient variability with the resulting statistical challenges in the analysis of such data. In addition, ethical considerations about patient risks rightfully prohibit exploratory experiments. Hence, as common in basic biological research, foundational knowledge needs to be acquired in simplified biological model systems. This requirement is reason and justification for the development of colony formation assays [6, 7]. In the following I will describe the idea, the underlying assumptions and notable outcomes of such assays.

The importance of colony formation assays

A main interest of radiobiological research is the determination of dose-response relationships. Responses of interest range from intracellular aspects like types and severity of DNA damage [8], the extent and details of DNA repair processes [9], changes in transcription or translation [10, 11] or shifts in metabolic processes [12] to cell population characteristics like loss of organ function [13] or tissue regeneration [14].

Since the pivotal goal of radiotherapy in cancer treatment is to stop proliferation of a certain type of malicious cells, the response of interest is the fraction of cancerous cells losing their viability, i.e. their capacity to proliferate. Therefore, assays that allow the quantification of loss of proliferative capacity (or inversely, preservation of proliferative capacity) are main tools used in this field.

It is important to note that in the context of colony formation assays, the preservation of proliferative capacity is often referred to as *clonogenicity* or simply survival. This reflects the implicit assumption that radiation has a binary outcome on single cells: either a single cell keeps its proliferative capacity = stays clonogenic = survives, or it loses its proliferative capacity = loses clonogenicity = dies.

In addition, colony formation assays rely on the assumption that this binary outcome can be determined based on the size of colonies which grew from irradiated single cells. Through irradiation of a known number of single cells and counting of *sufficiently large* colonies grown from those cells after a certain *incubation time*, the level of survival can be determined. Note that *sufficiently large* and the *incubation time* are aspects of the readout which need to be chosen by the executing scientist based on their experience or established protocols, introducing a certain level of arbitrariness. Yet, performing the same readout for multiple samples at different doses yields dose-survival curves describing how clonogenicity changes as a function of dose.

This allows to address multiple different questions:

- Performing this under different environmental conditions (O_2 concentration, temperature...) allows to gauge the influence of these conditions on the dose-response relationship.
- Applying this concept to different cell lines or cell types under identical conditions allows to compare radioresistance/radiosensitivity between different cell types.
- Administering identical doses using different radiation modalities (photon irradiation at different energies, protons, different heavy ion species) allows to compare the relative biological effectiveness (RBE) of these modalities.

Hence, after Puck and Marcus established a method to culture single eukaryotic cells in vitro [6], radiobiologists applied this method to their respective fields of work, resulting in large numbers of dose-survival curves being published. A detailed

description on how these curves are calculated is given in section 4.1.4.

It is important to note that aggregate studies show that the resulting dose-survival relationships vary substantially between different studies for many cell lines, even for results from the same research groups [15, 16, 17]. Several factors responsible for this variability were identified in the literature, including seeding densities [18], seeding times [16], cell culture conditions [19] and human error [20, 21].

Differences in readout procedure, namely the choice of abovementioned definitions of *sufficiently large* colonies and the *incubation time* were not covered in these studies. However, results acquired in our group before and during my Master Thesis project indicate that these choices indeed have an influence on the resulting dose-survival relationships. This motivated the idea to develop a time-resolved version of the in vitro clonogenic assay (IVCA) to inspect how and why these readout choices influence the resulting survival curves and whether they introduce systematic errors in viability quantification.

Earlier time-resolved studies on cell colony growth provide additional rationale to develop the method used in this project. Rather than quantifying the growth characteristics of whole populations of single cells, they focus on the demonstration of specific cell death modes [22, 23, 24] and specific morphologies [25], or are limited by manual analysis of a few distinct colonies [26, 27] as technologies for automated analysis of large data were not available yet.

In the following chapter, the general idea, underlying assumptions as well as noteworthy aspects of our new method are presented.

A new concept: studying cell colony growth in vitro

Here I will conceptualize an idealized experiment to examine cell colony growth in vitro. *Idealized* here means that independent of the actual implementation of this concept, we can expect certain types of results that are inherent to this kind of experiment. As we will see, these results motivate and justify the application of this concept.

I will outline the general procedure and the foundational assumptions of this experiment. In addition, I will describe the types of analyses available for the resulting data. The different types of resulting data will be highlighted in **bold** font and the necessary analyses are subdivided into *quantitative*, *categorical* and *parametric* types. I will outline their necessary assumptions as well as expected insights.

Based on this theoretical description, I will derive crucial aspects of pre-experimental procedure, of experimental design and execution as well as aspects of data processing and data analysis. These aspects need to be considered to make the application of this concept possible and useful.

Finally, I'll expand on the potential research questions that can be addressed with this new approach.

Procedure, assumptions and expected insight of the idealized experiment

In essence, we want to observe how single cells grow into colonies (or not) dependent on their condition. Here, *condition* can represent any aspect of the cells' internal states (cell cycle status, genetic attributes, etc.), surroundings (temperature, chemical concentrations etc.), or treatment (different substances, concentration of substances, irradiation doses and modalities etc.).

Since cell colony growth is a dynamic process, the experimental procedure demands time-course observations to capture these dynamics. The fundamental quantity to record here is colony size at any given point during the observation duration. Here we must rely on the assumption that our method can truthfully measure across a range of sizes and allows us to observe colony growth without disturbing it. If this is established, the method yields **colony size distributions** across the duration of the observation period for all conditions of interest. This result alone is already highly informative, since, in contrast to the standard IVCA, it allows us to extract

- the size development of the whole cell colony population through time, enabling qualitative, visual comparisons between different conditions.
- descriptive aspects of the single distributions themselves:
 - shape of the size distributions (unimodal, multimodal...),
 - mean, median and extreme sizes,
 - variance of colony sizes within a condition.

Based on the *quantitative* result of colony sizes, further analyses can be performed. If we attribute the size of each colony at each time to their original single cell (progenitor), we obtain **single colony growth tracks**, which are single data series, one per colony, represented by size values over time. We refer to this process as colony tracking. A necessary assumption here is that the underlying data allows sufficiently accurate assignment of colony size measurements to their respective progenitor. The resulting tracks open up two additional types of analysis with their respective resulting data:

- A *parametric* description of growth in terms of **growth rates**
- A *categorical* description of growth in terms of **growth types**

To gain **growth rates** from the growth tracks, we need to assume a model of growth where the growth rate represents an informative parametrization of growth. Assuming an unlimited proliferation of cells in a colony, exponential growth

$$S(t) = S(0)e^{\gamma t}, \quad (3.1)$$

with $S(t)$ as size at time t and γ as the growth rate, represents the most obvious description, with γ directly corresponding to the speed at which cells replicate within the colony. Independent of the chosen growth model, retrieval of growth rate values requires a curve fitting method which finds the best estimate of the parameter given the time series size data.

For the definition and detection of **growth types**, we must assume that we can find distinct, qualitatively different growth behaviors, represented by different models of growth. Naively, there seems to emerge a conflict between the two types of analysis: If colonies follow different growth models, these colonies cannot be directly compared in terms of a quantitative parameters. Hence, either only colonies exhibiting the same type of growth behavior can be compared, or parametric models of growth need to be chosen that preserve intercomparability despite different underlying descriptions. To summarize: In contrast to the standard IVCA, a time-resolved approach to colony formation analysis yields abundant results. These results are useful to:

1. perform *quantitative* analyses via **colony size distributions** to observe growth dynamics of colony populations
2. perform *parametric* analyses via **growth rates** to quantify growth speed of single colonies and of colony populations
3. perform *categorical* analyses via distinct **growth types** and quantify their prevalence as a function of condition/treatment

When translating this concept into a practical application, we must ensure the validity of the assumptions made as described above. In addition, technical aspects and limitations need to be considered to ensure robustness of the approach and therefore comparability of results within and between experiments. In the following paragraphs, I will highlight important aspects to consider when employing this type of analysis.

Important aspects of practical implementation

The recording and analysis of cell colony growth as described above is a multi-step process requiring various decisions in terms of preparation, execution and analysis of the experiment. These decisions are made based on available resources, feasibility of handling and the research focus at hand. Following the chronology above, I will discuss the aspects where decisions are necessary. Where applicable, I will suggest reasonable choices for these aspects. In addition, I will classify their influence on the result as either *obvious*, *critical* or *unclear*.

Obvious aspects have a clear influence on biological processes and therefore the experimental results. These parameters need to be controlled strictly.

Critical aspects most likely influence the experimental results either through biological processes or by introducing bias in the analysis. These should be controlled but require careful experiment design and planning to control.

Unclear aspects might have an influence through either biological or technical processes, but the significance and extent of these influences was not yet examined.

Pre-experimental aspects:

Choice of cell line: This choice should reflect the research question at hand and is an *obvious* aspect. Different tissue types react differently to different treatments [17] and show differences in morphology and growth. Hence, the choice of cell line is an *obvious* parameter, not least because some cell lines do not grow into distinct colonies. Suspension cells or highly motile cell types which do not form colonies can not be analyzed regarding their cell colony growth. This is true for the standard IVCA as well [7].

Pre-experimental culture conditions: These are *critical* aspects, since nutrient and space availability influence growth behavior in the absence of any interfering treatment. Therefore, cells should be harvested, seeded and treated in comparable conditions which reflect the "normal" state of the cell population. In the case of immortalized cancer cell lines, this is the exponential growth phase where neither nutrient nor space limitations restrict growth.

Cell line passage: The passage number is an *unclear* aspect whose influence depends on the genetic stability of the cell line at hand as well as the difference in cell line age. For genetically stable cell lines which do not suffer from selection bias during passaging, potential differences in experimental results might arise from large differences in passage number. Cell lines which are genetically unstable and/or suffer from selection biases during passaging potentially show changes in experimental results after only a few passages. This needs to be considered and possibly tested for each cell line before drawing conclusions on interexperimental differences.

Number of seeded cells / seeding density: These are *critical* aspects. Since irradiation effects on single cells are stochastic, sufficient numbers of single cells need to be recorded to capture the population behavior without losing infrequent behaviors and without compromising on statistical power. This demands for a lower bound on the number of seeded cells. However, assuming a constant available space for growth, increasing the number of seeded cells translates to increased seeding density. This can

introduce problems such as *colony fusions*, where neighbouring colonies merge and are subsequently indistinguishable, or density-dependent growth inhibition. These effects will vary between conditions where the level of growth varies, introducing condition-dependent biases in the analysis. Alternatively, the number of seeded cells can be increased at constant seeding density by increasing the available space for growth. This translates to bigger or more cell culture containers, longer imaging cycles, more image data per experiment and more cell culture resources used. This creates an upper limit to the number of seeded cells which depends on available equipment and resources.

Experimental aspects:

Type of acquisition: This is an *obvious* aspect. As the observables of interest are colony size and potentially morphology, imaging is a plausible choice of recording colony growth. Still, different imaging modalities (brightfield, phase contrast, fluorescence) and different resolutions could potentially be used. Different subprocesses such as focus finding and sample alignment routines add to the potentially used variations of measurements. However, these choices influence the results in several ways:

1. Not all cell lines grow robustly under the stress of repeated imaging rounds. Especially frequent application of high-energy fluorescent microscopy over long periods of time might influence the viability and growth behavior of colonies. This should be tested.
2. Different imaging modalities and resolutions yield data with different characteristics, which in turn require different image analysis methods to robustly extract information from the images.

Therefore, full comparability between different datasets can only be achieved if identical measurement procedures are adhered to.

Number of measurements per time / per experiment: This is a *critical* aspect, since it influences the quality of subsequent growth analysis. Ideally, we would be able to record the size of every single colony at any given point during the experiment. This would require instantaneous recording of the full sample, which to my knowledge is not physically possible with any imaging system available. Additionally, this level of acquisition rates would result in an infinitely large record of data. In reality, the frequency of measurements is limited by the duration of an imaging cycle and potentially technical limitations of the imaging equipment such as overheating of the sample due to excessive acquisition frequencies, data processing speed or data storage space. It is generally true that more time points help to capture characteristics of growth curves more clearly and hence improve classifiability with regards to growth behaviors. This needs to be weighed against the fact that depending on the imaging modality and equipment, every imaging cycle represents a physical perturbation of the observed biological systems by introduction of radiation and heat by illumination or potential movements of the microscopy stage. These perturbations might alter the cells' viability and growth behaviors. Fortunately, cell colony growth is a slow

process in relation to the acquisition process, which suggest that we should be able to capture the characteristic aspects of growth dynamics at reasonable, minimally perturbing acquisition rates.

Duration of observation: This is a *critical* aspect. The shorter the duration, the more likely it is to miss out on dynamics that would have occurred after the experiment already ended. The longer the duration, the more likely it is that the samples experience changes in environmental parameters caused by overcrowded cell culture containers, nutrient depletion or prolonged physical stress due to imaging. For different conditions, the magnitude of these changes might differ, potentially leading to condition-dependent biases.

Environmental conditions during incubation: These are *obvious* aspects, since it is known that biological activity and hence cell colony growth strongly depends on temperature, nutrient availability and the concentration of CO_2 and O_2 . For comparability of experiments, these conditions must be identical.

Post-experimental aspects:

Image quantification method: This is a *critical* aspect. The successful extraction of colony sizes from image data requires a method that is validated to work robustly on the type of image data used. This requires that colony sizes are correctly recorded independently of colony size, colony age and potential differential treatments/conditions.

Colony tracking method: This is a *critical* aspect. While successful tracking depends on the choice of a suitable cell line, sufficiently sparse seeding and sufficiently high frequency of imaging, the method itself naturally influences the quality of resulting growth curves.

Growth rate quantification: This is a *critical* aspect. Firstly, in order for the extracted growth rates to meaningfully describe colony growth, the chosen growth model (see e.g. Eq. 3.1) must reflect the actual growth process. Secondly, especially in non-linear curve fitting problems, where a normal distribution of error in the data can not be assumed, an optimal choice for cost functions and optimization algorithms is not given. Hence, depending on the data, different choices in these aspects can lead to markedly different values as well as confidence intervals for the estimator.

Growth behavior classification: This is a *critical* aspect. Above all, the observed colonies must actually show different distinct growth behaviors which can be clearly categorized by a human observer. Only then it makes sense to attempt an automated classification of colonies into these distinct behaviors.

Similar to the other aspects of post-experimental analysis, the chosen method depends on choices made for the pre-experimental and experimental aspects. The method needs to be validated for the type of data at hand i.e. the resulting classifications should strongly agree with the classifications made by an experimental observer.

Potential applications of the time-resolved growth analysis

I identify the following three main potential benefits of the presented method:

Testing the robustness of standard IVCA:

As indicated above, some results of the standard IVCA might be biased depending

on the readout choices, namely the *threshold colony size* which qualifies colonies to be scored as viable and the *readout time* at which the experiment is stopped and scoring is performed on the dead and stained colonies present. Since the time-course acquisition yields all colony sizes at multiple timepoints, it allows us to perform this scoring using different size thresholds at these timepoints. This way, experimental designs can be tested for their robustness in relation to the readout procedure. Establishing whether certain cell lines and experimental setups are robust against readout related bias allows experimentalists to gauge whether they can trust results gathered using the standard readout. If this robustness can not be established, the researcher has a quantitative foundation to choose alternative methods of quantifying viability, for example by evaluating growth behavior as an indicator for viability, as follows.

Using growth behavior as indicator of viability: Instead of trusting that the size of a colony at a given time truthfully reflects their viability, one can classify a cell colony's viability based on its growth track representing the full growth history. For cell lines and conditions where the assumption *sufficiently large colony* = viable colony is not clearly established, the additional information contained in the growth history of a cell colony should allow more confident viability scoring as compared to the standard IVCA scoring.

Functional investigations regarding growth characteristics: Observing the influence of radiation or other treatments on growth behaviors and growth characteristics of cell colonies introduces questions about how these influences translate into functional changes within cells or cell populations. Are there driving regulatory factors determining the cells' reaction to treatments in term of growth? What are the influences of and interactions between cell cycle status, DNA damage repair or inter-cellular communication pathways on cell colony growth? Such questions were already tackled in relation to cell survival, often employing standard IVCA's [28, 29, 30] but access to colony growth dynamics opens new venues of research regarding intracellular regulation of growth and how it is influenced by different treatments. Combining image-driven colony growth analysis with fluorescence-based molecular imaging, fluorescence assisted cell sorting (FACS) analyses or proteomic analyses could reveal functional connections to help understand in detail how cells integrate radiation-induced changes in their regulatory processes.

Technical background

In this chapter, general technical concepts are introduced which are necessary to understand the presented work. Generally, these concepts are divided into aspects of irradiation, image processing and the analysis of time series data.

4.1 Irradiation and biological effects

In the context of radiotherapy, *irradiation* refers to the transfer of energy to a target tissue using ionizing radiation. The *ionizing* part is an important distinction from other types of radiation that add energy to a body without ionizing the receiving molecules. While conventional radiation employs high-energy electromagnetic waves such as X-rays, newer approaches also involve charged particles like protons, helium- or carbon-ions [31].

Independent of the type of irradiation, the biological impact results from the ionization of molecules within the irradiated tissue. These molecules either disintegrate or react with other molecules, potentially interfering with their functions. In an otherwise healthy cell, this activates repair mechanisms which try to revert these changes [1]. Here, damage to the DNA is shown to be a key contributor to the adverse effects of irradiation. Nevertheless, mechanisms unrelated to DNA have also been found to contribute, encompassing processes activated through membrane-associated signaling pathways or effects on unirradiated bystander cells close to irradiated cells, indicating intercellular communication [32].

4.1.1 Irradiation doses

In radiobiology, the term *dose* usually refers to the amount of ionizing energy E that is deposited per mass M of irradiated tissue:

$$D = \frac{E}{M}, \quad (4.1)$$

where D is the dose.

This can be thought of as a mass "concentration" of energy that the process of irradiation adds to the irradiated body. The SI unit to quantify doses is a Gray (Gy), where $1 \text{ Gy} = 1 \text{ J/kg}$. This definition is independent of the type of irradiation, may it be electromagnetic waves or particle-based. Nevertheless, the small-scale spatial distribution of this energy deposition varies substantially depending on the employed irradiation modality. Hence, the properties of irradiation processes can be further characterized using the following definitions.

4.1.2 Linear energy transfer (LET)

The LET quantifies how much energy is deposited in the receiving material per distance traversed by an ionizing particle. It is commonly quantified in units of

keV/ μm . Several studies have shown that the biological effectiveness of irradiation depends on this characteristic, owing to the differences in the spatial distribution of molecular alterations evoked by the incident irradiation [33, 34]. The energy deposition patterns of electromagnetic and particular irradiation modalities differ strongly, causing different patterns of damage to the receiving biological structures [35], which in turn is thought to result in different magnitudes of biological effects at identical doses. In addition to differences in penetration depth and energy deposition profiles along the direction of irradiation, this gives a theoretical justification of using particular modalities in patient treatment [1].

4.1.3 RBE

The abovementioned differences in biological effects between different irradiation modalities and LET characteristics are generally quantified by comparing physical dose of two modalities that cause the same biological effect. As exemplified in Figure ??, this biological effect is often measured in terms of cellular survival as described by dose-dependent survival curves. In this particular case, the RBE is calculated as

$$RBE_X = \frac{D_{R,E}}{D_{C,E}}, \quad (4.2)$$

where $D_{R,E}$ is the dose needed to cause the effect E using the reference modality and $D_{C,E}$ is the dose needed to cause the effect E using the comparing modality.

Historically, photon irradiation was and still is used as the main modality in radiotherapy and hence in biomedical irradiation research. Therefore, the quantification of RBE values usually uses photon irradiation as a reference modality. Studies have shown that RBE values depend on the comparing modality, but also on the type of biological effect and the magnitude of this effect [36], the examined biological specimen [34], experiment conditions such as oxygen levels [37] and cell cycle status [29].

4.1.4 Survival curves and the linear-quadratic model (LQM)

In the classical IVCA experiment, dose-dependent cell survival is obtained as follows: single cells are seeded in vitro and irradiated at different doses. After an incubation time t which is chosen by the experimentalist to allow the cells to grow into colonies, all cell culture containers are stained, usually with crystal violet. Then, the number of cells in each colony is counted and the colonies are scored as either clonogenic or not clonogenic based on a threshold size of n cells. From the numbers of colonies scored as clonogenic at dose D , cell survival is calculated as follows:

$$S_{D,t,n} = \frac{R_{D,t,n}}{R_{0,t,n}} = \frac{\frac{N_{D,t,n}}{N_{D,0,1}}}{\frac{N_{0,t,n}}{N_{0,0,1}}}, \quad (4.3)$$

where $R_{D,t,n}$ represents the quotient between the number of colonies $N_{D,t,n}$ that contain at least n cells after irradiation with dose D and all initially seeded cells

$N_{D,0,1}$ for that dose. To normalize for imperfect survival even without irradiation, survival $S_{D,t,n}$ is determined by division with $R_{0,t,n}$, which represents the fraction of sufficiently large colonies formed after mock irradiation ($D = 0$).

A plot of these dose-dependent survival values on a semi-log plot (see Figure 4.1) results in a characteristic survival curve that can be modeled by different mathematical functions. The most commonly used form is the two-parameter LQM

$$S(D) = e^{-\alpha D - \beta D^2}. \quad (4.4)$$

Here, the parameters α and β represent a linear and a quadratic contribution to loss of clonogenicity, respectively. This relatively simple description fits most dose-dependent survival data remarkably well. Nevertheless, biological counterparts in terms of distinct processes or target structures could not be ascribed to these parameters [38].

Independent of a lack of functional correlate to these parameters, they are frequently used in characterizing radiosensitivity of different cell types and tissues and the quotient $\frac{\alpha}{\beta}$ is successfully applied in radiotherapy treatment planning [39]. On one hand, this indicates that the determination of these values through the standard IVCA is an important step in allowing informed and hence safe radiotherapeutic treatments. On the other hand, considering the large variability of IVCA's outlined in chapter 2, scrutinizing the underlying method and potentially improving on the robustness of cell viability quantification is a worthwhile endeavour to further improve radiotherapy.

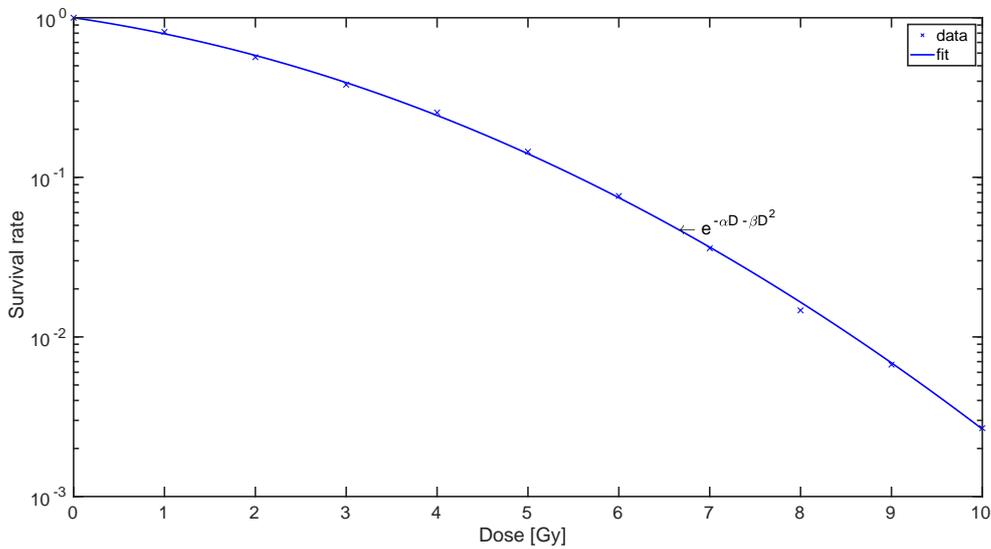


Figure 4.1: Example of a standard LQM fitted to survival ratios. The line represents curve fit to data (points). α and β as the fitted parameters as well as their quotient α/β traditionally characterize the radiosensitivity of cells.

4.2 Image processing

Quantitative analysis of objects encoded in image data requires standardized processing to extract these objects and measure their characteristics of interest. This encompasses object segmentation to detect these images and quantitative description of the identified objects.

In case of time-series image data, when location information is of essence, it is often necessary to preprocess the images so that spatial relationships between objects in different images can be represented in the same reference coordinates. In other words, a point present in multiple images needs to be described by the same coordinate values independent of the image it is found in. This process is called registration and generally includes a spatial transformation of one or multiple images to match a reference image.

