

Development of *in-vitro in-silico* technologies for modelling and analysis of haematological malignancies

A thesis submitted to Imperial College London for the degree of Doctor of Philosophy

By

José Carlos Ferreira Rodrigues de Morais

November 2021

Biological Systems Engineering Laboratory
Centre for Process Systems Engineering
Department of Chemical Engineering
Imperial College London
United Kingdom

Acknowledgements

When I started this journey roughly five years ago, finishing my studies and writing this thesis seemed hallmarks that were far too ahead in the future. I knew back then, as I know now, that at some point I would be staring at a blank page trying to write my Acknowledgements and I would struggle, as I am struggling now. Not because I am not utterly grateful, but because I knew that this would be a long, eventful journey, that would not have been completed with the invaluable help of so many bright individuals, to which I am, and will be, eternally grateful.

I would like to start by thanking my Professors at Técnico – Laura Ilharco, Joaquim Cabral and Jorge Leitão. They were, are, and will continue to be people that I look up to, not only because of their professional achievements, but most importantly because of their outstanding personal capabilities. With Laura, I learned what academic research truly means; with Joaquim, I was able to enhance my research capabilities, and with Jorge I had long, helpful and meaningful conversations. Thank you Laura, Joaquim and Jorge – this work is yours as well!

I would like to thank Professor Athanasios Mantalaris and Doctor Nicki Panoskaltsis for this outstanding opportunity. They taught me on how to work independently and showed me that data is the number one requisite to make any assumption, statement or conclusion. Above all, working under their supervision has inspired me to become better every day, and taught me that overcoming big obstacles requires taking even bigger challenges. Thank you Sakis and Nicki!

This work is a combined effort of uncountable hours of debate, discussion, laughter, banter, advice and complaints with my outstanding lab mates. Not only I have learned new techniques and skills from all of them – Mark, Asma, Susana, Joana, Michalis, António, Sophie, Spyro – but working alongside all of my lab mates was really what kept me going when the project seemed to be falling apart. I am eternally grateful for the continuous support, lessons, guidance and advice of the Greek-Portuguese alliance. Michalis, Sophie, Spyro, António, Joana – I am honored to have met you and I am sure that our paths will come across many times in the near or far future. Thank you very much!

A special note goes to the most outstanding, helpful and immensely kind people that I came across at Imperial and Surrey. Professor Nilay Shah, Doctor Karen Polizzi, Doctor Cleo Kontoravdi, Anusha Sri-Pathmanathan, Susi Underwood, Sarah Payne and the Velliou group – thank you ever so much not only for showing me that there is always light at the end of the tunnel, but also that actually it is much brighter and better on the other side. A very special mention to Doctor Karen Polizzi for all the long conversations, advice and support over the past couple of years, which I wholeheartedly appreciate and am grateful for. If I managed to get to this point, I truly owe it to you – thank you very much!

Finally, I would like to thank my beautiful, outstanding wife Cátia and our families for truly making this project their own as well. Cátia, as well as our moms, dads, sisters and grandparents were indefatigable on continuously providing support, help and, at times, just a helpful ear to listen to what was going wrong. In particular, Cátia was, is and will be (hopefully!) there every day with her beautiful smile, a quick, funny joke and arms wide open to discuss, help, support and nurture. Thank you... For everything!

Declarations

Declaration of Originality

The work presented in this thesis is original and produced by the author. Any other work has been referenced in an appropriate manner.

Copyright

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives license. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the license terms of this work.

Funding from the European Union Horizon 2020 SyMBioSys ITN (grant no. 675585) is greatly appreciated.

José C.F. Morais

List of Publications

Co-authored publications

1. Allenby MC, Tahlawi A, **Morais JCF**, Kang L et al, Ceramic hollow fibre constructs for continuous perfusion and cell harvest from 3D hematopoietic organoids. *Stem Cells Int* 2018; 6230214
2. Tahlawi A, Klontzas ME, Allenby MC, **Morais JCF** et al, RGD-functionalized polyurethane scaffolds promote umbilical cord blood mesenchymal stem cell expansion and osteogenic differentiation. *J Tissue Eng Regen Med* 2019; 13 (2), 232-243
3. Klontzas ME, Reakasame S, Silva R, **Morais JCF** et al, Oxidized alginate hydrogels with the GHK peptide enhance cord blood mesenchymal stem cell osteogenesis: A paradigm for metabolomics-based evaluation of biomaterial design. *Acta Biomaterialia* 2019; 88, 224-240

Abstract

Worldwide, haematological malignancies are responsible for roughly 6% of all the cancer-related deaths. Leukaemias are one of the most severe types of cancer, as only about 40% of the patients have an overall survival of 10 years or more. Myelodysplastic Syndrome (MDS), a pre-leukaemic condition, is a blood disorder characterized by the presence of dysplastic, irregular, immature cells, or blasts, in the peripheral blood (PB) and in the bone marrow (BM), as well as multi-lineage cytopenias.

We have created a detailed, lineage-specific, high-fidelity *in-silico* erythroid model that incorporates known biological stimuli (cytokines and hormones) and a competing diseased haematopoietic population, correctly capturing crucial biological checkpoints (EPO-dependent CFU-E differentiation) and replicating the *in-vivo* erythroid differentiation dynamics. In parallel, we have also proposed a long-term, cytokine-free 3D cell culture system for primary MDS cells, which was firstly optimized using easily-accessible healthy controls. This system enabled long-term (24-day) maintenance in culture with high (>75%) cell viability, promoting spontaneous expansion of erythroid phenotypes (CD71⁺/CD235a⁺) without the addition of any exogenous cytokines. Lastly, we have proposed a novel *in-vitro in-silico* framework using GC-MS metabolomics for the metabolic profiling of BM and PB plasma, aiming not only to discretize between haematological conditions but also to sub-classify MDS patients, potentially based on candidate biomarkers. Unsupervised multivariate statistical analysis showed clear intra- and inter-disease separation of samples of 5 distinct haematological malignancies, demonstrating the potential of this approach for disease characterization.

The work herein presented paves the way for the development of *in-vitro in-silico* technologies to better, characterize, diagnose, model and target haematological malignancies such as MDS and AML.

Contents

1. Introduction.....	14
1.1. Healthy Haematopoiesis	15
1.1.1. Erythropoiesis.....	18
1.2. Diseased Haematopoiesis	22
1.2.1. Myelodysplastic Syndrome	24
1.3. <i>In-vitro</i> models of haematological malignancies.....	39
1.3.1. <i>In-vitro</i> models of Myelodysplastic Syndrome	41
1.4. <i>In-silico</i> models of healthy and diseased haematopoiesis	45
2. Aims and Objectives	48
3. Materials and Methods.....	51
3.1. Mathematical modelling and simulations.....	51
3.2. Cell isolation	51
3.3. Porous polyurethane scaffold preparation and coating	52
3.4. Cell culture	53
3.5. Cell viability.....	53
3.6. Flow Cytometry.....	54
3.7. Confocal Microscopy	55
3.8. DNA quantification	56
3.9. GC-MS metabolomics.....	56
3.10. Metabolomic data processing.....	57
3.11. Statistical analyses.....	58
4. Development of a novel haematopoietic <i>in-silico</i> mathematical model	59
4.1. Introduction.....	60
4.1.1. Model Development.....	61
4.1.2. Modelling Healthy Erythropoiesis – A Case-Study	64
4.1.3. Modelling Diseased Erythropoiesis in MDS	73
4.2. Results and Discussion	75
4.3. Future Work.....	89
5. Long-term 3D cultures of primary MDS cells	93
5.1. Introduction.....	93
5.2. Results and Discussion	94
5.2.1. Healthy control cultures	95
5.2.2. Primary MDS cell cultures.....	111
5.3. Future work.....	123
6. Plasma Metabolomics.....	126
6.1. Introduction.....	127

6.2. Results and Discussion	129
6.2.1. Bone Marrow	129
6.2.2. Peripheral Blood	132
6.3. Future Work.....	136
7. Key Achievements and Future Developments	138
References	140
Appendix I – Model Assumptions and Biological Context.....	166
Appendix II – Proposed FC panel for the identification of healthy erythroid cells.....	178
Appendix III: Unique metabolites detected	179
Appendix IV: Code used	180

List of Tables

Table 1 – WHO classification of Myelodysplastic Syndrome. Except for MDS-u, disease severity and IPSS-R risk increases from top to bottom.	27
Table 2 – 2016-revised WHO classification of Myelodysplastic Syndrome. Except for MDS-u, disease severity and IPSS-R risk increases from top to bottom.	28
Table 3 – A summary of selected genes commonly mutated on MDS, their function and prognosis to MDS patients.	32
Table 4 – Studies using patient-derived BM-MNCs its cellular source, type and duration, cytokines used and study objectives. Legend: 2D LS - 2D Liquid Suspension; CR – Cytokine Responsiveness; CFS – Colony Formation Studies;	42
Table 5 – A summary of selected published mathematical models addressing both healthy and diseased haematopoiesis. Legend: CBA - Computational Biology Algorithm; DDE – Delayed Differential Equations; LSC – Leukaemic Stem Cell; ODE – Ordinary Differential Equations; ODE-S – Ordinary Differential Equations with Stochastic elements.	47
Table 6 – Nomenclature used for the mathematical description of a variety of cell processes (Figure 10) happening in different cell compartments, for a given generic cell type (A).	65
Table 7 – Nomenclature used for the mathematical description of the cellular fractions undergoing different cell processes (Figure 10) happening in different cell compartments, for a given generic cell type (A).	65
Table 8 - Rates of proliferation, differentiation and death according to each cell type. 70	
Table 9 – Maximum proliferation and differentiation fractions of different cell types.	71
Table 10 – Effects of SCF, IL3 and EPO on the proliferation and differentiation fractions of different cell types.	71
Table 11 - Rates of proliferation, differentiation and death according to each MDS cell type.	73
Table 12 - Maximum proliferation and differentiation fractions of different MDS cell types.	74
Table 13 – Summary of the experimental conditions used for the <i>in silico</i> simulations on the healthy erythroid population.	76
Table 14 - Summary of the experimental conditions used for the <i>in silico</i> simulations on the diseased erythroid population. Values in bold were kept constant.	81
Table 15 – Six different conditions tested using UCB-MNCs as healthy controls.	95
Table 16 - Proposed phenotypes for the haematopoietic stem and progenitor cells in the erythropoietic lineage.	178

List of Figures

Figure 1 – The bone marrow microenvironment and niche. Adapted from (Li L 2010).	16
Figure 2 – Overview of the classic definition of haematopoiesis. This diagram shows the different cell types that can arise from a multipotent HSC – on the top of this diagram. Both common myeloid and lymphoid progenitors (shown) are shared amongst different cell lineages, and each specific lineage has its own early and late progenitors (not shown). By A. Rad and M. Häggström. CC-BY-SA 3.0 license	17
Figure 3 – Erythropoiesis in the bone marrow. After enucleation, reticulocytes are released in the bloodstream where they further mature into an erythrocyte, or red blood cell. Adapted from (Immunopaedia 2017)	19
Figure 4 – The erythroblastic island. Adapted from (Chais JA 2008).	20
Figure 5 – Erythropoietic cell stages and selected cytokine receptor expression. Adapted from (Spivak 2005)	21
Figure 6 – Examples of cell dysplasia shown in the erythroid, megakaryocytic and granulocytic lineages. Ring sideroblasts are identified by using a Prussian Blue stain (Cazzola M 2011). Adapted from (Cazzola M 2013)	29
Figure 7 – Frequency of MDS somatic mutations. Data extracted and combined from Cazzola M 2013, Tothova Z 2013, Itzykson R 2013, Bejar R 2014, Haferlach T 2014, Harada H 2015 and Steensma 2015.	33
Figure 8 – A simplified schematic of the protective mechanism of Hsp70 on GATA-1, by preventing Caspase-3 dependant apoptosis. Adapted from (Weiss MJ 2009).	38
Figure 9 – Proposed framework for the in-vitro in-silico modelling of haematological malignancies.	48
Figure 10 – A schematic of cell death, differentiation and proliferation mechanisms.	62
Figure 11 – The proposed compartmentalization of haematopoiesis into the progenitor cell macro-compartment (in blue), the committed cell macro-compartment (in green) and the terminally differentiated cell macro-compartment (in red).	63
Figure 12 - Erythroid differentiation and the different identifiable intermediate stages. Legend: HSC – Haematopoietic Stem Cell; MPP – Multipotent Progenitor; CMP – Common Myeloid Progenitor; BFU-E – Erythroid burst-forming unit; CFU-E – Erythroid colony-forming unit; ProE – Proerythroblast; BasoE – Basophilic erythroblast; PolyE – Polychromatophilic erythroblast; OrthoE – Orthochromatophilic erythroblast; RBC – Red blood cell.	64
Figure 13 – An in-depth look at the inlet and outlet processes occurring at the BasoE cell population level.	67
Figure 14 – Effects of SCF, IL3 and EPO on cell proliferation (top) and differentiation (bottom). Positive, null and negative effects are represent by values of +1, 0 and -1 respectively.	72
Figure 15 - Effects of SCF and IL3 on the erythroid progeny, without any EPO added. The formation of BFU-E and CFU-E is independent of EPO, but further erythroid differentiation is impaired if EPO is not present.	77
Figure 16 - Effects of SCF and IL3 on the erythroid progeny with EPO added. IL3's inhibitory effect is strongly seen in the erythroid progeny (BasoE and RBC, bottom panel). The more mature a cell is, the stronger the inhibitory effect, typically as a consequence of differentiation arrest of earlier, immature cell types.	78
Figure 17 – Effects of SCF and EPO addition without (top panel) and with (bottom panel) IL3. The addition of IL3, provided under inhibitory concentrations, increases the numbers of both Committed Cells and RBC numbers.	79

Figure 18 – Adding SCF to a cocktail of EPO and IL3 delays IL3's inhibitory effect while considerably augmenting RBC numbers. The addition of no SCF (top-left), 10 ng/mL (top-right) and 50 ng/mL (bottom) of SCF yields a maximum RBC number of 5, 17 and 95 million cells, respectively.	80
Figure 19 – The presence of a competing dysplastic population shows a direct correlation between cytokine levels and inhibitory effects.	82
Figure 20 – Despite yielding a reduction of more than 2-fold of the progenitor cell numbers, the presence of MDS cells only seem to increase RBC sensitivity to higher concentrations of cytokines. However, these effects need further studies since the dependence on parameters like MDS(Ka) are very high.	83
Figure 21 – First-order sensitivity indices for HSCs, as a function of time.	86
Figure 22 – First-order sensitivity indices for BFU-E cells, as a function of time.	87
Figure 23 – First-order sensitivity indices for CFU-E cells, as a function of time.	87
Figure 24 – First-order sensitivity indices for RBCs, as a function of time.	88
Figure 25 – Proposed framework for Essential Future Work. Currently, this work has reached the Model Analysis stage, but it needs further refinement with better initial guesses for the parameters, as aforementioned.	90
Figure 26 – Supernatant cell viability was assessed for cells cultured at 0.5 million (A), 1 million (B) and 2 million (C) cells per scaffold. Across all conditions, no statistically significant differences were seen at the supernatant cell viability level. Slightly greater variability is seen when using collagen-containing, but within a negligible level. Results are shown as average $\pm$ SD of N=4, n=3.	96
Figure 27 – Fluorescence detected by dsDNA-PicoGreen cross-linking in-situ within scaffolds across all groups and time points, normalized per group at its Day 1 fluorescence levels. Results are shown as average $\pm$ SD of N=4, n=3.	97
Figure 28 – Sections of the scaffold core imaged using confocal microscopy. Green fluorescence is present on live cells due to Calcein AM Green, a viability dye. Results are shown as representative images of N=4, n=1. Scale bar shown represents 500 μ m. Aggregated data from all images is shown in Figure 30.	99
Figure 29 – Sections of the scaffold core imaged using confocal microscopy. Red fluorescence is present on dead cells due to the binding of Ethidium Homodimer-1 to DNA. Results are shown as representative images of N=4, n=1. Scale bar shown represents 500 μ m. Aggregated data from all images is shown in Figure 30.	100
Figure 30 – In-situ viability determined by Confocal Microscopy at the core of the scaffold. Viability was calculated as the ratio between viable cells (Calcein AM Green-positive) and total cells (viable and dead, verified by Ethidium homodimer-1-positive). Results are shown as average $\pm$ SD of N=4, n=1.	101
Figure 31 – Phenotypical changes of CD45 ⁺ CD36 ⁻ (leukocyte) supernatant cells, as a percentage of total viable cells, throughout a 24-day culture. There is a decrease in this phenotype from the initial time points (Day 0 – seeding – and Day 8) to the later time points. This decrease is inversely proportional to the initial seeding density. Results are shown as average $\pm$ SD of N=3. * P < 0.05; ** P < 0.01; *** P < 0.001; **** P < 0.0001.	102
Figure 32 – Phenotypical changes of CD45 ⁺ CD36 ⁻ (leukocyte) supernatant cells, as a percentage of total viable cells, throughout a 24-day culture. There is a decrease in this phenotype from the initial time points (Day 0 – seeding – and Day 8) to the later time points. This decrease is inversely proportional to the initial seeding density. Results are shown as average $\pm$ SD of N=3. * P < 0.05; ** P < 0.01; *** P < 0.001; **** P < 0.0001.	103

- Figure 34** – Phenotypical evolution, throughout the culture, of supernatant early erythroid cell populations (CD235a⁻ CD71⁺) expressed as a percentage of total viable cells. Results are shown as average $\pm$ SD of N=3. * P < 0.05; ** P < 0.01; *** P < 0.001..... 104
- Figure 34** – Phenotypical evolution, throughout the culture, of supernatant late erythroid cell populations (CD235a⁺ CD71⁻) expressed as a percentage of total viable cells. Results are shown as average $\pm$ SD of N=3. < 0.001 < p < 0.05 shows that there is a statistically significant decrease in this phenotype, across all conditions, from Day 0 to Day 8. No other statistically significant (ns) difference was found. 105
- Figure 35** – Phenotypical changes of CD235a⁺ CD71⁺ (mid erythroid progenitor) supernatant cells, as a percentage of total viable cells, throughout a 24-day culture. Collagen-coated scaffolds were seeded with 0.5 million (A), 1 million (B) and 2 million (C) cells respectively. Results are shown as average $\pm$ SD of N=3. * P < 0.05; ** P < 0.01; *** P < 0.001..... 106
- Figure 36** – Phenotypical changes of CD235a⁺ CD71⁺ (mid erythroid progenitor) supernatant cells, as a percentage of total viable cells, throughout a 24-day culture. Uncoated scaffolds were seeded with 0.5 million (A), 1 million (B) and 2 million (C) cells respectively. Results are shown as average $\pm$ SD of N=3. * P < 0.05; ** P < 0.01;..... 107
- Figure 37** – Representative flow cytometry plots of CD45⁺ CD36⁻ and erythroid progenitor supernatant cells, seeded at 1 million MNCs per uncoated scaffold. After excluding doublets, gating (Day 8 to Day 24) was performed on live cells (Calcein AM positive). Negative-positive phenotypical boundaries were done using isotype controls. 109
- Figure 38** – Viability of supernatant cells was measured across four independent primary MDS cultures and compared to that of UCB-MNCs (healthy control). Results are shown as average $\pm$ SD of N=3 (healthy control) and N=3 (primary MDS cells)..... 111
- Figure 39** – Fluorescence detected by dsDNA-PicoGreen cross-linking in-situ within scaffolds and normalised, per condition, to its Day 1 levels. Results are shown as average $\pm$ SD of N=3 (healthy control) and N=2-3 (primary MDS cells)..... 112
- Figure 40** – Sections of the core of the scaffold imaged by confocal microscopy. Green fluorescence is present on live cells due to Calcein AM Green, a viability dye. Red fluorescence is due to the DNA binding of the cell impermeable Ethidium Homodimer-1 stain. Qualitatively, scaffolds seeded with 1 million cells seem to have similar viable cell numbers at late culture time-points (left panel) while the MDS condition exhibits less dead cells (right panel). Results are shown as representative images of N=4, n=1 (Healthy UCB-MNC Control) and N=1, n=1 (Primary MDS sample). Scale bar shown represents 500 μ m 114
- Figure 41** – Isolated cells from primary MDS patients (top panel) and healthy UCB-MNCs (bottom panel) were stained with a FITC-labelled anti-human CD45 antibody immediately after isolation and prior to seeding in scaffolds. There are marked differences between independent MDS samples (top panel) both in population percentages and morphology (SSC-A levels), as opposed to the overall consistency of healthy UCB-MNC cell isolates (bottom panel). Variable blast percentages, as well as potential features of dysgranulopoiesis (dysplasia in the granulopoietic lineage) can be seen in MDS samples (top panel)..... 115
- Figure 42** – Bivariate CD45/CD36 plots of UCB-MNC (healthy control) and BM-MNCs (MDS). Legend: **UPN** – Unique Patient Number. 118
- Figure 43** – Bivariate CD235a/CD71 plots of UCB-MNC (healthy control) and BM-MNCs (MDS). Legend: **UPN** – Unique Patient Number. 119

Figure 44 – Phenotypical evolution, throughout the culture, of pan-myeloid (A) as well as early (B), mid (C) and late (D) erythroid progenitors, expressed as a percentage of total viable cells. Results are shown as average $\pm$ SD of N=3 (healthy control) and N=2 (MDS). 120

Figure 45 – 2D PCA analysis of BM samples. Every sample is obtained in four technical replicates. The colours indicate different characteristics/haematological conditions: **Red** – MDS; **Brown** – MDS pre-treated with folic acid; **Orange** – CMML-1; **Green** – ET; **Dark Blue** – CML at blast phase; **Navy Blue** – CML at restaging; **Purple** – Clonal CD8⁺ T Cell expansion. Meaningful groups discussed in the text below are circled and numerically identified in black. 129

Figure 46 – 2D PCA analysis of PB samples. Every sample is obtained in four technical replicates. The colours indicate different characteristics/haematological conditions: **Red** – MDS; **Orange** – CMML-1; **Green** – ET; **Dark Blue** – CML at blast phase. Meaningful groups discussed in the text below are circled and numerically identified in black. 132

List of Commonly Used Abbreviations

2D	Two-Dimensional
3D	Three-Dimensional
AML	Acute Myeloid Leukaemia
BFU-E	Burst-Forming Unit- Erythroid
BM	Bone Marrow
CFC	Colony-Forming Cell
CFU-E	Colony-Forming Unit- Erythroid
CFU-GEMM	Colony-Forming Unit- Granulocyte Erythroid Monocyte Macrophage
CFU-GM	Colony-Forming Unit - Granulocyte Monocyte
CML	Chronic Myeloid Leukaemia
CMML	Chronic Myelomonocytic Leukaemia
dsDNA	Double-stranded Deoxyribonucleic Acid
ECM	Extracellular Matrix
ET	Essential Trombocythemia
EPO	Erythropoietin
EPOR	Erythropoietin Receptor
FAB	French-American-British
FBS	Foetal Bovine Serum
G-CSF	Granulocyte Colony-Stimulating Factor
GM-CSF	Granulocyte Macrophage Colony-Stimulating Factor
HSC	Haematopoietic Stem Cell
HSPC	Haematopoietic Stem and Progenitor Cells
IFNγ	Interferon - Gamma
IL	Interleukin
IL3	Interleukin-3
IL3R	Interleukin-3 Receptor
IPSC	Induced Pluripotent Stem Cell
IPSS	International Prognostic Scoring System
MDS	Myelodysplastic Syndrome
MNC	Mononuclear Cells
MPN	Myeloproliferative Neoplasm
MSC	Mesenchymal Stem Cell
PB	Peripheral Blood
PBS	Phosphate-buffered Saline
RBC	Red Blood Cell
RNA	Ribonucleic Acid
SCF	Stem Cell Factor
TNFα	Tumor Necrosis Factor - Alpha
WBC	White Blood Cell
WHO	World Health Organization

1. Introduction

The bone marrow (BM) is one of the largest and more complex tissues in our organism, accounting for around 4% of the total mass of a healthy male adult – for a typical man weighting 80kg, it is expected that his bone marrow weights around 3.2kg (Rubin 2007). It is often called the “flexible inner tissue of bones” and is responsible for generating large quantities of blood cells in response to different stimuli. Within the bone marrow, cells divide, differentiate and migrate driven by chemical concentration gradients, cell-cell interactions and oxygen tension, undergoing a process commonly called haematopoiesis. This process occurs continuously within the bone marrow’s haematopoietic niche (Velten L 2017) and is mainly supported by Haematopoietic Stem Cells (HSC) – the starting stage for cellular growth and differentiation – and Mesenchymal Stem Cells (MSC), enabling cell-cell interactions and providing stromal support via a complex combination of chemical signals (Kaushansky K 2011).

These chemical signals are produced by BM-MSCs, fibroblasts and other stromal cells and are commonly known as cytokines, playing a key role in cell differentiation, tissue inflammation and pain. Cytokines range from chemical molecules to small proteins, enabling cell-cell communications, activating specific signalling pathways and pushing HSCs towards differentiation and/or proliferation (Zhang JM 2007). In homeostatic levels, cytokines are typically present in the pg/mL order of magnitude but can spike by up to 100-fold in response to a major infection by pathogenic agents or extreme tissue inflammation, known as a “cytokine storm”. This phenomena has been observed as an undesired side-effect of a clinical trial (Suntharalingam G 2006 and Yiu HH 2012) with both short-term and long-term effects. More recently, several studies report that patients infected with SARS-CoV-2 also experience cytokine storms, showing high concentrations of pro-inflammatory cytokines such as Interleukin-1, Interleukin-6 and Interferon-Gamma (IFN γ). These have been directly linked with COVID-19 disease

severity, poor prognosis, high mortality and even as a potential therapeutical target (Henderson LA 2020, Tang Y 2020, Hojyo S 2020 and Cron RQ 2021).

1.1. Healthy Haematopoiesis

The bone marrow and its haematopoietic niche have been extensively characterized using both human-derived xenograft as well as *in-vivo* mice models (Chen Y 2012, Morrison SJ 2014, Hamon P 2015, Boulais PE 2015, Calvi LM 2015 and Abarrategi A 2018). A human bone marrow model was proposed in 2010 by Li and Clevers (Figure 1) comprising the existence of an osteoblastic niche, under hypoxic conditions and further away from vascularization, and a vascular niche, around the bone marrow's vasculature (Li L 2010). In the osteoblastic niche, hypoxia-exposed cells are further away from direct nutrient supply whereas in the vascular niche cells have plenty of nutrients, as well as normoxic oxygen concentration levels.

Oxygen tension, nutrient availability and the surrounding biochemical, cellular and spatial microenvironment are known to affect cell cycle, quiescence, expansion and differentiation (Jing D 2012 and Gomariz A 2018). It is commonly accepted that quiescent, long-term HSCs (LT-HSCs) are mainly located within the osteoblastic niche. In order to maintain a constant supply of mature blood cells *in-vivo*, some of these LT-HSCs will divide asymmetrically giving rise to short-term HSCs (ST-HSCs) and rapidly-dividing haematopoietic progenitors. As HSCs start actively cycling and differentiating into haematopoietic progenitor cells, they migrate closer to the vasculature. When reaching terminal maturation stages, the haematopoietic progenitor cells will egress to the blood stream.

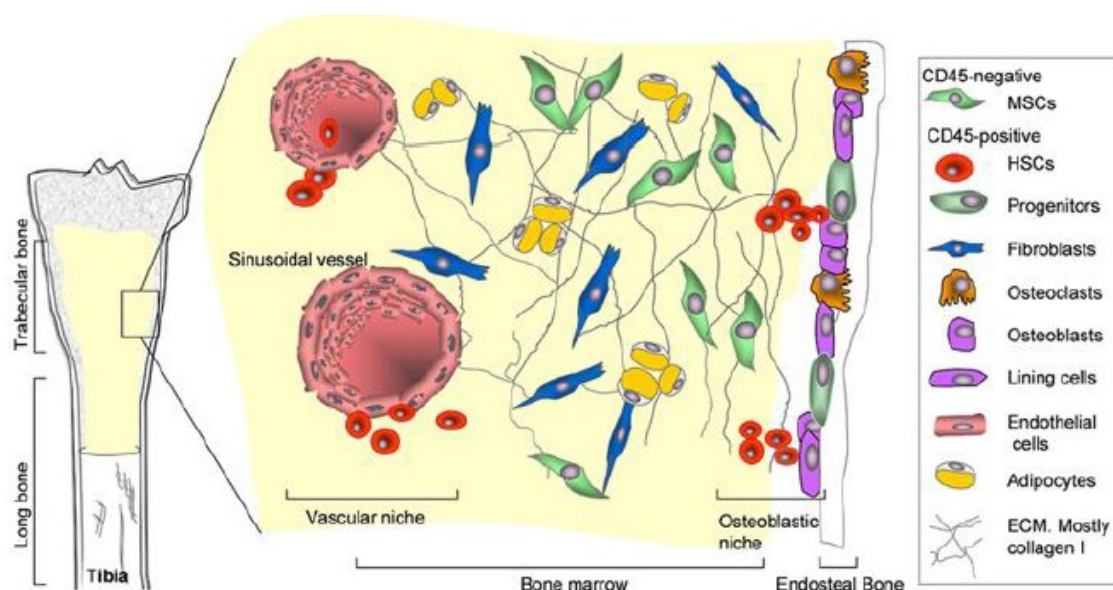


Figure 1 – The bone marrow microenvironment and niche. Adapted from (Li L 2010).

During early childhood, the vast majority of the bone marrow shows a red colour. This is believed to be related with a very high percentage of actively-cycling HSCs – and not arrested in G0 phase like most of the adult HSCs (Bowie MB 2006). With aging, the bone marrow slowly becomes yellow – rich in adipocytes, osteoblasts and fibroblasts. By the age of 18, the majority of the red bone marrow has been replaced by yellow bone marrow and the active haematopoietic bone marrow sites are located in the skull, vertebrae and ribs, pelvis and at the ends of both femur and humerus which are also called proximal epiphyses (Custer RP 1932 and Kaushansky K 2011). With aging, the prevalence of red bone marrow is mainly confined to the pelvic region, vertebrae, sternum and ribs.

Blood cell generation, or haematopoiesis, requires the existence of a heterogeneous pool of HSC that is continuously differentiating into the myeloid or lymphoid lineages while self-renewing in order to keep its total cell number relatively constant. These two common progenitors – myeloid and lymphoid – then further differentiate to give rise to a wide range of cell types, summarized in Figure 2.

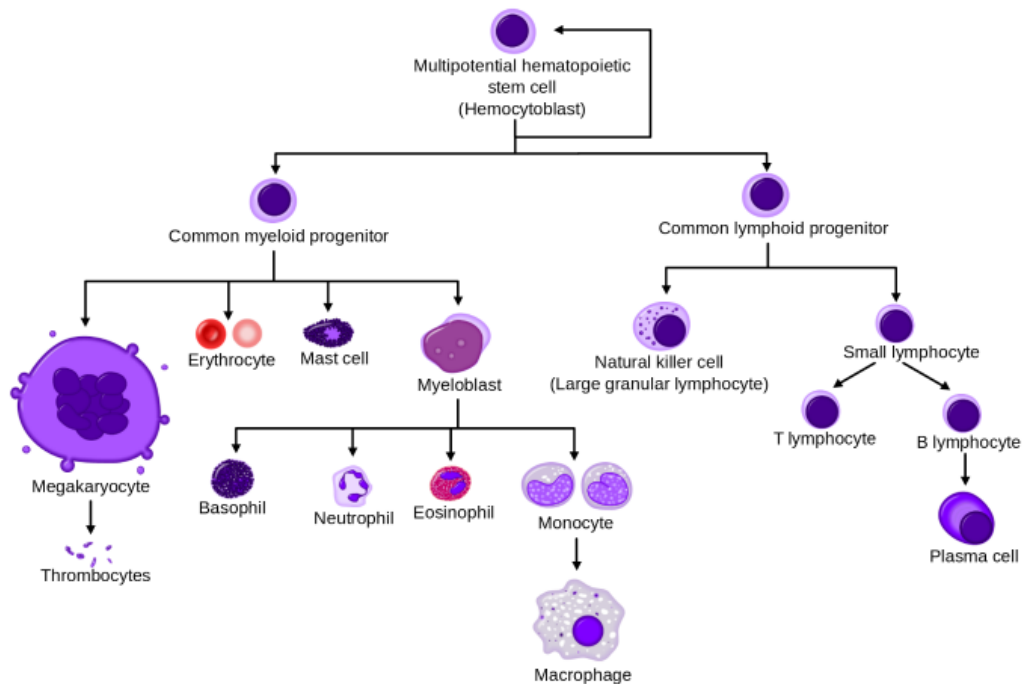


Figure 2 – Overview of the classic definition of haematopoiesis. This diagram shows the different cell types that can arise from a multipotent HSC – on the top of this diagram. Both common myeloid and lymphoid progenitors (shown) are shared amongst different cell lineages, and each specific lineage has its own early and late progenitors (not shown). By A. Rad and M. Häggström. CC-BY-SA 3.0 license

Haematopoiesis is a multi-level process with genetic, endocrine, paracrine and molecular players in an intricate regulatory network. The evolutionary fate of a progenitor HSC is heavily influenced by a myriad of macro- and micro-environmental factors, resulting in cell-lineage commitment, differential gene expression, morphological and physiological changes and cell apoptosis, differentiation and proliferation. Several studies addressed the relevance of bone marrow stromal cells (Anthony BA 2014), physical micro- and macro- environment (Zhang P 2019), genes and transcription factors (Nerlov C 2000 and Takahashi T 2000), signalling pathways (Varnum-Finney B 1998 and Nemeth MJ 2007), cytokines (Zhu J 2002, Buza-Vidas N 2006, Robb 2007 and Zhang CC 2008) mature cell numbers (Graaf CA 2010) and cell cycle (Pietras EM 2011) in haematopoietic regulation.

The denomination given to each blood cell lineage, or haematopoietic stem cell progeny, is closely related to its final cell type, or terminally differentiated cell. The two main

progenitor lineages are the myeloid lineage and the lymphoid lineage. Within the myeloid lineage, we can consider three main lineages, such as the erythroid, granulocytic and megakaryocytic lineages. The erythroid lineage generates erythrocytes, commonly known as Red Blood Cells (RBC), which are primarily responsible for the transport of oxygen and carbon dioxide *in vivo*. The granulocytic lineage generates granulocytes – a type of White Blood Cells (WBC) – like basophils, neutrophils, eosinophils and monocytes, with the latter migrating into tissues and aiding with phagocytosis. The megakaryocytic lineage gives rise to megakaryocytes, which are then further processed to generate platelets. The lymphoid lineage is primarily responsible for adaptive immunity and generates B and T lymphocytes, as well as Natural Killer Cells (Kaushansky K 2011).

Other regulatory factors – often called cytokines – are produced in the bone marrow microenvironment by stromal cells and fibroblasts (Zhao LR 2013) and play a key role in ensuring homeostasis. The Granulocyte Colony-Stimulating Factor (G-CSF) activates granulocyte production processes such as neutrophil production and terminal differentiation (Roberts 2005) whereas Stem Cell Factor (SCF) enhances the proliferation of Haematopoietic Stem and Progenitor Cells (HSPC) (Zhang CC 2008).

1.1.1. Erythropoiesis

Erythropoiesis is the process by which red blood cells are generated. This is a multi-step process starting at the HSC level and involving iron uptake and haemoglobin synthesis. Throughout erythropoiesis, the cellular intermediates undergo a dramatic decrease in cellular corpuscular volume and, consequently, cell size.

Erythropoiesis starts with two consecutive cellular intermediates called Burst-Forming Unit- Erythroid (BFU-E) and Colony-Forming Unit- Erythroid (CFU-E). Residing as rare cell populations in the bone marrow, BFU-E have a high proliferative capacity but their populations are largely latent, whereas CFU-E have limited proliferation capability but a higher percentage of cells actively undergoing cell cycle. Both *in-vitro* and *in-vivo*, BFU-

E generates large aggregates (bursts) of haemoglobinized erythroblasts (Tsiftoglou AS 2009) while CFU-E gets its designation from forming smaller, more organised colonies with the same types of cells but in lesser numbers – typically from 16 to 64 cells (Kaushansky K 2011). These two cell populations are the first identifiable cellular intermediates to be committed to the erythroid lineage. At the early, proerythroblast stage, cell volume is around 900 fL, decreasing sharply to around 150 fL at the late, normoblast stage (Kaushansky K 2011). Just after that, normoblasts enucleate in the bone marrow and the newly formed reticulocytes are released into the blood stream where they mature into erythrocytes, or red blood cells (Figure 3).

Erythropoiesis is mainly regulated by the production of Erythropoietin (EPO) in the kidneys. This glycoprotein promotes RBC generation (Koury MJ 1991) and its blood concentration is, in healthy conditions, inversely proportional to oxygen levels in the Peripheral Blood (PB). By activating a series of GATA-mediated signalling pathways, EPO interrupts the homeostatic mechanism of programmed cell death at the erythroid progenitor level (CFU-E) and promotes terminal erythroid differentiation within the bone marrow and further maturation in the peripheral blood (W. Jelkmann 2011). Without EPO, terminal maturation is hampered, resulting in premature death of erythroid progenitors.

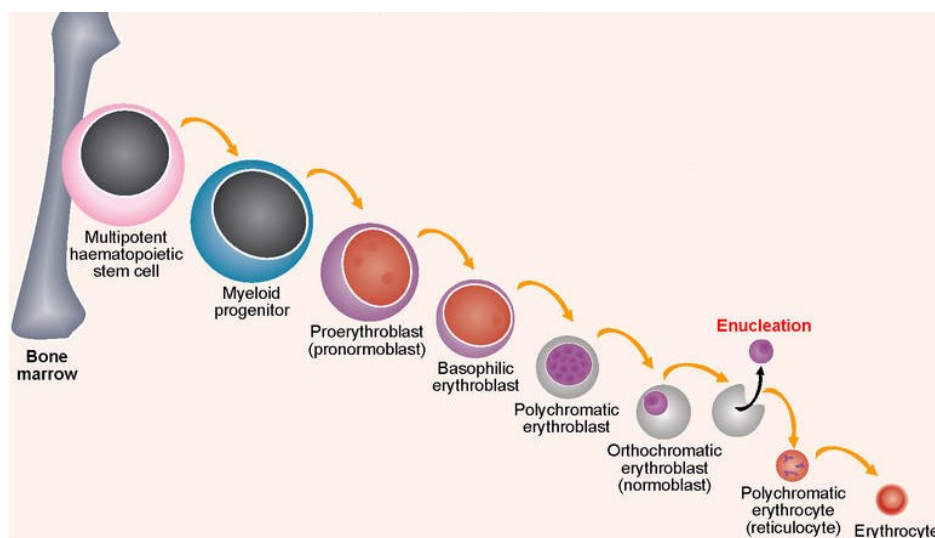


Figure 3 – Erythropoiesis in the bone marrow. After enucleation, reticulocytes are released in the bloodstream where they further mature into an erythrocyte, or red blood cell. Adapted from (Immunopaedia 2017)

In-vivo erythropoiesis is known to occur on recently characterized cellular complexes inside the red bone marrow called erythroblastic islands (Chais JA 2008) as shown in Figure 4. Formed by a central macrophage and erythroid progenitors at different stages of differentiation, this structure allows for maturation and development of erythroblasts into reticulocytes that are then released on the blood stream. Erythroid progenitors adhere to the macrophage at the CFU-E stage, and this central macrophage plays a pivotal role in supporting erythropoiesis via phagocytosis of extruded nuclei from mature erythroblasts. Murine co-cultures of macrophages with erythroblasts revealed a 3-fold increase in the numbers of erythroblasts when compared to erythroblast culture alone (Rhodes MM 2008), which further confirms the supporting functions of macrophages in erythropoiesis. Only very recently some groups were able to replicate and characterise human erythroblastic islands *in-vitro*, using Induced Pluripotent Stem Cells (IPSC) as well as healthy and diseased bone marrow aspirates (Lopez-Yrigoyen M 2019 and Elvarsdóttir EM 2020).

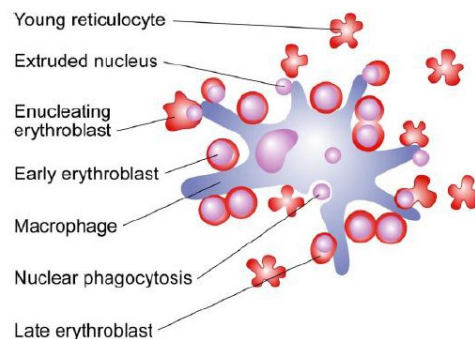


Figure 4 – The erythroblastic island. Adapted from (Chais JA 2008).

EPO is indispensable for achieving terminal differentiation and RBC generation. Other cytokines like SCF or Interleukin-3 (IL3) are known to also have an influence on the erythroid lineage by acting at immature cell populations. Analysing the expression of several cytokine receptors located on the cell membrane like c-Kit or EPOR can elucidate the effect of cytokines on the different stages of erythropoiesis, especially in the early

stages like BFU-E and CFU-E. While c-Kit recognizes SCF, EPOR is the membrane receptor for EPO and IL3-R recognizes IL3.

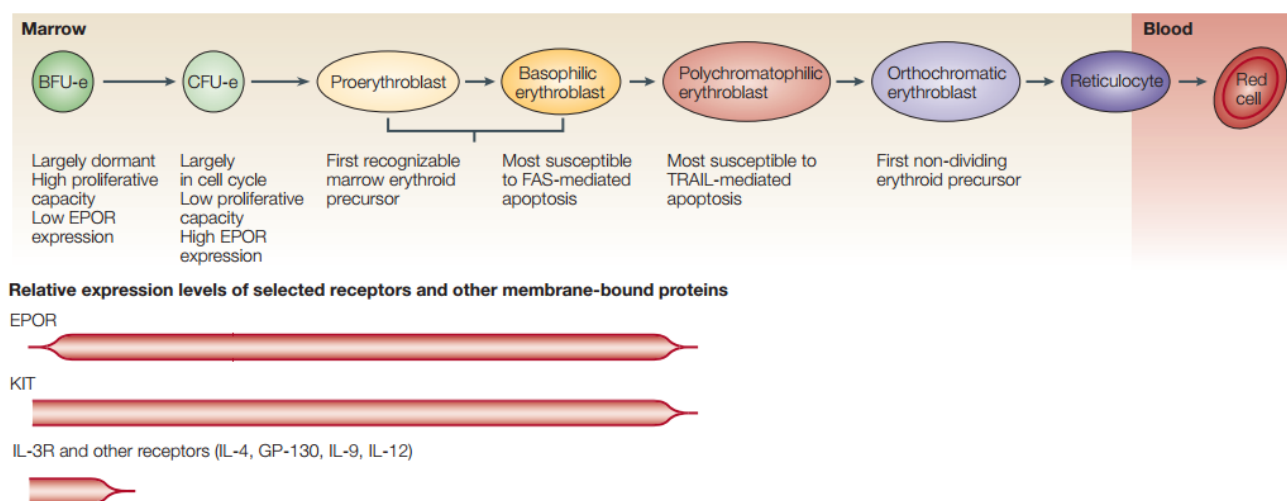


Figure 5 – Erythropoietic cell stages and selected cytokine receptor expression. Adapted from (Spivak 2005)

While both BFU-E and CFU-E have a high expression of c-Kit – associated with cell proliferation and survival (Keller JR 1995) – only CFU-E expresses considerable amounts of EPOR receptor – BFU-E also expresses EPOR, but at a concentration several hundred-fold lower. This higher EPOR expression means that CFU-E cells are more committed towards the erythroid lineage, when compared to the BFU-E one. A CFU-E is thought to arise from a single BFU-E daughter cell. On *in-vitro* 14-day Colony Forming Cell (CFC) assays, BFU-E cells replicate actively throughout the first seven days, forming the central colony. Then, some daughter cells start generating other types of colonies (called CFU-E) that reach full haemoglobinization over the remaining 7 days, where the original BFU-E reaches a similar state at the last day of the CFC assay. This correlates with the expression of EPO and c-Kit in the sense that BFU-E colonies have more cells than CFU-E colonies (higher BFU-E proliferative capacity) but CFU-E can reach the same maturation stage as BFU-E in half the time (7 days for full haemoglobinization on CFU-E versus 14 days for BFU-E). Further differentiation is dependent on correct iron uptake and haemoglobinization by proerythroblasts and their progeny (Nathan DG 1978).

1.2. Diseased Haematopoiesis

A wide variety of causes can lead disruption of haematopoiesis, affecting not only the bone marrow, but also the peripheral blood and the lymphatic system. Previous chemotherapeutic treatment, genetic disorders, exposure to carcinogenic chemicals, smoking, alcohol, obesity and genetic background are, among others, the common putative causes of any haematopoietic malignancy. The build-up of several disruptions can ultimately give rise to either a highly proliferative condition like Acute Myeloid Leukaemia (Jan M 2012), Myeloproliferative Neoplasms (Lundberg P 2014) and Polycythaemia Vera (Wolf BC 1988), or to slower malignancies such as Chronic Myeloid Leukaemia (Petzer AL 2003) or Myelodysplastic Syndrome (Li J 2013). These effects are not restricted to the myeloid lineage – in fact, the most common haematologic cancer in the US is non-Hodgkin's lymphoma (Horner MJ 2009), which primarily affects lymphoid tissues.

In many of these haematological malignancies, the bone marrow microenvironment is greatly affected by over-production of pro-inflammatory cytokines (Gañán-Gómez I 2015) with the resulting suppressing effects in the levels of supportive cytokines, insufficient levels of lineage-specific cytokines or an auto-immune response targeting specific cellular populations (Faquin WC 1992). Patients with aplastic anaemia, also called non-renal anaemia, display hypocellular bone marrows, likely to result from an auto-immune response targeting BM-HSCs. As a result, these patients display abnormally high EPO serum concentrations as the body attempts to counteract the peripheral blood anaemia, but remains unable to produce sufficient RBCs (Marsden 2006 and Medinger M 2018). Stroma and ECM remodelling (Tripodo C 2011) as well as the over-expression of certain surface glycoproteins (Jin L 2006 and Kumar R 2018) can also disrupt haematopoiesis.

Multi-factorial heterogeneity is ubiquitously present in all haematological malignancies, presenting a major challenge for patient-specific management, novel therapeutic

development, fundamental studies and clinical practice. Aberrant gene expression and complex karyotypes, concomitant with peripheral blood pancytopenias both in Myelodysplastic Syndrome (MDS) and Acute Myeloid Leukemia (AML), are the most studied features of these disease. Patients with haematological malignancies have one or more genetic mutations at the HSPC level, which then generate cellular clones (clonality), exhibiting combinations of these mutations. Due to competitive advantages and natural selection, malignant clones eventually take over the entire haematopoietic system. Clonality is indeed the main cause of therapy resistance, disease evolution and consecutive cycles of diagnosis-treatment-remission-relapse (Tuval A 2019). The accurate and correct characterization of a patient's clonality plays a major role in guiding clinicians on how best to tackle and manage leukaemia, both in children (Bertuccio SN 2021) and in adults (Yu J 2020 and Schmalbrock LK 2021).

Using cutting-edge techniques such as, for example, single-cell DNA sequencing, groups are able to identify unique clonal populations in AML patients and attempt to fully characterise patient trajectories, from disease onset to final outcome (Vosberg S 2019, Morita K 2020 and Romer-Seibert J 2021). Other groups have derived patient-specific iPSCs and replicated each patient's mutational profile *in-vitro* resorting to CRISPR/Cas9 techniques (Kotini AG 2017 and Deslauriers AG 2021).

There is an unmet need for patient-and-disease representative platforms for many of these diseases including MDS and AML. Particularly, the development of patient-specific *in-vitro* platforms is of great relevance, as this would allow researchers to test several hypotheses and better understand the underlying biological and molecular mechanisms of the disease, resistance to treatment and the influence of clonality on patient management and clinical outcomes. This will truly enable the concept of patient management from *bench-to-bedside*, where real-time *in-vitro* data is incorporated in the daily management of patients with haematological malignancies.

1.2.1. Myelodysplastic Syndrome

MDS can be defined as a clonal hematological malignancy comprised of several different subgroups of increased severity, with the ultimate progressive potential towards AML. A typical patient journey on the diagnosis of MDS starts by reports of nausea, dizziness, fatigue and loss of appetite. Clinicians can initially recommend preliminary full blood count tests, but any disease diagnosis is only definitive after a positive bone marrow aspiration and biopsy. The most common features of this so-called pre-leukemic state are cellular dysplasia – abnormal cellular shape and morphology, which hinders cell functionality – across one or more myeloid lineages, a paradoxal normo/hypercellular bone marrow composed up to 20% by blasts (immature cellular progenitors) and peripheral blood pancytopenias, or insufficient numbers of terminally-differentiated blood cells. Blasts are widely used as high-level biomarkers for an haematological malignancy, as they are typically not present in the peripheral blood, while making up to 5% of all cells in healthy bone marrows (Cancer.org 2018).

Within continental Europe, MDS has an incidence of roughly 3 per 100,000 (Lubeck DP 2016). In the United Kingdom, the incidence of MDS is estimated around 4 per 100,000 people, increasing by 7.5-fold for patients with 70 years of age and above (NICE 2012). 90% of the patients diagnosed with MDS are at least 60 years old. Between 2004 and 2016, there has been a steady increase in the yearly number of newly diagnosed cases, rising 8.53% from 1,993 (2004) to 2,163 in 2016 (ONS 2016).

Historically, MDS was sub-classified using the French-American-British (FAB) ranking system which was proposed in 1982 and was based mainly on morphologic criteria and blast (immature cell) numbers found in the bone marrow and peripheral blood (Bennett 2005). This classification system was then updated in the 2008 by the World Health Organization (WHO) (Vardiman JW 2009) by incorporating recent findings, with regards

to genetic mutations, karyotypes and disease progression. The 2008 MDS classification system proposed a novel MDS subtype – MDS del(5q) – which accounted for patients that exhibited a deletion in the long arm of chromosome 5. Particularly, patients with MDS del(5q) have later exhibited significant clinical improvements when treated with immunomodulatory drugs like lenalidomide, a thalidomide derivative (Komrokji RS 2016) (Krönke J 2015). Other major changes implemented by the 2008 WHO classification have included lowering the minimum blast percentage for AML diagnosis from 30% to 20%. Notably, patients exhibiting complex karyotypes such as translocations (t(15;17), t(8;21) or t(16;16)) and inversions (inv(16)) would now automatically be classified as having AML (Döhner H 2010).

MDS has been extensively studied over the past decade, which led to a recent revision to its WHO classification in 2016 (Arber DA 2016). The presence of a very specific gene mutation in the RNA spliceosome (SF3B1) resulted in a novel MDS sub-group, highlighting the presence of Ring Sideroblasts (Figure 7). The outdated “Refractory Anemia” nomenclature has been completely replaced by “MDS” – even though most of the patients exhibit anemia, this change clearly highlights the characterization of MDS as an haematological malignancy rather than a persistent, refractory pathology. The main features of the 2008 WHO classification and the 2016 revised version are summarized in Table 1 and 2, respectively. Several research and clinical groups worldwide still use, often interchangeably, the 2008 and 2016 WHO MDS classifications.

Originally published in 1997, the International Prognostic Scoring System (IPSS) quickly became one of the most used ranking systems for classifying MDS patients (Greenberg P 1997). Using the 2008 WHO classification, the IPSS was revised in 2012 (IPSS-R) by incorporating data collected from 7012 MDS patients and adding key features such as more cytogenetic subgroups, refined lower intervals for bone marrow blast percentages and depth of cytopenias, as well as patient age, serum ferritin and lactate dehydrogenase

levels – the latter three being used as prognostic features for survival, but not for AML progression (Greenberg 2012).

Clinical state-of-the-art diagnosis tools include a complete blood count, peripheral blood smears, bone marrow aspirations, cellular karyotyping and mutational analyses. Bone marrow biopsies are the gold standard for establishing accurate disease diagnosis in many haematological conditions. This procedure is typically done in the pelvic region and causes mild to severe discomfort and pain. For disease monitoring and re-staging, a patient is often required to undergo several aspirations. One of the biggest challenges in the field is to propose a series of robust, disease-identifying biomarkers and alternative cellular sources (e.g. peripheral blood smears) that could potentially reduce the need for confirmatory bone marrow biopsies, dramatically increasing patient's comfort and well-being.

Cell dysplasia is a common feature across all MDS sub-types. Dysplastic cells show aberrant morphologies and immature cellular development, with examples all of the myeloid lineages (Figure 6). Dysplastic cells are often not fully functional, which greatly affects homeostasis and normal blood cell production.

Scientific progress around MDS has focused on many different aspects like chromosomal aberrations (Andersen MK 1998 and Hasse 2008), epigenetic dysregulation (Issa 2010, Z. L.-N. Will B 2012, Santini V 2013 and Issa 2013), genetic mutations (Graubert T 2011, Papaemmanuil E 2013, T Haferlach 2014, Bejar R 2014 and Lindsley RC 2017), molecules and signalling pathways (Bernasconi 2008, Pellagatti A 201 and Follo MY 2012) and the characterization of the pro-inflammatory intramedullary environment with overexpression of cytokines like Tumor Necrosis Factor- α (TNF α) and Interleukins (IL) (Greenberg 2002, Liesveld JL 2004, Venkateswaran N 2016 and Moudra A 2016), in an attempt to better characterize disease onset, evolution and management.

Table 1 – WHO classification of Myelodysplastic Syndrome. Except for MDS-u, disease severity and IPSS-R risk increases from top to bottom.

Name	Abbreviation	Blood Findings	Bone Marrow Findings	Observations
Refractory Cytopenia with Unilineage Dysplasia (RCUD)	Refractory Anemia (RA)	- Unicytopenia, occasionally bicytopenia - No or rare blasts (<1%)	- Unilineage dysplasia (>10% of cells in one myeloid lineage) - <5% blasts - <15% of erythroid precursors are ring sideroblasts	- Miscellaneous driver mutations lead to this pathology
	Refractory Neutropenia (RN)			
	Refractory Thrombocytopenia (RT)			
Refractory Anemia with Ring Sideroblasts	RARS	- Anemia - No blasts	- >15% of erythroid precursors are ring sideroblasts - Erythroid dysplasia only - <5% blasts	- <u>Exclusively</u> SF3B1 (RNA Splicing) - In the case of RARS+RT, it is associated with a SF3B1 co-mutation with JAK2 (60%) or MPL
MDS associated with isolated del(5q)	Del(5q)	- Anemia - Usually normal or increased platelet count - No or rare blasts (<1%)	- Isolated 5q31 deletion - Normal to increased megakaryocytes with hypolobated nuclei - <5% Blasts - No Auer Rods	- MDS-del(5q) is associated only with the 5q31 deletion - More complex phenotypes/karyotypes can be present due to mutation build-up - Responsive to lenalidomide - Lower risk of AML progression
Refractory Cytopenia with Multilineage Dysplasia	RCMD	- Cytopenia - No or rare blasts (<1%) - No Auer rods - Monocytes <1x10⁹/L	- >10% of dysplastic cells in >2 myeloid lineages - <5% blasts - No Auer rods - ± 15% ring sideroblasts	- A mix of driver mutations on genes involved in <u>RNA splicing</u> (SRF2, U2AF1) or <u>DNA methylation</u> (TET2, DNMT3A) and <u>subclonal driver mutations</u> on genes like ASXL1, EZH2, RUNX1 and TP53.
Refractory Anemia with Excess Blasts, type 1	RAEB-1	- Cytopenia - <5% blasts - No Auer rods - Monocytes <1x10⁹/L	- Uni/multilineage dysplasia - 5-9% blasts - No Auer rods	- Miscellaneous driver mutations lead to this pathology, similar to RCMD
Refractory Anemia with Excess Blasts, type 2	RAEB-2	- Cytopenia - 5-19% blasts - ± Auer rods - Monocytes <1x10⁹/L	- Uni/multilineage dysplasia - 10-19% blasts - ± Auer rods	- Miscellaneous driver mutations lead to this pathology, similar to RCMD
MDS - unclassified	MDS-u	- Cytopenia - <1% blasts	- Minimal dysplasia but clonal cytogenetic abnormality considered evidence of MDS - <5% blasts	- Other del(5q) - del(7q) - +8 - del(Y)

Table 2 – 2016-revised WHO classification of Myelodysplastic Syndrome. Except for MDS-u, disease severity and IPSS-R risk increases from top to bottom.

