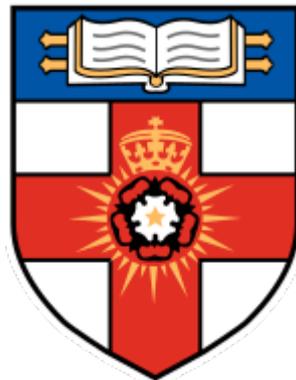


**Computational science-enabled  
radiological pathology for the non-invasive  
mapping of tumour heterogeneity  
in childhood neuroblastoma**

Konstantinos Zormpas-Petridis

A thesis submitted for the degree of  
Doctor of Philosophy  
of the  
University of London



September 2020  
Institute of Cancer Research, London  
University of London

## Declaration

I hereby declare that this thesis reports on my own original work. Any contribution made to this research by others whom I have worked with is explicitly acknowledged in the thesis.

Konstantinos Zormpas-Petridis

## Abstract

Neuroblastoma is a common childhood solid tumour that accounts for 15% of all cancer paediatric deaths. This thesis addresses key deficiencies in our ability to define, monitor and predict neuroblastoma heterogeneity for precision medicine. I used computational science to integrate the spatially-encoded phenotypic information provided by multi-parametric magnetic resonance imaging (MRI) with digital histopathology, demonstrating that MRI can provide non-invasive pathology to characterise neuroblastoma heterogeneity and provide biomarkers of response in clinically-relevant transgenic mouse models of high-risk disease. I first developed and demonstrated the application of novel computational pathology methodologies to enhance the quantitative assessment of tumour components from H&E-stained whole-slide images (WSI). These included two frameworks: *SuperCRF*, which fuses traditional machine learning with deep learning to model the way pathologists incorporate large-scale tissue architecture and context across spatial scales to significantly improve single-cell classification and, *SuperHistopath*, which combines the application of the SLIC superpixels algorithm on low-magnification WSIs (5x) with a convolutional neural network (CNN) for superpixels classification to accurately map tumour heterogeneity from low-resolution histology. I then developed an MRI-histopathology cross-validation pipeline which provides the rigorous validation needed to support the deployment of novel MRI scans in the neuroblastoma clinic. Using this platform, I demonstrated the sensitivity of susceptibility-, T<sub>1</sub>-Mapping- and diffusion-weighted- MRI to the cellular and microenvironmental hallmarks of high-risk neuroblastoma and their modulation by either vascular- or MYCN- targeted therapies. Finally, I used supervised machine learning classification- and regression-based approaches to show proof-of-concept that habitat imaging derived from these three scans can non-invasively provide quantitative data typically acquired from histological analysis, such as densities of specific cell populations. This thesis demonstrates the potential of multi-parametric MRI to deliver non-invasive “virtual” biopsies to enhance diagnostic and treatment monitoring for children with neuroblastoma and pave new ways in studying tumour as an evolving ecosystem.



## Acknowledgments

First and foremost, I would like to thank my supervisor Dr. Yann Jamin for the endless support, guidance, tutoring (and patience) that made this journey so extraordinary. For believing in me and for giving me this unique opportunity. For making me an independent scientist, allowing me to explore my own ideas (no matter how crazy), but at the same time always being available and giving me the right guidance in every step of the way. Thank you for showing me what true science is and for making it so fun. Your jokes could make me burst-out laughing even at the grimmest day! More than a supervisor, you have been a great mentor and friend. Words cannot really describe how much I appreciate everything you've done for me. I look forward to keep working with you and do "the cool stuff". Thank you! (and no, any Greek people that will be mentioned here are not my cousins!).

I would also like to thank my co-supervisor Dr. Yinyin Yuan for taking me into her lab, for her endless source of knowledge, motivation, enthusiasm and support, for always encouraging me to become a better scientist and of course for the best lab retreats! I am very proud of the work we did and I hope we can do even more in the future! I will forever be grateful to Dr. Matthew Blackledge for treating me like his own student from the beginning, for teaching me tons of exciting stuff (including how to scan a phantom with breast implants) and for always being there for me. Thank you for being a great friend. Collaborating with you has been an amazing experience, and now it will be a real honour to work for you. I look forward to even more amazing things! (When do we start teaching A.I. to Adam?). I would like to thank Dr. Simon Robinson for taking me as his own student, for his supportive and kind presence, for always being available and giving me a new point of view and for all his advice and on-point remarks (and *Blackadder* references). I would also like to thank my back-up supervisor Prof. Louis Chesler for his enthusiasm and support.

I am truly grateful to Rosetrees Trust and their generous donors for their financial support, encouragement and dedication that made this project possible.

I would like to thank all the members, past and present, of the Robinson and Yuan labs for all the fun and meaningful experiences, especially Jin, Elise, Emma, Jess, Mariama, and Khalid, Henrik, Andreas, Sidra, Violet, Yeman, Shan, Priya.

During this time, I have been extremely fortunate to have collaborated with many talented and incredible people. Thus, a huge thank you for all their help to: Dr. Henrik Failmezger for the machine learning projects we undertook together (and for the late nights having dinner at *Sofra* and then going back to editing the paper in time for MICCAI, as well as the Friday boxing gym sessions); Dr. Jin Li for all the work we did in the MRE paper, it has been a pleasure working with you; Dr. Ioannis Roxanis for his invaluable expertise, his endless knowledge in pathology, for his fresh ideas leading to awesome projects, for his genuine enthusiasm and passion for science and his tireless efforts to incorporate A.I. in the clinic. Working with you has been truly inspiring and I look forward to more exciting projects and collaborations; Dr. Khalid Abdul-Jabbar and Dr. Shan Raza for promptly and patiently helping me and answering any issue I had with the training of their fancy deep learning architectures; Dr. Matthew Clarke for always finding the time to explain any pathology-related question I had, for his genuine enthusiasm and for teaching me so much during these short meetings; Dr. Andreas Heindl, Dr. Sidra Nawaz and Dr. Violet Kovacheva for showing me the ropes in the beginning, as well as all the tricks around CRImage, I learnt a great deal from you; Prof. Dow-Mu Koh and Dr. Nina Tunariu for the quickDWI project, you have both been an inspiration to me; Prof. Rosa Noguera for welcoming us to her lab in Valencia; Dr. Debbie Hill for working with me on my first project in the ICR.

Cliché as it may sound, the foundation is the most important part of any construction. None of this would be possible without the tutoring and guidance of a few amazing individuals during my undergraduate and master's studies. I would like to send the most sincere thanks and appreciation to Prof. Evangelos Zygouris for all his patience and valuable teachings during our 4-year collaboration which have shaped me both as a scientist and as a person, and to Prof. Dimitris Bakalis and Dr. Athanasios Kalantzopoulos from the University of Patras.

A very special thanks to all the people involved in my parallel journey during the PhD, which is Kung-Fu. My instructors Dai-Sifu Keith, Sifu Rena, Dan, Laura, James and everyone who was training with me, thank you for the new world you opened for me and for helping me understand and learn so many things about myself.

A big thank you to my dear friends and flatmates of Kingswood Drive for taking this road with me and for all the great times.

To my mother and father, thank you for your unconditional love, your kindness and for always supporting me. Thank you for the values you taught me, for making me a better person, and for answering the million “Why this and why that?” I always had. Again words don’t really suffice. It’s all thanks to you and I will forever be grateful. Dad, you have been in my mind every single day during this journey and I will always carry you with me in every step of the way.

Last, but definitely not least (like the “last author”) I would like to thank all my friends (you know who you are) for being patient, standing alongside me through these years, making me laugh and generally “being there” for me. Thank you Roza for every single thing we’ve been through all this time.

# Table of Contents

<b>Declaration</b> .....	<b>2</b>
<b>Abstract</b> .....	<b>3</b>
<b>Acknowledgments</b> .....	<b>5</b>
<b>Table of Contents</b> .....	<b>8</b>
<b>List of Figures</b> .....	<b>13</b>
<b>List of Tables</b> .....	<b>24</b>
<b>List of abbreviations</b> .....	<b>27</b>
<b>Chapter 1 : Introduction</b> .....	<b>31</b>
<b>1.1 Understanding tumour heterogeneous evolving ecosystem to deliver precision medicine.</b> .....	<b>31</b>
1.1.1 Tumour heterogeneity and evolution are major obstacles to the delivery of precision medicine. ....	31
1.1.2 The tumour as an ecosystem .....	32
1.1.3 Understanding and quantifying tumour ecology .....	33
1.1.3.1 Computational pathology-enabled spatial analysis .....	33
1.1.3.2 Limitations of the approach: a role for functional imaging? .....	34
<b>1.2 Functional imaging biomarkers, radiomics and habitat imaging for the delivery of precision medicine.</b> .....	<b>35</b>
1.2.1 MRI imaging biomarkers to characterise the tumour microenvironment.....	35
1.2.2 Radiomics.....	38
1.2.3 Habitat imaging.....	40
<b>1.3 Functional imaging-guided delivery of precision medicine for children with neuroblastoma</b> .....	<b>43</b>
1.3.1 Neuroblastoma .....	43
1.3.1.1 Histological characteristics of peripheral neuroblastic tumours .....	47
1.3.2 Targeted therapies for high-risk refractory and relapse neuroblastoma .....	50
1.3.2.1 Targeting MYCN.....	50
1.3.2.2 The BEACON Neuroblastoma trial .....	50
1.3.2.3 The mouse hospital and co-clinical trial concept.....	51
1.3.2.3.1 Genetically engineered murine (GEM) models of neuroblastoma: an information-rich experimental platform.....	51
1.3.3 Biomarkers-guided delivery of precision medicine .....	54
1.3.3.1 Need for non-invasive biomarkers.....	54
1.3.3.2 Current radiological management of children with neuroblastoma.....	54
1.3.3.3 MRI-based functional imaging for neuroblastoma .....	55
1.3.3.4 Validation of imaging biomarkers.....	55
<b>1.4 The mouse hospital as a robust platform for the validation of imaging biomarkers of neuroblastoma</b> .....	<b>56</b>
<b>1.5 Aim of the project</b> .....	<b>56</b>
<b>1.6 Structure of the thesis</b> .....	<b>58</b>
<b>1.7 Contributions to the work presented in this thesis.</b> .....	<b>59</b>