4.2.1 Segmentation

Generally, segmentation procedures take an image (or a 3D image) and attribute a label from a finite set of labels to each pixel (or voxel) of this image. While the human brain is naturally capable of distinguishing different objects in a visual scene, a computational approach solves this task by making a labeling decision for each pixel based on

- properties of the pixel,
- properties of the pixel neighborhood,
- properties of the transformed and preprocessed pixel,

or a combination of the above. This can be implemented either through pre-defined rules or through rules that a supervised learning algorithm can deduce based on training data. For the latter, training data is created by experts who create segmentation maps for a set of images by correctly labeling them by hand. Feeding these pairs of images plus their segmentation maps to a learning algorithm, the algorithm should ideally be able to infer rules that allow to create segmentation maps for unseen data. User-defined rule-based procedures are often relatively simple and hence work well for data where the classes of interest show distinct features that allow to distinguish them from each other. The simplest example would be obvious differences in intensity or color values between the regions of interest. Other distinguishing features can be regular patterns or other structures such as edges of an object. Multiple dedicated morphological filters exist in the field of image processing which can detect and extract such structural features [40]. In most implementations, the last step of these procedures is to decide on the label for a pixel in a processed image or a combination of processed images based on threshold values chosen based on visual examination of the result.

However, in many applications such procedures do not sufficiently generalize across the data at hand. This can be due to global variations of intensity between images or local variations of intensity within an image. Another obstacle can be a high variability among the features of objects belonging to the same class. These chal-

Challenges can be overcome by approaches that are able to generalize over such varying characteristics.

To allow this, instead of relying on expert knowledge to define pixel classification rules, they learn those rules based on expert-labeled training data. Two concepts help in achieving strong generalization: First, the use of training data that already covers the high variability of characteristics present in the image data. Second, data augmentation steps which artificially introduce multiple types of variability into the training data, including spatial transformations (stretching, flipping, rotation of images), intensity distribution changes (brightness, contrast) or the introduction of different types of noise (gaussian noise, salt and pepper). This allows the rule inference algorithm to generalize over these variations even if they are not present in the original training data, reducing the necessary effort in creating representative training data.

The supervised segmentation approach used in this project is a neural-net based pixel classification.

Neural-net based classifiers are among the most successful tools in terms of image segmentation tasks [41]. The approach adapts concepts from biological visual processing and works by combining multiple neurons in a network of connections. Each neuron represents a function that takes some input and calculates an output. Input as well as output values can be multidimensional. The input can stem from multiple other neurons and the output can be input for multiple other neurons. The exact architecture of connections between neurons and the functions that are encoded in the neurons are design choices made by humans, while the strength of connections (i.e. weights) between the neurons are attributes of the network that are learned by optimization. This optimization usually involves multiple rounds of feeding the training data into the network and calculating an error metric by means of a loss function. In the case of image segmentation problems, this loss function quantifies the difference between the predicted output segmentation masks and the correct segmentation label masks (ground truth). A key feature of these networks is that the gradient of the loss function with respect to each individual weight in the network can be calculated by a process called backpropagation, allowing a subsequent adjustment of the weights along the gradient. This way, the error decreases with each round of updates.

Many network architectures build on, extend and adjust this general concept to tackle issues such as computational costs, class imbalances or requirements for large training data sets, leading to a plethora of published models [42]. One approach that very successfully applied a neural network for medical image segmentation is the U-Net [43], which implements a relatively parsimonious architecture that uses multi-level spatial context information but does not compromise on exact localization of labels. Due to comprehensive data augmentation, this model achieves high-quality segmentation with relatively few training data. A practicable implementation of this architecture was published by the Division of Medical Image Computing at Deutsches Krebsforschungszentrum (DKFZ), with a focus on automated model pre-configuration based on properties of the datasets in question [44].

4.2.2 Registration

In image processing tasks, it is often necessary to align identical structures within two or more images so that they can be referenced or measured in the same coordinate system. This process is called registration and can involve *multi-modality* approaches working on images which are acquired using different techniques, or *single-modality* approaches for images acquired by the same type of technique. Independent of the input data, registration tasks are generally defined as optimization problems where the objective function defines some metric of similarity calculated on the region of overlap between the input images. The optimization goal is to maximize this objective function by spatially transforming one of the input images (the *moving* image) in relation to the other image (the *reference* image). Depending on the types of spatial transformations allowed on the *moving* image, registration approaches can be distinguished as

translational registrations, where the only transformations allowed include shifts of the *moving* image in the image plane. In the context of 2D images, this implies two degrees of freedom, namely horizontal and vertical shifts.

rigid registrations, where in addition to translational shifts, rotation of the *moving* image about the image plane is also possible. This translates to three possible degrees of freedom, adding a rotation angle.

affine registrations, where in addition to translational and rotational transformations, scaling and shearing of the *moving* image is allowed. This corresponds to six possible degrees of freedom.

deformable registrations, where each point of the *moving* image can be translated to any other point independently. In theory this approach has $2 * N$ degrees of freedom, with N representing the number of points (pixels, control points) chosen for the image. Since this process is not well-defined, control terms need to be defined that penalize unreasonable transformations. In medical image registration, this is an area of active research [45].

An objective function often used to assess similarity for single-modality images in *translational* registration problems is the cross-correlation between *reference* and *moving* image. The global maximum of this function is found at the translational shift values which lead to a maximum overlap of structures between the two images.

4.3 Analysis of time series

Time series data analysis comes with a large set of methods to detect underlying trends, general shapes, seasonal changes and change points. In the course of this project, I applied two methods, one to estimate growth rates from fitting growth curves to the data and another to detect change points within the time series.

4.3.1 Curve fitting

Curve fitting problems assume that there is a mathematical model which describes the underlying process creating a series of data points as a function of some input variable. Some objective function can then be defined to assess the difference between the series of data points and the functional values resulting from evaluation of the mathematical model at the values of the input variable found at the data points, given a certain parametrization of the model. In the case of growth models, the input variable is commonly the time or duration of a process and the dependent variable describes the size, mass, volume, concentration or count of the growing quantity over that time.

Depending on the type of relation between independent and dependent variable as well as the assumptions that can be made about the process, different objective functions and optimization algorithms should be used to determine robust estimators of the model parameters. For example, linear relationships between independent and dependent variables with independent, normally distributed error terms allow for linear least-squares regression procedures with sum of squared errors (SSE) as objective function and R^2 as a valid goodness-of-fit measure. For non-linear relationships, these methods are not valid [46] and parameter estimation needs be performed by either non-linear curve fitting or by transforming the input data (e.g through log-transform) to allow the application of linear curve fitting methods.

4.3.2 Change points

Change points define values of the independent variable at which the relationship between independent and dependent variables changes. This can represent a change in the underlying type of relationship, or a change in the parametrization of the relationship while the type of relationship stays the same. Examples would be a switch from oscillating to constant behaviour or a switch from a strong oscillation to a weak oscillation, respectively.

Usually these change points are detected by splitting a time series at all data points and comparing statistical properties (mean, variance, root mean square, slope) of the split segments. Points that lead to a high difference in these properties between the split segments are then flagged as change points.

Part II

Material & Methods

Cell Culture	23
Irradiation Treatment	25
Imaging	27
Image Analysis	29
Postprocessing	37

In the course of this project, five different cell lines were used. Their origin and culturing specifics are listed in Table 5.1.

Table 5.1: Cell lines used.

cell line	organism	tissue	pathology
H460	human	lung	large cell lung carcinoma
H3122	human	lung	adenocarcinoma
SAT	human	oral cavity	squamous cell carcinoma
UT-SCC 5	human	tongue	squamous cell carcinoma
RENCA	mouse	kidney	renal adenocarcinoma

Independent of cell line and experimental setup, all cells were cultured as adherent cells in T25 cell culture flasks at 37°C and 5% CO_2 concentration in RPMI1640 medium with 10 % FCS and 100 U/mL PenStrep added. Regular passaging was performed at 60-80% confluence with dilution ratios of 1:10 to 1:100, depending on the growth speed of the cell line at hand. This procedure was adjusted in preparation of time-course experiments, as described in the following chapter.

Irradiation Treatment

For the old Incucyte data which I was given permission to work on by Dr. Ivana Dokic and Dr. Quanxiang Wei, no dedicated pre-treatment experimental procedures were recorded. For the data that I acquired myself, the same pre-treatment practice was used: In the last passaging before treatment, identical numbers ($\sim 10^3$) of cells were plated in T12.5 cell culture flasks. This sparse seeding allowed for a long growth period before singularizing and seeding single cells for treatment. This way, potential synchronization effects at irradiation time could be minimized.

Independent of the type of irradiation, treatments were performed according to the “plating after treatment” protocol detailed by Franken et al. [7].

6.1 Photon irradiation

Photon irradiation was applied with a Faxitron Multirad 225 (Faxitron Bioptics, Tucson, Arizona, USA), using 200 kV X-rays, a 0.5 mm Cu filter, resulting in a dose rate of roughly 1.0 Gy/min. Unirradiated samples were mock treated by putting them into the irradiation chamber without using the machine.

6.2 Carbon Ion irradiation

Carbon ion irradiation was performed at the experimental beamline at Heidelberg Ion Beam Therapy Center (HIT). The dose-averaged LET was $100 \text{ keV}/\mu\text{m}$. Due to the scanning beam approach employed at HIT, the dose rate was not stable over time, but the energy was deposited in 2-5 "bursts", depending on the total dose. Nevertheless, aggregating over the whole irradiation time, an average dose rate of roughly 2.0 Gy/min can be determined.

In contrast to the photon irradiation, the incident beam line at HIT is horizontal, which required the irradiated flasks to be positioned upright for homogeneous energy deposit across the growth area. In addition, to avoid interface effects during irradiation, the flasks were completely filled with medium shortly before irradiation. Subsequently, the superfluous medium was evacuated by pipette.

Analogously to the photon irradiation procedure, unirradiated samples were mock treated.

Imaging

In general, two different imaging procedures were performed to create the data used in this work. One relies on manual imaging of cell culture flasks in an inverted light microscope (Zeiss Axio CellObserver). The other one relies on the IncuCyte incubated imaging system, which automatically takes images of different types of cell culture containers at user-defined imaging intervals. From the IncuCyte acquisition we received *old* datasets that were acquired on 96-well-plate, with each well representing a technical replicate. For newer data that I acquired on the IncuCyte, I used 6-well-plates to minimize well rim effects as well as to provide sufficient space to grow for the colonies. Figure 7.1 shows example images for all three modalities.

7.1 IncuCyte

Imaging using the IncuCyte aims at making the process convenient for the experimentalist, requiring no user input or activity during the imaging duration, except for medium changes or for the introduction of planned perturbations.

New experiments are registered and customized in the beginning. This includes the choice of imaging modality (phase contrast and/or fluorescence channels) as well as image resolution and imaging schedules. After this setup, the software handles the whole process, including focus finding, stitching, postprocessing and file naming.

Imaging frequency is limited by the number of concurrently imaged containers, the types of containers, as well as the resolution chosen for these images. For the data examined in this project, imaging intervals of 3 h (old IncuCyte data) or 6 h (new IncuCyte data) were possible.

7.2 Zeiss Axio CellObserver

Semi-auto imaging as performed on the CellObserver microscope required more manual work, since the following steps need to be taken for each image of a given container:

1. transfer the container from the incubator into the imaging stage
2. find the focus at four support points within the imaged area
3. start the imaging process and wait (4 min)
4. remove the container and put it back into the incubator

For a standard dataset of 30 flasks, this requires roughly 2.5 h and hence could be performed once a day, resulting in imaging intervals of 24 h.

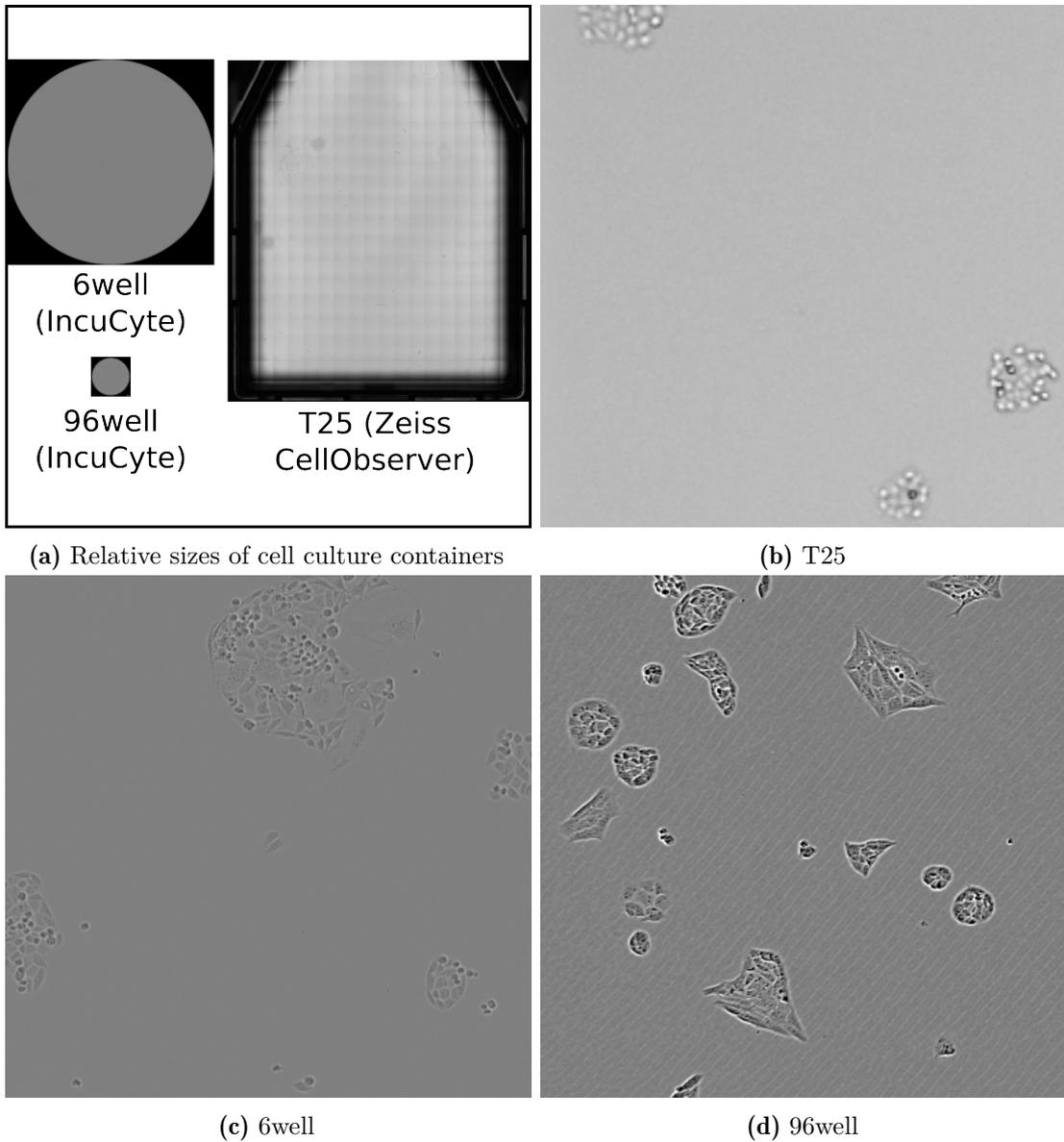


Figure 7.1: Cell culture containers. (a) The types of containers used, illustrating their relative sizes. Growth areas: T25 - 25 cm²; 6well - 9.6 cm²; 96well - 0.32 cm². (b)-(d) example sections of the cell culture containers, each representing a 1x1 mm area.

Image Analysis

To retrieve quantitative results from the images, a multi-step analysis framework is necessary which include the following main tasks:

- stitching of subimages (only for CellObserver data, since the image output is unstitched)
- registration of images within a time series
- segmentation of cell/colony objects in all images
- quantification of the segmented objects

8.1 Stitching

When imaging large areas of e.g. a cell culture flask, imaging systems need to scan over the whole area and capture subimages which are later merged into a large image that contains the whole area. The most common procedure here is to acquire subimages on a regular, evenly spaced grid, with a regular overlap between the images that allow subsequent fusion of these subimages into a complete image. Since the IncuCyte system does this internally and returns already stitched full images, this procedure is not necessary for IncuCyte image data.

In contrast, stitching is necessary for the Axio CellObserver data. There is an option to return stitched images within the ZEN Blue software available for the Axio CellObserver. However, this was of no use for me, as the overlap between subimages that could be chosen in the software did not agree with the actual overlap in the resulting images. Since the software uses the overlap value of choice in the stitching procedure, the resulting stitched images were not properly aligned. Therefore, I extracted the unstitched *.czi*-files and used an *ImageJ* macro to implement this procedure correctly. The macro script uses the MIST stitching algorithm [47]. Since the subimages were taken on a regular grid, no optimization routine was needed to align and fuse them into the final image; instead the stitching was performed based on fixed overlaps.

In addition to the stitching work, the script also sorts and names the resulting files in accordance with the subsequent analysis steps. Further, preprocessing of the subimages can be performed before stitching commences, which is not an option within the ZEN Blue software. While this procedure is tailored to data coming from this specific imaging setup and the returning file/folder structure is tailored for the subsequent analysis pipeline, it can in principle be adjusted to other datasets and subsequent usages.

8.2 Registration

Successful tracking of cell colonies depends on accurate location information. Mechanical inaccuracies of imaging systems, namely the relative positioning of microscope objective and imaged sample, can introduce shifts and rotations over time. This causes misalignments of cell culture containers and therefore misalignments between the colony objects. To correct these errors, all images representing the same cell culture container need to be aligned using registration methods.

Again, I used multiple different approaches in consideration of the specific properties of different image data types:

For the old Incucyte data (96 well plate), rotational displacements were not present, allowing for an approach that corrects only translational displacements. The most efficient and effective method finds the optimal transformations by locating the maximum of cross-correlation between the two images to be registered. For computational efficiency, the cross correlation is calculated after Fourier transforming the input images. The underlying algorithm is presented in [48] and implemented in the MATLAB function `dfregistration`¹.

The Axio CellObserver images needed a different registration approach, since repeated manual positioning of the container in the imaging platform introduced translational shifts as well as rotational differences between the images. Correcting for those, this data requires an approach that, in contrast to `dfregistration`, can compensate for rotation as well. In addition, as apparent in Figure 7.1, uneven illumination of subimages creates a regular grid-like pattern. This structure represents a strong attractor for the tried registration optimization procedures that outweighs the structures of interest (cell colonies). The location of this grid is independent of the positioning of the flask and therefore always at exactly the same position on the stitched images, even if the underlying flask position is shifted. Therefore, I registered the Axio CellObserver using the elastix toolbox, allowing rigid transformations as well as optimization based on specific regions of an image.

The newer Incucyte data (6well plate) could not be registered with the same approach, since for this data the algorithm tends to align the edges of the circular structures which represent artificially cut out regions created by the Incucyte software, not actual physical structures. Hence, instead of registration of the original image data, the necessary transformations were calculated based on already segmented images. The segmentation step extracts the relevant structures (cell colonies) within the images while leaving out the artificial circular structure. The original images as well as the segmentation masks were then registered by performing the determined transformations on all pairs of original images and their segmentation masks.

¹<https://de.mathworks.com/matlabcentral/fileexchange/18401-efficient-subpixel-image-registration-by-cross-correlation>

8.3 Segmentation

Throughout the project I used two different approaches to segment single colonies from the image data:

1. morphological filters and subsequent intensity threshold based segmentation
2. neural net based pixel classification using the *nnUNet* toolbox

In the following, I describe the different approaches and their area of application.

8.3.1 intensity thresholding

The initial method was described in detail in [49] and was used for the old Incucyte data (96well format). The approach relies on distinct structural differences between the contrast-rich cell colony regions and the uniform, low-contrast background. I used tophat and bottomhat filters to enhance the bright, outer regions of colonies as well as their dark, inner regions. The resulting images were merged by calculating pixel-wise maxima between the two images for each pixel. This resulting image was then binarized using an intensity threshold to segment cell colonies. As the image data is encoded in the unsigned 8-bit data format, the range of possible values is [0, 255]. Binarization is performed as follows:

$$M(x, y) = \begin{cases} 0 & | I(x, y) < t \\ 1 & | I(x, y) \geq t, \end{cases} \quad (8.1)$$

with x, y representing the pixel coordinates in the image and $t = 38$ a global intensity threshold.

Additional morphological operations on this binary image were used to remove small regions that represent specks of dust (morphological opening) and to fill false-negative holes in the colony regions (morphological closing). A visual outline of this procedure is shown in Figure 8.1.

I validated the segmentation quality by determining the proportion of objects with substantial (>10%) error on the determined areas by eye. To avoid potential systematic biases throughout the data, I repeated this procedure on 100 images across all doses and multiple time points to test for potential biases within the data set.

8.3.2 neural net in *nnUNet*

As the intensity-based segmentation turned out to fail for the less homogeneous Axio Cell Observer data, I employed a neural-net based segmentation approach implemented in the *nnUNet* toolbox [44]. Training data was chosen semi-randomly and consisted of image sections including the full range of doses, different cell lines (H460, UTSCC5 and RENCA), different doses (0 to 8 Gy) at both available irradiation modalities (photon and carbon) and different incubation times. Semi-randomly here means that inclusion of the full range of parameters mentioned before was enforced, but the location within the image was chosen randomly. The variability in data

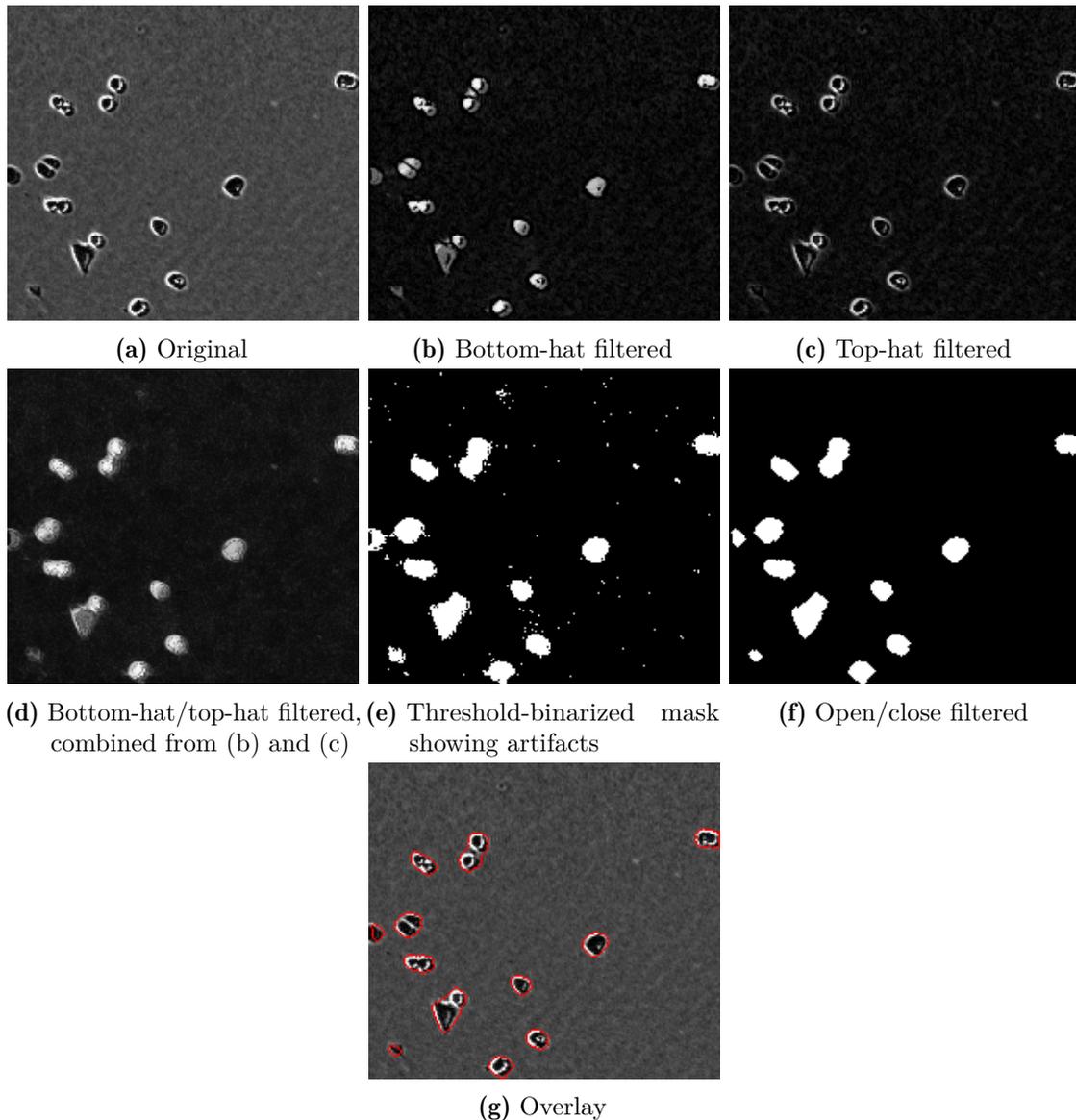


Figure 8.1: Workflow of intensity-based segmentation. The original image (a) is transformed by bottom-hat ((b)) as well as top-hat filtering ((c)). The combined result ((d)) is achieved by taking the maximum value of (b) and (c) at every pixel. The top-hat/bottom-hat filtered image (d) is turned into a binary mask (e) by thresholding intensity values using $T=38$. Morphological opening removes small false-positive spots in the background and morphological closing removes holes in the foreground objects ((f)). Panel (g) shows an overlay of the original image ((a)) with the outline of the segmentation mask ((f)).

sources was included to obtain a well-generalizing model. I performed the labeling by manual "painting" of cell colonies within the interactive visualization framework napari [50]. To this end, I contributed some functionality to the napari framework by allowing more convenient adjustment of labeling "paint brush sizes" ².