Name	Abbreviation	Dysplasias/Cytopenias	Ring Sideroblasts & BM/PB Blasts	Cytogenetics/karyotypes
MDS with single lineage dysplasia	MDS-SLD	- 1 dysplastic lineage - 1 or 2 cytopenias	- Ring Sideroblasts <15% - BM Blasts <5% - PB Blasts <1% - No Auer rods	- Any, unless it fulfills all criteria for MDS with isolated del(5q)
MDS with multilineage dysplasia	MDS-MLD	- 2 or 3 dysplastic lineages - 1-3 cytopenias	- Ring Sideroblasts <15% - BM Blasts <5% - PB Blasts <1% - No Auer rods	- Any, unless it fulfills all criteria for MDS with isolated del(5q)
MDS with ring sideroblasts	MDS-RS	- 1-3 dysplastic lineages - 1-3 cytopenias - RS-SLD or RS-MLD	- Ring Sideroblasts >15% - BM Blasts <5% - PB Blasts <1% - No Auer rods	- Any, unless it fulfills all criteria for MDS with isolated del(5q)
MDS with isolated del(5q)	MDS-5q	- 1-3 dysplastic lineages - 1-2 cytopenias	- Ring Sideroblasts: None or any - BM Blasts <5% - PB Blasts <1% - No Auer Rods	- del(5q) alone or with 1 additional abnormality, except del(7) or del(7q)
MDS with Excess Blasts 1	MDS-EB-1	- 0-3 dysplastic lineages - 1-3 cytopenias	- Ring Sideroblasts: None or any - BM Blasts 5-9% - PB Blasts 2-4% - No Auer Rods	- Any
MDS with Excess Blasts 2	MDS-EB-2	- 0-3 dysplastic lineages - 1-3 cytopenias	- Ring Sideroblasts: None or any - BM Blasts 10-19% - PB Blasts 5-19% - No Auer Rods	- Any
MDS - unclassified	MDS-U	- 0-3 dysplastic lineages - 1-3 cytopenias	- Ring Sideroblasts: None or any - BM Blasts <5% - PB Blasts <1% - No Auer Rods	- Any

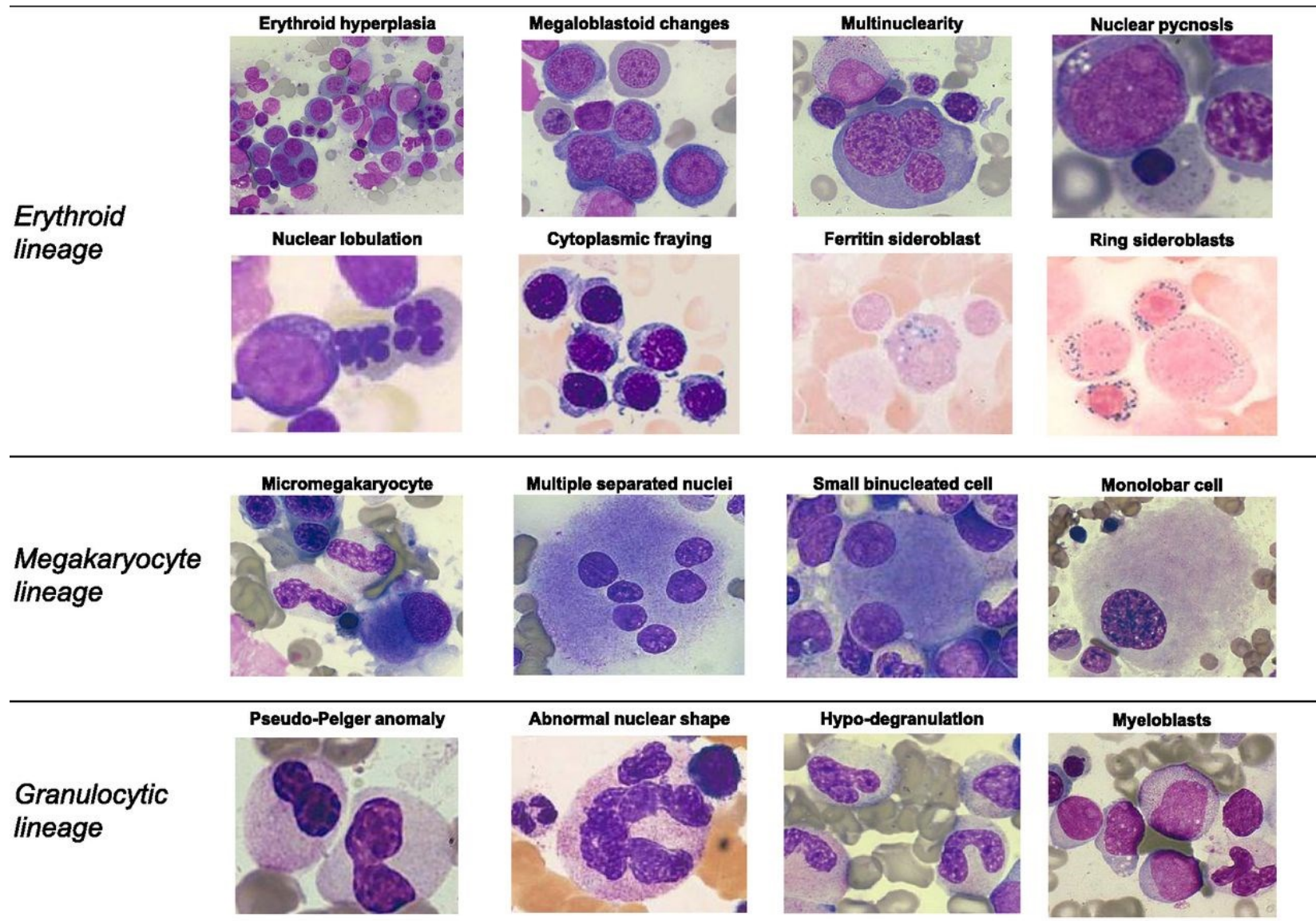


Figure 6 – Examples of cell dysplasia shown in the erythroid, megakaryocytic and granulocytic lineages. Ring sideroblasts are identified by using a Prussian Blue stain (Cazzola M 2011). Adapted from (Cazzola M 2013)

1.2.1.1. Genetics

Across all disease subgroups, more than 90% of MDS patients carry one or more somatic (acquired) mutations (Cazzola M 2013) in genes involved in several biological functions as diverse as RNA splicing, DNA methylation and repair control, chromatin modification and kinase signalling (Bejar R 2014). Multiple studies (Cazzola M 2013, Tothova Z 2013, Itzykson R 2013, Bejar R 2014, Haferlach T 2014, Harada H 2015 and Steensma 2015) report the prevalence of these mutations (Figure 7). The most commonly mutated genes seem to be part of epigenetic regulation (TET2, ASXL1, DNMT3A), RNA splicing (SF3B1, SRSF2, U2AF1), transcription (RUNX1) as well as genes responsible for triggering DNA repair and apoptosis, with tumor protein p53 (TP53) being the most prominent example.

TET2 and DNMT3A are involved in DNA methylation. TET2 mutations yield a loss of function, and unlike mutations in RNA splicing genes, are not mutually exclusive of mutations in other epigenetic regulators (Bejar R 2011). TET2 mutations seem to be highly present (50-60%) in patients with Chronic Myelomonocytic Leukemia (CMML) (Kosmider O 2009), a myelodysplastic/myeloproliferative condition. These patients have abnormal-looking cells (dysplasia) in their bone marrow as well as an abnormally high number of white blood cells (myeloproliferative) with blast levels well below 20% (ACS 2017). On the other hand, DNMT3A mutations are less frequent than TET2 ones, but patients with such mutations often demonstrate a more aggressive clinical scenario, increased AML transformation rate and shorter overall survival (Walter MJ 2011).

RNA splicing is a processing step where pre-messenger RNA (pre-mRNA) transcripts have their introns removed while exons are joined together to form mature mRNA. Genes like SF3B1, U2AF1 and SRSF2 code for proteins involved in the splicing process, or spliceosome. SF3B1 and U2AF1 are part of the U2 spliceosome unit, which recognizes the 3' splice site at intron-exon

junctions (Padgett 2012). SF3B1 is of the most common mutations (>60%) in MDS patients with ring sideroblasts (MDS-RS) while being mutated in 10-15% of other MDS subtypes (Patnaik MM 2012). Patients with SF3B1 mutations seem to have a longer event-free survival (Malcovati L 2015) but little is known about patients with U2AF1 mutations (Zhang L 2015). SRSF2 encodes a protein crucial for pre-mRNA splicing, being responsible for marking exon/intron boundaries and redefining splice sites, which can lead to the formation of alternative transcripts (Makishima H 2012). Most of the MDS patients harbouring SRSF2 have at least one additional mutation (Kao HW 2011) and seem to have lower overall survival (Wu SJ 2012).

RUNX1 is one of the Runt-related transcription factor (RUNX) family of genes and is specifically involved in regulating HSC differentiation to mature blood cells (Okuda T 2001). RUNX1 mutations were first described in patients with AML (Song WJ 1999) and seem to be found at a much higher frequency in patients with high-risk MDS than with low-risk MDS (Harada H 2004). Patients exhibiting RUNX1 mutations are often linked with poor MDS prognosis (Dicker F 2010).

Tumor protein p53 is encoded by TP53 and binds directly to DNA, determining if the cell containing damaged DNA should undergo repair or apoptosis. Mutations in TP53 have a negative impact on prognosis for all MDS subtypes, accounting for a 11-fold shorter overall survival period when compared to wtTP53 (Caceres G 2013).

The aforementioned genes, their functions and the relevance to the prognosis of MDS patients are summarized in Table 3.

Table 3 – A summary of selected genes commonly mutated on MDS, their function and prognosis to MDS patients.

Category	Gene	Protein function	Prognosis
Epigenetic Regulators	TET2	Hydroxylation of methylated cytosine	Neutral
	DMNT3A	DNA methyltransferase	Adverse
Spliceosome and pre-mRNA processing	SF3B1	U2 spliceosome component	Favorable
	U2AF1	U2 auxiliary protein	Unknown
	SRSF2	Marking exon/intron boundaries	Adverse
Transcription factor	RUNX1	Crucial for HSC differentiation towards mature blood cells	Adverse
DNA repair	TP53	DNA repair and apoptosis	Adverse

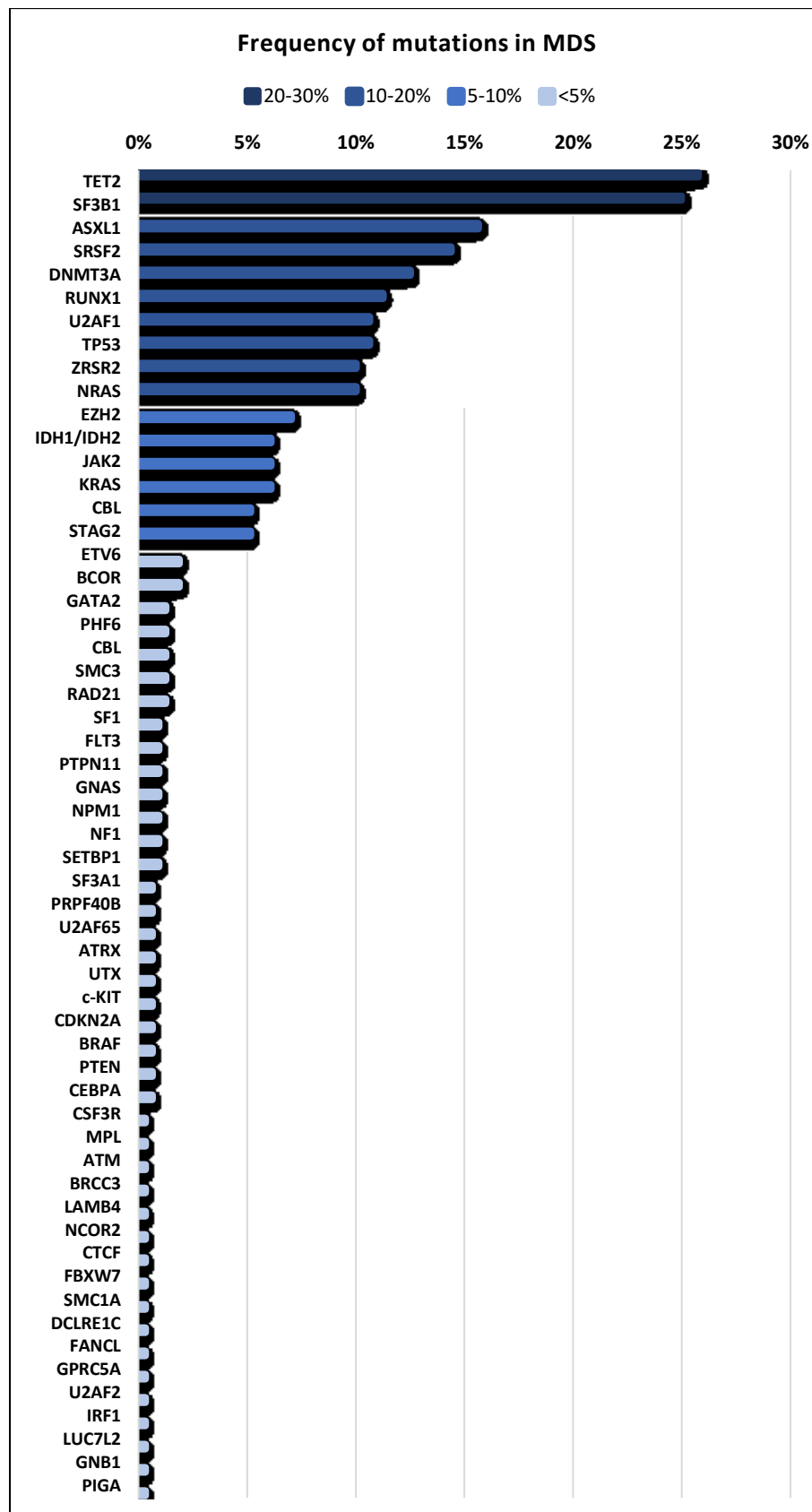


Figure 7 – Frequency of MDS somatic mutations. Data extracted and combined from Cazzola M 2013, Tothova Z 2013, Itzykson R 2013, Bejar R 2014, Haferlach T 2014, Harada H 2015 and Steensma 2015.

1.2.1.2. Cytokines and Stroma

Cytokines and chemokines are present in abnormal concentrations in the blood serum of MDS patients. Despite great advances in the field of cytokine profiling of MDS patients and its comparison to healthy levels, to date there has not yet been established a clinical diagnostic framework that relies solely on cytokine levels. The pro-inflammatory intramedullary environment present in most of the MDS patients seems to be translated in excessive apoptosis within the bone marrow. This phenomena is clearly observed in low-risk MDS patients, perhaps as the body's mechanism of defense against the mutated cells starting to build-up in the bone marrow. As the disease evolves, there seems to be an over-production of proliferative cytokines, whose concentrations peak in high-risk MDS patients, muting the apoptotic effort and ultimately leading to clonal expansion and disease progression.

Several studies consistently report increased intramedullary concentrations of TNF α and Interferon Gamma (IFN γ) (Mundle SD 1996, Greenberg 2001 and Sloand EM 2010) as well as IL32, a TNF α inducer (Marcondes AM 2008). High levels of CXCL10, IL6 and IL7 were shown to decrease patient median survival by 2-fold, when compared with patients with normal levels of all of the three cytokines (Pardanani A 2012). A study showed that IL1 Receptor Accessory Protein (IL1RAP) is overexpressed in MDS HSPCs and has a clear implication on the IPSS-R scoring by increasing risk levels. This protein is constitutively expressed in patients with AML and high-risk MDS, but its targeted downregulation seems to translate in an inhibitory effect on the clonogenic capacity of the disease (Barreyro L 2012). IL-10 concentration seems to positively correlate with MDS risk and may be crucial for disease progression (Kordasti SY 2009). Recent studies have shown an increase in pro-inflammatory/apoptotic cytokine concentrations on patients undergoing azacytidine treatment (Moudra A 2016) which directly implies an upregulation of T cell cytotoxicity and apoptosis induction (D'Silva S 2019).

Stroma also seems to play a role in the pathophysiology of the disease. A pioneering study using murine osteoprogenitors showed that deleting the protein encoded by DICER1, crucial for RNA interference mechanisms, would disrupt normal haematopoiesis. These deletion mutants show reduced SBDS expression (Orelia C 2011), and the targeted deletion of SBDS triggers both bone marrow dysfunction with myelodysplasia (Raaijmakers MHGP 2010) and erythroid progenitor apoptosis (Sen S 2011).

Studies have shown that MDS-derived MSCs impair the *in-vitro* growth of HSPCs, leading to epigenetic and phenotypical abnormalities (Poon Z 2018). MDS MSCs themselves exhibit major transcriptomic differences when compared to age-matched controls. A study reports that about 200 genes are differentially expressed, with a demarked downregulation of early adipogenic differentiation and upregulation of extracellular matrix (ECM) remodeling (Jann JC 2016). Recent reports show that MDS-derived MSCs cultured *in-vitro* express statistically significant higher levels of pro-inflammatory cytokines like IL-6, IL-10 and TNF α when compared to healthy controls. When treated with 5-Azacytidine, IL-6 production levels decrease sharply to those of healthy controls (Boada M 2021). This is extremely interesting as 5-Azacytidine has been widely used as the standard-of-care therapy for patients with high-risk MDS (Buckstein R 2011), in an attempt to halt disease progression towards AML. Determining the exact causality between dysregulated stroma and impaired haematopoiesis and dysplasia, the relevant underlying mechanisms and the route towards disease progression may pave avenues for future studies and therapeutical targets.

1.2.1.3. Dyserythropoiesis

Diseased and dysplastic erythropoiesis, commonly known as dyserythropoiesis, is a major hallmark of MDS. Being the most common feature across every MDS subtype and patient, it can primarily manifest itself as fatigue and dizziness (Goldberg SL 2010), paleness (Foucar K 1985) and an impaired sense of wellbeing (Jansen AJ 2003). According to the WHO, dyserythropoiesis

is diagnosed when patients show low haemoglobin levels (<10g/dL) coupled with morphological evidences of a change in shape and format (dysplasia) of >10% of the erythroid progenitors on a bone marrow smear, with less than 1% of blasts on a peripheral blood smear (Swerdlow SH 2008). The causes of dyserythropoiesis are multiple and very hard to untangle, but are most likely comprised of two main sources: abnormal inflammatory signaling and the build-up of genetic mutations.

As discussed above, several groups report excessive inflammatory cytokine levels in patients with MDS, which are directly linked with adverse symptomology like fatigue, disease progression and survival (Gañán-Gómez I 2015). Inflammatory cytokines TNF α (Shetty V 1996 and Sawanobori M 2003), IFN γ (Allampallam K 1999) and Interleukin-1-alpha/beta (IL1 α /IL1 β) (Feng X 2011) are excessively present on the blood serum of MDS patients. On the other hand, TNF α (Rusten LS 1994), IFN γ (Thawani N 2006) and IL1 α /IL1 β (Johnson CS 1991) are known for suppressing both *in-vitro* and *in-vivo* erythropoiesis (Nemeth E 2014). Moreover, TNF α and IFN γ are believed to inhibit the *ex-vivo* proliferation of human haematopoietic stem and progenitor cells (Dybedal I 2001) while suppressing haematopoietic colony formation (Selleri C 1995) and differentiation of haematopoietic cells across all lineages (Zhang Y 1995 and Schuettpeitz LG 2013). The overexpression of these cytokines in MDS patients could be one of the causes for ineffective erythropoiesis.

Mutations in one or more of the six GATA-family transcription factors found in vertebrates (Shimizu R 2005) are also a putative cause for dyserythropoiesis. GATA1, GATA2 and GATA3 are mainly expressed in haematopoietic tissues, being the so-called “haematopoietic GATA factors” (Shimizu R 2016). While GATA1 and GATA2 play a role on both the lymphoid and the myeloid lineage, GATA3 has a restricted effect on the lymphoid lineage only (Collin M 2015). GATA2 is highly expressed at the HSC and early haematopoietic progenitor level, activating genes crucial for pluripotency, cell survival and proliferation. During erythropoiesis, the GATA2 locus is effectively shut down and the GATA1 expression levels spike, in a process called “GATA-

switch” (Shimizu R 2016). The GATA switch is a major regulatory mechanism in the erythropoietic differentiation (Suzuki M 2013) and mutations in one of these two transcription factors can lead to major erythroid dysplasia, impaired erythropoiesis and early progenitor apoptosis (Collin M 2015 and Romano O 2020). Interestingly, GATA2 promotes the expression of c-Kit (SCF receptor expressed at HSPC level and responsible for HSC proliferation and differentiation arrest), whereas GATA1 heavily downregulates the expression of c-Kit. It is believed that this GATA-switch drives cells from proliferation (GATA2-dominant) to differentiation (GATA1-dominant) (Bresnick EH 2005), and GATA1 mutations can induce erythroblast maturation arrest (Ribeil JA 2007).

GATA1 is strongly associated with a chaperone called Heat Shock Protein 70 (HSP70), which is constitutively present and highly expressed in the cytoplasm and nucleus of healthy human erythroblasts. Under apoptotic pressure and caspase activation, this chaperone translocates to the nucleus to protect GATA1 from cleavage, maturation arrest and programmed cell death. Interestingly, a group concluded that HSP70 trafficking and gene expression is altered in MDS and its restitution to normal expression and function levels managed to rescue GATA1 expression thus resulting in further erythroid maturation, despite not alleviating apoptosis (Frisan E 2012). This may mean that alternative apoptotic mechanisms, unrelated to GATA1 cleavage, induce apoptosis in MDS erythroblasts, either due to genetic mutations, constitutive activation of the caspase complex, toxicity of the pro-inflammatory intramedullary microenvironment or even mutations in cytokine receptors and the associated signalling pathways.

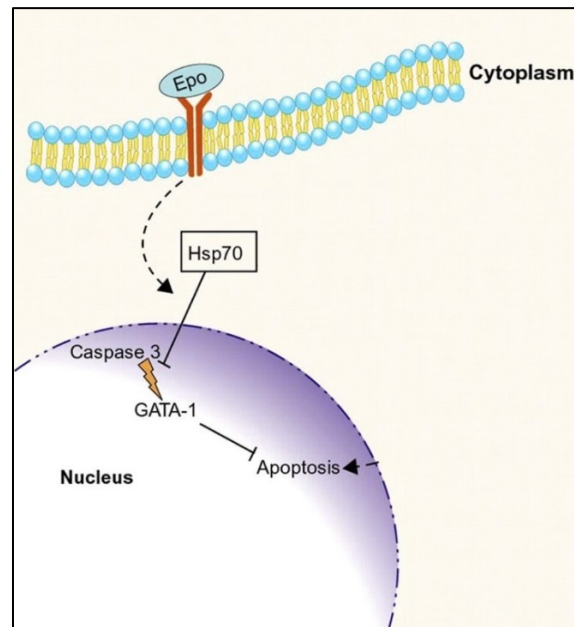


Figure 8 – A simplified schematic of the protective mechanism of Hsp70 on GATA-1, by preventing Caspase-3 dependant apoptosis. Adapted from (Weiss MJ 2009).

Current research in the field has employed multi-omics techniques, focusing in areas as diverse as genomics (Schwartz JR 2017, Nassereddine S 2017 and Ogawa 2019), transcriptomics (Gerstung M 2015, Shiozawa Y 2017 and Im H 2018) and proteomics (Aivado M 2006, Majek P 2014 and Ragon BK 2016), but to date, limited work has focused metabolomics (Khalid A 2018). Metabolomics typically relies on the detection of very small molecules such as glucose, amino-acids and small lipids, which can also be easily measured in routine blood scans. There is an unmet research need for MDS work using metabolomics, as this could potentially unveil candidate biomarkers for disease sub-group and risk stratification, as well as for leukaemia progression.

1.3. *In-vitro* models of haematological malignancies

While culturing human cells *in-vitro*, every researcher aims to create representations of the *in-vivo* scenario, in order to study topics as diverse as fundamental biology, molecular mechanisms or response to biochemical cues. The closer the *in-vitro* representation is from the real *in-vivo* scenario, the more accurate potential the conclusions will be. More often than not, researchers have to compromise and make significant *in-vitro* approximations compared to the *in-vivo* scenario, in order to optimise experimental costs, complexity and outputs without compromising system fidelity and reproducibility.

Cancer cell lines – “immortalized” cells whose progeny yields the exact same genomic information over multiple cell divisions – are often used as good approximations of primary oncological patient cells. Extensive work has been done using the NCI60 cancer cell line repository, that comprises cells from 10 main cancer types (NCI, NCI-60 Human Tumor Cell Lines Screen 2015) for areas as diverse as drug discovery and screening (Shoemaker 2006 and Wilding JL 2014), multi-omics (Su G 2011 and Park ES 2010) and karyotyping (Roschke AV 2003 and Lorenzi PL 2009). Despite good initial cellular models, cancer cell lines do not fully represent the wide phenotypical spectra of cancer patients.

In-vitro cell culture protocols have primarily relied on two-dimensional (2D) culture systems, often using culture medium supplemented with foetal bovine serum (FBS) to provide additional nutrients, proteins and minerals. Some protocols also incorporate the addition of a supporting feeder (stromal) layer, composed primarily of fibroblasts or MSCs. 2D systems are easily maintained and highly scalable but unfortunately are not fully representative of the delicate, complex *in-vivo* microenvironments and interactions, crucial for disease maintenance and progression (Panoskaltsis N 2005 and Duval K 2017).

2D liquid suspension (LS) cultures replace stromal support by abnormal amounts of cytokines which can lead to cell mutations in culture; 2D LS co-cultures where stromal support is used typically use exogenous stromal cells, inserting human or animal antigens on cell cultures, which can skew the results (Panoskaltsis N 2005). Primary patient cells cultured in 2D seem to be more prone to apoptosis as opposed to their culture in 3D (Imamura Y 2015). This is even more relevant in the field of cancer research, as it is believed that the patient-specific *in-vivo* tumour microenvironment, ECM density and composition, 3D conformation, and nutrient/oxygen gradients play a pivotal role in cancer homeostasis, metastasis and expansion (Whiteside 2008, Wu M 2014, Wang M 2017, Lee SW 2018 and Kim H 2018). Several studies use 3D culture systems to study different types of cancer, as they can potentially incorporate several *in-vivo* tumour characteristics such as vasculature, ECM composition, cellular distribution and mechanical conformation (Riedl A 2017, Kapalczyńska M 2018 and Melissaridou S 2019).

3D cell culture systems have gained importance over the past decade in diverse applications, ranging from drug discovery and repositioning (Langhans 2018), cell-based biosensors (Edmondson R 2014) and cancer tumor growth (Thoma CR 2014). 3D scaffolds have been used to culture cell types as different as MSCs (Cesarz Z 2015), skin fibroblasts (Sun T 2006 and Randall MJ 2018) and pancreatic cancer cells (Gupta P 2020 and Wishart G 2021), which also naturally form 3D scaffold-free cell aggregates in culture. While designing scaffolds for 3D cell culture, researchers can tweak both the scaffold material composition as well as adding extracellular matrix proteins as coating to promote cell adhesion. Molecules such as alginate (Chen CY 2015, Cavo M 2018 and Bidarra SJ 2019) and chitosan (Cho MO 2016, Kafi A 2019 and Chang PH 2021) are widely used, often coupled with ECM protein coatings like collagen (Guo J 2019) or commercially-available mixtures like Matrigel (Hongisto V 2013). Using biocompatible materials and trying to resemble as much as possible the native cellular *in vivo* environment is paramount (Ravi M 2015).

Previous work in our lab was primarily focused in culturing AML cells and developed a porous polyurethane scaffold mimicking the natural 3D environment and conformation of the bone marrow. This culture system has allowed for long-term, cytokine-free cultures of both healthy human cells (M. A. Mortera-Blanco T 2011) as well as leukaemic cell lines (Mortera-Blanco T 2010). This culture system is now used in the work herein presented as a starting point to sustain long-term 3D culture of primary MDS cells.

1.3.1. *In-vitro* models of Myelodysplastic Syndrome

During the 1990s and early 2000s, several scientific groups developed outstanding efforts in trying to better understand MDS. In 1990, researchers cultured bone marrow cells isolated from MDS patients for 48 hours supplementing the culture medium with either Granulocyte Colony-Stimulating Factor (G-CSF), Granulocyte-Macrophage Colony-Stimulating Factor (GM-CSF) or EPO for chromosome analysis, suggesting that a robust *in-vitro* cell culture protocol would be the way to proceed for further exploration and studies (Ohayashiki K 1990). The response of bone marrow progenitors to haematopoietic cytokines was further investigated in a study including G-CSF, GM-CSF and IL-3 which showed a similar impairment on colony growth in most MDS patients, suggesting that these cells might not be fully able to respond to cytokine stimuli, even after using over-saturating cytokine concentrations (Merchav S 1991). A retrospective study published in the same year analysed the colony forming capabilities of MDS cells, confirming results published earlier in 1980 that correlated *in-vitro* abnormal growth with abnormal differentiation in MDS patients (Raymakers R 1991). A study performed in 1994 included SCF in the previously used cytokine cocktail (G-CSF, GM-CSF and IL3) and demonstrated its positive effect by enhancing *in-vitro* erythropoiesis, but only in MDS patients without severe anaemia, while having a neglectable effect on megakaryopoiesis (Piacibello W 1994).

The first *in-vitro* 2D stroma-free long-term culture system (28 days) of BM Mononuclear Cells (MNC) derived from MDS patients was established in 1996, obtaining a significant response (CFU-GM numbers increased at Week 4) in half of the patients studied (19/38), whereas a third of the total patients (12/38) was cytokine-irresponsive (Soligo DA 1996). The same group repeated its successful *in-vitro* culture protocol one year later, now more focused on analyzing karyotypic evolution throughout a 4 week culture. Despite studying only 6 patients, this study shows that after a leukaemic transformation in culture (n=1), the number of clones containing trisomy 8 decreased sharply – suggesting that cells with abnormal karyotype might not be single-handedly responsible for disease progression after stimulation with growth factors (Bossolasco P 1997). Another study analysed both colony-formation capacity and bcl-2 expression, a protein encoded by BCL2 and considered to be a strong apoptosis inhibitor. This study found that bcl-2 has a low expression in MDS MNCs when compared to healthy controls, while the MDS group had an higher percentage of apoptotic cells. These results confirmed what we know now for a fact: early stage MDS cells seem to have a higher-than-normal apoptosis rate while keeping normal proliferation rates, which is the putative cause of peripheral blood cytopenias observed in most patients (Bincoletto C 1998). A summary of these studies is shown in Table 4.

Table 4 – Studies using patient-derived BM-MNCs its cellular source, type and duration, cytokines used and study objectives. **Legend:** **2D LS** - 2D Liquid Suspension; **CR** – Cytokine Responsiveness; **CFS** – Colony Formation Studies;

Type	Duration	Cytokines Used	Objectives	Main Findings	Reference
2D LS	48h	G-CSF, GM-CSF, EPO	CR, CFS	Direct correlation between cell numbers and cytokine concentrations; differentiation arrest	(Ohyashiki K 1990)
2D LS	N/A	G-CSF, GM-CSF, IL3	CR	Demonstrates that MDS cells do not seem to be able to fully respond to cytokines, even when present at saturating concentrations	(Merchav S 1991)

2D LS	7d	M-CSF	CFS	First correlation established between <i>in vitro</i> abnormal growth and <i>in vivo</i> abnormal differentiation in MDS patients	(Raymakers R 1991)
2D LS	<14d	G-CSF, GM-CSF, IL3, EPO, SCF	CR, CFS	Supplementation with SCF had a positive effect in <i>in-vitro</i> erythropoiesis, but only in patients without severe anemia	(Piacibello W 1994)
2D LS & feeder-free	28d	G-CSF, GM-CSF, IL3, SCF, IL6	Long-term culture, CR	First long-term, 2D, feeder-free culture of MDS cells. One third of samples showed irresponsiveness to cytokines	(Soligo DA 1996)
2D LS	28d	G-CSF, GM-CSF, IL3, SCF, IL6	Long-term culture, karyotype analysis	Proposes that cells with abnormal karyotypes might not be single-handedly responsible for disease progression <i>in vivo</i>	(Bossolasco P 1997)
2D LS	14d	No cytokines used	CFS, bcl-2 expression	High BCL2 (apoptosis inhibitor) expression vs healthy controls	(Bincoletto C 1998)

Coupled with *in-vitro* models, several groups have utilised *in-vivo* models to better understand disease initiation and evolution mechanisms. Using murine models, researchers have evaluated leukaemia-initiating factors such as chemical components (Alfred LJ 1983), radiation (Finch 2007) and viral infections (Singer D 1974). Moreover, immunocompromised mice (NOD/SCID) have been widely used for human patient xenograft transplantation assessing cultured cell engraftment and leukaemia-initiating capabilities using primary MDS cells (Ito M 2002, Thanopolou E 2004 and Theocharides AP 2016). The *in-vivo* systems play a key part in understanding how leukaemia is initiated and sustained and are an invaluable tool for researchers in the onco-haematological field. The main limitations of *in-vivo* systems are their costly maintenance and the need to use high-end biosafety facilities and employ highly-skilled researchers. When executing *in-vitro* work,

consumable and facility costs are several fold smaller and researchers can start working immediately with minimal training. However, the main limitations of *in-vitro* systems are its reproducibility and robustness when trying to translate results to clinical practice, as *in-vitro* systems are indeed just an adaptation of the *in-vivo* scenario.

Historically, a vast majority of the *in-vitro* studies with MDS studies were focused in assessing the *in-vitro* clonogenic potential and response to cytokines, with very limited work focused in establishing long-term, 3D patient-representative cultures. The lack of an appropriate MDS cell line (Drexler HG 2009) could potentially be one of the causes, as researchers have to rely on valuable primary patient samples to test their cell culture systems. This poses another problem, as data from 2014 shows that the incidence of MDS in the UK is roughly 3-fold smaller than that of all leukaemias (CRUK 2014), limiting the numbers of patients that can potentially be enrolled in clinical trials and voluntarily provide samples of their cells.

To date and to the author's best knowledge, one group has successfully developed a 3D system to culture primary MDS cells using cytokine concentrations at least ten-fold higher than those present *in-vivo* (Elvarsdóttir EM 2020). Despite showing extremely promising and relevant results, the cell culture protocol applied could potentially have a limited application in patient-specific, personalised medicine since it uses xenogeneic components (FBS) and high cytokine concentrations, which are not fully representative of the *in-vivo* diseased environment.

There is, at the time of writing, an unmet need for the creation of patient-specific, serum-and-xeno-free long-term 3D culture systems – which represent as much as possible the *in-vivo* environment – and that can adequately mimic MDS features such as apoptosis, differentiation arrest and low response to cytokines. These will form the basis of a platform for drug discovery and optimisation, fundamental studies as well as candidate biomarker discovery.

1.4. *In-silico* models of healthy and diseased haematopoiesis

In-silico modelling of healthy and diseased haematopoiesis has become increasingly popular due to the infinite scenarios and study possibilities that computer-based models can provide. However, it is quite hard to capture biological the healthy homeostatic heterogeneity by mathematical equations or functions, let alone modelling haematological malignancies, which are defined by extreme complexity and have intrinsic patient-to-patient variability.

Work as early as 1977 focused on blood cell kinetics by performing labelling with radioactive hydrogen and chromium isotopes to measure their *in-vivo* production rates. Despite pioneering its field, this work only focused on healthy blood production and the computed rates are obviously subject to external factors like diet, smoking, alcohol consumption and gender, to name a few, which are not considered explicitly in this study (Finch CA 1977). Mackey and Glass pioneered haematopoietic modelling of white blood cell production by studying a periodic form of chronic myelogenous leukaemia, introducing delayed differential equations in haematopoietic mathematical modelling and paving the way for future work in the field (Mackey MC 1977). After the discovery of G-CSF by Metcalf in 1985, groups started modelling its effects on blood progenitors (Bungart B 1990). With the rise and global availability of computers in the new millenium, increased computational power and refined modelling tools and platforms, more models emerged, both for healthy and diseased haematopoiesis. The action of cytokines on haematopoietic progenitors was studied using G-CSF in 2004, simulating the effects of continued G-CSF therapeutic injections (Obeyesekere 2004).

Haematopoiesis was also modelled stochastically resorting to compartmentalization of the haematopoietic differentiation cascade into small bins. Despite being simple, this model shows agreement with the clinical data provided for PIG-A mutants (Dingli D 2007). One study further elucidated how to mathematically describe the effects of activation and inhibition mechanisms both on differentiation and proliferation of cells, which are known to happen *in-vivo* as feedback loops (Lander AD 2009). Further work was performed using the concept of compartmentalized

haematopoiesis on modelling inhibitory and activating feedback loops; this was experimentally validated showing relationships between self-renewal, differentiation, and the concentration of stimulatory factors (Kirouac DC 2009). A theoretical work was produced at the same time modelling the effects of a generic signal that was dependant on the terminally-differentiated cell population number, on differentiation, self-renewal (proliferation) or both, and managed to conclude that self-renewal seems to be crucial for efficient repopulation, when compared with trends seen in leukocyte repopulation kinetics after chemotherapy and subsequent HSC transplantation (Marciniak-Czochra A 2009). Other authors developed models using stochastic elements (Werner B 2011), delayed differential equations (DDE) (Afenya E 2010), ordinary differential equations (ODE), including competition between healthy and diseased cells (Stiehl T 2012) (MacLean AL 2012) and ordinary differential equations with stochastic elements (ODE-S) (Werner B 2016) to account for dynamics of leukaemic and healthy cells.

Several groups developed mathematical models for different haematological malignancies like Chronic Myeloid Leukaemia (CML) (Colijn C 2005 and Michor F 2005) and AML (Andersen LK 2001 and Stiehl T 2014). With respect to MDS, the existent mathematical models are focused on studying bone marrow-peripheral blood dynamics (Obeyesekere MN 2005), competition between healthy and dysplastic stem and progenitor cells (Walenda T 2014), bone marrow haematopoietic cell maturation delay and dynamics (Afenya EK 2016) and the effects of different chemotherapy on MDS patients, including lenalidomide and hypomethylating agents (HMA) alone and combined (Drusbosky L 2017). A summary of the aforementioned publications can be seen in Table 5. Exciting developments have seen the creation of an AML patient-specific mathematical model that can capture chemotherapeutic outputs in the peripheral blood, which can potentially pave the way for *in-silico* therapy optimisation (Quiroga-Campano AL 2017). To the best of our knowledge, no single model of MDS taking into account specific cytokines or looking deeply into each of the myeloid lineages affected has been published so far.

Table 5 – A summary of selected published mathematical models addressing both healthy and diseased haematopoiesis. Legend: **CBA** - Computational Biology Algorithm; **DDE** – Delayed Differential Equations; **LSC** – Leukaemic Stem Cell; **ODE** – Ordinary Differential Equations; **ODE-S** – Ordinary Differential Equations with Stochastic elements.

Equations Used	Objective of Study	Reference
Mathematical Modelling of Healthy Hematopoiesis		
Algebraic	Kinetics of formed blood cell elements	(Finch CA 1977)
DDE	Neutrophils in periodic CML	(Mackey MC 1977)
ODE	G-CSF effects	(Bungart B 1990)
ODE	Therapeutic G-CSF effects	(Obeyesekere 2004)
Stochastic	Haematopoiesis	(Dingli D 2007)
ODE	Feedback loops and control	(Lander AD 2009)
ODE-S	Activation and inhibition feedback loops	(M. G. Kirouac DC 2009)
ODE	HSC self-renewal post Chemotherapy	(Marciniak-Czochra A 2009)
ODE	Haematopoietic stem and progenitor cell kinetics	(Bast L 2021)
Mathematical Modelling of Diseased Hematopoiesis		
DDE	Chronic Myeloid Leukaemia	(Colijn C 2005)
ODE-S	Chronic Myeloid Leukaemia	(Michor F 2005)
DDE	Acute Myeloid Leukaemia	(Andersen LK 2001)
ODE-S	Acute Myeloid Leukaemia	(Stiehl T 2014)
Stochastic	LSC/HSC dynamics	(Werner B 2011)
DDE	LSC/HSC dynamics	(Afenya E 2010)
ODE	LSC/HSC dynamics	(Stiehl T 2012)
ODE	LSC/HSC dynamics	(MacLean AL 2012)
ODE-S	LSC/HSC dynamics	(Werner B 2016)
ODE	Myelodysplastic Syndrome	(Obeyesekere MN 2005)
ODE	Myelodysplastic Syndrome	(Walenda T 2014)
ODE	Myelodysplastic Syndrome	(Afenya EK 2016)
ODE	Chronic and Acute Myeloid Leukaemia	(Parajdi LG 2020)
ODE	Myeloproliferative Neoplasms	(Sajid 2021)

2. Aims and Objectives

The work herein described aims to develop several *in-vitro*, *in-silico* technologies to model and analyse haematological malignancies. Using MDS as a starting point, we aimed to better understand the disease's pathophysiology and how it compared to the normal healthy state. When procuring patient samples, we aimed to maximize the outputs extracted by combining a series of *in-vitro* techniques – cell culture, characterization and omics – and feeding them to state-of-the-art *in-silico* mathematical models (Figure 9).

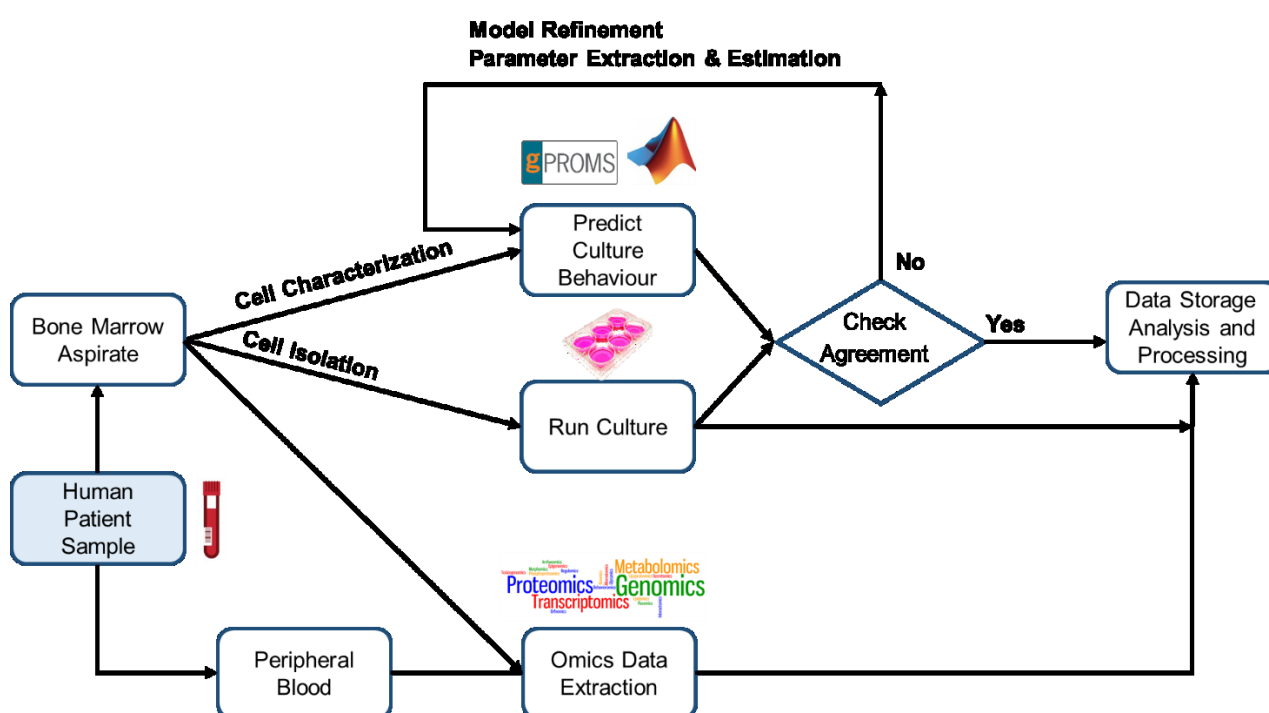


Figure 9 –Proposed framework for the *in-vitro in-silico* modelling of haematological malignancies.

This work is comprised of three main Objectives detailed below, with each objective being presented as an individual Results chapter.

Objective 1: Development of a novel haematopoietic *in-silico* mathematical model (Chapter 4)

There is an unmet need for mathematical models that can potentially assist clinicians in making well-informed decisions, enabling quick and informed selection of the most effective, patient-specific therapeutics. Such models need to be empirically validated and subject to clinical trials prior to their full use in the clinic. However, as a starting step, these models should be based on first-principle mechanisms and include as much known biological information as possible. Previous work focused on the creation of healthy (Obeyesekere MN 2005, Lander AD 2009 and M. G. Kirouac DC 2009) and diseased (Colijn C 2005, Stiehl T 2014 and Afenya EK 2016) haematopoietic models, but often lacked an in-depth description of the biological perspectives of the problem, disregarding several cellular populations or accounting for undefined stimuli.

Objective 2: Establishment of a long-term 3D cell culture platform for primary MDS cells (Chapter 5)

Long-term 3D cell culture platforms for primary patient samples have been highly sought after, as they can potentially allow for the study and development of targeted therapeutics, molecular studies and *in-vitro* hypothesis-generation. Despite several efforts devoted to culturing primary MDS cells, projects have been limited to 2D cultures using abnormal cytokine concentrations, often resorting to exogenous feeder layers for stromal support which leads to a misrepresentation of the features of the disease (Piacibello W 1994) (Soligo DA 1996) (Bossolasco P 1997). Recently, one group has successfully developed a 3D system to culture primary MDS cells using cytokine concentrations at least ten-fold higher than those present *in-vivo* (Elvarsdóttir EM 2020), which could have limited applicability when trying to test near-physiological conditions for patient-specific therapeutics. Our group has recently developed viable, long-term 3D cell culture platforms both for AML (Rende M 2012) and CLL (Santos JD 2017) without the addition of any exogenous cytokines, and this work builds on top of that existing expertise.

Objective 3: Identification of candidate biomarkers for intra-/inter-disease discretization (Chapter 6)

Aiming to gradually replace the need for bone marrow aspirates, we have performed GC-MS metabolomics analyses in plasma obtained from both BM and PB of diseased patients. To date, very limited research has focused in plasma metabolomics of MDS patients, as a vast majority of the groups are focused in intra-cellular metabolomics instead (Stevens BM 2018 and Poulaki A 2020). Using GC-MS metabolomics, this work partially fills the existing gap in the literature as we are able to clearly discretize samples of the same and different haematological malignancies using unsupervised multivariate statistical analyses.

With a fully validated mathematical model, researchers and clinicians would be able to i) optimise cell culture protocols for healthy cell expansion; ii) test several scenarios on the co-existence of healthy and diseased cells; iii) perform large-scale simulations that could help driving insights, first-line therapeutics and better care for patients with haematological malignancies.

Both the *in-vitro* cell culture system and the metabolomics workstreams will aid researchers in exploring and analysing MDS cells, with the overall aim to i) uncover novel therapeutical targets based on *in-vitro* results; ii) propose, test and validate candidate biomarkers that can speed up disease diagnosis and therefore increase life expectancy; iii) enable omics-data-driven large-scale projects in the MDS field, which are to date primarily focused on genomics.

3. Materials and Methods

3.1. Mathematical modelling and simulations

Mathematical models were implemented in gPROMS (v4.2, Process Systems Enterprise). Model simulations were performed using a machine equipped with a Intel(R) Core (TM) i7-3770 CPU @ 3.40GHz, 16.0 GB of RAM, 256 GB of SSD and 1TB of HDD storage. Using the gO:MATLAB tool, simulations were performed in gPROMS and visualized as 3D surfaces in MATLAB (version R2014a, Mathworks), using custom-made scripts. For GSA, GUI-HDMR was used as previously published (Ziehn T 2009).

3.2. Cell isolation

Umbilical cord blood mononuclear cells (UCB-MNC) were isolated from cord blood obtained directly from NHS Blood and Transplant (NHSBT) blood bank (Colindale, London). Cord blood not suitable for transplant was procured from consenting mothers, and according to the relevant ethical approval (05/Q0405/20; NRES Committee, Harrow, London). Primary patient samples were obtained from BM aspirates of consenting patients undergoing diagnosis, treatment or follow-up at either Ealing Hospital or Northwick Park & St. Mark's Hospital. These samples were procured according to the relevant ethical approval (ERC funded BioBlood, No. 340719, IRAS No. 125451). All human samples were processed, isolated and cultured according to the principles set out in the Declaration of Helsinki.

MNCs from both UCB and primary patient samples were isolated not more than 24 hours after collection from the relevant site by a density gradient centrifugation. Briefly, the original sample volume was diluted in four parts of Phosphate-Buffered Saline (PBS) and layered on top of a Ficoll-Paque (GE Healthcare, UK) with 15mL Ficoll per 35mL of diluted blood) Cells were then spun at 450g for 30 minutes with minimum acceleration and no deceleration. Buffy coats were collected and diluted in PBS (4x dilution, or 1:3 dilution ratio), centrifuged at 250g for 7 minutes,

further resuspended in IMDM medium supplemented with 10% FBS (ThermoFisher Scientific, UK) and kept in a 37°C, 5% CO₂ incubator prior to seeding.

3.3. Porous polyurethane scaffold preparation and coating

Porous polyurethane scaffolds were prepared using a Thermally-Induced Phase Separation (TIPS) process, as previously described elsewhere (Mortera-Blanco T 2010). Briefly, polyurethane pellets were dissolved in dioxane (5% w/v) under continuous heating (50°C) and magnetic stirring overnight. 60mL of such dioxane solution were transferred to a Petri dish, frozen at -80°C for at least 2h and freeze-dried in an ethylene glycol bath (-15°C), with continuous solvent removal in a liquid nitrogen cold trap. This allowed for the production of a disk of a porous polyurethane scaffold, with 90-95% porosity and average pore sizes ranging from 100-250µm. The polyurethane disk was then cut into 5x5x5mm cubes which were then used for further experimentation.

Collagen-coated scaffolds were produced as previously described (Mortera-Blanco T 2010) (Mortera-Blanco T 2011). Briefly, 30-40 cubes were washed in 30mL of ethanol 70% (v/v) (VWR, UK), squeezed and transferred to a conical tube containing 30mL of filtered PBS for 10 minutes. Scaffolds were then centrifuged at 220g for 10 minutes, transferred to a conical tube containing 30mL of a collagen type-I solution (Sigma-Aldrich, UK) at a concentration of 62.5µg/mL and centrifuged for 20 minutes at 200g. Finally, scaffold cubes were transferred to a new conical tube with 30mL of filtered PBS and centrifuged at 180g for 10 minutes. Collagen-coated scaffolds were then transferred to a new conical tube containing 30mL of filtered PBS and kept at 4°C prior to cell seeding. For uncoated scaffolds, a similar protocol was followed, replacing the collagen type-I solution for filtered PBS. PBS was prepared fresh by the addition of a single tablet (Gibco, UK) to 500mL of deionized water, and filtered using a 0.22µm membrane filter (Merck, UK).

3.4. Cell culture

Prior to cell seeding, uncoated and collagen-coated scaffolds were sterilized by immersion in ethanol 70%(v/v) for 2h in 24-well plates (Corning, UK) followed by vacuum aspiration of both the scaffold and the supernatant. Scaffolds were then washed twice with PBS and UV-sterilized for 15 minutes. PBS was then removed and scaffolds were dried using vacuum suction, and left inside a 37°C, 5% CO₂ incubator for thermal balance prior to cell seeding.

Isolated cells were resuspended in IMDM supplemented with 10% FBS (Life Technologies, UK) at the desired seeding density per scaffold, accounting for a seeding volume of 50µL. Concentrated cell suspensions were seeded in dried scaffolds and placed inside a 37°C, 5% CO₂ incubator for 45 minutes prior to any medium addition. 1.5mL of pre-warmed (37°C) IMDM+10%FBS was then added to each well of a 24-well plate, containing only one scaffold per well.

For all cell culture experiments, medium was fully replenished every two days by aspiration of the old medium and addition of 1.5mL of pre-warmed (37°C) IMDM+10%FBS medium to each well. Supernatants from three independent wells were collected in microcentrifuge tubes and centrifuged at 900g for 5 minutes, followed by snap-freezing in liquid nitrogen for posterior analyses.

3.5. Cell viability

At every feeding day, cells from three independent supernatant wells were collected by centrifugation of 1.5mL of supernatant medium at 900g for 5 minutes in microcentrifuge tubes. Cell pellets were then washed with 1mL of filtered PBS, centrifuged again at 900g for 5 minutes in microcentrifuge tubes and resuspended in 100µL of fresh, filtered PBS. Cell viability was then assessed using Guava ViaCount (Merck Millipore, UK) using the protocol recommended by the manufacturer. Briefly, 20µL of concentrated cell suspension were mixed with 180µL of ViaCount stain in triplicates in a 96-well round-bottom plate, and analysed using a pre-defined, flow cytometry-based program (ViaCount, Merck Millipore, UK). The composition of the

staining solution is commercially protected, but is composed of two DNA stains – one cell-impermeable and the other cell-permeable. Dead cells were stained positively by a cell-impermeable DNA stain, whereas all nuclear DNA was simultaneously stained using a cell-permeable stain. Dead cells were gated out and viable cells (live cells over total cells) and total viable nucleated cell numbers were calculated. Final results were computed as average $\pm$ SD of N=3 (three independent scaffolds), n=3 (three readings per independent scaffold) read in technical duplicates.

3.6. Flow Cytometry

At Day 0, cells were stained immediately after isolation. At other culture timepoints (Day 8, 16 and 24) supernatant medium was combined in a 50mL conical tube and centrifuged at 250g for 10 minutes. Cell pellets were then washed with 3mL of fresh Cell Staining Buffer (CSB) and centrifuged again at 250g for 10 minutes. CSB was prepared by adding 1% w/v of bovine serum albumin (Sigma Aldrich) and 0.01% w/v of NaN_3 (Sigma-Aldrich, UK) to filtered PBS. After centrifugation, cells were concentrated to a final volume of 300 μL of CSB and divided in 3x100 μL aliquotes in three independent wells of a round bottom 96-well plate, to create unstained and isotype controls, alongside stained cell samples. A staining solution at a concentration of 19.05 $\mu\text{g/mL}$ of Calcein AM Violet in CSB (ThermoFisher Scientific, UK), prepared according to the manufacturer's recommendations, was added at 5 μL per well for viability assessment. Cell suspensions were then incubated at room temperature, in the darkness, for 45 minutes.

Immediately after incubation, cells in the round-bottom 96-well plate were centrifuged for 5 minutes in a standard 96-well plate centrifuge (VWR, UK) and cell pellets were washed with 100 μL of CSB, followed by a 5 minute centrifugation. After washing with 100 μL of CSB, 20 μL of FBS were added to each well to enhance stain specificity. Cells were then stained with 5 μL of mouse anti-human monoclonal antibodies CD45-FITC, CD36-APC, CD71-PECy5 and 1.5 μL of CD235a-PE (BD Biosciences, UK). Similarly, isotype controls were done by adding 5 μL of IgG1-FITC, IgG2a-PECy5, IgM-APC and 1.5 μL of IgG2b-PE (BD Biosciences, UK). Staining

was done for 1h at 4°C, in the darkness. Prior to data acquisition, cells were centrifuged for 5 minutes, washed with 100µL of CSB, centrifuged once more for 5 minutes and resuspended in 300µL of PBS.

Sample acquisition was done on a BD LSR Fortessa (BD Biosciences, UK) equipped with FACSDiva software. A minimum of 20,000 live events were captured. Fluorescence was recorded for Calcein AM (laser 405, 450-50 filter midpoint-bandwidth), CD45-FITC (488 530-30), CD235a-PE (561, 582-15), CD71-PECy5 (561, 660-20) and CD36-APC (640, 670-14). Single-stain controls were performed using positive and negative compensation beads (ThermoFisher Scientific, UK) and unstained controls were performed with unstained cells. Post-run analysis was done using FlowJo (v10, Becton Dickinson, UK).

3.7. Confocal Microscopy

In-situ cell viability was assessed using the LIVE/DEAD Viability/Cytotoxicity Kit for mammalian cells (ThermoFisher Scientific, UK) according to the manufacturer protocol. Briefly, scaffolds were removed from culture and washed twice with PBS prior to its incubation in a staining solution containing 2µM Calcein AM and 4 µM Ethidium Homodimer-1 (EthD-1) for 1h, in the dark. Scaffolds were then snap-frozen in liquid nitrogen and sectioned using sharp stainless steel surgical blades. A section of the core of the scaffold was then imaged in a Leica SP5 inverted confocal microscope, using 10µm steps between Z-stacks and 10x magnification (Leica Microsystems, UK), capturing images with 512x512 pixels of resolution. Using the stitching capabilities of Leica LAS, we collated several of those 512x512 images to be able to capture whole-scaffold section images. The number of viable (green – Calcein AM positive) and dead (red – EthD-1 positive) cells was quantified in raw, whole-scaffold images using the 3D Object Counter plugin of ImageJ (Bolte S 2006 and Schindelin J 2012). Cell viability was computed as the percentage of total cells found to be Calcein AM positive. All confocal images shown are 2D Z-projections of maximum intensity at each Z-stack.

3.8. DNA quantification

The quantification of dsDNA was used as an indirect measure of cell numbers in 2D and 3D cultures, according to other studies (Grayson WL 2010 and Klontzas ME 2017). A working solution of PicoGreen (ThermoFisher Scientific, UK) was prepared according to the manufacturer's instructions. Scaffolds were collected and washed twice with PBS prior to freezing using liquid nitrogen. Scaffolds were then sectioned in slices and incubated overnight at 37°C in 1mL of quantification buffer containing 10mM Tris, 1mM EDTA, 0.1% Triton-X (Sigma-Aldrich, UK) and 0.1mg/mL Proteinase K (ThermoFisher Scientific, UK). The resulting solution was then diluted 10x in PBS prior to sampling of 100µL and mixing with 100µL of PicoGreen (in triplicates) followed by a 5 minute incubation in the dark. Finally, fluorescence was measured with excitation at 485nm and emission at 530nm using a fluorescence plate reader (Glomax Multi-Detection System, Promega, UK). For every group/patient sample, fluorescence levels were normalized to that of Day 1.