<b>Chapter 2 : Development of computational pathology methodologies in whole-slide digital histology images.....</b>	<b>60</b>
<b>2.1 The new era of digital pathology.....</b>	<b>60</b>
2.1.1 Computer-aided diagnosis.....	60
2.1.2 Machine learning.....	61
2.1.2.1 Machine learning methods.....	62
2.1.2.2 Machine learning approaches .....	62
2.1.2.3 Deep learning .....	63
2.1.3 Aim of the chapter .....	68
<b>2.2 Cell classification using morphological features .....</b>	<b>69</b>
2.2.1 Introduction .....	69
2.2.1.1 Aim of the section.....	69
2.2.2 Materials and Methods .....	69
2.2.2.1 Dataset.....	69
2.2.2.2 Cell segmentation and classification.....	70
2.2.2.3 Evaluation metrics for classification.....	71
2.2.3 Results .....	71
2.2.4 Discussion .....	73
2.2.5 Conclusion .....	74
<b>2.3 Capturing global spatial context to improve cell classification.....</b>	<b>75</b>
2.3.1 Introduction .....	75
2.3.1.1 Aim of the section.....	75
2.3.2 Methodology and Results.....	77
2.3.2.1 Data.....	77
2.3.2.2 Superpixel-based tumour region classification.....	77
2.3.2.3 Cell Classification based on nuclear features .....	81
2.3.2.4 Incorporating global context in cell classification .....	82
2.3.3 Discussion and Conclusion.....	83
<b>2.4 SuperCRF: An enhanced deep learning approach for incorporating global and local context in whole-slide image analysis.....</b>	<b>85</b>
2.4.1 Introduction .....	85
2.4.1.1 Aim of the section.....	85
2.4.2 Materials and Methods .....	85
2.4.2.1 Datasets.....	85
2.4.2.2 Single-Cell Classification Using a Spatially Constrained Convolutional Neural Network.....	86
2.4.2.3 Superpixel-Based Tumour Region Classification.....	86
2.4.2.4 SuperCRF .....	87
2.4.2.5 Survival analysis.....	91
2.4.3 Results .....	91
2.4.3.1 SuperCRF improves accuracy of cell classification.....	91
2.4.3.2 SuperCRF's increased accuracy of cell classification improves confidence in stromal cell ratio as a predictive feature of survival in melanoma .....	97
2.4.3.3 Combining cell and region classification: Location of the immune infiltrate is predictive of survival in melanoma.....	97
2.4.4 Discussion .....	99
2.4.5 Conclusion .....	102
<b>2.5 SuperHistopath: A Deep Learning Pipeline for Mapping Tumour Heterogeneity on Low-Resolution Whole-Slide Digital Histopathology Images .....</b>	<b>103</b>
2.5.1 Introduction .....	103
2.5.1.1 Aim of this section .....	104
2.5.2 Materials and Methods .....	105
2.5.2.1 Ethical statement.....	105
2.5.2.2 Datasets.....	105

2.5.2.3 Region classification.....	106
2.5.2.4 Training of the convolutional neural networks.....	110
2.5.2.5 Application of SuperHistopath for the quantification of clinical features of interest ...	112
2.5.2.5.1 Melanoma.....	112
2.5.2.5.2 Neuroblastoma.....	112
2.5.3 Results .....	113
2.5.3.1 SuperHistopath can accurately map the complex histological heterogeneity of tumours .....	113
2.5.3.1.1 Melanoma.....	113
2.5.3.1.2 Breast cancer.....	117
2.5.3.1.3 Neuroblastoma.....	119
2.5.3.2 SuperHistopath pipeline for the analysis of low-resolution WSI affords significant speed advantages.....	122
2.5.3.3 SuperHistopath can provide robust quantification of clinically relevant features.....	122
2.5.3.3.1 Stroma-to-tumour ratio and clusters of lymphocytes abundance as predictive markers of survival in melanoma.....	122
2.5.3.3.2 Necrosis quantification.....	123
2.5.3.3.3 Quantification of neuroblastoma differentiation.....	124
2.5.4 Discussion .....	126
2.5.5 Conclusion .....	129
<b>Chapter 3 : Development of a MRI-histopathology cross-validation pipeline.....</b>	<b>130</b>
<b>3.1 Introduction .....</b>	<b>130</b>
3.1.1 Aim of the chapter .....	130
<b>3.2 MRI-histology alignment.....</b>	<b>131</b>
3.2.1 Materials and Methods .....	131
3.2.1.1 Animals.....	131
3.2.1.2 T <sub>2</sub> -weighed MRI.....	131
3.2.1.3 T <sub>2</sub> -weighted MRI guided tumour excision and cutting.....	131
3.2.1.4 Sample processing and slide digitization.....	132
3.2.1.5 Supervised visual correspondence and identification of slices of interest.....	132
3.2.2 Results .....	133
3.2.3 Challenges for MRI-histology comparison .....	134
<b>3.3 Matching the resolution of histological parametric maps to the MRI resolution .....</b>	<b>135</b>
<b>3.4 MRI-histology registration.....</b>	<b>136</b>
<b>3.5 Discussion and conclusion.....</b>	<b>138</b>
<b>Chapter 4 : Validation of R<sub>2</sub>* and fractional Blood Volume (fBV) biomarkers for the imaging of the hemodynamic vasculature of neuroblastoma .....</b>	<b>140</b>
<b>4.1 Introduction .....</b>	<b>140</b>
4.1.1 Vascular-targeted therapies for high risk neuroblastoma .....	140
4.1.2 Non-invasive functional imaging of tumour vasculature.....	140
4.1.3 Aim of this chapter .....	141
<b>4.2 Materials and Methods.....</b>	<b>142</b>
4.2.1 Animals.....	142
4.2.2 In vivo study design .....	142
4.2.3 MRI .....	143
4.2.4 Computational pathology/digital histology .....	143
4.2.5 Clinical investigations .....	145
<b>4.3 Results.....</b>	<b>146</b>
4.3.1 Baseline tumour fBV predicts response to cediranib in the Th-MYCN GEM model of neuroblastoma .....	146

4.3.2 Baseline tumour $R_2^*$ predicts response to cediranib in the Th-MYCN GEM model of neuroblastoma .....	148
4.3.3 Spatial heterogeneity in fBV reflects regional variations in the microvasculature of neuroblastoma .....	150
4.3.4 Regional heterogeneity in $R_2^*$ reflects variations in RBC distribution in neuroblastoma.....	154
4.3.5 Innate resistance to cediranib is associated with a differentiating phenotype .....	156
4.3.6 MYCN-amplified childhood neuroblastomas are haemorrhagic.....	160
4.3.7 Initial experience with IS-MRI in children with neuroblastoma .....	161
<b>4.4 Discussion.....</b>	<b>163</b>
<b>4.5 Conclusion .....</b>	<b>167</b>
<b><i>Chapter 5 : Exploring the histological correlates of native <math>T_1</math> mapping in neuroblastoma pathology and its response to MYCN-targeted therapy in vivo .....</i></b>	<b><i>168</i></b>
<b>5.1 Introduction .....</b>	<b>168</b>
5.1.1 MYCN-targeted therapies for high-risk MYCN-driven neuroblastoma .....	168
5.1.2 Reduction in the native spin lattice relaxation time $T_1$ as a generic biomarker of response to therapy.....	168
5.1.3 Aim of this chapter .....	169
<b>5.2 Materials and Methods.....</b>	<b>170</b>
5.2.1 Animals, imaging and drug treatment schedule .....	170
5.2.2 MRI .....	170
5.2.3 Computational pathology/digital histology .....	171
5.2.4 Statistical Analysis .....	172
<b>5.3 Results.....</b>	<b>174</b>
5.3.1 Alisertib and vistusertib elicit significant anti-tumour activity associated with a decrease in native $T_1$ .....	174
5.3.2 Low native tumour $T_1$ correlates with high tumour red blood cell content.....	178
5.3.3 High native tumour $T_1$ correlates with high density of undifferentiated neuroblasts and low density of apoptotic neuroblasts.....	179
5.3.4 Regions rich in differentiating neuroblasts are associated with lower $T_1$ values. ....	185
<b>5.4 Discussion.....</b>	<b>186</b>
<b>5.5 Conclusion .....</b>	<b>192</b>
<b><i>Chapter 6 : Evaluating the sensitivity of ADC to childhood neuroblastoma pathology using gaussian mixture models and computational pathology in vivo.....</i></b>	<b><i>193</i></b>
<b>6.1 Introduction .....</b>	<b>193</b>
6.1.1 Aim of this chapter .....	193
<b>6.2 Materials and methods .....</b>	<b>194</b>
6.2.1 Gaussian mixture models for tumour clustering:.....	194
<b>6.3 Results.....</b>	<b>195</b>
6.3.1 Median tumour ADC and ADC compartments are not sensitive to tissue response following treatment with vistusertib. ....	198
6.3.2 Low native tumour ADC correlates with high red blood cell content.....	200
6.3.3 Increasing cell death is associated with a gradual reduction in ADC.....	204
6.3.4 ADC is sensitive to areas of therapy induced maturation and areas rich in differentiating neuroblasts.....	209
<b>6.4 Discussion.....</b>	<b>211</b>
<b>6.5 Conclusion .....</b>	<b>215</b>