²<https://github.com/napari/napari/pull/5086>

Training of the model within the nnUNet framework was performed using 5-fold cross validation by default. The loss function optimized during training is the sum of cross-entropy and Dice score. The Dice score (also known as F1 score) is defined as follows for the binary segmentation as performed in this project:

$$S_{Dice} = \frac{2TP}{2TP + FP + FN}, \quad (8.2)$$

where TP is the number of pixels labeled as foreground in the training data as well as in the model prediction. FP represents the number of pixels labeled as foreground in the model prediction, but not the training data, while FN represents the number of pixels labeled as foreground in the training data, but not the model prediction. This score is a metric for segmentation accuracy and can be deduced from a combination of precision and recall.

The evaluation metric is also the Dice score. The final resulting model used for inference is an ensemble model from all five folds, using majority votes to decide on the label for each pixel.

8.4 Quantification

8.4.1 Extraction of colony objects and their properties from segmentation masks

Independent of the data set at hand, all segmentation objects in all images of a were saved into tables using MATLAB scripts employing the `regionprops` function from the Image Processing Toolbox. This function works on binary images, which by definition only contain the values 0 or 1, or label images, which by definition contain only positive integers from a set of labels. The function finds every contiguous region of a certain value, that is a set of neighboring pixels that all have the same value. For these regions, many different characteristics can be determined and calculated, and regions can be filtered depending on those characteristics. In my implementation, the following attributes of regions outlined in Table 8.1 were extracted:

For the old 96well IncuCyte data, segmented using the intensity thresholding after morphological filtering, a substantial number of false positive segmentation objects were found. These structures include specks of dirt, scratches in the container surface or cell debris. This intrinsic error could not be remedied by changes in the filtering steps or adjustments of the threshold value, so I used additional filter ranges on some of the region properties listed in Table 8.1 to filter out these objects. Threshold values for this procedure were found by inspection of the image data by means of a visualization function. I implemented this function to mark objects depending on whether a given property is within a defined range of values. Table 8.2 lists the ranges used and Figure 8.2 shows example objects filtered based on different properties.

Table 8.1: Colony object properties extracted from segmentation results.

Property	Definition	Type
Area	Number of pixels defining the region	scalar
Centroid	Center of mass of region	2-element vector
Perimeter	Distance along the boundary of the region	scalar
Minor axis length	Length of the minor axis of the ellipse having the same normalized second central moments as the region	scalar
Major axis length	Length of the major axis of the ellipse having the same normalized second central moments as the region	scalar
Circularity	$\frac{4\pi * Area}{(Perimeter + \pi)^2}$	scalar
Aspect ratio	$\frac{MajorAxisLength}{MinorAxisLength}$	scalar

Table 8.2: Region property ranges to filter false positive segmentation objects. For each property, a lower bound defines the minimum value allowed and the upper bound defines the maximum value allowed for an object to be considered a colony.

Property	lower bound	upper bound
Area	55	Inf
Circularity	0.2	1
Aspect ratio	1	4

8.4.2 Mapping colony area to cell number

As the traditional IVCA uses the number of cells in a colony as a measure of clonogenicity, I aimed to extract this quantity from the image data as well. However, since the number of cells in a colony is often not discernible from the image data, I instead used a mapping from the measurable area of colonies to their cell number as an estimator. I achieved this by visualizing randomly chosen colonies and storing the number of cells as well as the colony area for 300 colonies where cells were clearly countable. Figure 8.3 shows examples of such colonies. Since the relationship between colony area and cell count is approximately linear in the range of colony sizes traditionally used to score clonogenicity, I extracted the mapping from a simple linear regression, as depicted in Figure 8.4. I did this for each cell line and each image acquisition modality independently, since cells differ in their physical sizes and the different microscopes work on different resolutions.

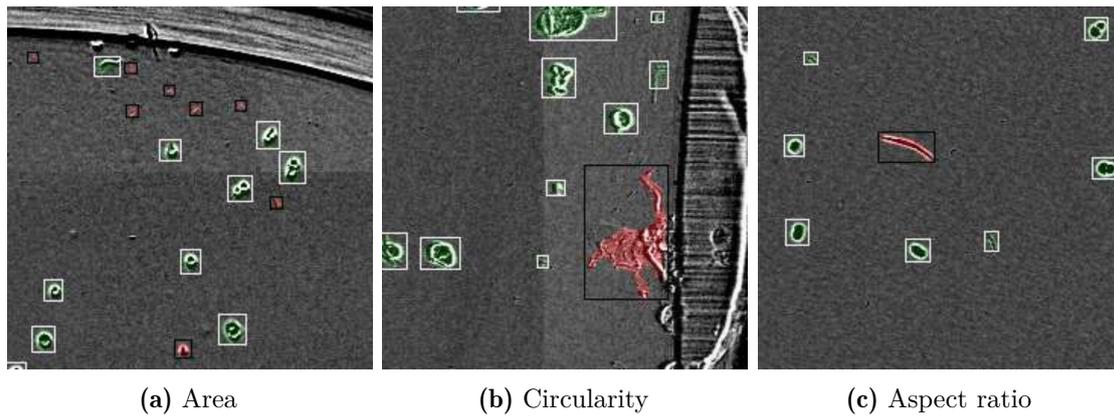


Figure 8.2: Filtered objects based on different properties. In each panel, green shaded objects surrounded by white boxes are accepted based on the property mentioned. Red shaded objects surrounded by black boxes are rejected. (a) Small objects (dust particles or dead cells) are filtered out. (b) Highly irregular shapes indicate cell debris or other dirt, which is filtered out using the circularity property. (c) Cell or cell colony objects show a certain regularity in terms of their aspect ratio. A high value in this property is almost exclusively found in scratches or fibrous dirt particles.

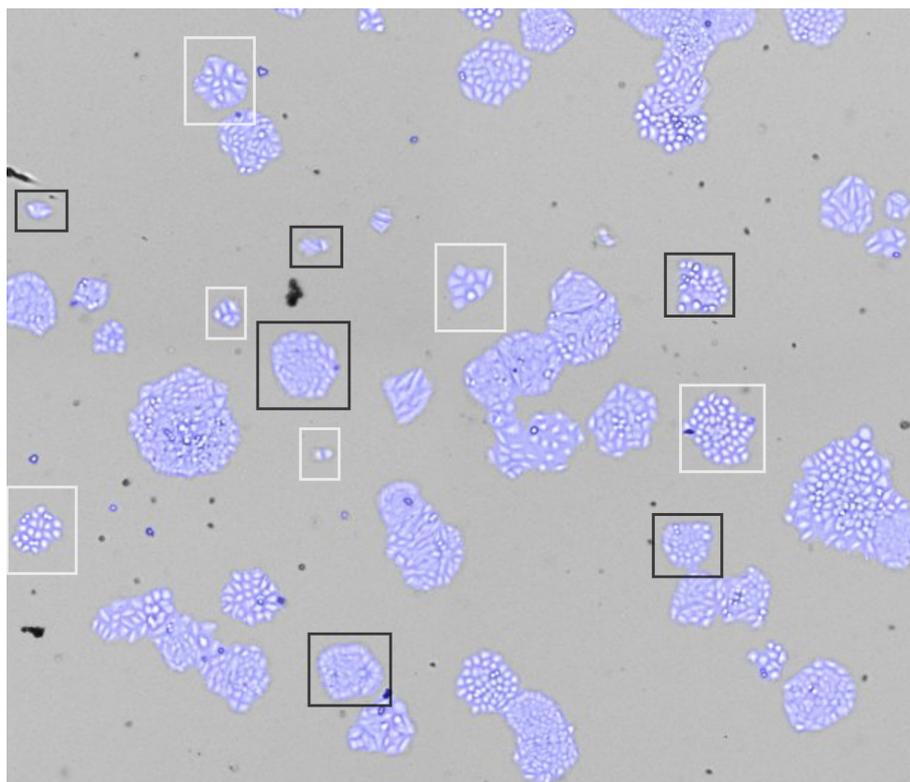


Figure 8.3: Examples of colonies with countability indicated. Colonies with a white rim are countable, colonies with a black rim are not.

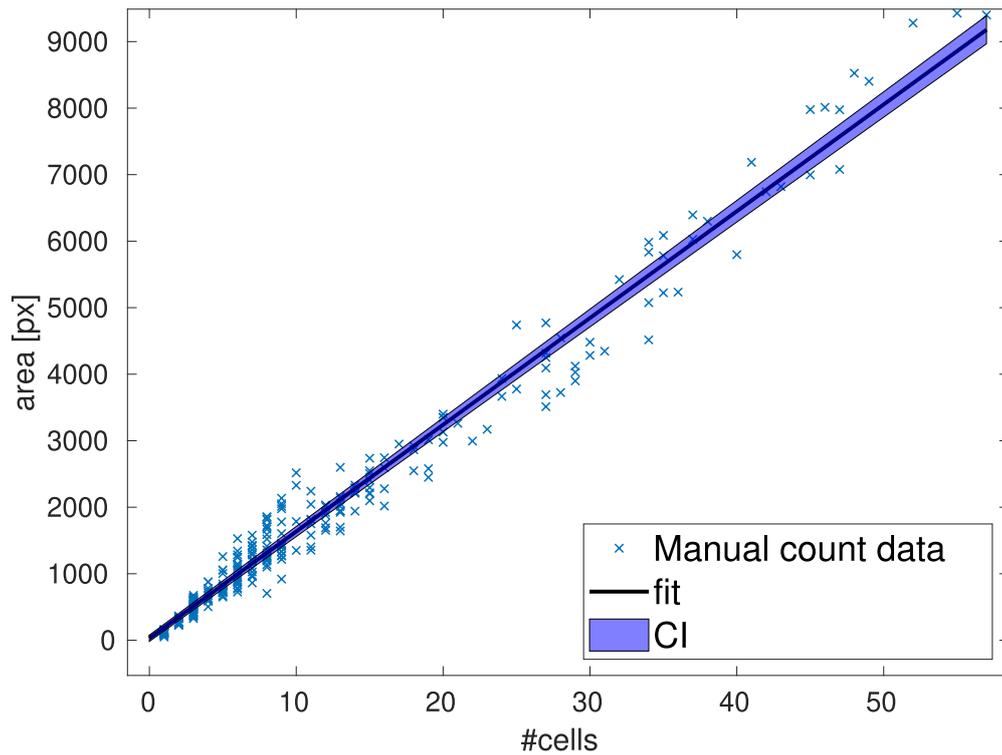


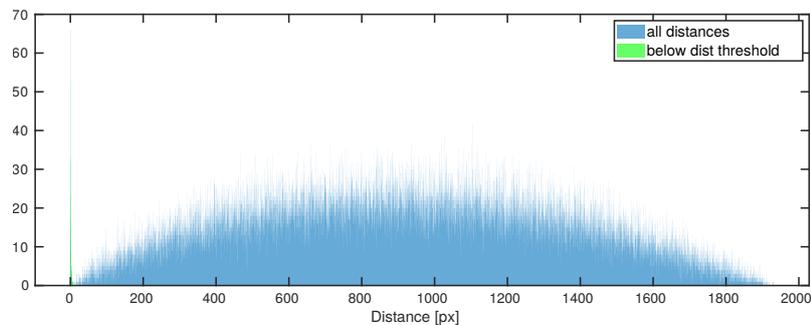
Figure 8.4: Mapping of colony area to cell count. Points indicate the area in pixels for manually counted cell numbers. Cell line: H460; Modality: Axio Cell Observer, T25 flasks

9.1 Tracking

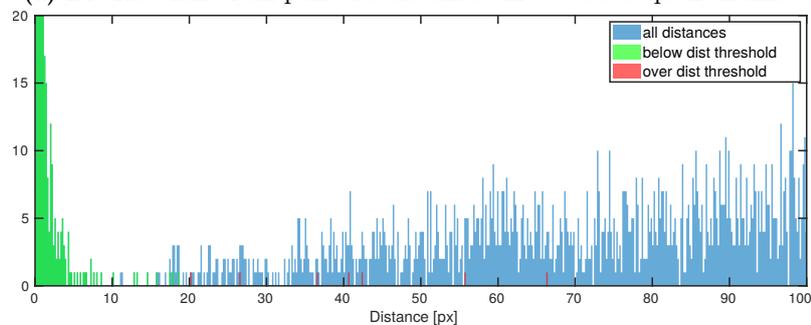
Tracking of colonies was performed using location-based information, attributing a colony object C_{f+1} in frame $f + 1$ to a colony object C_f in frame f if

1. their centroids have the mutually closest distance between each other,
2. the object C_f is already part of a track,
3. the distance between the objects does not exceed a user-defined threshold distance D_t of 20 pixels.

The threshold D_t was chosen by examining the distribution of mutual distances between all objects within a data set, as shown in Fig 9.1.



(a) Distribution of all pairwise distances in two subsequent frames



(b) Small distances as shown in (a)

Figure 9.1: Pairwise distances between all objects in two subsequent frames. This plot was used to find a robust displacement threshold t_D . Panel (a) shows all pairwise distances between regions in two subsequent frames. Panel (b) is a zoom-in of (a) focussing on small distances. Green bars represent distances where conditions 1 and 3 above are met. Red bars also show distances where condition 1 is met, but condition 3 is violated.

9.2 Growth rate quantification

The growth curves resulting from tracking of independent cell colonies were quantified based on the assumption of exponential growth. The following procedure was only applied to IncuCyte data, as the large time intervals of the Axio CellObserver data did not allow robust fitting and growth type classification.

I used the Theil-Sen slope estimator [51, 52] on log-transformed size values of all growth curves to robustly extract growth rates for each single colony. It is noteworthy however, that not all cell colonies follow an exponential growth behavior throughout the incubation time, as I have observed in the old 96well IncuCyte data. For this reason, I applied a changepoint detection algorithm to find for each growth track the time at which a change in behavior was most likely, using the MATLAB function `findchangepts`. Subsequently, two growth rates were determined for the two parts divided by the change point. After growth curve classification (see next section), the growth rate for each track was defined as either the original Theil-Sen estimator if it was classified as *exponentially growing* or *initially abortive*, or as the Theil-Sen estimator for the first part, if the growth curve was classified as *delayed abortive*.

9.3 Growth curve classification

In each data set I found the following three distinct growth behaviors: *exponentially growing*, *delayed abortive* and *initially abortive*. Due to differences in data properties, the classification of growth curves into these behaviors differed between the IncuCyte data and the Axio CellObserver data.

For the IncuCyte data where sufficiently dense growth tracks were available, I classified the growth curves utilizing information gathered from change point detection and the partwise Theil-Sen slope estimates, among other characteristics. Table 9.1 lists the attributes for each growth curve that were used as predictors to inform the classifiers.

I trained classifiers for each individual cell line, owing to their different growth characteristics. For each cell line, I scored 800 growth tracks into one of the three distinct growth behaviors, employing a visualization and labeling function I implemented in MATLAB. Using these labeled tracks, I used the MATLAB-based classification learner app to find the best performing classifiers based on model accuracy. Robust assessment of model performance was ensured by using a 5-fold cross validation scheme. For all datasets, bagged-tree classifiers performed well and the trained classifiers were used on the remaining data to predict their classes.

Since the Axio CellObserver data with 24h intervals did not contain sufficient datapoints to robustly determine either changepoints or Theil-Sen slope estimates, classification of these growth curves needed a new approach. Here I restricted the predictors to direct descriptors of the growth curves, namely the maximum size of a colony and the time at which this maximum is reached. Here I labeled 250 growth tracks per dataset into the distinct behavior classes and used the classification learner app to find the best performing classifiers. For these datasets, simple decision

Table 9.1: Predictor variables used for classification of any growth curve G . After finding the most probable changepoint tc , Theil-Sen estimates for both the whole growth curve as well as for the parts on either side of the changepoint are determined. The following predictors result from this procedure.

Predictor	Description
Maximum size of growth curve	$\max(G)$
Relative position of maximum size	Position of $\max(G)$ along G normalized to interval $[0,1]$
Size at changepoint	$G(tc)$
Growth rate estimate total fit	Theil-Sen slope estimate using whole G
Growth rate estimate initial fit	Theil-Sen slope estimate using G from start to tc
Growth rate estimate end fit	Theil-Sen slope estimate using G from tc to end
Z-value initial vs end slopes	Z-value of ranksum test comparing all pairwise slopes between points before tc with all pairwise slopes between points after tc
Mean squared error of total fit	$\frac{1}{n} \sum_1^n (G - \hat{G})^2$, where \hat{G} is the growth curve estimate based on the Theil-Sen estimate for the whole curve
R^2 of total fit	$1 - \frac{\sum_1^n (G - \hat{G})^2}{\sum_1^n (G - \bar{G})^2}$, where \hat{G} is the growth curve estimate based on the Theil-Sen estimate for the whole curve and \bar{G} the naïve model using the mean of G

trees showed the best performance and were trained analogously to the IncuCyte classifiers. Prediction accuracies for the resulting classifiers were 95.6 ± 0.3 % for H460, 96.1 ± 0.7 % for RENCA, and 93.7 ± 0.4 % for UTSCC-5. The accuracies values are calculated as mean \pm SD of 1000 cross-validation accuracies. Each accuracy was obtained by performing 5-fold crossvalidation based on independent, random splits of the training data. Again, the resulting classifiers were applied to all remaining growth curves to determine their respective growth behaviors. This procedure is also described in Koch,2023[53].

9.4 Survival rate calculation and survival curve fitting

9.4.1 Survival rates based on colony size thresholds

As adapted from the standard IVCA readout [7], colony size-based survival rates are acquired from object counts. In the following, the necessary calculations to arrive at dose-dependent survival rates are detailed. All numbers N refer to numbers of colony objects. In the special case of initially seeded cells, the numbers N_{seeded} still refers to numbers of detected colony objects, even though each colony object only contains a single cell. It is important to mention, that in contrast to the standard IVCA, we have access to single cell colony counts immediately after seeding, while the standard approach estimates this number based on dilution protocols.

The absolute survival $S_{ab,D,i}$ at dose D in a single replicate i is calculated as the fraction of initially seeded cells ($N_{seeded,D,i}$) in that replicate which grew into colonies larger than the size threshold ($N_{large,D,i}$):

$$S_{ab,D,i} = \frac{N_{large,D,i}}{N_{seeded,D,i}} \quad (9.1)$$

To get a baseline survival for unirradiated cells, also called plating efficiency (PE), the same is done for replicates j at 0 Gy:

$$S_{ab,0,i} = \frac{N_{large,0,i}}{N_{seeded,0,i}}. \quad (9.2)$$

The normalization factor PE is calculated as the mean absolute survival over all replicates at 0 Gy:

$$PE_{mean} = \frac{1}{n} \sum_{j=1}^n S_{ab,0,j} \quad (9.3)$$

Relative survival, i.e. the survival rate for each replicate at a given dose D is then calculated as the ratio between the absolute survival $S_{ab,D,i}$ and the PE:

$$S_{D,i} = \frac{S_{ab,D,i}}{PE_{mean}} = \frac{\left(\frac{N_{large,D,i}}{N_{large,0,i}}\right)}{\frac{1}{n} \sum_{j=1}^n S_{ab,0,j}} \quad (9.4)$$

Finally, the scalar survival rate S_D at a given dose D is calculated as the mean relative survival over all replicates:

$$S_D = \frac{1}{n} \sum i = 1nS_{D,i} \quad (9.5)$$

9.4.2 Survival rates based on colony growth behavior

A new, alternative approach to score dose-dependent colony survival is based on colony growth behaviors. As described in Koch, 2023 [53], absolute dose-dependent survival rates are calculated as

$$F_{g,D,i} = \frac{N_{c,D,i}}{N_{t,D,i}}, \quad (9.6)$$

where $N_{c,D,i}$ is the number of growth curves classified as clonogenic at dose D in replicate i , while $N_{t,D,i}$ is the of the total number of growth curves tracked at dose D in replicate i .

Analogously to the size threshold-based quantification, a PE can be calculated by applying Equation 9.6 to non-irradiated samples and calculating the mean value over all samples j :

$$PE_{g,mean} = \frac{1}{n} \sum_{j=1}^n F_{g,0,j} \quad (9.7)$$

The relative survival $S_{g,D,i}$ at dose D in replicate i is then calculated by normalizing to the mean PE:

$$S_{g,D,i} = \frac{F_{g,D,i}}{PE_{g,mean}}, \quad (9.8)$$

Finally, the scalar growth behavior-based survival rate $S_{g,D}$ at dose D can be calculated by averaging over all relative survival values at that dose:

$$S_{g,D} = \frac{1}{n} \sum i = 1nS_{g,D,i}. \quad (9.9)$$

9.4.3 LQM-based survival curves

For both colony size-based survival rates and growth behavior-based survival rates, survival curves, representing fits of the LQM to survival rates, were acquired from fits to replicate survival rates $S_{D,i}$ or $S_{g,D,i}$, respectively. LQ parameters were fit by robust linear fitting after log-transformation of survival ratios using the `fit` function implemented in Matlab, minimizing the sum of absolute residuals as defined by the robust least absolute residuals routine.

9.5 RBE calculation

RBE values were calculated as a function of the reference radiation dose D_R as

$$RBE(D_R) = \frac{-2\beta_C D_R}{\alpha_C - \sqrt{\alpha_C^2 - 4\beta_C * (-\alpha_R D_R - \beta_R D_R^2)}} \quad (9.10)$$

with α_R and β_R as LQ parameters of the reference curve and α_C and β_C as LQ parameters of the comparison curve. For the data used in our studies, the comparison modality is carbon ion irradiation with an LET of 100 keV/ μm . Equation 9.10 can be derived from the standard definition of RBE as described in section 4.1.3.

Part III

Results

Registration	45
Segmentation	47
Tracking	51
Growth at different seeding densities	53
Growth of different culture passages	57
Photon irradiation influences growth rates	61
Photon irradiation influences growth behaviors	63
Readout choices influence survival curves	67
Readout choices influence RBE results	69

For the old 96well-based IncuCyte data, I checked registration success by visual inspection of the resulting image stacks. For all three available datasets, all time-course stacks were loaded in ImageJ and examined for systematic shifts. No such shifts were found after registration, even though replacements of up to 200 px were present before. Figure 10.1 shows an example of a shift correction by comparing two subsequent frames.

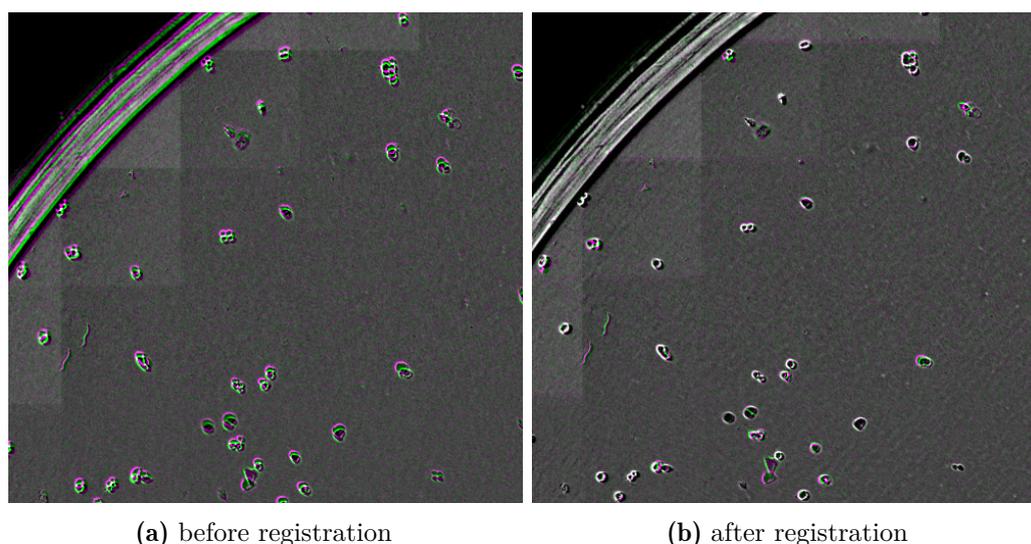


Figure 10.1: Registration of subsequent frames. The left panel shows a false color overlay of a section of two subsequent frames. The left panel shows the same sections after registration, with the same objects now overlapping optimally.