3.9. GC-MS metabolomics

GC-MS metabolomics analysis was performed as an adaptation of the work performed by Kanani et al. (Kanani HH 2007 and Kanani HH 2008). Plasma samples were obtained in a clinical setting by centrifuging peripheral blood and bone marrow samples originally collected in EDTA-coated Vacutainer tubes (Beckton Dickinson, UK) at 2,000g for 10 minutes. 100µL of plasma samples were added to 3mL of cold (-20°C) methanol (Sigma-Aldrich, UK). Ribitol (0.1mg/100µL plasma) and U-¹³C Glucose (0.2mg/100µL plasma) were added as internal standards. The resulting solutions were then incubated for 20 minutes at 72°C. After incubation, 3mL of ultra-pure HPLC-grade water (Alfa-Aesar, UK) were added and the resulting solution was mixed thoroughly. Samples were then centrifuged at 10,000g for 10 minutes at 4°C. Supernatants were collected and re-centrifuged until no residual pellet was observed. The final mixture of dry polar metabolites was obtained by vacuum-drying the samples at 50°C overnight in conical glass vials (Corning, UK).

The dry polar metabolite mixture was derivatized for 1.5h in 50µL of a 20mg/mL solution of Methoxyamine Hydrochloride (MeOx) (Sigma-Aldrich, UK) in pyridine (Alfa Aesar, UK), followed by 6h in N-methyl-N-(trimethylsilyl)trifluoroacetamine (Alfa Aesar, UK). Finally, derivatized metabolite mixtures were run in quadruplicate injections in a Shimadzu QP2010 Plus GC-MS instrument (Shimadzu, UK), using a previously published method (Kanani HH 2008). For a complete list of the unique metabolites detected with this method, please refer to Appendix III.

3.10. Metabolomic data processing

Metabolite profiles were analysed with Shimadzu GCMSolution Postrun Analysis Software (v4.11, Shimadzu, UK) according to previously published methods (Kanani HH 2007 and Kanani HH 2008). Metabolite peaks were identified based on their retention time and marker ion pattern. Peak areas were quantified based on the marker ion peak areas and normalized (relative peak areas, RPA) to that of the internal standard (ribitol, marker ion m/z 103). Glucose labelled with heavy carbon atoms (carbon-13) – glucose-C-13 – was added as a second internal standard, to detect intra and inter-sample deviations due to unexpected, unrequested changes in operating conditions of the equipment during or between sample runs. Within the analysed dataset, samples that had an absolute variation of the RPAs of the glucose-C-13 peaks greater than 25% - accounting for less than 5% of all samples and runs analysed – were re-run. Metabolite identification was performed using an in-house custom made method, which was able to detect 121 distinct metabolite peaks (82 known, 39 unknown peaks, including >1 trimethylsilyl derivatives of the metabolites containing at least one amine group and two methoxime derivatives of ketone-containing molecules). Peaks with a signal-to-noise ratio of less than 3 were automatically filtered out of the analysis. Peaks with an absolute coefficient of variation higher than 25% between injections ($n=4$) were removed from further analyses, as previously described in the literature (Kanani HH 2007 and Kanani HH 2008).

For multivariate statistical analyses, standardized peak areas were used. For every unique sample, these were computed by subtracting the mean RPA of each metabolite (across all

samples) and dividing by its standard deviation (across all samples). Unsupervised Principal Component Analysis (PCA) was performed in R (v3.4.2) using scripts created in-house (Appendix IV).

3.11. Statistical analyses

Statistical analysis was performed using ANOVA and Tukey's post-hoc test. Significance was defined by a P-value smaller than 0.05. Analysis were only performed between groups that had at least three independent replicates (N=3). Results were expressed as average $\pm$ SD. Graphs were prepared with GraphPad Prism (v6, GraphPad Software, USA).

4. Development of a novel haematopoietic *in-silico* mathematical model

Chapter Summary

This chapter proposed a compartmentalized model of haematopoiesis into the three main myeloid lineages: erythroid, megakaryocytic and granulocytic. Since dyserythropoiesis is an MDS hallmark, the initial focus of this work was to study normal and abnormal erythropoiesis, featuring healthy-diseased cell competition and cytokine effects across progenitors in erythroid lineage progenitors.

In-house MATLAB scripts were created (Appendix IV) combining the gO:MATLAB basic code interface and MATLAB's graphical capabilities to produce 3D surface plots for the maximum nominal output value, given two inputs (time-independent) as well as time-course videos of the aforementioned plots (videos not shown). Fine-tuning of MATLAB scripts (reducing the number of sampling points by 10-fold) allowed for a 18-fold reduction in total simulation time – for 2500 simulations, total simulation time decreased from 370 minutes to around 20 minutes without any noticeable loss in resolution. Global Sensitivity Analysis (GSA) was performed as a proof-of-concept using GUI-HDMR and sensitivity indices are shown as time-dependant.

4.1. Introduction

As previously discussed in Chapter 1, MDS is a complex haematological malignancy with cytopenias in one or multiple haematopoietic lineages as one of the major disease hallmarks. Excessive apoptosis, cell dysplasia (i.e. abnormal shape and functionality) and high numbers of haematopoietic progenitor cells (blasts) are present in the BM and PB. MDS cells are characterized by the presence of one or more genetic mutations and/or karyotypic alterations. Overall, one can characterize MDS as a pre-leukemic state, despite the fact that only about 30-40% of all patients end up evolving to Acute Myeloid Leukemia (AML). The remaining 60-70% pass away due to increasingly severe and aggravated cytopenias.

The main assumption of the model is that healthy and diseased haematopoiesis co-occur in MDS patients (Walenda T 2014). MDS patients have to retain some form of healthy haematopoiesis to a degree, as only a small fraction of their terminally-differentiated cells are dysplastic or exhibit mutational defects. The diseased haematopoietic process competes with the healthy one for space and nutrients and ultimately inhibiting healthy cell growth. Other relevant assumptions are detailed in Appendix I.

With respect to the healthy haematopoietic cell population:

1. **Healthy hematopoiesis *in-vivo* is a well-defined and balanced system.**

Deterministic models are composed by a set of differential equations and describe the state of a system over a given time without random phenomena (Whichard ZL 2010). Perturbations in this time-dependent definition of the system can be inserted, but their mechanism needs to be known. As a well-defined and step-wise process, one can look at haematopoiesis as a complex intricate network of small production plants (hematopoietic lineages) that are connected by pipeline tubes (signaling molecules and cytokines). In this production plant, a perfect balance between input/demand and output/supply is necessary (daily production of terminally differentiated cells) and the

inter-dependence between production plants is evident (e.g. relatively constant *in-vivo* myeloid-to-lymphoid ratio).

2. **Healthy haematopoiesis is an intricate net of interdependent process that are not randomly regulated.** Even though not all haematopoietic differentiation pathways have unequivocally been described in the literature, it is widely accepted that haematopoiesis is a complex combination of signaling cascades, tight regulation by cytokines and cell concentrations, which exhibit different phenotypes and gene expression levels. Ideally, one would be able to break down haematopoiesis and its process into simple molecular, step-wise reactions, but that would be very time consuming and over-descriptive, as a simpler model might have similar predictive capabilities and functionality than a fully-descriptive, detailed model at a fraction of the computational cost.

Therefore, if one considers a simple approach where haematopoiesis is a production plant of terminally differentiated cells from haematopoietic stem cells, it is possible to breakdown its processes by a lineage-specific approach. The work herein described accounts for the development of a mathematical model describing erythropoiesis as a case-study, initial approach. Upon validating the combined *in-vitro in-silico* methodology, it will be expanded to other blood lineages. At the final stage, the final ensemble model will represent haematopoiesis as a whole process by linking the erythroid, megakaryocytic and granulocytic models.

4.1.1. Model Development

The model considers cytokines and their effect on the different erythroid lineage cells, as well as three main cellular processes: proliferation, differentiation and death. Some HSPCs also undergo assymetric division, where a cell divides into two cells, one of the same “stemness” and another one primed for differentiation. Notably, this is the mechanism that enables the *in-*

vivo maintenance of stem cell populations (López-Gómez S 2014). Being hard to track experimentally, asymmetric cell division is not considered in the model.

A proliferating cell is considered to give rise to two identical daughter cells, retaining the characteristics of the mother cell. A differentiating cell is considered to progress to a further differentiated state, without cell number expansion. Mathematically, this is implemented by creating several sub-compartments, denoted by the letter i , accounting for cells with an increasing degree of cellular maturity. Each sub-compartment hosts one and only one cellular population and cells migrate from sub-compartment i to sub-compartment $i+1$ (differentiation) or divide in sub-compartment i and remain in sub-compartment i (proliferation):

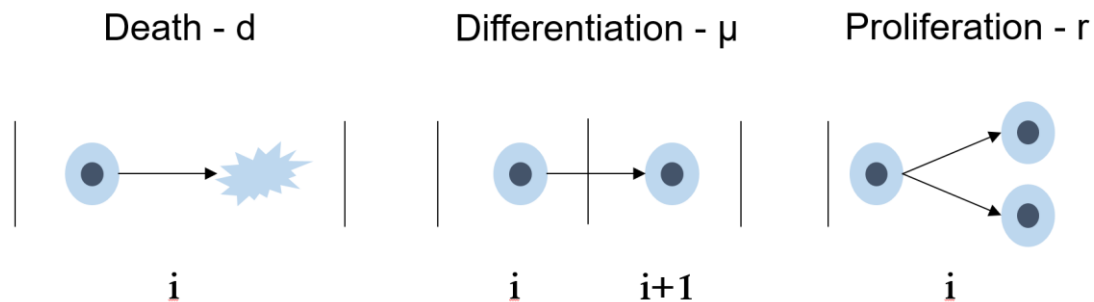


Figure 10 – A schematic of cell death, differentiation and proliferation mechanisms.

Common to every haematopoietic blood lineage represented in the model (the so-called sub models) are **three** different macro-compartments, where cells have different characteristics according to their phenotype, shape, functionality and proliferation capabilities.

The first compartment is the **Progenitor** cell (P) macro-compartment, where more immature and lineage-unspecific forms of haematopoietic stem and progenitor cells reside. Generically, this macro-compartment is composed by HSCs and progenitor cells such as Multipotent Progenitors (MPP). As progenitor cells mature, they become lineage-specific, giving rise to a new macro-compartment which includes the **Committed** progenitor cells (C). Biologically, the commitment of cells to a specific lineage and the differentiation towards more mature cells

yields a higher proliferation rate (Young, et al. 1996 and M. G. Kirouac DC 2009). The final macro-compartment is where **Terminally** differentiated cells (T) reside. White blood cells, platelets and red blood cells reside in this compartment. As a good rule of thumb, these are the cell populations observed in a common blood test.

For clarity, we refer to macro-compartment as the progenitor, committed and terminally differentiated cellular compartments. We also mention sub-compartments, which host one and only one cell type. Finally, blood lineages are the three main myeloid cellular populations, such as the erythroid lineage (Red Blood Cells), the granulocytic lineage (White Blood Cells) and the megakaryocytic lineage (Platelets).

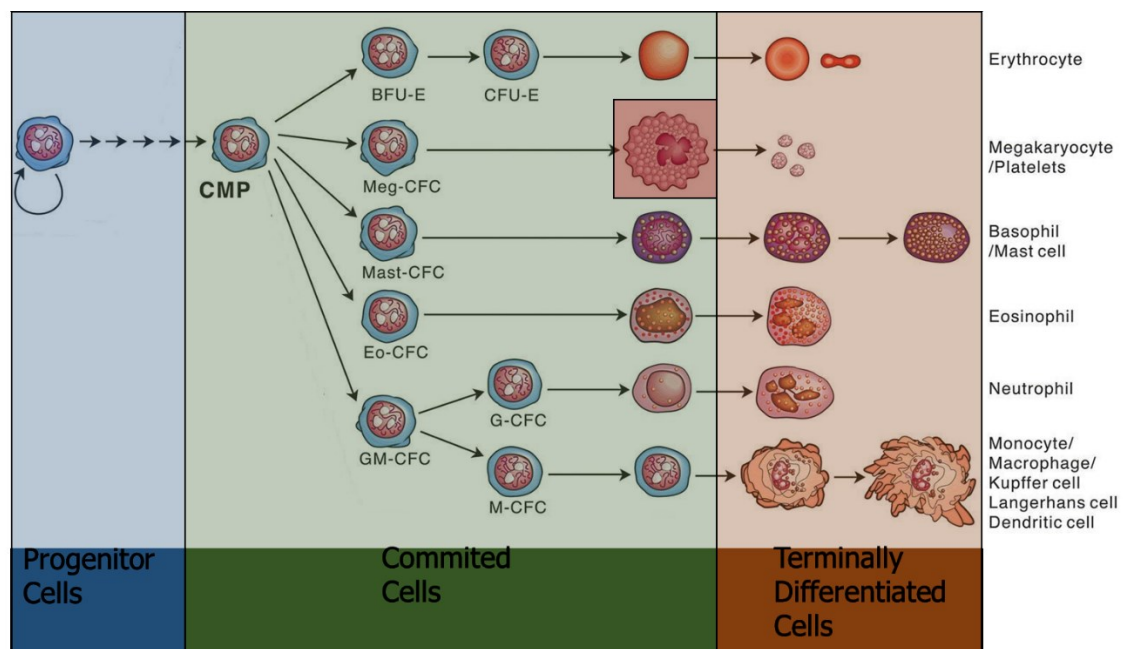


Figure 11 – The proposed compartmentalization of haematopoiesis into the progenitor cell macro-compartment (in blue), the committed cell macro-compartment (in green) and the terminally differentiated cell macro-compartment (in red).

It is known that HSCs and immature progenitors privilege proliferation over differentiation towards more mature cell forms, except if prompted to do so by external stimuli such as cytokine concentrations (Kondo, et al. 2003) or selective pressure due to the lack of terminally differentiated cells (Obeyesekere, et al. 2004). We therefore propose that:

1. A blood lineage-independent number of sub-compartments will be needed for the Progenitor cell compartment, accounting for each and every single one of the immature

haematopoietic progenitors. In essence, since all committed cells originate from HSCs and MPPs, there will be a defined number of sub-compartments within the progenitor cell macro-compartment, common to each blood lineage.

2. The number of sub-compartments within the Committed and Terminally Differentiated macro-compartments will be lineage-specific. These will account for all the intermediate cellular stages between the most immature, lineage-committed progenitor cell and a terminally-differentiated cell. The total number of sub-compartments to be considered will differ between each blood lineage as a function of the total number of cellular intermediates in that specific lineage.

4.1.2. Modelling Healthy Erythropoiesis – A Case-Study

Cells undergoing erythroid differentiation need to evolve through a series of maturation stages. Erythropoiesis starts with the differentiation of an HSC and finishes with the enucleation and terminal maturation of Orthochromatophilic Erythroblasts (OrthoE) into RBCs. All the identifiable erythroid cell intermediates are described in the figure below:

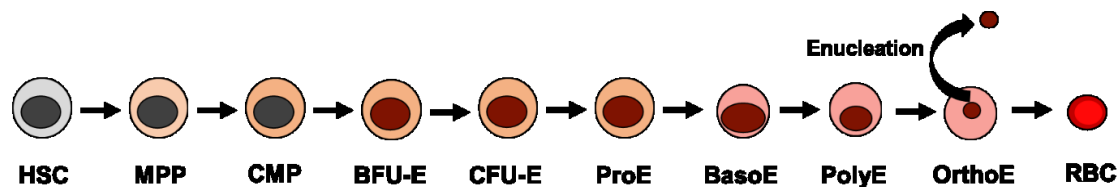


Figure 12 - Erythroid differentiation and the different identifiable intermediate stages. Legend: **HSC** – Haematopoietic Stem Cell; **MPP** – Multipotent Progenitor; **CMP** – Common Myeloid Progenitor; **BFU-E** – Erythroid burst-forming unit; **CFU-E** – Erythroid colony-forming unit; **ProE** – Proerythroblast; **BasoE** – Basophilic erythroblast; **PolyE** – Polychromatophilic erythroblast; **OrthoE** – Orthochromatophilic erythroblast; **RBC** – Red blood cell.

Using mass-action kinetics to macroscopically describe the cellular processes a given cell population can undergo, mathematical expressions were derived accordingly for each cellular process (Figure 10), for all the identifiable erythroid intermediate stages (Figure 12).

The mathematical nomenclature of the rates used to describe the three cell processes considered are summarized in Table 6.

Table 6 – Nomenclature used for the mathematical description of a variety of cell processes (Figure 10) happening in different cell compartments, for a given generic cell type (A).

Process	Expression	Description
Proliferation	r_A	Proliferation rate, cell type A
Differentiation	μ_A	Differentiation rate, cell type A
Death	d_A	Death rate, cell type A

The cellular processes represented in Figure 10 are mutually exclusive – one cell can only perform one of those cellular processes at a time. Also, according to their degree of maturation and the external stimuli provided, some cells will prefer to proliferate, differentiate or die. We then define three cell fractions which account for the number of cells, relative to the total number of cells, undergoing a given process at a given time. These fractions are summarized in Table 7.

Table 7 – Nomenclature used for the mathematical description of the cellular fractions undergoing different cell processes (Figure 10) happening in different cell compartments, for a given generic cell type (A).

Process	Expression	Description
Proliferation	a_{r_A}	Proliferation fraction, cell type A
Differentiation	a_{μ_A}	Differentiation fraction, cell type A
Death	a_{d_A}	Death fraction, cell type A

This model considers that cells are only able to proliferate, differentiate or die. Since every cell will have to do one of these three processes, it is trivial to conclude that:

$$a_{r_A} + a_{\mu_A} + a_{d_A} = 1 \quad (\text{Eq.1})$$

These cell fractions depend on the external stimuli (cytokines) provided to the cells. For convenience, we determine that cytokines influence proliferation and differentiation cell fractions, while the death fraction depends on the other two fractions, as per Eq. 1. In future iterations and expansions of this work stimuli that drive a cell to die (e.g. excessive apoptotic signals) can be added to allow for a better biological representation. Mathematically, the fraction variation with a stimulus P is given by Hill Kinetics (Eq.2), which have been widely used in mathematical modelling of biological systems (Weiss 1997 and Santillan 2008). The result of a stimulus can either be positive (**s**) or negative (**si**), resulting in promoting (Eq. 3) or inhibitory (Eq. 4) effects on cells.

$$a_A(S) = a_A \max \times s(P) \quad (\text{Eq.2})$$

$$s(P) = \frac{P^n}{K_a^n + P^n} \quad (\text{Eq.3})$$

$$si(P) = \frac{K_a^n}{K_a^n + P^n} \quad (\text{Eq.4})$$

K_a and n are the half-saturation constant and Hill coefficient, respectively. When $P = K_a$, the signal will have half of its maximum intensity. On the other hand, n defines the steepness of the response and the cooperativity of the substrate-active site binding (Hill 1910). When $n = 1$, this equation becomes the well-known Michaelis-Menten enzyme-substrate kinetic equation. Cytokine concentrations are assumed to be time-invariant, and therefore their signal is constant over time, depending only on the values given to K_a and n . In future iterations of this work, degradation of cytokines (thermal or other) can also be incorporated accordingly.

Using mass-action kinetics, we can take a deeper look into one of the erythroid cell populations and perform a mass balance to a generic cell state, with its cell number given by A :

$$Acumulation = \frac{dA}{dt} \quad (\text{Eq.5})$$

$$Differentiation = a_{\mu_A} \times \mu_A \times A \quad (\text{Eq.6})$$

$$Proliferation = a_{r_A} \times r_A \times A \quad (\text{Eq.7})$$

$$Death = a_{d_A} \times d_A \times A \quad (\text{Eq.8})$$

An example using the ProE, BasoE and PolyE cell populations is shown on Figure 13, where the BasoE cells can either **proliferate**, **die** or **differentiate** (mature) into PolyE cells, with the additional effect of ProE cells **differentiating** into BasoE, which increases BasoE cell numbers.

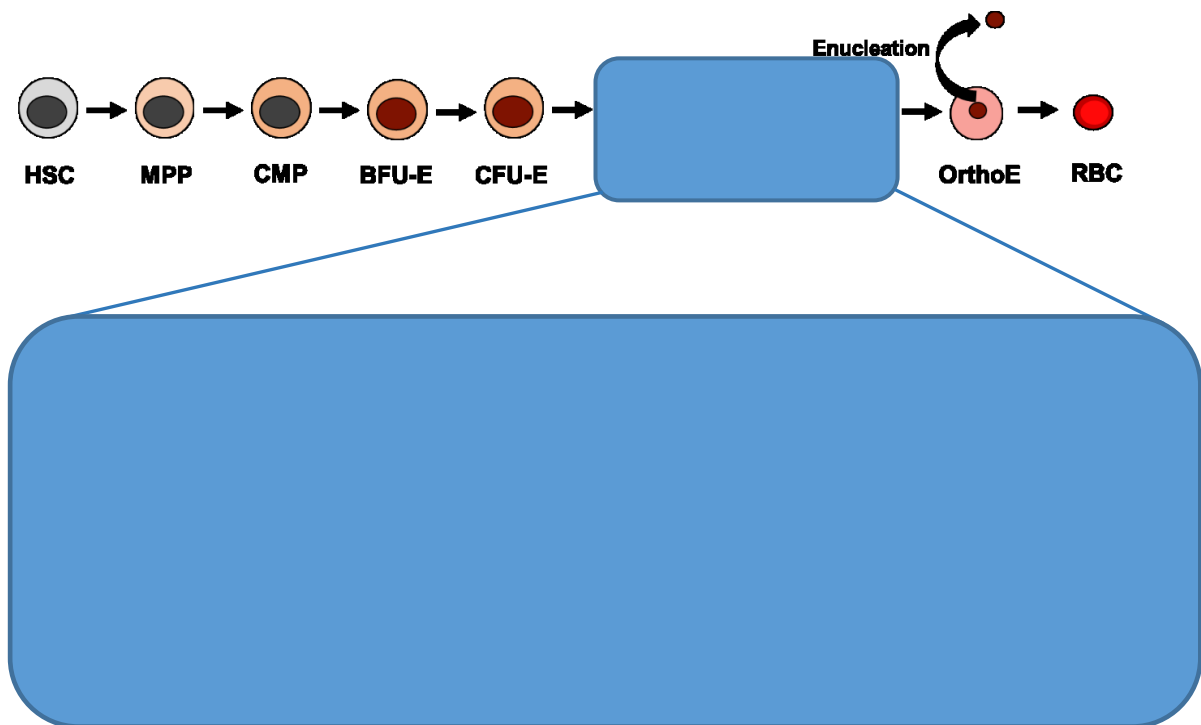


Figure 13 – An in-depth look at the inlet and outlet processes occurring at the BasoE cell population level.

We can write the following generic mass-balance equation:

$$Acumulation = Differentiation + Proliferation - Maturation - Death \quad (Eq.9)$$

Performing a mass balance to the BasoE cell population and identifying with specific colors the relevant processes yields the below expression:

$$\frac{dBasoE}{dt} = a\mu_{ProE} \cdot \mu_{ProE} \cdot ProE + 2ar_{BasoE} \cdot r_{BasoE} \cdot BasoE - ad_{BasoE} \cdot d_{BasoE} \cdot BasoE - a\mu_{BasoE} \cdot \mu_{BasoE} \cdot BasoE$$

The **differentiation of ProE** into BasoE increases BasoE cell numbers, whereas BasoE cells themselves **differentiate into PolyE**, which decreases BasoE cell numbers.

For each cell type in the erythroid lineage, a similar mass balance is created. Values for death, differentiation and proliferation rates, as well as cell fractions, will differ between each cell type. This mathematical construction ensures that all the cell compartments and populations are interconnected. It is relevant to note that the model does not consider any back-flux of cells from a more mature state to a less mature state, since cell differentiation is a one-way, irreversible process.

It is relatively hard to obtain the relevant parameter values due to the lack of information in the literature. However, the literature review described in depth in Chapter 1 and Appendix I allows for some educated guesses and assumptions about the variation of some parameters throughout the erythroid cell lineage:

1. Proliferation rate increases with the degree of maturity (Möbest D 1999);
2. Erythroblast differentiation starts with an increase in proliferation (Akker EVD 2010);
3. We assume that differentiation rates are higher the more mature the cell is; this assumption is based on the fact that erythroid-committed progenitors will, under the correct conditions, always aim for terminal differentiation (Huang P 2020);
4. The generation of BFU-E/CFU-E is not EPO/EPOR dependent; however, EPO is required for further erythroid differentiation (Wu H 1995);

5. SCF acts on hematopoietic stem cells and multipotent haematopoietic progenitors (Leary AG 1992), stopping its effect after the CFU-E stage, where c-kit expression decays rapidly (Testa U 1996);
6. IL3 increases BFU-E numbers while having an inhibitory effect on CFU-E and its erythroid progeny (Leonard JP 1988);
7. When using either SCF or IL3 alone, the effects on CD34⁺ expansion were comparable (Brugger W 1993);
8. The half-saturation constants for EPO/EPOR, SCF/c-kit and IL3/IL3RA were determined to be 350 mU/mL (Becker V 2010), 10 ng/mL (Dai CH 1994) and 49.11 ng/mL (Itoh N 1990), respectively. This will be used as the half-saturation constants for Hill Kinetics, whereas *n* is assumed to be 2 for all cytokines, based on the two described EPO binding sites to EPOR (Syed RS 1998);
9. HSC and lineage-unspecific progenitors prefer to proliferate instead of differentiate (Cheshier SH 1999);
10. All the erythroid progenitors are proliferative up to the OrthoE stage (Spivak 2005).

Table 8 summarizes the assumed values for the relevant parameters, according to the assumptions stated above. These values are educated assumptions which serve as a first set of parameters to test model structure and functionality. For model-based applications, these rates will need to be, at a later stage, estimated very accurately using well-designed experiments.

Table 8 - Rates of proliferation, differentiation and death according to each cell type.

<u>Cell Type</u>	<u>Proliferation</u> <i>r</i>	<u>Differentiation</u> <i>μ</i>	<u>Death</u> <i>d</i>
HSC	0.01	0.025	0.01
MPP	0.01	0.025	0.01
CMP	0.01	0.025	0.01
BFU-E	0.05	0.17	0.04
CFU-E	0.025	0.18	0.04
ProE	0.1	0.19	0.02
BasoE	0.1	0.2	0.02
PolyE	0.1	0.22	0.02
OrthoE	0.0	0.4	0.01
RBC	--	--	0.001

The proposed maximum cell fractions and the effects of SCF, IL3 and EPO on the different cell populations are summarized in Tables 9 and 10 respectively. Positive, activating effects are denoted by (+) whereas negative, inhibitory effects are denoted by (-). 0 represents a null effect.

Table 9 – Maximum proliferation and differentiation fractions of different cell types.

<u>Cell Type</u>	<u>Proliferation Fraction</u> $\max(ar)$	<u>Differentiation Fraction</u> $\max(a\mu)$	<u>Assumption Number</u>
HSC	0.5	0.45	9
MPP	0.5	0.45	9
CMP	0.5	0.45	9
BFU-E	0.6	0.35	2, 10
CFU-E	0.35	0.6	2, 10
ProE	0.35	0.6	2, 10
BasoE	0.35	0.6	2, 10
PolyE	0.35	0.6	2, 10
OrthoE	0	0.95	2, 10
RBC	--	--	--

Table 10 – Effects of SCF, IL3 and EPO on the proliferation and differentiation fractions of different cell types.

<u>Cell Type</u>	<u>Proliferation Fraction</u> ar SCF / IL3 / EPO	<u>Differentiation Fraction</u> $a\mu$ SCF / IL3 / EPO	<u>Assumption Number</u>
HSC	+ / + / 0	0 / 0 / 0	4, 5
MPP	+ / + / 0	0 / 0 / 0	4, 5
CMP	+ / + / 0	0 / 0 / 0	4, 5
BFU-E	+ / + / 0	0 / 0 / 0	4, 5, 6
CFU-E	+ / - / 0	0 / - / +	4, 5, 6
ProE	0 / - / +	0 / - / +	1, 4
BasoE	0 / - / +	0 / - / +	1, 4
PolyE	0 / - / +	0 / - / +	1, 4
OrthoE	0 / - / +	0 / - / +	1, 4
RBC	--	--	--

The cytokine effects on all the erythroid progenitors are easier to understand when represented visually, as it allows researchers to understand cytokine dynamics within erythropoiesis. Assuming a positive, null and negative effect have values of +1, 0 and -1 respectively, we obtain the following graphs that show the combined effect of the cytokines in the erythroid progeny.



Figure 14 – Effects of SCF, IL3 and EPO on cell proliferation (top) and differentiation (bottom). Positive, null and negative effects are represented by values of +1, 0 and -1 respectively.

4.1.3. Modelling Diseased Erythropoiesis in MDS

Modelling diseased erythropoiesis, or dyserythropoiesis, was done by creating a parallel model structurally similar to the healthy one. This diseased haematopoietic population interacts with its healthy counterpart by inhibiting its normal growth. This is achieved by implementing a new inhibitory Hill function dependent on the total MDS cell number (Eq. 5) whereas the “half-saturation constant” K_a is assumed to be 1×10^4 cells – this means that the inhibitory signal will have an intensity of half the maximum if MDS cell numbers reach those levels.

The response of MDS cells to cytokines is known to be relatively lower than their healthy counterparts (Greenberg 2006). Also, it is assumed that MDS cells will proliferate faster and differentiate slower, when compared to the healthy hematopoietic population. This would contribute to explain disease evolution *in-vivo*, as the dysplastic cell populations slowly but surely outcompete the healthy ones. Death rates are also higher than the healthy ones due to excessive apoptosis, an MDS hallmark (Walenda T 2014)

Tables 11 and 12 summarize the assumed values for proliferation, differentiation and death rates as well as the maximum cell fractions, respectively.

Table 11 - Rates of proliferation, differentiation and death according to each MDS cell type.

<u>Cell Type</u>	<u>Proliferation</u> r	<u>Differentiation</u> μ	<u>Death</u> d
HSC	0.012	0.012	0.05
MPP	0.012	0.014	0.05
CMP	0.012	0.014	0.05
BFU-E	0.024	0.02	0.04
CFU-E	0.02	0.02	0.04
ProE	0.02	0.03	0.02
BasoE	0.02	0.04	0.02
PolyE	0.02	0.06	0.02
OrthoE	0.02	0.01	0.01
RBC	--	--	0.01

Table 12 - Maximum proliferation and differentiation fractions of different MDS cell types.

<u>Cell Type</u>	<u>Proliferation Fraction</u> $\max(ar)$	<u>Differentiation Fraction</u> $\max(a\mu)$
HSC	0.5	0.45
MPP	0.5	0.45
CMP	0.5	0.45
BFU-E	0.35	0.6
CFU-E	0.3	0.65
ProE	0.1	0.45
BasoE	0.1	0.45
PolyE	0.1	0.45
OrthoE	0.1	0.45
RBC	--	--

4.2. Results and Discussion

Simulations were run to test the influence of the three cytokines studied on the healthy and diseased erythroid populations. At the model development stage, one aims to understand if the model is able to correctly capture the relevant biological problem, testing its responsiveness to scenarios where the outcome is somewhat predicted or expected. The mathematical model and all the simulations were executed in gPROMS ModelBuilder 4.2.0. To perform multiple simulations with varying parameters, the gO:Matlab interface was used, enabling gPROMS to communicate with Matlab R2014a. The model has 110 equations (89 algebraic and 21 differential) as well as 100 variables and 109 parameters.

For the healthy erythroid population, we have assessed the cytokine effects on early (BFU-E, CFU-E) and committed (ProE to OrthoE) erythroid progenitors. Simultaneously, we are trying to understand how affecting early erythroid progenitors might influence RBC numbers. For the diseased (MDS) erythroid population, we have implemented a competition-based model where the dysplastic, diseased population inhibits the growth of its healthy counterpart. We aimed to understand how a dysplastic (diseased) cell population affects the behaviour and growth of healthy cells.

EPO, SCF and IL3 were studied in concentrations ranging from 0-70% of their maximum effect on cells, resulting in signal intensities between 0 and 0.7, respectively. The maximum effect was derived from relevant literature (Appendix I). Each scenario (Table 13) is composed of 2,500 independent model runs. The concentrations **in bold** were kept constant across all of the 2,500 simulations, whereas the other two cytokine concentrations were varied between the lower and upper bounds of their respective ranges by adding incremental amounts of 1/50 of the maximum range value, ensuring that all the possible combinations were tested.

Table 13 – Summary of the experimental conditions used for the *in silico* simulations on the healthy erythroid population.

<u>Scenario</u> <u>Number</u>	<u>SCF</u> ng/mL	<u>IL3</u> ng/mL	<u>EPO</u> mU/mL
1	0-20	0-100	0
2	0-20	0-100	500
3	0-50	0-250	500
4	0-20	0	0-500
5	0-20	50	0-500
6	0	0-100	0-500
7	10	0-100	0-500
8	50	0-100	0-500

Results are shown in 3D surface plots generated using the built-in Matlab graphic capabilities, where the x and y axes show the different cytokine concentrations used and the z axis shows the maximum cell number, for each cell type, achieved throughout the duration of each model run. It is worth noting that cytokine concentrations do not decay as a function of the model run time and are, therefore, assumed to be constant throughout each model run. The heat-map colouring scale is specific to each surface plot and helps identifying local and global minimum and maximum z axis values.

EPO is necessary for CFU-E progeny to be generated

BFU-E and CFU-E generation is EPO-independent; however EPO is needed for terminal erythroid differentiation all the way to RBC. Biologically, there is an “EPO-checkpoint” at the CFU-E level which prevents further erythroid differentiation in the absence of EPO. This has been correctly incorporated in the model – Figure 15 shows that neither ProE nor RBC are generated in the absence of EPO (bottom panel), despite CFU-E numbers reaching close to 3.5×10^6 cells (SCF = 20 ng/mL, IL3 = 100 ng/mL).

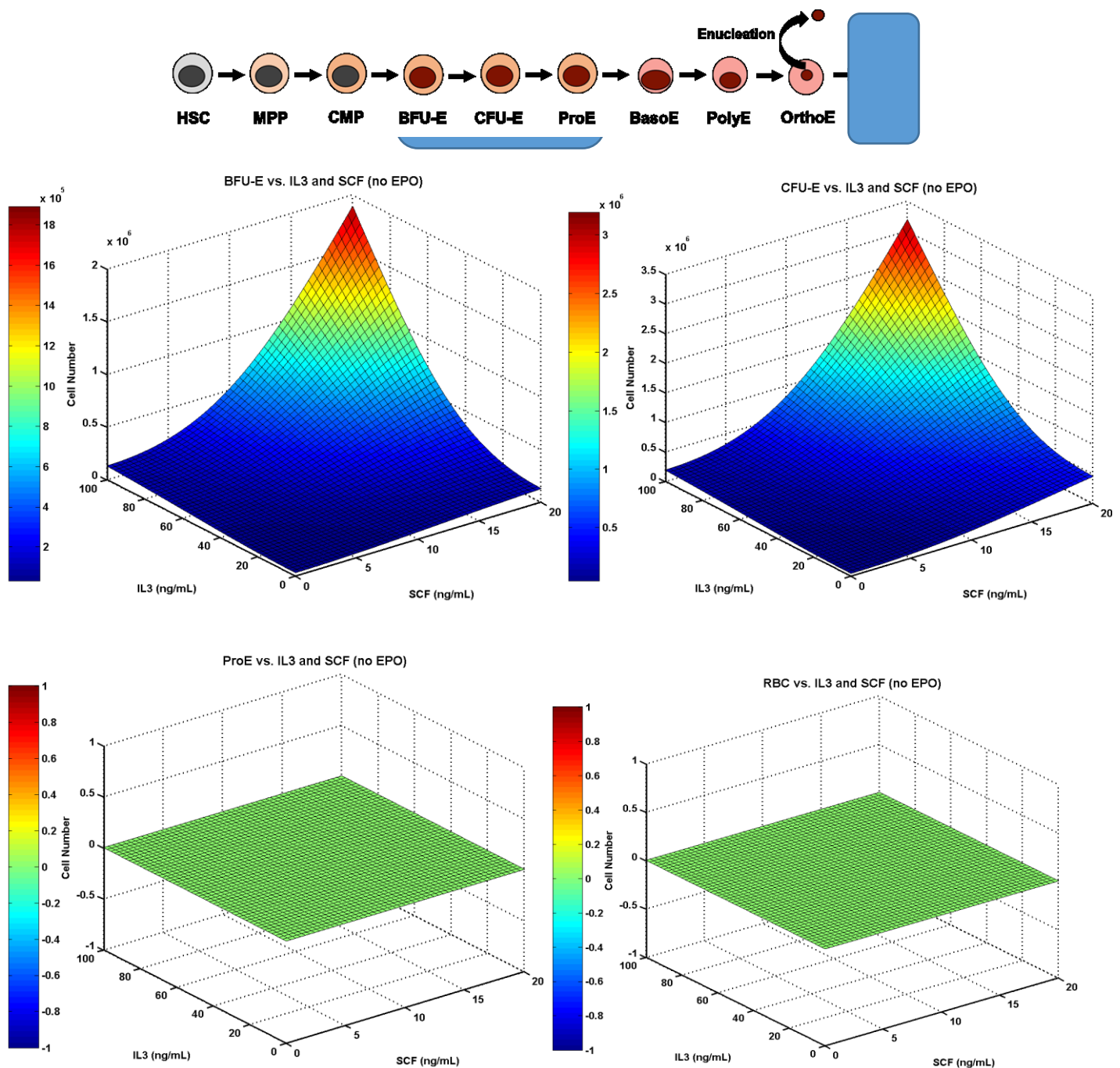


Figure 15 - Effects of SCF and IL3 on the erythroid progeny, without any EPO added. The formation of BFU-E and CFU-E is independent of EPO, but further erythroid differentiation is impaired if EPO is not present.

IL3 increases BFU-E and CFU-E cell numbers, but is a strong inhibitor of their progeny

IL3 yields a positive proliferative stimulus to BFU-E cells and inhibits erythroid progenitor growth and differentiation. Biologically, this inhibitory effect is only seen in the CFU-E progeny. More mature cells will be more sensitive to IL3 due to a cascade effect – the more mature the cell is, the lower the concentration of IL3 necessary for it to become inhibitory, at which point the total number of cells decays. In the bottom panel, we can see that IL3 becomes inhibitory for BasoE in concentrations above 80 ng/mL, where concentrations above 40 ng/mL have an inhibitory effect in the total RBC number.

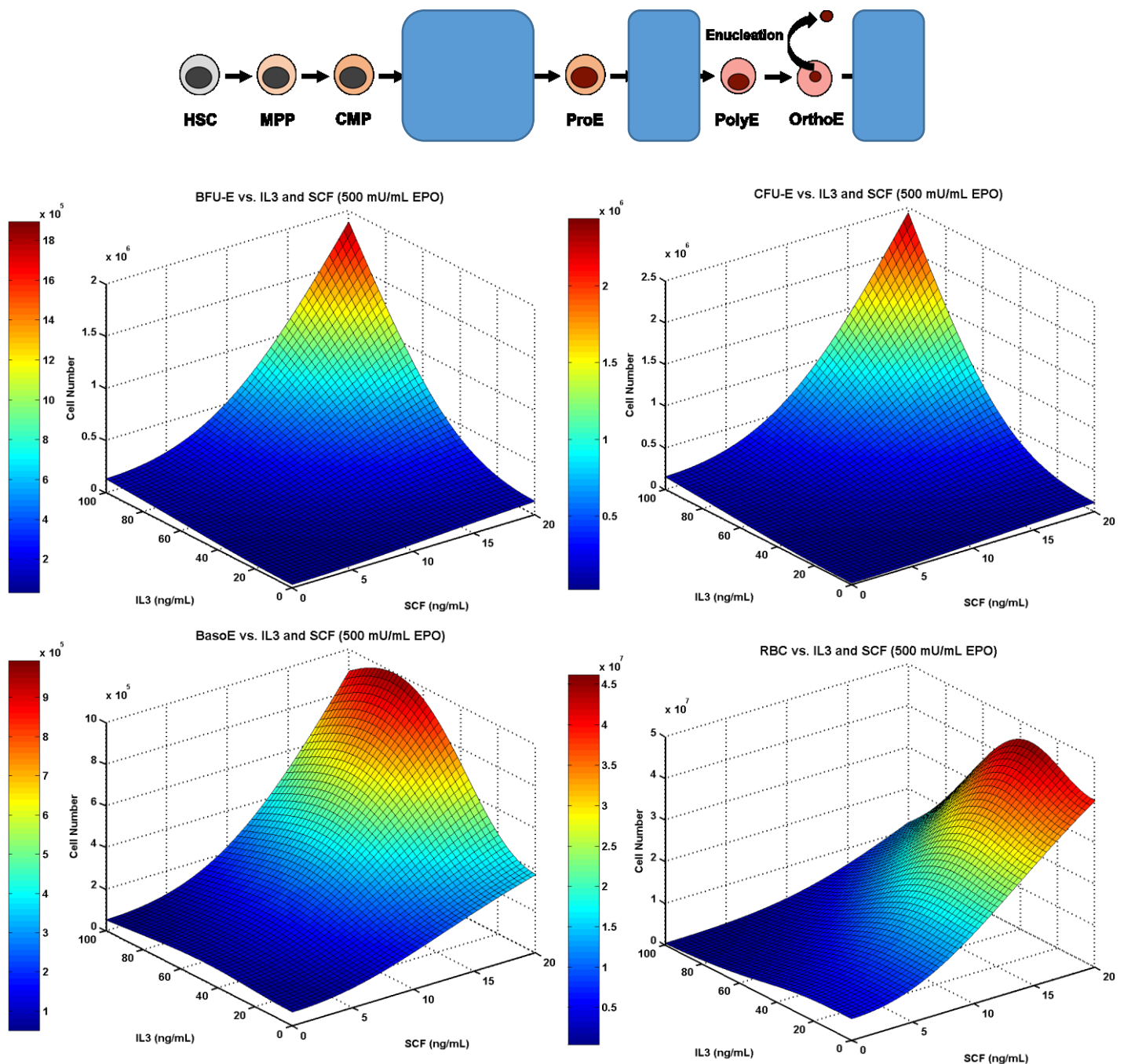


Figure 16 - Effects of SCF and IL3 on the erythroid progeny with EPO added. IL3's inhibitory effect is strongly seen in the erythroid progeny (BasoE and RBC, bottom panel). The more mature a cell is, the stronger the inhibitory effect, typically as a consequence of differentiation arrest of earlier, immature cell types.

A combination of EPO and SCF enhances erythroid progenitor numbers

Without IL3, adding EPO and SCF increases the cell numbers of BFU-E and its progeny, excluding RBC (Committed Cells), whereas when IL3 is added in non-inhibiting concentrations, its early-acting proliferative effects also augment the maximum number of Committed and RBC by a factor of 2-fold (Top vs Bottom panel, left-hand side) and 1.125-fold (Top vs Bottom panel, right-hand side) respectively, as shown in Figure 17. The increase in RBC numbers with IL3 at 50ng/mL is likely to be a direct consequence of having a 2-fold higher number of Committed Cells.

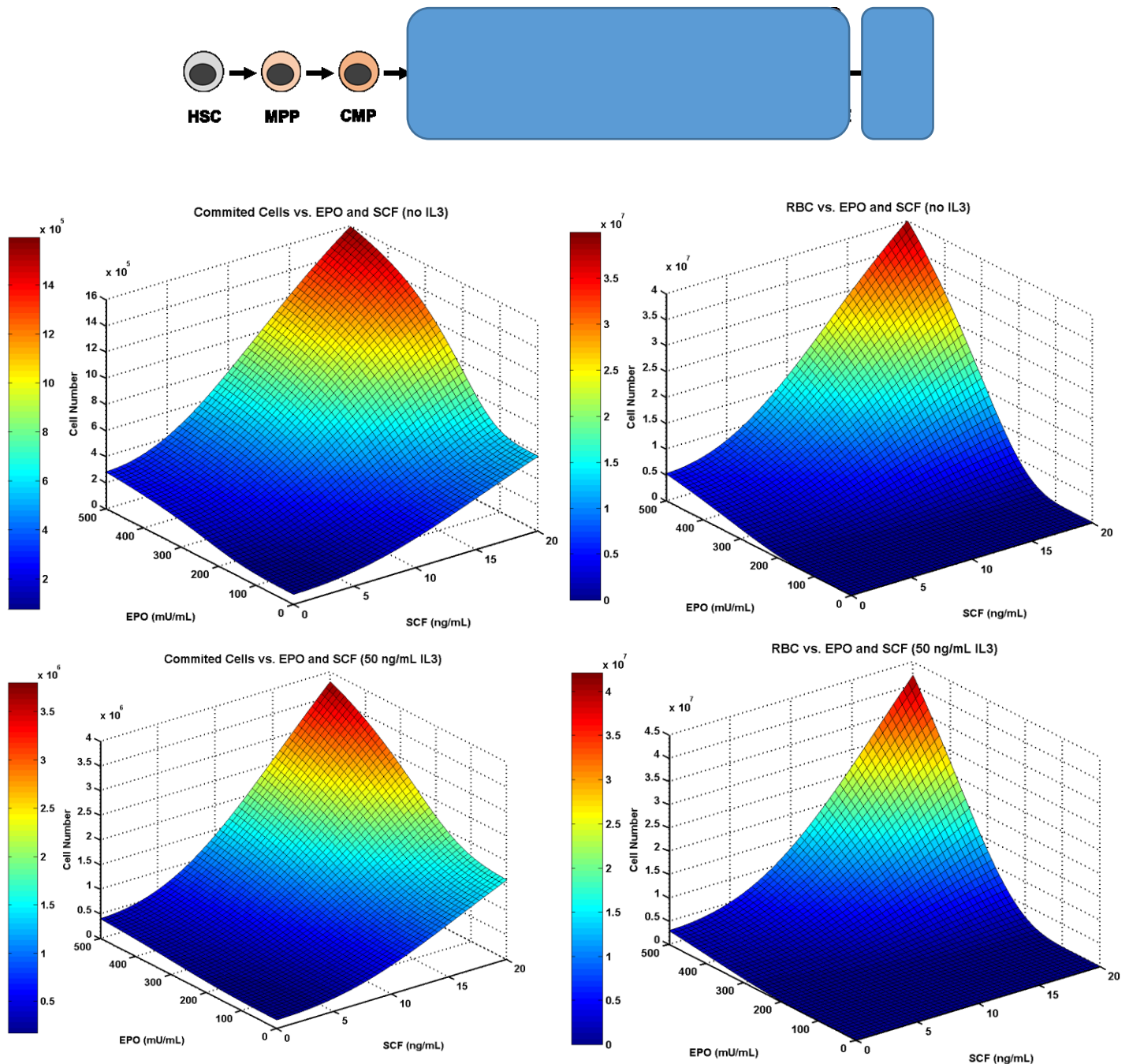


Figure 17 – Effects of SCF and EPO addition without (top panel) and with (bottom panel) IL3. The addition of IL3, provided under inhibitory concentrations, increases the numbers of both Committed Cells and RBC numbers.

Adding SCF to a combination of EPO and IL3 delays the inhibitory effect of IL3 and yields a higher RBC number

The addition of SCF to a cocktail of EPO and IL3 has a strong delaying effect on the inhibition caused by high concentrations of IL3. Adding 10 ng/mL SCF increases the IL3 inhibitory concentration and the maximum RBC numbers to 30 ng/mL and by 3.4-fold (when compared to the “No SCF” scenario) respectively. When adding 50 ng/mL SCF, the IL3 inhibitory concentration is now around 50 ng/mL and maximum RBC numbers increase by 19-fold versus the “No SCF” condition. This is another example of the multiple effects that different cytokine combinations can have in RBC numbers.

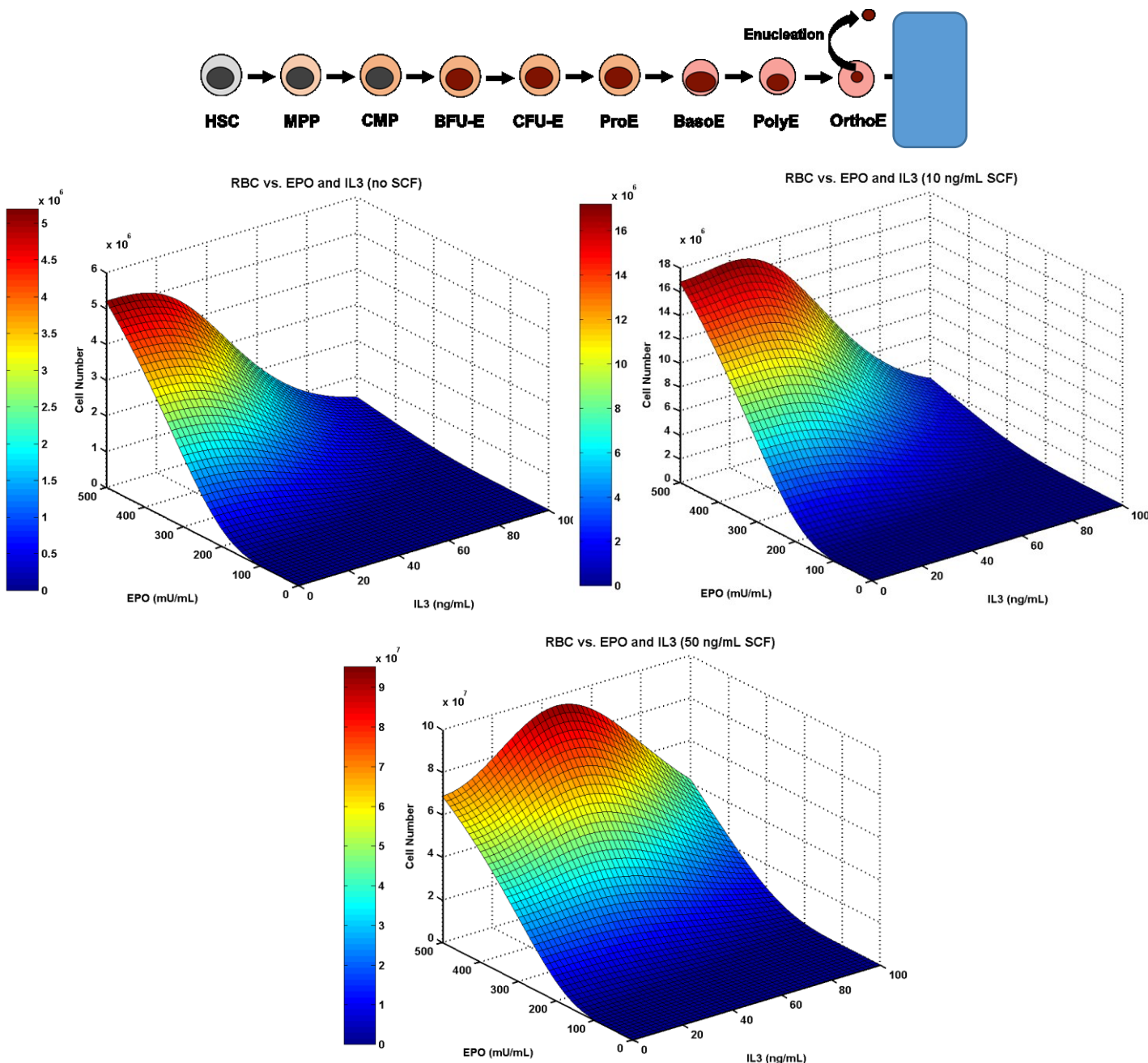


Figure 18 – Adding SCF to a cocktail of EPO and IL3 delays IL3’s inhibitory effect while considerably augmenting RBC numbers. The addition of no SCF (top-left), 10 ng/mL (top-right) and 50 ng/mL (bottom) of SCF yields a maximum RBC number of 5, 17 and 95 million cells, respectively.

Simulations were also performed after implementing a diseased MDS population, at the HSC level. These cells, despite being dysplastic, will differentiate and compete for nutrients and space with their healthy counterparts, inhibiting their growth. For each condition, 500 MDS HSC were added – corresponding to 0.025% of the initial total cell population – to assess their inhibitory effect on the healthy population.

When a dysplastic population is present, its inhibitory effects on the healthy population are translated by a Hill Kinetic equation (Eq. 4) with an initial value for the half-saturation constant set arbitrarily at 1×10^4 cells. This parameter will need to be experimentally determined and is crucial to correctly capture the healthy-diseased dynamics.

Table 14 shows a summary of all the cytokine values and ranges used for the diseased population study.

Table 14 - Summary of the experimental conditions used for the *in silico* simulations on the diseased erythroid population. Values in **bold** were kept constant.

<u>Simulation Number</u>	<u>SCF</u> ng/mL	<u>IL3</u> ng/mL	<u>EPO</u> mU/mL	<u>MDS HSC</u> Cell No.
9	0-20	0-100	0	500
10	0-20	0-100	500	500
11	0-50	0-250	500	500
12	0-20	0	0-500	500
13	0-20	50	0-500	500
14	0	0-100	0-500	500
15	10	0-100	0-500	500
16	50	0-100	0-500	500

The inhibitory effect of the MDS population is highly visible on the committed cell compartment

Graphics on the left describe Committed cell numbers in the absence of any dysplastic population (100% healthy), where the graphics on the right look just at the Healthy Committed cell numbers, in the presence of a dysplastic population. Comparisons are made left versus right and look at the maximum cell number on each graph.

For the maximum SCF and IL3 concentrations (Simulation 10, top panel) there is a 1.14-fold decrease in committed cell numbers due to the presence of a dysplastic population. If one increases SCF and IL3 concentrations to near-saturating values (Simulation 11, bottom panel, x and y axis maximum values now changed) there is a dramatic 43.3-fold decrease in the maximum committed cell numbers, due to overall bigger numbers of diseased cells.

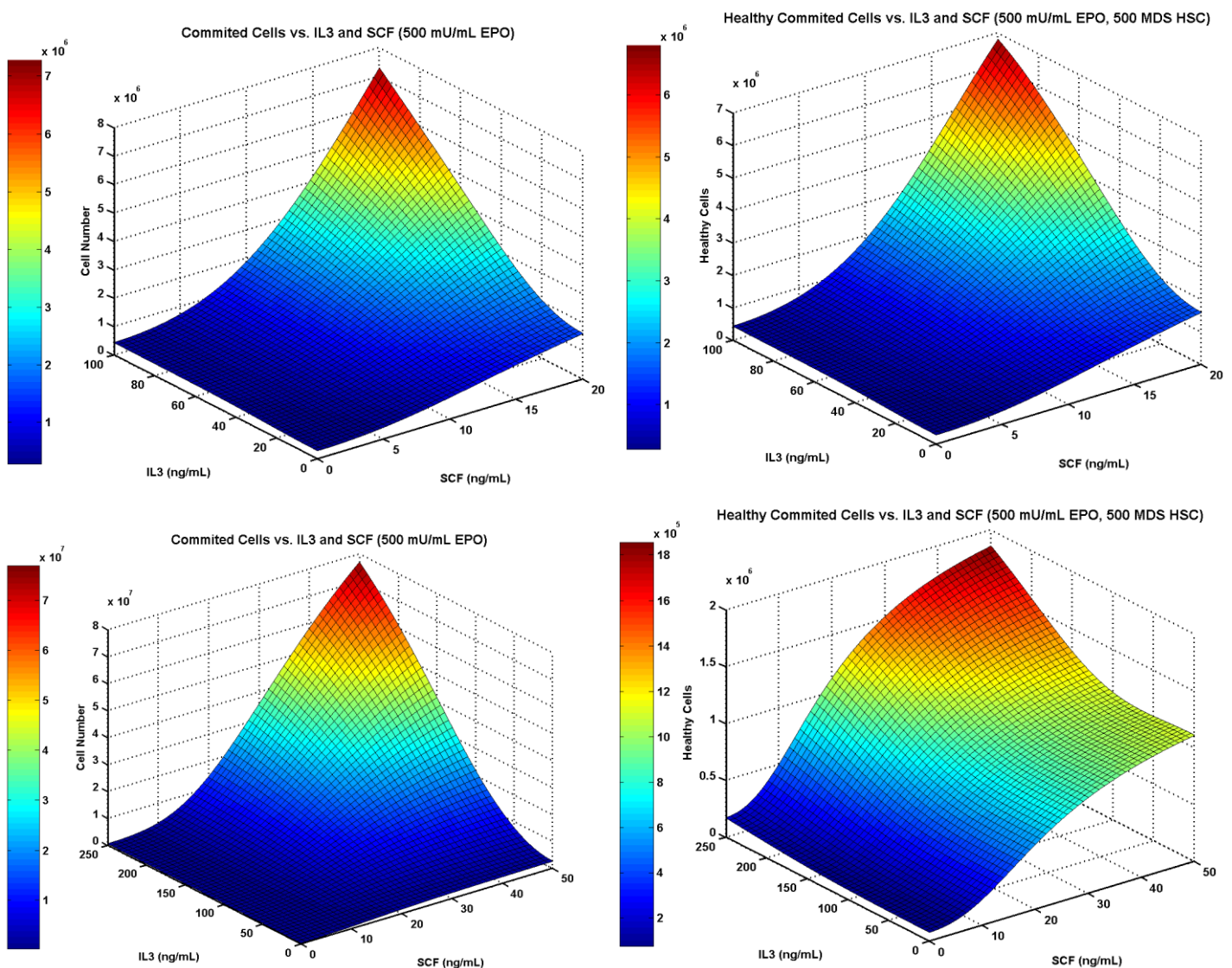


Figure 19 – The presence of a competing dysplastic population shows a direct correlation between cytokine levels and inhibitory effects.