<b>Chapter 7 : Habitat imaging of childhood neuroblastoma using multi-parametric MRI and co-registered histopathology .....</b>	<b>216</b>
<b>7.1 Aim of this chapter .....</b>	<b>216</b>
<b>7.2 Materials and Methods .....</b>	<b>217</b>
7.2.1 Classifying tumour regions based on restrictions.....	217
7.2.2 Habitat characterisation based on supervised machine learning .....	218
7.2.2.1 Spatial registration of MRI and SuperHistopath region classifier .....	218
7.2.2.2 Training dataset .....	219
7.2.2.3 Features .....	219
7.2.2.4 Machine learning algorithms.....	219
7.2.3 Prediction of key histological features using a regression approach.....	220
<b>7.3 Results .....</b>	<b>222</b>
7.3.1 Empirical thresholds can confidently characterise a large proportion of the tumour.....	222
7.3.2 Classification of neuroblastoma tumour habitats from multi-parametric MRI.....	224
7.3.3 Multi-parametric MRI can map the density of different cell populations .....	228
<b>7.4 Discussion.....</b>	<b>233</b>
<b>7.5 Conclusion .....</b>	<b>235</b>
<b>Chapter 8 : Conclusions .....</b>	<b>236</b>
<b>8.1 Summary of results .....</b>	<b>236</b>
8.1.1 Novel computational approach for enhanced quantitative histopathology.....	236
8.1.1.1 Superpixel-based Conditional Random Fields (SuperCRF): Incorporating global and local context for enhanced deep learning in melanoma histopathology .....	236
8.1.1.2 SuperHistopath: A Deep Learning Pipeline for Mapping Tumour Heterogeneity on Low-Resolution Whole-Slide Digital Histopathology Images.....	236
8.1.2 A robust MRI-histopathology cross-validation pipeline.....	237
8.1.3 Imaging the Hemodynamic Vasculature of Neuroblastoma with Susceptibility MRI for the non-invasive Prediction of Response to Anti-angiogenic Treatment .....	237
8.1.4 Non-invasive MRI native T <sub>1</sub> -mapping detects response to MYCN-targeted therapies in the Th-MYCN model of neuroblastoma .....	238
8.1.5 ADC is sensitive to the microstructure of poorly- or un-differentiated neuroblastoma but is a not an early biomarker of response to apoptosis-inducing MYCN-targeted therapies.....	238
8.1.6 Habitat imaging: Towards non-invasive pathology for children with neuroblastoma.....	239
<b>8.2 Future Directions.....</b>	<b>240</b>
8.2.1 A mouse hospital and co-clinical approach.....	240
8.2.1.1 <b>MAGNIFY</b> : comprehensive <b>M</b> apping of high-risk <b>N</b> euroblastoma heterogeneity to predict and target <b>r</b> eFractor <b>Y</b> disease .....	241
8.2.2 Incorporating MRI biomarkers of the tumour microenvironment.....	241
8.2.3 Refining the MRI-Histopathology pipeline.....	242
8.2.3.1 A larger and more diverse dataset.....	242
8.2.3.2 Improving registration.....	242
8.2.3.3 Improving cell classification for the enhanced characterisation of the tumour ecosystem .....	242
<b>Publications arising from the work presented in this thesis and collaborations .....</b>	<b>266</b>
<b>Appendix.....</b>	<b>272</b>

## List of Figures

**Figure 1.1** Tumour heterogeneity and its role in clinical progression. **A.** Sequential accumulation of somatic events by the malignant clone, drives disease progression and metastasis. **B.** The presence of genetically heterogeneous sub-clones with inherent differences in drug sensitivity can also affect response to treatment and cause clinical resistance. Further evolution of the dominant clone in response to chemotherapy, may lead to relapse after initial treatment response. **C.** The tumour microenvironment, apart from harbouring carcinoma cells, consists of various components that have a major role in influencing the outcome of the malignancy [Adapted from Kleppe et al (4) and Pattabiraman et al (5)]......32

**Figure 1.2** Non-invasive MR biomarkers **A. Tumour Perfused OxyR: a spatial biomarker of hypoxia.** Representative maps showing oxygen-enhancing (OxyE) and non-enhancing (OxyR) voxel and perfused and non-perfused voxels following oxygen breathing (OE-T<sub>1</sub>-weighted MRI and Dynamic contrast enhanced MRI respectively). Combined maps show that perfused OxyR spatially correlate with pimonidazole adduct formation, the gold-standard for the immunohistopathological detection of hypoxia. **B. Tumour apparent diffusion coefficient (ADC): a biomarker of disease progression in the TRAMP GEM model of prostate cancer.** Parametric maps of ADC allow, in the prostate of the TRAMP GEM model, the detection of malignant tissue from well-differentiated, due to increase cellularity, which restricts water diffusion. More aggressive undifferentiated disease, shows a marked lower ADC compared to well-differentiated disease. **C. Native tumour spin-lattice relaxation time T<sub>1</sub>: A biomarker of response to therapy in GEM model of Neuroblastoma.** Treatment of the Th-MYCN model of neuroblastoma with cyclophosphamide, leads to a reduction in native T<sub>1</sub>. **D. The absolute value of the complex shear modulus |G\*| measured by MR elastography: a biomarker of tissue integrity.** Parametric maps show a decrease in |G\*|, 24h following treatment, of a SW620 colon carcinoma tumour xenograft with the vascular-disrupting agent ZD6126, which caused massive central necrosis. **E. Identification of tumours harbouring the ALK<sup>F1174L</sup> mutation in MYCN-driven transgenic mice with Intrinsic Susceptibility (IS)–MRI.** Native R<sub>2</sub>\* maps and ΔR<sub>2</sub>\* oxygen-air maps obtained following continuous inhalation of 100% oxygen, were able to discriminate between tumours arising in the Th-ALK<sup>F1174L</sup>/Th-MYCN mouse from tumours arising in the Th-MYCN mouse, due to a differential vascular phenotype characterised by the absence or presence of haemorrhage respectively. [Figure adapted from **A.** O'Connor et al (39), **B.** Hill et al (40), **C.** Jamin et al (41), **D.** Li et al (42) **E.** Jamin et al (43)]. .....37

**Figure 1.3** A radiomics process pipeline **A.** Acquisition of multi-parametric MR images from standard exams, in this case prostate: T<sub>2</sub>-weighted MRI, diffusion weighted imaging (DWI) and the calculated apparent diffusion coefficient (ADC) maps, dynamic contrast enhanced (DCE)–MRI; **B.** manual or preferably semi-automatic/automatic segmentation of regions of volumes of interest. For the prostate, these volumes are: prostate, peripheral zone (PZ) [and subsequently transition zone (TZ)], urethra, normal appearing tissues in PZ and TZ and tumour regions of interest (ROI); **C.** Extraction and quantification of imaging features, such as intensity volume histogram (first order features), features related to volume/shape, texture features (second order features) and transform analysis features; **D.** radiomic data could then be integrated with clinical, genomic, proteomic and metabolomics data; **E.** the final integrated dataset is mined to develop diagnostic, predictive and prognostic models [Adapted from Stoyanova et al (49)]. .....39

**Figure 1.4** Workflow for radiomics analysis with “classical” machine learning (left, “hand-crafted features”) and deep learning approaches (right) [Adapted from Avanzo et al (54)].  
.....40

**Figure 1.5** Primary distribution of neuroblastoma in children, a tumour of the sympathetic nervous system. <https://www.cancer.net/cancer-types/neuroblastoma-childhood/medical-illustrations>.  
.....44

**Figure 1.6 Childhood neuroblastoma** **A.** picture illustrates a child <1 year old with stage IVs neuroblastoma with extensive metastatic disease, which spontaneously regressed over a period of 6 months without any therapeutic intervention. **B.** <sup>123</sup>I-meta-iodobenzylguanidine [123I-MIBG] scan showing extensive bone and bone marrow metastasis. **C.** Whole-body MRI showing a large primary abdominal tumour distorting the abdominal aorta and a secondary tumour in the upper thorax (73). **D.** Children with neuroblastoma presenting with typical orbital metastasis **E.** Overall survival rates of children with high-risk neuroblastoma from the time of diagnosis (74).  
.....45

**Figure 1.7** Multi-modal imaging management of children with neuroblastoma. As neuroblastoma is often asymptomatic children usually develop large abdominal tumours detectable by physical palpation usually confirmed by ultrasonography before the child is transferred to a specialised oncology centre. Example of Magnetic resonance imaging (MRI) **A.** axial T<sub>1</sub>-weighted **B.** axial T<sub>2</sub>-weighted showing large abdominal mass **C.** axial T<sub>1</sub>-weighted post-contrast images showing heterogeneous tumour enhancement **D.** apparent diffusion coefficient (ADC) map, **E.** fused <sup>123</sup>I-MIBG scan with SPECT-CT and to **F.** area of calcification on computerized tomography scan. This child presented with a rare case of ganglioneuroblastoma nodular characterised by a 1cm “circular” neuroblastoma nodule in the context of a large ganglioneuroma mass. Note that the neuroblastic lesion showed area of restricted diffusion, MIBG positivity and calcification consistent with the phenotype of undifferentiated neuroblastoma [Adapted from Jain et al (75)].  
.....46

**Figure 1.8 Histological characteristics of neuroblastic tumours.** **A-B.** Immature neuroblastic tissue. Neuroblastoma (Schwannian stroma-poor), poorly differentiated subtype, composed of undifferentiated neuroblastic cells with clearly recognisable neuropil. (**A.** original magnification x100; **B.** x200). **C.** Ganglioneuroma (Schwannian stroma-dominant), maturing subtype (original magnification x100). Note that there are some neuroblastic cells that have still not reached full maturation to ganglion cells. Those neuroblastic and ganglion cells are individually scattered and have not formed nests. **D.** Ganglioneuroma (Schwannian stroma-dominant), mature subtype (original magnification x200). Fully mature ganglion cells (arrows) with satellite cells are individually embedded in Schwannian stroma. [Adapted from Shimada et al (80)].  
.....49