As described in section 8.2 and shown in Figure 7.1(b), the Axio Cell Observer data shows a distinct artificial grid introduced by non-homogeneous lighting. In addition, the shifts between images of the same flask include rotational transformations, motivating me to allow rigid transformations for registration of this data.

The elastix optimizer is able to only take dedicated regions of the image into account. By defining a binary map representing the rim regions of the cell culture flasks as allowed input for the optimizer, successful registration could be achieved without interference by the grid-like artificial structure. Again, success of registration was established by visual inspection of the complete image time courses for every data set.

Visual inspection of 6well IncuCyte data registered by the two approaches mentioned above (`dftregistration` and rigid registration using elastix and a mask defining the image areas to take into account) shows that both approaches do not work robustly, especially at early time points when the images contain only a few single cells or small cell colonies. Here, switching the order of registration and segmentation

yields robust results: Segmentation of unregistered images, followed by elastix-based registration applied to segmentation masks shows robust and correct segmentation. Analogously to the other approaches, visual inspection of complete data sets was used to validate registration success across all time course image series.

Segmentation quality for the intensity threshold-based approach on IncuCyte datasets was assessed by visual inspection of randomly chosen individual cell colonies. For each dataset, 2500 objects were examined using a custom visualization function implemented in MATLAB. For each individual object, segmentation quality was scored based on the agreement of the segmentation mask with the actual colony object. Objects were classified as either overestimated, if the area of the segmentation mask exceeded the actual colony by more than 10 %, as underestimated if the area of the segmentation mask was more than 10 % smaller than the actual object, or as correct, if segmented area and actual area agreed within a 10 % range. Table 11.1 shows the aggregate results for the three 96well-based IncuCyte datasets.

Table 11.1: Segmentation results for 96well-based IncuCyte datasets. For each dataset, 2500 individual colony objects were visually assessed. Entries represent absolute values and relative frequencies of segmentation qualities in brackets for each dataset. Adapted from Koch,2021[49].

Dataset	Correct	Overestimated	Underestimated
H3122 A	2440 [97.6%]	23 [0.9%]	37 [1.5%]
H3122 B	2453 [98.1%]	24 [1%]	23 [0.9%]
RENCA	2436 [97.4%]	30 [1.2%]	34 [1.4%]

Segmentation quality for the nnUNet-based approach was assessed through performance metrics of the model on the training data as well as visual inspection of overlays of segmentation masks with the original data. Figure 11.1 shows how the loss improves during training. Figure 11.2 show non-trivial segmentation tasks and how the trained model successfully handles these cases. Since I labeled training data from multiple cell lines, I trained the final model on all data together and the resulting model generalizes well over all cell lines, colony sizes and morphologies. In Figure 11.3, edge cases are presented where the model fails. The Dice score evaluated on the whole training data was 0.913 for the final model.

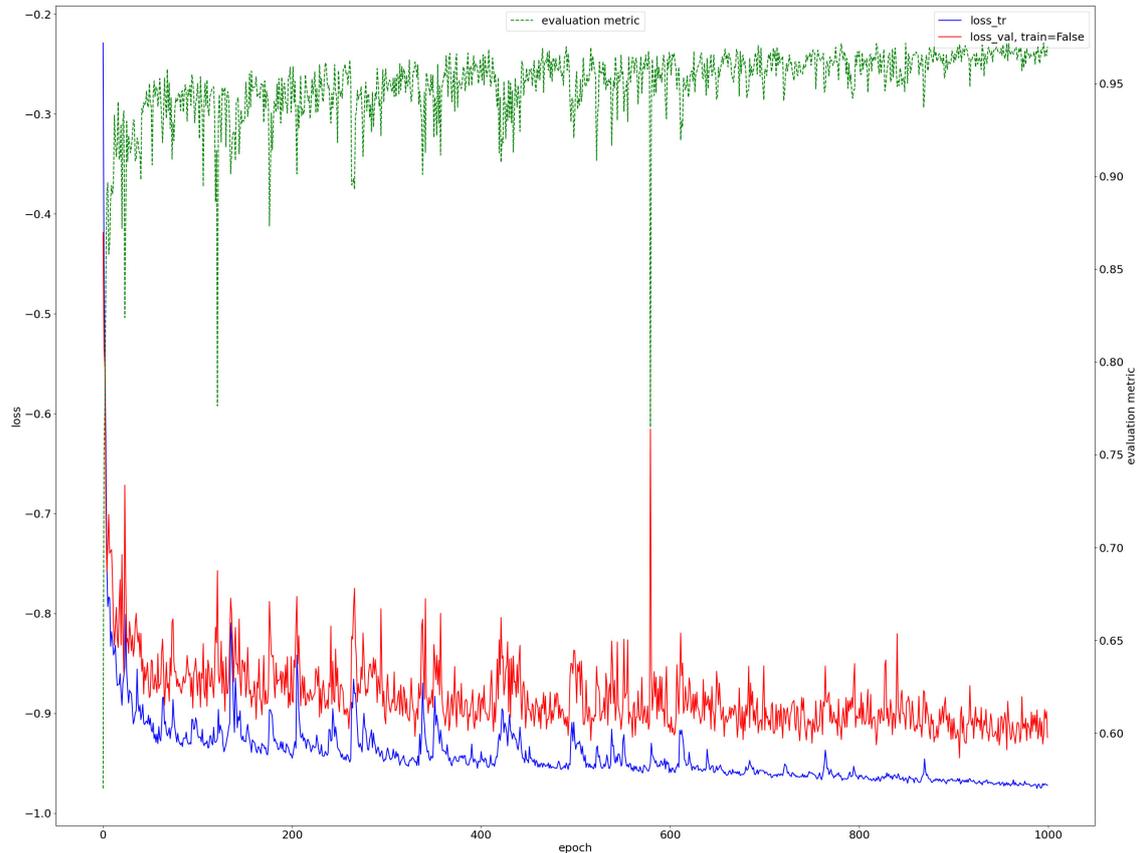
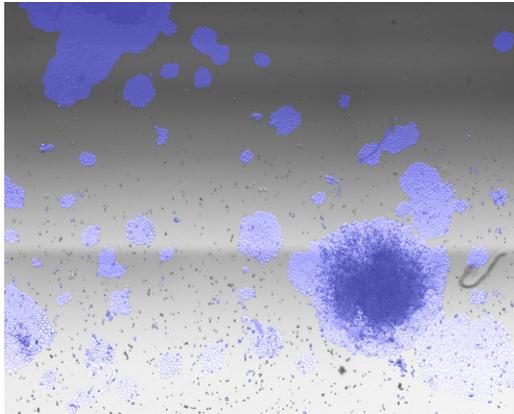
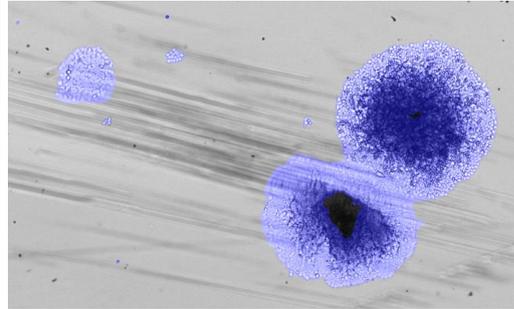


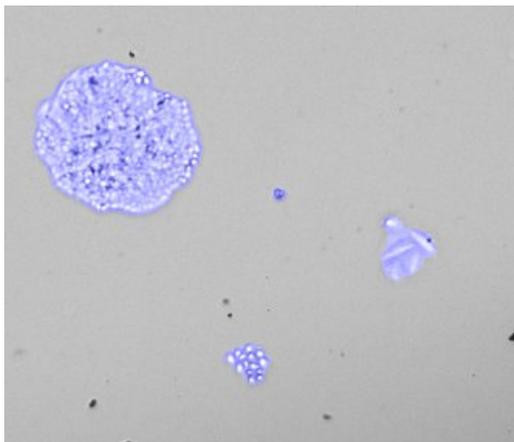
Figure 11.1: Development of loss function and evaluation metric during nnUNet segmentation model training. The blue curve represents the progress of the training loss (sum of cross-entropy and Dice score) through the training epochs, the red curve shows the validation loss (as measured on the hold-out fold making up 1/5 of the training data) and the green curve shows the evaluation metric (Dice Score). This plot is created by the nnUNet tool during training.



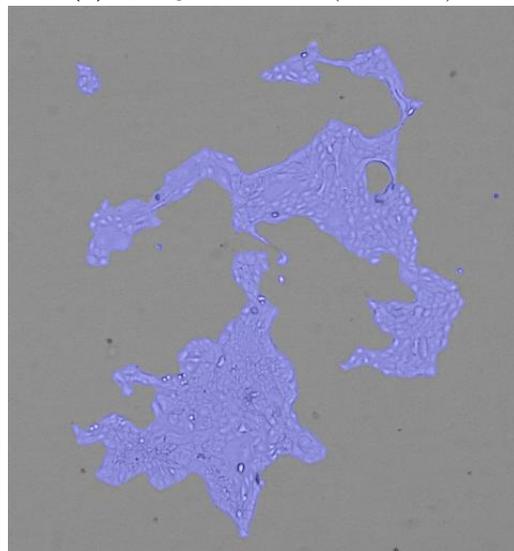
(a) Across different brightness and gradient



(b) Background noise (scratches)



(c) Different cell morphologies



(d) Irregular colony shapes

Figure 11.2: Examples of hard segmentation tasks successfully solved by the nnUNet model. Blue shaded areas represent the segmentation masks. The segmentation works remarkably well in a multitude of challenging circumstances. Brightness differences and gradients (a) which would lead to failed segmentations for intensity-based thresholding approaches, superposition of colonies with background structures such as scratches (b), varying cell morphologies (c) and irregular colony shapes (d) are all handled successfully.

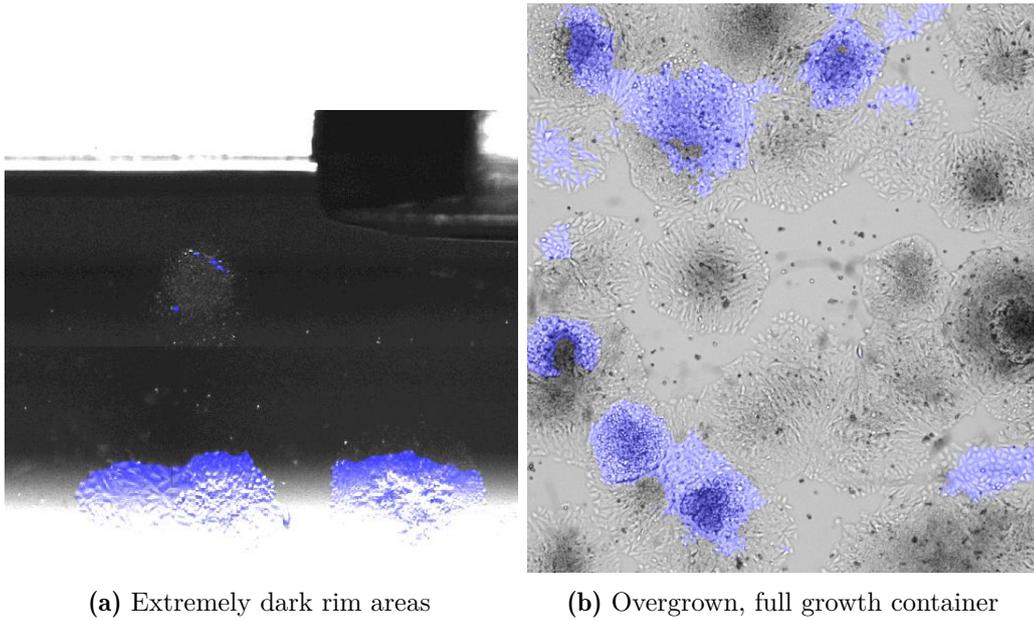


Figure 11.3: Examples of hard segmentation tasks where the model fails. Blue shaded areas represent the segmentation masks. Some areas in the rim region of T25 flasks are too dark for the model to detect the colonies successfully (a). Towards very late time points, the flasks are overcrowded which leads to segmentation failures (b). Both of these cases are not present in the training data.

To assess tracking quality, I implemented a visual inspection functionality that randomly chooses a certain track and displays the corresponding image data to the user. By using this feature, I looked at 300 randomly chosen tracks for each dataset and scored those tracks as either *correct* if the track contains no errors, or as prematurely terminated if a track ends despite the fact that the colony could still be visually identified in later frames, or as *fused*, if the track contains any colony object that originates from two or more initial cells. Figure 12.1 exemplifies this quality analysis grouped by different doses for an old IncuCyte 96well plate dataset of RENCA cells.

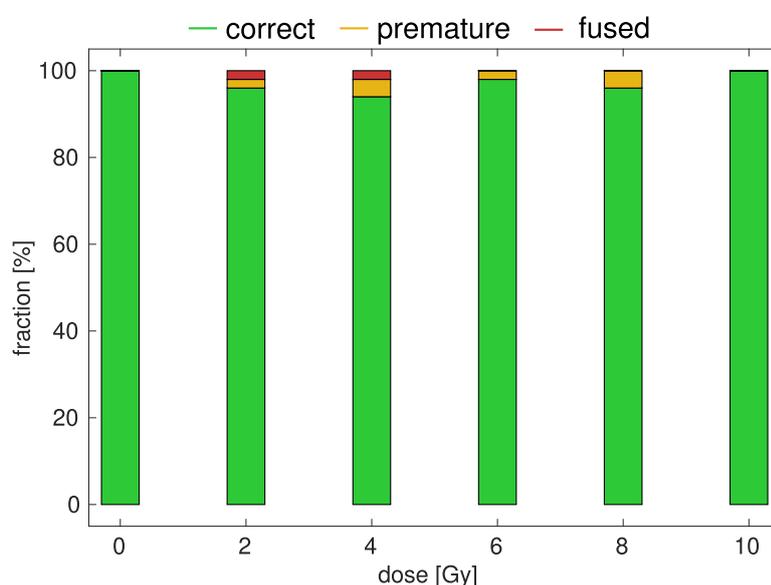


Figure 12.1: Relative fractions of track quality labels for a RENCA dataset. A total of 300 tracks were investigated and scored for their quality. Relative fractions are shown per dose. Adapted from Koch,2021[49].

The implemented tracking method does only contain tracks that can be traced to an initial mother cell. In addition, some tracks end before the end of the observation duration. This leads to two consequences: First, for each time point there are more objects present in a sample than there are tracked objects. Secondly, the number of tracked objects per sample (well or flask) either decreases or stays the same. The data analysis pipeline yields plots to monitor this procedure, allowing for quality control and for exploration of the influence of changes in the data analysis. Figures 12.2 and 12.3 exemplify this and show the fractions of tracked objects from all present objects as well as the number of tracked objects through time, respectively.

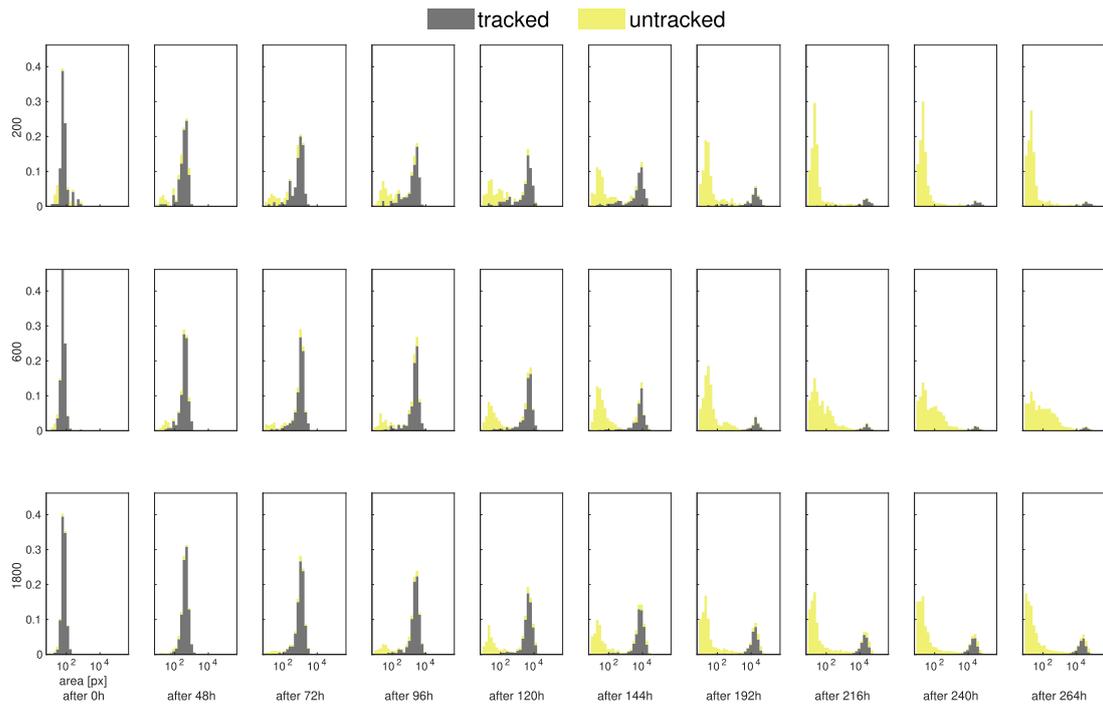


Figure 12.2: Relative frequency of colony sizes, grouped by tracked and untracked colonies. This data stems from an experiment to investigate the influence of seeding densities on growth without irradiation. The tracked portion corresponds to the distributions shown in 13.1. Cell line: H460.

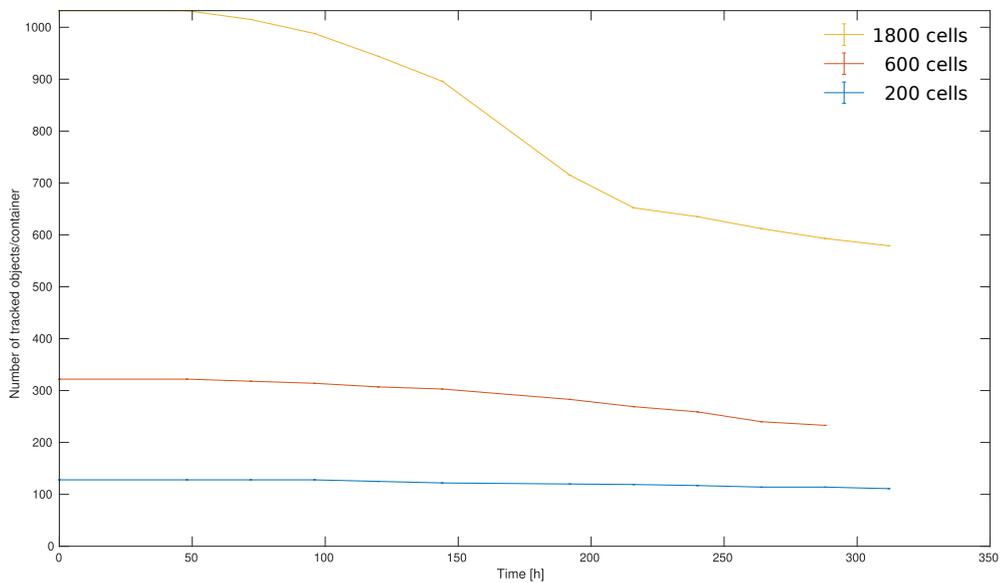


Figure 12.3: The number of tracked objects over time for three different conditions. As in Figure 12.2, this visualizes the number of tracked objects for three different seeding densities. Cell line: H460.

Growth at different seeding densities

As recently shown, seeding densities have an influence on the result of clonogenic assays[18]. To evaluate whether this parameter has a distinct effect on the cell lines used in this work, I compared unirradiated growth of cell colonies at three different concentrations for the four cell lines H460, RENCA, SAT and UTSCC-5. Figures 13.1, 13.2, 13.3 and 13.4 show size distributions through time for these cell lines.

Comparing these overview plots, it is clear that the cell lines show very different growth behaviors, even without irradiation treatment. H460 behaves according to the assumptions of the standard IVCA, showing a relatively narrow distribution of cell colony sizes at all time points, with a majority of colonies shifting towards larger sizes. The minority of smaller colonies seems to be distinguishable from the growing colonies at later time points, as the size distributions stretches into a relatively bimodal distribution. RENCA shows a very comparable behavior, potentially including more colonies which are "left behind" at smaller sizes. SAT and UTSCC-5 behave differently, with broader size distributions especially towards later time points. In some cases (UTSCC-5, 600 cells/flask, time points > 192 h and to lesser extent UTSCC-5, 200cells/flask, time points > 192 h) the distribution seems to split into a bimodal shape as well. For the highest concentration however, this is not the case. For SAT, the distribution stays relatively broad and a bimodal shape can not be discerned at any time point or concentration.

Absolute numbers of cells, as indicated within the panels, are always substantially lower (around 40 % too low) than the targeted numbers, indicating that either there was a systematic error in cell counting pre-seeding, or that a substantial fraction of the seeded cells did not attach before the first imaging session. Comparing the number of tracked objects in the last panel with the number of initially tracked objects in the first panel, another trend emerges: For all cell lines, higher concentrations have a lower fraction of initially tracked colonies that are still tracked at later points. Visual inspection of the image data reveals that at higher concentrations, more neighbouring colonies fuse, leading to track termination, as the resulting fused object can not be associated with a single initial cell anymore.

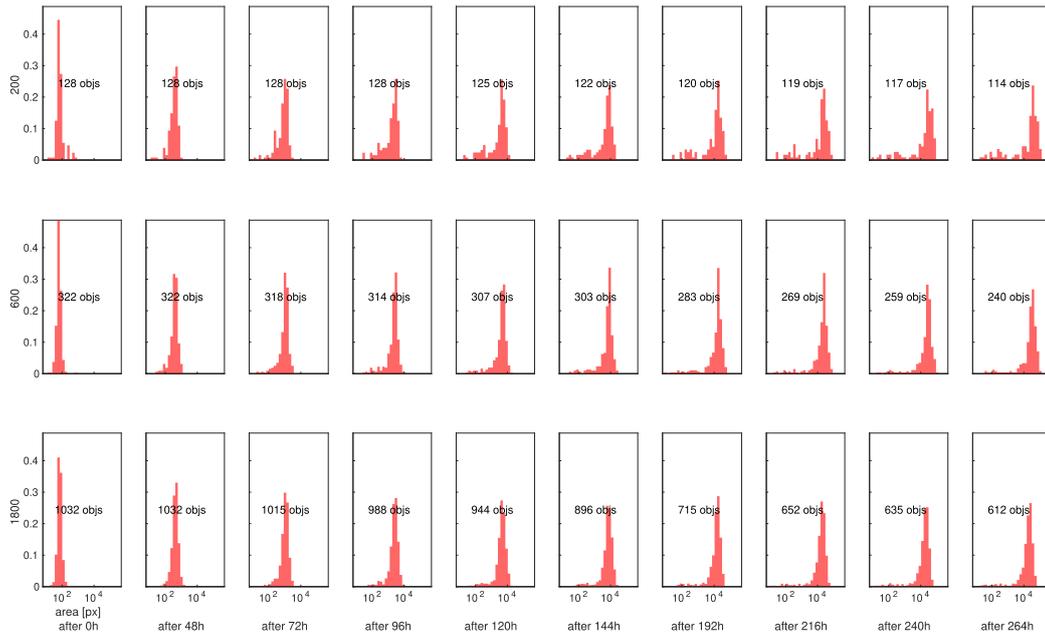


Figure 13.1: Colony sizes of tracked H460 cells at different seeding concentrations. Relative frequencies of cell colony sizes are shown at different times. Absolute cell numbers are given within each panel. The rows indicate the targeted number of cells per flask.