High cytokine concentrations do not greatly affect maximum RBC numbers yet they become more sensitive to inhibitory cytokines like IL3

Across all the simulations there was no visible decrease in RBC numbers (Figure 20) due to the presence of MDS cells. This will have later to be adjusted, and we believe it to be strongly correlated with the chosen value for the MDS K_a inhibition of 1×10^4 cells. However, the presence of MDS cells made RBC numbers decay faster with lower IL3 concentrations, showing that the presence of these diseased populations may affect normal RBC generation by reducing progenitor levels by around 2.2-fold (top panel). This effect then cascades down to the RBC level by increasing their sensitivity to high concentrations of inhibitory cytokines like IL3, but without greatly affecting their maximum numbers (bottom panel).

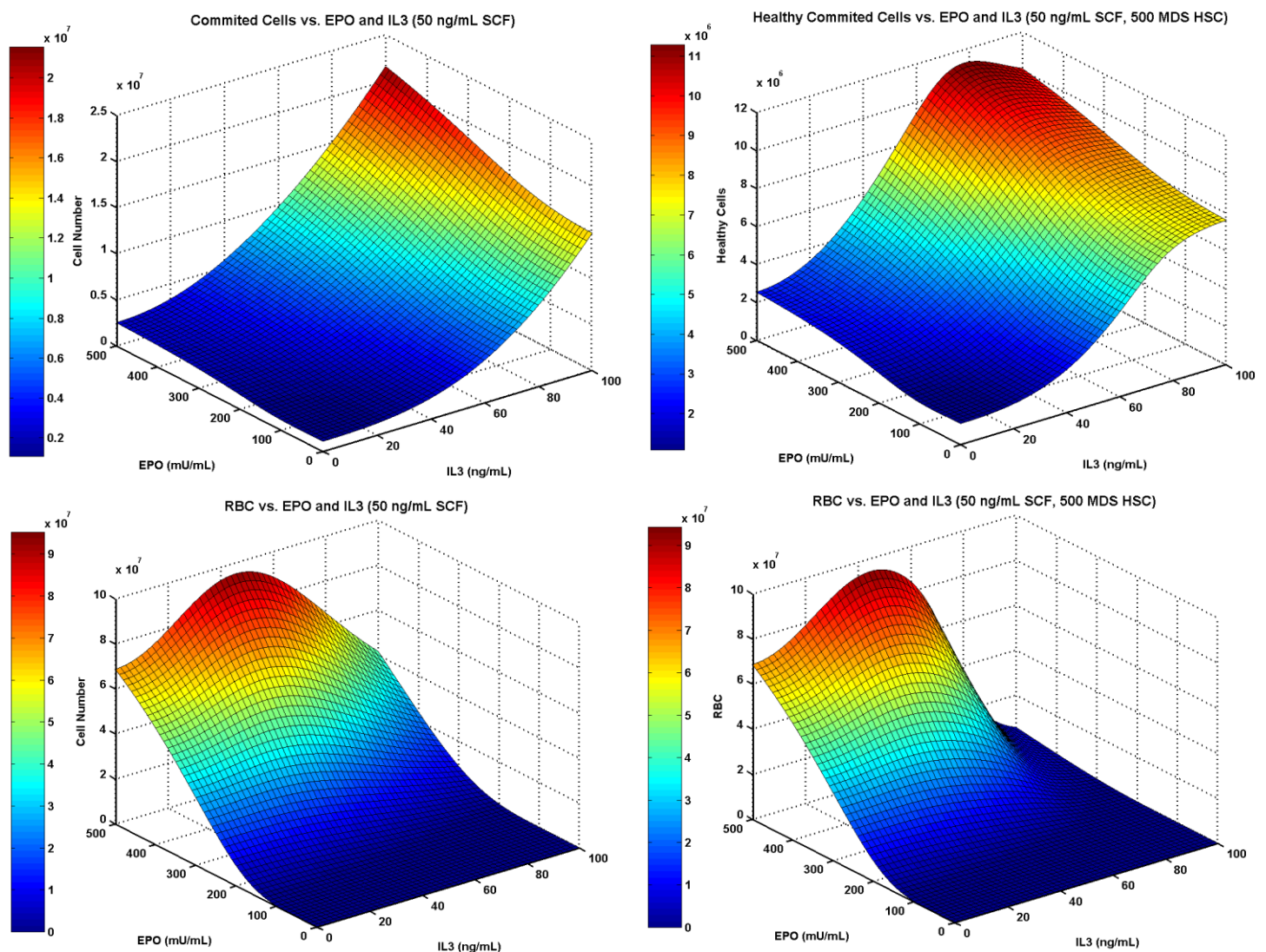


Figure 20 – Despite yielding a reduction of more than 2-fold of the progenitor cell numbers, the presence of MDS cells only seem to increase RBC sensitivity to higher concentrations of cytokines. However, these effects need further studies since the dependence on parameters like $MDS(K_a)$ are very high.

HSCs will differentiate spontaneously until the CFU-E stage in the absence of EPO (Figure 15). At that stage, there is a “check-point” and cells only differentiate further to the erythroid when EPO is present, and this is indeed the case *in-vivo*. IL3 is a known HSPC stimulant and has a positive effect on the proliferative capability of these cells. Above a certain concentration, IL3 becomes inhibitory, drastically reducing cell numbers. (Figure 16). Its inhibitory effect is however mitigated by the addition of SCF as it dramatically increases the number of Committed cells (Figure 18). Despite not affecting RBCs directly, greater concentrations of IL3 will yield higher RBC numbers, probably because of higher committed progenitor cell numbers (Figure 17). These simulations are in good agreement with the model’s underlying assumptions (Appendix I) and provide an accurate representation of the biological problem.

Incorporating a dysplastic MDS HSC population has a direct effect on the healthy erythroid population, decreasing its numbers by competing for nutrients and space. Mathematically, the presence of a dysplastic population has been implemented via a Hill Equation, which can be too rough of an approximation. The *in-vivo* competition mechanisms between these two populations are largely unknown, but seem centered in competition for space and nutrients in the bone marrow. Nonetheless, the model presented herein is able to capture some of that complexity, with inhibitory effects seen in the decreased number of committed cells (Figure 19) as well as that of RBC (Figure 20).

All the simulations performed show that the model structure appears to capture correctly the biological complexities of the erythroid differentiation. The perturbations input into the system by adding MDS cells also display the expected effects. Following the framework proposed by Kontoravdi et al., Global Sensitivity Analysis (GSA) should be performed immediately after the model development stage (Kontoravdi C 2005). GSA, and in particular the Sobol Indices Method (Sobol 2001) can be used to systematically assess the relevance of model parameters to the model outputs. Parameters that are significant will need to be estimated with a greater degree of

accuracy, as small variations in their nominal values can greatly impact the model outputs. Non-significant parameters have little to no influence in the outputs and can often be left at their nominal values. One can arbitrarily define a threshold value for the sensitive index, which is often set at 0.1 and subject to change according to the modeller's perception of the relevance of the model outputs (Kiparissides A 2008 and Kotidis P 2019). The first order indices cannot, by design, be bigger than 1.

As a proof-of-concept, we have used GUI-HDMR to perform GSA and compute Sobol's first order sensitivity indices for the differentiation, death and proliferation rates of all the erythroid cellular populations (model outputs). These first-order indices are shown on a heatmap as a function of time, which greatly helps to draw conclusions from the model structure. Different parameters may be significant to the same output at different simulation times, elucidating the dynamic structure of the model.

For HSCs, the differentiation and proliferation rates are significant throughout the simulation time (Figure 21). This is in good agreement with the model structure and the intrinsic biology of the system. As these cells are at the top of the differentiation cascade, their numbers over time are highly dependent on the velocity which they either reproduce themselves (proliferation) or differentiate to a more mature cell type. For BFU-E cells, the most significant parameter is its differentiation rate, as well as the differentiation rate of the previous cellular intermediates (Figure 22). This pattern is also observed for all intermediate cellular stages, and highlights the linear structure of the model. It is impossible for one cell to become a BFU-E without first having been an HSC, MPP or CMP – thus, the significance of the differentiation rates of the previous cellular intermediates.

Interestingly, for CFU-E cells, one of the most significant parameters is its death rate. As mentioned above, there is a biological check-point at the CFU-E stage during erythroid differentiation; in the absence of EPO, cells will undergo apoptosis and stop their differentiation.

It is quite interesting that GSA can capture that “biological bottleneck” (Figure 23), which is reflected in the model by the implementation of Hill Equations. Lastly, for RBCs, most of the differentiation rates of previous cellular intermediates are significant, apart from the differentiation rate from PolyE to OrthoE, potentially because these are late, committed progenitors which normally will differentiate all the way to RBCs (Figure 24).

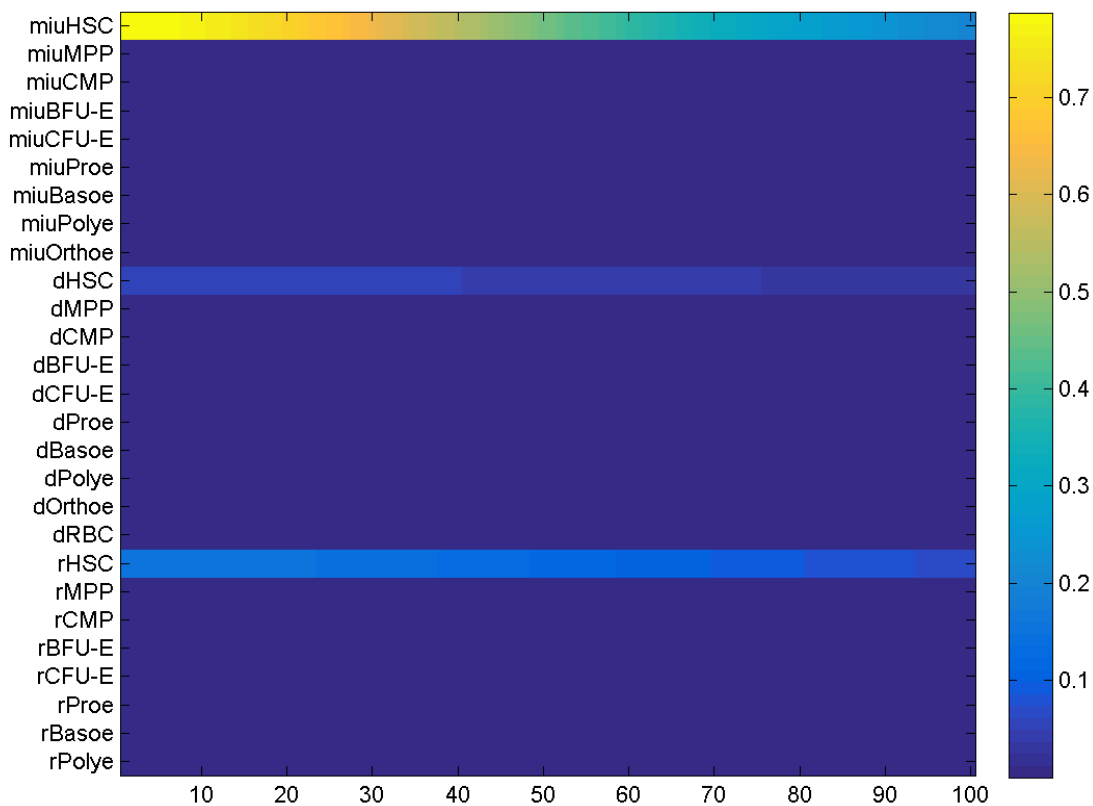


Figure 21 –First-order sensitivity indices for HSCs, as a function of time.

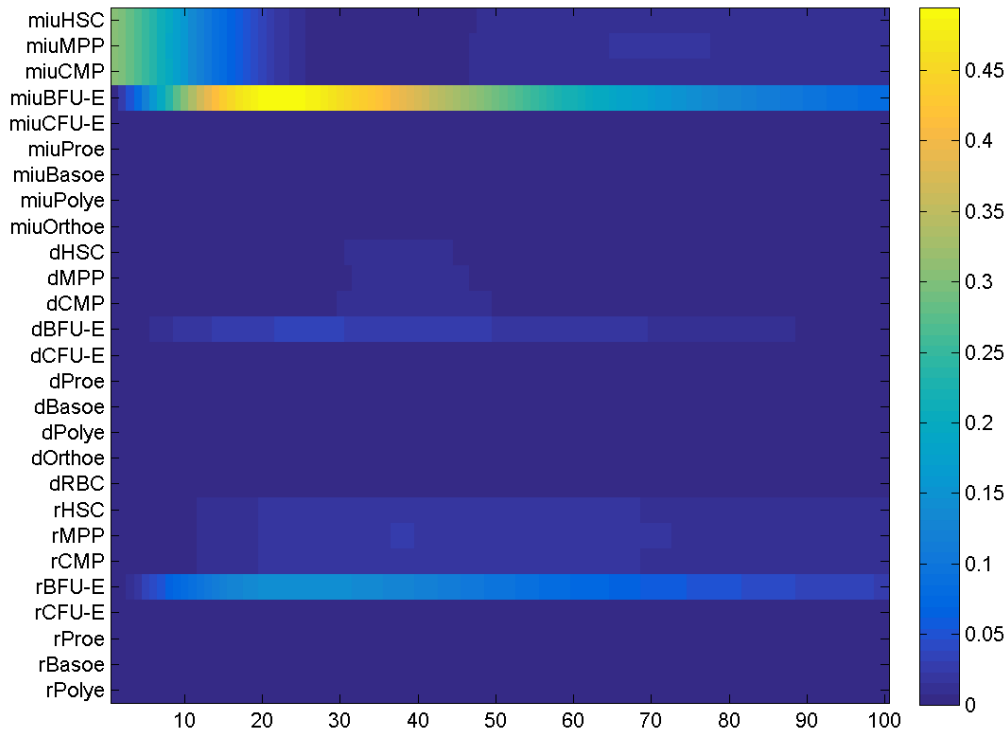


Figure 22 – First-order sensitivity indices for BFU-E cells, as a function of time.

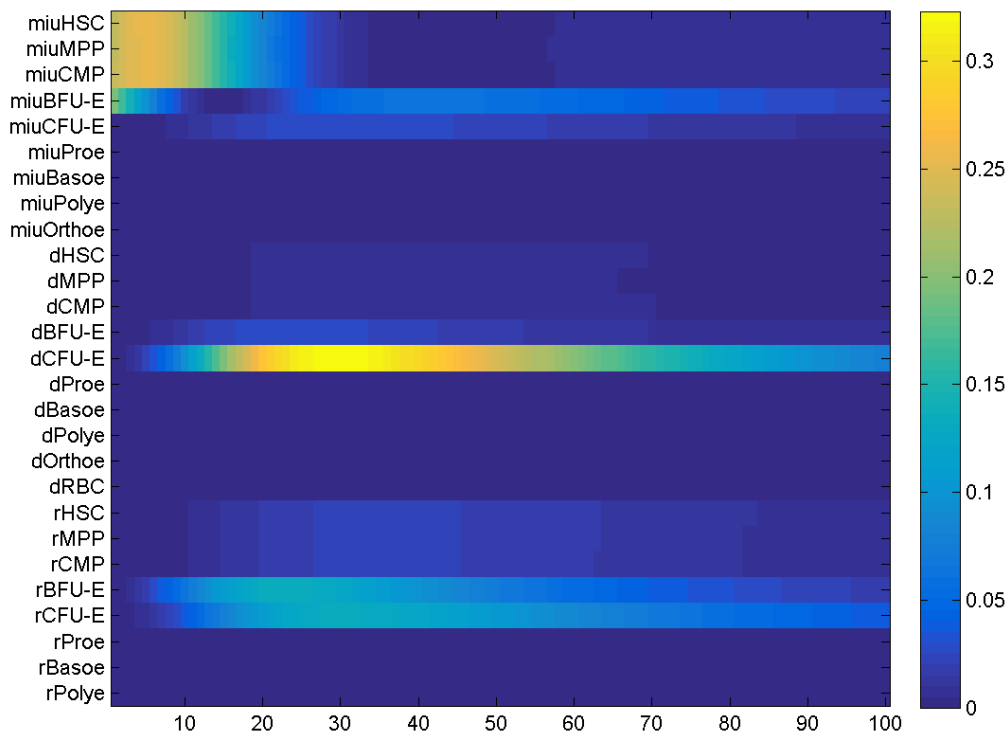


Figure 23 – First-order sensitivity indices for CFU-E cells, as a function of time.

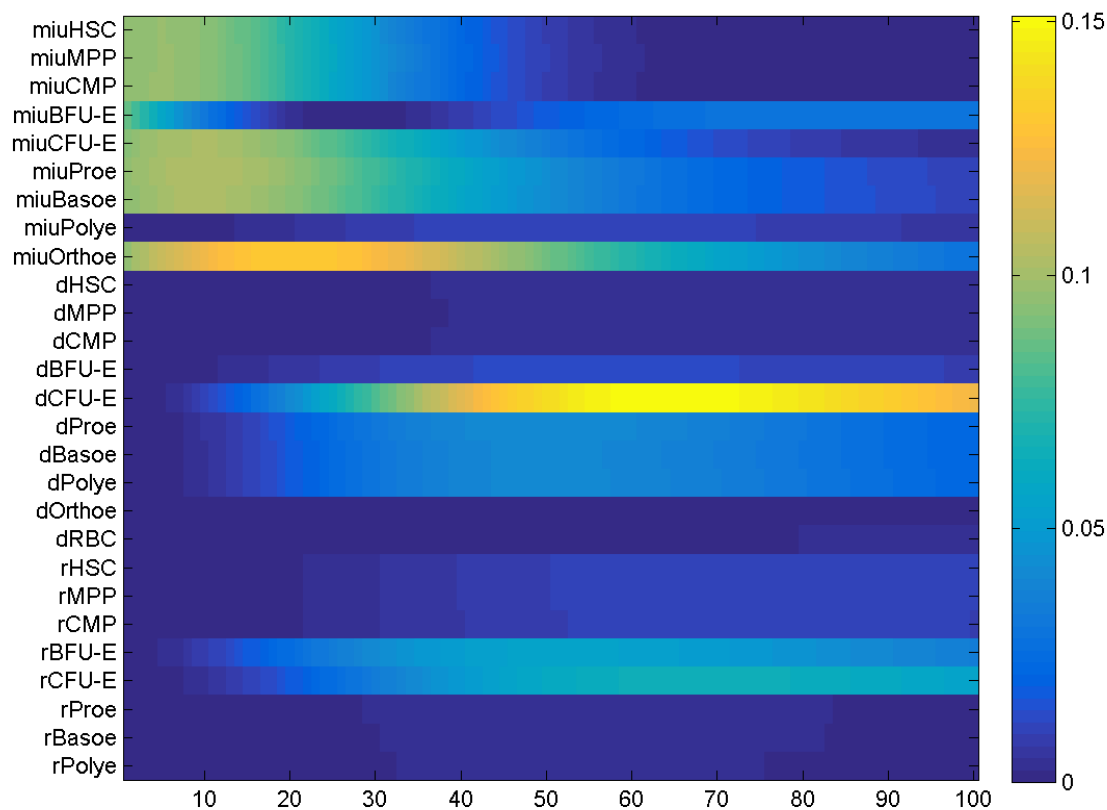


Figure 24 – First-order sensitivity indices for RBCs, as a function of time.

Using Sobol's method for GSA, one generates a Monte-Carlo combination of parameter values, usually within $\pm 100\%$ of their nominal value. This directly implies that the “real” value for your parameter is within that $\pm 100\%$ range, rendering the GSA useless otherwise. From here, it follows that one needs a relatively good initial guess for the parameter values in order to proceed with GSA. If that is not the case, preliminary experiments are required to obtain better initial guesses for the parameters via parameter estimation algorithms. The parameter values herein are no more than an educated guess based on what is described in the literature (Appendix I) as they have been set based on qualitative considerations rather than quantitative ones. Employing GSA techniques using such parameter values might be dangerous, as the derived conclusions might not yield when a better estimate for the parameter values is obtained. It is clear that GSA is a

useful tool for model structure analysis while it helps constraining the experimental work required to produce data for parameter estimation.

4.3. Future Work

We have herein created a novel, detailed, macroscopic healthy erythropoietic model that incorporates well-defined biochemical stimuli as driving forces for cells to proliferate, differentiate or die. This work serves as a proof-of-concept for hypothesis-generation and model-based optimisation scenarios. Upon successful experimental *in-vitro* validation, this modelling strategy can potentially be applied to other blood cellular lineages. The model includes competition between a pre-leukaemic, diseased, erythroid cell population and its healthy counterpart as an independent module, which can be used for simulations of diseased erythroid development.

The erythropoietic mathematical model herein is at a preliminary stage. Despite showing good responsiveness and biological agreement with several scenarios, the parameter values used are no more than an educated guess, which can potentially skew the results. Future work can be divided into two categories: Essential and Refinement. The “Essential” future work is required for Model-Based applications such as, for example, cell culture protocol optimisations. The “Refinement” stage requires a fully-validated model that will then be tweaked to be a better representation of the reality.

Essential Future Work

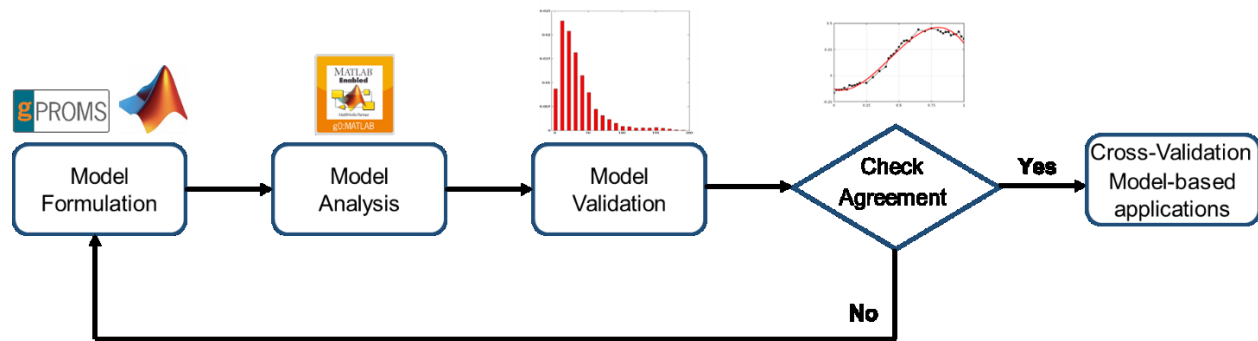


Figure 25 – Proposed framework for Essential Future Work. Currently, this work has reached the Model Analysis stage, but it needs further refinement with better initial guesses for the parameters, as aforementioned.

1) Obtain a better first approximation of the cell differentiation, death and self-renewal rates and fractions. Typically, mathematical modellers infer the first approximations for their parameters from previous experimental work already performed. Most of the existing work done in *in-vitro* erythropoiesis either uses serum-containing cell culture protocols (Lim M 2011) or does not provide time-series data on the evolution of all the cell populations described in this model (Giarratana MC 2011 and Trakarnsanga K 2017).

The work herein developed concerns a model of erythropoiesis whose purpose is mainly to describe erythroid differentiation *in-vitro* – it can also account for *in-vivo* erythropoiesis provided that additional feedback loops are implemented, such as homeostatic regulation and maintenance of relatively constant cell numbers across erythroid populations within the bone marrow. Therefore, an initial experimental dataset is required to obtain, via parameter estimation, more accurate parameter values. This would require starting with a well-defined number of HSCs and following their differentiation over a couple of weeks, ensuring that EPO, SCF and IL3 concentrations are kept constant throughout. A detailed flow cytometry panel for the identification of erythroid progenitors is detailed in Appendix II and could be used to support these experiments.

2) Rerun GSA and determine which parameters are significant. With a more accurate initial values for the rates and fractions obtained in Step 1, one will rerun GSA for all of the model parameters to highlight which parameters are significant, for each and every single model output. At this stage it might be easier to aggregate model outputs – potentially into Progenitor, Committed and Terminally-differentiated cell numbers – to make the GSA problem computationally more tractable. Significant parameters will have to be estimated with a high degree of precision, whereas the non-significant ones can be left at their nominal values.

3) Parameter Estimation and Model Validation. Without the correct validation, any model does not have an adequate utility. Model validation will be done by running several independent experiments, varying initial number of HSCs and cytokine concentrations and comparing the *in-vitro* results with the *in-silico* simulations. Parameters can be re-estimated at this stage to provide adequate agreement, whereby the model results do not deviate by more than 20% of the experimental results. If the gap between *in-vitro* results and *in-silico* simulations seems too big to bridge, we might need to go back to the drawing board and add additional features to the model, to better describe the reality.

Refinement Future Work

1) Model-based optimisation. A fully validated, representative model can be further used for several model-based applications. In particular, model-based optimisation could potentially open up avenues for the *in-silico* optimisation of the *in-vitro* erythroid differentiation process, in areas such as optimal medium design for the mass production of RBCs. Cytokines are herein considered as constant stimuli over time, but in reality they are time-variant, as they thermally degrade in culture, for example. Implementing cytokine degradation and feeding schedules as functions of time and running optimisation algorithms can potentially minimize process costs, maximize RBC yield and provide a more accurate description of the biological problem.

2) Model expansion. The erythroid model herein presented is part of a bigger project, towards comprehensively modelling haematopoiesis (Figure 11). Future work will involve model development focused on other blood lineages (granulocytic and megakaryocytic) and providing the appropriate link between the models, as they all start from the same cellular stage (HSC). For the purpose of both healthy and diseased (MDS) haematopoiesis modelling we have not consider lymphopoiesis. The existing model can be expanded to include it using a similar framework, by adding the relevant equations, regulatory cytokines and feedback loops.

3) Transition to 3D and/or *in-vivo* modelling. For more detailed *in-vitro* applications, it would potentially be useful to expand the work herein to consider spatial distribution of nutrients, cells and cytokines within 3D culture systems (Allenby MC 2019). Moreover, for human patient applications, this work can potentially be expanded to describe *in-vivo* haematopoiesis. Feedback loops between early progenitor cells and terminally differentiated cells need to be implemented. Parameters such as rates and fractions will also need to be re-estimated, to better represent the *in-vivo* scenario.

5. Long-term 3D cultures of primary MDS cells

Chapter Summary

To date, only one group has successfully developed a 3D system to culture primary MDS cells. Despite showing extremely promising and relevant results, the cell culture protocol applied could potentially have a limited application in patient-specific, personalised medicine since it uses xenogeneic components (FBS) and high cytokine concentrations – at least ten-fold higher than those present *in-vivo* (Elvarsdóttir EM 2020).

There is, at the time of writing, an unmet need for the creation of patient-specific, serum-and-xeno-free long-term 3D culture systems that can adequately mimic MDS features, providing a platform for drug discovery and optimisation, fundamental studies as well as candidate biomarker discovery. This chapter focuses in the attempt of developing such 3D cell culture system using primary MDS cells isolated from BM aspirates.

5.1. Introduction

The work herein paves the way for the establishment of a long-term 3D culture system that, while maintaining high cell viability throughout 24-day cultures, retains phenotypical MDS characteristics as seen upon sample collection. Our long-term 3D cell culture system (Chapter 3.3 and 3.4) relies on the assumption that stromal cells present in the original bone marrow sample will provide the adequate support for haematopoiesis.

Experiments with cells from healthy controls were performed both to screen the optimal cell seeding density as well as identify the best coating for the 3D scaffolds. These cultures were performed using Mononuclear Cells derived from Umbilical Cord Blood (UCB-MNCs). This cell source was selected due to its wide availability, despite not being from adult, age-matched healthy controls (e.g. adult healthy BM cells).

The cell seeding density is a crucial factor, as a higher number of cells will promote stronger cell-cell interactions while constraining the information one can extract from these cultures, due to the limited cell numbers available per sample. Carefully selecting an appropriate seeding density aims to maximize the number of time-points and the different analyses, while ensuring that cells can communicate and proliferate within the 3D culture system.

We have tested several combinations of coating and seeding density using healthy controls, hoping to determine what would be the best, optimal combination to be used for primary MDS samples. After selecting the best condition for cell seeding and scaffold coating, we cultivated three independent primary MDS samples and compared the culture outcomes to that of the healthy controls, performed using UCB-MNCs.

5.2. Results and Discussion

This work has been conducted in two distinct phases. Firstly, we used healthy UCB-MNCs to optimise culture conditions, aiming to maximize the information extracted while minimizing the number of cells seeded per scaffold. We used supernatant cells as an accessible probe to assess culture evolution inside the scaffold. This methodology was the standard approach in our research group, being also followed by several other authors (dos Santos SB 2016, Santos JD 2017 and Allenby MC 2019). With regards to cell viability, the results herein described confirm the applicability of using supernatant cells as a probe for culture fitness and evolution over time. *In-situ* confocal microscopy would need to be performed to further confirm the phenotypical dynamics observed in supernatant cells throughout the culture.

Secondly, we obtained BM-MNCs from consenting MDS patients and followed the optimal culture conditions determined earlier, comparing healthy and diseased cell culture outcomes. This was done at cell viability and phenotype level, with interesting results especially in the phenotypes observed throughout the culture.

5.2.1. Healthy control cultures

Healthy control experiments using UCB-MNCs were done using four independent replicates, screening six different conditions:

Table 15 – Six different conditions tested using UCB-MNCs as healthy controls.

Condition Name	Seeding Density per scaffold	Scaffold Coating
0.5M Unc	0.5 million UCB-MNCs	Uncoated
1M Unc	1 million UCB-MNCs	Uncoated
2M Unc	2 million UCB-MNCs	Uncoated
0.5M Col	0.5 million UCB-MNCs	Collagen
1M Col	1 million UCB-MNCs	Collagen
2M Col	2 million UCB-MNCs	Collagen

Firstly, we evaluated supernatant cell viability by flow cytometry at every medium replenishment (every other day) in order to determine whether the scaffold coating had an effect in cell viability throughout the culture. Across all conditions, coating did not seem to exert an influence on the supernatant cell viability as there are no significant differences between each group compared – same seeding density, different coating – at any time point (Figure 26).

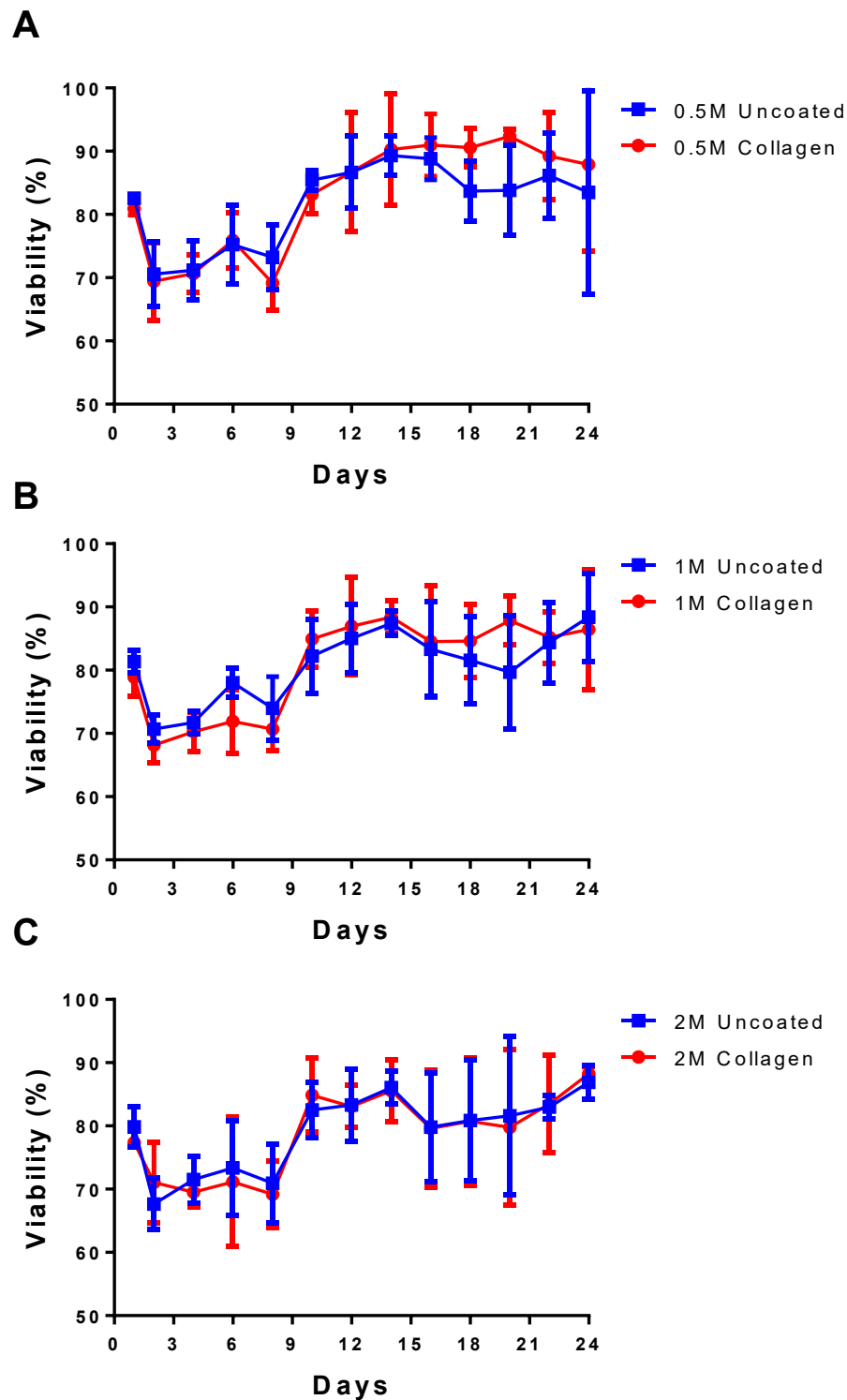


Figure 26 – Supernatant cell viability was assessed for cells cultured at 0.5 million (A), 1 million (B) and 2 million (C) cells per scaffold. Across all conditions, no statistically significant differences were seen at the supernatant cell viability level. Slightly greater variability is seen when using collagen-containing, but within a negligible level. Results are shown as average $\pm$ SD of N=4, n=3.

Next, we indirectly assessed *in-situ* cell numbers within the 3D scaffolds by fluorescent detection of dsDNA using PicoGreen (Figure 27). Fluorescence measured at Day 1 served as a baseline for each culture condition, and results throughout the 24-day culture are shown as relative-to-baseline fold differences. Apart from 2M Uncoated, all conditions saw a statistically significant decrease in cell numbers from Day 1 to Day 8. Conditions with lower seeded cell numbers (0.5M and 1M Collagen) show statistically significant differences up to Day 24, when compared to Day 1. Across all conditions, no differences were seen when comparing Day 8 with later days. Within each time point, there were no statistically significant differences between culture conditions. Taken altogether, these results show that coating and seeding density do not seem to play a role in influencing cell growth over a 24-day culture. Across conditions, the trend seems to be a sharp decay between Day 1 and Day 8, followed by a period of cell number maintenance in culture.

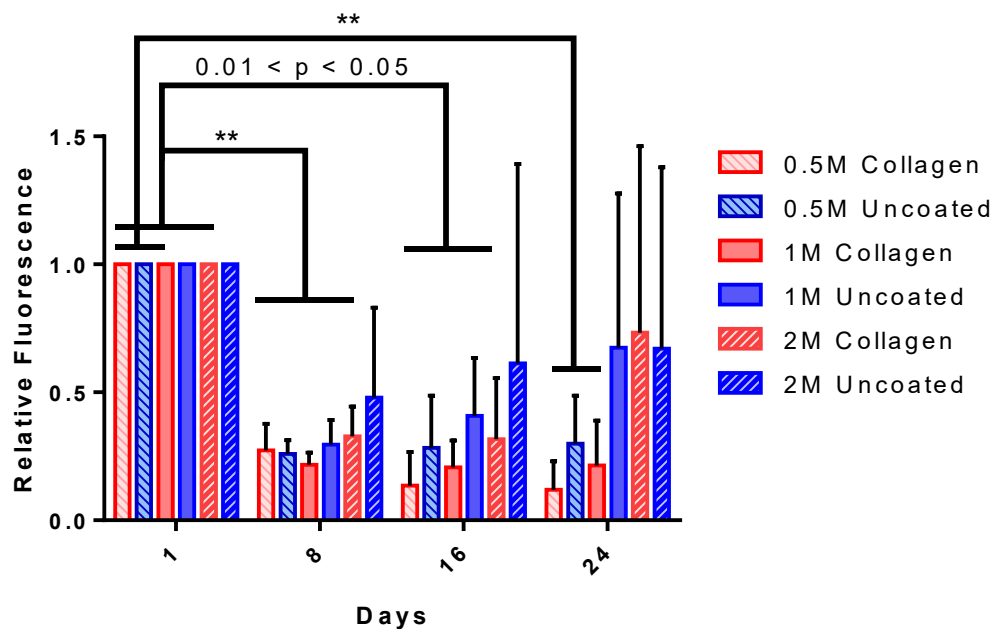


Figure 27 – Fluorescence detected by dsDNA-PicoGreen cross-linking *in-situ* within scaffolds across all groups and time points, normalized per group at its Day 1 fluorescence levels. Results are shown as average $\pm$ SD of N=4, n=3.

We next measured *in-situ* cell viability at the core of the scaffold, using thin-cut slices stained with Calcein AM (live cells stained green) and Ethidium-Homodimer 1 (dead cells stained red). It is

worth noting that the core of the scaffold is where nutrient concentrations are hypothetically lower, and the build-up of toxic metabolic by-products is hypothetically higher, as this is the zone of the scaffold furthest away from the fresh media surrounding it. The general trend for every condition seems to be a steady increase in viability measured *in situ*, directly at a thin slice of the core of the scaffold, (Figure 28) which is consistent with the supernatant viability data. Scaffold coating and cell seeding density do not seem to play any role in influencing *in-situ* cell viability. Conditions where more cells are seeded exhibit greater viable cell absolute numbers at Day 24 (Figure 29), whereas it seems that all conditions show a significant number dead cells at Day 1, 24h after isolation (Figure 30). This is potentially due to the seeding process, which might be stressful for cells, despite cell viability immediately after isolation by Ficoll-Paque gradient being around 80-85%.

There is good agreement between the *in-situ* viability measurements (Figure 30) and those performed with supernatant cells (Figure 26), as the supernatant measurements reflect correctly the overall culture viability dynamics throughout the 24-day time horizon. This strengthens the hypothesis that the easily-accessible supernatant cells can be used as a good first approximation for culture fitness.

.

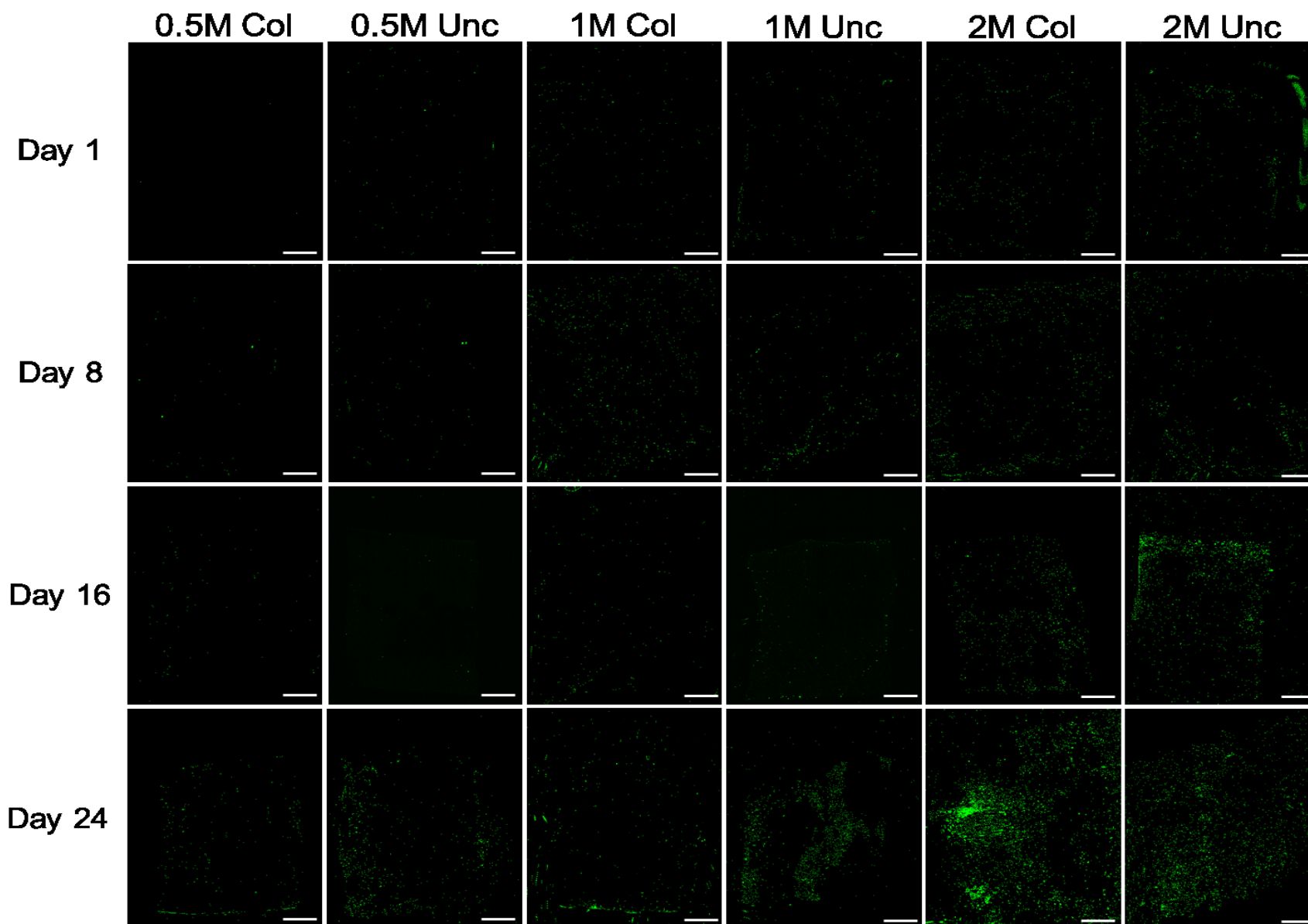


Figure 28 – Sections of the scaffold core imaged using confocal microscopy. Green fluorescence is present on live cells due to Calcein AM Green, a viability dye. Results are shown as representative images of N=4, n=1. Scale bar shown represents 500µm. Aggregated data from all images is shown in Figure 30.

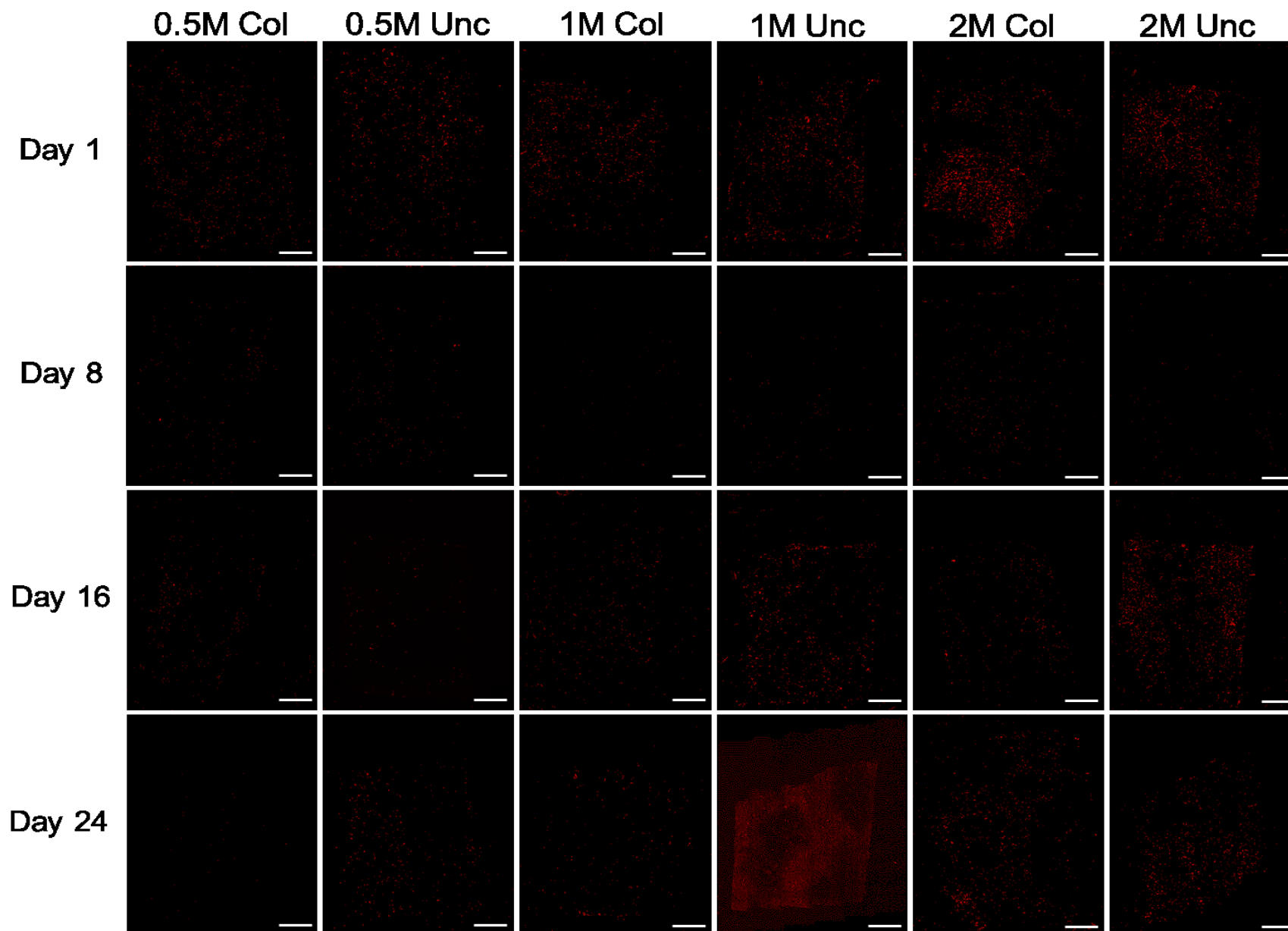


Figure 29 - Sections of the scaffold core imaged using confocal microscopy. Red fluorescence is present on dead cells due to the binding of Ethidium Homodimer-1 to DNA. Results are shown as representative images of N=4, n=1. Scale bar shown represents 500µm. Aggregated data from all images is shown in Figure 30

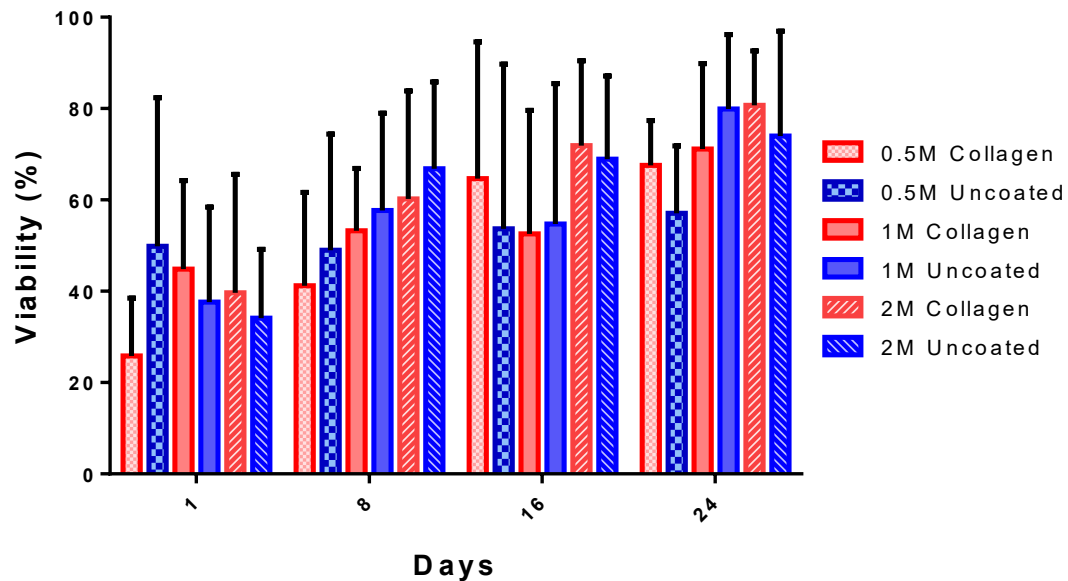


Figure 30 – *In-situ* viability determined by Confocal Microscopy at the core of the scaffold. Viability was calculated as the ratio between viable cells (Calcein AM Green-positive) and total cells (viable and dead, verified by Ethidium homodimer-1-positive). Results are shown as average $\pm$ SD of N=4, n=1.

We next examined the phenotypes of several cell populations of interest present in the supernatant via flow cytometry. All of the flow cytometry events shown here have had doublets excluded (two cells aggregated) by exclusion gating in FSC-A vs FSC-H, and are live-gated (staining positive for a viability dye).

We first assessed the population dynamics of leukocytes throughout the culture, identified by the expression of the CD45⁺ CD36⁻ phenotype. Where CD45 is a common leukocyte marker, lineage-committed healthy haematopoietic immature progenitors also express CD45, increasing CD36 expression as they mature (Landberg N 2018). CD45 is also a crucial marker for immunophenotyping of several haematological malignancies such as AML (Lacombe F 1997) and MDS (Westers TM 2012).

All the conditions show a statistically significant decrease in supernatant CD45⁺ CD36⁻ cells towards the end of the culture, and this effect seems to be more pronounced in the conditions with lower cell seeding densities in both the Uncoated (Figure 31) and Collagen (Figure 32) conditions. This decrease may be explained by the loss of CD45⁺ phenotype as lineage-committed progenitors mature in culture.

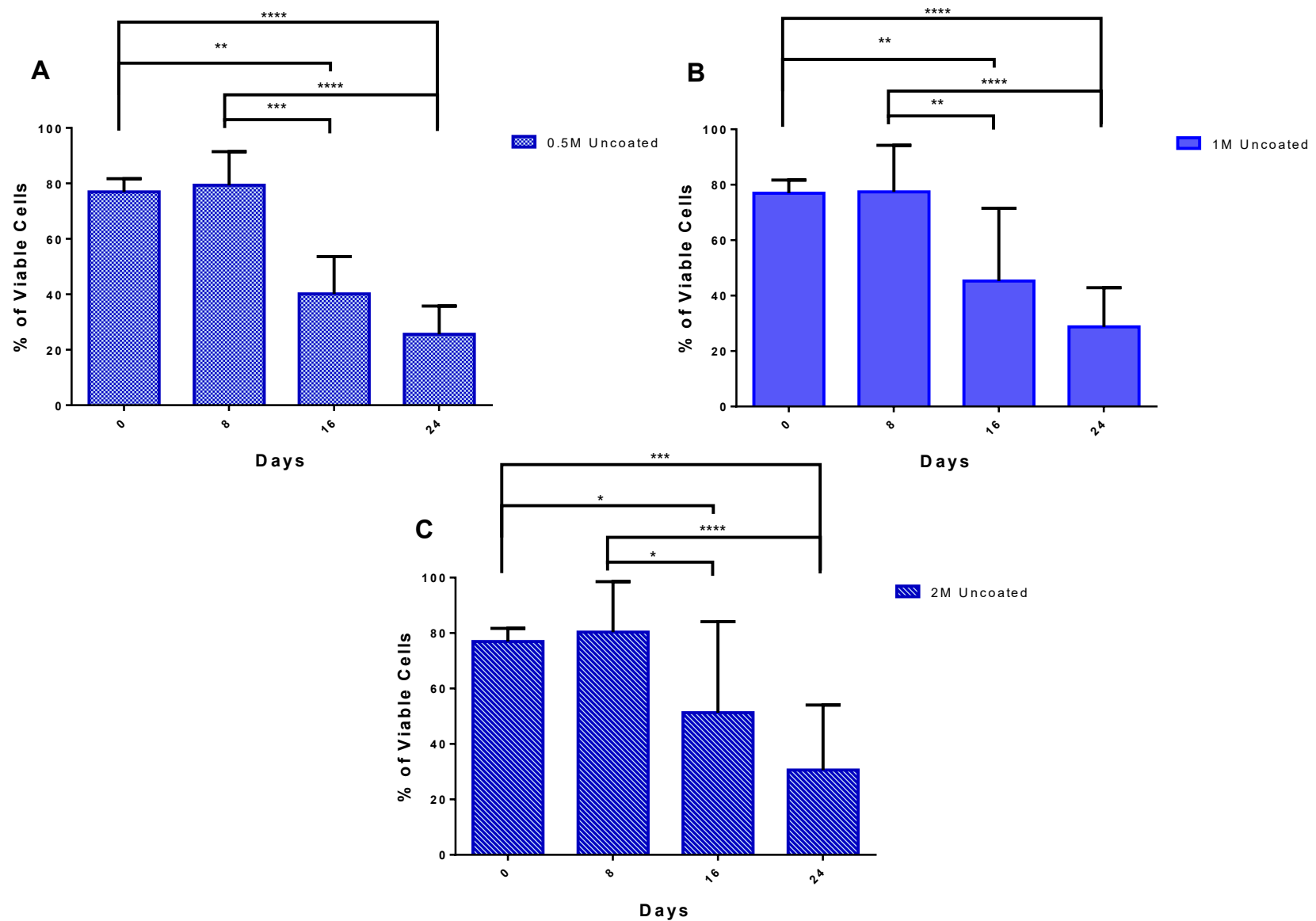


Figure 31 – Phenotypical changes of CD45⁺ CD36⁻ (leukocyte) supernatant cells, as a percentage of total viable cells, throughout a 24-day culture. There is a decrease in this phenotype from the initial time points (Day 0 – seeding – and Day 8) to the later time points. This decrease is inversely proportional to the initial seeding density. Results are shown as average $\pm$ SD of N=3. * P < 0.05; ** P < 0.01; *** P < 0.001; **** P < 0.0001.

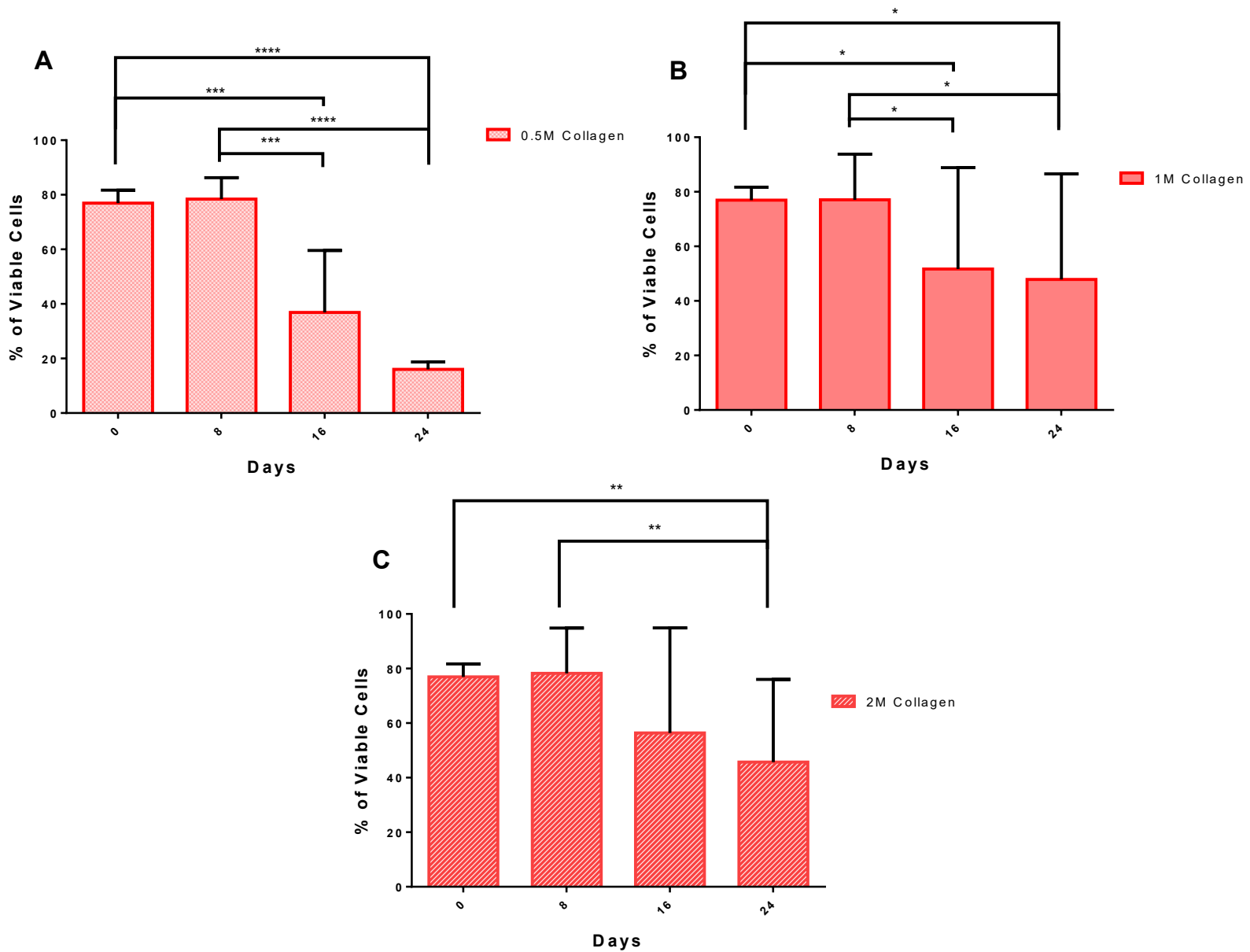


Figure 32 – Phenotypical changes of CD45⁺ CD36⁺ (leukocyte) supernatant cells, as a percentage of total viable cells, throughout a 24-day culture. There is a decrease in this phenotype from the initial time points (Day 0 – seeding – and Day 8) to the later time points. This decrease is inversely proportional to the initial seeding density. Results are shown as average $\pm$ SD of N=3. * P < 0.05; ** P < 0.01; *** P < 0.001; **** P < 0.0001.