**Figure 1.9 Childhood neuroblastoma.** Anatomical **A.** transverse and **B.** coronal T<sub>2</sub>-weighted MRI image from children with abdominal neuroblastoma. **C.** Sagittal T<sub>2</sub>-weighted MRI showing a case of thoracic neuroblastoma growing around the spine with local invasion. **D, E, F** show corresponding representative images of tumours arising in the Th-MYCN mouse. Tumours cause displacement of major abdominal organs such as the kidneys (k), small-bowel (s.b.) and the displacement of the abdominal aorta away from the spine (s) [Clinical images adapted from Elsayes et al (96), Goo et al (73), and Nour-Eldin et al (97)].  
.....52

**Figure 1.10 Histological Profile of Tumours from MYCN Transgenic Mice.** Haematoxylin and eosin (H&E) staining of tumour arising from the Th-MYCN genetically engineered model of neuroblastoma. Overall, tumours resembled high-risk childhood neuroblastoma **A.** Islands of undifferentiated neuroblastoma cancer cells separated by

fibrous septa **B.** Apoptotic cells distinguishable by the fragmentation of nuclear material and condensation of chromatin within them. **C.** Islands of larger differentiated neuroblastoma cells **D.** Neuropil (arrowed) among undifferentiated tumour **E.** Tingible body macrophages (arrows) **F.** Region of high haemorrhage in Th-MYCN GEM model .....53

**Figure 1.11** Developing radiologic pathology for the non-invasive characterisation of tumour heterogeneity: overview of the methodology. **A.** The primary abdominal tumours in the GEM model of neuroblastoma present with numerous anatomical landmarks (arrows, k, kidney) that are detectable by conventional anatomical MRI, which can be used to ensure precise registration between the MRI-derived parametric maps (example:  $R_2^*$ , sensitive to deoxyhaemoglobin concentration and thus a proxy for haemodynamic vasculature) and the histopathology slices, which is crucial for validation and calibration of the methodology. Imaging analysis and spatial statistics tools will be developed to **B.** evaluate the spatial correlation between the parametric maps of each selected MRI biomarker and the associated histopathological features and **C.** to integrate the different MRI biomarker maps into a spatial map of sub-regions (habitats) and evaluate their spatial correlation with the distribution of sub-regions defined by histopathology.....57

**Figure 2.1** Digitisation of a whole-slide image. The glass-slide is scanned by a digital slide scanner and the images are stored in a pyramid-like structure containing different magnification levels. At the “cell-level” (40x-20x) information about cell distribution and morphology are derived, while at the “tissue-level” (10x-1.25x) the different tissue structures are visible. ....61

**Figure 2.2** A deep convolutional neural network (CNN) trained for face recognition. Visualisation of the first hidden layer reveals the detected edge features. In the next hidden layer where the context is larger, the network captures shapes and face parts. The next hidden layer where the context is even larger, the network captures bigger part of the faces. The output layer predicts (classifies) the probability the input image belongs into a pre-defined set of categories (classes) (120). Adapted from Artificial Intelligence in Medical Imaging: Opportunities, Applications and Risks, page 28, figure 3.4.....66

**Figure 2.3** The cell segmentation and subsequent feature extraction process from H&E images using the CRImage pipeline. ....71

**Figure 2.4** Computational analysis of a digitized whole-slide histological image of a Th-MYCN neuroblastoma. Cells were segmented and classified into 5 categories with an overall accuracy of 94.3%: undifferentiated neuroblasts (98.61% accuracy, green), differentiating neuroblasts (96.79%, purple), apoptotic cells (95.41%, yellow), lymphocytes (96.15%, blue), stromal cells (84.54%, red). ....72

**Figure 2.5** Representative results of the computational analysis of a digitized whole-slide histological image of a Th-MYCN neuroblastoma; Cell outlines: green – undifferentiated neuroblasts, blue – lymphocytes, red – stromal cells, yellow – apoptotic/dying cells, purple – differentiated neuroblasts. ....73

**Figure 2.6** The complex nature of melanoma architecture. **A.** Heterogeneous tumour stroma makes accurate cell classification a difficult task. **B.** Superpixels captures tumour global architecture by delineating the boundaries among heterogeneous tumour components, including haemorrhage area, fatty tissues, stromal regions, epidermis and cancer nests. **C.** Current classification scheme assigned these components accurately to their respective superpixel classes. Top: example images; Bottom: segmentation and classification using SLIC superpixels.....76

**Figure 2.7** Proposed hierarchical framework to project tumour global context onto single cell classification by integrating superpixel segmentation and classification. ....79

**Figure 2.8** Superpixels provide global context for single cell classification. **A.** Representative superpixel classification of tumour regions overlaid with ground-truth cell annotations. **B.** ROC curves (cancer vs. all) illustrate the improvement in classification accuracy by adding superpixel context as additional features. **C.** Representative images comparing ground truth annotation and single cell classifiers with and without superpixels.....81

**Figure 2.9** Overview of the SuperCRF framework for analysing H&E-stained pathological images of melanoma. **A.** Major histological features of melanoma architecture. **B.** Projection of regional classification results using superpixels from various scales to the 20x magnification for the improvement of single-cell classification. **C.** Graphical representation of node dependencies (cells and superpixels) across different scales. **D.** Region classification scheme using a superpixel based machine-learning method in whole-slide images (5x and 1.25x magnification). **E.** Single-cell classification using a state-of-the-art spatially constrained-convolution neural network (SC-CNN) classifier **F.** representative results of the SC-CNN cell classifier alone and combined with our SuperCRF system. Note the misclassification of various stromal cells by the SC-CNN, which are corrected by our model. ....90

**Figure 2.10** Representative examples of both superpixel and single-cell classification with or without SuperCRF. **A.** Superpixels-based regional classification on representative whole slide images (5x magnification) of melanoma. Green: tumour area, Red: stroma area, Blue: normal epidermis, Yellow: lymphocyte cluster. **B.** Representative images showing cell classification using a state-of-the-art spatially constrained-convolution neural network (SC-CNN) and four conditional random fields (CRF) models. Note the mislabelling of many cancer and stromal cells as epidermis cells when using the SC-CNN and the gradual increase in classification accuracy with the best accuracy achieved with the SuperCRF. Green, cancer cells; Red, stromal cells; Blue, lymphocytes; Yellow, epidermis cells.....96

**Figure 2.11** Associations between survival outcomes and SuperCRF-define risk groups in the Cancer Genome Atlas (TCGA) cohorts of patients with melanoma. **A.** Kaplan-Meier Survival curves for patients in the high-risk group (blue) and low risk group classified by stromal cells ratio derived from SuperCRF (left) and using only the SC-CNN classifier. Note the difference in the p-value using the two methods. **B.** Kaplan-Meier Survival curves for patients in the high-risk group (blue) and low risk group classified by immune phenotype based on spatial distribution of lymphocytes in different tumour compartments derived from SuperCRF. ....98

**Figure 2.12** Representative examples of the SLIC superpixels segmentation and ground-truth annotations in TCGA melanoma samples **A.** Whole-slide image segmentation using the SLIC superpixels algorithm. Note how the superpixels adhere to the boundaries of the different components of the tumour with each superpixel containing a single type of tissue **B.** Ground-truth annotations are provided by the pathologists by marking samples of the region components (the different colors represent different regions) without the need for delineating the boundaries of the tumour components.108

**Figure 2.13** Architecture of our custom-designed convolutional neural network for the classification of superpixels into different tissue-level categories.....111

**Figure 2.14 A-F.** Representative examples of the results obtained from the application of the SuperHistopath pipeline in whole-slide images of tumours (5x) of the Cancer

Genome Atlas (TCGA) melanoma dataset (**G**. Magnified regions of interest). Note the important clinically-relevant phenotypes characterized by clusters of lymphocytes infiltrating the tumour in samples **B** and **D** or the majority of clusters of lymphocytes residing just outside the tumour area (left and central part) with only a few clusters infiltrating the tumour (right part) in sample **C**. .....116

**Figure 2.15 A-F.** Representative examples of the results obtained from the application of the SuperHistopath pipeline in whole-slide images of tumours (5x) of the triple-negative breast cancer (**G**. Magnified region of interest). Note the important clinically-relevant features, such as the amount of tumour necrosis inside tumours **A** and **B**, lymphocytes which, are infiltrating the tumour in large number in samples **C** and **D**, but are surrounding the stroma barrier without infiltrating the tumour in samples **A**, **B**, **E**, **F**. .....118

**Figure 2.16** Representative examples of the results obtained from the application of the SuperHistopath pipeline in whole-slide images of tumours (5x) arising in genetically-engineered mouse models of high-risk neuroblastoma (**B**. Magnified region of interest). .....121

**Figure 2.17** Associations between survival outcomes and SuperHistopath-defined risk groups in the Cancer Genome Atlas (TCGA) cohorts of patients with melanoma. **A**. Kaplan-Meier Survival curves for patients in the high-risk group (blue) and low risk group (red) classified by stromal cells ratio derived from SuperHistopath and **B**. Kaplan-Meier Survival curves for patients in the high-risk group (blue) and low risk group (red) classified by immune infiltrate based on lymphocytes cluster ratio derived from SuperHistopath. ....123

**Figure 2.18** SuperHistopath-based quantification of tumour phenotype in genetically-engineered mouse model of high-risk neuroblastoma. **A**. Representative SuperHistopath-segmented whole-slide images (5x) and pie chart showing the SuperHistopath quantified mean composition of the tumours arising in Th-MYCN (n=6) and Th-ALK<sup>F1174L</sup>/MYCN (n=7) mouse models of high-risk neuroblastoma. Note the marked difference of phenotype induced by the expression of the ALK<sup>F1174L</sup> mutation characterized by **B**. a significantly increased neuroblastoma differentiation neuroblasts and the total abrogation of the characteristic haemorrhagic phenotype of Th-MYCN tumours. ....125