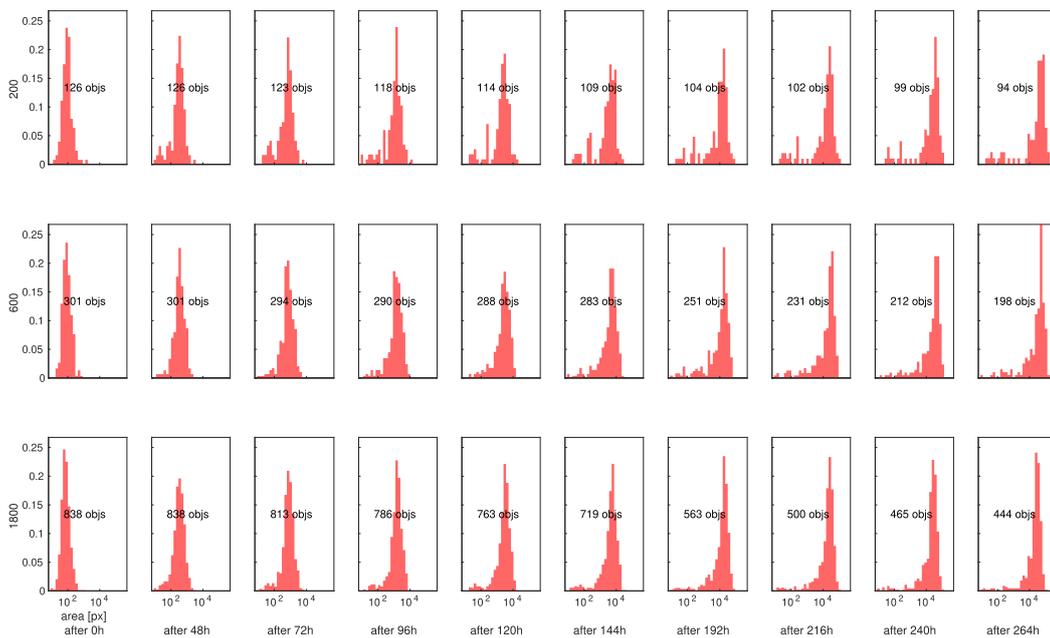


Figure 13.2: Colony sizes of tracked RENCA cells at different seeding concentrations. Relative frequencies of cell colony sizes are shown at different times. Absolute cell numbers are given within each panel. The rows indicate the targeted number of cells per flask.

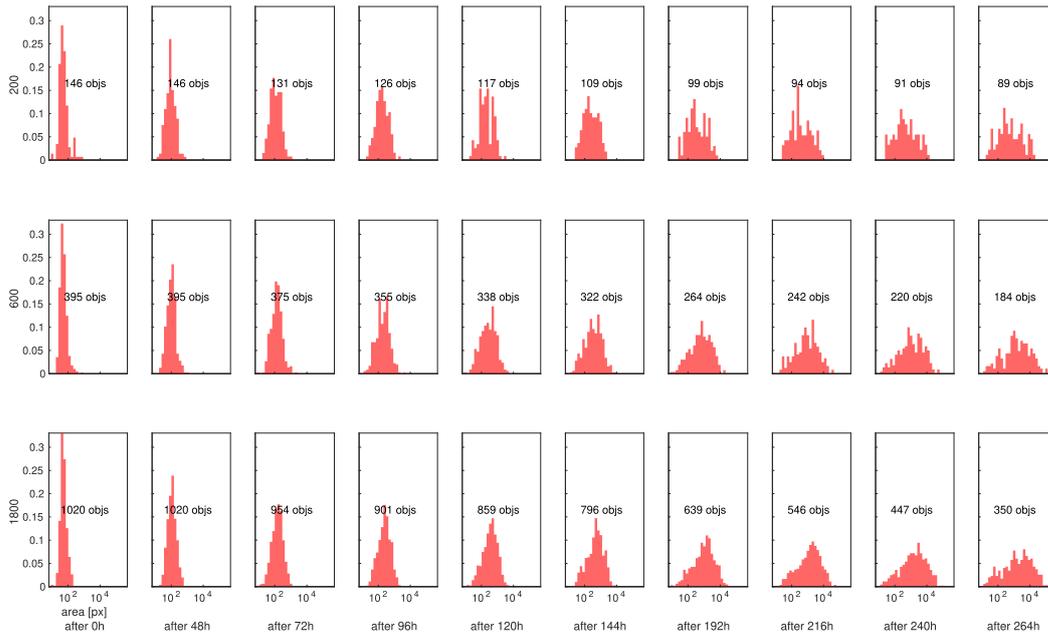


Figure 13.3: Colony sizes of tracked SAT cells at different seeding concentrations. Relative frequencies of cell colony sizes are shown at different times. Absolute cell numbers are given within each panel. The rows indicate the targeted number of cells per flask.

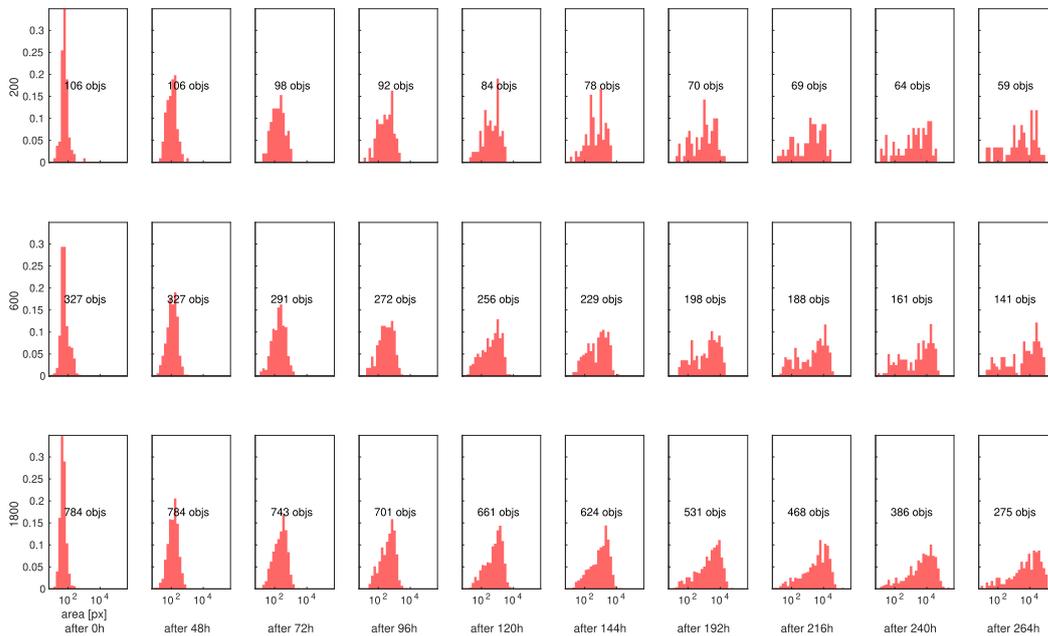


Figure 13.4: Colony sizes of tracked UTSCC-5 cells at different seeding concentrations. Relative frequencies of cell colony sizes are shown at different times. Absolute cell numbers are given within each panel. The rows indicate the targeted number of cells per flask.

Growth of different culture passages

It is known that prolonged cultivation and the concomitant repeated passaging rounds can change cell line characteristics, influencing metabolic processes [54] and DNA damage repair responses [55]. Hence I examined the growth dynamics of SAT and UTSCC-5 cell lines in a long-term experiment including eleven and nine consecutive passages, respectively. Since the volume of concurrently cultivated and imaged flasks was high, the number of time points was reduced to three: Directly after seeding, after four days and after ten days. In contrast to the regularly imaged samples for seeding density analysis (see section 13), tracking of colonies here was not reliable due to the long intervals between image cycles. Hence, Figures 14.1 and 14.2 depict the size distribution of all objects present in the flasks. For both cell lines, a global trend is discernible: The size distribution spreads towards larger sizes after four days, with slightly more objects present at that time point. After ten days however, the number of objects has approximately doubled for almost all samples, with the distribution showing a large peak at sizes corresponding to very small colonies of one to ten cells. This indicates the emergence of secondary colonies, which form by cells which detached from their original colony and settled at a different location. Visual inspection of the image data confirmed this observation, as shown in Figure 14.3. For both cell lines, the size distributions do not show a clear difference for different passage numbers. The only clear exception is passage "P7(2020)" from the UTSCC-5 cell line, which was seeded from a culture flask that had not received fresh medium for a period over two weeks. These cells clearly did not grow healthily. However, their successor passage "P8(2020)" seemed to have fully recovered from this "treatment". The SAT cells of passage "P9(2020)", sharing the same medium deprivation, did not show a clear reduction in growth.

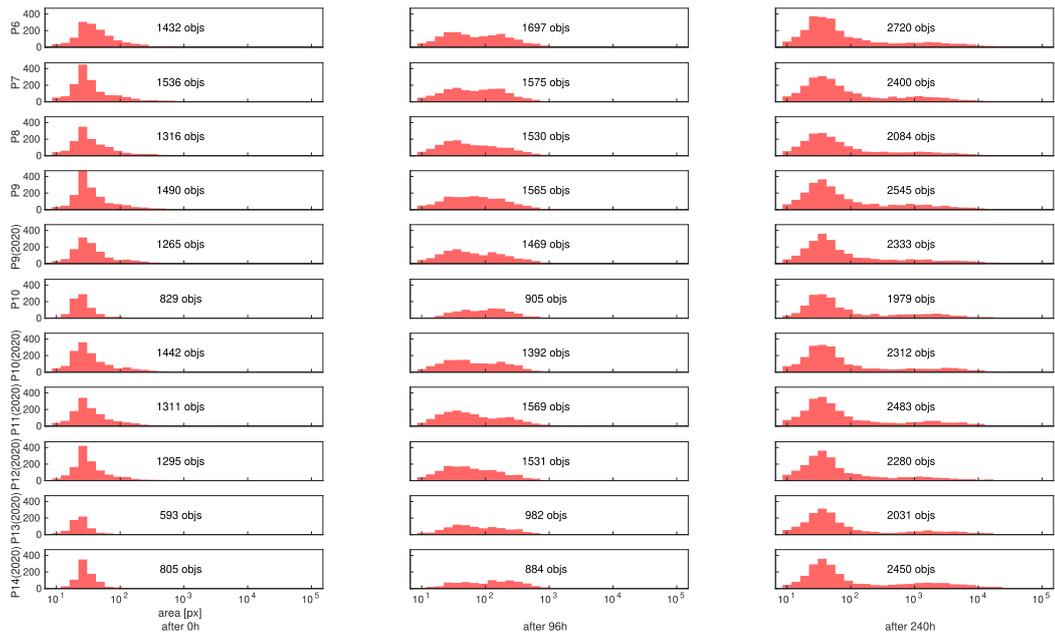


Figure 14.1: Colony sizes of SAT cells from different passages. All passage descriptions containing a "(2020)" string represent samples that are successors of cells which had a two-week period without medium changes over the christmas break.

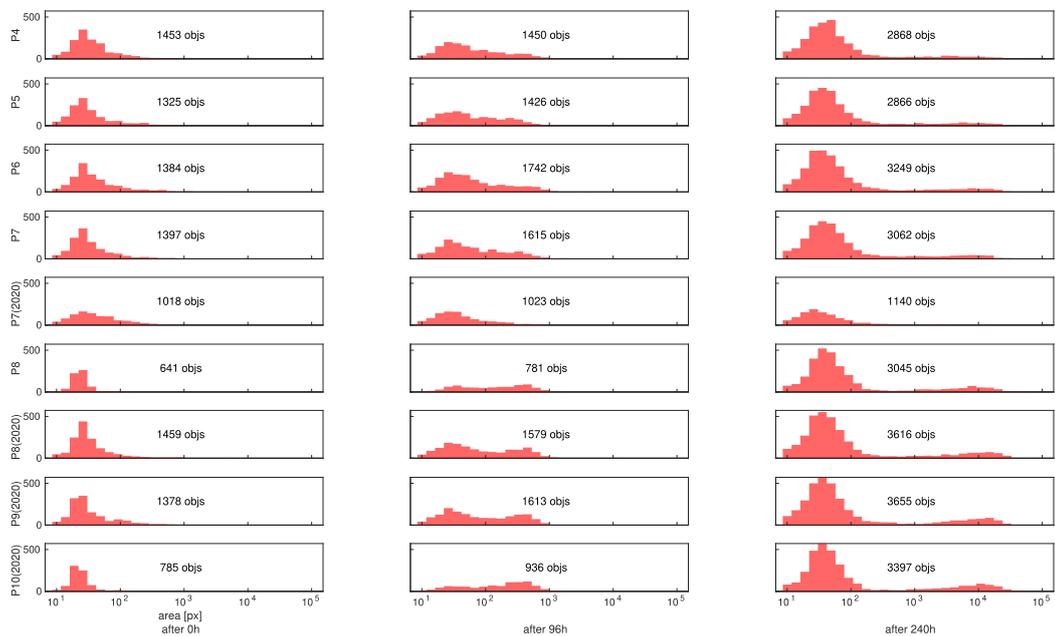


Figure 14.2: Colony sizes of UTSCC-5 cells from different passages. All passage descriptions containing a "(2020)" string represent samples that are successors of cells which had a two-week period without medium changes over the christmas break.

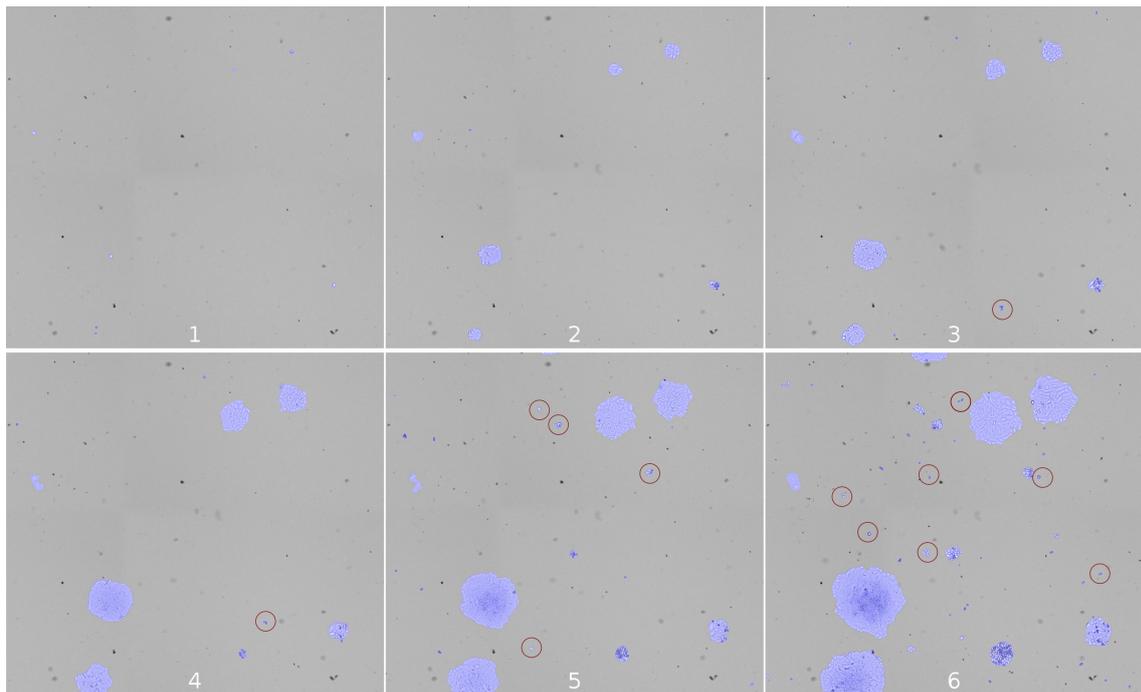


Figure 14.3: Emergence of secondary colonies. The numbered panels represent the same section of a culture flask at successive time points. Throughout the incubation time, secondary colonies (marked by red circles) can emerge, some of which disappear while others stay adherent and grow into colonies.

Photon irradiation influences growth rates

Growth rate extraction as described in section 9.2 revealed interesting quantitative results regarding the influence of photon irradiation on colony growth. These results were previously published in Koch,2021[49]. As shown in panels a) and b) of Figure 15.1, growth rates for H3122 and RENCA are normally distributed, independent of irradiation dose, across doses commonly used in clonogenic assays. In addition, the variance σ^2 of these distributions is large: The doubling rates corresponding to the growth rate values at $\mu - \sigma$ and $\mu + \sigma$ are 27 h and 48 h for H3122 and 20 h and 27 h for RENCA, respectively. Here, μ corresponds to the mean parameter of the Gaussian fits to the growth rate distributions as shown in Figure 15.1.

In addition, the growth rate distributions as well as Figure 16.3 show that with increasing dose, the proportion of delayed abortive colonies increases. Interestingly, these delayed abortive colonies show a similar distribution of growth rates with their exponentially growing counterparts for identical doses. In addition, the mean μ of the growth rate distributions for exponentially growing colonies decreases approximately linearly, as shown in Figure 15.1(c).

Initially abortive colonies show growth rates slightly below zero, reflecting the tendency of dead cells to shrink rather than stay at their original size.

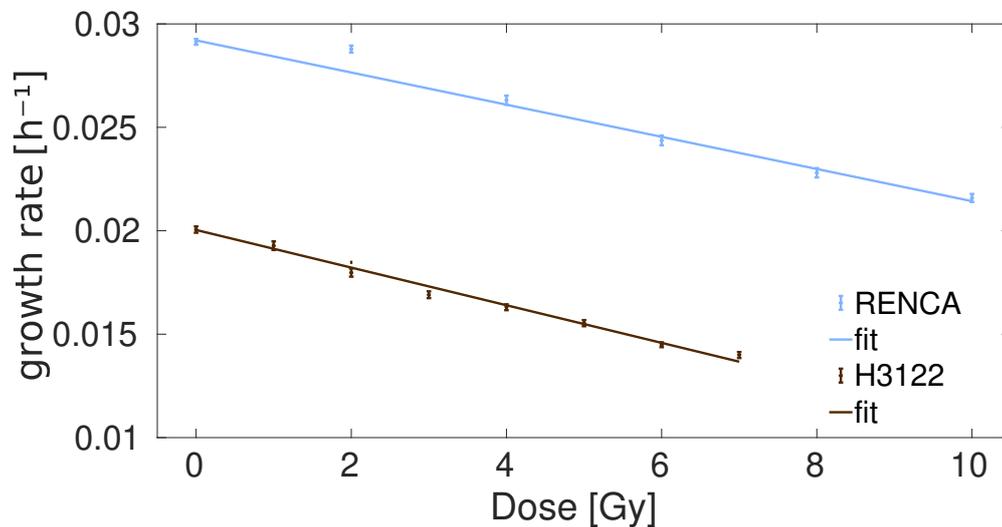
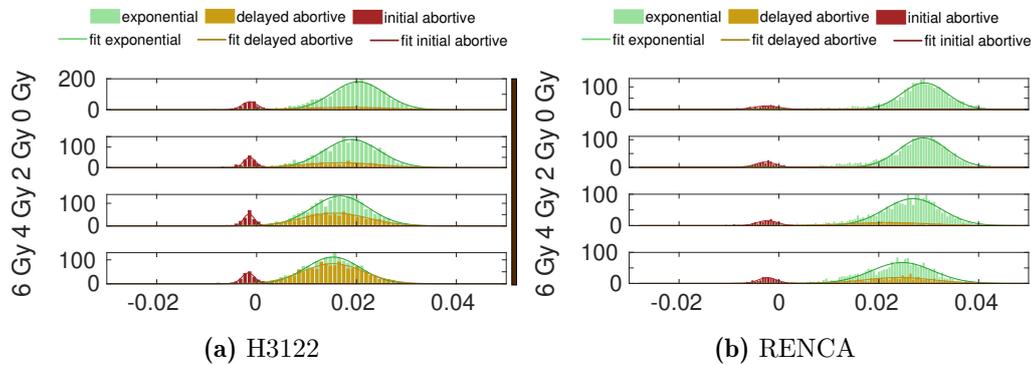


Figure 15.1: Distributions of growth rates at different doses. The depicted growth rates are grouped by three distinct growth behaviors as indicated in the legends. a) and b) show distributions for H3122 and RENCA cell lines, respectively, with Gaussian distributions fit to each distribution of growth rates. c) shows the mean values of the fitted distributions for exponential growth with error bars depicting the 95 % confidence intervals. Adapted from Koch,2021[49].

Photon irradiation influences growth behaviors

As already apparent from Figure 15.1, irradiation of H3122 or RENCA cells leads to the occurrence of delayed abortive cell colonies. Such colonies start out growing exponentially, but abruptly stop their growth after a certain time, after which they tend to shrink. Increasing photon irradiation doses increase the fraction of delayed abortive colonies, as depicted in Figure 16.2. While the level of this increase differs between H3122 and RENCA, both cell lines show substantial shares of colonies with this growth type. It is noteworthy that this behavior is present at relatively low levels (approximately 10%) in unirradiated H3122 samples and is generally more prevalent for this cell line at all comparable doses.

When matching tracked colonies with their growth behavior in size distribution plots as depicted in Figure 16.3, the explicit influence of delayed abortive colonies on traditional size based viability classification becomes apparent: At all doses and all time points, the distributions of exponentially growing and delayed abortive colonies overlap. Any attempt to distinguish non-viable, delayed abortive colonies from viable, exponentially growing colonies by means of a fixed size threshold will introduce a certain level of error. As the distributions differ between different doses, this error will most likely not occur to the same extent for all doses, introducing a dose-dependent bias in the resulting dose-depending survival curves. Figure 16.4 illustrates this by showing the level of falsely classified colonies depending on dose and readout time. While the misclassification severity between doses converges towards later timepoints, the effect is always there and always dose-dependent.

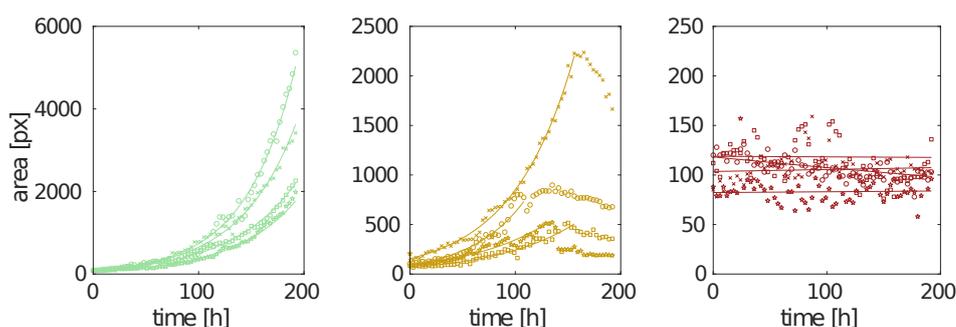
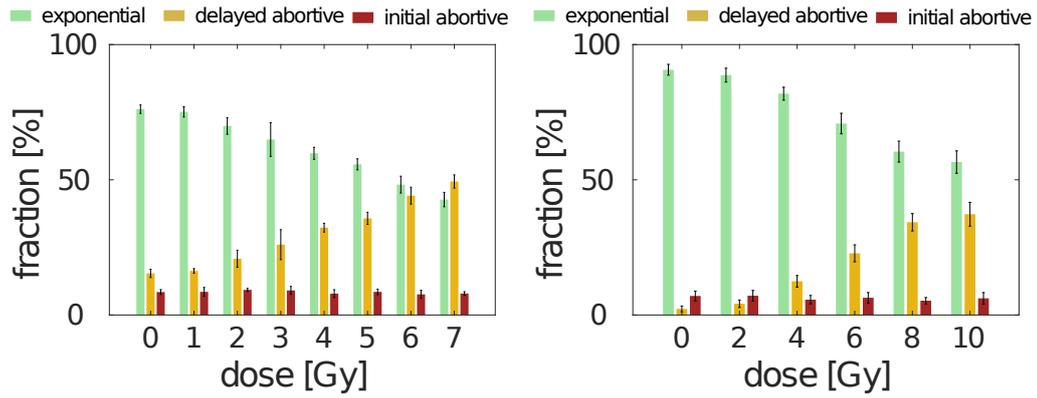


Figure 16.1: Examples of growth behaviors discovered. Crosses indicate colony sizes, lines represent fits to the exponential parts of the curves as described in section 9.2. RENCA as well as H3122 cells show three distinct growth behaviors: They either grow exponentially (left, green), or grow exponentially for a while before aborting growth, in this work called *delayed abortive* (middle, yellow), or they do not grow at all, in this work called *initial abortive* (right, red). Note the different y-axes.



(a) H3122

(b) RENCA

Figure 16.2: Dose-dependent fractions of growth behaviors. With increasing doses, the fraction of delayed abortive colonies increases, while the fraction of initially abortive colonies stays approximately constant. This is true for both H3122 (a) and RENCA (b) cell lines. Adapted from Koch,2021[49].