Our 3D culture system has previously been demonstrated to have multi-lineage differentiation potential towards multiple haematopoietic lineages without the addition of any endogenous cytokine factors (Mortera-Blanco T 2011). Because impaired erythropoiesis is a major hallmark of MDS patients; we then assessed the erythroid differentiation potential of this culture system to then be able to compare later with that of primary MDS cells.

In the supernatant cells, the expression of an early erythroid cell phenotype (CD235a⁺ CD71⁺) increased significantly throughout the culture, especially in the collagen-containing conditions (red bars in Figure 33). While the uncoated conditions did not show a statistically significant increase, the collagen-containing conditions showed greater standard deviations. This can be a consequence of, for example, an heterogeneous collagen distribution along the several scaffolds cultured.

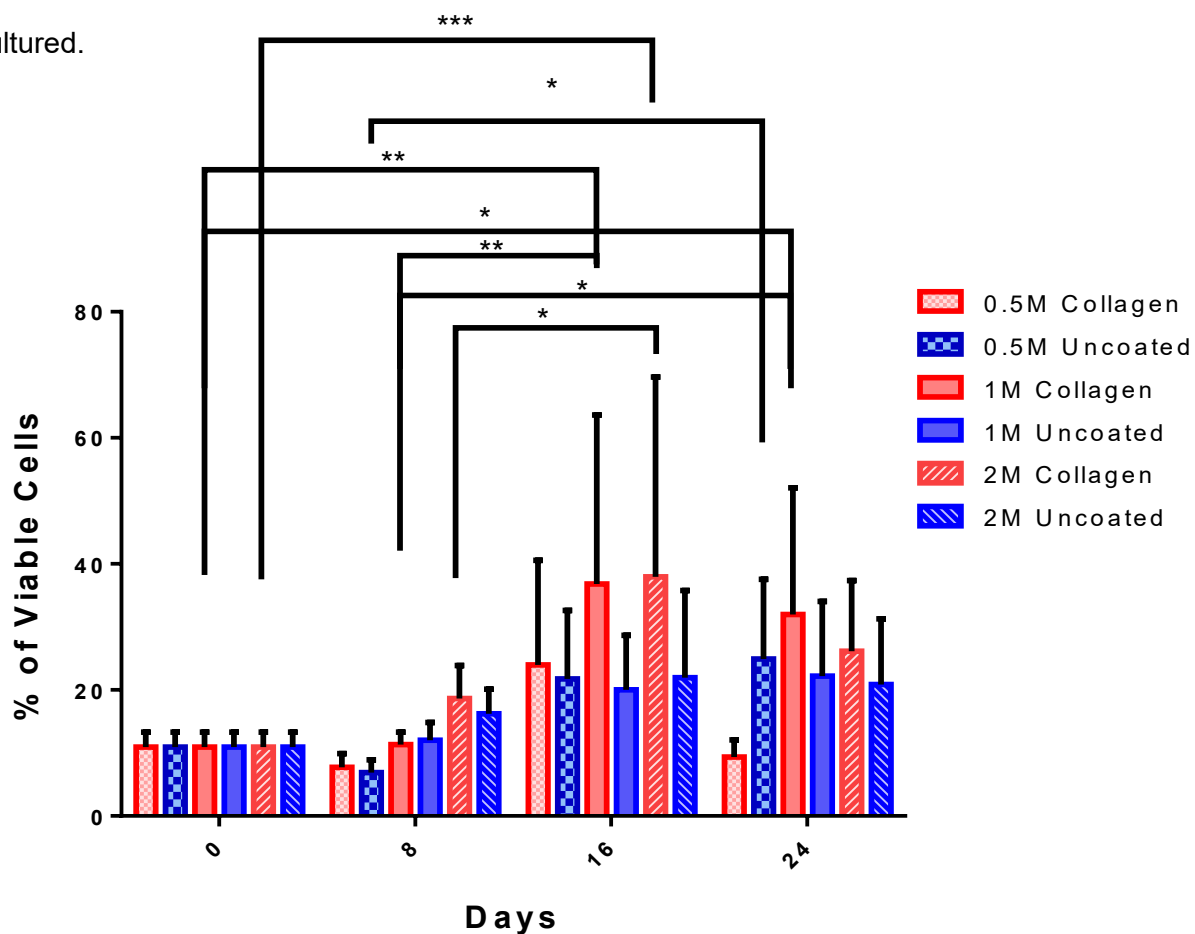


Figure 33 – Phenotypical evolution, throughout the culture, of supernatant early erythroid cell populations (CD235a⁺ CD71⁺) expressed as a percentage of total viable cells. Results are shown as average $\pm$ SD of N=3. * P < 0.05; ** P < 0.01; *** P < 0.001.

Late erythroid progenitors – and potentially enucleated RBC progenitors – expressing CD235a⁺ CD71⁻ were seen at cell isolation, probably as contaminants; however, their numbers sharply decreased across all conditions until Day 8 (Figure 34). No further expansion of this phenotype is seen throughout the culture. This result is consistent with not adding any EPO to the cultures, hampering terminal maturation of erythroid progenitors.

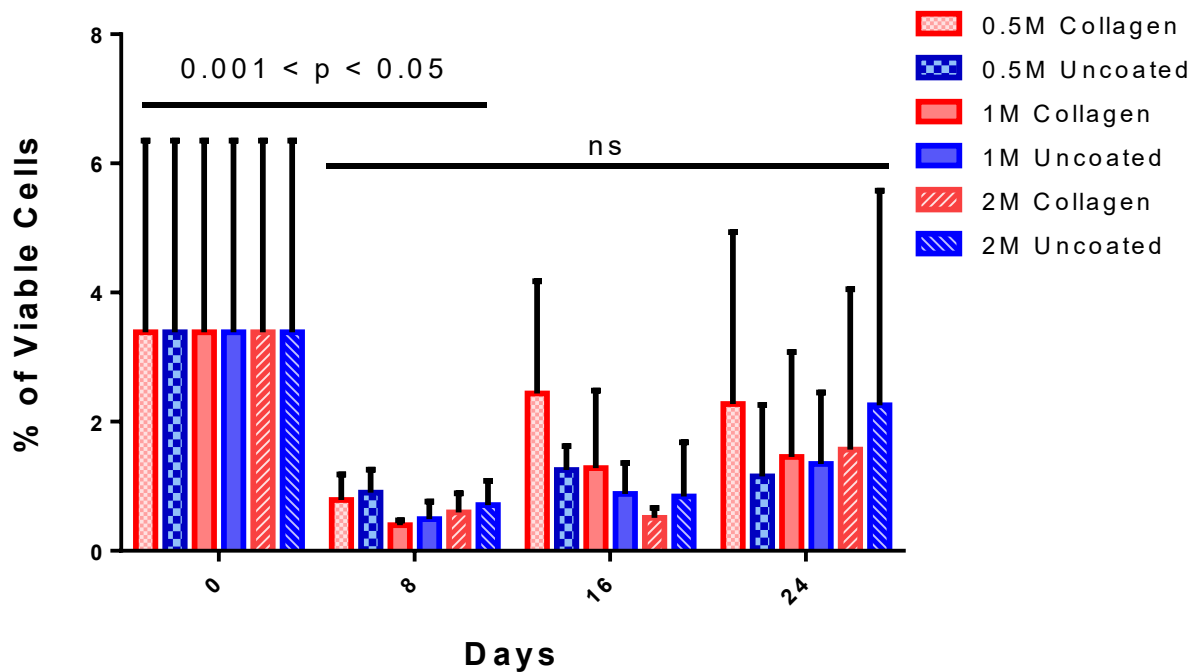


Figure 34 – Phenotypical evolution, throughout the culture, of supernatant late erythroid cell populations (CD235a⁺ CD71⁻) expressed as a percentage of total viable cells. Results are shown as average $\pm$ SD of N=3. $< 0.001 < p < 0.05$ shows that there is a statistically significant decrease in this phenotype, across all conditions, from Day 0 to Day 8. No other statistically significant (ns) difference was found.

Interestingly, in most of the conditions, the numbers of mid erythroid progenitors (CD235a⁺ CD71⁺) decreased when comparing Day 0 to the last time-points, while remaining constant from Day 8 onwards for both Collagen (Figure 35) and Uncoated (Figure 36) conditions. This may potentially highlight the intrinsic priming of our 3D system to support early-to-mid erythropoiesis without the addition of any endogenous cytokines. Measuring EPO levels in the culture medium, as well as replicating this experiment with the supplementation of EPO would be very interesting future steps.

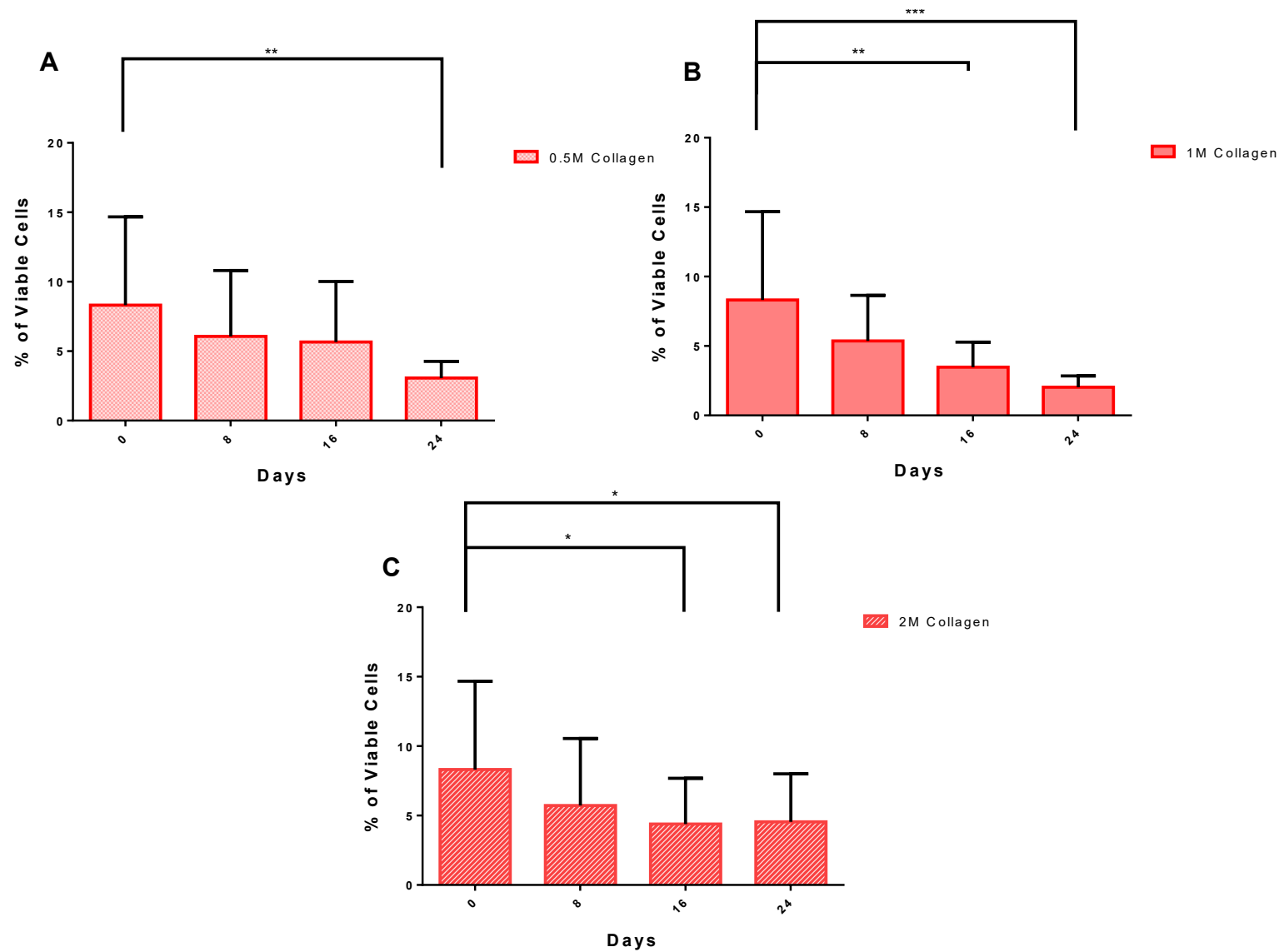


Figure 35 – Phenotypical changes of CD235a⁺ CD71⁺ (mid erythroid progenitor) supernatant cells, as a percentage of total viable cells, throughout a 24-day culture. Collagen-coated scaffolds were seeded with 0.5 million (A), 1 million (B) and 2 million (C) cells respectively. Results are shown as average $\pm$ SD of N=3. * P < 0.05; ** P < 0.01; *** P < 0.001.

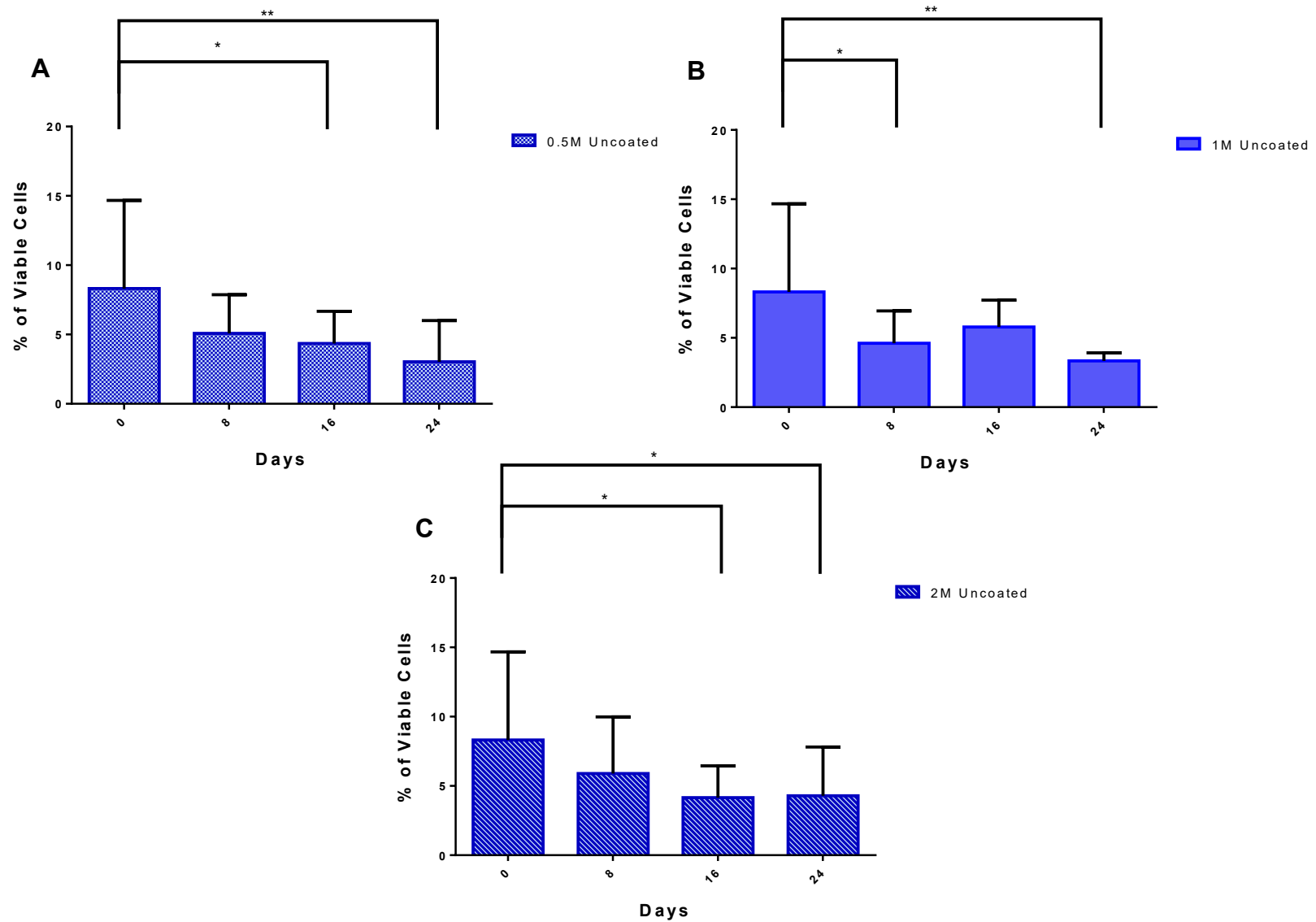


Figure 36 – Phenotypical changes of CD235a⁺ CD71⁺ (mid erythroid progenitor) supernatant cells, as a percentage of total viable cells, throughout a 24-day culture. Uncoated scaffolds were seeded with 0.5 million (A), 1 million (B) and 2 million (C) cells respectively. Results are shown as average $\pm$ SD of N=3. * P < 0.05; ** P < 0.01;

When looking at combined bi-variate flow cytometry plots, we gain insights on how the cellular phenotypes change throughout the culture. Forward Scatter (FSC) versus Side Scatter (SSC) show a visual representation of the cell size (FSC) and internal complexity (SSC). CD36 vs CD45 show the evolution of the immature myeloid progenitors (CD45⁺ CD36⁻) with the gain of CD36 expression as they mature and loss of CD45 upon full lineage commitment. Finally, CD71 vs CD235a shows the evolution of erythroid-specific progenitors, which start by exhibiting a CD71⁺ CD235a⁻ phenotype, followed by sequential gain of CD235a expression and loss of CD71 expression.

Figure 37 shows a representative example of a bi-variate plot analysis for a representative replicate. Each row represents an individual time point with the three bi-variate plots as described above. For CD36 vs CD45 (central panel), the phenotypical changes seen as cells mature towards the erythroid lineage would translate into some of the dots slowly shifting clockwise, from the top-left corner to the bottom-right corner. For CD71 vs CD235a, maturation would result in a counter-clockwise movement from the bottom-right corner to the top-left corner.

At Day 0, a large cell population displays a pan-myeloid phenotype (CD45⁺ CD36⁻). This is largely expected as these are MNCs isolated from UCB, which would be primarily composed of leukocytes. The percentage of total cells expressing an early erythroid phenotype (CD71⁺ CD235a⁻) increased almost 3-fold between Day 0 (9.61%) and Day 24 (29.3%), which further confirms the capability of this culture system to drive, to an extent, lineage commitment without the addition of endogenous cytokines.

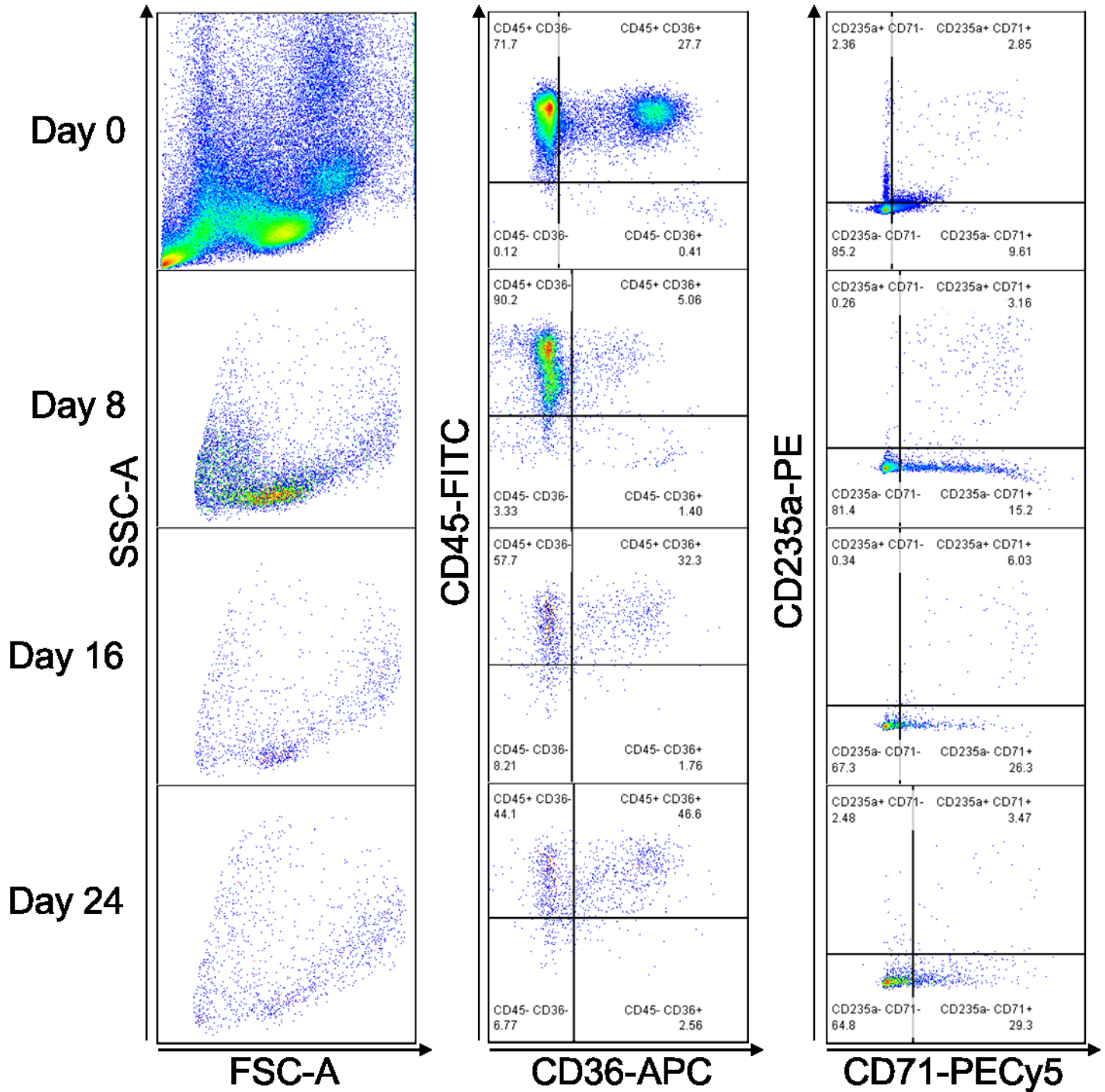


Figure 37 – Representative flow cytometry plots of CD45⁺ CD36⁻ and erythroid progenitor supernatant cells, seeded at 1 million MNCs per uncoated scaffold. After excluding doublets, gating (Day 8 to Day 24) was performed on live cells (Calcein AM positive). Negative-positive phenotypic boundaries were done using isotype controls.

Combining *in-situ* viability with supernatant cell viability does not highlight any clear differences between any of the tested seeding density and coating conditions. The phenotypical results across all conditions do not favour the selection of a specific seeding density or coating for the scaffolds either. Results from our group with primary AML and CLL cells show smaller intra-scaffold variations when using uncoated scaffolds, compared to that of collagen-coated scaffolds (data not shown). With regards to the cell seeding density, there is a fine balance between the number of cells seeded in each scaffold and the variety of analyses to be performed. Typically, the main bottleneck of culturing primary cancer cells is the rather limited cell numbers extracted from patient samples.

Despite being a good first approximation, the use of UCB-MNC cells as a healthy control should be done with caution. As opposed to “virgin” UCB-MNCs, cells derived from MDS patients are adult BM-MNCs, which have probably acquired several somatic mutations and are relatively aged. The most appropriate healthy control would be age-, comorbidity-matched adult healthy BM-MNCs. However, these cells are relatively hard to obtain, mainly because they require voluntary healthy donations and further bioprocessing in labs to obtain meaningful cell numbers. In the future, we plan to procure such cells and perform adequate healthy control experiments, especially when trying to compare metabolism in culture.

Based on all of the above, we chose to seed 1 million primary MDS cells per uncoated scaffold, which we believe to maximize the number of useful analyses during the culture while still not disrupting cell-cell interactions.

5.2.2. Primary MDS cell cultures

We cultured 3 independent primary MDS samples, seeding 1 million cells per uncoated scaffold. These patients have been clinically diagnosed with MDS and cells were obtained prior to any therapeutic approach.

At medium exchange, we assessed supernatant viability by flow cytometry and compared it with healthy control cultures. For primary MDS cell cultures, supernatant viability was never below 75% and stabilised at circa 80% after Day 10, which is a good indicator of culture health overall. It is important to note that these results are not influenced by cells that “preferred” to stay within the environment (culture medium) that surrounded the 3D culture system, as these were fully aspirated at every medium exchange. Instead, they are as egressed cells that left the 3D scaffold directly to the culture medium between each medium exchange.

Figure 38 shows a direct comparison between the UCB-MNC Healthy Control culture and the MDS culture, with similar viability percentages achieved throughout the 24-day cultures.

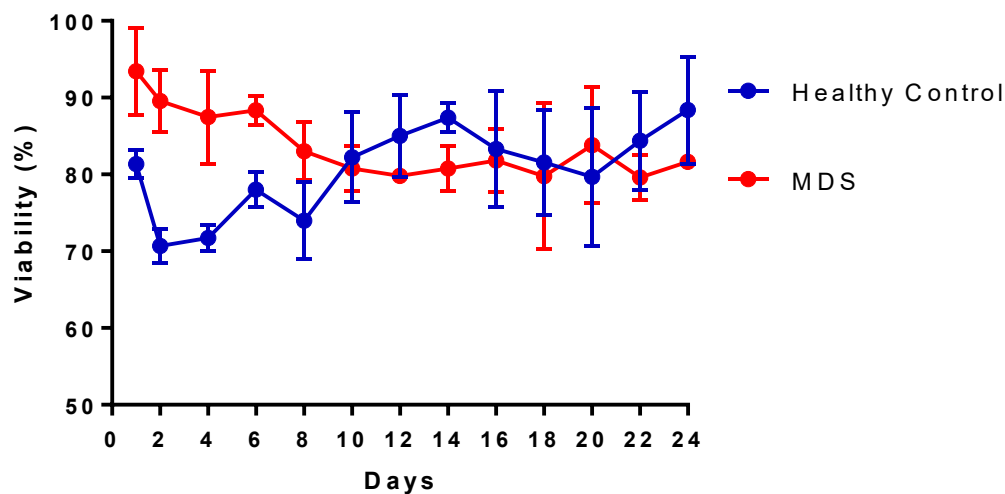


Figure 38 – Viability of supernatant cells was measured across four independent primary MDS cultures and compared to that of UCB-MNCs (healthy control). Results are shown as average $\pm$ SD of N=3 (healthy control) and N=3 (primary MDS cells).

Next, we indirectly assessed *in-situ* cell numbers within the 3D scaffolds by fluorescent detection of dsDNA using PicoGreen (Figure 39). Throughout a 24-day culture, all the experiments using primary MDS cells show a 50% drop in fluorescence (vs Day 1) by Day 8 remaining then relatively constant throughout the culture. This trend had already been observed across all healthy control conditions (Figure 27). One would expect cell growth throughout the culture, but maintenance of cell numbers can potentially mean that the culture system requires a long adaptation period – where cell numbers will initially decay and stabilise – to then sustain steady-state cell production and long-term growth. Longer cultures (~50 day) of both healthy and diseased cells can potentially verify this assumption.

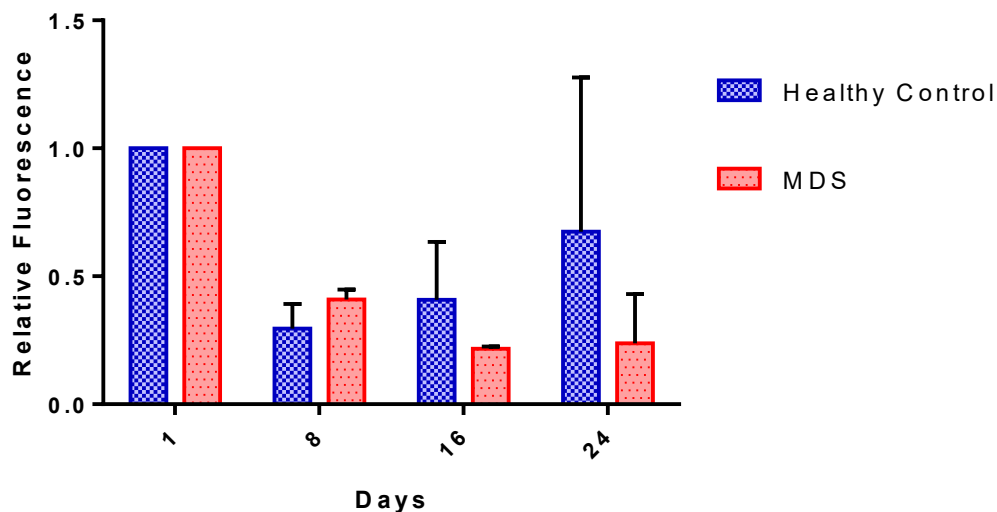


Figure 39 – Fluorescence detected by dsDNA-PicoGreen cross-linking *in-situ* within scaffolds and normalised, per condition, to its Day 1 levels. Results are shown as average $\pm$ SD of N=3 (healthy control) and N=2-3 (primary MDS cells).

In-situ viability by confocal microscopy was assessed in just one of the three primary MDS cultures due to the low number of cells extracted from the other two patient samples. When comparing to the healthy controls, these preliminary results seem to show similar viability profiles, coupled with lower dead cell numbers at Day 24, when culturing primary MDS cells (Figure 40). However, the lack of quantitative information and the insufficient number of independent primary MDS cultures

hinder any quantitative analysis that could potentially show differences within the culture as well as between the healthy control and the primary MDS cultures.

Phenotypically, independent primary MDS cell isolates have a wider sample-to-sample variability when compared to that of healthy UCB-MNC isolates (Figure 41). When stained with an anti-human CD45 FITC-labelled antibody, primary MDS cells show high variability in internal complexity (SSC-A), potentially with features of granulocytic dysplasia (top panel, region labelled as “Granulocytes”). This would need to be confirmed by further assessing cell morphology using histology techniques such as May-Grünwald Giemsa stains, for instance. UCB-MNCs, on the other hand, are much more consistent throughout the three biological replicates shown in the bottom panel.

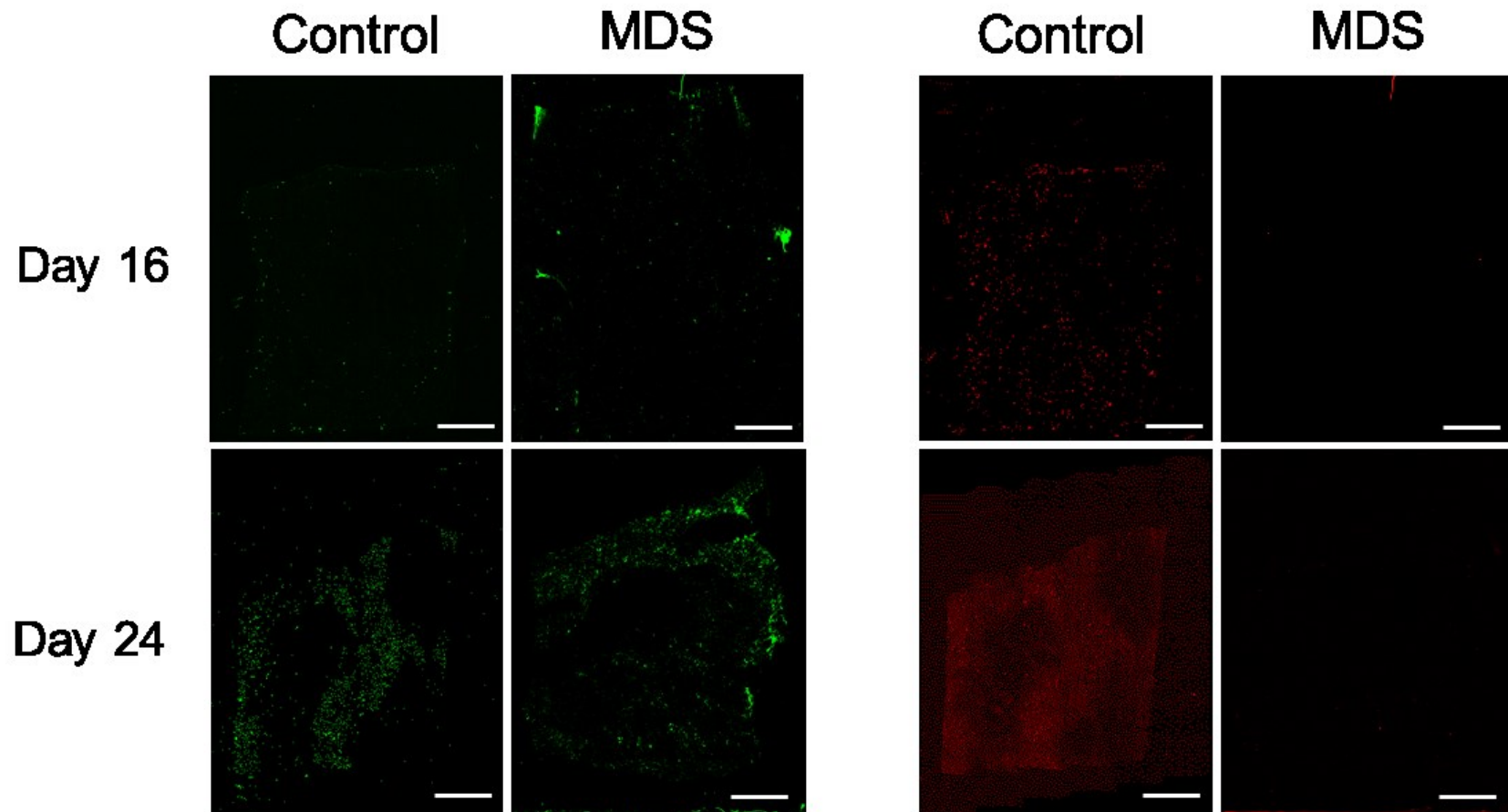


Figure 40 –Sections of the core of the scaffold imaged by confocal microscopy. Green fluorescence is present on live cells due to Calcein AM Green, a viability dye. Red fluorescence is due to the DNA binding of the cell impermeable Ethidium Homodimer-1 stain. Qualitatively, scaffolds seeded with 1 million cells seem to have similar viable cell numbers at late culture time-points (left panel) while the MDS condition exhibits less dead cells (right panel). Results are shown as representative images of N=4, n=1 (Healthy UCB-MNC Control) and N=1, n=1 (Primary MDS sample). Scale bar shown represents 500µm

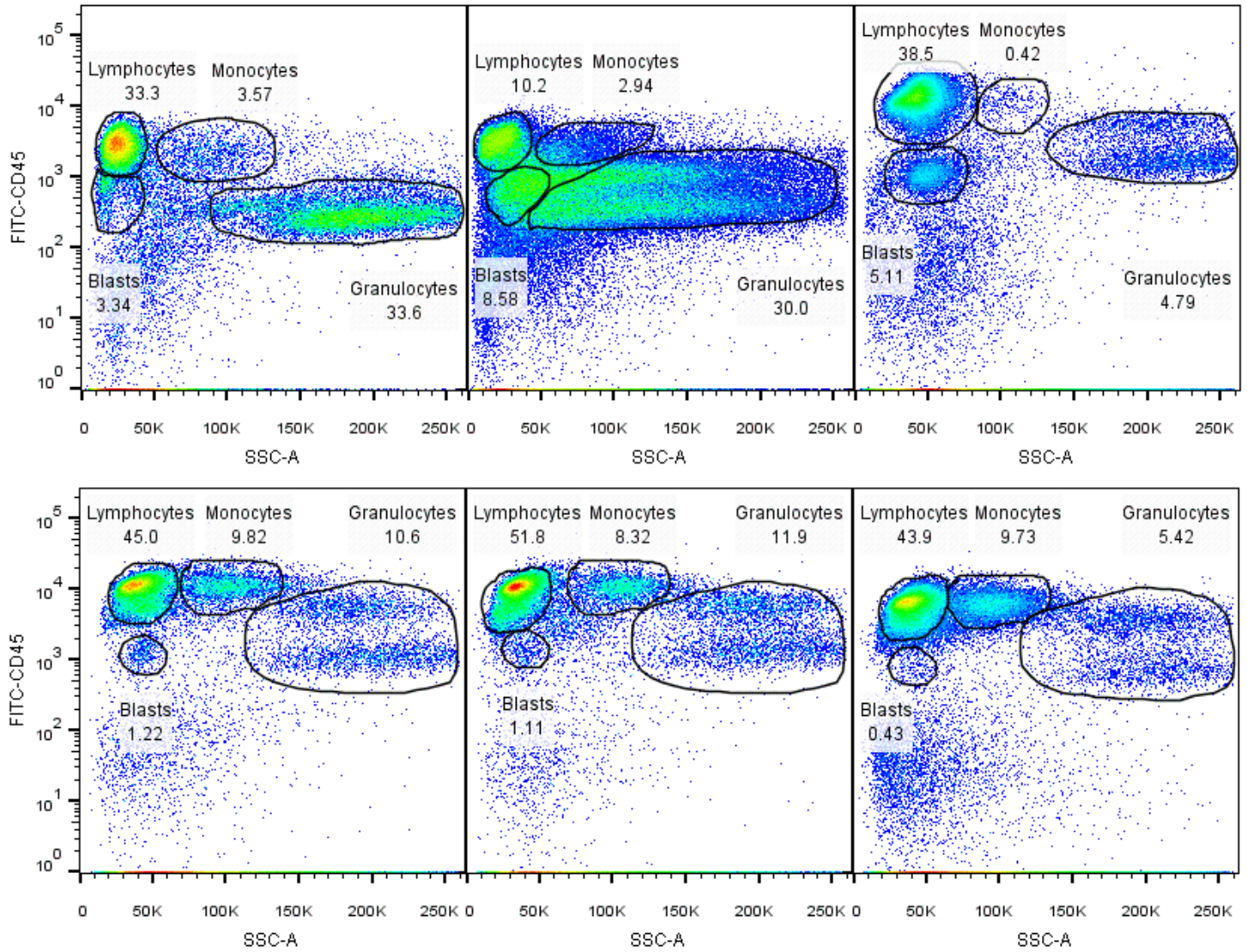


Figure 41 – Isolated cells from primary MDS patients (top panel) and healthy UCB-MNCs (bottom panel) were stained with a FITC-labelled anti-human CD45 antibody immediately after isolation and prior to seeding in scaffolds. There are marked differences between independent MDS samples (top panel) both in population percentages and morphology (SSC-A levels), as opposed to the overall consistency of healthy UCB-MNC cell isolates (bottom panel). Variable blast percentages, as well as potential features of dysgranulopoiesis (dysplasia in the granulopoietic lineage) can be seen in MDS samples (top panel).

These cell isolates may not be fully comparable, as primary MDS cells are adult BM-derived and UCB-MNCs are of foetal origin. Furthermore, it is widely accepted that within MDS there is a high-degree of intra-disease as well as patient-to-patient variability. Additional experimentation is required with healthy, adult BM-MNCs to be able to make more appropriate comparisons with the primary MDS cultures, despite the expected sample-to-sample variability. Figure 41 further illustrates how complex and patient-specific MDS is, with a high degree of variability in the biological starting material. Going forward, developing creative, cutting-edge cell culture protocols is paramount to enable an accurate *in-vitro* biomimicry of the *in-vivo* environment – a possible avenue to explore in future experimentation would be to supplement cultures with patient-specific blood plasma, extracted from a sample of peripheral blood. Accurate biochemical characterization of such plasma is crucial, evaluating for example cytokine and nutrient levels, to be able to understand patient-specific behaviours and pathophysiology. Altogether, this highlights the huge degree of variability upstream of our culture system, stressing the need for increased analytical power (more primary MDS cells to perform independent replicates) and the need to use a high-fidelity control (healthy adult BM-MNCs).

Primary MDS cell isolates seemed to contain roughly 4-fold lower CD45⁺ CD36⁻ cells than UCB-MNCs (Figure 42, top-left panel). This could be explained by the different biological sources herein analysed – while the Healthy Control cells were isolated from umbilical cord blood, MDS MNCs were isolated from bone marrow samples. From a total mononuclear cell stand-point, the bone marrow will be poorer in non-committed leukocytes (CD45⁺ CD36⁻) and richer in, for example, erythroid-committed phenotypes (CD71⁺). In turn, leukocytes compose most of the mononuclear cell fraction of the peripheral blood, which justifies the 4-fold higher percentages of CD45⁺ CD36⁻ cells, when compared to the BM-derived MDS MNCs. This further illustrates the differences in cell sources upstream of the culture, stressing the need for using healthy BM controls in further experimentation.

Figures 42 and 43 show the evolution of the cellular phenotypes assessed during the culture of a representative healthy UCB experiment as well as two experiments with primary MDS cells. There is a high degree of variability in the expression levels of the phenotypes present, especially within MDS samples. UCB-MNC cultures seemed to slightly promote growth of CD235a⁻ CD71⁺ early erythroid progenitors, whereas MDS cultures seem to exhibit a decay of these progenitor numbers until Day 8, with phenotypic maintenance until the end of the culture (Figure 44, top-right panel). Remarkably, mid (CD235a⁺ CD71⁺) and late (CD235a⁺ CD71⁻) erythroid progenitors are maintained throughout the culture, as opposed to the decay seen in these phenotypes in UCB-MNC cultures. For primary MDS cultures, mid and late progenitors seem to peak at Day 8 and Day 16, respectively. The BM is far richer in stromal cells and HSPCs than UCB, which can potentially explain both the higher initial numbers of erythroid progenitors as well as their maintenance in culture. The latter may potentially be related to the endogenous production of EPO by BM-derived stromal cells – this hypothesis would have to be confirmed with additional experimentation.

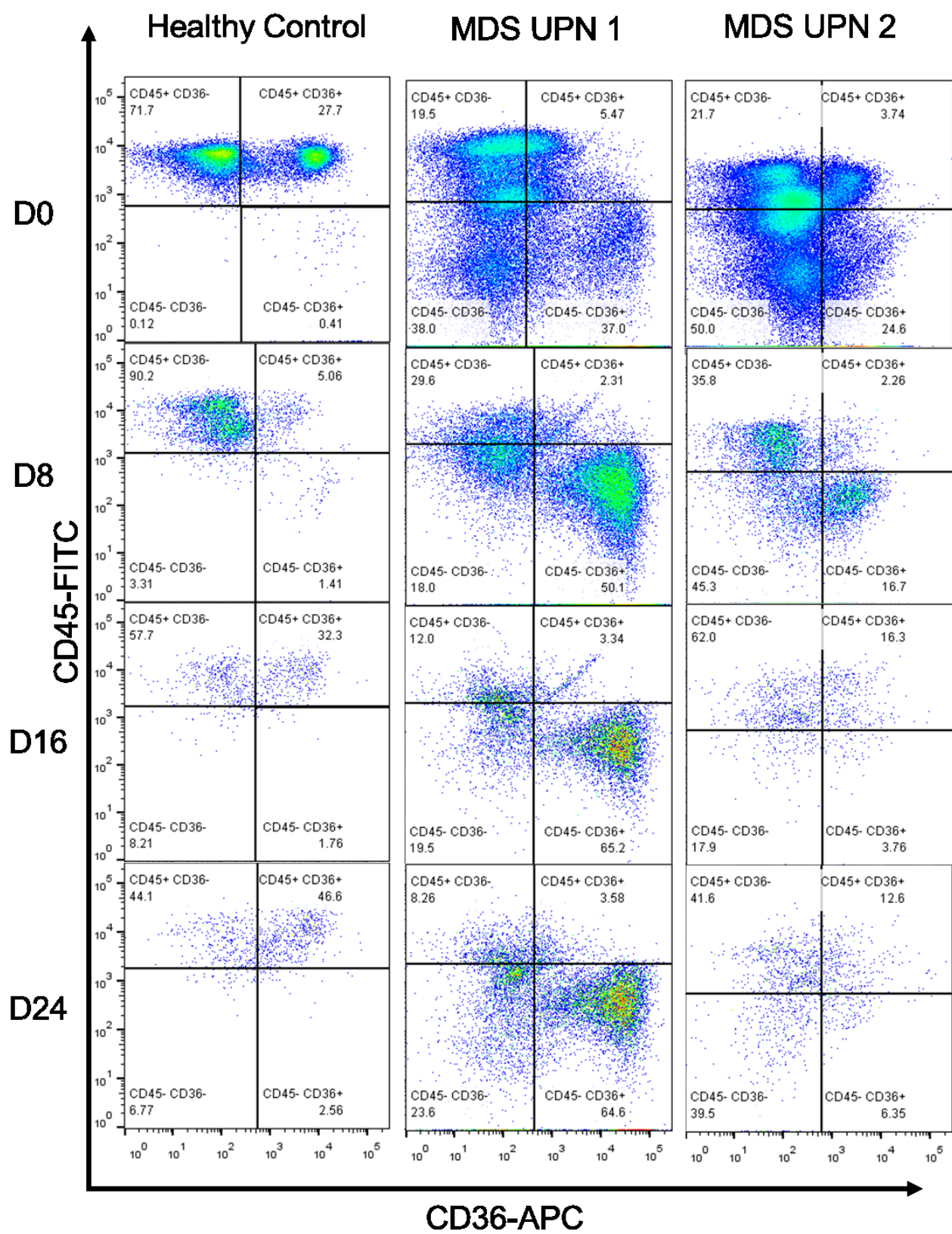


Figure 42 – Bivariate CD45/CD36 plots of UCB-MNC (healthy control) and BM-MNCs (MDS). Legend: **UPN** – Unique Patient Number.

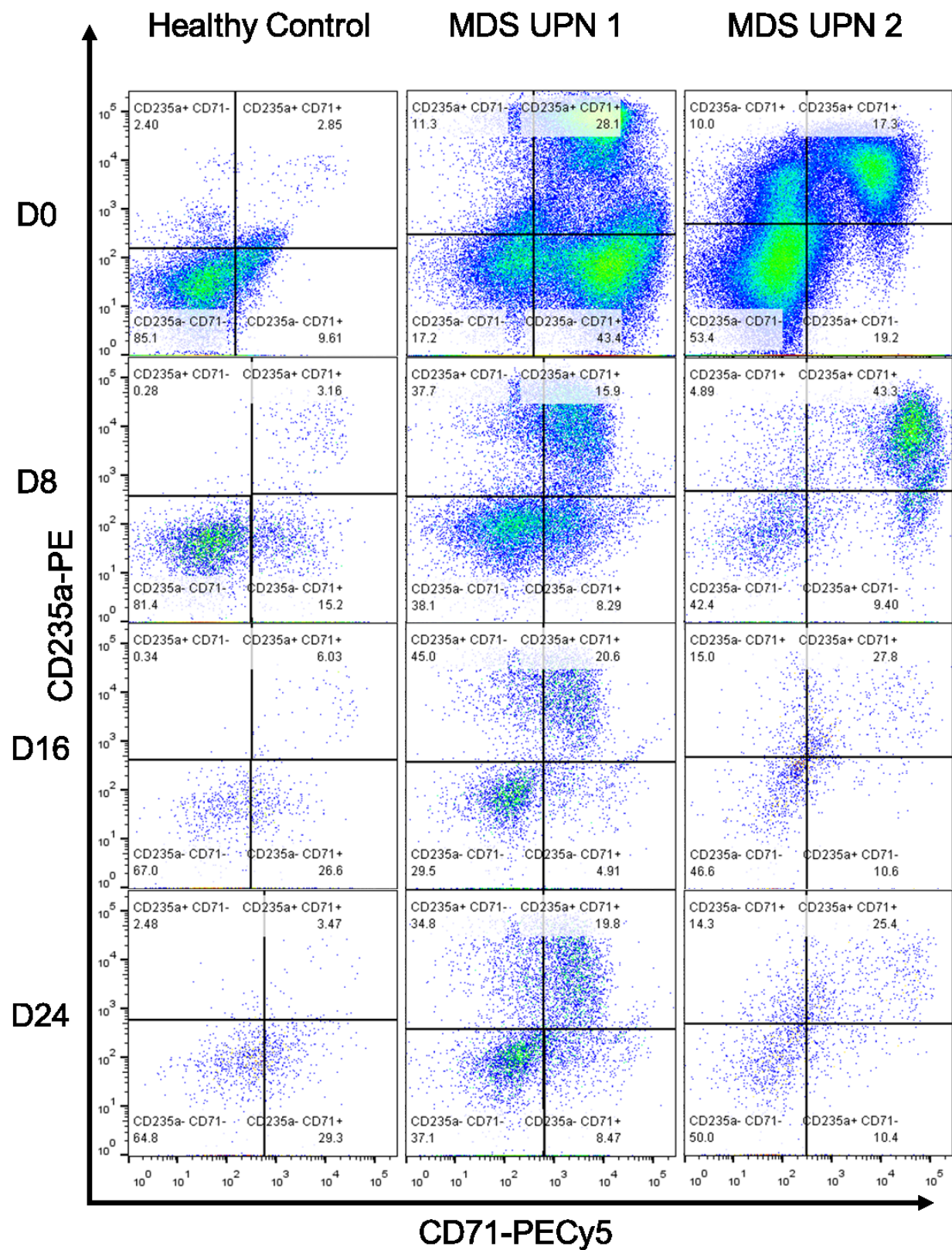


Figure 43 – Bivariate CD235a/CD71 plots of UCB-MNC (healthy control) and BM-MNCs (MDS). Legend: **UPN** – Unique Patient Number.

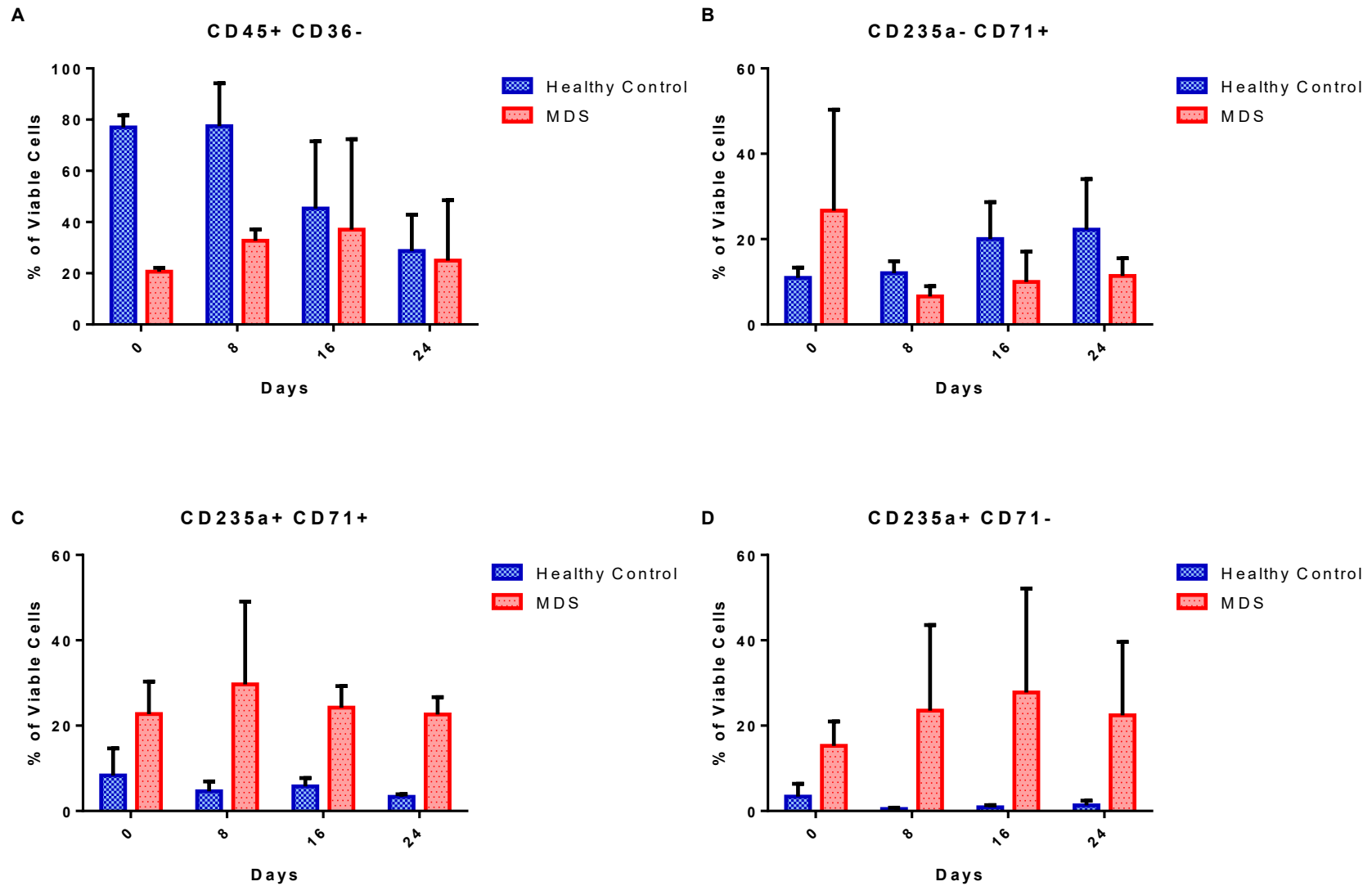


Figure 44 – Phenotypical evolution, throughout the culture, of pan-myeloid (A) as well as early (B), mid (C) and late (D) erythroid progenitors, expressed as a percentage of total viable cells. Results are shown as average $\pm$ SD of N=3 (healthy control) and N=2 (MDS).

The work herein, despite being promising, is at the proof-of-concept stage, requiring further experimentation, validation and optimisation. We have demonstrated that healthy UCB-MNC controls can be cultured for 24 days without the addition of any exogenous cytokines. Upon selection of the optimal seeding density and scaffold coating, we have then proceeded to culture 3 independent primary MDS samples. Herein we have shown that primary MDS cells can be maintained in 3D long-term cultures, with high cell viability throughout. We have optimised culture conditions and cell seeding densities based on healthy controls, keeping the fine balance between the limited amount of diseased raw material obtained from patient samples and the multiple analyses performed. Unfortunately, the vast majority of the analyses performed lack statistical significance, as they were performed just with two independent cell sources given the difficulties with obtaining patient samples.

Previous work has also stressed the need for developing robust 3D systems that enable long-term cell culture of haematopoietic stem and progenitor cells. While a conventional 2D cell culture system was unable to sustain cell cultures, leading to total cell death 14 days after seeding, several 3D cell culture systems enabled cell culture and growth up to 28 days, with cell viability not dropping below 75% (Mortera-Blanco T 2011). Recent unpublished results obtained in our group show that using collagen coating (versus uncoated scaffolds) does not lead to a significant difference in cell growth, viability and maintenance when culturing cancer cells derived from primary patient samples. Also, the collagen used comes from calf skin, which potentially can interfere with the culture. The isolated UCB-MNCs are rich in haematopoietic stem and progenitor cells which do not require collagen or other ECM proteins to proliferate – as they are typically cultured in liquid suspension. This might partially justify why coating seems to have no influence in the analysed cell culture outcomes.

As shown in Figures 42 and 43, there is a great degree of variability within cultured MDS samples, which starts right from the cell isolation step. The main bottlenecks of primary MDS cultures are

sample scarcity as well as low cellular yields obtained from each sample. To tackle both the intra-disease, patient-to-patient variability as well as sample scarcity and low cellular yields, one requires a relatively high (at least 10-15) number of independent primary MDS cell donors to be able to try to derive clinically-meaningful conclusions.