**Figure 3.1** Representative T<sub>2</sub>-weighted images through the abdomen of two neuroblastoma-bearing Th-MYCN mice showing the numerous anatomical landmarks detected by conventional anatomical T<sub>2</sub>-weighted MRI (abdominal aorta: a, renal artery: r, lymph node: ly, kidney: k) which can be used in addition to tumour shape for the precise alignment of MRI with H&E- stained whole-slide high-resolution histology. ....133

**Figure 3.2** Representative examples of satisfactory correspondence between T<sub>2</sub>-weighted MRIs with their corresponding histopathological slices from tumours arising in Th-MYCN and Th-ALK<sup>F1174L</sup>/ Th-MYCN transgenic mouse models. Some of the intra-tumoural features under investigation (e.g. the artery, haemorrhage, lymph nodes, necrotic areas) are marked (yellow arrowheads). ....134

**Figure 3.3** Automatic extraction of red blood cells (RBC) from H&E images in Th-MYCN tumours. Whole-slide images of RBC were converted into binary and processed to match MRI resolution (234x234µm), with the fraction of pixels occupied by RBC within 518x518 pixel-regions representing a single pixel in the final RBC map.....136

**Figure 3.4** Representative example of the application of the coherent point drift (CPD) algorithm. Firstly, derived-from-histology density maps of all the segmented cells (“Cellularity”) were non-rigidly registered to the MRI T<sub>1</sub> images based on features extracted by a Canny edge detector. The same transformation is subsequently applied to the density maps of each classified cell category. ....137

**Figure 4.1** Susceptibility-contrast MRI-derived fBV predicts response to VEGFR (panVEGFR) inhibitor cediranib in the Th-MYCN model of neuroblastoma. **A.** Representative tumour parametric fBV maps in the Th-MYCN mice prior to (day 0) and 24 hours (day 2) after treatment with 6 mg/kg cediranib or vehicle. **B** and **C.** Changes (**B**) and relative changes (**C**) in tumour median fBV prior to (day 0) and 24 hours (day 2) after treatment with 6 mg/kg cediranib or vehicle. **D.** Representative T<sub>2</sub>-weighted anatomical MRI taken through the abdomen of tumour-bearing Th-MYCN mice prior to (day 0) and 24 hours (day 2) after treatment with 6 mg/kg cediranib or vehicle. **E** and **F.** Changes (**E**) and relative changes (**F**) in tumour volume prior to (day 0) and 24 hours (day 2) after treatment with 6 mg/kg cediranib or vehicle. Data are the individual median value for each tumour and the cohort mean  $\pm$  1 s.e.m. (\*\*\*, P < 0.005, two-tailed paired Student t test; \*\*, P < 0.01; \*\*\* P < 0.005, two-tailed unpaired Student t test, 5% level of significance). **G.** Representative T<sub>2</sub>-weighted anatomical MRI and tumour fBV parametric map of a Th-MYCN mouse prior, 24 hours, and 7 days after daily treatment with 6 mg/kg cediranib. **H** and **I.** Changes in tumour fBV (**H**) and tumour volume (**I**) in Th-MYCN mice prior, 24 hours, and 7 days after daily treatment with 6 mg/kg cediranib. **J.** Average tumour median fBV correlated with change in fBV over the 7 days of the trial (r = -0.83; P = 0.0008). **K.** Tumour median fBV correlated with change in tumour volume over the 7 days of the trial (r = -0.65; P = 0.02). ....147

**Figure 4.2** Intrinsic susceptibility MRI-derived transverse relaxation rate R<sub>2</sub>\* predicts response to VEGFR (panVEGFR) inhibitor cediranib in the Th-MYCN model of neuroblastoma. **A.** Waterfall plot documenting relative changes in tumour volume in the Th-MYCN mouse model of neuroblastoma following 7-day treatment with daily dose of 6 mg/kg cediranib or vehicle (DV<sub>cediranib\_D0-D7</sub> = -16  $\pm$  5% vs. DV<sub>vehicle\_D0-D7</sub> = 142  $\pm$  11%; P < 0.0001, two-tailed unpaired Student t test with a 5% level of significance). **B.** Baseline R<sub>2</sub>\* value (day 0) of responsive and nonresponsive tumours. Data are the individual median value for each tumour and the cohort mean  $\pm$  1 SEM (P < 0.0001, two-tailed unpaired Student t test with a 5% level of significance). **C.** Representative baseline native R<sub>2</sub>\* maps for responsive and progressive tumours in Th-MYCN transgenic mice (assessed by anatomical MRI) following daily treatment with 6 mg/kg cediranib for 7 days. **D.** Native tumour median R<sub>2</sub>\* (day 0) correlated with relative change in tumour volume following daily treatment with 6 mg/kg cediranib for 7 days (r = -0.72, P < 0.001). ....149

**Figure 4.3** Pipeline for the generation of computed maps of RBC and endomucin staining from histopathologic images in Th-MYCN tumours. **A.** RBCs were automatically extracted from H&E-stained whole section slides, with an accuracy of 99.74%. **B.** Endomucin staining was extracted from chromogenic immunohistochemistry whole section slides, with an accuracy of 97%. Note that the contrast of the whole slide extracted endomucin was enhanced, solely for printing visibility. Whole-section images were subsequently processed to match the MRI resolution, with the fraction of pixels occupied by RBCs (**A**) and endomucin (**B**) within 518x518 pixels of the original images, representing a single pixel in the final maps. ....151

**Figure 4.4** Susceptibility-contrast MRI-derived fBV correlates with endothelial cell marker endomucin fraction in tumours of the Th-MYCN model of neuroblastoma. **A.** Representative T<sub>2</sub>-weighted anatomical MRI images and their corresponding H&E-stained histology, parametric fBV maps, and computed areas of vascular endothelial cell

marker endomucin staining, KDE hotspot maps of high values (above the 85th percentile of each tumour sample) of fBV and endomucin fraction in the Th-MYCN model of neuroblastoma, 24 hours and 7 days after daily treatment with cediranib or vehicle (24 hours). Note the different scale for fBV and endomucin maps across the groups to better capture the intra-tumoural heterogeneity. **B** and **C**. Median tumour fBV (**B**) and median tumour endomucin fraction area (**C**) in the Th-MYCN GEM model 24 hours and 7 days after treatment with cediranib or vehicle (\*\*\*,  $P < 0.001$ , two-tailed unpaired Student t test). **D**. Tumour median fBV correlated with endomucin fraction area ( $r = 0.85$ ,  $P < 0.0001$ ).....152

**Figure 4.5** Intrinsic susceptibility MRI-derived transverse relaxation rate  $R_2^*$  correlates with tumour RBC content. **A**. Representative  $T_2$ -weighted MRI images and their corresponding H&E histology, maps of the transverse relaxation rate  $R_2^*$  and computed area of RBC detected from H&E-stained images, KDE hotspot maps of high values (above the 85th percentile of each tumour sample) of  $R_2^*$ , and RBC fraction in tumours arising in control Th-MYCN transgenic mouse. **B**. Tumour median  $R_2^*$  positively correlated with the mean value of RBC fraction detected from H&E-stained tumours arising in the Th-MYCN transgenic mouse model ( $r = 0.87$ ;  $P = 0.0003$ ).  $R_2^*$  values (**C**) and mean value of RBC fraction (**D**) also correlated with endomucin staining ( $r = 0.53$ ,  $P = 0.08$  and  $r = 0.64$ ,  $P = 0.02$ ).....155

**Figure 4.6** Calibration of intrinsic susceptibility MRI-derived tumour  $R_2^*$  phenotype with histopathologic phenotype in the Th-MYCN genetically engineered murine model of neuroblastoma. **A–F**. Susceptibility-weighted MRI-derived maps of the tumour transverse relaxation  $R_2^*$  and their corresponding H&E-stained whole-section slides (k., kidney), arranged by increasing median tumour  $R_2^*$  value. Tumours with very low median value of  $R_2^*$  ( $< 60 \text{ second}^{-1}$ ; **A**) present with a differentiating phenotype characterized by large areas of differentiating neuroblasts (**G**), with enlarged cytoplasm and more prominent nucleoli, surrounded by large amounts of neuropil (1.), alongside dense regions of undifferentiated neuroblasts (2.). Note that haemorrhage is absent from these tumours and the vasculature is characterized by sparse capillaries-like blood vessels as revealed by IHC staining for the vascular endothelial cells marker endomucin. For a median value of  $R_2^*$  of  $\sim 70 \text{ second}^{-1}$  and above, tumours present with dense area of poorly differentiated neuroblasts presenting with numerous mitotic features (**H**), yet tumour median  $R_2^*$  is determined by the increasing ratio of region with large area of haemorrhage and high vascular density (4., ---,  $\triangle$ ) over region with lower density of capillary like vessel (3., ---,  $\blacktriangleright$ ). Note that the extravasated RBC appear intact. Additionally, tumours with median  $R_2^*$  over  $110 \text{ second}^{-1}$  show vast area of cellular damage (necrosis) filled with intact and damaged RBCs (**I**; ---, 5.). Necrotic area with low RBC content (---, 6.) only present with small regional increase in  $R_2^*$  (as seen in **A**).158