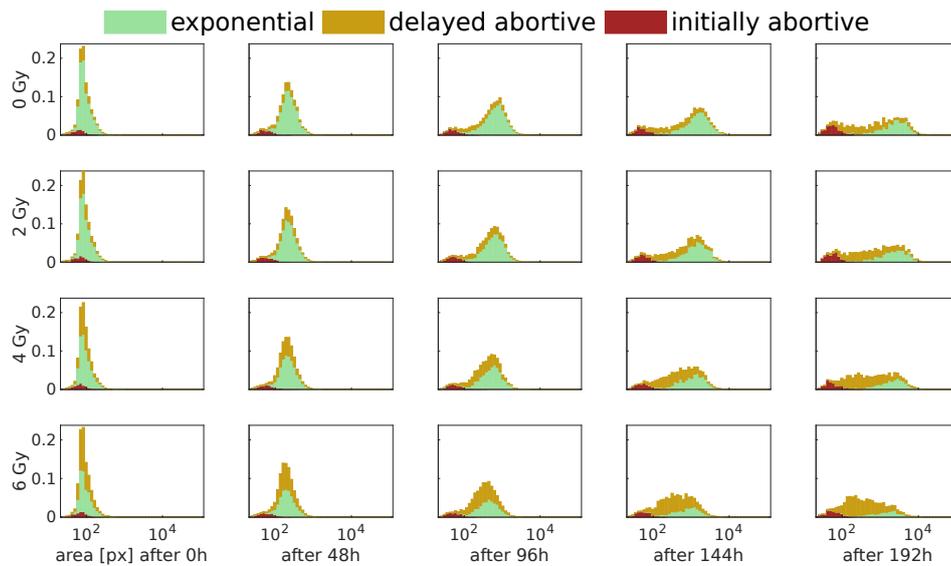


Figure 16.3: Cell colony sizes grouped by the three distinct growth behaviors. By mapping their respective growth behavior to each tracked colony, the influence of different growth types on size distributions is visualized. Cell line: H3122. Adapted from Koch,2021[49].

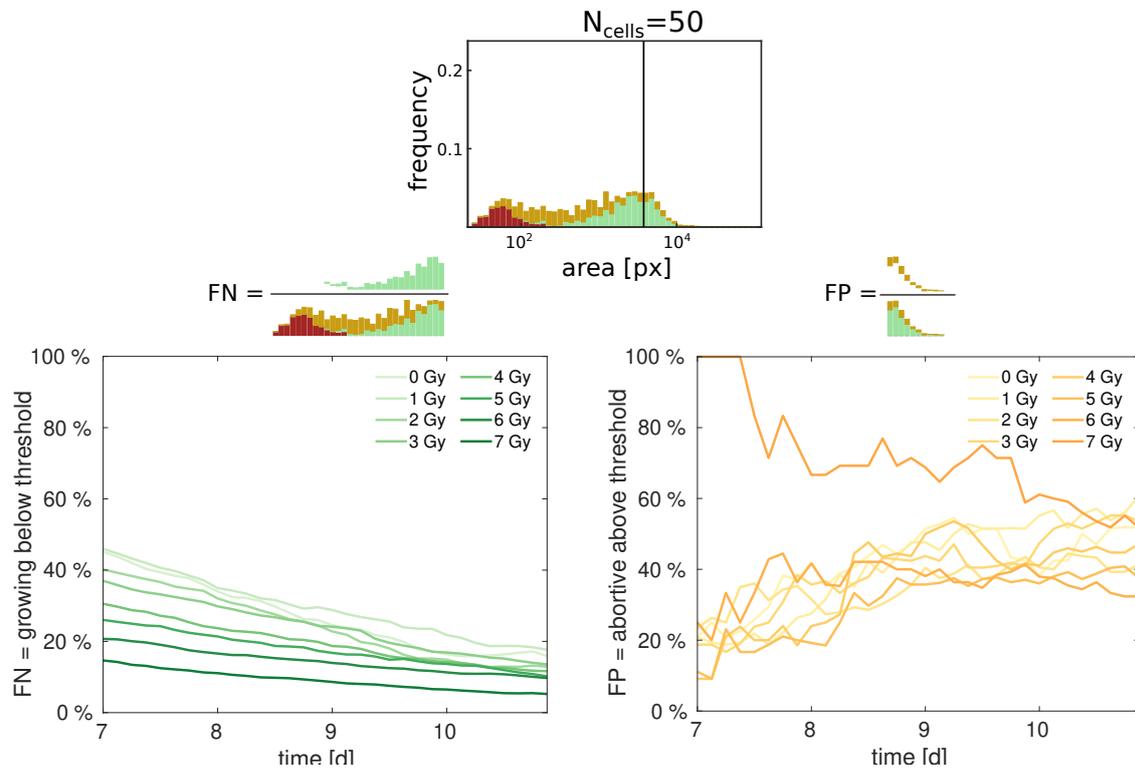


Figure 16.4: Illustration of dose- and readout-dependent viability misclassification by fixed size thresholds. Based on H3122 data as shown in Figure 16.3. A colony size threshold corresponding to approximately 50 cells splits the complete distribution of cell colonies into two parts, with the left side considered non-viable due to insufficient size and the right side viable since they are above the threshold. Growth behavior classification shows that some colonies on the left side are exponentially growing, while some colonies on the right side are abortive, despite their size. As depicted graphically, these two populations correspond to false negatives (FN) and false positives (FP), respectively. The plots show how these values change for different doses depending on the chosen readout time.

Readout choices influence survival curves

As indicated in Figure 16.4, the readout time affects the number of colonies considered viable. As depicted in Figures 17.1, 17.2 and 17.3, I calculated survival rates and fitted corresponding LQM curves for data from three cell lines, each represented two data sets, one of which was irradiated using photons, the other with carbon ions. Varying readout times as well as colony size thresholds introduces substantial variability in the resulting survival curves.

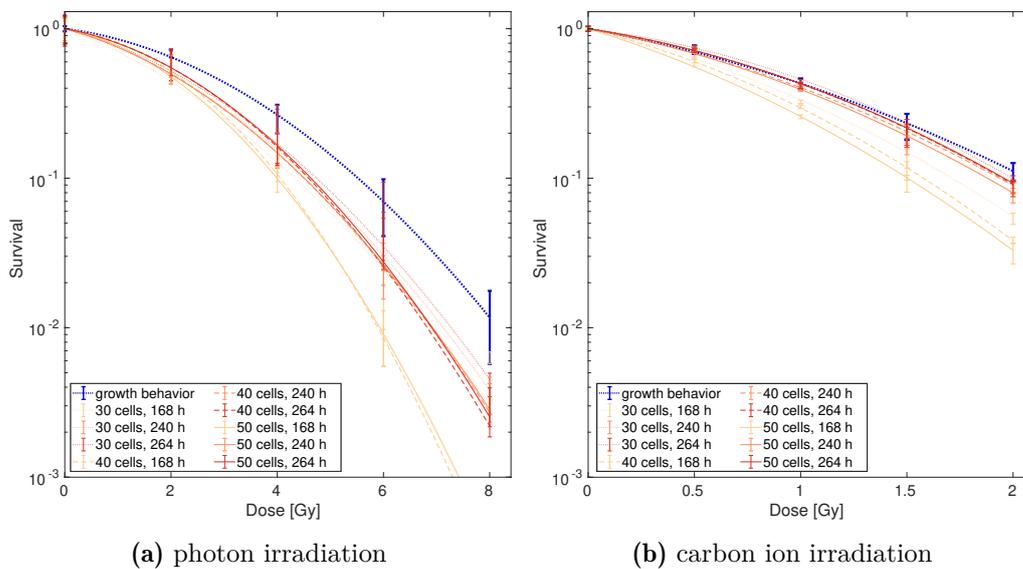


Figure 17.1: Survival fractions and fitted LQM curves for H460. Different readout times and cell colony size thresholds change the apparent dose-dependent survival of cells after photon (a) and carbon ion (b) irradiation. Growth behavior-based survival fractions and curves take into account the complete growth dynamics of colonies throughout the incubation time. Error bars represent mean \pm standard deviation of all samples. Adapted from Koch,2023[53].

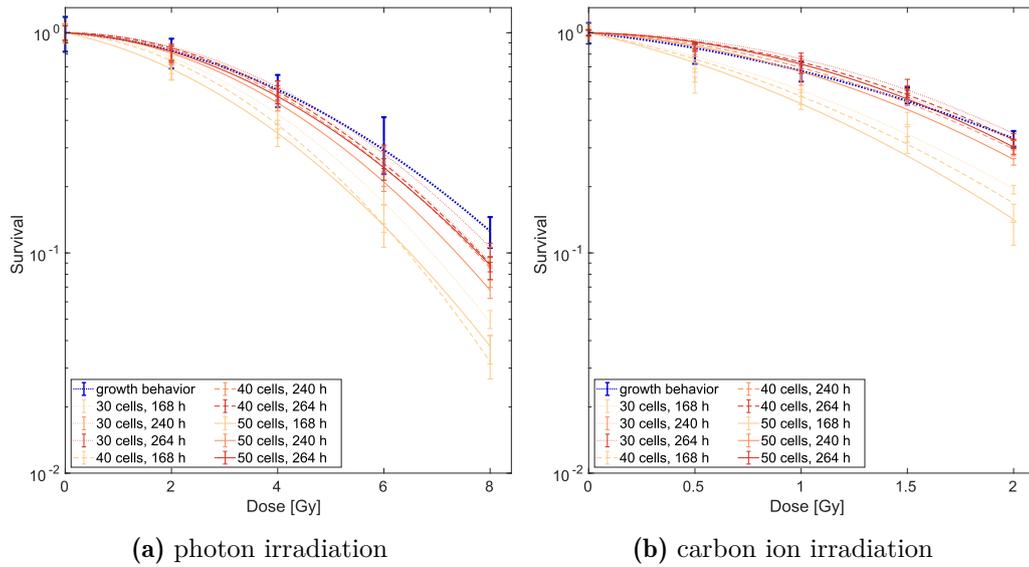


Figure 17.2: Survival fractions and fitted LQM curves for RENCA. Different readout times and cell colony size thresholds change the apparent dose-dependent survival of cells after photon (a) and carbon ion (b) irradiation. Growth behavior-based survival fractions and curves take into account the complete growth dynamics of colonies throughout the incubation time. Error bars represent mean \pm standard deviation of all samples. Adapted from Koch,2023[53].

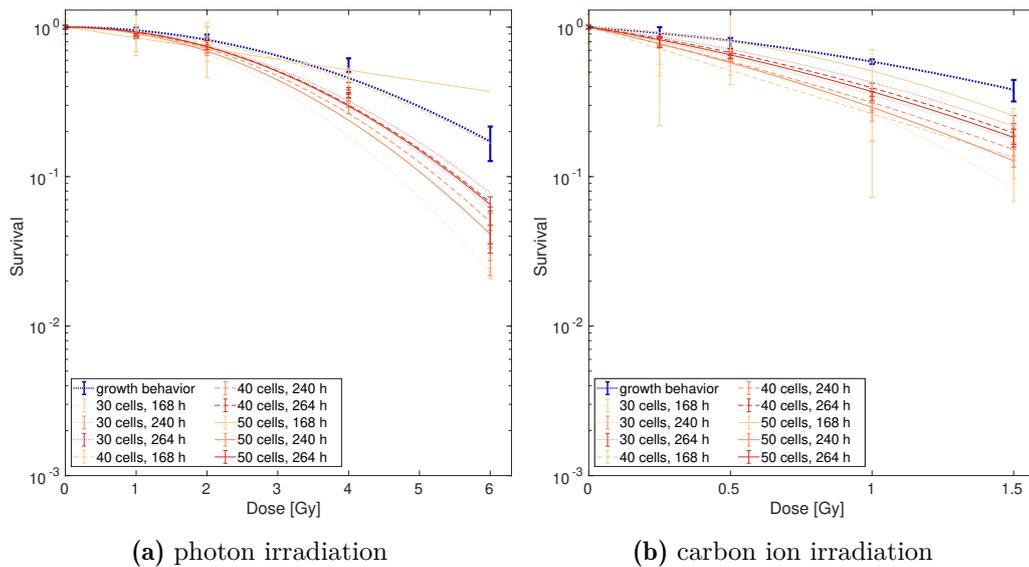


Figure 17.3: Survival fractions and fitted LQM curves for UTSCC-5. Different readout times and cell colony size thresholds change the apparent dose-dependent survival of cells after photon (a) and carbon ion (b) irradiation. Growth behavior-based survival fractions and curves take into account the complete growth dynamics of colonies throughout the incubation time. Error bars represent mean \pm standard deviation of all samples. Adapted from Koch,2023[53].

Readout choices influence RBE results

For the three cell lines discussed in the previous section, the characterization and quantification of survival rates and corresponding survival curves allows an examination of the influences of readout choices on downstream RBE quantification. As depicted in Figure 18.1, RBE values and their dependence on reference dose vary substantially between different choices for readout parameters. Figure 18.2 shows that in relation to the RBE curves acquired through the growth behavior-based approach, threshold-based estimates can both under- or overestimate RBE values. There is no clear trend discernible for all cell lines, but at the clinically relevant 2 Gy level a large variability can be found with values ranging from 2.8 to 3.6 for H460, from 2.5 to 3.5 for RENCA and from 2.0 to 6.1 for UTSCC-5. The low 2.0 RBE value for UTSCC-5, based on the combination of readout time 7 d and size threshold 50 cells is an outlier, with all other RBE dynamics behaving more similarly and more like the growth-behavior-based results, as shown in Figure 18.1(c).

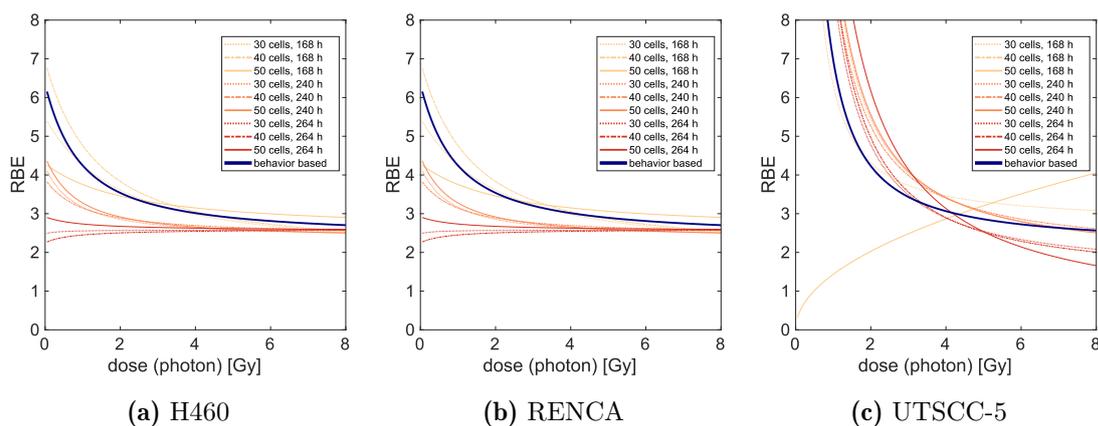


Figure 18.1: RBE of carbon ion irradiation as a function of reference dose for three cell lines. RBE trends are calculated from survival curve fits shown in Figures 17.1, 17.2 and 17.3 using Equation 9.10. Adapted from Koch,2023[53].

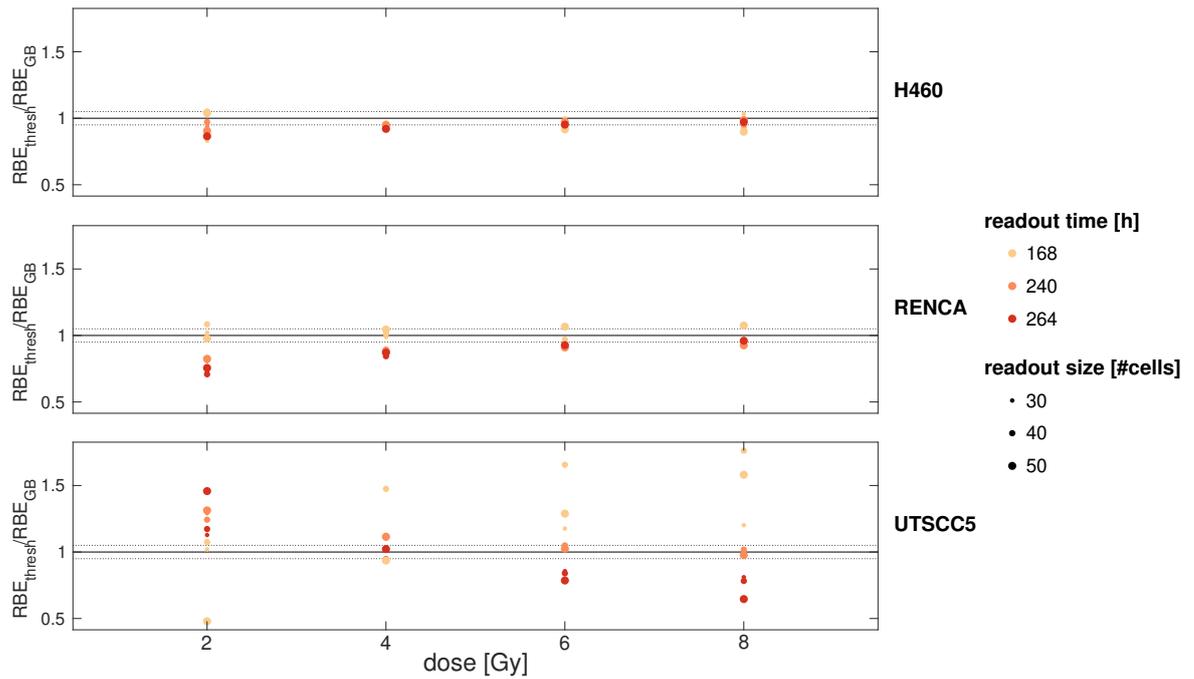


Figure 18.2: Threshold-based RBE values relative to growth-behavior based RBE values for all three cell lines. RBE_{thresh} defines threshold-based RBE values, RBE_{GB} defines growth behavior-based RBE values. For all reference doses and all cell lines, the quotient RBE_{thresh}/RBE_{GB} is depicted for each readout choice. The solid line in the middle represents identity and the dashed lines above and below show a $\pm 5\%$ interval. Adapted from Koch,2023[53].

Part IV

Discussion & Conclusion

**New perspectives from the
time-resolved IVCA method**
73

**Implications of biological
results** 75

**Potential, challenges & limi-
tations of the method** ... 79

**Subsequent / additional
analyses** 83

New perspectives from the time-resolved IVCA method

In my opinion, the greatest benefit of the presented method lies in the various new visual means to scrutinize cell colony growth in vitro. This includes the explorative inspection of single colonies throughout their complete growth history. This aspect allowed me to identify distinct growth behaviors and to gain hands-on (and eyes-on) insights into how these colonies behave. It allows researchers to browse through large numbers of colonies and develop ideas and hypotheses on the processes playing out in the cell culture dishes. In contrast to traditional seed-and-wait approaches, the storage of image data in connection with segmentation and tracking information allows this exploration retrospectively, calmly and repeatedly.

Naturally, visually observing growing colonies only helps so much in quantifying the observations. Therefore, the variety of new types of plots help to make sense of the bigger picture:

Colony size distribution plots (e.g Figure 13.1) represent an overview over complete datasets, without stripping away essential details of the growth dynamics. The representation of size distributions migrating from small, single cell colonies to larger colony sizes over time is an intuitive depiction of how the whole colony population behaves. Direct visual comparison between different conditions allows insights into differential population behavior depending on external factors. The option to map colony growth behaviors to these plots allowed me to understand why the presence and frequency of delayed abortive colonies has such profound effects on traditional IVCA readouts, potentially introducing dose-dependent biases.

Even though the initial motivation of this project was to scrutinize and find potential flaws and systemic biases in the traditional IVCA approach, I was allowed multiple additional findings which were only possible through detailed data extraction and quantitative analysis of the processes during incubation: I could only discover the significance of highly variable growth rates leading to large variability in resulting colony sizes by gathering these growth rates systematically and visualizing their distributions as presented in Figure 15.1. Again, direct comparison between different conditions, in this case doses, unveiled a surprisingly clear and yet undiscovered connection between irradiation dose and rates of growth.

Further visual representations such as Figures 12.2 and 14.1 allowed to detect an additional process that can potentially strongly influence quantitative analyses of cell culture flasks: Large numbers of secondary colonies forming after some days, as discussed in section 21.1.

Depictions of the number of tracked cells throughout the experiment duration can be used as a quality control tool for the experimental steps preceding the tracking: Irregularities in imaging, image registration or colony segmentation will undoubtedly show up as irregularities in the plot in the form of steep changes in the number of

tracked objects, as exemplified in figure 19.1.

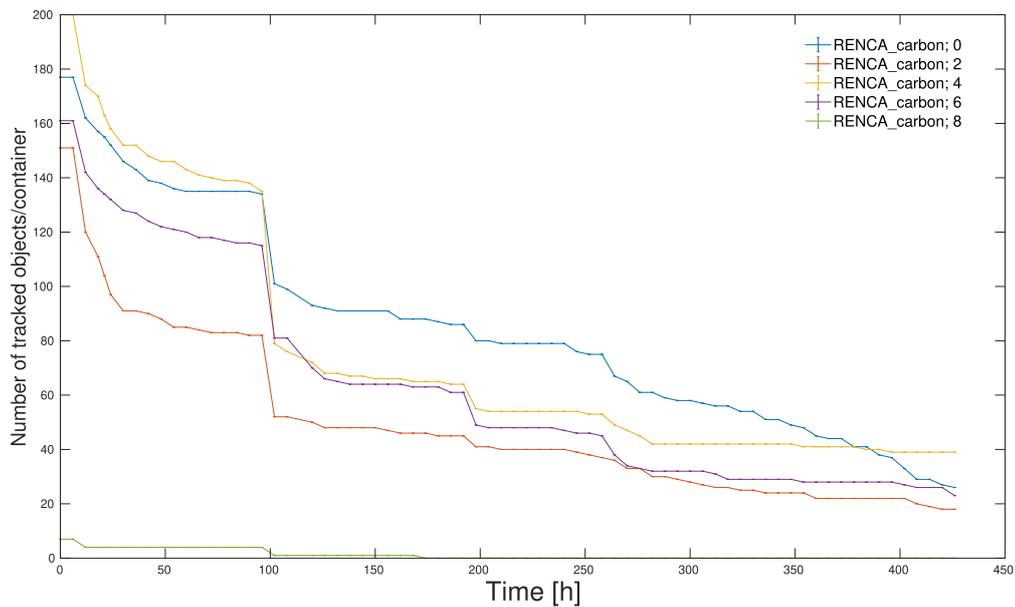


Figure 19.1: Example of quality control by examination of numbers of tracked objects. A sharp drop in tracked objects across all conditions indicates a problem in one of the preceding steps.

In this case, visual inspection of the image data revealed that at the time of the steep change across all conditions, the IncuCyte did not find the focus plane for one of three replicates for all conditions, leading to a complete dropout of these replicates. This way, problems are detected that might otherwise go unnoticed. In an ideal scenario, such quality control steps would happen "online", so that faulty samples could be sorted out immediately or fixed. This would avoid the expense of additional physical and computational resources needed to create and process ultimately useless data.

Additional varieties of quantification and visualization not presented here could be informative in some applications. These could be distributions of temporary growth rates, relating the size of each tracked colony to their size in the previous and/or following frame. Another option is the display of growth rate distributions or colony size distributions grouped by track length or spatial position within wells. All of these opportunities are feasible within this new analysis framework.

Implications of biological results

20.1 Variability between cell lines

As apparent from sections 13 and 14, different cell lines show very different growth behaviors, even without irradiation. This indicates that for each cell line that can potentially be used in this time-resolved growth analysis or the standard IVCA, a thorough inspection of growth behavior should be performed. To this end, the presented method delivers a valuable toolbox to scrutinize cell colony growth as described in section 19. It also shows that some cell lines, such as SAT or UTSCC-5 are potentially more prone to systematic bias in the traditional IVCA, since their growth characteristics appear to be not as distinct as for example for H460 or RENCA. While the two latter cell lines grow fast, with low variability and high PE. SAT and UTSCC on the other hand seem to grow slower, with higher variability among the individual growth rates and a lesser PE. Unfortunately, further studies of these cell lines including irradiated samples were not performed, but it would be interesting to see whether these trends seen for unirradiated samples are attenuated or potentially amplified under irradiation treatment.