To provide a more accurate comparison, age and co-morbidity matched healthy controls would need to be procured, cultured and analysed using similar protocols. While frozen BM MNCs are widely available from well-known biological suppliers, the extent to which cryo-freezing affects sample integrity and quality – when compared to fresh BM – is unknown and could potentially skew the experimental results. The procurement of fresh, healthy bone marrow samples is extremely complicated, which could be aided by forming a partnership with an Orthopaedic Department is established – more often than not, during surgical procedures, orthopaedists can easily access the bone marrow, from which they could potentially sample.

The work herein described demonstrates that it is indeed possible to successfully culture MDS cells *in-vitro* for 24 days, with high viability (>75%) throughout the whole culture, confirmed both *in situ* (Confocal Microscopy) and using the supernatant cell viability as a probe for overall culture fitness. Terminal erythroid phenotypes (CD71⁺ CD235a⁺) were observed in the supernatant of both healthy and diseased cell cultures, and our hypothesis is that this is due to endogenous EPO production within the cultured cells. Further experimentation is crucial to validate this hypothesis and to explore other avenues, as described in the following chapter.

5.3. Future work

This chapter requires future work following two different avenues. There is the need to work with samples from **healthy controls**, ensuring that we repeat the experiments using age-, co-morbidity-matched adult BM-derived cells, which is a high-fidelity healthy control. Secondly, there is the need for further procurement of **primary MDS samples**, which can potentially help to establish and validate this novel, long-term, patient-representative 3D cell culture system. However, most of the work concerns both healthy and diseased cell sources, as the culture methodologies employed and analyses performed are the same.

Healthy Controls

1) Repeat the experiments with high-fidelity adult BM cells. All the work performed with healthy controls needs to be replicated using adult BM cells. These results will then be compared to the ones obtained with UCB-MNCs to check for agreement. It is expected that the results will differ greatly at the phenotype level, but potentially not at the cell expansion and growth level. The latter are potentially more related to medium composition and nutrient availability rather than cell source.

Primary diseased cells

1) Obtain more independent primary MDS samples, ideally from patients with different disease risks. Obtaining more independent primary MDS samples will allow for more independent experiments, allowing to extract meaningful, statistically-significant results from diseased cell cultures. No statistical analysis was performed in Chapter 5.2.2 due to the low number of independent replicates (N=3, with N=2 in most of the analyses). The MDS samples cultured were all of medium-risk, so it would be definitely interesting and valuable to procure samples of both low-risk and high-risk MDS, comparing their culture outcomes. Being able to sub-

classify culture results based on disease risk can reveal interesting conclusions that will directly be applicable to the clinical practice.

Healthy and diseased cell cultures

1) Implement serum-free medium Serum/xeno-free media are always preferred, as their chemical composition is well-defined, eliminating batch-to-batch variability. This can potentially lead to more reproducible results and ultimately open the avenue for clinically-meaningful results, paving the way for a 3D cell culture system that truly represents the physiological *in-vivo* scenario.

2) *In-situ* confocal microscopy searching for cell phenotypes. Similar to the work done to assess *in-situ* viability, confocal microscopy can be used to detect up to 4 cellular phenotypes per scaffold slice (Allenby MC 2019), allowing for the simultaneous assessment of abundance and spatial distribution of different phenotypes. This could help unveil novel cell-cell interactions/phenotypes that are differentially expressed in healthy and diseased cell cultures.

3) Flow cytometry analysis of extracted cells. The flow cytometry work herein described relies on the assumption that cells present on the outside of the 3D scaffold are representative, at that time-point, of what is happening inside the scaffold. That assumption needs to be validated using alternative approaches. While confocal microscopy can help, flow cytometric analyses of cells extracted from the scaffold is a faster and more robust alternative, when researchers are purely interested in cellular phenotype. Mechanically and chemically (with a trypsin-analogue like TrypLE) extracting cells from the scaffold will enable multicolour flow cytometry analyses of the cellular phenotypes present within the scaffold. While optimising this protocol, one needs to ensure that a minimum residual number of cells, if any, remains within the scaffold after the extraction. This can be confirmed by various methods, but using PicoGreen as done in this thesis is a viable option.

4) Medium composition analysis. Techniques such as HPLC can potentially help to understand which nutrients are preferentially used by the cells, which may later be subjected to chemically-defined medium design and optimisation. On the other hand, measuring toxic metabolites like ammonia and lactate using a Bioprofiler can give further indication of the culture's wellbeing.

5) Intracellular metabolomics. Using the techniques described in Chapter 6, one could potentially assess intracellular metabolites throughout the culture. Examination of the metabolic differences between healthy UCB-MNCs, healthy BM-MNCs and primary MDS cells in culture has not been attempted at of the time of writing, and could potentially lead to a high-impact publication.

6. Plasma Metabolomics

Several groups have applied omics techniques to gain deeper insights into the pathophysiology of MDS. As previously mentioned, current research in the field made good use of multi-omics techniques, such as genomics (Schwartz JR 2017, Nassereddine S 2017 and Ogawa 2019), transcriptomics (Gerstung M 2015, Shiozawa Y 2017 and Im H 2018) and proteomics (Aivado M 2006, Majek P 2014 and Ragon BK 2016), but limited work has focused metabolomics (Khalid A 2018).

DNA is transcribed into RNA which is then translated into proteins. Some of these proteins, called enzymes, are involved in a series of chemical (metabolic) reactions and pathways that are vital for the survival of every human cell. Some of the most important, well-known metabolic pathways are the TCA cycle, Oxidative Phosphorylation (OxPhos) and Glycolysis. Metabolites are therefore the end product of these metabolic pathways, and act as regulatory molecules in some of these chemical reactions. ATP and pyruvate, for example, are involved in glycolysis but tightly regulate the metabolic flux through that pathway (Berg JM 2002). To date, it is relatively unknown the extent of how the healthy metabolome is affected in MDS patients. Moreover, proteins and metabolites are usually good candidates for disease-specific biomarkers. A remarkable example is the Prostate-specific antigen (PSA), a protein overexpressed by patients with prostate disorders and cancer (Saini 2016). Research in fields as diverse as early Lyme disease (Pegalajar-Jurado A 2018), hepatocellular carcinoma (Tan W 2018) and AML (Chen WL 2014) have identified candidate biomarkers for intra-disease classification, stratification and diagnosis.

There is an unmet research need for work in MDS using metabolomics that could potentially unveil candidate biomarkers for disease sub-group and risk stratification, as well as for leukaemia progression. In the particular case of haematological malignancies, performing a bone marrow biopsy is the gold standard for diagnosis. Identifying such candidate biomarkers in easily-

accessible, alternative body fluids such as peripheral blood or urine could potentially change the paradigm and replace the need to perform the current gold standard, and often painful procedure such as a bone marrow biopsy.

6.1. Introduction

The work herein described paves the way for a novel approach using GC-MS metabolomics for the characterization of bone marrow and peripheral blood plasma of MDS patients. We have obtained a total of 23 samples from 19 unique patients, divided between 14 peripheral blood samples and 9 bone marrow samples. Out of the 23 samples, we have managed to simultaneously collect both peripheral blood and bone marrow samples from four unique patients (“paired samples”).

Out of the 23 samples obtained, we have managed to procure samples from different haematological conditions, such as MDS (4 BM samples, 10 PB samples), CMML-1 (1 BM, 2 PB), ET (1 BM, 1 PB), CML (2 BM, 1 PB) and Clonal CD8⁺ T Cell Expansion (1 BM). We have analysed the metabolomics data using multivariate statistical techniques such as Principal Component Analysis (PCA) and Hierarchical Clustering (HCI).

Metabolite peaks were identified based on their retention time and marker ion pattern, according to previously published methods (Kanani HH 2007 and Kanani HH 2008) and in-house created methods. Peak areas were quantified based on the marker ion peak areas and normalized (relative peak areas, RPA) to that of the internal standard (ribitol, marker ion m/z 103). Glucose labelled with heavy carbon atoms (carbon-13) – glucose-C-13 – was added as a second internal standard, to detect intra and inter-sample deviations due to unexpected, unrequested changes in operating conditions of the equipment during or between sample runs. Within the analysed dataset, samples that had an absolute variation of the RPAs of the glucose-C-13 peaks greater than 25% - accounting for less than 5% of all samples and runs analysed – were re-run. Metabolite

identification was performed using an in-house custom made method, which was able to detect 121 distinct metabolite peaks (82 known, 39 unknown peaks, including >1 trimethylsilyl derivatives of the metabolites containing at least one amine group and two methoxime derivatives of ketone-containing molecules). Peaks with a signal-to-noise ratio of less than 3 were automatically filtered out of the analysis. Peaks with an absolute coefficient of variation higher than 25% between injections (n=4) were removed from further analyses.

For multivariate statistical analyses, standardized peak areas were used. For every unique sample, these were computed by subtracting the mean RPA of each metabolite (across all samples) and dividing by its standard deviation (across all samples). Unsupervised Principal Component Analysis (PCA) was performed in R using scripts created in-house (Appendix IV). This multivariate statistical technique seeks to increase interpretability of datasets by reducing the dimensionality – herein, dimensionality being the number of unique features observed – and minimising information loss. It does so by proposing a new set of variables – or Principal Components – that are uncorrelated to each other (Jolliffe IT 2016).

All of the four technical replicates of each biological sample analysed were used to generate the PCA plots below. Broadly speaking, PCA clusters together samples that are similar and further away samples that are dissimilar.

6.2. Results and Discussion

6.2.1. Bone Marrow

We have obtained 9 bone marrow samples from 5 different haematological malignancies. PCA results are shown in 2D plots for the first three Principal Components (PC), which account for 67.42% of the total variance. Circled regions are discussed in detail below.

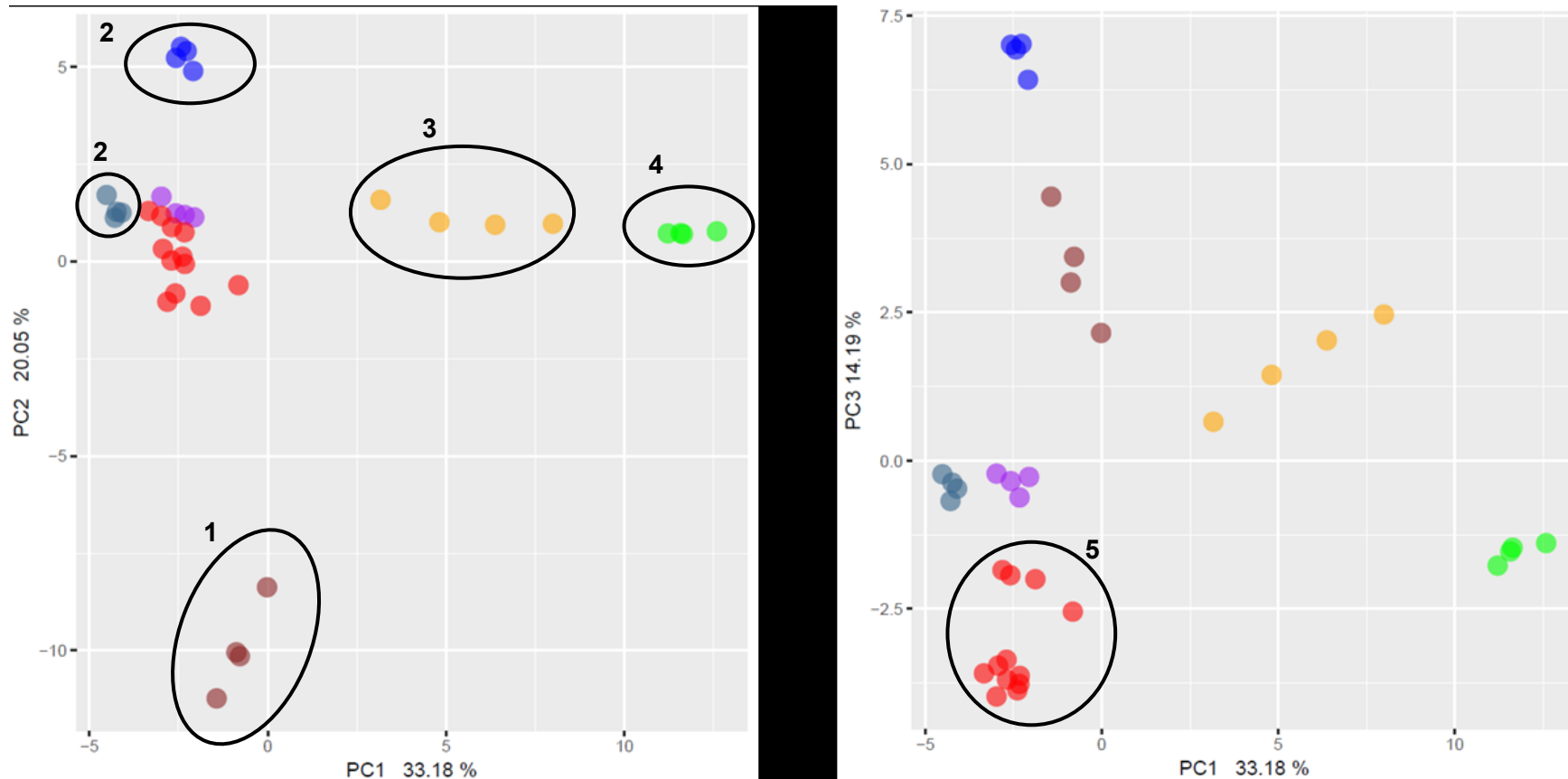


Figure 45 – 2D PCA analysis of BM samples. Every sample is obtained in four technical replicates. The colours indicate different characteristics/haematological conditions: **Red** – MDS; **Brown** – MDS pre-treated with folic acid; **Orange** – CMML-1; **Green** – ET; **Dark Blue** – CML at blast phase; **Navy Blue** – CML at restaging; **Purple** – Clonal CD8⁺ T Cell expansion. Meaningful groups discussed in the text below are circled and numerically identified in black.

The three first PCs account for 67.42% of the total variance between these samples. Within this dataset, we analysed BM-derived plasma of 3x MDS patients (**red**), 1x MDS patient pre-treated with folic acid (**brown**), 1x patient with CMML-1 (**orange**), 1x patient with Essential Thrombocythemia (**green**), 1x patient with CML (**dark blue**), 1x patient with CML at restaging (**navy blue**) and 1x patient with Clonal CD8⁺ T Cell expansion (**purple**).

Region 1 – PCA can capture pre-existing therapeutic intervention

This patient was admitted into hospital for further examination after having mild peripheral blood cytopenias, and at the time of sample procurement had been taking folic acid supplements for some weeks, which is a known BM stimulant (NHS 2019). Interestingly, the unsupervised PCA is able to capture differences between patients with untreated MDS (in **red**) and patients with MDS who already had some form of clinical intervention (in **brown**).

Region 2 – CML sub-group classifications are discretized by PCA

CML is a slow-growing form of myeloid leukaemia. CML is commonly divided into two main phases: the chronic phase and the blast phase, also called blast crisis. About 90% of the patients are at the slow, chronic phase when diagnosed with CML (Cancer.net 2018) and their peripheral blood and bone marrow findings often resemble some biological features of MDS. CML patients can have WBC counts up to 50-fold higher than normal values (ASH 2008). The disease can then evolve to a blast phase, where controlling WBC numbers becomes challenging. At this stage, the disease shows characteristics often seen in AML.

The **dark blue** sample was obtained at the diagnosis of a patient which was later confirmed to have CML at the blast phase. Conversely, the **navy blue** sample was obtained 3 months after an original CML-at-blast-phase diagnosis, which was confirmed to have reverted back to the chronic phase. Remarkably, the PCA can capture the differences between these two disease sub-groups by clustering them separately, with further separation seen on the PC1-PC3 plot (Figure 45, right).

Moreover, CML at the blast phase is known to have slow-growing, MDS-like features. This is confirmed by the PCA by clustering this patient sample closer with the MDS samples.

Regions 3 and 4 – Different haematological malignancies are clustered further away

CMML-1 is a rare form of leukaemia, characterized by the over-production of monocytes. Despite sharing common features with MDS and myeloproliferative neoplasms (MPNs), CMML-1 is often treated differently than these two pathologies, as patients can have low (CMML-1) or high (CMML-2) blast numbers, as well as dysplastic (MDS-like) or proliferative (MPN-like) disease forms (Bloodwise 2017). Essential Thrombocythemia (ET) is a MPN-like condition characterized by the excessive production of platelets (NIH, 2019).

Both CMML-1 (in **orange**) and ET (in **green**) are clustered further away from every other sample, being clearly isolated in the PC1-PC2 plot (Figure 45, left). On the PC1-PC3 plot, they are even further apart. This shows the capability of PCA to discretize different haematological malignancies purely based on metabolic features.

Region 5 – Low-resolution when discretizing MDS sub-groups

When using PCA, discretizing between different MDS sub-groups seems to be difficult. This may be mainly because of the presence of many different haematological malignancies in the same 2D PCA plot, which are the driving force for the separation, keeping in mind that PCA clusters samples together that are similar and further away samples that are dissimilar. For intra-MDS characterization one would need to increase sample size as well as removing other pathologies from the analysis. Interestingly, the sample coming from a patient diagnosed with Clonal CD8⁺ T Cell expansion (in **purple**) clusters together with the MDS samples. Being a lymphoid-related malignancy, this may mean that the myeloid features of this patient resemble that of MDS.

6.2.2. Peripheral Blood

We have obtained 14 peripheral blood samples from 4 different haematological malignancies. PCA results are shown in 2D plots for the first three Principal Components (PC), which account for 69.85% of the total variance. Circled regions are discussed in detail below.

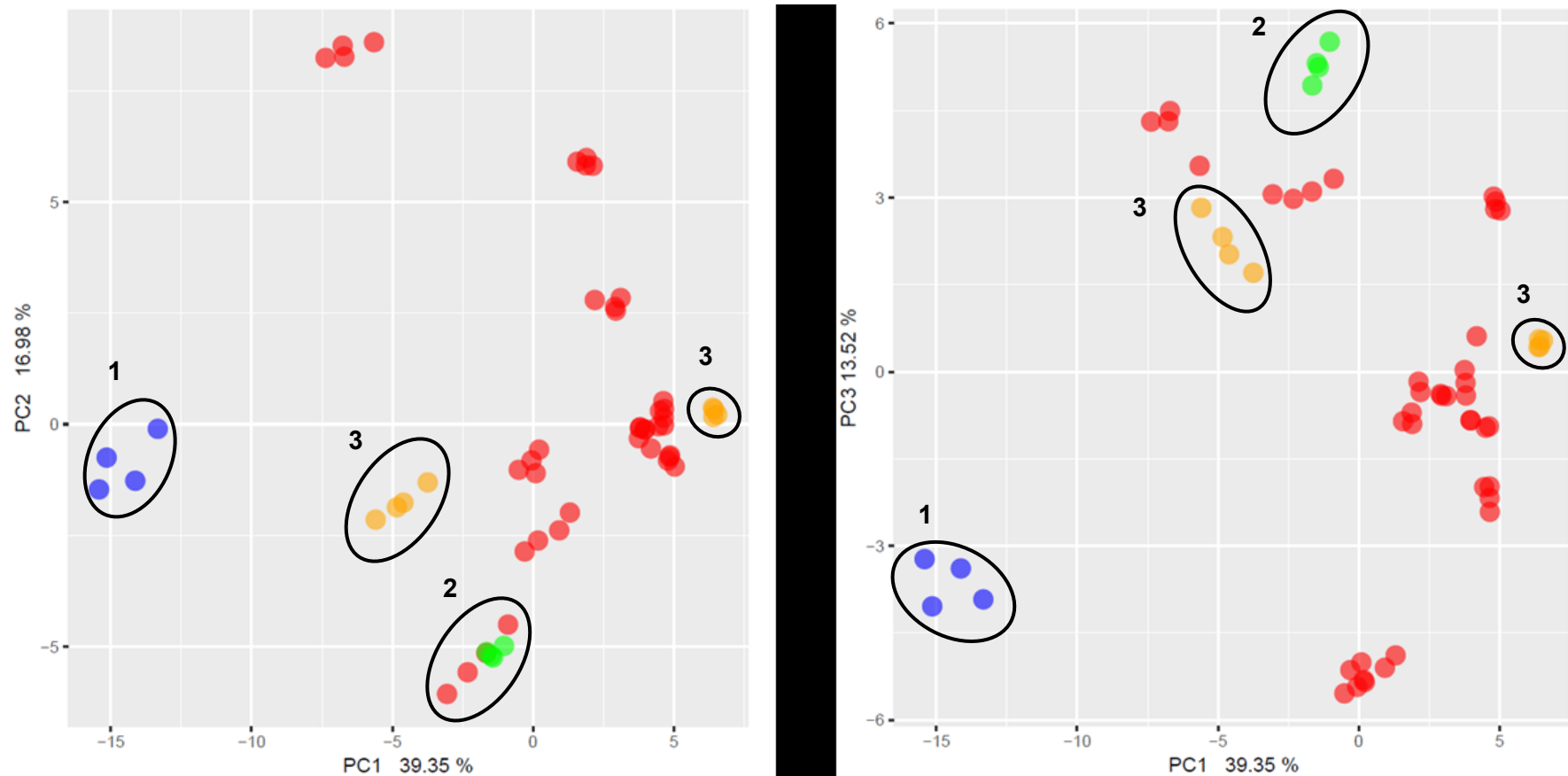


Figure 46 – 2D PCA analysis of PB samples. Every sample is obtained in four technical replicates. The colours indicate different characteristics/haematological conditions: **Red** – MDS; **Orange** – CMML-1; **Green** – ET; **Dark Blue** – CML at blast phase. Meaningful groups discussed in the text below are circled and numerically identified in black.

The three first PCs account for 69.85% of the total variance between these samples. Within this dataset, we analysed plasma of 10x MDS patients (**red**), 2x patients with CMML-1 (**orange**), 1x patient with ET (**green**) and 1x patient with CML(**dark blue**).

Region 1 – CML is clustered further from any other haematological malignancy

Similar to what had been observed in with the BM samples, CML (at blast phase, in **dark blue**) clusters further away from any other haematological malignancies. This further highlights the ability of PCA to, in an unsupervised fashion, discretize different haematological malignancies. Interestingly, the same separation was seen both in BM and PB samples, which may indicate that features present in the PB alone might be sufficient to separate CML from other pathologies. Within CML, other work in the field has found differences, at the metabolome level, between therapy-responding and therapy-resistant patients (Sweet K 2013 and Karlíková R 2016) as well as between CML patients and healthy controls (Yang B 2018). Taken altogether, this opens up an interesting avenue for further experimental work towards proposing a well-defined set of biomarkers that are CML-specific, easily accessible in the peripheral blood and ultimately that could initially complement and potentially replace the need for a bone marrow aspiration.

Region 2 – ET clusters with MDS on PC1-PC2, but is further separated on PC1-PC3

The ET sample clusters together with an MDS on the PC1-PC2 plot, but is separated in the PC1-PC3 plot. This highlights the importance of investigating the three first principal components when using PCA, to ensure maximization of the extracted information.

Region 3 – CMML-1 is clearly identifiable

PCA is able to clearly discretize CMML-1 from any other haematological malignancy. Interestingly, one of the CMML-1 samples is further away from ET both on PC1-PC2 and PC1-PC3 plots, as opposed to what we had seen in the BM PCA analysis. The three CMML-1 samples (1x BM, 2x PB) are from three independent patients, which can have different disease sub-forms

(e.g. MDS- or MPN-like). Potentially, CMML-1 samples clustering closer to ET are likely to be MPN-like – this would be a very interesting assumption to validate.

Despite the close-clustering of some MDS samples, Figure 46 clearly highlights the intrinsic variability present in MDS patients. It is worth noting that PB plasma is greatly affected by dietary and lifestyle factors, as well as any pre-existing condition or medication (e.g. diabetes). While analysing the plasma of patients with haematological malignancies, it is often complicated to distinguish between the disease consequences and lifestyle and patient-specific factors, even for bone marrow samples. For the correct proposal and *in-vitro* validation of candidate biomarkers, a larger sample number would be required.

Our methods were able to capture differences within patients with MDS that, at the time of disease diagnosis, were already undergoing a pre-existing therapeutical intervention.- Remarkably, we were able to discretize one patient that was taking folic acid, a known BM stimulant, prior to their bone marrow aspirate. The clear distinction between different haematological malignancies is remarkable; however, it seems that the PCA alone is not able to perform intra-MDS disease discretization. Greater sample numbers are required to validate all of these assumptions.

Out of the 23 samples, we have managed to simultaneously collect both peripheral blood and bone marrow samples from four unique patients (“paired samples”). Unfortunately, the low number of paired samples does not allow for any statistically relevant analyses. With greater numbers of paired samples, it would be interesting to assess PCA clustering patterns – for example, will samples from the same patient/tissue cluster together?

It is important to note that the outcome of these unsupervised statistical techniques are heavily dependent on the original dataset provided. PCA, for example, may be able to distinguish between two very different haematological malignancies (MDS vs CML), but potentially does not have the resolution to distinguish between MDS sub-types in the same PCA plots. However, if we

remove the CML samples from the original dataset and re-do the analysis, PCA might be able to further discretize samples from different MDS sub-types and risk groups. Here the decision was made to analyse all the samples from all haematological malignancies in a set of PCA plots, as finding differences between these pathologies, at the metabolomic level, is of great scientific interest.

Comparing the results obtained herein with healthy controls published elsewhere is, unfortunately, not possible. The GC-MS method used relies on relative peak areas, which measure a metabolite's signal intensity (area under their marker ion curve) against that of an internal standard (ribitol), being therefore a relative, rather than absolute, quantitative method. It would be extremely interesting to procure healthy bone marrow and peripheral blood plasma samples and run them using the same *in-vitro* and *in-silico* methodologies. This is also crucial for the clinical applicability of the findings herein described, as the ultimate goal of this chapter would be to provide a set of candidate disease biomarkers for rapid, first-line disease diagnostic.

6.3. Future Work

In this chapter, the main challenge is sample procurement. Currently, it seems to be possible to discretize between different haematological conditions purely using either BM or PB samples, despite with limited clinical applicability given the lack of biological samples. However, for intra-MDS sub-group discretization, this does not seem possible nor clear.

1) Increase sample number and number of multivariate analyses. Procuring and adding more samples of both BM and PB will help to further discretize between diseases, simultaneously accounting for more patient-to-patient variability and allowing for additional analyses. A collaboration with colleagues from NTUA using Linear Discriminant Analysis (LDA), using PB data, that these haematological malignancies could be discretized by the differential expression of four different metabolites (data not shown), but this analysis requires at least 15-20-fold higher sample numbers to ensure sufficient statistical power (Gromski PS 2014). Techniques like Significance Analysis of Microarrays (SAM), that correct for the false detection rate (FDR) in multivariate statistics, also require greater sample numbers (Tusher VG 2001).

2) Procure paired samples. To achieve the ultimate goal to propose potential candidate biomarkers for disease identification, one will require paired samples to do so. If the desired disease discretization is seen both at the PB level and confirmed at the BM level, then PB alone may suffice to confirm the disease. Then, LDA or SAM techniques could potentially indicate a set of metabolites that specifically and independently characterize a haematological malignancy.

3) Healthy controls. To be able to determine the differences between a healthy individual and a cancer patient, one will need to procure samples from healthy controls. These have to be adult, age and co-morbidity matched healthy donors. However, we predict that it would be very difficult to procure healthy BM samples, unless a collaboration with an Orthopaedic Department, for

instance, could be made (i.e. during orthopaedic surgeries, there is often free access to the bone marrow, where a sample could then be taken).

4) Increase the number and diversity of metabolic features

Our current metabolomics method is able to capture roughly 50 unique metabolites, being limited to the polar fraction of the metabolome and to molecules smaller than cholesterol. Analysing the non-polar metabolome (lipidomics) could potentially help to further discretize these haematological malignancies, as this would increase the total number of features considered in the analysis. Also, alternative metabolomics techniques (GC-MS/MS, LC-MS, GC-TOF) could dramatically increase the number of samples, positively affecting the sensitivity and specificity of these techniques.

7. Key Achievements and Future Developments

The work herein paves the way for frameworks integrating *in-vitro* and *in-silico* modelling of haematological malignancies.

Starting with the erythroid lineage as a case-study, we have developed high-fidelity *in-silico* mathematical model from first principles that considers several cellular intermediates, integrates cytokine signaling and its cell-specific effect on the erythroid differentiation pathway. Upon successful experimental validation, the proposed model can be then expanded to comprise not only other blood lineages (e.g. granulocytic or megakaryocytic) as well as fully-validated competing (diseased) cellular population dynamics that greatly impairs the differentiation and expansion of the existing healthy population. The major limitation of having such a detail model can be its experimental validation, as it will require measuring very specific signal intensities, cell populations and differentiation dynamics.

We have also proposed a framework for the long-term, cytokine-free 3D culture of primary MDS cells. This work is in line with our lab's approach towards similar goals using primary AML and CLL cells. The proposed culture system needs to be expanded to serum-free conditions in an attempt to better model the *in-vivo* scenario while avoiding serum batch-to-batch variability. Looking at patient-specific phenotype conservation throughout the culture (e.g. clonality, transcriptome, proteome and metabolome) is also of extreme relevance, as these findings could generate further insights into the pathophysiology of MDS. Simultaneously, this system can potentially enable *in-vitro* chemotherapeutic studies, looking into ways to optimize drug dosage and administration. Coupling these findings with recent *in-silico* models developed by our group (Quiroga-Campano AL 2017) will truly shorten the bench-to-bedside timescale, generating a positive impact in the clinical management and treatment of haemato-oncological patients.

Lastly, we have proposed a framework to characterize the metabolome of patients with haematological malignancies. In future, procuring more patient samples will confer statistical power to the analyses performed, while accounting for patient-to-patient and intra-disease variability. With clinically relevant sample numbers, unsupervised and supervised multivariate statistical methods can potentially propose candidate biomarkers, not only for disease diagnostics (e.g. healthy vs diseased) but also for intra-disease discretization (e.g. MDS sub-groups, risk score). With a set of clinically-relevant disease biomarkers, clinicians could potentially do quicker first-line disease diagnosis and risk-stratification as well as deploying first-line supportive measures, while waiting for confirmatory bone marrow biopsy results. This would have an extremely positive impact on the lives of patients battling against haematological malignancies.

References

- A Ganser, G Seipelt, A Lindemann, OG Ottmann, S Falk, M Eder, F Herrmann, R Becher, K Hoffken and T Buchner. 1990. "Effects of recombinant human interleukin-3 in patients with myelodysplastic syndromes." *Blood* 76:455.
- Abarrategi A, Mian SA, Passaro D, Rouault-Pierre K et al. 2018. "Modeling the human bone marrow niche in mice: From host bone marrow engraftment to bioengineering approaches." *J Exp Med* Vol 215, 729.
- ACS. 2017. *Chronic myelomonocytic leukemia*. Accessed April 2019.
<https://www.cancer.org/cancer/chronic-myelomonocytic-leukemia/about/what-is-chronic-myelomonocytic.html>.
- Afenya E, Mundle S. 2010. "Hematologic disorders and bone marrow - peripheral blood dynamics." *Math Model Nat Phenom* Vol 5, No 3, 5-15.
- Afenya EK, Ouifki R, Camara BI, Mundle SD. 2016. "Mathematical modeling of bone marrow - peripheral blood dynamics in the disease state based on current emerging paradigms, part I." *Math Bio* Vol 274, 83-93.
- Aivado M, Spentzos D, Germing U, Alterovitz G et al. 2006. "Serum proteome profiling detects myelodysplastic syndromes and identifies CXC chemokine ligands 4 and 7 as markers for advanced disease." *Proc Nat Ac Sci USA* Vol 104, 1307-1312.
- Akker EVD, Satchwell TJ, Pellegrin S, Daniels G, Toye AM. 2010. "The Majority Of The In Vitro Erythroid Expansion Potential Resides In CD34- Cells, Outweighing The Contribution Of CD34+ Cells And Significantly Increasing The Erythroblast Yield From Peripheral Blood Samples." *Haematologica* 95:1594.
- Alfred LJ, Wojdani A, Nieto M, Perez R et al. 1983. "A chemical carcinogen, 3-methylcholanthrene, alters T-cell function and induces T-suppressor cells in a mouse model system." *Immunology* Vol 50, 207-213.
- Allampallam K, Shetty V, Hussaini S, Mazzoran L et al. 1999. "Measurement of mRNA expression for a variety of cytokines and its receptors in bone marrows of patients with myelodysplastic syndromes." *Anticancer Res* Vol 19, 5323-5328.
- Allenby MC, Panoskaltsis N, Tahlawi A, Dos Santos SB et al. 2019. "Dynamic human erythropoiesis in a three-dimensional perfusion bone marrow biomimicry." *Biomaterials* Vol 188, 24-37.
- Andersen LK, Mackey MC. 2001. "Resonance in periodic chemotherapy: a case study of acute myelogenous leukemia." *J Theor Biol* Vol 209, 113-130.
- Andersen MK, Johansson B, Larsen SO, Pedersen-Bjergaard J. 1998. "Chromosomal abnormalities in secondary MDS and AML. Relationship to drugs and radiation with specific emphasis on the balanced rearrangements." *Haematologica* Vol 83, 483-488.
- Anthony BA, Link DC. 2014. "Regulation of hematopoietic stem cells by bone marrow stromal cells." *Trends Immunol* Vol 35, 32-37.

- Arakawa T, Yphantis DA, Lary JW, Narhi LO, Lu HS, Prestrelski SJ, Clogston CL, Zsebo KM, Mendiaz EA, Wypych J, Langley KE. 1991. "Glycosylated and unglycosylated recombinant-derived human stem cell factors are dimeric and have extensive regular secondary structure." *J Biol Chem* 266:18942.
- Arber DA, Orazi A, Hasserjian R, Thiele J et al. 2016. "The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia." *Blood* Vol 127, 2391-2405.
- ASH. 2008. *Targeted therapy for chronic myeloid leukemia*. Dec. Accessed Apr 2019. <https://www.hematology.org/About/History/50-Years/1529.aspx>.
- Ashman LK, Cambareri AC, To LB, Levinsky RJ, Juttner CA. 1991. "Expression of the YB5.B8 antigen (c-kit proto-oncogene product) in normal human bone marrow." *Blood* 78:30.
- Aul C, Arning M, Runde V, Schneider W. 1991. "Serum erythropoietin concentrations in patients with myelodysplastic syndromes." *Leuk Res* 15:571.
- Barreyro L, Will B, Bartholdy B, Zhou L et al. 2012. "Overexpression of IL-1 receptor accessory protein in stem and progenitor cells and outcome correlation in AML and MDS." *Blood* Vol 120, 1290-1298.
- Bast L, Buck MC, Hecker JS, Oostendorp RAJ, Gotze KS, Marr C. 2021. "Computational modeling of stem and progenitor cell kinetics identifies plausible hematopoietic lineage hierarchies." *iScience* Vol 24, 102120.
- Becker V, Schilling M, Bachmann J, Baumann U, Raue A, Maiwald T, Timmer J, Klingmüller U. 2010. "Covering a broad dynamic range: information processing at the erythropoietin receptor." *Science* 328:1404.
- Bejar R, Steensma DP. 2014. "Recent developments in myelodysplastic syndromes." *Blood* Vol 124, 2793-2803.
- Bell, CR. 1990. "Growth of *Agrobacterium tumefaciens* under octopine limitation in chemostats." *Appl Environ Microbiol* 56:1775.
- Bennett, JM. 2005. "A comparative review of classification systems in myelodysplastic syndromes (MDS)." *Semin Oncol* Vol 32, S3-10.
- Berg JM, Tymoczko JL, Stryer L. 2002. "The glycolytic pathway is tightly controlled." In *Biochemistry*, by Tymoczko JL, Stryer L Berg JM. New York: WH Freeman.
- Bernasconi, P. 2008. "Molecular pathways in myelodysplastic syndromes and acute myeloid leukemia: relationships and distinctions-a review." *Br J Haematol* Vol 142, 695-708.
- Bertuccio SN, Anselmi L, Masetti R, Lonetti A et al. 2021. "Exploiting Clonal Evolution to Improve the Diagnosis and Treatment Efficacy Prediction in Pediatric AML." *Cancers* Vol 13, 1995.
- Beru N, McDonald J, Lacombe C, Goldwasser E. 1986. "Expression of the erythropoietin gene." *Mol Cell Biol* 6:2571.

- Bidarra SJ, Barrias CC. 2019. "3D Culture of Mesenchymal Stem Cells in Alginate Hydrogels." *Meth Mol Biol* Vol 2002, 165-180.
- Bincoletto C, Saad ST, Soares da Silva E, Queiroz ML. 1998. "Autonomous proliferation and bcl-2 expression involving haematopoietic cells in patients with myelodysplastic syndrome." *Br J Cancer* Vol 78, 621-624.
- Bloodwise. 2017. *Chronic myelomonocytic leukaemia (CMML)*. Jan. Accessed Apr 2019. <https://bloodwise.org.uk/info-support/chronic-myelomonocytic-leukaemia>.
- Boada M, Echarte L, Guillermo C, Diaz Lilián et al. 2021. "5-Azacytidine restores interleukin 6-increased production in mesenchymal stromal cells from myelodysplastic patients." *Hematol Transfus Cell Ther* Vol 43, 35-42.
- Bolte S, Cordelieres FP. 2006. "A guided tour into subcellular colocalisation analysis in light microscopy." *J Microscopy* Vol 224, 213-232.
- Bossolasco P, Soligo D, Servida F, Romitti L et al. 1997. "Response of myelodysplastic syndrome bone marrow cells to multiple cytokine stimulation in liquid cultures: an in situ hybridization study." *Haematologica* Vol 82, 532-6.
- Boulais PE, Frenette PS. 2015. "Making sense of hematopoietic stem cell niches." *Blood* Vol 125, 2621-2629.
- Bowie MB, McKnight KD, Kent DG, McCaffrey L et al. 2006. "Hematopoietic stem cells proliferate until after birth and show a reversible phase-specific engraftment defect." *J Clin Invest* Vol 116, 2808-2816.
- Bray LJ, Binner M, Korner Y, von Bonin M et al. 2017. "A three-dimensional ex vivo tri-culture model mimics cell-cell interactions between acute myeloid leukemia and the vascular niche." *Haematologica* Vol 102, 1215-1226.
- Bresnick EH, Martowicz ML, Pal S, Johnson KD. 2005. "Developmental control via GATA factor interplay at chromatin domains." *J Cell Physiol* Vol 205, 1-9.
- Broudy VC, Lin N, Zsebo KM, Birkett NC, Smith KA, Bernstein ID, Papayannopoulou T. 1992. "Isolation and characterization of a monoclonal antibody that recognizes the human c-kit receptor." *Blood* 79:338.
- Broudy, VC. 1997. "Stem Cell Factor and Hematopoiesis." *Blood* 90:1345.
- Brugger W, Möcklin W, Heimfeld S, Berenson RJ, Mertelsmann R, Kanz L. 1993. "Ex vivo expansion of enriched peripheral blood CD34+ progenitor cells by stem cell factor, interleukin-1 beta (IL-1 beta), IL-6, IL-3, interferon-gamma, and erythropoietin." *Blood* 81:2579.
- Buckstein R, Yee K and Wells RA. 2011. "5-Azacytidine in myelodysplastic syndromes: a clinical practice guideline." *Cancer Treat Rev* Vol 37, 160-167.
- Bungart B, Loeffler M, Goris H, Dontje B et al. 1990. "Differential effects of rekombinant human colony stimulating factor (rh-G-CSF) on stem cells in marrow, spleen and peripheral blood in mice." *Br J Haematol* Vol 76, 174-179.

- Buza-Vidas N, Antonchuk J, Qian H, Mansson R et al. 2006. "Cytokines regulate postnatal hematopoietic stem cell expansion: opposing roles of thrombopoietin and LNK." *Genes Dev* Vol 20, 2018-2023.
- Caceres G, McGraw K, Yip BH, Pellagatti A et al. 2013. "TP53 suppression promotes erythropoiesis in del(5q) MDS, suggesting a targeted therapeutic strategy in lenalidomide-resistant patients." *Proc Natl Acad Sci USA* Vol 110, 16127-32.
- Calvi LM, Link DC. 2015. "The hematopoietic stem cell niche in homeostasis and disease." *Blood* Vol 126, 2443-2451.
- Cancer.net. 2018. *Leukemia - Chronic Myeloid - CML: Phases*. Mar. Accessed Apr 2019. <https://www.cancer.net/cancer-types/leukemia-chronic-myeloid-cml/phases>.
- Cancer.org. 2018. *Tests for Acute Myeloid Leukemia (AML)*. August. Accessed November 2021. [https://www.cancer.org/cancer/acute-myeloid-leukemia/detection-diagnosis-staging/how-diagnosed.html#:~:text=\(In%20normal%20bone%20marrow%2C%20the,doesn't%20reach%2020%25](https://www.cancer.org/cancer/acute-myeloid-leukemia/detection-diagnosis-staging/how-diagnosed.html#:~:text=(In%20normal%20bone%20marrow%2C%20the,doesn't%20reach%2020%25).
- Cavo M, Caria M, Pulsoni I, Beltrame F, Fato M, Scaglione S. 2018. "A new cell-laden 3D Alginate-Matrigel hydrogel resembles human breast cancer cell malignant morphology, spread and invasion capability observed "in vivo". " *Sci Rep* Vol 8, 5333.
- Cesarz Z, Tanama K. 2015. "Spheroid Culture of Mesenchymal Stem Cells." *Stem Cell Inter* Vol 2016, 976357.
- Chais JA, Mohandas N. 2008. "Erythroblastic islands: niches for erythropoiesis." *Blood* Vol 112, 470-478.
- Chang PH, Chao HM, Chern E, Hsu SH. 2021. "Chitosan 3D cell culture system promotes naïve-like features of human induced pluripotent stem cells: A novel tool to sustain pluripotency and facilitate differentiation." *Biomaterials* Vol 268, 120575.
- Chen CY, Ke CJ, Yen KC, Hsieh HC, Sun JS, Lin FH. 2015. "3D Porous Calcium-Alginate Scaffolds Cell Culture System Improved Human Osteoblast Cell Clusters for Cell Therapy." *Theranostics* Vol 5, 643-655.
- Chen WL, Wang JH, Zhao AH, Xu X et al. 2014. "A distinct glucose metabolism signature of acute myeloid leukemia with prognostic value." *Blood* Vol 10, 1645-1654.
- Chen Y, Jacamo R, Shi YX, Wang RY et al. 2012. "Human extramedullary bone marrow in mice: a novel in vivo model of genetically controlled hematopoietic microenvironment." *Blood* Vol 119, 4971-4980.
- Chen, K, J Liu, S Heck, JA Chasis, X An, and N Mohandas. 2009. "Resolving the distinct stages in erythroid differentiation based on dynamic changes in membrane protein expression during erythropoiesis." *PNAS* 17413-17418.
- Cheng H, Zheng Z, Cheng T. 2020. "New paradigms on hematopoietic stem cell differentiation." *Protein & Cell* Vol 11, 34-44.

- Cheshier SH, Morrison SJ, Liao X, Weissman IL. 1999. "In vivo proliferation and cell cycle kinetics of long-term self-renewing hematopoietic stem cells." *Proc Natl Acad Sci USA* 96:3120.
- Cho MO, Li Z, Shim HE, Nurunnabi M et al. 2016. "Bioinspired tuning of glycol chitosan for 3D cell culture." *NPG Asia Mat* 8, e309.
- Colijn C, Mackey MC. 2005. "A mathematical model of hematopoiesis-I. Periodic chronic myelogenous leukemia." *J Theor Biol* Vol 237, 117-132.
- Collin M, Dickinson R, Bigley V. 2015. "Haematopoietic and immune defects associated with GATA2 mutation." *Br J Haematol* Vol 169, 173-187.
- Cotes, PM. 1982. "Immunoreactive erythropoietin in serum. I. Evidence for the validity of the assay method and the physiological relevance of estimates." *Br J Haematol* 50:427.
- Cox, RA. 2004. "Quantitative relationships for specific growth rates and macromolecular compositions of Mycobacterium tuberculosis, Streptomyces coelicolor A3(2) and Escherichia coli B/r: an integrative theoretical approach." *Microbiology* 150:1413.
- Cron RQ. 2021. "COVID-19 cytokine storm: targeting the appropriate cytokine." *The Lancet Rheumatology* Vol 3, E236-E237.
- CRUK. 2014. *Leukaemia (all subtypes combined) incidence statistics*. April. Accessed April 2019. <https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/leukaemia/incidence#heading-Four>.
- Custer RP, Ahlfeldt FE. 1932. "Studies on the structure and function of the bone marrow." *J Lab Clin Med*.
- Dai CH, Krantz SB, Koury ST, Kollar K. 1994. "Polycythaemia vera. IV. Specific binding of stem cell factor to normal and polycythaemia vera highly purified erythroid progenitor cells." *Br J Haematol* 88:497.
- Dalal N, Hooberman A, Mariani R, Sirota R et al. 2015. "Copper deficiency mimicking myelodysplastic syndrome." *Clin Case Rep* Vol 3, 325-327.
- De Smet D, Trullemans F, Jochmans K, Renmans W, Smet L, Heylen O, Bael AM, Schots R, Leus B, De Waele M. 2012. "Diagnostic potential of CD34+ cell antigen expression in myelodysplastic syndromes." *Am J Clin Pathol* 138:732.
- Deslauriers AG, Kotini AG, Papapetrou EP. 2021. "Modeling Leukemia Stem Cells with Patient-Derived Induced Pluripotent Stem Cells." *Methods Mol Biol* Vol 2185, 411-422.
- Dessypris E, Graber SE, Krantz SB, Stone WJ. 1988. "Effects of recombinant erythropoietin on the concentration and cycling status of human marrow hematopoietic progenitor cells in vivo." *Blood* 72:2060.
- Dicker F, Haferlach C, Sundermann J, Wendland N et al. 2010. "Mutation analysis for RUNX1, MLL-PTD, FLT3-ITD, NPM1 and NRAS in 269 patients with MDS or secondary AML." *Leukemia* Vol 24, 1528-1532.

- Dingli D, Traulsen A, Pacheco JM. 2007. "Compartmental Architecture and Dynamics of Hematopoiesis." *PLoS One* e435.
- Döhner H, Estey EH, Amadori S, Appelbaum FR et al. 2010. "Diagnosis and management of acute myeloid leukemia in adults: recommendations from an international expert panel, on behalf of the European LeukemiaNet." *Blood* Vol 115, 453-74.
- Dorssers L, Burger H, Bot F, Delwel R, Geurts van Kessel AH, Löwenberg B, Wagemaker G. 1987. "Characterization of a human multilineage-colony-stimulating factor cDNA clone identified by a conserved noncoding sequence in mouse interleukin-3." *Gene* 55:115.
- dos Santos SB, Allenby MC, Mantalaris A, Panoskaltsis N. 2016. "Early Erythroid Development Is Enhanced with Hypoxia and Terminal Maturation with Normoxia in a 3D Ex Vivo Physiologic Erythropoiesis Model." *Blood* Vol 128, 2453.
- Drexler HG, Dirks WG, MacLeod RA. 2009. "Many are called MDS cell lines: One is chosen." *Leuk Res* Vol 33, 1011-1016.
- Drusbosky L, Medina C, Martuscello R, Hawkins KE et al. 2017. "Computational drug treatment simulations on projections of dysregulated protein networks derived from the myelodysplastic mutanome match clinical response in patients." *Leuk Res* Vol 52, 1-7.
- D'Silva S, Rajadhyaksha SB, Singh M. 2019. "Immune dysregulation in MDS: the role of cytokines and immune cells." In *Recent developments in Myelodysplastic Syndromes*, by Ota Fuchs. IntechOpen.
- Duval K, Grover H, Han LH, Mou Y et al. 2017. "Modeling physiological events in 2D vs. 3D cell culture." *Physiology* Vol 32, 266-277.
- Dybedal I, Bryder D, Fossum A, Rusten LS et al. 2001. "Tumor necrosis factor (TNF)-mediated activation of the p55 TNF receptor negatively regulates maintenance of cycling reconstituting human hematopoietic stem cells." *Blood* Vol 98, 1782-1791.
- Edmondson R, Broglie JJ, Adcock AF, Yang L. 2014. "Three-Dimensional Cell Culture Systems and Their Applications in Drug Discovery and Cell-Based Biosensors." *Assay Drug Dev Technol* Vol 12, 207-218.
- Egrie JC, Cotes PM, Lane J, GainesDas RE, Tam RC. 1987. "Development of radioimmunoassays for human erythropoietin using recombinant erythropoietin as traced and immunogen." *J Immunol Methods* 99:235.
- Elvarsdóttir EM, Mortera-Blanco T, Dimitriou M, Boudierlique T et al. 2020. "A three-dimensional in vitro model of erythropoiesis recapitulates erythroid failure in myelodysplastic syndromes." *Leukemia* Vol 34, 271-282.
- Fajtova, Michaela, Anna Kovarikova, Peter Svec, Esko Kankuri, and Jan Sedlak. 2013. "Immunophenotypic profile of nucleated erythroid progenitors during maturation in regenerating bone marrow." *Leukemia & Lymphoma* 2523-2530.
- Faquin WC, Schneider TJ, Goldberg MA. 1992. "Effect of inflammatory cytokines on hypoxia-induced erythropoietin production." *Blood* Vol 93, 1987-94.

- Feng X, Scheinberg P, Wu CO, Samsel L et al. 2011. "Cytokine signature profiles in acquired aplastic anemia and myelodysplastic syndromes." *Haematologica* Vol 96, 602-606.
- Finch CA, Harker LA, Cook JD. 1977. "Kinetics of the formed elements of human blood." *Blood* Vol 50, 699-707.
- Finch, SC. 2007. "Radiation-induced leukemia: lessons from history." *Best Pract Res Clin Haematol* Vol 20, 109-118.
- Follo MY, Mongiorgi S, Clissa C, Paolini S et al. 2012. "Activation of nuclear inositide signalling pathways during erythropoietin therapy in low-risk MDS patients." *Leukemia* Vol 26, 2474-2782.
- Foucar K, Langdon RM, Armitage JO, Olson DB et al. 1985. "Myelodysplastic syndromes. A clinical and pathologic analysis of 109 cases." *Cancer* Vol 56, 553-561.
- Frisan E, Vandekerckhove J, de Thonel A, Pierre-Eugène C et al. 2012. "Defective nuclear localization of Hsp70 is associated with dyserythropoiesis and GATA-1 cleavage in myelodysplastic syndromes." *Blood* Vol 9, 1532-1542.
- Fritsche-Polanz R, Jordan JH, Feix A, Sperr WR, Sunder-Plassmann G, Valent P, Födinger M. 2001. "Mutation analysis of C-KIT in patients with myelodysplastic syndromes without mastocytosis and cases of systemic mastocytosis." *Br J Haematol* 113:357.
- Gañán-Gómez I, Wei Y, Starczynowski DT, Colla S et al. 2015. "Deregulation of innate immune and inflammatory signaling in myelodysplastic syndromes." *Leukemia* Vol 29, 1458-1469.
- Garland JM, Crompton S. 1983. "A preliminary report: Preparations containing interleukin-3 (IL-3) promote proliferation of multipotential stem cells (CFU) in the mouse." *Exp Hematol* 11:757.
- Gavosto F, Pileri A, Bachi C, Pegararo L. 1964. "Proliferation and maturation defect in acute leukemia cells." *Nature* 203:92.
- Gerstung M, Pellagatti A, Malcovati L, Giagounidis A et al. 2015. "Combining gene mutation with gene expression data improves outcome prediction in myelodysplastic syndromes." *Nat Comm* Vol 6, 5901.
- Giarratana MC, Rouard H, Dumont A, Kiger L, Safeukui I, LePennec P-Y, François S, Trugnan G, Peyrard T, Marie T, Jolly S, Hebert N, Mazurier C, Mario N, Laurence H, Lapillone H, Devaux J-Y, Douay L. 2011. "Proof of principle for transfusion of in vitro-generated red blood cells." *Blood* 118:5071.
- Glinsmann-Gibson B, Spier C, Baier M, Taetle R, Broudy VC, List AF. 1994. "Mast cell growth factor (c-kit ligand) restores growth of multipotent progenitors in myelodysplastic syndrome." *Leukemia* 8:827.
- Goldberg SL, Chen E, Corral M, Guo A et al. 2010. "Incidence and clinical complications of myelodysplastic syndromes among United States Medicare beneficiaries." *J Clin Oncol* Vol 28, 2847-2852.

- Gomariz A, Helbling PM, Isringhausen S, Suessbier U et al. 2018. "Quantitative spatial analysis of haematopoiesis-regulating stromal cells in the bone marrow microenvironment by 3D microscopy." *Nat Comm* Vol 9, 2532.
- Gorczyca, Wojciech, Zhong-Yi Sun, William Cronin, Xiaoyu Li, Sophal Mau, and Sorina Tugulea. 2011. "Immunophenotypic pattern of myeloid populations by Flow Cytometry analysis." *Methods in Cell Biology* 221-266.
- Graaf CA, Kauppi M, Baldwin T, Hyland CD et al. 2010. "Regulation of hematopoietic stem cells by their mature progeny." *Proc Natl Acad Sci USA* Vol 107, 21689-21694.
- Graubert T, Walter MJ. 2011. "Genetics of myelodysplastic syndromes: new insights." *Hematology Am Soc Hematol Educ Program* 543-549.
- Grayson WL, Frohlich M, Yeager K, Bhumiratana S. 2010. "2010." *Proc Natl Acad Sci* Vol 107, 3299-3304.
- Greenberg P, Cox C, LeBeau MM, Fenaux P et al. 1997. "International Scoring System for Evaluating Prognosis in Myelodysplastic Syndromes." *Blood* Vol 89, 2079-2088.
- Greenberg, PL. 2006. *Myelodysplastic Syndromes: Clinical and Biological Advances*. Cambridge, United Kingdom: Cambridge University Press.
- Grinenko T, Eugster A, Thielecke L, Ramasz B et al. 2018. "Hematopoietic stem cells can differentiate into restricted myeloid progenitors before cell division in mice." *Nature Communications* Vol 9, 1898.
- Gromski PS, Xu Y, Correa E, Ellis DI et al. 2014. "A comparative investigation of modern feature selection and classification approaches for the analysis of mass spectrometry data." *Analytica Chimica Acta* Vol 829, 1-8.
- Guo J, Zhao C, Yao R, Sui A et al. 2019. "3D culture enhances chemoresistance of ALL Jurkat cell line by increasing DDR1 expression." *Exp Ther Med* Vol 17, 1593-1600.
- Gupta P, Pérez-Mancera PA, Kocher H, Nisbet A, Schettino G, Velliou EG. 2020. "A Novel Scaffold-Based Hybrid Multicellular Model for Pancreatic Ductal Adenocarcinoma—Toward a Better Mimicry of the *in vivo* Tumor Microenvironment." *Front Bioeng Biotechnol* Vol 8, Article 290.
- Haferlach T, Nagata Y, Grossmann V, Okuno Y et al. 2014. "Landscape of genetic lesions in 944 patients with myelodysplastic syndromes." *Leukemia* Vol 28, 241-247.
- Hagiwara S, Sato T, Itoh Y. 1997. "Effects of recombinant human interleukin-3 on maturation of megakaryocytic cell line, CMK." *Leuk Res* 21:153.
- Hamon P, Rodero MP, Combadière C, Boissonnas A. 2015. "Tracking mouse bone marrow monocytes *in vivo*." *J Vis Exp* Vol 27, e52476.
- Hara T, Miyajima A. 1996. "Function and Signal Transduction Mediated by the Interleukin 3 Receptor System in Hematopoiesis." *Stem Cells* 14:605.
- Harnan S, Ren S, Gomersall T, Everson-Hock ES, Sutton A, Dhanasiri S, Kulasekararaj A. 2016. "Association between Transfusion Status and Overall Survival in Patients with

- Myelodysplastic Syndromes: A Systematic Literature Review and Meta-Analysis." *Acta Haematol* 136:23.
- Hasse, D. 2008. "Cytogenetic features in myelodysplastic syndromes." *Ann Hematol* Vol 87, 515-526.
- Hattangadi SM, Wong P, Zhang L, Flygare J, Lodish HF. 2011. "From stem cell to red cell: regulation of erythropoiesis at multiple levels by multiple proteins, RNAs and chromatin modifications." *Blood* 118:6258.
- Hellström-Lindberg, E. 1995. "Efficacy of erythropoietin in the myelodysplastic syndromes: a meta-analysis of 205 patients from 17 studies." *Br J Haematol* 89:67.
- Henderson LA, Canna SW, Schulert GS, Volpi S et al. 2020. "On the Alert for Cytokine Storm: Immunopathology in COVID-19." *Arthritis Rheumatol* Vol 72, 1059-1063.
- Henschler R, Brugger W, Luft T, Frey T, Mertelsmann R, Kanz L. 1994. "Maintenance of transplantation potential in ex vivo expanded CD34+-selected human peripheral blood progenitor cells." *Blood* 84:2898.
- Hibner U, Coutinho A. 1994. "Signal Antonymy: A mechanism for apoptosis induction." *Cell Death Differ* 1:33.
- Hill, AV. 1910. "The possible effects of the aggregation of the molecules of haemoglobin on its dissociation curves." *J Physiol*.
- Hoek PV, Dijken JPV, Pronk JT. 1998. "Effect of specific growth rate on fermentative capacity of baker's yeast." *Appl Environ Microbiol* 64:4226.
- Hofer T, Rodewald HR. 2018. "Differentiation-based model of hematopoietic stem cell functions and lineage pathways." *Blood* Vol 132, 1106-1113.
- Hojyo S, Uchida M, Tanaka K, Hasebe R et al. 2020. "How COVID-19 induces cytokine storm with high mortality." *Inflamm Regen* Vol 40, 37.
- Hongisto V, Jernstrom S, Fey V, Mpindi JP et al. 2013. "High-Throughput 3D Screening Reveals Differences in Drug Sensitivities between Culture Models of JIMT1 Breast Cancer Cells." *PLOS One* Vol 8, e77232.
- Horner MJ, Ries LAG, Krapcho M, Neyman N. 2009. "SEER Cancer Statistics Review, 1975-2006." *Surv Epidem End Results*.
- Huang E, Nocka K, Beier DR, Chu T-Y, Buck J, Lahm H-W, Wellner D, Leder P, Besmer P. 1990. "The hematopoietic growth factor KL is encoded by the SI locus and is the ligand of the c-kit receptor, the gene product of the W locus." *Cell* 63:225.
- Huang P, Zhao Y, Zhong J, Zhang X et al. 2020. "Putative regulators for the continuum of erythroid differentiation revealed by single-cell transcriptome of human BM and UCB cells." *PNAS* Vol 117, 12868-12876.
- Huff JD, Keung YK, Thakuri M, Beaty MW et al. 2007. "Copper deficiency causes reversible myelodysplasia." *Am J Hematol* Vol 82, 625-630.