**Figure 4.7** Haemorrhage in MYCN-amplified childhood neuroblastoma MRI and initial experience with SW-MRI in the neuroblastoma clinic. **A–C**. show abdominal axial fat-suppressed STIR  $T_2$ -weighted-MRI images, fat-suppressed SPAIR  $T_1$ -weighted MRI images, before and after administration of gadolinium (Gd)-based contrast agent, and diffusion-weighted MRI-derived ADC maps of children with neuroblastoma at the time of diagnosis. **A–C**. MYCN-amplified high-risk neuroblastoma in a 10-month old boy (**A**), a 2-year old boy (**B**), a 8-month-old boy (**C**), **D**. Comparison between a GNB in a 4-year-old female patient and GNB in a 5-year-old male patient. Note the presence of the neuroblastic nodule in the GNBn easily identified on ADC maps (arrowhead). **E**. Proportion of patients with MYCN-amplified neuroblastoma ( $n = 19$ ) presenting with a MRI phenotype, suggestive of the presence or absence of a hemorrhagic phenotype, based on the well-established knowledge of the appearance of aging blood in hematoma on conventional MRI (242), as illustrated in **F**. **G**. Abdominal  $T_2$ -weighted and  $T_1$ -weighted images (not fat-suppressed), ADC maps, intrinsic susceptibility MRI-derived

transverse relaxation  $R_2^*$  map of tumours, and post Gd contrast-enhanced  $T_1$ -weighted MRI images in children with refractory/relapsing neuroblastoma. (Patient 1, 7-year-old male; Patient 2, 5-year-old male; Patient 3, 6-year-old female). **H.** Tumour median  $R_2^*$  values for each individual tumour shown in **G** ( $\pm$ SD). Note that  $R_2^*$  values increased monotonically and approximately linearly with magnetic field strength-dependent and, as such, clinical  $R_2^*$  value at 1.5T was estimated to be four times lower than if they were measured at 7T (243).....162

**Figure 5.1** Representative  $T_2$ -weighted anatomical MR images of tumour-bearing Th-MYCN mice and associated parametric maps of the tumour spin-lattice relaxation time  $T_1$ , transverse relaxation rate  $R_2^*$ , spin-spin relaxation rate  $R_2$  ( $=1/T_2$ ), apparent diffusion coefficient (ADC) and magnetization transfer ratio (MTR), prior to and 24 hours following treatment with 25mg/kg vistusertib, 30mg/kg alisertib or vehicle.....177

**Figure 5.2** Summary of the tumour volumes quantified using  $T_2$ -weighted MRI at the time of enrolment (D0). There was no significant difference in tumour volume at D0 between the different treatment cohorts ( $P > 0.01$ , one-way ANOVA with a 5% level of significance).....177

**Figure 5.3** Waterfall plots documenting relative changes in tumour volume in the Th-MYCN mouse model of neuroblastoma 24 hours following treatment with **A.** 25mg/kg vistusertib, **B.** 30mg/kg alisertib, or their respective vehicle.....178

**Figure 5.4 A.** Scatter graph of the native spin lattice relaxation time  $T_1$  against native transverse relaxation rate  $R_2^*$  from 71 untreated tumours arising in genetically-engineered murine models of neuroblastoma. Linear regression analysis and associated 95% confidence and prediction intervals are shown. A highly significant negative correlation was obtained ( $r = -0.59$ ,  $P < 0.0001$  with Bonferroni correction [ $n=5$ ]). **B.** Box-and-whisker plot showing the difference in native  $T_1$  in sub-regions categorized by low ( $< 70s^{-1}$ ), intermediate ( $70s^{-1} < R_2^* < 250s^{-1}$ ) and high ( $> 250s^{-1}$ ) values of  $R_2^*$  measured in Th-MYCN tumours treated with vehicle ( $n= 13$ ). Data are medians and interquartile range. **C.** Scatter graph of relative changes in native tumour  $R_2^*$  ( $\Delta R_2^*$ ) and relative changes in native  $T_1$  ( $\Delta T_1$ ) 24 hours following treatment with either alisertib or vistusertib. Bold lines represent linear regression with crossed dots indicating outliers determined using the robust regression and outlier removal approach. Grey shaded area indicates the 95% confidence intervals while dashed lines indicate 95% prediction confidence. A significant negative correlation was obtained ( $r = -0.78$ ,  $P = 0.002$  with Bonferroni correction [ $n=5$ ]). .....179

**Figure 5.5** Representative MRI-derived parametric maps of the tumour spin-lattice relaxation time  $T_1$  and transverse relaxation rate  $R_2^*$ , and registered histopathology-derived parametric maps of cell density including undifferentiated and apoptotic neuroblasts in the Th-MYCN model of neuroblastoma, 24 hours following treatment with either vehicle control, 25mg/kg vistusertib or 30mg/kg alisertib. ....180

**Figure 5.6 A.** Box-and-whisker plot showing the difference in native  $T_1$  values in sub-regions categorized by low and high density of undifferentiated neuroblasts. Dichotomization was achieved using either median, Otsu or 85<sup>th</sup> percentile thresholds on registered histopathology-derived parametric maps of segmented and classified undifferentiated neuroblasts in vehicle control Th-MYCN tumours ( $n=13$ ). **B.** Box-and-whisker plot showing the difference in undifferentiated neuroblast density in sub-regions categorized by low and high native  $T_1$  values, defined using either median, Otsu or  $T_1 > 1900ms$  thresholds in vehicle control Th-MYCN tumours. Data are medians and interquartile range. ( $P$ , Wilcoxon signed rank test with a 5% level of significance) **C.** Proportion of undifferentiated and apoptotic neuroblasts relative to all cells derived from

cell segmentation and classification from haematoxylin and eosin (H&E) stained histopathology from Th-MYCN tumours 24 hours following treatment with either vehicle control, 25mg/kg vistusertib or 30mg/kg alisertib. **D.** Scatter graph showing that the reduction in native tumour  $T_1$  over 24h treatment with either alisertib or vistusertib correlated with an increased proportion of apoptotic cells present in the tumour at the time of excision ( $r= 0.55$ ,  $P=0.04$ ). **E,F.** Scatter graphs showing that median tumour native  $T_1$  in the treated and vehicle control cohorts positively correlated with the proportion of undifferentiated neuroblasts ( $r= 0.70$ ,  $P < 0.0001$ ), and negatively correlated with the proportion of apoptotic neuroblasts ( $r= -0.63$ ,  $P = 0.006$ ). Grey shaded area indicates 95% confidence intervals while dashed lines indicate 95% prediction confidence.....182

**Figure 5.7** Quantification of cleaved caspase-3 staining in tumours arising in the Th-MYCN mouse model of neuroblastoma 24h after treatment with 30mg/kg alisertib or vehicle.....184

**Figure 5.8 A.** Scatter graph showing that the proportion of apoptotic cells in tumours arising in the Th-MYCN mouse model of neuroblastoma 24h after treatment with 30mg/kg alisertib or vehicle and quantified using the cell classification algorithm significantly correlated ( $r = 0.88$ ,  $P < 0.0001$ ) with the proportion of cells staining positively for cleaved caspase 3 on adjacent IHC slides. The grey shaded area indicates the 95% confidence intervals while the dashed lines indicate 95% prediction confidence **B.** Histopathology-derived parametric maps showing apoptotic cells density determined using the cell classification algorithm and the density of cells staining positively for cleaved caspase 3 in Th-MYCN tumours 24 hours following treatment with 30mg/kg alisertib or vehicle.....184

**Figure 5.9** Three cases of differentiating tumours in the Th-MYCN model of neuroblastoma. **A.** Representative MRI-derived parametric maps of the tumour spin-lattice relaxation time  $T_1$  and transverse relaxation rate  $R_2^*$ , and registered representative pathology-derived parametric maps of tumour cell density including undifferentiated, apoptotic and differentiating neuroblasts.....185

**Figure 6.1** Representative maps of tumour ADC values and ADC compartments derived from Gaussian mixture models with registered histopathology-derived maps of cell density including undifferentiated and apoptotic neuroblasts in the Th-MYCN model of neuroblastoma, 24h following treatment with vehicle control or vistusertib. ....195

**Figure 6.2** Gaussian mixture modelling of the apparent diffusion coefficient distribution in the tumours of the Th-MYCN model of neuroblastoma prior and after treatment with vistusertib or vehicle control. ....196

**Figure 6.3** Representative histopathological characteristic of the ADC compartments derived from Gaussian mixture modelling of the apparent diffusion coefficient distribution in the Th-MYCN model of neuroblastoma.....197

**Figure 6.4 A.** Changes in tumour median apparent diffusion coefficient (ADC) in the Th-MYCN model of neuroblastoma 24h following treatment with vehicle control or vistusertib. Note that there was not significant change in tumour median ADC over treatment. Interestingly there was significant difference in tumour median ADC between treated and vehicle control tumours at the study endpoint (24h). **B.** Difference in tumour median cell density, the ratio of undifferentiated neuroblasts and apoptotic cells to all cells in tumours 24h after treatment with vehicle control or vistusertib.....199

**Figure 6.5** Relationship between the change in median tumour ADC and the change in tumour median magnetisation transfer ratio (MTR) in tumours of the Th-MYCN model of neuroblastoma 24h after treatment with vistusertib.....199

**Figure 6.6** Relationships between median tumour ADC, median tumour  $R_2^*$  and median tumour cell density in neuroblasts in tumours of the Th-MYCN model of neuroblastoma 24h after treatment with vistusertib or vehicle control. ....201

**Figure 6.7** Relationships between ADC compartments derived from Gaussian mixture models and median tumour cell density in neuroblasts in tumours of the Th-MYCN model of neuroblastoma 24h after treatment with vistusertib or vehicle control.....202

**Figure 6.8 A.** Sub-regional analysis demonstrates that several ADC compartments have significantly different median  $R_2^*$ . \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.005$ , Wilcoxon signed rank test. **B.** Box-and-whisker plot showing the difference in ADC in sub-regions categorized by low ( $< 70s^{-1}$ ), intermediate ( $70s^{-1} < R_2^* < 250s^{-1}$ ) and high ( $> 250s^{-1}$ ) values of  $R_2^*$  measured in Th-MYCN tumours treated with vehicle ( $n = 13$ ). Data are medians and interquartile range. ....203

**Figure 6.9** Sub-regional analysis defining the cell density, the fraction of undifferentiated neuroblasts and apoptotic cells, and native  $T_1$  value associated with the ADC compartment \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.005$ , Wilcoxon signed rank test. ....207