20.2 Seeding densities

The influence of seeding densities on growth characteristics as assessed via colony size distributions in section 13 is negligible for the four tested cell lines. It needs to be mentioned however, that only a range of around one order of magnitude was covered in the presented experiments. Hence we cannot exclude the possibility that influences become apparent at seeding densities outside of this range. One crucial aspect to consider here is that in standard IVCA, samples receiving high doses are often seeded in densities some orders of magnitude higher than for the unirradiated samples. This strategy aims at retaining sufficient numbers of surviving colonies to score, even at doses that leave a large majority of cells non-clonogenic. When applying such a strategy, one should be careful and assess the potential influences of seeding densities on colony growth behavior beforehand. The presented method qualifies as an appropriate tool to achieve that.

Another important aspect to consider is that the higher the seeding density, the more like it is for colonies to fuse. Even in the relatively sparsely seeded culture flasks used in the respective experiments, the number of fusions increased markedly at higher densities, reducing the number of colonies where the full growth history is available.

20.3 Passage numbers

The influence of passage numbers on growth characteristics of unirradiated cells as outlined in 14 is negligible for the tested cell lines and the range of passage numbers. Potential influences on irradiated cells or influences of larger differences between the passage numbers can not be excluded, but the results suggest that experiments based on immortalized cell lines can be compared, if the difference in age is not more than two months.

A bycatch finding here is that for some cell lines (i.e. UTSCC-5), harsh environmental stress as caused by prolonged deprivation from fresh medium can lead to transient changes in growth behavior, in this case growth arrest. This is in line with literature findings that point out the influence of cell culture conditions on cell viability assessment [19].

20.4 Growth rates

I think the discovery of normally distributed growth rates exhibiting large variability is one of the most impactful findings of this work. Given the large variance among growth rates from the same cell type, complete classification of viable, growing cells by means of their final colony size requires very long incubation times for even the slowest growing colonies to reach this threshold. This is likely to introduce difficulties in scoring of the fastest growing colonies: Due to their large size, they potentially fuse with other colonies, biasing the subsequent colony count. To avoid this, only sparse seeding helps, which in turn reduces the sample size and hence the statistical power of the experimental results. Hence, the high variability in growth rates forces the experimenter to find a compromise including incubation time, seeding density and size threshold.

The observation that these growth rate distributions shift towards slower rates as a function of irradiation dose only aggravates this situation. It effectively broadens the range of growth rates present in any given experiment. In addition to above-mentioned need to find a suitable compromise, the standard IVCA now also faces the risk of dose-dependent bias in clonogenicity scoring, as depicted in Figure 16.4. Nevertheless, the growth rate findings are limited to two cell lines (H3122 and RENCA) and could only be observed in data from relatively densely seeded 96well plates, with images acquired in the IncuCyte. Since other types of data acquired in this project are not suited to extract growth rates in the same manner due to insufficiently dense data acquisition (see sections 9.2 and 21.5) these results could not be reproduced in other cell lines, for less dense seeding conditions or different cell culture containers. However, as literature findings indicate, highly variable growth rate distributions are found in other biological systems as well [56] and the increase of generation times after irradiation was shown previously for small numbers of manually inspected cell colonies [57, 58, 59]. Hence I expect these trends to be found in other cell types and conditions as well, once the practical limitations are overcome.

20.5 Growth behaviors

The observation of the delayed abortive growth type is another major outcome within this projects. As opposed to the growth rate results, I could confirm this finding in all five tested cell lines. The seminal IVCA paper by Puck and Marcus already describes how HeLa cell colonies continue to grow for some time before they abort growth [6]. In addition, Endlich et al. also describes the presence of cell colonies that only stop growing after a couple of generations in three additional cell lines[23]. Both of those findings are qualitative descriptions based on examination of single colonies. In my first publication, I quantified the frequency of this growth behavior as a function of dose for the first time [49].

The presence of delayed abortive colonies and their dose-dependent prevalence adds to the list of factors that complicate the standard readout procedure, as in my investigations these colonies can grow large enough to be considered viable based on their size.

Regarding the biological mechanisms inducing this behavior, a probable explanation is the cell's attempt to repair DNA damage inflicted by irradiation. A fascinating observation is the concerted stop of proliferation among all cells within a colony. If we assumed that the decision to switch from a state of repair and continued proliferation to growth termination is based on stochastic regulatory processes within single cells, I would expect the decision to stop to be spread across a certain time period for all cells within a colony. Based on their common progenitor cell, this decision would most likely be correlated, but would still result in a gradual reduction in proliferation of a colony due to different cels aborting at different times. In contrast, we see abrupt switches from exponentially growing colonies to stagnant colonies. This tight temporal correlation of growth abort indicates that either the cells within a colony share a communicating regulatory network that allows this switch-like behavior, or that the decision is already predestined based on the internal state of the progenitor cell after irradiation.

Potential, challenges & limitations of the method

The presented method uncovers a lot of so far mostly unnoticed, but valuable information. This information can help us to understand technical limitations of the well-established IVCA. It can also warn us about biases and lack of robustness introduced by inappropriate readouts. However, in my opinion this is just the start: Tapping into this information by time-resolved image analysis is a first methodological step towards a more general line of thinking: We can appreciate single colonies as individual entities that can make concerted decisions on their fate based on internal regulation mechanisms as well as external influences. This empirical observation uncovers a multitude of follow-up questions: How does a heap of cells decide on whether they should abort growth or not? How does it decide on the time of abort? What are the key influences on these decisions? How stochastic/determined is this behavior? If highly determined, are there predictors that can be clearly associated with growth abort? How early during growth can we identify these predictors? These questions might be answered by combining the presented method with additional techniques, as described in section 22.

Despite the demonstrated utility of this method and the potential extensions, the approach itself encompasses a multi-step process, with many technical and experimental design decisions that certainly influence its reliability and usability. In the following sections I will touch upon some of the challenges and limitations that come with this method.

21.1 Cell culture

Standardized cell culture procedures are a key factor to ensure robust experimental results and comparability between experiments. Hence, the following aspects need to be considered when attempting to use this method successfully. Most of those aspects are shared with the traditional IVCA: Cell lines that can be used in this approach are restricted to adherently growing types which form clear colonies.

Seeded cells need to be single cells. If a substantial number of duplets or cell clumps containing more cells are seeded, the result is biased.

The number of cells seeded needs to strike a balance between sufficiently large sample sizes and sufficient space per colony to grow without fusing with neighboring colonies. An important additional factor to achieve this is to ensure even distribution of cells after seeding. This can be optimized by avoiding medium convection within cell culture containers, by special shaking methods and by avoiding seeding from high viscosity media into low viscosity media.

I have observed the presence of secondary colonies after some days of incubation. In the traditional assay, these colonies have the capacity to be scored as viable colonies,

if they grow fast enough. Based on the high variability in growth rates, it is likely that some of those colonies surpass the threshold and are indistinguishable from slow growing, primary colonies. The influence of this process was not systematically examined in this work and it is possible that the formation of secondary colonies is a peculiar attribute of the presented setup, where repeated imaging and transport from and to the incubator lead to detachment of single cells due to mechanical stress. The presented tracking approach filters out those colonies by design, as it only considers tracks starting from single cells present in the first image. However, the presence of secondary colonies can still disturb the tracking of primary colonies if they fuse.

21.2 Imaging

Working with both the IncuCyte and the Axio CellObserver imaging systems, I had the opportunity to experience benefits and shortcomings of both highly automated, "convenience first" and manual, semi-automated, "customizability first" image acquisition approaches. Hence I was able to learn a lot about focus finding procedures, time management, and optimization of imaging procedures in terms of image quality but also efficiency.

As the Axio CellObserver system was highly customizable, initial problems with focus finding and incorrectly stitched images could be solved by switching to manual focus finding procedures and by implementing an ImageJ macro to outsource the stitching task and allow consistent image quality. Nevertheless, this approach required substantial amounts of time from the user, approximately 15 imaging hours per standard data set encompassing five conditions imaged in 24 h intervals over ten days.

In contrast, the same dataset in the IncuCyte required approximately 10 of setup time, with no need for additional work from the experimenter, despite substantially shorter imaging intervals (3 to 6 h). In this case however, there was no way to adjust either the focus finding procedures or aspects of image processing, such as stitching. While the stitching worked robustly, the the IncuCyte could not find the focus for sparsely seeded cell samples. This prohibited the useful application of the IncuCyte to the type of sparsely seeded, long-term imaging experiment I planned for. In addition, free time slots for imaging were rarely available for the IncuCyte system. In an ideal world, an imaging system combining the ease of use of the IncuCyte with the customizability of the Axio CellObserver would bring together the benefits of both. Until then, I would suggest focusing on automated imaging systems such as the IncuCyte or the CellCyte system, as they allow dense imaging intervals, which benefits the quality of all subsequent data processing and analysis steps, from registration to tracking, change point detection and growth behavior classification. A potential option to overcome the prohibitive lack of focus finding in the IncuCyte could be the introduction of inert, high-contrast structures in the samples to attract the focus finding procedure even at low seeding densities. Examples for this could be marks on the growth surface or adherent metal beads. This was not tested in this project, but could present a future option to allow robust *and* convenient imaging.

Nevertheless, if these systems are not available or can not be used as intended, manual semi-automated imaging at 24 h intervals is still a valid option to examine cell colony growth in vitro.

21.3 Image analysis

As described above, the data in this project came from two different imaging systems, using different imaging modalities, magnification factors and cell culture containers. Therefore, I developed two highly independent postprocessing pipelines. To allow consistent production of high quality data and focused development and extension of the analysis framework, commitment towards one system is highly recommended. This also facilitates comparability between different data sets, since differences due to data acquisition and postprocessing characteristics can be avoided.

For new cell lines, the performance of the established segmentation model needs to be thoroughly tested on this cell line, as the model might not have encountered specific characteristics of this cell line in the former training data. If segmentation is suboptimal, two solutions are possible: Either adapt the existing model by retraining it on a combination of the old training data and newly created training data which includes the specific characteristics of the new cell line. Alternatively, train a completely new model based solely on training data for the new cell line. The first approach favours generalizability of the model, but needs to be reevaluated on the old data as well. The second approach will most likely perform better on the new data, but lacks generalizability. When working with many cell lines, this approach leads to a set of many small, specialized segmentation models.

21.4 Threshold-based survival rate and cell counts

If the time-resolved clonogenic assay should be used to examine results as if they were produced by the standard IVCA, the mapping between colony sizes and numbers of cells in a colony is accurate. Since different cell lines have different sizes, especially when growing flatly on a surface, the mapping should be newly determined for every new cell line. One should also be aware that the linear mapping between colony sizes and cell numbers as postulated in this work does not necessarily work for all cell lines. Investigation of larger colonies also indicates that at some point the cells start to "heap" in a three-dimensional mound. At this point the latest, the linear mapping is not accurate anymore. Even though the linear mappings are quite accurate in the size ranges considered, the regression data (e.g. 8.4) still shows considerable variation for all sizes.

I would argue that the extent of this uncertainty is small in comparison to the uncertainties introduced by delayed abortive colonies and slow growing exponential colonies. Nevertheless, an imaging modality and segmentation procedure that would allow to gain actual cell counts from the image data would increase the discriminative power of the presented approach. In contrast to our estimated cell numbers, human

observers working on stained colonies are most likely more accurate in their cell count assessment.

21.5 Growth curve analysis

As I have experienced with the Axio CellObserver data, sparse sampling (every 24 h) of colony growth dynamics still allows tracking as well as growth type classification when using the methods established here. The general dynamics of growth are still captured for the examined cell lines. Nevertheless, since the detection of change points does not work robustly at this level of sample density, the extraction of meaningful growth rates is not possible for this type of data. Therefore, I suggest that future applications of this method employ sampling at least every 12 h or shorter.

Another point to consider for new cell lines is the fact that the assumption of finding three distinct growth behaviors might not hold true. Therefore, explorative inspection of individual cell colony image data in combination with their growth curves is necessary to test whether these behaviors are indeed present for the new cell line. In case the assumption holds true, analogously to the segmentation model, the growth behavior classification needs to be adjusted to this cell line. Again, either by combining new training data from the new cell line and old training data from the old cell line to retrain a generalizing classifier, or by training a new, cell line specific classifier on new training data. Based on the variability of growth rates and abort time found in the data analysed so far, I assume that the second approach is more promising.

Subsequent / additional analyses

22.1 Examining the influence of cell cycle status at irradiation

Multiple sources have shown an influence of the cell cycle status at irradiation time on cell survival [28, 60]. Since this difference might very well be mediated by a change in growth behaviors after irradiation, the idea to apply the presented method to cells irradiated at different cell cycle phases. Obtaining cells at different cell cycle states can generally be achieved by two approaches. One way is to perform synchronization of whole cell populations by chemical blocks, the other is sorting a population of cells according to their current cell cycle status. Chemical synchronization techniques arrest cells in a certain state, as they are hindered from progressing further through the cell cycle [61]. However, some authors argue that this artificial block does not really synchronize cells but rather align them according to a certain property that does not reflect a "natural" cell cycle status [62]. The second approach requires a method to determine the cell cycle status for single cells and sort them accordingly. Usually this is done by FACS after staining cells for cell cycle state-specific markers. However, these stains are usually lethal. Since my goal is to cultivate cells after sorting, the method has to be as gentle as possible, ideally not impairing cell viability at all. According to the literature, staining with Hoechst33342 for subsequent DNA content analysis is the best option to retain viability in cells after staining [63]. Based on the stoichiometric binding of the stain to DNA, the relative DNA content of stained cells can be determined and used to distinguish G1, S and G2 phases as depicted in Figure 22.1.

In some pilot experiments the staining as well as sorting worked well, but for cell lines RENCA, SAT and UTSCC-5, the cells did not attach and grow afterwards. This indicates that either the staining or the FACS procedure drastically decreased their viability. For H460, the sorted cells grew, but the imaging procedure did not work as the IncuCyte could not find the focus. Therefore, unfortunately no results could be presented for this type of analysis.

22.2 Discovery of predictive proteomic profiles

Another potential extension of the presented method aims at elucidating candidate proteins or proteomic profiles with predictive power towards the growth behavior fate of a cell colony. The idea is to perform the time-resolved imaging method after irradiation of single colonies as described in this work. The extension consist of single cell extraction from small colonies (4-8 cells) by physical picking from the container they grow in. The origin colony for each picked cell is recorded and imaging is continued on the sample. The picked cell is immediately transferred into lysis buffer

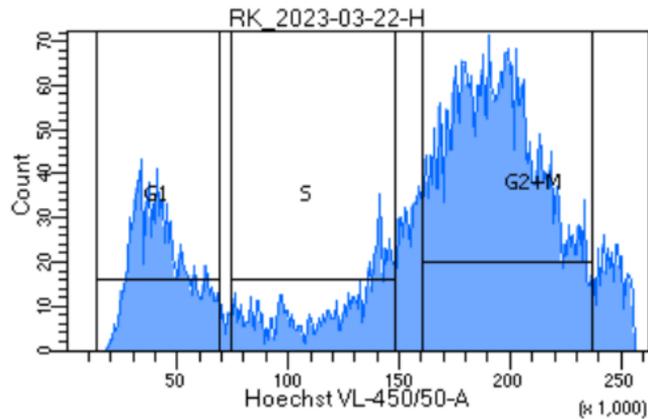


Figure 22.1: Histogram of Hoechst intensity in stained H460 cells. The vertical lines represent thresholds splitting the population into three subgroups of G1, S and G2/M phase. Based on these thresholds, single cells can be sorted into different containers for subsequent experiments.

and analyzed for its proteomic profile. By correlating the resulting proteomic profile with growth characteristics such as growth rate, growth behavior, time of delayed abort in case of delayed abortive colonies or other attributes of the origin colony, some predictive power for the colony fate might be found in the proteomic profile. Such findings might be used to form hypotheses on which regulatory processes and states are involved in the cell colony's fate.

Pilot experiments for single cell picking were performed using the CellCelector Flex (Sartorius, Göttingen, Germany) and single cells could indeed be picked while keeping the origin colony intact.

Part V

Appendix and Bibliography

Bibliography	87
List of Figures	95
List of Tables	97
Acronyms	99

Bibliography

- [1] Eric J. Hall and Amato J. Giaccia. *Radiobiology for the radiologist*. en. 7th ed. OCLC: 752207373. Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins, 2012. ISBN: 978-1-60831-193-4.
- [2] Stephen Joseph McMahon. “The linear quadratic model: usage, interpretation and challenges”. en. In: *Physics in Medicine and Biology* 64.1, Dec. 2018, 01TR01. ISSN: 1361-6560. DOI: [10.1088/1361-6560/aaf26a](https://doi.org/10.1088/1361-6560/aaf26a). URL: <http://stacks.iop.org/0031-9155/64/i=1/a=01TR01?key=crossref.f43b33f838087719610cef2c8061440b> (visited on 02/05/2020).
- [3] Stephen J McMahon and Kevin M Prise. “Mechanistic Modelling of Radiation Responses”. en. In: *Cancers* 11.2, Feb. 2019, page 205. ISSN: 2072-6694. DOI: [10.3390/cancers11020205](https://doi.org/10.3390/cancers11020205). URL: <http://www.mdpi.com/2072-6694/11/2/205> (visited on 05/05/2021).
- [4] Theresa Suckert et al. “Models for Translational Proton Radiobiology—From Bench to Bedside and Back”. en. In: *Cancers* 13.16, Aug. 2021, page 4216. ISSN: 2072-6694. DOI: [10.3390/cancers13164216](https://doi.org/10.3390/cancers13164216). URL: <https://www.mdpi.com/2072-6694/13/16/4216> (visited on 08/27/2021).
- [5] H Nikjoo et al. “Energy deposition in small cylindrical targets by ultrasoft X-rays”. en. In: *Physics in Medicine and Biology* 34.6, June 1989, pages 691–705. ISSN: 0031-9155, 1361-6560. DOI: [10.1088/0031-9155/34/6/005](https://doi.org/10.1088/0031-9155/34/6/005). URL: <https://iopscience.iop.org/article/10.1088/0031-9155/34/6/005> (visited on 05/12/2021).
- [6] Theodore T. Puck and Philip I. Marcus. “Action of X-rays on Mammalian Cells”. en. In: *The Journal of Experimental Medicine* 103.5, May 1956, pages 653–666. ISSN: 1540-9538, 0022-1007. DOI: <https://doi.org/10.1084/jem.103.5.653>. URL: <https://rupress.org/jem/article/103/5/653/2233/ACTION-OF-XRAYS-ON-MAMMALIAN-CELLS> (visited on 02/05/2020).
- [7] Nicolaas A P Franken et al. “Clonogenic Assay of Cells in vitro”. en. In: *Nature Protocols* 1.5, Dec. 2006, pages 2315–2319. ISSN: 1754-2189, 1750-2799. DOI: [10.1038/nprot.2006.339](https://doi.org/10.1038/nprot.2006.339). URL: <http://www.nature.com/articles/nprot.2006.339> (visited on 02/05/2020).
- [8] W. Friedland et al. “Comprehensive track-structure based evaluation of DNA damage by light ions from radiotherapy-relevant energies down to stopping”. en. In: *Scientific Reports* 7.1, May 2017, page 45161. ISSN: 2045-2322. DOI: [10.1038/srep45161](https://doi.org/10.1038/srep45161). URL: <http://www.nature.com/articles/srep45161> (visited on 05/04/2021).
- [9] Atsushi Shibata and Penny Jeggo. “A historical reflection on our understanding of radiation-induced DNA double strand break repair in somatic mammalian cells; interfacing the past with the present”. en. In: *International Journal of Radiation Biology* 95.7, July 2019, pages 945–956. ISSN: 0955-3002, 1362-3095. DOI: [10.1080/09553002.2018.1564083](https://doi.org/10.1080/09553002.2018.1564083). URL: <https://www.tandfonline.com/doi/full/10.1080/09553002.2018.1564083> (visited on 02/05/2020).

- [10] Steffen Nielsen et al. “Differential gene expression in primary fibroblasts induced by proton and cobalt-60 beam irradiation”. en. In: *Acta Oncologica* 56.11, Nov. 2017, pages 1406–1412. ISSN: 0284-186X, 1651-226X. DOI: [10.1080/0284186X.2017.1351623](https://doi.org/10.1080/0284186X.2017.1351623). URL: <https://www.tandfonline.com/doi/full/10.1080/0284186X.2017.1351623> (visited on 02/05/2020).
- [11] Steffen Nielsen et al. “Comparison of Coding Transcriptomes in Fibroblasts Irradiated With Low and High LET Proton Beams and Cobalt-60 Photons”. en. In: *International Journal of Radiation Oncology Biology Physics* 103.5, Apr. 2019, pages 1203–1211. ISSN: 03603016. DOI: [10.1016/j.ijrobp.2018.11.065](https://doi.org/10.1016/j.ijrobp.2018.11.065). URL: <https://linkinghub.elsevier.com/retrieve/pii/S0360301618340677> (visited on 02/05/2020).
- [12] Reingard Senekowitsch-Schmidtke et al. “Tumor Cell Spheroids as a Model for Evaluation of Metabolic Changes After Irradiation”. en. In: 1998.
- [13] Maria Saager et al. “Determination of the proton RBE in the rat spinal cord: Is there an increase towards the end of the spread-out Bragg peak?” en. In: *Radiotherapy and Oncology* 128.1, July 2018, pages 115–120. ISSN: 01678140. DOI: [10.1016/j.radonc.2018.03.002](https://doi.org/10.1016/j.radonc.2018.03.002). URL: <https://linkinghub.elsevier.com/retrieve/pii/S0167814018301348> (visited on 06/11/2021).
- [14] L. O. Jacobson et al. “Recovery from Radiation Injury”. en. In: *Science* 113.2940, May 1951, pages 510–511. ISSN: 0036-8075, 1095-9203. DOI: [10.1126/science.113.2940.510](https://doi.org/10.1126/science.113.2940.510). URL: <https://www.science.org/doi/10.1126/science.113.2940.510> (visited on 08/21/2023).
- [15] Toshiaki Matsui et al. “Robustness of Clonogenic Assays as a Biomarker for Cancer Cell Radiosensitivity”. en. In: *International Journal of Molecular Sciences* 20.17, Aug. 2019, page 4148. ISSN: 1422-0067. DOI: [10.3390/ijms20174148](https://doi.org/10.3390/ijms20174148). URL: <https://www.mdpi.com/1422-0067/20/17/4148> (visited on 08/19/2022).
- [16] Endang Nuryadi et al. “Inter-assay precision of clonogenic assays for radiosensitivity in cancer cell line A549”. en. In: *Oncotarget* 9.17, Mar. 2018, pages 13706–13712. ISSN: 1949-2553. DOI: [10.18632/oncotarget.24448](https://doi.org/10.18632/oncotarget.24448). URL: <https://www.oncotarget.com/lookup/doi/10.18632/oncotarget.24448> (visited on 01/22/2021).
- [17] Harald Paganetti. “Relative biological effectiveness (RBE) values for proton beam therapy. Variations as a function of biological endpoint, dose, and linear energy transfer”. en. In: *Physics in Medicine and Biology* 59.22, Nov. 2014, R419–R472. ISSN: 0031-9155, 1361-6560. DOI: [10.1088/0031-9155/59/22/R419](https://doi.org/10.1088/0031-9155/59/22/R419). URL: <http://stacks.iop.org/0031-9155/59/i=22/a=R419?key=crossref.65e4182cbde19f19bc18ee5fba783bef> (visited on 02/05/2020).