- Im H, Rao V, Sridhar K, Bentley J et al. 2018. "Distinct transcriptomic and exomic abnormalities within myelodysplastic syndrome marrow cells." *Leuk Lymphoma* Vol 59, 2952-2962.
- Imamura Y, Mukohara T, Shimono Y, Funakoshi Y et al. 2015. "Comparison of 2D- and 3D-culture models as drug-testing platforms in breast cancer." *Oncol Rep* Vol 33, 1837-1843.
- Immunopaedia. 2017. Accessed Apr 2019. immunopaedia.org.
- Ito M, Hiramatsu H, Kobayashi K. 2002. "NOD/SCID/gamma(c)(null) mouse: an excellent recipient mouse model for engraftment of human cells." *Blood* Vol 10, 3175-3182.
- Itoh N, Yonehara S, Schreurs J, Gorman DM, Maruyama K, Ishii A, Yahara I, Arai K, Miyajima A. 1990. "Cloning of an Interleukin-3 Receptor Gene: A Member of a Distinct Receptor Gene Family ." *Science* 247:324.
- Itzykson R, Kosmider O, Fenaux P. 2013. "Somatic mutations and epigenetic abnormalities in myelodysplastic syndromes." *Best Pract Res Clin Haematol* Vol 26, 355-364.
- Jan M, Snyder TM, Corces-Zimmerman MR, Vyas P et al. 2012. "Clonal evolution of preleukemic hematopoietic stem cells precedes human acute myeloid leukemia." *Sci Transl Med* Vol 29, 149.
- Jann JC, Mossner M, Nolte F, Boch T et al. 2016. "RNA Sequencing of Bone Marrow Derived Mesenchymal Stromal Cells (MSCs) of Healthy Young, Healthy Old and Patients with Myelodysplastic Syndrome (MDS) Reveals Transcriptomic Changes upon Aging of the Niche and Onset of MDS." *Blood* Vol 126. 4300.
- Jansen AJ, Essink-Bot ML, Beckers EA, Hop WC et al. 2003. "Quality of life measurement in patients with transfusion-dependent myelodysplastic syndromes." *Br J Haematol* Vol 121, 270-274.
- Jelkmann, W. 1992. "Erythropoietin: Structure, Control of Production and Function." *Phys Rev* 72:449.
- Jelkmann, Wolfgang. 2011. "Regulation of erythropoietin production." *J Physiol* Vol 589, 1251-1258.
- Jin L, Hope KJ, Zhai Q, Smadja-Joffe F et al. 2006. "Targeting of CD44 eradicates human acute myeloid leukemic stem cells." *Nat Med* Vol 12, 1167.
- Jing D, Wobus M, Poitz DM, Bornhauser M, Ehninger G. 2012. "Oxygen tension plays a critical role in the hematopoietic microenvironment in vitro." *Haematologica* Vol 97, 331-339.
- Johnson CS, Pourbohloul SC, Furmanski P. 1991. "Negative regulators of in vivo erythropoiesis: interaction of IL-1 alpha and TNF-alpha and the lack of a strict requirement for T or NK cells for their activity." *Exp Hematol* Vol 19, 101-105.
- Jolliffe IT, Cadima J. 2016. "Principal component analysis: a review and recent developments." *Phil Trans R Soc A* Vol 374, 20150202.
- Jordan CT, Upchurch D, Szilvassy SJ, Guzman ML, Howard DS, Pettigrew AL, Meyerrose T, Rossi R, Grimes B, Rizzieri DA, Luger SM, Phillips GL. 2000. "The interleukin-3 receptor

- alpha is a unique marker for human acute myelogenous leukemia stem cells." *Leukemia* 14:1777.
- Kafi A, Aktar K, Todo M, Dahiya R. 2019. "Engineered chitosan for improved 3D tissue growth through Paxillin-FAK-ERK activation." *Regen Biomat* Vol 7, 141-151.
- Kanani HH, Klapa MI. 2007. "Data correction strategy for metabolomics analysis using gas chromatography-mass spectrometry." *Metab Eng* Vol 9, 39-51.
- Kao HW, Sanada M, Liang DC, Lai CL et al. 2011. "A high occurrence of acquisition and/or expansion of C-CBL mutant clones in the progression of high-risk myelodysplastic syndrome to acute myeloid leukemia." *Neoplasia* Vol 13, 1035-1042.
- Kapałczyńska M, Kolenda T, Przybyła W, Zajączkowska M et al. 2018. "2D and 3D cell cultures – a comparison of different types of cancer cell cultures." *Arch Med Sci* Vol 14, 910-919.
- Karlíková R, Šírká J, Friedecký D, Faber E. 2016. "Metabolite Profiling of the Plasma and Leukocytes of Chronic Myeloid Leukemia Patients." *J Proteome Res* Vol 15, 3158-3166.
- Kaushansky K, Lichtman MA, Beutler E, Kipps TJ, Seligsohn U, Prchal JT. 2011. *Williams Hematology, 8th Edition*. McGrawHill Medical.
- Kaushansky, K, MA Lichtman, E Beutler, TJ Kipps, U Seligsohn, and JT Prchal. 2011, 8th Edition. *Williams Hematology*. McGrawHill Medical.
- Kee N, Sivalingam S, Boonstra R, Wojtowicz JM. 2002. "The utility of Ki-67 and BrdU as proliferative markers of adult neurogenesis." *J Neurosci Methods* 30:97.
- Keller JR, Ortiz M, Ruscetti FW. 1995. "Steel factor (c-kit ligand) promotes the survival of hematopoietic stem/progenitor cells in the absence of cell division." *Blood (Blood)* Vol 86, 1757-1764.
- Kerbaui DB, Deeg JH. 2007. "Apoptosis and anti-apoptotic mechanisms in the progression of MDS." *Exp Hematol* 35:1739.
- Khalid A, Siddiqui AJ, Huang JH, Shamsi T et al. 2018. "Alteraiton of serum free fatty acids are indicators for progression of pre-leukaemia diseases to leukaemia." *Sci Rep* Vol 8, 14883.
- Killman, SA. 1968. "Acute Leukemia: The kinetics of leukemic blasts in man. An analytic review." *Ser Hematol* 1:38.
- Kim H, Lin Q, Glazer PM, Yun Z. 2018. "The hypoxic tumor microenvironment in vivo selects the cancer stem cell fate of breast cancer cells." *Breast Cancer Res* Vol 20, 16.
- Kiparissides A, Rodriguez-Fernandez M, Kucherenko S, Mantalaris A et al. 2008. "Appkication of Global Sensitivity Analysis to Biological Models." *18th European Symposium on Computed Aided Process Engineering*. Elsevier B.V. 689-694.
- Kirouac DC, Madlambayan GJ, Yu M, Sykes EA et al. 2009. "Cell-cell interaction networks regulate blood stem and progenitor cell fate." *Mol Syst Biol* Vol 5, 293.
- Kirouac DC, Madlambayan GJ, Yu M, Sykes EA, Ito C, Zandstra PW. 2009. "Cell-cell interaction networks regulate blood stem and progenitor cell fate." *Mol Syst Biol* 293:1.

- Kitamura T, Sato N, Arai K, Miyajima A. 1991. "Expression cloning of the human IL-3 receptor cDNA reveals a shared beta subunit for the human IL-3 and GM-CSF receptors." *Cell* 66:1165.
- Klontzas ME, Vernardis SI, Heliotiis M, Tsiridis E et al. 2017. "Metabolomics analysis of the osteogenic differentiation of umbilical cord blood mesenchymal stem cells reveals differential sensitivity to osteogenic agents." *Stem Cells Dev* Vol 26, 723-733.
- Koike K, Ihle JN, Ogawa M. 1986. "Declining sensitivity to interleukin 3 of murine multipotential hemopoietic progenitors during their development. Application to a culture system that favors blast cell colony formation." *J Clin Invest* 77:894.
- Komrokji RS, List AF. 2016. "Short- and long-term benefits of lenalidomide treatment in patients with lower-risk del(5q) myelodysplastic syndromes." *Ann Oncol* Vol 27, 62-68.
- Kondo, M, AJ Wagers, MG Manz, SS Prohaska, DC Scherer, GF Beilhack, JA Shizuru, and IL Weissman. 2003. "Biology of hematopoietic stem cells and progenitors: implications for clinical application." *Annu Rev Immunol* 759-806.
- Kontoravdi C, Asprey SP, Pistikopoulos EN, Mantalaris A. 2005. "Application of global sensitivity analysis to determine goals for design of experiments: an example study on antibody-producing cell cultures." *Biotechnol Prog* Vol 21, 1128-1135.
- Kordasti SY, Afzali B, Lim Z, Ingram W et al. 2009. "IL-17 producing CD4(+) T cells, pro-inflammatory cytokines and apoptosis are increased in low risk myelodysplastic syndrome." *Br J Haematol* Vol 145, 64-72.
- Kosmider O, Gelsi-Boyer V, Ciudad M, Racœur M et al. 2009. "TET2 gene mutation is a frequent and adverse event in chronic myelomonocytic leukemia." *Haematologica* Vol 94, 1676-1681.
- Kotidis P, Demis P, Goey CH, Correa E et al. 2019. "Constrained global sensitivity analysis for bioprocess design space identification." *Comp Chem Eng*.
- Kotini AG, Chang CJ, Chow A, Yuan H et al. 2017. "Stage-Specific Human Induced Pluripotent Stem Cells Map the Progression of Myeloid Transformation to Transplantable Leukemia." *Cell Stem Cell* Vol 20, 315-328.
- Koury MJ, Bondurant MC. 1991. "The mechanism of erythropoietin action." *Am J Kidney Dis* Vol 18, 20-23.
- Koury, MJ. 2015. "Tracking erythroid progenitor cells in times of need and times of plenty." *Exp Hematol*.
- Krönke J, Fink EC, Hollenbach PW, MacBeth KJ. 2015. "Lenalidomide induces ubiquitination and degradation of CK1 α in del(5q) MDS." *Nature* Vol 523, 183-188.
- Krosi J, Damen JE, Krystal G, Humphries RK. 1996. "Interleukin-3 (IL-3) inhibits erythropoietin-induced differentiation in Ba/F3 cells via the IL-3 receptor alpha subunit." *J Biol Chem* 271:27432.
- Kumar R, Godavarthy PS, Krause DS. 2018. "The bone marrow microenvironment in health and disease at a glance." *J Cell Sci* Vol 131.

- Lacombe F, Durrieu F, Briais A, Dumain P et al. 1997. "Flow cytometry CD45 gating for immunophenotyping of acute myeloid leukemia." *Leukemia* Vol 11, 1878-1886.
- Landberg N, Palffy SV, Askmyr M et al. 2018. "CD36 defines primitive chronic myeloid leukemia cells less responsive to imatinib but vulnerable to antibody-based therapeutic targeting." *Haematologica* Vol 3, 447-455.
- Lander AD, Gokoffski KK, Wan FYM, Nie Q et al. 2009. "Cell lineages and the logic of proliferative control." *PLoS Biology* 7:e1000015.
- Langhans, SA. 2018. "Three-Dimensional in Vitro Cell Culture Models in Drug Discovery and Drug Repositioning." *Front. Pharmacol* Vol 9, Article 6.
- Leary AG, Zeng HQ, Clark SC, Ogawa M. 1992. "Growth factor requirements for survival in G0 and entry into the cell cycle of primitive human hemopoietic progenitors." *Proc Natl Acad Sci USA* 89:4013.
- Lee SW, Kwak HS, Kang MH, Park YY et al. 2018. "Fibroblast-associated tumour microenvironment induces vascular structure-networked tumouroid." *Sci Rep* Vol 8, 2365.
- Lehner B, Sandner B, Marschallinger J, Lehner C, Furtner T, Couillard-Despres S, Rivera FJ, Brockhoff G, Bauer HC, Weidner N, Aigner L. 2011. "The dark side of BrdU in neural stem cell biology: detrimental effects on cell cycle, differentiation and survival." *Cell Tissue Res* 345:313.
- Leonard JP, May S, Ihle JN, Pettit GR, Sharkis SJ. 1988. "Regulation of Hematopoiesis IV: The Role of Interleukin-3 and Bryostat 1 in the Growth of Erythropoietic Progenitors From Normal and Anemic W/W^u Mice." *Blood* 72:1492.
- Leth IK, McDonald KA. 2017. "Growth kinetics and scale-up of *Agrobacterium tumefaciens*." *Appl Microbiol Biotechnol*.
- Levenspiel, O. 1980. "The Mond Equation: A Revisit and a Generalization to Product Inhibition Situations." *Biotech Bioeng* 22:1671.
- Li L, Clevers H. 2010. "Coexistence of quiescent and active adult stem cells in mammals." *Science (Science)* Vol 327, 542-545.
- Li LJ, Tao JL, Fu R, Wang HQ, Jiang HJ, Yue LZ, Zhang W, Liu H, Shao ZH. 2014. "Increased CD34⁺ CD38⁻ CD123⁺ cells in myelodysplastic syndrome displaying malignant features similar to those in AML." *Int J Hematol* 100:60.
- Li, J, J Hale, P Bhagia, F Xue, L Chen, J Jaffray, H Yan, et al. 2014. "Isolation and transcriptome analyses of human erythroid progenitors: BFU-E and CFU-E." *Blood* 123:3636-3645.
- Liesveld JL, Jordan CT, Phillips GL. 2004. "The hematopoietic stem cell in myelodysplasia." *Stem Cells* Vol 22, 590-599.
- Lim M, Panoskaltsis N, Ye H, Mantalaris A. 2011. "Optimization of in vitro erythropoiesis from CD34⁺ cord blood cells using design of experiments (DOE)." *Biochem Eng Jour* Vol 55, 154-161.

- Lindsley RC, Saber W, Mar BG, Redd R et al. 2017. "Prognostic mutations in myelodysplastic syndrome after stem-cell transplantation." *N Engl J Med* Vol 376, 536-547.
- López-Gómez S, Lerner RG, Petritsch C. 2014. "Asymmetric cell division of stem and progenitor cells during homeostasis and cancer." *Cell Mol Life Sci* Vol 71, 575-597.
- Lopez-Yrigoyen M, Yang CT, Fidanza A, Cassetta A et al. 2019. "Genetic programming of macrophages generates an in vitro model for the human erythroid island niche." *Nat Comm* Vol 10, 881.
- Lorenzi PL, Reinhold WC, Varma S, Hutchinson AA et al. 2009. "DNA fingerprinting of the NCI-60 cell line panel." *Mol Cancer Ther* Vol 9, 713-724.
- Lorenzo F, Nishii K, Monma F, Kuwagata S, Usui E, Shiku H. 2006. "Mutational analysis of the KIT gene in myelodysplastic syndrome (MDS) and MDS-derived leukemia." *Leuk Res* 30:1235.
- Lubeck DP, Danese M, Jennifer D, Miller K et al. 2016. "Systematic literature review of the global incidence and prevalence of myelodysplastic syndrome and acute myeloid leukemia." *Blood* Vol 128, 5930.
- Lundberg P, Takizawa H, Kubovcakova L, Guo G et al. 2014. "Myeloproliferative neoplasms can be initiated from a single hematopoietic stem cell expressing JAK2-V617F." *J Exp Med* 211:2213.
- Lyons AB, Parish CR. 1994. "Determination of lymphocyte division by flow cytometry." *J Immunol Meth* 171:131.
- Mackey MC, Glass L. 1977. "Oscillation and chaos in physiological control systems." *Science* Vol 197, 287-289.
- MacLean AL, LoCelso C, Stumph MPH. 2012«3. "Population dynamics of normal and leukaemia stem cells in the haematopietic stem cell niche show distinct regimes where leukaemia will be controlled." *J R Soc Interface* 10:20120968.
- Majek P, Riedelova-Reicheltova Z, Suttner J, Pecankova K et al. 2014. "Proteome Changes in the Plasma of Myelodysplastic Syndrome Patients with Refractory Anemia with Excess Blasts Subtype 2." *Dis Markers* Vol 2014, 178709.
- Makishima H, Visconte V, Sakaguchi H, Jankowska AM et al. 2012. "Mutations in the spliceosome machinery, a novel and ubiquitous pathway in leukemogenesis." *Blood* Vol 119, 3203-3210.
- Malcovati L, Karimi M, Papaemmanuil E, Ambaglio I. 2015. "SF3B1 mutation identifies a distinct subset of myelodysplastic syndrome with ring sideroblasts." *Blood* Vol 126, 233-241.
- Marciniak-Czochra A, Stiehl T, Ho AD, Jager W et al. 2009. "Modeling of asymmetric cell division in hematopoietic stem cells - regulation of self-renewal is essential for efficient repopulation." *Stem Cel Dev* Vol 18, 377-385.
- Marcondes AM, Mhyre AJ, Stirewalt DL, Kim SH et al. 2008. "Dysregulation of IL-32 in myelodysplastic syndrome and chronic myelomonocytic leukemia modulates apoptosis and impairs NK function." *Proc Natl Acad Sci USA* Vol 105, 2865-2870.

- Marr, AG. 1991. "Growth Rate of Escherichia coli." *Microbiol Rev* 55:316.
- Marsden, JT. 2006. "Erythropoietin - measurement and clinical applications." *Ann Clin Biochem* Vol 43, 97-104.
- Mauer AM, Murphy AB. 1979. "Kinetic studies of cells in childhood leukemias." *Am J Clin Path* 72:753.
- McGill J, Legge KL. 2009. "Contribution of lung-resident T cell proliferation to the overall magnitude of the antigen-specific CD8 T cell response in the lungs following murine influenza virus infection." *J Immunol* 183:4177.
- McGraw KL, Basiorka AA, Johnson JO, Clark J, Caceres G, Padron E, Heaton R, Ozawa Y, Wei S, Sokol L, List AF. 2014. "Lenalidomide Induces Lipid Raft Assembly to Enhance Erythropoietin Receptor Signaling in Myelodysplastic Syndrome Progenitors." *PLoS One* 3:e114249.
- Medinger M, Drexler B, Lengerke C, Passweg J. 2018. "Pathogenesis of Acquired Aplastic Anemia and the Role of the Bone Marrow Microenvironment." *Front Oncol* Vol 8, 587.
- Melissaridou S, Wiechec E, Magan M, Jain MV et al. 2019. "The effect of 2D and 3D cell cultures on treatment response, EMT profile and stem cell features in head and neck cancer." *Cancer Cell Int* Vol 19, 16.
- Merchav S, Wagemaker G, Souza LM, Tatarsky I. 1991. "Impaired response of myelodysplastic marrow progenitors to stimulation with recombinant haemopoietic growth factors." *Leukemia* Vol 5, 340-346.
- Merchuk JC, Asenjo JA. 1995. "The Monod Equation and Mass Transfer." *Biotech Bioeng* 45:91.
- Metcalf, D. 2008. "Hematopoietic cytokines." *Blood* 111:485.
- Michor F, Hughes TP, Iwasa Y, Branford S et al. 2005. "Dynamics of chronic myeloid leukaemia." *Nature* Vol 435, 1267-1270.
- Migliaccio G, Migliaccio AR, Druzin ML, Giardina P-JV, Zsebo KM, Adamson JW. 1992. "Long-term generation of colony-forming cells in liquid culture of CD34+ cord blood cells in the presence of recombinant human stem cell factor." *Blood* 79:2620.
- Miller CL, Rebel VI Helgason CD, Lansdorp PM, Eaves CJ. 1997. "Impaired steel factor responsiveness differentially affects the detection and long-term maintenance of fetal liver hematopoietic stem cells in vivo." *Blood* 89:1214.
- Mittelman M, Gardyn J, Carmel M, Malovani H, Barak Y, Nir U. 1996. "Analysis of the erythropoietin receptor gene in patients with myeloproliferative and myelodysplastic syndromes." *Leuk Res* 20:459.
- Miyake T, Goldwasser E. 1977. "Purification og human erythropoietin." *J Biol Chem* 252:5558.
- Möbest D, Goan SR, Junghahn I, Winkler J, Fichtner I, Hermann M, Becker M, de Lima-Hahn E, Henschler R. 1999. "Differential kinetics of primitive hematopoietic cells assayed in vitro

- and in vivo during serum-free suspension culture of CD34+ blood progenitor cells." *Stem Cells* 17:152.
- Monod, J. 1958. *Recherches sur la Croissance des Cultures Bactérienne*. Paris: Hermann et Cie.
- Morales H, Robert J. 2008. "In vivo and in vitro techniques for comparative study of antiviral T-cell responses in the amphibian *Xenopus*." *Biol Proceed Online* 10:1.
- Morita K, Wang F, Jahn K, Hu T et al. 2020. "Clonal evolution of acute myeloid leukemia revealed by high-throughput single-cell genomics." *Nat Comms* Vol 11. 5327.
- Morrison SJ, Scadden DT. 2014. "The bone marrow niche for haematopoietic stem cells." *Nature* Vol 505, 327-334.
- Mortera-Blanco T, Mantalaris A, Bismarck A, Aqel N et al. 2011. "Long-term cytokine-free expansion of cord blood mononuclear cells in three-dimensional scaffolds." *Biomaterials* Vol 32, 9263-9270.
- Mortera-Blanco T, Mantalaris A, Bismarck A, Panoskaltsis N et al. 2010. "The development of a three-dimensional scaffold for ex vivo biomimicry of human acute myeloid leukaemia." *Biomaterials* Vol 31, 2243-2251.
- Moudra A, Hubackova S, Machalova V, Vancurova M et al. 2016. "Dynamic alterations of bone marrow cytokine landscape of myelodysplastic syndromes patients treated with 5-azacytidine." *Oncoimmunology* Vol 5, e1183860.
- Mundle SD, Venugopal P, Cartlidge JD, Pandav DV et al. 1996. "Indication of an involvement of interleukin-1 beta converting enzyme-like protease in intramedullary apoptotic cell death in the bone marrow of patients with myelodysplastic syndromes." *Blood* Vol 88, 2640-2647.
- Musto P, Sanpaolo G, D'Arena G, Scalzulli PR, Matera R, Falcone A, Bodenizza C, Perla G, Carotenuto M. 2001. "Adding growth factors or interleukin-3 to erythropoietin has limited effects on anemia of transfusion-dependent patients with myelodysplastic syndromes unresponsive to erythropoietin." *Haematologica* 86:44.
- Nakamura Y, Komatsu N, Nakauchi H. 1992. "A truncated erythropoietin receptor that fails to prevent programmed cell death of erythroid cells." *Science* 257:1138.
- Nand S, Sosman J, Godwin JE, Fisher RI. 1994. "A phase I/II study of sequential interleukin-3 and granulocyte-macrophage colony-stimulating factor in myelodysplastic syndromes." *Blood* 83:357.
- Nassereddine S, Nishihori T, Padron E, Mahfouz R et al. 2017. "Integrating genomics in myelodysplastic syndrome to predict outcomes after allogeneic hematopoietic cell transplantation." *Clin Lymph Myel Leuk* Vol 17, 7-13.
- Nathan DG, Chess L, Hillman DG, Clarke B et al. 1978. "Human erythroid burst-forming unit: T-cell requirement for proliferation in vitro." *J Exp Med* (J Exp Med) Vol 147, 324-339.
- NCI. 2019. *NCI Dictionary of Cancer Terms*. Accessed April 2019.
<https://www.cancer.gov/publications/dictionaries/cancer-terms/def/bone-marrow-biopsy>.

- . 2015. *NCI-60 Human Tumor Cell Lines Screen*. August. Accessed April 2019. https://dtp.cancer.gov/discovery_development/nci-60/.
- Nemeth E, Ganz T. 2014. "Anemia of Inflammation." *Hematol Oncol Clin North Am* Vol 28, 671-681.
- Nemeth MJ, Bodine DM. 2007. "Regulation of hematopoiesis and the hematopoietic stem cell niche by Wnt signaling pathways." *Cell Res* Vol 17, 746-758.
- Nerlov C, Querfurth E, Kulesa H, Graf T. 2000. "GATA-1 interacts with the myeloid PU.1 transcription factor and represses PU.1-dependent transcription." *Blood* Vol 15, 2543-2551.
- NHS. 2019. *Folic acid*. Mar. Accessed Apr 2019. <https://www.nhs.uk/medicines/folic-acid/>.
- NICE. 2012. *National Institute for Health and Clinical Excellence*. October. Accessed Apr 2019. <https://www.nice.org.uk/guidance/ta322/documents/myelodysplastic-syndrome-deletion-5q-lenalidomide-final-scope2>.
- NIH. 2019. *Essential thrombocythemia*. Apr. Accessed Apr 2019. <https://ghr.nlm.nih.gov/condition/essential-thrombocythemia>.
- . 2008. "The Role of Erythropoietin in Myelodysplastic Syndrome." *Clinical Trials*. July 24. Accessed May 18, 2017. <https://clinicaltrials.gov/ct2/show/study/NCT00723112>.
- Nygren JM, Bryder D. 2008. "A Novel Assay to Trace Proliferation History In Vivo Reveals that Enhanced Divisional Kinetics Accompany Loss of Hematopoietic Stem Cell Self-Renewal." *PLoS One* 3:e3710.
- Obeyesekere MN, Spicer PP, Korbling M. 2005. "A mathematical model for the progression of an abnormality in the hematopoietic system." *LNCIS* Vol 321, 257-267.
- Obeyesekere, MN, RW Berry, PP Spicer, and M Korbling. 2004. "A mathematical model of haemopoiesis as exemplified by CD34+ cell mobilization into the peripheral blood." *Cell Prolif* 279-294.
- Ogawa, Seishi. 2019. "Genetics of MDS." *Blood* In Press.
- Ohyashiki K, Fujieda H, Iwabuchi A, Ohyashiki JH. 1990. "In vitro cytogenetic effects of recombinant human hematopoietic growth factors on cells derived from myelodysplastic syndromes." *Cancer Genet Cytogenet* Vol 48, 169-178.
- Okuda T, Nishimura M, Nakao M, Fujita Y. 2001. "RUNX1/AML1: a central player in hematopoiesis." *Int J Hematol* Vol 74, 252-257.
- ONS. 2016. *Cancer Registration Statistics, England*. Accessed April 2019. <https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/cancerregistrationstatisticsengland/2016>.
- Oostendorp RAJ, Audet J, Eaves CJ. 2000. "High-resolution tracking of cell division suggests similar cell cycle kinetics of hematopoietic stem cells stimulated in vitro and in vivo." *Blood* 95:855.

- Orelia C, van der Sluis RM, Verkuijlen P, Nethe M. 2011. "Altered intracellular localization and mobility of SBDS protein upon mutation in Shwachman-Diamond Syndrome." *PLoS One* Vol 6, e20727.
- Padgett, RA. 2012. "New connections between splicing and human disease." *Trends Genet* Vol 28, 147-154.
- Panoskaltis N, Mantalaris A, Wu JD. 2005. "Engineering a mimicry of bone marrow tissue ex vivo." *J Biosci Bioeng* Vol 100, 28-35.
- Papaemmanuil E, Gerstung M, Malcovati L, Tauro S. 2013. "Clinical and biological implications of driver mutations in myelodysplastic syndromes." *Blood* Vol 122, 3616-3627.
- Parajdi LG, Precup R, Bonci EA, Tomuleasa C. 2020. "A Mathematical Model of the Transition from Normal Hematopoiesis to the Chronic and Accelerated-Acute Stages in Myeloid Leukemia." *Mathematics* Vol 8, 376.
- Pardanani A, Finke C, Lasho TL, Al-Kali A et al. 2012. "IPSS-independent prognostic value of plasma CXCL10, IL-7 and IL-6 levels in myelodysplastic syndromes." *Leukemia* Vol 26, 693-699.
- Parish, CR. 1999. "Fluorescent dyes for lymphocyte migration and proliferation studies." *Immunol Cell Biol* 77:499.
- Park ES, Rabinovsky R, Carey M, Hennessy BT et al. 2010. "Integrative analysis of proteomic signatures, mutations, and drug responsiveness in the NCI60 cancer cell lines set." *Molec Cancer Therap* Vol 9, 257-267.
- Patnaik MM, Lasho TL, Hodnefield JM, Knudson RA et al. 2012. "SF3B1 mutations are prevalent in myelodysplastic syndromes with ring sideroblasts but do not hold independent prognostic value." *Blood* Vol 119, 569-572.
- Pegalajar-Jurado A, Fitzgerald BL, Islam N, Blisle JT et al. 2018. "Identification of urine metabolites as biomarkers of early Lyme disease." *Sci Rep* 12204.
- Pellagatti A, Cazzola M, Giagounidis A, Perry J et al. 2010. "Deregulated gene expression pathways in myelodysplastic syndrome hematopoietic stem cells." *Leukemia* Vol 24, 756-764.
- Petzer AL, Gunsilius E. 2003. "Hematopoietic stem cells in chronic myeloid leukemia." *Arch Med Res* Vol 34, 496-506.
- Phosphosite. 2017. *Interleukin-3*. Accessed May 9, 2017.
<http://www.phosphosite.org/proteinAction?id=14564800&showAllSites=true>.
- Piacibello W, Sanavio F, Bresso P, Severino A et al. 1994. "Stem cell factor improvement of proliferation and maintenance of hemopoietic progenitors in myelodysplastic syndromes." *Leukemia* Vol 8, 250-257.
- Pietras EM, Warr MR, Passegué E. 2011. "Cell cycle regulation in hematopoietic stem cells." *J Cell Biol* Vol 195, 709-720.

- Poon Z, Dighe N, Venkatesan SS, Cheung AMS et al. 2018. "Bone marrow MSCs in MDS: contribution towards dysfunctional hematopoiesis and potential targets for disease response to hypomethylating therapy." *Leukemia*.
- Poulaki A, Katsila T, Stergiou IE, Giannouli S et al. 2020. "Bioenergetic Profiling of the Differentiating Human MDS Myeloid Lineage with Low and High Bone Marrow Blast Counts." *Cancers* Vol 12, 3520.
- Prystowsky MB, Otten G, Naujokas MF, Vardiman J, Ihle JN, Goldwasser E, Fitch FW. 1984. "Multiple hemopoietic lineages are found after stimulation of mouse bone marrow precursor cells with interleukin 3." *Am J Pathol* 17:171.
- Quesenberry PJ, Ihle JN, McGrath E. 1985. "The effect of interleukin-3 and GM-CSA-2 on megakaryocyte and myeloid clonal colony formation." *Blood* 65:214.
- Quiroga-Campano AL, Mantalaris A, Kostoglou M, Georgiadis M et al. 2017. "Precision chemotherapy for patients with acute myeloid leukemia using the π Chemo prediction and optimization tool." *Blood* Vol 130, 2074.
- Raaijmakers MHGP, Mukherjee S, Guo S, Zhang S. 2010. "Bone progenitor dysfunction induces myelodysplasia and secondary leukaemia." *Nature* Vol 464, 852-857.
- Ragon BK, Hoff FW, Hu CW, Qiu Y et al. 2016. "Functional proteomic profiling of myelodysplastic syndromes reveals distinct protein expression signatures and identifies candidate therapeutic targets." *Blood* Vol 128, 1975.
- Randall MJ, Jungel A, Rimann M, Wuertz-Kozak K. 2018. "Advances in the Biofabrication of 3D Skin in vitro: Healthy and Pathological Models." *Front Bioeng Biotechnol* Vol 6, 154.
- Ravi M, Paramesh V, Kaviya SR, Anuradha E, Solomon FDP. 2015. "3D Cell Culture Systems: Advantages and Applications." *J Cell Physiol* Vol 230, 16-26.
- Raymakers R, de Witte T, Joziassse J, van der Lely N et al. 1991. "In vitro growth pattern and differentiation predict for progression of myelodysplastic syndromes to acute nonlymphocytic leukaemia." *Br J Haematol* Vol 78, 35-41.
- Reinhart D, Damjanovic L, Kaisermayer C, Kunert R. 2015. "Benchmarking of commercially available CHO cell culture media for antibody production." *Appl Microbiol Biotechnol* 99:4645.
- Rende M, Mortera-Blanco T, Aqel N, Ryley S et al. 2012. "Ex vivo three-dimensional (3D) functional biomimicry for the long-term cytokine-free culture of human primary acute myeloid leukaemia (AML)." *Blood* Vol 120, 4896.
- Rhodes MM, Kopsombut P, Bondurant MC, Price JO et al. 2008. "Adherence to macrophages in erythroblastic islands enhances erythroblast proliferation and increases erythrocyte production by a different mechanism than erythropoietin." *Blood* Vol 111, 1700-1708.
- Ribeil JA, Zermati Y, Vandekerckhove J, Cathelin S et al. 2007. "Hsp70 regulates erythropoiesis by preventing caspase-3-mediated cleavage of GATA-1." *Nature* Vol 445, 102-105.

- Riedl A, Schleder M, Pudenko K, Stadler M et al. 2017. "Comparison of cancer cells in 2D vs 3D culture reveals differences in AKT–mTOR–S6K signaling and drug responses." *J Cell Sci* Vol 130, 203-218.
- Robb, L. 2007. "Cytokine receptors and hematopoietic differentiation." *Oncogene* Vol 26, 6715-6723.
- Roberts, Andrew. 2005. "G-CSF: A key regulator of neutrophil production, but that's not all!" *Growth Factors* Vol 23, 33-41.
- Romano O, Petiti L, Felix T, Meneghini V et al. 2020. "GATA Factor-Mediated Gene Regulation in Human Erythropoiesis." *iScience* Vol 23, 101018.
- Romer-Seibert J, Meyer SE. 2021. "Genetic heterogeneity and clonal evolution in acute myeloid leukemia." *Curr Opin Hematol* Vol 28, 64-70.
- Roschke AV, Tonon G, Gehlhaus KS, McTyre N et al. 2003. "Karyotypic complexity of the NCI-60 drug-screening panel." *Cancer Res* Vol 15, 8634-8647.
- Rubin. 2007. *Rubin's Pathology: Clinicopathologic Foundations of Medicine*. Lippincott Williams & Wilkins.
- Rusten LS, Jacobsen FW, Lesslauer W, Loetscher H, Smeland EB, Jacobsen SE. 1994. "Bifunctional effects of tumor necrosis factor alpha (TNF alpha) on the growth of mature and primitive human hematopoietic progenitor cells: involvement of p55 and p75 TNF receptors." *Blood* Vol 83, 3152-3159.
- Saini, Sharanjot. 2016. "PSA and beyond: alternative prostate cancer biomarkers." *Cell Oncol (Dordr)* Vol 36, 97-106.
- Sajid, Zamra. 2021. "Mathematical Modelling of the Dynamics of Pathological Hematopoiesis during Treatments." *Institut for Naturvidenskab og Miljø Matematik og Fysik (IMFUFA) - PhD Thesis*.
- Santillan, M. 2008. "On the Use of the Hill Functions in Mathematical Models of Gene Regulatory Networks." *Math Mod Nat Phen* 3:85.
- Santini V, Melnick A, Maciejewski JP, Duprez E et al. 2013. "Epigenetics in focus: Pathogenesis of myelodysplastic syndromes and the role of hypomethylating agents." *J Crit Rev Onc* Vol 88, 3152-3159.
- Santos JD, Enfield L, Santos SBD, Allenby MC et al. 2017. "Primary chronic lymphocytic leukemia cells can be maintained long-term in serum-free, cytokine-free 3D culture." *Blood* Vol 130, 2989.
- Sawanobori M, Yamaguchi S, Hasegawa M, Inoue M et al. 2003. "Expression of TNF receptors and related signaling molecules in the bone marrow from patients with myelodysplastic syndromes." *Leuk Res* Vol 27, 583-591.
- Schindelin J, Arganda-Carreras I, Frise E, Kaynig V et al. 2012. "Fiji: an open-source platform for biological-image analysis." *Nat Methods* Vol 9, 676-682.

- Schmalbrock LK, Dolnik A, Cocciardi S, Sträng E et al. 2021. "Clonal evolution of acute myeloid leukemia with FLT3-ITD mutation under treatment with midostaurin." *Blood* Vol 137, 3093-3104.
- Schuettpeitz LG, Link DC. 2013. "Regulation of hematopoietic stem cell activity by inflammation." *Front Immunol* Vol 4, 204.
- Schwartz JR, Ma J, Lamprecht T, Walsh M et al. 2017. "The genomic landscape of pediatric myelodysplastic syndromes." *Nat Comm* Vol 8, 1557.
- Selleri C, Sato T, Anderson S, Young NS. 1995. "Interferon-gamma and tumor necrosis factor-alpha suppress both early and late stages of hematopoiesis and induce programmed cell death." *J Cell Physiol* Vol 165, 538-546.
- Sen S, Wang H, Nghiem CL, Zhou K et al. 2011. "The ribosome-related protein, SBDS, is critical for normal erythropoiesis." *Blood* Vol 118, 6407-6417.
- Shetty V, Mundle S, Alvi S, Showel M et al. 1996. "Measurement of apoptosis, proliferation and three cytokines in 46 patients with myelodysplastic syndromes." *Leuk Res* Vol 20, 891-900.
- Shiozawa Y, Malcovati L, Galli A, Pellagatti A et al. 2017. "Gene expression and risk of leukemic transformation in myelodysplasia." *Blood* Vol 130, 2642-2653.
- Shoemaker, RH. 2006. "The NCI60 human tumor cancer cell line anticancer drug screen." *Nature Rev Cancer* Vol 6, 813-823.
- Singer D, Cooper M, Maniatis GM, Marks PA et al. 1974. "Erythropoietic differentiation in colonies of cells transformed by Friend virus." *Proc Natl Acad Sci USA* Vol 71, 2668-2670.
- Sloand EM, Barrett AJ. 2010. "Immunosuppression for myelodysplastic syndrome: how bench to bedside to bench research led to success." *Hematol Oncol Clin North Am* Vol 24, 331-341.
- Snoep JL, Mrwebi M, Schuurmans JM, Rohwer JM, Mattos MJT. 2009. "Control of specific growth rate in *Saccharomyces cerevisiae*." *Microbiology* 155:1699.
- Sobol, IM. 2001. "Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates." *Math Comp Simul* Vol 55, 271-280.
- Soligo DA, Campiglio S, Servida F, Bossolasco P et al. 1996. "Response of myelodysplastic syndrome marrow progenitor cells to stimulation with cytokine combinations in a stroma-free long-term culture system." *Br J Haematol* Vol 92, 548-558.
- Song WJ, Sullivan MG, Legare RD, Hutchings S et al. 1999. "Haploinsufficiency of CBFA2 causes familial thrombocytopenia with propensity to develop acute myelogenous leukaemia." *Nat Genet* Vol 23, 166-175.
- Steensma, DP. 2015. "Myelodysplastic syndromes: diagnosis and treatment." *Mayo Clin Proc* Vol 90, 969-983.

- Stein, RS. 2003. "The role of erythropoietin in the anemia of myelodysplastic syndrome." *Clin Lymphoma* 4:S36.
- Stevens BM, Khan N, D'Alessandro A, Nemkov T et al. 2018. "Characterization and targeting of malignant stem cells in patients with advanced myelodysplastic syndromes." *Nature Communications* Vol 9, 3694.
- Stiehl T, Baran N, Ho AD, Marciniak-Czochra A. 2014. "Clonal selection and therapy resistance in acute leukaemias: mathematical modelling explains different proliferation patterns at diagnosis and relapse." *J R Soc Interface* Vol 94, 20140079.
- Stocking C, Ostertag Q. 1990. "Interleukin 3: A multilineage Hematopoietic Growth Factor." In *Growth Factors, Differentiation Factors, and Cytokines*, by A Habenicht, 115-128. Berlin: Springer-Verlag.
- Su G, Burant CF, Beecher CW, Athey BD et al. 2011. "Integrated metabolome and transcriptome analysis of the NCI60 dataset." *BMC Bioinformatics* Vol 12. S36.
- Sun T, Jackson S, Haycock JW, MacNeil S. 2006. "Culture of skin cells in 3D rather than 2D improves their ability to survive exposure to cytotoxic agents." *J Biotechnol* Vol 122, 372-381.
- Suntharalingam G, Perry MR, Ward S, Brett SJ, Castello-Cortes A, Brunner MD, Panoskaltsis N. 2006. "Cytokine Storm in a Phase 1 Trial of the Anti-CD28 Monoclonal Antibody TGN1412." *N Engl J Med* Vol 355, 1018-1028.
- Suzuki M, Kobayashi-Osaki M, Tsutsumi S, Pan X et al. 2013. "GATA factor switching from GATA2 to GATA1 contributes to erythroid differentiation." *Genes Cells* Vol 18, 921-933.
- Sweet K, Zhang L, Pinilla-Ibarz J. 2013. "Biomarkers for determining the prognosis in chronic myelogenous leukemia." *J Hematol Oncol* Vol 6, 54 .
- Swerdlow SH, Campo E, Harris NL, Jaffe ES et al. 2008. *WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues, Fourth Edition*. Lyon, France: IARC Press.
- Syed RS, Reid SW, Cheetham JC, Aoki KH et al. 1998. "Efficiency of signalling through cytokine receptors depends critically on receptor orientation." *Nature* Vol 395, 511-516.
- T Haferlach, Y Nagata, V Grossmann, Y Okuno et al. 2014. "Landscape of genetic lesions in 944 patients with myelodysplastic syndromes." *Leukemia* Vol 28, 241-247.
- Takahashi T, Suwabe N, Dai P, Yamamoto M et al. 2000. "Inhibitory interaction of c-Myb and GATA-1 via transcriptional co-activator CBP." *Oncogene* Vol 19, 134-140.
- Takamoto M, Sugane K. 1995. "Synergism of IL-3, IL-5, and GM-CSF on eosinophil differentiation and its application for an assay of murine IL-5 as an eosinophil differentiation factor." *Immunol Lett* 45:43.
- Takeshita A, Shinjo K, Naito K, Ohnishi K, Higuchi M, Ohno R. 2002. "Erythropoietin receptor in myelodysplastic syndrome and leukemia." *Leuk Lymphoma* 43:261.

- Takizawa H, Regoes RR, Boddupalli CS, Bonhoeffer S, Manz MG. 2011. "Dynamic variation in cycling of hematopoietic stem cells in steady state and inflammation." *J Exp Med* 208:273.
- Tan W, He J, Deng J, Yang X et al. 2018. "Small molecule metabolite biomarkers for hepatocellular carcinoma with bile duct tumor thrombus diagnosis." *Sci Rep* 3309.
- Tang Y, Liu J, Zhang D, Xu Z et al. 2020. "Cytokine Storm in COVID-19: The Current Evidence and Treatment Strategies." *Front Immunol* Vol 11, 1708.
- Thanopolou E, Cashman J, Kakagianne T, Eaves A et al. 2004. "Engraftment of NOD/SCID-beta2 microglobulin null mice with multilineage neoplastic cells from patients with myelodysplastic syndrome." *Blood* Vol 103, 4285-4293.
- Thawani N, Tam M, Chang KH, Stevenson MM. 2006. "Interferon-gamma mediates suppression of erythropoiesis but not reduced red cell survival following CpG-ODN administration in vivo." *Exp Hematol* Vol 34, 1451-1461.
- Theocharides AP, Rongvaux A, Fritsch K, Flavell RA, et al. 2016. "Humanized hemato-lymphoid system mice." *Haematologica* Vol 101, 5-19.
- Thoma CR, Zimmermann M, Agarkova I, Kelm JM, Krek W. 2014. "3D cell culture systems modeling tumor growth determinants in cancer target discovery." *Adv Drug Del Rev* Vol 69, 29-41.
- Tothova Z, Steensma DP, Ebert BL. 2013. "New strategies in myelodysplastic syndromes: application of molecular diagnostics to clinical practice." *Clin Cancer Res* Vol 19, 1637-1643.
- Trakarnsanga K, Griffiths RE, Wilson MC, Blair A et al. 2017. "An immortalized adult human erythroid line facilitates sustainable and scalable generation of functional red cells." *Nat Comms* 14750.
- Tripodo C, Sangaletti S, Piccaluga PP, Prakash S et al. 2011. "The bone marrow stroma in hematological neoplasms - a guilty bystander." *Nat Rev Clin Oncol* Vol 29, 456-466.
- Tsiftoglou AS, Vizirianakis IS, Strouboulis J. 2009. "Erythropoiesis: model systems, molecular regulators, and developmental programs." *IUBMB Life* Vol 61, 800-830.
- Tusher VG, Tibshirani R, Chu G. 2001. "Significance analysis of microarrays applied to the ionizing radiation response." *PNAS* Vol 98, 5116-5121.
- Tuval A, Shlush LI. 2019. "Evolutionary trajectory of leukemic clones and its clinical implications." *Haematologica* Vol 109, 872-880.
- van Galen, P, A Kreso, N Mbong, DG Kent, T Fitzmaurice, JE Chambers, S Xie, et al. 2014. "The unfolded protein response governs integrity of the haematopoietic stem-cell pool during stress." *Nature* 268-271.
- Vardiman JW, Thiele J, Arber DA, Brunning RD et al. 2009. "The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes." *Blood* Vol 114, 937-951.

- Varnum-Finney B, Purton LE, Yu M, Brashem-Stein C et al. 1998. "The Notch ligand, Jagged-1, influences the development of primitive hematopoietic precursor cells." *Blood* Vol 91, 4084-4091.
- Velten L, Haas SF, Raffel S, Blaszkiewicz S et al. 2017. "Human haematopoietic stem cell lineage commitment is a continuous process." *Nat Cell Bio* Vol 19, 271-281.
- Venkateswaran N, Venkatachalam P, Vettriselvi V, Paul SFD et al. 2016. "Overexpression of pro-inflammatory cytokines in myelodysplastic syndrome (MDS-RA)." *J Hematol Thrombo Dis* Vol 4, 231.
- Vergara M, Becera S, Berrios J, Osses N, Reyes J, Rodríguez-Moyá M, Gonzalez R, Altamirano C. 2014. "Differential Effect of Culture Temperature and Specific Growth Rate on CHO Cell Behavior in Chemostat Culture." *PLoS One* 9:e93865.
- Vosberg S, Greif PA. 2019. "Clonal evolution of acute myeloid leukemia from diagnosis to relapse." *Genes Chromosomes Cancer* Vol 58, 839-849.
- Walenda T, Stiehl T, Braun H, Frobel J et al. 2014. "Feedback Signals in Myelodysplastic Syndromes: Increased Self-Renewal of the Malignant Clone Suppresses Normal Hematopoiesis." *PLoS Comput Biol* Vol 10, e1003599.
- Wallvik J, Stenke L, Bernell P, Nordahl G, Hippe E, Hast R. 2002. "Serum erythropoietin (EPO) levels correlate with survival and independently predict response to EPO treatment in patients with myelodysplastic syndromes." *Eur J Haematol* 68:180.
- Walter MJ, Ding L, Shen D, Shao J et al. 2011. "Recurrent DNMT3A mutations in patients with myelodysplastic syndromes." *Leukemia* Vol 25, 1153-1158.
- Wang M, Zhao J, Zhang L, Wei F et al. 2017. "Role of tumor microenvironment in tumorigenesis." *J Cancer* Vol 8, 761-773.
- Weiss MJ, dos Santos CO. 2009. "Chaperoning erythropoiesis." *Blood* Vol 113, 2136-2144.
- Weiss, JN. 1997. "The Hill equation revisited: uses and misuses." *FASEB J* 11:835.
- Westers TM, Ireland R, Kern W, Alhan C et al. 2012. "Standardization of flow cytometry in myelodysplastic syndromes: a report from an international consortium and the European LeukemiaNet Working Group." *Leukemia* Vol 26, 1730-1741.
- Weston SA, Parish CR. 1990. "New fluorescent dyes for lymphocyte migration studies. Analysis by flow cytometry and fluorescence microscopy." *J Immunol Meth* 133:87.
- Whichard ZL, Sarkar CA, Kimmel M, Corey SJ. 2010. "Hematopoiesis and its disorders: a systems biology approach." *Blood* Vol 115, 2339-2347.
- Whiteside, TL. 2008. "The tumor microenvironment and its role in promoting tumor growth." *Oncogene* Vol 27, 5904-5912.
- Wilding JL, Bodmer WF. 2014. "Cancer cell lines for drug discovery and development." *Cancer Res* Vol 74, 2377-2384.
- Wilkinson, D. 2006. *Stochastic Modelling for Systems Biology*. New York: NY: Chapman&Hall/CRC.

- Will B, Zhou L, Vogler TO, Ben-Neriah S et al. 2012. "Stem and progenitor cells in myelodysplastic syndromes show aberrant stage-specific expansion and harbor genetic and epigenetic alterations." *Blood* Vol 120, 2076-2086.
- Will B, Zhou L, Vogler TO, Ben-Neriah S, Schinke C, Tamari R, Yu Y, Bhagat TD, Bhattacharyya S, Barreyro L, Heuck C, Mo Y, Parekh S, McMahon C, Pellagatti A, Boulwood J, Montagna C, Silverman L, Maciejewski J, Grealley JM, Ye BH, List AF, SC, SU, Verma A. 2012. "Stem and progenitor cells in myelodysplastic syndromes show aberrant stage-specific expansion and harbor genetic and epigenetic alterations." *Blood* 120:2076.
- Wishart G, Gupta P, Schettino G, Nisbet A, Velliou EG. 2021. "3d tissue models as tools for radiotherapy screening for pancreatic cancer." *Br J Radiol* 92:20201397.
- Wolf BC, Banks PM, Mann RB, Neiman RS. 1988. "Splenic hematopoiesis in polycythemia vera. A morphologic and immunohistologic study." *Am J Clin Pathol* Vol 89, 69-75.
- Wu H, Liu X, Jaenisch R, Lodish HF. 1995. "Generation of committed erythroid BFU-E and CFU-E progenitors does not require erythropoietin or the erythropoietin receptor." *Cell* 83:59.
- Wu M, Swartz MA. 2014. "Modeling tumor microenvironments in vitro." *J Biomech Eng* Vol 136, 0210111-0210117.
- Wu SJ, Kuo YY, Hou HA, Li LY et al. 2012. "The clinical implication of SRSF2 mutation in patients with myelodysplastic syndrome and its stability during disease evolution." *Blood* Vol 120, 3106-3111.
- Yanagawa S, Hirade K, Ohnata H, Sasaki R, Chiba H, Ueda M, Goto M. 1984. "Isolation of human erythropoietin with monoclonal antibodies." *J Biol Chem* 259:2707.
- Yang B, Wang C, Xu L, Wu X et al. 2018. "Monitoring tyrosine kinase inhibitor therapeutic responses with a panel of metabolic biomarkers in chronic myeloid leukemia patients." *Cancer Sci* Vol 109, 777-784.
- Yiu HH, Graham AL, Stengel RF. 2012. "Dynamics of a Cytokine Storm." *PLoS One* Vol 7, e45027.
- Young JC, Varma A, DiGiusto D, Backer MP. 1996. "Retention of quiescent hematopoietic cells with high proliferative potential during ex vivo stem cell culture." *Blood* 87:545.
- Young, JC, A Varma, D DiGusto, and MP Backer. 1996. "Retention of quiescent hematopoietic cells with high proliferative potential during ex-vivo stem cell culture." *Blood* 545-556.
- Yu J, Jiang P, Sun H, Zhang X et al. 2020. "Advances in targeted therapy for acute myeloid leukemia." *Biomarker Research* Vol 8, 17.
- Zhang CC, Lodish HF. 2008. "Cytokines regulating hematopoietic stem cell function." *Curr Opin Hematol* Vol 15, 307-311.
- Zhang JM, An J. 2007. "Cytokines, Inflammation and Pain." *Int Anesthesiol Clin* Vol 45, 27-37.

- Zhang L, Padron E, Lancet J. 2015. "The molecular basis and clinical significance of genetic mutations identified in myelodysplastic syndromes." *Leuk Res* Vol 39, 6-17.
- Zhang P, Zhang C, Li J, Han J et al. 2019. "The physical microenvironment of hematopoietic stem cells and its emerging roles in engineering applications." *Stem Cell Research & Therapy* Vol 327, 1422-1427.
- Zhang Y, Harada A, Bluethmann H, Wang JB et al. 1995. "Tumor necrosis factor(TNF) is a physiologic regulator of hematopoietic progenitor cells: increase of early hematopoietic progenitor cells in TNF receptor p55-deficient mice in vivo and potent inhibition of progenitor cell proliferation by TNFalpha in vivo." *Blood* Vol 86, 2930-2937.
- Zhao LR, Piao CS, Murikinati SR, Gonzalez-Toledo ME. 2013. "The Role of Stem Cell Factor and Granulocyte-Colony Stimulating Factor in Treatment of Stroke." *Recent Pat CNS Drug Discov* Vol 8, 2-12.
- Zhu J, Emerson SG. 2002. "Hematopoietic cytokines, transcription factors and lineage commitment." *Oncogene* Vol 21, 3295-3313.
- Ziehn T, Tomlin AS. 2009. "GUI-HDMR - A software tool for Global Sensitivity Analysis of complex models." *J Env Mod Soft*.

Appendix I – Model Assumptions and Biological Context

For a given cell type, the model considers a fraction of the cells undergoing each one of these processes, while each process is associated with a specific rate. For *in-vitro* scenarios these specific rates are not available in literature, since they are very dependent on parameters such as the culture medium used (serum/xeno-free vs serum/xeno-containing), feeding regime (batch vs replenishment), cytokine supplementation (with or without cytokines) and starting cell population (pure progenitors or mononuclear cells), among other factors. At a first approximation and when available, one would use the rates measured *in-vivo* or, when unavailable, will use *in-vitro* parameters and/or produce educated guesses based on the literature.

Normal Erythropoiesis

This model considers the action of three different cytokines on the healthy erythroid cell lineage. SCF and IL3 are known to be early-acting, unspecific cytokines, whereas EPO is known to be the main erythropoietic cytokine, indispensable for lineage commitment and terminal maturation. Due to its action later on the differentiation process, EPO is also classified as a late-acting cytokine (Metcalf 2008). The proposed preliminary version of the model includes three different types of cytokines: an early-acting, unspecific cytokine with a positive proliferative effect (SCF), an early-acting, unspecific cytokine with both positive and negative proliferative effects (IL3) and a late-acting, specific cytokine with positive effects in the differentiation and proliferation of erythroid progenitors (EPO).

With respect to the fractions and rates, there is very limited information in the literature. The majority of the existing papers concerning erythroid differentiation aim either to produce RBC without measuring any fractions/rates (Giarratana MC 2011) or focus on fundamental studies in developmental biology (Hattangadi SM 2011) as well as homeostatic and stress erythropoiesis (Koury 2015). Gathering insights into the biology of the cytokines and their action mechanisms

may potentially help one to produce educated guesses as what the values for cell rates and fractions should be.

Stem Cell Factor (SCF)

- SCF is a dimeric protein produced by fibroblasts and endothelial cells with about 18,500 Da in its soluble form (Huang E 1990).
- Healthy human serum contains SCF at a concentration of approximately 3.3 ng/mL (Broudy 1997);
- When expressed in CHO cells, this yields a protein with 28,000 to 40,000 Da due to heavy and non-homogeneous glycosylation (Arakawa T 1991);
- SCF is known to act mainly on haematopoietic stem cells and multipotent hematopoietic progenitors, by accelerating their entry into cell cycle (Leary AG 1992);
- It is believed that a continuous and stable *in-vivo* stem cell self-renewal is dependent on SCF (Miller CL 1997);
- About 60 to 75% of all human CD34⁺ cells express the *c-kit* receptor, to which SCF binds specifically (Ashman LK 1991) (Broudy VC 1992);
- SCF's receptor saturation constant was measured in patients with Polycythemia Vera and compared with healthy controls by Scatchard plot analysis with 125I-labelled SCF. The value of 10 ng/mL was found to be the same despite possessing a healthy or diseased condition (Dai CH 1994).
- Used alone, SCF has a modest effect on increasing both the size and the number of colonies (Broudy 1997);
- Used *in-vitro* cooperatively with IL3, it increases the numbers of BFU-E from 20-fold (Migliaccio G 1992) to about 200-fold, when combined with four other cytokines (Henschler R 1994).