**Figure 6.10** Relationships between ADC compartments derived from Gaussian mixture models and the fraction of undifferentiated and apoptotic neuroblasts in tumours of the Th-MYCN model of neuroblastoma 24h after treatment with vistusertib or vehicle control. ....208

**Figure 6.11**  $ADC_h$  compartment derived from Gaussian mixture modelling spatially associates with area dense in differentiating neuroblasts in 3 cases of differentiating tumours arising in the Th-MYCN model of neuroblastoma.....210

**Figure 7.1 An overview of the methodology for habitat imaging.** The supervised machine learning algorithm was trained using the values from registered histology-derived maps as labels to characterise distinct tumour sub-regions (classification) or predict the histology values (regression) from multi-parametric MRI ( $T_1$ ,  $R_2^*$ , ADC) data. ....221

**Figure 7.2** Representative examples of classified tumours using empirical thresholds for  $R_2^*$ ,  $T_1$  and ADC. Note the unclassified intra-tumour areas in black.....223

**Figure 7.3** Overview of the 10-fold cross-validation classification accuracies between the MRI-derived and histology-derived data achieved by the linear regression (LR), linear discriminant analysis (LDA), k-nearest neighbour (KNN), decision trees (CART), Gaussian naïve Bayes (NB), support vector machines (SVM), Adaboost (AB), gradient boosting machines (GBM), random forests (RF) and extra trees (ET) algorithms.....225

**Figure 7.4** Representative examples of multi-parametric MRI (mpMRI)-derived habitat maps and the region classification maps from histology (SuperHistopath). Note that many areas of haemorrhage are classified as tissue damage after the processing into MRI resolution of the SuperHistopath results, since areas of tissue damage very often include extensive haemorrhage.....227

**Figure 7.5** Overview of the mean square error between the MRI-derived and histology-derived data achieved after the 10-fold cross-validation by the linear regression (LR), k-

nearest neighbour (KNN), decision trees (CART), support vector regression (SVR), XGBoost, light gradient boosting machines (LGBost), random forests (RF) and extra trees (ET) algorithms.....228

**Figure 7.6** Representative examples of multi-parametric MRI (mpMRI)-derived and histology derived maps of different cell population densities.....232

**Figure 8.1** Towards non-invasive pathology: summary of the results.....240

## List of Tables

<b>Table 1.1</b> Non-invasive and clinically available functional MRI biomarkers of the tumour microenvironment.....	36
<b>Table 1.2</b> Applications of habitat imaging .....	42
<b>Table 1.3</b> Summary of favourable and unfavourable histological characteristics (Shimada classification system) [Adapted from Shimada et al (77)] .....	49
<b>Table 2.1</b> Common elements of convolutional neural networks (CNN).....	67
<b>Table 2.2</b> Confusion matrix of the cell classification results in the Th-MYCN mouse model of neuroblastoma .....	72
<b>Table 2.3</b> Accuracy matrix of the classification of superpixels for single sets of features (left) and various combinations (right).....	80
<b>Table 2.4</b> Confusion matrix of the Single Cell Classifiers, based on nuclei morphology (left), adding of the global context as extra feature in the feature set (middle) and implementation of the voting scheme (right). C: cancer, E: epidermis, L: lymphocyte, S: stromal. Note the confusion between cancer cells and epidermis cells highlighted in red. ....	83
<b>Table 2.5</b> Classification performance for the three classifiers. Both approaches of incorporating the global context led to markedly better results, with the voting scheme achieving the highest score.....	83
<b>Table 2.6</b> Summary of the data used to train and test the different parts of the SuperCRF system, as well as to study the cancer-immune-stroma interface. ....	86
<b>Table 2.7</b> Evaluation of different conditional random fields (CRF) versions and a state-of-the-art spatially constrained-convolution neural network (SC-CNN) deep learning cell-classifier.....	94
<b>Table 2.8</b> Confusion matrix of the classified cells from the spatially constrained-convolution neural network (SC-CNN) deep learning cell-classifier. C: cancer cells, E: epidermis cells, L: lymphocytes, S: Stromal cells. ....	94
<b>Table 2.9</b> Confusion matrix of the classified cells from singleCellCRF architecture, where the trained conditional random field (CRF) combines the classified cells from the spatially constrained-convolution neural network (SC-CNN), with local context information (neighbouring classified cells). C: cancer cells, E: epidermis cells, L: lymphocytes, S: Stromal cells.....	94
<b>Table 2.10</b> Confusion matrix of the classified cells from CRF1.25x architecture, where the trained conditional random field (CRF) combines the classified cells from the spatially constrained-convolution neural network (SC-CNN), with the region classification information from the 1.25x magnification whole-slide images. C: cancer cells, E: epidermis cells, L: lymphocytes, S: Stromal cells. ....	94
<b>Table 2.11</b> Confusion matrix of the classified cells from CRF5x architecture, where the trained conditional random field (CRF) combines the classified cells from the spatially	

constrained-convolution neural network (SC-CNN), with the region classification information from the 5x magnification whole-slide images. C: cancer cells, E: epidermis cells, L: lymphocytes, S: Stromal cells.....95

**Table 2.12** Confusion matrix of the classified cells from SuperCRF architecture, where the trained conditional random field (CRF) combines the classified cells from the spatially constrained-convolution neural network (SC-CNN), with the region classification information from the 1.25x and 5x magnification whole-slide images and local context information (neighbouring classified cells). C: cancer cells, E: epidermis cells, L: lymphocytes, S: Stromal cells.....95

**Table 2.13** Summary of the datasets used.....106

**Table 2.14** Summary of the datasets used for training and testing the convolutional neural network. Note that the testing datasets consisted of whole-slide images from different patients from the training dataset.....109

**Table 2.15** Evaluation metrics of the different neural network architectures in the TCGA melanoma test dataset. ....114

**Table 2.16** Confusion matrix of the classification of superpixels using the optimized Xception network in melanoma patients in 6 categories: tumour, stroma, normal epidermis, cluster of lymphocytes (Lym), fat and empty/white space (separate test set of 5 whole-slide images). Overall accuracy = 98.8%, average precision = 96.9%, average recall = 98.5%. ....114

**Table 2.17** Confusion matrix of the classification of superpixels using the custom-made CNN in melanoma patients in 6 categories: tumour, stroma, normal epidermis, cluster of lymphocytes (Lym), fat and empty/white space (independent test set of 5 whole-slide images). Overall accuracy = 96.7%, average precision = 93.6%, average recall = 93.6%. ....115

**Table 2.18** Confusion matrix of the classification of superpixels using the optimized Xception network in triple-negative breast cancer patients in 6 categories: tumour, necrosis, cluster of lymphocytes (Lym), stroma, fat and lumen/empty space (separate test set of 5 whole-slide images). Overall accuracy = 93.1%, average precision = 93.9%, average recall = 93.6%.....117

**Table 2.19** Confusion matrix of the classification of superpixels using the custom-made CNN in triple-negative breast cancer patients in 6 categories: tumour, necrosis, cluster of lymphocytes (Lym), stroma, fat and lumen/empty space (independent test set of 5 whole-slide images). Overall accuracy = 92%, average precision = 92.7%, average recall = 92.2%.....117

**Table 2.20** Confusion matrix of the classification of superpixels using the optimized Xception network in the Th-MYCN and Th-ALK<sup>F1174L</sup>/MYCN mouse models in 8 categories: region of undifferentiated neuroblasts, necrosis, cluster of lymphocytes (Lym), haemorrhage (blood), empty/white space, muscle tissue and kidney (separate test set of 16 whole-slide images). Overall accuracy = 98.3%, average precision = 98.5%, average recall = 98.4%.....119

**Table 2.21** Confusion matrix of the classification of superpixels using the custom-made CNN in the Th-MYCN and Th-ALK<sup>F1174L</sup>/MYCN mouse models in 8 categories: region of undifferentiated neuroblasts, necrosis, cluster of lymphocytes (Lym), haemorrhage (blood), empty/white space, muscle tissue and kidney (independent test set of 16 whole-

slide images). Overall accuracy = 96.8%, average precision = 97.2%, average recall = 97.1%.....120

**Table 4.1** Summary of Mantel test results showing the highest correlation of distance matrices obtained between MRI fractional blood volume fBV and computed area of vascular endothelial cell marker endomucin staining maps following manual, non-rigid and rigid registration.....153

**Table 4.2** Summary of Mantel test results showing the highest correlation of distance matrices obtained between MRI transverse relaxation  $R_2^*$  and computed red blood cell maps following manual, non-rigid and rigid registration. ....156

**Table 4.3** Transverse relaxation rate  $R_2^*$  and pathological characteristics of the Th-MYCN tumours.....159

**Table 5.1** Summary of the response of the Th-MYCN transgenic model of neuroblastoma to vistusertib. Data are presented as mean of tumour median value  $\pm$  1 s.e.m. <sup>§</sup> Student's two-tailed paired t-test, <sup>‡</sup> Student's two-tailed unpaired t-test, both incorporating a Bonferroni correction (n=6) and assuming a 1% level of significance. The difference in the number of mice associated with the different parameters reflects that it was not possible to acquire the full protocol in all cases. ....175

**Table 5.2** Summary of the response of the Th-MYCN transgenic model of neuroblastoma to alisertib. Data are presented as mean of tumour median value  $\pm$  1 s.e.m. <sup>§</sup> Student's two-tailed paired t-test, <sup>‡</sup> Student's two-tailed unpaired t-test, both incorporating a Bonferroni correction (n=6), and assuming a 1% level of significance. The difference in the number of mice associated with the different parameters reflect that it was not possible to acquire the full protocol in all cases. ....176

**Table 5.3** Summary of sub-regional analysis for the comparison of  $T_1$  with segmented cell density and classified undifferentiated neuroblasts density. ....181