- [18] Nikko Brix et al. “The clonogenic assay: robustness of plating efficiency-based analysis is strongly compromised by cellular cooperation”. en. In: *Radiation Oncology* 15.1, Dec. 2020, page 248. ISSN: 1748-717X. DOI: [10.1186/s13014-020-01697-y](https://doi.org/10.1186/s13014-020-01697-y). URL: <https://ro-journal.biomedcentral.com/articles/10.1186/s13014-020-01697-y> (visited on 01/22/2021).
- [19] Peter Buhl Jensen et al. “Reduced variation in the clonogenic assay obtained by standardization of the cell culture conditions prior to drug testing on human small cell lung cancer cell lines”. en. In: *Investigational New Drugs* 7, Nov. 1989, pages 307–315. DOI: <https://doi.org/10.1007/BF00173760>.
- [20] Matthew Armstrong Lumley et al. “Colony counting is a Major Source of Variation in CFU-GM Results Between Centres”. en. In: *British Journal of Haematology* 97.2, May 1997, pages 481–484. ISSN: 0007-1048, 1365-2141. DOI: [10.1046/j.1365-2141.1997.492695.x](https://doi.org/10.1046/j.1365-2141.1997.492695.x). URL: <https://onlinelibrary.wiley.com/doi/abs/10.1046/j.1365-2141.1997.492695.x> (visited on 02/05/2020).
- [21] Marie-Claude Biston et al. “An Objective Method to Measure Cell Survival by Computer-Assisted Image Processing of Numeric Images of Petri Dishes”. en. In: *Physics in Medicine and Biology* 48.11, June 2003, pages 1551–1563. ISSN: 0031-9155, 1361-6560. DOI: [10.1088/0031-9155/48/11/305](https://doi.org/10.1088/0031-9155/48/11/305). URL: <http://stacks.iop.org/0031-9155/48/i=11/a=305?key=crossref.67368adf70154e27862cbd82d80b0a56> (visited on 02/05/2020).
- [22] Kenneth Chu et al. “Computerized Video Time Lapse Study of Cell Cycle Delay and Arrest, Mitotic Catastrophe, Apoptosis and Clonogenic Survival in Irradiated 14-3-3 and CDKN1A (p21) Knockout Cell Lines”. en. In: *Radiation Research* 162.3, Sept. 2004, pages 270–286. ISSN: 0033-7587, 1938-5404. DOI: [10.1667/RR3221](https://doi.org/10.1667/RR3221). URL: <http://www.bioone.org/doi/10.1667/RR3221> (visited on 02/05/2020).
- [23] Brian Endlich et al. “Computerized Video Time-Lapse Microscopy Studies of Ionizing Radiation-Induced Rapid-Interphase and Mitosis-Related Apoptosis in Lymphoid Cells”. en. In: *Radiation Research* 153.1, Jan. 2000, pages 36–48. ISSN: 0033-7587, 1938-5404. DOI: [10.1667/0033-7587\(2000\)153\[0036:CVTLMS\]2.0.CO;2](https://doi.org/10.1667/0033-7587(2000)153[0036:CVTLMS]2.0.CO;2). URL: <http://www.bioone.org/doi/abs/10.1667/0033-7587%282000%29153%5B0036%3ACVTLMS%5D2.0.CO%3B2> (visited on 02/05/2020).
- [24] Helen B Forrester et al. “Using Computerized Video Time Lapse for Quantifying Cell Death of X-irradiated Rat Embryo Cells Transfected with c-myc or c-Ha-ras”. en. In: *Cancer Research* 59, Feb. 1999, pages 931–939.
- [25] Geraldine Prieur-Carrillo et al. “Computerized Video Time-Lapse (CVTL) Analysis of the Fate of Giant Cells Produced by X-Irradiating EJ30 Human Bladder Carcinoma Cells”. en. In: *Radiation Research* 159.6, June 2003, pages 705–712. ISSN: 0033-7587, 1938-5404. DOI: [10.1667/RR3009](https://doi.org/10.1667/RR3009). URL: <http://www.bioone.org/doi/abs/10.1667/RR3009> (visited on 02/05/2020).

- [26] M. M. Elkind, Harriet Sutton, and Willie B. Moses. “Postirradiation Survival Kinetics of Mammalian Cells Grown in Culture”. en. In: *Journal of Cellular and Comparative Physiology* 58.S1, Dec. 1961, pages 113–134. ISSN: 0095-9898, 1553-0809. DOI: [10.1002/jcp.1030580412](https://doi.org/10.1002/jcp.1030580412). URL: <http://doi.wiley.com/10.1002/jcp.1030580412> (visited on 02/05/2020).
- [27] M.M. Elkind, Antun Han, and Kathleen W. Volz. “Radiation Response of Mammalian Cells Grown in Culture - IV. Dose Dependence of Division Delay and Postirradiation Growth of Surviving and Nonsurviving Chinese Hamster Cells”. In: *Journal of The National Cancer Institute* 30.4, Apr. 1963, pages 705–721. DOI: <https://doi.org/10.1093/jnci/30.4.705>.
- [28] Warren K. Sinclair. “Cyclic X-Ray Responses in Mammalian Cells *in Vitro*¹”. en. In: *Radiation Research* 178.2, Aug. 2012, AV112–AV124. ISSN: 0033-7587, 1938-5404. DOI: [10.1667/RRAV09.1](https://doi.org/10.1667/RRAV09.1). URL: <http://www.bioone.org/doi/10.1667/RRAV09.1> (visited on 05/10/2022).
- [29] Ramon Lopez Perez et al. “DNA damage response of clinical carbon ion versus photon radiation in human glioblastoma cells”. en. In: *Radiotherapy and Oncology* 133, Apr. 2019, pages 77–86. ISSN: 01678140. DOI: [10.1016/j.radonc.2018.12.028](https://doi.org/10.1016/j.radonc.2018.12.028). URL: <https://linkinghub.elsevier.com/retrieve/pii/S0167814019300027> (visited on 02/05/2020).
- [30] Aisling B. Heeran, Helen P. Berrigan, and Jacintha O’Sullivan. “The Radiation-Induced Bystander Effect (RIBE) and its Connections with the Hallmarks of Cancer”. en. In: *Radiation Research* 192.6, Oct. 2019, page 668. ISSN: 0033-7587. DOI: [10.1667/RR15489.1](https://doi.org/10.1667/RR15489.1). URL: <https://bioone.org/journals/radiation-research/volume-192/issue-6/RR15489.1/The-Radiation-Induced-Bystander-Effect-RIBE-and-its-Connections-with/10.1667/RR15489.1.full> (visited on 10/27/2022).
- [31] Michael Joiner and Albert van der Kogel, editors. *Basic Clinical Radiobiology*. en. 4th edition. London: Hodder Arnold, 2009. ISBN: 978-0-340-92966-7.
- [32] Kevin M Prise et al. “New insights on cell death from radiation exposure”. en. In: *The Lancet Oncology* 6.7, July 2005, pages 520–528. ISSN: 14702045. DOI: [10.1016/S1470-2045\(05\)70246-1](https://doi.org/10.1016/S1470-2045(05)70246-1). URL: <https://linkinghub.elsevier.com/retrieve/pii/S1470204505702461> (visited on 02/05/2020).
- [33] G. W. Barendsen. “Responses Of Cultured Cells, Tumours, And Normal Tissues To Radiations Of Different Linear Energy Transfer.” English. In: *pp 293-356 of Current Topics in Radiation Research. Vol. IV. Ebert, Michael Howard, Alma (eds.). New York, John Wiley and Sons, Inc., 1968.*, Oct. 1968. Institution: Health Research Organization TNO, Rijswijk, Netherlands. URL: <https://www.osti.gov/biblio/4500126> (visited on 06/05/2023).

- [34] Elisabeth Mara et al. “Investigating the impact of alpha/beta and LETd on relative biological effectiveness in scanned proton beams: An in vitro study based on human cell lines”. en. In: *Medical Physics* 47.8, 2020. _eprint: <https://aapm.onlinelibrary.wiley.com/doi/pdf/10.1002/mp.14212>, pages 3691–3702. ISSN: 2473-4209. DOI: <https://doi.org/10.1002/mp.14212>. URL: <https://aapm.onlinelibrary.wiley.com/doi/abs/10.1002/mp.14212> (visited on 01/15/2021).
- [35] N. T. Henthorn et al. “Clinically relevant nanodosimetric simulation of DNA damage complexity from photons and protons”. en. In: *RSC Advances* 9.12, 2019, pages 6845–6858. ISSN: 2046-2069. DOI: [10.1039/C8RA10168J](https://doi.org/10.1039/C8RA10168J). URL: <http://xlink.rsc.org/?DOI=C8RA10168J> (visited on 05/05/2021).
- [36] Lawrence Bronk et al. “Mapping the Relative Biological Effectiveness of Proton, Helium and Carbon Ions with High-Throughput Techniques”. en. In: *Cancers* 12.12, Dec. 2020, page 3658. ISSN: 2072-6694. DOI: [10.3390/cancers12123658](https://doi.org/10.3390/cancers12123658). URL: <https://www.mdpi.com/2072-6694/12/12/3658> (visited on 06/10/2021).
- [37] J.J. Broerse. “Survival of Cultured Human Cells after Irradiation with Fast Neutrons of Different Energies in Hypoxic and Oxygenated Conditions”. en. In: *International Journal of Radiation Biology and Related Studies in Physics, Chemistry and Medicine* 13.6, Jan. 1968, pages 559–572. ISSN: 0020-7616. DOI: [10.1080/09553006814550621](https://doi.org/10.1080/09553006814550621). URL: <http://www.tandfonline.com/doi/full/10.1080/09553006814550621> (visited on 08/24/2023).
- [38] Larry Bodgi et al. “Mathematical models of radiation action on living cells: From the target theory to the modern approaches. A historical and critical review”. en. In: *Journal of Theoretical Biology* 394, Apr. 2016, pages 93–101. ISSN: 00225193. DOI: [10.1016/j.jtbi.2016.01.018](https://doi.org/10.1016/j.jtbi.2016.01.018). URL: <https://linkinghub.elsevier.com/retrieve/pii/S0022519316000515> (visited on 05/06/2021).
- [39] C. M. van Leeuwen et al. “The alfa and beta of tumours: a review of parameters of the linear-quadratic model, derived from clinical radiotherapy studies”. en. In: *Radiation Oncology* 13.1, Dec. 2018, page 96. ISSN: 1748-717X. DOI: [10.1186/s13014-018-1040-z](https://doi.org/10.1186/s13014-018-1040-z). URL: <https://ro-journal.biomedcentral.com/articles/10.1186/s13014-018-1040-z> (visited on 02/05/2020).
- [40] Martin Oberholzer et al. “Methods in quantitative image analysis”. en. In: *Histochemistry and Cell Biology* 105.5, May 1996, pages 333–355. ISSN: 0948-6143, 1432-119X. DOI: [10.1007/BF01463655](https://doi.org/10.1007/BF01463655). URL: <http://link.springer.com/10.1007/BF01463655> (visited on 02/05/2020).
- [41] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. “ImageNet classification with deep convolutional neural networks”. en. In: *Communications of the ACM* 60.6, May 2017, pages 84–90. ISSN: 0001-0782, 1557-7317. DOI: [10.1145/3065386](https://doi.org/10.1145/3065386). URL: <https://dl.acm.org/doi/10.1145/3065386> (visited on 08/28/2023).

- [42] Laith Alzubaidi et al. “Review of deep learning: concepts, CNN architectures, challenges, applications, future directions”. en. In: *Journal of Big Data* 8.1, Mar. 2021, page 53. ISSN: 2196-1115. DOI: [10.1186/s40537-021-00444-8](https://doi.org/10.1186/s40537-021-00444-8). URL: <https://journalofbigdata.springeropen.com/articles/10.1186/s40537-021-00444-8> (visited on 08/28/2023).
- [43] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. “U-Net: Convolutional Networks for Biomedical Image Segmentation”. en. In: *arXiv:1505.04597 [cs]*, May 2015. arXiv: 1505.04597. URL: <http://arxiv.org/abs/1505.04597> (visited on 04/06/2020).
- [44] Fabian Isensee et al. “nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation”. en. In: *Nature Methods* 18.2, Feb. 2021, pages 203–211. ISSN: 1548-7091, 1548-7105. DOI: [10.1038/s41592-020-01008-z](https://doi.org/10.1038/s41592-020-01008-z). URL: <http://www.nature.com/articles/s41592-020-01008-z> (visited on 03/01/2021).
- [45] Bastien Rigaud et al. “Deformable image registration for radiation therapy: principle, methods, applications and evaluation”. en. In: *Acta Oncologica* 58.9, Sept. 2019, pages 1225–1237. ISSN: 0284-186X, 1651-226X. DOI: [10.1080/0284186X.2019.1620331](https://doi.org/10.1080/0284186X.2019.1620331). URL: <https://www.tandfonline.com/doi/full/10.1080/0284186X.2019.1620331> (visited on 08/28/2023).
- [46] Andrej-Nikolai Spiess and Natalie Neumeyer. “An evaluation of R2 as an inadequate measure for nonlinear models in pharmacological and biochemical research: a Monte Carlo approach”. en. In: *BMC Pharmacology* 10.1, Dec. 2010, page 6. ISSN: 1471-2210. DOI: [10.1186/1471-2210-10-6](https://doi.org/10.1186/1471-2210-10-6). URL: <http://link.springer.com/10.1186/1471-2210-10-6> (visited on 02/05/2020).
- [47] Joe Chalfoun et al. “MIST: Accurate and Scalable Microscopy Image Stitching Tool with Stage Modeling and Error Minimization”. en. In: *Scientific Reports* 7.1, Dec. 2017, page 4988. ISSN: 2045-2322. DOI: [10.1038/s41598-017-04567-y](https://doi.org/10.1038/s41598-017-04567-y). URL: <http://www.nature.com/articles/s41598-017-04567-y> (visited on 08/29/2022).
- [48] Manuel Guizar-Sicairos, Samuel T. Thurman, and James R. Fienup. “Efficient subpixel image registration algorithms”. en. In: *Optics Letters* 33.2, Jan. 2008, page 156. ISSN: 0146-9592, 1539-4794. DOI: [10.1364/OL.33.000156](https://doi.org/10.1364/OL.33.000156). URL: <https://www.osapublishing.org/abstract.cfm?URI=ol-33-2-156> (visited on 02/05/2020).
- [49] Robin A Koch et al. “A framework for automated time-resolved analysis of cell colony growth after irradiation”. en. In: *Physics in Medicine and Biology* 66.3, Feb. 2021, page 035017. ISSN: 0031-9155, 1361-6560. DOI: [10.1088/1361-6560/abd00d](https://doi.org/10.1088/1361-6560/abd00d). URL: <https://iopscience.iop.org/article/10.1088/1361-6560/abd00d> (visited on 02/26/2021).
- [50] napari contributors. “napari: a multi-dimensional image viewer for python”. In: <https://napari.org/stable/>, 2019. DOI: [10.5281/zenodo.3555620](https://doi.org/10.5281/zenodo.3555620).

- [51] Pranab Kumar Sen. “Estimates of the Regression Coefficient Based on Kendall’s Tau”. In: *Journal of the American Statistical Association* 63.324, Dec. 1968, pages 1379–1389. ISSN: 0162-1459. DOI: [10.1080/01621459.1968.10480934](https://doi.org/10.1080/01621459.1968.10480934). URL: <https://www.tandfonline.com/doi/abs/10.1080/01621459.1968.10480934> (visited on 02/06/2020).
- [52] Henri Theil and Henri Theil. “A Rank-Invariant Method of Linear and Polynomial Regression Analysis”. en. In: *Henri Theil’s Contributions to Economics and Econometrics*. Edited by Baldev Raj and Johan Koerts. Volume 23. Advanced Studies in Theoretical and Applied Econometrics. Dordrecht: Springer, 1992. DOI: [10.1007/978-94-011-2546-8_20](https://doi.org/10.1007/978-94-011-2546-8_20). URL: http://link.springer.com/10.1007/978-94-011-2546-8_20 (visited on 02/06/2020).
- [53] Robin A Koch et al. “A time-resolved clonogenic assay for improved cell survival and RBE measurements”. en. In: *Clinical and Translational Radiation Oncology* 42, Sept. 2023, page 100662. ISSN: 24056308. DOI: [10.1016/j.ctro.2023.100662](https://doi.org/10.1016/j.ctro.2023.100662). URL: <https://linkinghub.elsevier.com/retrieve/pii/S2405630823000873> (visited on 08/03/2023).
- [54] Jianhua Cao et al. “Uncovering the Effect of Passage Number on HT29 Cell Line Based on the Cell Metabolomic Approach”. en. In: *Journal of Proteome Research* 20.3, Mar. 2021, pages 1582–1590. ISSN: 1535-3893, 1535-3907. DOI: [10.1021/acs.jproteome.0c00806](https://doi.org/10.1021/acs.jproteome.0c00806). URL: <https://pubs.acs.org/doi/10.1021/acs.jproteome.0c00806> (visited on 09/05/2023).
- [55] Chin-Mei Chang-Liu and Gayle E Woloschak. “Effect of passagenumber on cellular response to DNA-damaging agents: cell survival and gene expression”. en. In: 1997.
- [56] Michel Godin et al. “Using buoyant mass to measure the growth of single cells”. en. In: *Nature Methods* 7.5, May 2010, pages 387–390. ISSN: 1548-7091, 1548-7105. DOI: [10.1038/nmeth.1452](https://doi.org/10.1038/nmeth.1452). URL: <http://www.nature.com/articles/nmeth.1452> (visited on 12/03/2021).
- [57] W K Sinclair and R A Morton. “X-Ray Sensitivity during the Cell Generation Cycle of Cultured Chinese Hamster Cells”. en. In: *Radiation Research* 29.3, Nov. 1966, pages 450–474.
- [58] G. F. Whitmore, J. E. Till, and S. Gulyas. “Radiation-Induced Mitotic Delay in L Cells”. en. In: *Radiation Research* 30.1, Jan. 1967, page 155. ISSN: 00337587. DOI: [10.2307/3572041](https://doi.org/10.2307/3572041). URL: <https://www.jstor.org/stable/3572041?origin=crossref> (visited on 02/05/2020).
- [59] G.P. Joshi et al. “Discrimination of Slow Growth from Non-survival Among Small Colonies of Diploid Syrian Hamster Cells After Chromosome Damage Induced by a Range of X-ray Doses”. en. In: *International Journal of Radiation Biology and Related Studies in Physics, Chemistry and Medicine* 42.3, Jan. 1982, pages 283–296. ISSN: 0020-7616. DOI: [10.1080/09553008214551201](https://doi.org/10.1080/09553008214551201). URL: <http://www.tandfonline.com/doi/full/10.1080/09553008214551201> (visited on 02/05/2020).

- [60] Timothy M. Pawlik and Khandan Keyomarsi. “Role of cell cycle in mediating sensitivity to radiotherapy”. en. In: *International Journal of Radiation Oncology*Biography*Physics* 59.4, July 2004, pages 928–942. ISSN: 03603016. DOI: [10.1016/j.ijrobp.2004.03.005](https://doi.org/10.1016/j.ijrobp.2004.03.005). URL: <https://linkinghub.elsevier.com/retrieve/pii/S0360301604003888> (visited on 02/05/2020).
- [61] Penny K. Davis, Alan Ho, and Steven F. Dowdy. “Biological Methods for Cell-Cycle Synchronization of Mammalian Cells”. en. In: *BioTechniques* 30.6, June 2001, pages 1322–1331. ISSN: 0736-6205, 1940-9818. DOI: [10.2144/01306rv01](https://doi.org/10.2144/01306rv01). URL: <https://www.future-science.com/doi/10.2144/01306rv01> (visited on 04/27/2022).
- [62] S. Cooper. “Rethinking synchronization of mammalian cells for cell cycle analysis”. en. In: *Cellular and Molecular Life Sciences* 60.6, June 2003, pages 1099–1106. ISSN: 1420-682X. DOI: [10.1007/s00018-003-2253-2](https://doi.org/10.1007/s00018-003-2253-2). URL: <http://link.springer.com/10.1007/s00018-003-2253-2> (visited on 04/27/2022).
- [63] Zbigniew Darzynkiewicz, Xuan Huang, and Hong Zhao. “Analysis of Cellular DNA Content by Flow Cytometry”. en. In: *Current Protocols in Immunology* 119.1, Nov. 2017. ISSN: 1934-3671, 1934-368X. DOI: [10.1002/cpim.36](https://doi.org/10.1002/cpim.36). URL: <https://currentprotocols.onlinelibrary.wiley.com/doi/10.1002/cpim.36> (visited on 09/06/2023).

List of Figures

4.1	Example of a standard LQM fitted to survival ratios	15
7.1	Cell culture containers	28
8.1	Workflow of intensity-based segmentation	32
8.2	Filtered objects based on different properties	35
8.3	Examples of colonies with countability indicated	35
8.4	Mapping of colony area to cell count	36
9.1	Pairwise distances between all objects in two subsequent frames	37
10.1	Registration of subsequent frames	45
11.1	Development of loss function and evaluation metric during nnUNet segmentation model training	48
11.2	Examples of hard segmentation tasks successfully solved by the nnUNet model	49
11.3	Examples of hard segmentation tasks where the model fails	50
12.1	Relative fractions of track quality labels for a RENCA dataset	51
12.2	Relative frequency of colony sizes, grouped by tracked and untracked colonies	52
12.3	The number of tracked objects over time for three different conditions	52
13.1	Colony sizes of tracked H460 cells at different seeding concentrations	54
13.2	Colony sizes of tracked RENCA cells at different seeding concentrations	54
13.3	Colony sizes of tracked SAT cells at different seeding concentrations	55
13.4	Colony sizes of tracked UTSCC-5 cells at different seeding concentrations	55
14.1	Colony sizes of SAT cells from different passages	58
14.2	Colony sizes of UTSCC-5 cells from different passages	58
14.3	Emergence of secondary colonies	59
15.1	Distributions of growth rates at different doses	62
16.1	Examples of growth behaviors discovered	63
16.2	Dose-dependent fractions of growth behaviors	64

16.3	Cell colony sizes grouped by the three distinct growth behaviors	64
16.4	Illustration of dose- and readout-dependent viability misclassification by fixed size thresholds	65
17.1	Survival fractions and fitted LQM curves for H460	67
17.2	Survival fractions and fitted LQM curves for RENCA	68
17.3	Survival fractions and fitted LQM curves for UTSCC-5	68
18.1	RBE of carbon ion irradiation as a function of reference dose for three cell lines	69
18.2	Threshold-based RBE values relative to growth-behavior based RBE values for all three cell lines	70
19.1	Example of quality control by examination of numbers of tracked objects	74
22.1	Histogram of Hoechst intensity in stained H460 cells	84

List of Tables

- 5.1 Cell lines used 23
- 8.1 Colony object properties extracted from segmentation results 34
- 8.2 Region property ranges to filter false positive segmentation objects 34
- 9.1 Predictor variables used for classification of any growth curve G 39
- 11.1 Segmentation results for 96well-based IncuCyte datasets 47

Acronyms

DKFZ	Deutsches Krebsforschungszentrum
FACS	fluorescence assisted cell sorting
HIT	Heidelberg Ion Beam Therapy Center
IVCA	in vitro clonogenic assay
LET	linear energy transfer
LQM	linear-quadratic model
PE	plating efficiency
RBE	relative biological effectiveness
SSE	sum of squared errors

Acknowledgements

First of all, I would like to thank Markus Alber and Emanuel Bahn, who both believed that I could tackle this challenging, yet (hence?) exciting project. If you did not believe that this work was worth pursuing and that I was able to handle the various challenges, I would have missed a lot of valuable lessons acquired in the last four years. You both gave me numerous ideas on how to write papers, how to give meaningful presentations, how to convince reviewers and audiences, how to formulate ideas, how to illustrate results, how to continue through the rough patches and much more. New approaches naturally prompt resistance within established structures and established ways of thinking. Thank you for backing me and helping me to trust my path when the wind was blowing from ahead.

Additionally, I want to thank Marc Boucsein, my PhD ally through thick and thin. Despite our differences in training, political views or preferred lunch locations, I feel we could always rely on each other when it came to navigating the organizational jungle of a UKHD/DKFZ cooperation unit, finding solutions to segmentation/registration/tracking/plotting problems or figuring out why this world and their inhabitants are as marvelously insane as they are.

I am grateful to Lina Renner and Viktoria Kiehl for their diligent work at the microscope during their respective times as HiWis. Thanks to you, I was able to spend more time and energy on experiment planning and data analysis. Likewise, I want to thank the Master students Max Kilian and Nicolas Lichti for their time and effort spent on their subprojects. You helped me to better understand the growth dynamics in our original data as well as the chances and limitations of ilastik as a segmentation tool.

A really big Dankeschön goes raus an Christoph den alten Aal. I think our constant work-life-balance-challenges during our mutual Bachelor study experience were a great preparation for the rest of our lives. What is studying hard without a rewarding Lasagne? What is Lasagne without some händische Fehlerfortpflanzungsrechnung to earn it? Without you I would be somewhere else.

In addition, I want to thank all my good friends out there, who make up the social net that is necessary to feel like this craze of a life has actual meaning. Cristian, Morgan, Vini, Mert, Francisco, Hannes, Tasso, Max, Maxi, Anton, Eric, Clemens, Oskar, Robin G. I love you!

Finally I want to thank my wife, Carla. Who leaves NY to be with a Dorfboy? You were and are always there when challenges become obstacles. When doubt and insecurity take over, you are a source of stability, a reminder of beauty, a kind soul. You help me see the big picture and listen to my heart. With all due respect to the hard facts and rational thinking: Without the guts, the heart, the soul, what are we even thinking for? I love you!

I learned a lot of fascinating facts, methods and concepts through the last four years. However, I think the biggest insights considered myself: Why do I do the things I do? What are my strengths? What are my weaknesses? What do I need to thrive? Equipped with (partial) answers to these questions, I am excited to start a new chapter and feel ready for new adventures.

Erklärung

Ich versichere, dass ich diese Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Heidelberg, den 6. September 2023,

Robin-André Koch