- SCF does not influence cells after the CFU-E stage and its receptor expression decays rapidly on the erythroid lineage (Testa U 1996).

Interleukin-3 (IL3)

- IL3 is a monomer with approximately 17,250 Da (Phosphosite 2017) and is produced by activated T-cells (Dorssers L 1987);
- IL3, once called multi-colony stimulating factor (multi-CSF) supports the survival and proliferation of the myeloid progeny (Stocking C 1990);
- When used *in-vitro*, it increases the numbers of CFU-GEMM and CFU-GM cells (Garland JM 1983), as well as cell numbers and maturation stage across the granulocytic (Koike K 1986), monocyte-macrophage (Prystowsky MB 1984), eosinophil (Takamoto M 1995), megakaryocytic (Hagiwara S 1997) and mast cell lineages (Quesenberry PJ 1985);
- With respect to the erythroid lineage, IL3 is shown to increase the numbers of BFU-E cells while having an inhibitory effect on the maturation and proliferation of CFU-E cells and their progeny (Leonard JP 1988);
- IL3's receptor (IL3R) is an heterodimeric cytokine receptor, with an alpha subunit (IL3R α) comprising the specific substrate binding locus (low-affinity) while the beta subunit (IL3R β), necessary for signal transduction, is shared with IL-5 and GM-CSF receptors (Kitamura T 1991) (Hara T 1996);
- IL3R's expression on different haematopoietic lineages and progeny decays rapidly with maturation on the erythroid and megakaryocytic lineages, while having a more moderate decrease on the monocytic lineage and being conserved throughout the granulocytic lineage (Testa U 1996);
- Using a Scatchard plot analysis with ¹²⁵I-labelled IL3 the receptor's half-saturation constant was reported to be 2.85nM (Itoh N 1990) or 49.11405 ng/mL;

- EPOR differentiating signals are inhibited by IL3 via the IL3R α subunit (Krosi J 1996);
- When used separately and in combination with SCF for the culture of isolated CD34⁺ cells from healthy peripheral blood, IL3-containing medium yielded twice as much cells than the SCF-only containing medium, while IL3-only medium yielded the same number of cells as the SCF-only condition (Brugger W 1993). This may provide evidence that the effect of SCF and IL3 alone on CD34⁺ cells has the same order of magnitude, and that its combined effect yields an effect twice as intense.

Erythropoietin (EPO)

- Erythropoietin is a 30,400 Da protein (Miyake T 1977) produced mainly in the kidneys (W. Jelkmann 1992);
- Its concentration *in-vivo* is inversely proportional to the O₂ concentration in the blood (Beru N 1986) and its serum concentration in healthy individuals ranges from 6-32 mU/mL (Cotes 1982) (Egrie JC 1987);
- Its specific activity is believed to be between 70,000 – 80,000 U/mg, as per isolation and assessment of human EPO from urine (Yanagawa S 1984);
- When used *in-vivo*, EPO infusions increase BFU-E and CFU-E numbers. Studies using 150-300 U/kg intravenously administered three times a week shown a 3.37 and 4.15-fold increase on the BFU-E and CFU-E absolute numbers, respectively (Dessypris E 1988);
- EPO Receptor (EPOR) is poorly expressed on CD34⁺ cells, but its expression starts increasing at the BFU-E stage and is kept constant throughout the erythropoietic progeny (Testa U 1996);
- The homeostatic generation of BFU-E and CFU-E cells *in-vivo* is not dependent of EPO nor EPOR; EPO's effect on a given erythroid progenitor correlates well with its EPOR expression. (Wu H 1995);

- EPOR's half-saturation constant was determined using BaF3 cells (murine IL3-dependant cell line) and its value is 164 ± 28 pM, which works out as 350 ± 60 mU/mL (Becker V 2010);

Fractions and Rates

The growth (proliferation) rate is a well-studied concept for multiple organisms, with vast literature focusing on prokaryotic cells like *E.coli* (Marr 1991) (Cox 2004) and *A. tumefaciens* (Bell 1990) (Leth IK 2017) as well as eukaryotic cells like *S. cerevisiae* (Hoek PV 1998) (Snoep JL 2009) and even mammalian cells like Chinese Hamster Ovary (CHO) (Vergara M 2014) (Reinhart D 2015), to name a few. Despite being very different organisms, determining their growth rate is constrained by the same conditions – it is strongly dependent on the culture medium composition temperature, pH, cell number and concentration as well as feeding regime. To address these issues Monod developed the well-known Monod kinetic model (Monod 1958) which correlates the growth rate of a given organism with one (the limiting) substrate and the maximum growth rate. His work was further extended to include the influence of inhibitory components (Levenspiel 1980) and it is widely accepted as a very fair approximation of cell growth kinetics (Merchuk JC 1995). However, this concept is not easily applicable to stem cell cultures, since typically their growth does not only depend on one substrate – which makes it more difficult to describe growth with Monod kinetics. Spontaneous differentiation also occurs on stem cell cultures, which adds another degree of complexity to the whole process.

This does not necessarily mean it is impossible to determine, in a relatively accurate fashion, the relevant rates and parameters for these cells, given a set of experimental conditions. Generically speaking, bromodeoxyuridine (BrdU), a thymidine analogue, is a common chemical compound used to detect proliferating cells *in-vivo/in-vitro* (Lehner B 2011). It is incorporated by actively cycling cells during S phase and can be passed on to daughter cells (Kee N 2002). Carboxyfluorescein succinimidyl ester (CFSE) is another chemical commonly used to track cell division and daughter cell generations, using flow cytometry (Parish 1999). Due to its cell permeability and covalent binding between its succinimidyl group and intracellular molecules, CFSE is stable for many months (Weston SA 1990) and can be used to track up to 8 cell divisions

due to the progressive halving of CFSE signal intensity with each cell division (Lyons AB 1994). The combination of BrdU, CFSE and cell-specific membrane stains allows identifying and tracking a specific cell population's division and proliferation behavior, which is useful to determine cell fractions and differentiation/proliferation rates.

When considering the haematopoietic lineage, little information is known about specific cell fractions and/or rates. Studies reveal that the proliferation rate increases with the degree of maturity (Möbest D 1999), as cells tend to cycle faster (Young JC 1996). Also, *in-vivo* differentiation of erythroblasts typically starts with an early increase in proliferation (Akker EVD 2010) which eventually decreases and ends at the orthochromatic erythroblast (OrthoE) stage, from which the cell enucleates and exits the bone marrow as a reticulocyte (Spivak 2005). Both BrdU and CFSE labelling have been widely used for antigen-specific CD8 T-cell identification in different animals (Morales H 2008) (McGill J 2009); with respect to tracking haematopoietic cell populations, the majority of the work is performed in murine platforms (Oostendorp RAJ 2000) (Nygren JM 2008) (Takizawa H 2011). Interestingly, a rate of one cell division each 57 days for murine long-term hematopoietic stem cells (LT-HSC) was established by *in-vivo* BrdU labelling of murine hematopoietic stem cells coupled with cell cycle analysis (Cheshier SH 1999), which works out as a rate of proliferation of 0.01754 divisions/day. The same group has also proposed that half of the total LT-HSC will proliferate, while the remainder will either enter apoptosis or differentiate, according to the signals they are provided with.

Diseased Erythropoiesis (MDS)

Stem Cell Factor (SCF)

Literature on SCF and *c-kit* on diseased patients, despite limited, agrees that *c-kit* is not commonly disrupted in MDS. When looking for mutations on the *c-kit* gene of 75 MDS patients, only 8 patients (10.6%) had an aberrant genetic sequence, either on the extracellular, juxtamembrane or tyrosine kinase domain of the *c-kit* gene. Interestingly, the same group found that none of the nine patients with Refractory Anemia (low-risk MDS) had any mutation on *c-kit* (Lorenzo F 2006). Another group analysed a cohort of 28 MDS and 12 mastocytosis patients and verified that, despite being present in all the patients suffering from mastocytosis, none of the MDS patients had any mutation in *c-kit* (Fritsche-Polanz R 2001). Two other studies report that adding SCF to *in-vitro* cultures of bone marrow cells isolated from 25 MDS patients increases the formation of CFU-GEMM and BFU-E by up to 27-fold on 60 and 80% of the patients, respectively (Glinsmann-Gibson B 1994) whereas another study with BM cells from 63 MDS patients, adding SCF (10ng/mL) lead to up to a 26-fold increase on BFU-E colony numbers, particularly in RA patients (Soligo DA 1994). When combined, this information leads to the relatively reasonable assumption that the pair SCF/*c-kit* has comparable functionality in the MDS cells than that of their healthy counterparts.

Interleukin-3 (IL3)

IL3 was evaluated on phase I/II clinical trials as a plausible supportive care therapy in MDS patients. After an initial study where improvements, especially in the granulocytic lineage, were seen in nine myelodysplastic patients (A Ganser 1990) the side therapeutic effects were determined to be relatively severe (Nand S 1994). Another study analysed the administration of both EPO and IL3, yielding an increased reticulocyte count in 5 (22.7%) out of 22 patients. When used in coordination with G-CSF or GM-CSF, it significantly improved neutrophil counts in 96%

and 78.9% of the recipients, respectively (Musto P 2001); however, its efficacy when combined with EPO did not significantly alleviate symptoms nor guaranteed transfusion independence.

IL3R α (identified by CD123) is known to be over-expressed in about 45% CD34⁺ cells of AML patients, conferring higher proliferation and bone marrow cellularity and strongly correlating with a poor prognosis (Testa U 2002) (Testa U 2004). CD123 was also established as a marker for AML Leukaemic Stem Cells (LSCs), since CD34⁺ CD38⁻ CD123⁺ cells were able to start and maintain a leukaemic state when transplanted to immunocompromised mice (Jordan CT 2000) and is also generally overexpressed in Acute Lymphocytic Leukemia and in Blastic Plasmacytoid Dendritic Cell Neoplasm (BPDCN) (Testa U 2014). With respect to CD123 expression in MDS, a study reports no significant differences between CD123 expression in MDS cells when compared to healthy controls while showing significant differences ($P < 0.05$) with CD123 expression in AML cells (De Smet D 2012). More recently, another study demonstrated that CD123 expression increases with disease progression ultimately leading to a similar CD123 AML expression at high-risk MDS stages, unveiling the possibility for CD123 to be a potential disease progression marker (Li LJ 2014). To our best knowledge, no CD123 functionality studies were performed so far using MDS cells, which also opens a research opportunity. Thus, we assume that the effects of IL3 are similar on MDS cells than that of their healthy counterparts, and that CD123 functionality is not affected, despite being overexpressed in late, high-risk MDS stages.

Erythropoietin (EPO)

A study as early as 1991 was able to conclude that anemia, a hallmark of MDS, is not caused by an endogenous EPO deficiency (Aul C 1991). This pathology is commonly addressed by administration of rhEPO or similar compounds; however, protocols including only EPO result often in a low response rate (16 to 33%) (Stein 2003). A review compiling 205 patients from 17 different studies reported serum EPO levels seem to inversely correlate with response rate, but this factor alone was not sufficient to predict response to EPO administration (Hellström-Lindberg 1995).

This result was confirmed and extended about 10 years later, proposing serum EPO, disease stage and transfusion dependence/independence as indicators for response. The response rate, defined as an increase on haemoglobin by a factor of at least 1.5 g/dL or elimination of red blood cell transfusion dependence was proposed to be 79% (53-93% with 95% CI) for patients with serum EPO concentrations under 50 U/L, RA/RARS disease state and independence of RBC transfusion (Wallvik J 2002). Due to the multiplicity of disease states, genetic mutations, complex karyotypes and clone number/aggressivity, it is still very hard to determine wide correlations for predicting response to EPO therapeutics. A more recent meta-analysis study involving around sixty independent studies was able to conclude that transfusion-independent patients have a 60% reduction on mortality when compared with transfusion-dependent patients (Harnan S 2016).

With respect to the EPO Receptor (EPOR), several studies were performed in order to determine possible causes of EPO irresponsiveness on MDS patients. The restriction of its gene with four different enzymes results in fragments with the same restriction pattern, consistent with healthy controls and the erythroleukaemic cell line K562, yielding that the EPOR gene is intact in MDS patients (Mittelman M 1996). Also, EPOR mRNA expression in MDS cells was not found to be statistically different when compared to healthy controls (Takeshita A 2002). However, a study in 1999 cloned a truncated EPOR sequence in a leukaemic cell line, showing that cells expressing this truncated form (EPOR-T) would undergo apoptosis even in the presence of EPO. A comparison between nine patients with MDS and healthy controls showed that 7 (77.8%) patients expressed this truncated version whereas none of the healthy controls did (Shimizu R 1999). This confirmed the results proposed a decade before where alternative splicing of the EPOR gene gave origin to two splice variants (EPOR-T and EPOR-F), with differential expression across the erythroid lineage (early and late progenitors, respectively), mitogenic effect (transduced by EPOR-T but not by EPOR-F) and predisposition for apoptosis (higher when EPOR-T is present and blockage of the apoptotic signalling, independent of the mitogenic signal transduction when

EPOR-F is present) (Nakamura Y 1992). A clinical trial was set up to assess EPOR sequence and its expression in MDS erythroid progenitors back in 2007, but no study results were published so far (NIH, The Role of Erythropoietin in Myelodysplastic Syndrome 2008). While the causes are unclear at the moment, it is certain that erythroid differentiation is impaired due to a lower response, at the EPOR complex, of EPO-dependant signalling, despite appropriate growth factor production, cell receptor density and display (McGraw KL 2014).

Fractions and Rates

Impaired differentiation and excessive intramedullary apoptosis are one of the pathophysiological hallmarks of MDS (Kerbaui DB 2007). Work performed in the 1960's paved the way for future studies involving labelling of proliferative cells, both *in-vitro* and *in-vivo* (Gavosto F 1964) (Killman 1968). While the majority of the work focused on AML/ALL patients (Mauer AM 1979) (Raza A 1987), the time-consuming nature of the protocols hindered the study of a high number of patients. In 1992, cell cycle kinetics of MDS (5) and AML (5) patients were studied using *in-vivo* administration of BrdU, which allowed for a direct comparison the duration of S-phase and total cell cycle in both pathologies. Despite the low number of patients, this study established a reproducible correlation between the parameters measured using bone marrow biopsies and bone marrow aspirations, within patients with the same pathology (Raza A 1992). Three years later, a similar protocol using *in situ* end-labelling (ISEL) of DNA both for proliferative and apoptotic markers was applied to 35 MDS patients and 15 AML patients, which remarkably concluded that MDS myeloblasts have a higher cycling rate than their AML counterparts, while the apoptotic rates on MDS are higher than the ones found on AML. Ultimately, double-labelling for proliferative and apoptotic markers allowed to determine that the majority of the cells undergoing DNA synthesis (S phase) were simultaneously undergoing apoptosis (Raza A 1995). This work was further expanded two years later with 68 MDS patients, confirming the cell “antonymy” hypothesis proposed earlier by Hibner and Coutinho, where a cell actively synthesizing DNA that receives a

signal to enter quiescence due to any DNA fault or other damage, would stop doing so and enter apoptosis. (Hibner U 1994). Across a representative sample of 10 patients out of 68, Raza and colleagues found that 30 to 90% of MDS cells in S-phase are also undergoing apoptosis whereas AML cells have lost the capacity to do so, generating the hypothesis that this is a self-defence mechanism against disease progression, and its elimination on MDS patients could lead to true remission (Raza A 1997).

Ultimately, it is known that the normal haematopoietic structure is somewhat conserved in MDS (Z. L.-N. Will B 2012), but evidence seems to point towards MDS cells having higher apoptosis rates while having a similar proliferative potential than their healthy counterparts. The paradox between MDS patients' normo/hypercellular bone marrow and peripheral blood cytopenias is a famous hallmark in continuous study by different groups.

Appendix II – Proposed FC panel for the identification of healthy erythroid cells

In-vitro identification of erythroid cellular intermediates is possible by creating complex multicolour Flow Cytometry panels, both for HSCs (van Galen, et al. 2014) and erythroid progenitors (Kaushansky, et al. 2011, 8th Edition) (Gorczyca, et al. 2011) (Fajtova, et al. 2013) (Li, et al. 2014) on different stages of maturation. Studies also report using CD44 as a marker for differential selection of erythroblasts in different degrees of maturation (Chen, et al. 2009). This membrane stain allows separating PolyE from OrthoE since its expression falls considerably from one stage to the other.

Table 16 - Proposed phenotypes for the haematopoietic stem and progenitor cells in the erythropoietic lineage.

Compartments	Cell	Phenotype
<u>Progenitor (P)</u>	HSC	CD34 ⁺ CD38 ⁻ CD90 ⁺
	Multipotent Prog. (MPP)	CD34 ⁺ CD38 ⁻ CD90 ⁻
	Common Myeloid Prog. (CMP)	CD34 ⁺ CD38 ⁺ CD71 ⁻
<u>Committed (C)</u>	BFU-E	CD34 ⁺ CD38 ⁺ CD45 ⁺ CD71 ^{Low}
	CFU-E	CD34 ⁻ CD38 ⁺ CD45 ⁺ CD71 ⁺
	Erythroblasts - Gating	CD45 ^{Low} CD71 ⁺ vs SSC
	Proerythroblasts (ProE)	CD44 ⁺ CD45 ^{Low} FSC ^{Hi}
	Basophilic Erythroblasts (BasoE)	CD44 ⁺ CD45 ⁻ FSC ^{Hi}
	Polychromatophilic Erythroblasts (PolyE)	CD44 ⁺ FSC ^{Med}
	Orthochromatophilic Erythroblasts (OrthoE)	CD44 ⁺ FSC ^{Low}
<u>Terminally Differentiated (T)</u>	Erythrocyte (Red Blood Cell)	CD71 ⁻ CD235a ⁺

Appendix III: Unique metabolites detected

2,4-bis-hydroxybutanoic acid	glutamate	methyl benzoate
2-hydroxybutanoic acid	glutamine	monomethyl P
alanine	glycerol	myoinositol
asparagine	glycerol 2P	N-acetylglutamic acid
aspartate	glycerol 3P	octadecanoic acid
benzoate	glycine	ornithine
cholesterol	glycolic acid	phenylalanine
citrate	hexadecanoic acid	proline
cysteine	homoserine	pyruvate
dopamine	isocitrate	serine
ethanolamine	isoleucine	succinate
fumarate	lactic acid	threonine
GABA	leucine	tryptophan
gluconate	lysine	tyrosine
glucopyranose	malate	urea
glucose	methionine	valine

Appendix IV: Code used

GO:Matlab – Interface between gPROMS and Matlab

```
tic

Counter = 0;

% matrix with min and max values V1 on first row, V2 on 2nd row
Vars=[0 200; 0 500];

step = 50
stepV2 = 50
stepV1 = 50
%step = (min(Vars(:,2)))/10;

% number of model simulations
no_sims = (step+1)^2;

V2 = zeros(step+1,1); %V2 is EPO
for m=2:step+1

    %V2 (m)=m-1;
    V2 (m)= V2 (m-1) + (Vars(2,2) - Vars(2,1))/stepV2;
end

V1 = zeros(step+1,1); %V1 is IL3
for n=2:step+1

    V1(n) = V1(n-1) + (Vars(1,2) - Vars(1,1))/(stepV1);
end

%Cumulative Sampling time vector. For 4000 days sampling each 10 days.
%Last value must be zero.
d=400; % number of days (or time units)
sampling_time = 10*ones(1,d+1);
sampling_time (1,d+1)=0; %last coordinate must always be zero. This tells it
to stop.

no_function_calls = size(sampling_time,2); %The number two means the second
coordinate of the "size" which is the number of columns

%Create a helper for each required output you want to test
Output01Helper = zeros(1,no_function_calls);
Output02Helper = zeros(1,no_function_calls);
Output03Helper = zeros(1,no_function_calls);
Output04Helper = zeros(1,no_function_calls);
Output05Helper = zeros(1,no_function_calls);
Output06Helper = zeros(1,no_function_calls);
Output07Helper = zeros(1,no_function_calls);
Output08Helper = zeros(1,no_function_calls);
Output09Helper = zeros(1,no_function_calls);
```

```
Output10Helper = zeros(1,no_function_calls);
Output12Helper = zeros(1,no_function_calls);
Output13Helper = zeros(1,no_function_calls);

%Again create new lines for the outputs you want to test
Output01 = zeros(no_sims,no_function_calls);
Output02 = zeros(no_sims,no_function_calls);
Output03 = zeros(no_sims,no_function_calls);
Output04 = zeros(no_sims,no_function_calls);
Output05 = zeros(no_sims,no_function_calls);
Output06 = zeros(no_sims,no_function_calls);
Output07 = zeros(no_sims,no_function_calls);
Output08 = zeros(no_sims,no_function_calls);
Output09 = zeros(no_sims,no_function_calls);
Output10 = zeros(no_sims,no_function_calls);
Output11 = zeros(no_sims,no_function_calls);
Output12 = zeros(no_sims,no_function_calls);
Output13 = zeros(no_sims,no_function_calls);

ModelTime = zeros(1,no_function_calls); % Never delete or change this line.

Input = zeros(no_sims,15); %This will store the inputs and the max. output

j=0;

for k=1:no_sims

    gOMATLAB('startONLY');
    %Note: Model4SA is both the name of project and process. Do so too!!!
    gOMATLAB('select', 'Erythropoiesis', 'Erythropoiesis');
    % This function starts the gPROMS simulation. If started successfully the
    % simulation will pause when gPROMS requests inputs from MATLAB
    gOMATLAB('simulate', 'Erythropoiesis');

    s=rem(k,step+1);
    if s==0
        s=step+1;
    end
    if s==1
        j=j+1;
    end

    for i=1:no_function_calls

        % The last coordinate is number of outputs + 1 (modeltime)
        ReturnVector = gOMATLAB('evaluate',[ sampling_time(i) ...
            V1(j) V2(s) ],13);

        ModelTime(i) = ReturnVector(1,1);

        % Again create more lines for each of the required output.
        % The second coordinate in the ReturnVector is Number of outputs + 1
        Output01Helper(i) = ReturnVector(1,2); %RBC
        Output02Helper(i) = ReturnVector(1,3); %Total_C
        Output03Helper(i) = ReturnVector(1,4); %BFU-E %mdsRBC
```

```
Output04Helper(i) = ReturnVector(1,5); %CFU-E %mdsC
Output05Helper(i) = ReturnVector(1,6); %Proe
Output06Helper(i) = ReturnVector(1,7); %Basoe
Output07Helper(i) = ReturnVector(1,8); %Polye
Output08Helper(i) = ReturnVector(1,9); %Orthoe
Output09Helper(i) = ReturnVector(1,10); %MDS_C Cells
Output10Helper(i) = ReturnVector(1,11); %MDS_RBC
Output12Helper(i) = ReturnVector(1,12); %Healthy MPP
Output13Helper(i) = ReturnVector(1,13); %Healthy CMP

end

% Again add more outputs if you have those
Output01(k,:) = [Output01Helper];
Output02(k,:) = [Output02Helper];
Output03(k,:) = [Output03Helper];
Output04(k,:) = [Output04Helper];
Output05(k,:) = [Output05Helper];
Output06(k,:) = [Output06Helper];
Output07(k,:) = [Output07Helper];
Output08(k,:) = [Output08Helper];
Output09(k,:) = [Output09Helper];
Output10(k,:) = [Output10Helper];
Output12(k,:) = [Output12Helper];
Output13(k,:) = [Output13Helper];

gOMATLAB('stop');

Input(k,1)=V1(j);
Input(k,2)=V2(s);

time_passed = toc;

Counter = Counter + 1;

if rem(Counter, 5) == 0

    disp(['Iteration number ' num2str(Counter) ' finshed after '
num2str(time_passed/60) ' mins'])

end

end

clear Output01Helper Output02Helper Output03Helper Output04Helper
Output05Helper Output06Helper Output07Helper Output08Helper Output09Helper
Output10Helper Output11Helper Output12Helper Output13Helper ReturnVector i k

%% Preparing data for processing

for k=1:no_sims
    Input(k,3)=max(Output01(k,:)); %RBC
    Input(k,4)=max(Output02(k,:)); %Total_C
    Input(k,5)=max(Output03(k,:)); %BFUe
    Input(k,6)=max(Output04(k,:)); %CFUe
    Input(k,7)=max(Output05(k,:)); %Proe
```

```

        Input(k,8)=max(Output06(k,:)); %Basoe
        Input(k,9)=max(Output07(k,:)); %Polye
        Input(k,10)=max(Output08(k,:)); %Orthoe
        Input(k,11)=max(Output09(k,:)); %MDS_C Cells
        Input(k,12)=max(Output10(k,:)); %MDS_RBC

Input(k,13)=(max(Output09(k,:))*100/(max(Output09(k,:))+max(Output02(k,:))));
        Output11(k,:)=(Output09(k,:))*100 ./
        ((Output09(k,:))+max(Output02(k,:)));
        Output11(:,1)=zeros(no_sims,1);

        Input(k,14)=max(Output12(k,:)); %Healthy MPP
        Input(k,15)=max(Output13(k,:)); %Healthy CMP
end

%% 3D Plot of the data: max output vs. inputs

SpecsTitle = {'IL3 and EPO (50 ng/mL SCF)'};

SpecsXAxis = {'IL3 (ng/mL)'};

SpecsYAxis = {'EPO (mU/mL)'};

SpecsZAxis = {'Cell Number'};

X=reshape(Input(:,1),step+1,step+1);
Y=reshape(Input(:,2),step+1,step+1);
Z=reshape(Input(:,3),step+1,step+1);
ZA=reshape(Input(:,4),step+1,step+1);
ZB=reshape(Input(:,5),step+1,step+1);
ZC=reshape(Input(:,6),step+1,step+1);
ZD=reshape(Input(:,7),step+1,step+1);
ZE=reshape(Input(:,8),step+1,step+1);
ZF=reshape(Input(:,9),step+1,step+1);
ZG=reshape(Input(:,10),step+1,step+1);
ZH=reshape(Input(:,11),step+1,step+1);
ZI=reshape(Input(:,12),step+1,step+1);
ZJ=reshape(Input(:,13),step+1,step+1);
ZK=reshape(Input(:,14),step+1,step+1);
ZL=reshape(Input(:,15),step+1,step+1);

%3D profiles
figure
surf(X,Y,Z)
title(['RBC vs. ' char(SpecsTitle)],'fontweight','bold','fontsize',12)
xlabel([char(SpecsXAxis)],'fontweight','bold')
ylabel([char(SpecsYAxis)],'fontweight','bold')
zlabel('Cell Number','fontweight','bold')
set(gcf,'Color','White')
set(gca,'linewidth',1.5,'fontweight','bold','outerposition',[0.05 0 0.9 1])
colorbar ('fontweight','bold','position',[0.01 0.1 0.05 0.8])
savefig('RBC')
print('RBC', '-dpng')

```

```
figure
surf(X,Y,ZA)
title(['Committed Cells vs. ' char(SpecsTitle)],'fontweight','bold','fontsize',12)
xlabel([char(SpecsXAxis)],'fontweight','bold')
ylabel([char(SpecsYAxis)],'fontweight','bold')
zlabel('Cell Number','fontweight','bold')
set(gcf,'Color','White')
set(gca,'linewidth',1.5,'fontweight','bold','outerposition',[0.05 0 0.9 1])
colorbar ('fontweight','bold','position',[0.01 0.1 0.05 0.8])
savefig('Committed Cells')
print('Committed Cells', '-dpng')
```

```
figure
surf(X,Y,ZB)
title(['BFU-E vs. ' char(SpecsTitle)],'fontweight','bold','fontsize',12)
xlabel([char(SpecsXAxis)],'fontweight','bold')
ylabel([char(SpecsYAxis)],'fontweight','bold')
zlabel('Cell Number','fontweight','bold')
set(gcf,'Color','White')
set(gca,'linewidth',1.5,'fontweight','bold','outerposition',[0.05 0 0.9 1])
colorbar ('fontweight','bold','position',[0.01 0.1 0.05 0.8])
savefig('BFU-E')
print('BFU-E', '-dpng')
```

```
figure
surf(X,Y,ZC)
title(['CFU-E vs. ' char(SpecsTitle)],'fontweight','bold','fontsize',12)
xlabel([char(SpecsXAxis)],'fontweight','bold')
ylabel([char(SpecsYAxis)],'fontweight','bold')
zlabel('Cell Number','fontweight','bold')
set(gcf,'Color','White')
set(gca,'linewidth',1.5,'fontweight','bold','outerposition',[0.05 0 0.9 1])
colorbar ('fontweight','bold','position',[0.01 0.1 0.05 0.8])
savefig('CFU-E')
print('CFU-E', '-dpng')
```

```
figure
surf(X,Y,ZD)
title(['ProE vs. ' char(SpecsTitle)],'fontweight','bold','fontsize',12)
xlabel([char(SpecsXAxis)],'fontweight','bold')
ylabel([char(SpecsYAxis)],'fontweight','bold')
zlabel('Cell Number','fontweight','bold')
set(gcf,'Color','White')
set(gca,'linewidth',1.5,'fontweight','bold','outerposition',[0.05 0 0.9 1])
colorbar ('fontweight','bold','position',[0.01 0.1 0.05 0.8])
savefig('ProE')
print('ProE', '-dpng')
```

```
figure
surf(X,Y,ZE)
title(['BasoE vs. ' char(SpecsTitle)],'fontweight','bold','fontsize',12)
```

```
xlabel([char(SpecsXAxis)], 'fontweight', 'bold')
ylabel([char(SpecsYAxis)], 'fontweight', 'bold')
zlabel('Cell Number', 'fontweight', 'bold')
set(gcf, 'Color', 'White')
set(gca, 'linewidth', 1.5, 'fontweight', 'bold', 'outerposition', [0.05 0 0.9 1])
colorbar ('fontweight', 'bold', 'position', [0.01 0.1 0.05 0.8])
savefig('BasoE')
print('BasoE', '-dpng')
```

```
figure
surf(X,Y,ZF)
title(['PolyE vs. ' char(SpecsTitle)], 'fontweight', 'bold', 'fontsize', 12)
xlabel([char(SpecsXAxis)], 'fontweight', 'bold')
ylabel([char(SpecsYAxis)], 'fontweight', 'bold')
zlabel('Cell Number', 'fontweight', 'bold')
set(gcf, 'Color', 'White')
set(gca, 'linewidth', 1.5, 'fontweight', 'bold', 'outerposition', [0.05 0 0.9 1])
colorbar ('fontweight', 'bold', 'position', [0.01 0.1 0.05 0.8])
savefig('PolyE')
print('PolyE', '-dpng')
```

```
figure
surf(X,Y,ZG)
title(['OrthoE vs. ' char(SpecsTitle)], 'fontweight', 'bold', 'fontsize', 12)
xlabel([char(SpecsXAxis)], 'fontweight', 'bold')
ylabel([char(SpecsYAxis)], 'fontweight', 'bold')
zlabel('Cell Number', 'fontweight', 'bold')
set(gcf, 'Color', 'White')
set(gca, 'linewidth', 1.5, 'fontweight', 'bold', 'outerposition', [0.05 0 0.9 1])
colorbar ('fontweight', 'bold', 'position', [0.01 0.1 0.05 0.8])
savefig('OrthoE')
print('OrthoE', '-dpng')
```

```
figure
surf(X,Y,ZH)
title(['MDS Comm. Cells vs. '
char(SpecsTitle)], 'fontweight', 'bold', 'fontsize', 12)
xlabel([char(SpecsXAxis)], 'fontweight', 'bold')
ylabel([char(SpecsYAxis)], 'fontweight', 'bold')
zlabel('MDS Cells', 'fontweight', 'bold')
set(gcf, 'Color', 'White')
set(gca, 'linewidth', 1.5, 'fontweight', 'bold', 'outerposition', [0.05 0 0.9 1])
colorbar ('fontweight', 'bold', 'position', [0.01 0.1 0.05 0.8])
savefig('MDS Committed Cells')
print('MDS Committed Cells', '-dpng')
```

```
figure
surf(X,Y,ZI)
title(['MDS RBC vs. ' char(SpecsTitle)], 'fontweight', 'bold', 'fontsize', 12)
xlabel([char(SpecsXAxis)], 'fontweight', 'bold')
ylabel([char(SpecsYAxis)], 'fontweight', 'bold')
zlabel('MDS Cells', 'fontweight', 'bold')
set(gcf, 'Color', 'White')
set(gca, 'linewidth', 1.5, 'fontweight', 'bold', 'outerposition', [0.05 0 0.9 1])
```

```
colorbar ('fontweight','bold','position',[0.01 0.1 0.05 0.8])
savefig('MDS Committed Cells')
print('MDS RBC', '-dpng')

figure
surf(X,Y,ZJ)
title(['Blast Percentage vs. '
char(SpecsTitle)], 'fontweight','bold','fontsize',12)
xlabel([char(SpecsXAxis)], 'fontweight','bold')
ylabel([char(SpecsYAxis)], 'fontweight','bold')
zlabel('Blast Percentage', 'fontweight','bold')
set(gcf, 'Color', 'White')
set(gca, 'linewidth', 1.5, 'fontweight','bold','outerposition',[0.05 0 0.9 1])
colorbar ('fontweight','bold','position',[0.01 0.1 0.05 0.8])
savefig('Blast Percentage')
print('Blast Percentage', '-dpng')

figure
surf(X,Y,ZK)
title(['Multipotent Prog. vs. '
char(SpecsTitle)], 'fontweight','bold','fontsize',12)
xlabel([char(SpecsXAxis)], 'fontweight','bold')
ylabel([char(SpecsYAxis)], 'fontweight','bold')
zlabel('Cell Number', 'fontweight','bold')
set(gcf, 'Color', 'White')
set(gca, 'linewidth', 1.5, 'fontweight','bold','outerposition',[0.05 0 0.9 1])
colorbar ('fontweight','bold','position',[0.01 0.1 0.05 0.8])
savefig('MPP')
print('MPP', '-dpng')

figure
surf(X,Y,ZL)
title(['Comm. Myeloid Prog. vs. '
char(SpecsTitle)], 'fontweight','bold','fontsize',12)
xlabel([char(SpecsXAxis)], 'fontweight','bold')
ylabel([char(SpecsYAxis)], 'fontweight','bold')
zlabel('Cell Number', 'fontweight','bold')
set(gcf, 'Color', 'White')
set(gca, 'linewidth', 1.5, 'fontweight','bold','outerposition',[0.05 0 0.9 1])
colorbar ('fontweight','bold','position',[0.01 0.1 0.05 0.8])
savefig('CMP')
print('CMP', '-dpng')

save('simulation')

%SurfPlotter

%BarPlotter

%StackPlotter

%test
```

Matlab code for animated, time-dependent 3D surface plot videos

```
%%SURFACE PLOTTER

VideoNames = {'RBC','Committed Cells','BFU-E','CFU-  
E','ProE','BasoE','PolyE','OrthoE','MDS Committed Cells','MDS RBC','Blast  
Percentage'};

%do Ctrl+F and find either 8 or 11.  
%define 8 for healthy simulations, 11 for diseased simulations

for bla = 1:8

tic

Counter = 0;

A = zeros(no_sims,no_function_calls);

A(:,1) = Input(:,1);  
A(:,2) = Input(:,2);

if bla < 10  
eval(['w = Output0' num2str(bla) ';''])  
else  
eval(['w = Output' num2str(bla) ';''])  
end

A(:,3:size(A,2)) = w(:,1:size(A,2)-2);

v = VideoWriter(['Video_Surf_' char(VideoNames(bla)) '.avi']);  
v.FrameRate = 27;  
v.Quality = 100;

open(v);

for a=1:(size(A,2)-2)  
AX=reshape(A(:,1),step+1,step+1);  
AY=reshape(A(:,2),step+1,step+1);  
AZ=reshape(A(:,a+2),step+1,step+1);  
fig(a) = figure('Visible','off');  
set(fig(a),'PaperPositionMode','auto');  
hold on  
set(gcf, 'Position', [1 1 1280 800]); % Maximize figure.  
surf(AX,AY,AZ)  
title([char(VideoNames(bla)) ' vs. ' char(SpecsTitle)]  
, 'fontweight','bold','fontsize',12)  
xlabel([char(SpecsXAxis)], 'fontweight','bold')
```

```
ylabel([char(SpecsYAxis)], 'fontweight', 'bold')
xlabel('Cell Number', 'fontweight', 'bold')
set(gcf, 'Color', 'White')
set(gca, 'linewidth', 1.5, 'fontweight', 'bold', 'outerposition', [0.05 0 0.9 1])
colorbar('fontweight', 'bold', 'position', [0.01 0.1 0.05 0.8])
grid on
zlim([0 max(max(A(:, 3:size(A, 2))))])
caxis([0 max(max(A(:, 3:size(A, 2))))])
view(3)
annotation(gcf, 'textbox', ...
    [0.8 0.867479674796748 0.14 0.032520325203252], ...
    'String', {'For Time = ' num2str(a*10)}, 'EdgeColor', [1 1 1]);
figinfo = hardcopy(fig(a), '-dzbuffer', '-r0');
writeVideo(v, im2frame(figinfo));
close all

time_passed = toc;

    Counter = Counter + 1;

        if rem(Counter, 50) == 0

            disp(['Iteration number ' num2str(Counter) ' finished after '
num2str(time_passed/60) ' mins for Video ' num2str(bla)])

        end
    end

close(v);
close all

end
```

Matlab code for stacked, animated, time-dependent 3D surface plot videos

```
tic
Counter = 0;

SA = zeros(no_sims,no_function_calls);
SB = zeros(no_sims,no_function_calls);
SC = zeros(no_sims,no_function_calls);

SA = Output12;
SB = Output13;
SC = Output03;

maxS = [max(max(SA)) max(max(SB)) max(max(SC))];

ssurfv = VideoWriter(['Video_SSsurf_MPP_CMP_BFU-E.avi']);
ssurfv.FrameRate = 14;
ssurfv.Quality = 100;

open(ssurfv);

for ssurf = 1:(no_function_calls)/2;

    SAr = reshape(SA(:,ssurf),step+1,step+1);
    SBr = reshape(SB(:,ssurf),step+1,step+1);
    SCr = reshape(SC(:,ssurf),step+1,step+1);

    fig(ssurf) = figure('Visible','off');
    set(fig(ssurf),'PaperPositionMode','auto');
    hold on
    set(gcf,'Position',[1 1 1280 800]);
    set(gcf,'Color','White')

    annotation(gcf,'textbox',...
        [0.826470588235292 0.94018527670378 0.14 0.0325203252032522],...
        'String',{'For Time = ' num2str(ssurf*10)},'EdgeColor',[1 1 1]);

    subplot(3,1,1);
    lines = surf(X,Y,SAr);
    annotation(gcf,'textbox',...
        [0.754117647058814 0.678492785642993 0.0870588235294163
        0.0325203252032523],...
        'String',{'[char(SpecsXAxis)]'},...
        'EdgeColor',[1 1 1],'fontweight','bold','fontsize',10);
    annotation(gcf,'textbox',...
        [0.170588235294113 0.680614358944542 0.0870588235294164
        0.0325203252032523],...
        'String',{'[char(SpecsYAxis)]'},...
        'EdgeColor',[1 1 1],'fontweight','bold','fontsize',10);
    zlabel('MPP','fontweight','bold','fontsize',10);
    title(['MPP vs. ' char(SpecsTitle)], 'fontweight','bold','fontsize',12);
    zlim([0 max(maxS)]);
```

```
subplot(3,1,2);
surf(X,Y,SBr);
annotation(gcf,'textbox',...
    [0.754117647058814 0.378492785642993 0.0870588235294163
    0.0325203252032523],...
    'String',{[char(SpecsXAxis)]},...
    'EdgeColor',[1 1 1],'fontweight','bold','fontsize',10);
annotation(gcf,'textbox',...
    [0.170588235294113 0.380614358944542 0.0870588235294164
    0.0325203252032523],...
    'String',{[char(SpecsYAxis)]},...
    'EdgeColor',[1 1 1],'fontweight','bold','fontsize',10);
xlabel('CMP','fontweight','bold','fontsize',10);
title(['CMP vs. ' char(SpecsTitle)], 'fontweight','bold','fontsize',12);
xlim([0 max(maxS)]);

subplot(3,1,3);
surf(X,Y,SCr);
annotation(gcf,'textbox',...
    [0.754117647058814 0.078492785642993 0.0870588235294163
    0.0325203252032523],...
    'String',{[char(SpecsXAxis)]},...
    'EdgeColor',[1 1 1],'fontweight','bold','fontsize',10);
annotation(gcf,'textbox',...
    [0.170588235294113 0.080614358944542 0.0870588235294164
    0.0325203252032523],...
    'String',{[char(SpecsYAxis)]},...
    'EdgeColor',[1 1 1],'fontweight','bold','fontsize',10);
xlabel('BFU-E','fontweight','bold','fontsize',10);
title(['BFU-E vs. ' char(SpecsTitle)], 'fontweight','bold','fontsize',12);
xlim([0 max(maxS)]);

figinfo = hardcopy(fig(ssurf),'-dzbuffer','-r0');
writeVideo(ssurfv,im2frame(figinfo));
close all

time_passed = toc;

    Counter = Counter + 1;

    if rem(Counter, 50) == 0

        disp(['Iteration number ' num2str(Counter) ' finshed after '
num2str(time_passed/60) ' mins for Video ' ])

    end
end

close(ssurfv);
close all
```

R Code used for Unsupervised Multivariate Statistical Analysis

```
rm(list=ls())
library(readr)
library(ggplot2)
library(reshape2)
library(readr)
library(tidyr)
library(readxl)
library(geneplotter)
library(data.table)
library(GGally)
library(tidyverse)
library(corrplot)
library(RColorBrewer)
library(gridExtra)
library(ggfortify)
library(gplots)
library(made4)
library(pheatmap)
library(heatmap.plus)
library(cowplot)
library(magick)
library(pca3d)
library(heatmap3)
library(Hmisc)
library(samr)
library(muma)

#Make sure this working directory has all the functions like
#multiplot, volcano, magicplot, etc

setwd('D:/Morais José/Experimental/R')

#Use Transposed data
#First column should be named "Condition"
#Metabolites as column titles and conditions as first column
#Lines - sample runs
#Columns - metabolites

#Paste path here for the transposed dataset file
path <- "D:/Google Drive/Imperial PC/Morais
José/Experimental/2018_01_GCMS_MDS/Samples/Test.csv"

#read_csv if document is created in English windows, due to separation being
made with ,
#read_csv2 if document is created in Portuguese Windows, due to separation
being made with ;

data <- as.data.frame(read_csv(path))
dimdata <- as.matrix(dim(data))
dimdata <- as.numeric(dimdata[2,])
#data_24 <- data[grep("1M", data$Condition),]
data[, -c(1)] <- as.data.frame(scale(data[, c(2:dimdata)]))
```

```
#data_24 <- data[grepl("24", data$Condition),]
#data_24[, -c(1)] <- as.data.frame(scale(data_24[, c(2:dimdata)]))

tdata <- data %>%
  gather(key, Value, -Condition) %>%
  separate(key, "Measurement", "\\.")
tdata$Condition <- as.factor(tdata$Condition)

##### Multivariate Analysis #####

##### PCA #####

pca <- prcomp(data[, -c(1)])
summary(pca)$importance[, 1:6]
cols <- character(nrow(data))
cols[] <- "black"

##1. Edit here with all the condition names

condnames =
as.matrix(c('05M_Unc', '05M_Col', '1M_Unc', '1M_Col', '2M_Unc', '2M_Col'))

##2. Add as many lines as the number of your conditions
#cols[data$Condition == "condname[1,]"] <- "blue"
#cols[data$Condition == "condname[2,]"] <- "red"
#cols[data$Condition == "condname[3,]"] <- "green"

scores <- pca$x[, 1:6]

km <- kmeans(scores, centers = 6, nstart=20)
ggdata <- data.frame(scores, Cluster=km$cluster, Species=data$Condition)

##3. Calculating each PC percentage (here we do for 3)
npcs <- as.numeric(3) #number of Principal Components
povstd <- as.matrix(pca$sdev) #Collect STD of each PC
pov <- as.matrix(c(1:npcs))

for (i in 1:npcs) {
  pov[i,] <- (povstd[i,]^2)/(sum(povstd^2))*100
}

if (i == npcs ) {
  print('PCA % Calculated Successfully')
} else {
  print ('Error Number of PCS do not match iterations')}
pov <- round(pov, digits=2)

##4. Create as many lines as needed
c1 <- toString(pov[1,])
c2 <- toString(pov[2,])
```

```
c3 <- toString(pov[3,])

##5. These are the actual axis labels
pc1 <- paste("PC1 ", c1, "%", sep=" ")
pc2 <- paste("PC2 ", c2, "%", sep=" ")
pc3 <- paste("PC3", c3, "%", sep=" ")

#scale_color_manual(values=c("deepskyblue2","deepskyblue2", "tomato",
"deepskyblue2","deepskyblue2", "tomato", "deepskyblue2","deepskyblue2",
"tomato", "deepskyblue2","deepskyblue2", "tomato",
"deepskyblue2","deepskyblue2", "tomato", "deepskyblue2","deepskyblue2",
"tomato"))

#scale_color_manual(values=c("deepskyblue2","green","deepskyblue2","magenta4",
"deepskyblue2","dodgerblue1","deepskyblue2","coral","deepskyblue2","tomato",
"deepskyblue2","goldenrod4"))

##6.Ploting PCA

p1<-ggplot(ggdata) +
  geom_point(aes(x=PC1, y=PC2, color=factor(data$Condition)), size=7, alpha =
0.6, shape=20) +
  guides(color=guide_legend("Cluster"),fill=guide_legend("Cluster"))+
  theme_gray()+
  xlab(pc1)+
  ylab(pc2)+
  scale_color_manual(values=c("red", "green", "red", "red", "red", "yellow",
"red", "blue", "red", "red", "red", "red", "red", "yellow", "red"))

#scale_color_manual(values=c("red","green","red","red","red","yellow","red","
red","red","red","blue","red","red","red","red","red","red","yellow","red"))

p1

#scale_color_manual(values=c("red","green","green","red","red","red","red","r
ed","yellow","pink","yellow","red","red","red","red","red","blue","blue","blu
e","red","red","red","red","red","red","red","red","red","red","yellow","red","red"
))

#scale_color_manual(values=c("red", "blue", "green", "magenta1", "red",
"blue", "green", "gold4", "red", "blue", "green", "darkorange3", "red",
"blue", "green", "coral4", "red", "blue", "green", "tan3", "red", "blue",
"green", "darkviolet"))

p2<-ggplot(ggdata) +
  geom_point(aes(x=PC2, y=PC3, color=factor(data$Condition)), size=7, alpha =
0.6, shape=20) +
  guides(color=guide_legend("Cluster"),fill=guide_legend("Cluster"))+
  theme_gray()+
  xlab(pc2)+
  ylab(pc3)+
```

```
scale_color_manual(values=c("red","green","green","red","red","red","red","red","red","yellow","pink","yellow","red","red","red","red","red","blue","blue","blue","red","red","red","red","red","red","red","red","red","red","red","yellow","red","red"))

#scale_color_manual(values=c("hotpink1", "hotpink2", "hotpink3", "hotpink4",
"magenta1", "magenta2", "magenta3", "magenta4", "red1", "red2", "red3",
"red4", "blue1", "blue2", "blue3", "blue4", "slateblue1", "slateblue2",
"slateblue3", "slateblue4", "aquamarine1", "aquamarine2", "aquamarine3",
"aquamarine4"))

p3<-ggplot(ggdata) +
  geom_point(aes(x=PC1, y=PC3, color=factor(data$Condition)), size=7, alpha =
0.6, shape=20) +
  guides(color=guide_legend("Cluster"),fill=guide_legend("Cluster"))+
  theme_gray()+
  xlab(pc1)+
  ylab(pc3)+

scale_color_manual(values=c("red","green","green","red","red","red","red","red","red","yellow","pink","yellow","red","red","red","red","red","blue","blue","blue","red","red","red","red","red","red","red","red","red","red","red","yellow","red","red"))

#scale_color_manual(values=c("hotpink1", "hotpink2", "hotpink3",
"hotpink4", "magenta1", "magenta2", "magenta3", "magenta4", "red1", "red2",
"red3", "red4", "blue1", "blue2", "blue3", "blue4", "slateblue1",
"slateblue2", "slateblue3", "slateblue4", "aquamarine1", "aquamarine2",
"aquamarine3", "aquamarine4"))

pdf(file.path('MDS_ALL.pdf'), height = 12, width = 12, paper = 'special')
source('multiplot_Jose.R')
multiplot(p1,p3,p2, cols = 1)
dev.off()

multiplot(p1)

#rm(c1,c2,c3,i,npcs,pov,povstd)

##### HCI #####

#data <- as.data.frame(read_csv(path))
row.names(data) <- make.unique(as.character(data$Condition), sep = "_")
cn <- data$Condition
trdata <- as.data.frame(t(data))
trdata <- trdata[-c(1),]
rn <- row.names(trdata)
trdata <- as.data.frame(apply(trdata, 2, as.numeric))
row.names(trdata) <- rn
pdf(file.path('PB_HCI.pdf'), height = 12, width = 24, paper = 'special')
heatplot(trdata, dend=c("column"), distfun = "manhattan", scale = "none",
classvec = cn)
```

```
dev.off()

cn <- values("MDS")
#"midnightblue","pink"
#"darkgoldenrod","darkorchid4","lightcoral"

##### Correlation Matrix #####

#Cor function works with data frames and automatically
#segments data based on the Condition argument.

#1. Edit here your condition names
condnames = as.matrix(c('D0','D8_Unc','D8_Col'))

#2. Create as many lines as conditions
data1 <- subset(data, Condition == condnames[1,])
data2 <- subset(data, Condition == condnames[2,])
data3 <- subset(data, Condition == condnames[3,])

#3. Concatenate the specific combinations that you want to compare.
#Pay attention t
data12 <- subset(data, Condition == condnames[1,] | Condition ==
condnames[2,])
data23 <- subset(data, Condition == condnames[2,] | Condition ==
condnames[3,])
data13 <- subset(data, Condition == condnames[1,] | Condition ==
condnames[3,])

#4. Find critical correlation coefficient.
#Note - This is sample-size, n, dependant. It will give you
#the correlation coefficient at which the observations are
#statistically different at an alpha (p-value) of "x".
#If the observed correlation coefficient is smaller than the critical
#then this is not significantly different from random.

critical.r <- function( n, alpha = .05 ) {
  df <- n - 2
  critical.t <- qt(alpha/2, df, lower.tail = F)
  critical.r <- sqrt( (critical.t^2) / ( (critical.t^2) + df ) )
  return(critical.r)
}

criticalR <- critical.r( dimdata )

cormat <- cor(data2,data3)

for (i in 1:(dimdata-1)) {
  for (j in 1:(dimdata-1)) {
    if (cormat[i,j] <= criticalR & cormat[i,j] >= -criticalR) {
      cormat[i,j] = NA
    }
  }
}
```

```
heatmap(cormat)

##### Volcano Plots #####

#1. Edit here your condition names
condnames = as.matrix(c('D0', 'D8_Unc', 'D8_Col'))

#2. Volcano plots compare means -> Apply the mean for each dataset
means <- aggregate(data[,2:dimdata], list(data$Condition), mean)

#3. Define what to compare

FC1 <- means[means$Group.1 == "D8_Unc",]
FC2 <- means[means$Group.1 == "D8_Col",]

FC1Aux <- as.matrix(FC1[, -c(1)])
FC2Aux <- as.matrix(FC2[, -c(1)])

#4. FC of features

##### d24 only #####
##### PCA #####

pca <- prcomp(data_24[, -c(1)])
summary(pca)$importance[, 1:6]
cols <- character(nrow(data_24))
cols[] <- "black"

##1. Edit here with all the condition names

condnames =
as.matrix(c('05M_Unc', '05M_Col', '1M_Unc', '1M_Col', '2M_Unc', '2M_Col'))

##2. Add as many lines as the number of your conditions
#cols[data_24$Condition == "condname[1,]"] <- "blue"
#cols[data_24$Condition == "condname[2,]"] <- "red"
#cols[data_24$Condition == "condname[3,]"] <- "green"

scores <- pca$x[, 1:6]

km <- kmeans(scores, centers = 6, nstart=20)
ggdata_24 <- data.frame(scores, Cluster=km$cluster,
Species=data_24$Condition)

##3. Calculating each PC percentage (here we do for 3)
npcs <- as.numeric(3) #number of Principal Components
povstd <- as.matrix(pca$sdev) #Collect STD of each PC
pov <- as.matrix(c(1:npcs))
```

```
for (i in 1:npcs) {
  pov[i,] <- (povstd[i,]^2)/(sum(povstd^2))*100
}

if (i == npcs ) {
  print('PCA % Calculated Successfully')
} else {
  print ('Error Number of PCS do not match iterations')}}
pov <- round(pov, digits=2)

##4. Create as many lines as needed
c1 <- toString(pov[1,])
c2 <- toString(pov[2,])
c3 <- toString(pov[3,])

##5. These are the actual axis labels
pc1 <- paste("PC1 ", c1, "%", sep=" ")
pc2 <- paste("PC2 ", c2, "%", sep=" ")
pc3 <- paste("PC3", c3, "%", sep=" ")

#scale_color_manual(values=c("deepskyblue2","deepskyblue2", "tomato",
"deepskyblue2","deepskyblue2", "tomato", "deepskyblue2","deepskyblue2",
"tomato", "deepskyblue2","deepskyblue2", "tomato",
"deepskyblue2","deepskyblue2", "tomato", "deepskyblue2","deepskyblue2",
"tomato"))

#scale_color_manual(values=c("deepskyblue2","green","deepskyblue2","magenta4",
"deepskyblue2","dodgerblue1","deepskyblue2","coral","deepskyblue2","tomato",
"deepskyblue2","goldenrod4"))

##6. Plotting PCA

p1<-ggplot(ggdata_24) +
  geom_point(aes(x=PC1, y=PC2, color=factor(data_24$Condition)), size=7,
alpha = 0.6, shape=20) +
  guides(color=guide_legend("Cluster"),fill=guide_legend("Cluster"))+
  theme_gray()+
  xlab(pc1)+
  ylab(pc2)+
  scale_color_manual(values=c("red", "blue", "green", "magenta1", "red",
"blue", "green", "gold4", "red", "blue", "green", "darkorange3", "red",
"blue", "green", "coral4", "red", "blue", "green", "tan3", "red", "blue",
"green", "darkviolet"))

p1
p2<-ggplot(ggdata_24) +
  geom_point(aes(x=PC2, y=PC3, color=factor(data_24$Condition)), size=7,
alpha = 0.6, shape=20) +
  guides(color=guide_legend("Cluster"),fill=guide_legend("Cluster"))+
  theme_gray()+
  xlab(pc2)+
  ylab(pc3)+
  scale_color_manual(values=c("hotpink1", "hotpink2", "hotpink3", "hotpink4",
"magenta1", "magenta2", "magenta3", "magenta4", "red1", "red2", "red3",
```

```
"red4", "blue1", "blue2", "blue3", "blue4", "slateblue1", "slateblue2",
"slateblue3", "slateblue4", "aquamarine1", "aquamarine2", "aquamarine3",
"aquamarine4"))

p3<-ggplot(ggdata_24) +
  geom_point(aes(x=PC1, y=PC3, color=factor(data_24$Condition)), size=7,
alpha = 0.6, shape=20) +
  guides(color=guide_legend("Cluster"),fill=guide_legend("Cluster"))+
  theme_gray()+
  xlab(pc1)+
  ylab(pc3)+
  scale_color_manual(values=c("hotpink1", "hotpink2", "hotpink3", "hotpink4",
"magenta1", "magenta2", "magenta3", "magenta4", "red1", "red2", "red3",
"red4", "blue1", "blue2", "blue3", "blue4", "slateblue1", "slateblue2",
"slateblue3", "slateblue4", "aquamarine1", "aquamarine2", "aquamarine3",
"aquamarine4"))

pdf(file.path('Full.pdf'), height = 12, width = 12, paper = 'special')
source('multiplot_Jose.R')
multiplot(p1,p3,p2, cols = 1)
dev.off()

#rm(c1,c2,c3,i,npcs,pov,povstd)

##### HCI #####

data_24 <- as.data.frame(read_csv(path))
row.names(data_24) <- make.unique(as.character(data_24$Condition), sep = "_")
cn <- data_24$Condition
trdata_24 <- as.data.frame(t(data_24))
trdata_24 <- trdata_24[-c(1),]
rn <- row.names(trdata_24)
trdata_24 <- as.data.frame(apply(trdata_24, 2, as.numeric))
row.names(trdata_24) <- rn
pdf(file.path('HCI_D24.pdf'), height = 12, width = 24, paper = 'special')
heatplot(trdata_24, dend=c("column"), distfun = "manhattan", scale = "none")
dev.off()
```