**Table 7.1** A summary of the empirical thresholds used to confidently characterise specific tumour regions in Th-MYCN mouse model of neuroblastoma.....218

**Table 7.2** Tumour percentages that were classified with high certainty after the application of empirical thresholds in all mice.....222

**Table 7.3** Summary of pixel-to-pixel similarity between the MRI-derived habitats and the region classification maps from histology. To measure similarity the accuracy metric was used (note that accuracy equals to the Jaccard similarity index for a multi-label task such as this). ....226

**Table 7.4** Summary of the mean square error (MSE) and structural similarity index (SSIM) between the MRI-derived maps and the histology-derived maps for all Th-MYCN tumours using the multi-output random forests algorithm. ....230

**Table 7.5** Summary of the mean square error (MSE) and structural similarity index (SSIM) between the MRI-derived maps and the histology-derived maps for all Th-MYCN tumours using the single-output random forests algorithm.....231

**Table A.1** Summary of the MRI sequences and parameters used in this study.....276

## List of abbreviations

ADC	Apparent diffusion coefficient
AI	Artificial intelligence
AIK	Akaike information criterion
ALK	Anaplastic lymphoma kinase
ANN	Artificial neural network
ASL	Arterial spin labelling
BET	Bromodomain extra terminal
BIC	Bayesian information criterion
BOLD	Blood oxygen level dependent
CART	Classification and regression trees
CE	Contrast enhanced
CNN	Convolutional neural network
CoV	Coefficient of variation
CPD	Coherent point drift
CRF	Conditional random field
CT	Computed tomography
DCE	Dynamic contrast enhanced
DCNN	Deep convolutional neural network
DW	Diffusion-weighted
DWI	Diffusion-weighted imaging
ETPS	Efficient topology preserving segmentation
fBV	Fractional blood volume
FCN	Fully-connected network
FDG	[F-18]2-fluoro-2-deoxyglucose
FFPE	Formalin-fixed and paraffin-embedded
FISP	Fast imaging with steady-state precession
FLAIR	Fluid-attenuated inversion-recovery
GBM	Gradient boosting machines
Gd	Gadolinium
GEM	Genetically-engineered mouse
GOSH	Great Ormond Street Hospital

GMM	Gaussian mixture model(ing)
GNB	Ganglioneuroblastoma intermixed
GNBn	Ganglioneuroblastoma nodular
H&E	Haematoxylin and eosin
IHC	Immunohistochemistry
INRC	International Neuroblastoma Response Criteria
IS	Intrinsic susceptibility
IVIM	Intravoxel incoherent motion
KDE	Kernel density estimation
KNN	K-nearest neighbour
MGE	Multi-gradient echo
MIBG	Metaiodobenzylguanidine
MOLLI	Modified Look-Locker inversion recovery
MRI	Magnetic resonance imaging
MRS	Magnetic resonance spectroscopy
MSE	Mean squared error
MTR	Magnetisation transfer rate
NB	Neuroblastoma
OE	Oxygen-enhanced
OxyE	Oxygen-enhanced voxels
OxyR	Oxygen-refractory voxels
PCA	Principal component analysis
PET	Positron emission tomography
pMRI	Perfusion MRI
RARE	Rapid acquisition with refocused echoes
RBC	Red blood cells
RBF	Radial basis function
RECIST	Response Evaluation Criteria in Solid Tumours
ReLU	Rectified linear unit
RILBP	Rotation-invariant local binary patterns
ROC	Receiver operating characteristic
ROI	Region of interest
SC	Susceptibility contrast
SC-CNN	Spatially constrained convolutional neural network

SFTA	Segmentation-based fractal texture analysis
SLIC	Simple linear iterative clustering
SSIM	Structural similarity index metric
SVM	Support vector machine
SVR	Support vector regression
SW	Susceptibility-weighted
TCGA	The Cancer Genome Atlas
TE	Echo time
TIL	Tumour-infiltrating lymphocytes
TME	Tumour microenvironment
TR	Repetition time
TRACERx	TRACKing Cancer Evolution through Therapy (Rx)
USPIO	Ultra-small paramagnetic iron oxides
VEGF	Vascular endothelial growth factor
VERDICT	Vascular, extracellular and restricted diffusion for cytometry in tumour
WSIs	Whole-slide images

*"Do or do not... there is no try!"*

Yoda, Jedi Master

*The Empire Strikes Back (1980)*

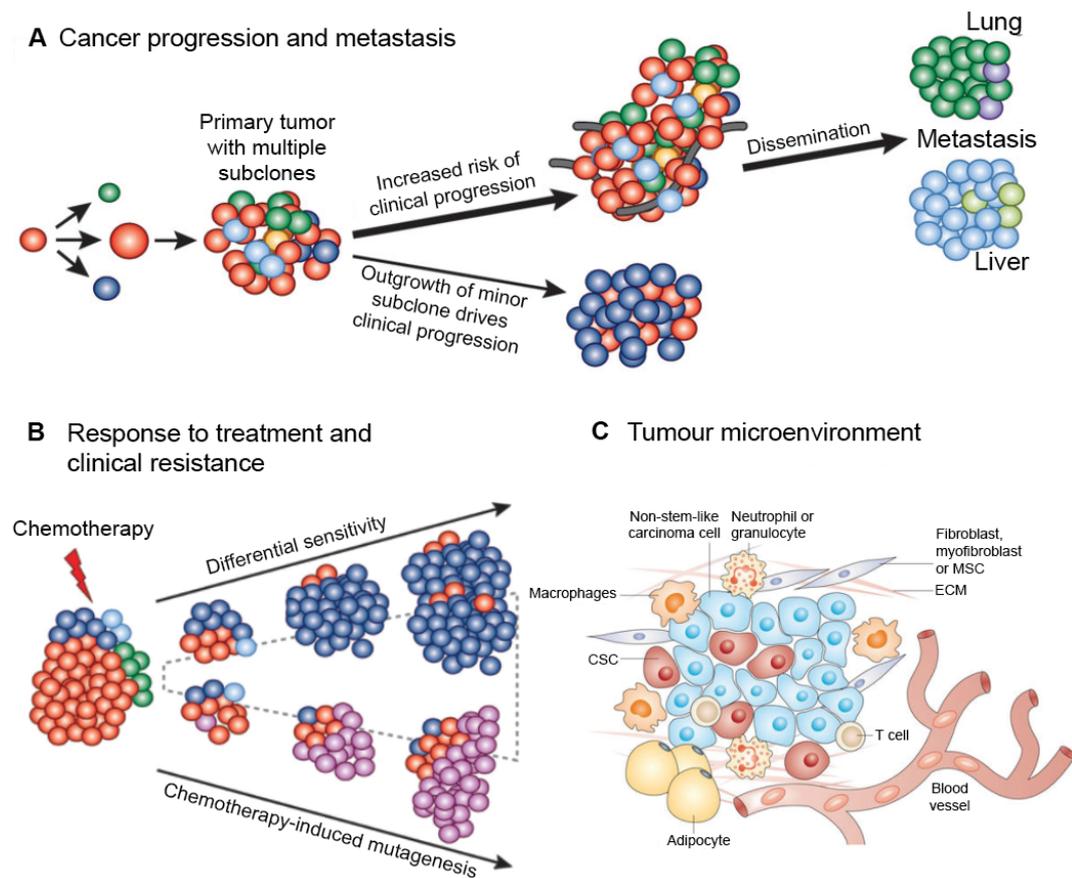
## **Chapter 1 : Introduction**

### **1.1 Understanding tumour heterogeneous evolving ecosystem to deliver precision medicine.**

#### *1.1.1 Tumour heterogeneity and evolution are major obstacles to the delivery of precision medicine.*

Tumours are intrinsically heterogeneous, dynamic and open systems across multiple scales (1). They consist of many interacting components, which can change over time, while simultaneously interacting to their surroundings including distant host tissue. Intra-tumoural heterogeneity is a major cause of resistance to therapy, has severe implications for the development of both novel therapeutics and biomarkers, and represents a major obstacle to the delivery of precision medicine (1) (Figure 1.1).

Molecular pathology has unravelled the genetic complexity of malignant tumours, with increasing evidence for the presence of enclaves of cells with distinct genetic signatures within the same tumour, which, in accordance with the Darwinian theory of evolution, are the result of selective microenvironmental pressures within intra-tumoural microhabitats. It is the interactions between microenvironmental selective pressures and the cells, i.e. the phenotype not genotype, that dictate cancer cell evolution (2, 3).



**Figure 1.1** Tumour heterogeneity and its role in clinical progression. **A.** Sequential accumulation of somatic events by the malignant clone, drives disease progression and metastasis. **B.** The presence of genetically heterogeneous sub-clones with inherent differences in drug sensitivity can also affect response to treatment and cause clinical resistance. Further evolution of the dominant clone in response to chemotherapy, may lead to relapse after initial treatment response. **C.** The tumour microenvironment, apart from harbouring carcinoma cells, consists of various components that have a major role in influencing the outcome of the malignancy [Adapted from *Kleppe et al* (4) and *Pattabiraman et al* (5)].

### 1.1.2 The tumour as an ecosystem

Cancer can also be considered as an ecological process (6). Each tumour can be viewed as a unique evolving ecosystem where distinct tumour regions (“habitats”) are interacting with each other and different cell populations are adapting under different selective pressures (including the immune system, access to nutrients, therapy) to gain evolutionary fitness advantages.

Understanding the ecologic mechanisms, especially the role of extracellular pH in the progression of prostate cancer allowed *Gatenby and colleagues* to control the proliferation of a more aggressive cell population *in vivo* and steer the tumour to a less invasive pathway by inducing small increase in the intraprostatic pH (7).

From an ecological perspective, the task of studying such a system would require both microscopic and macroscopic approaches. Landscape ecology methods can be exploited on digitized histopathological slides to investigate the tumour microenvironment (TME) and uncover the link between different spatial scales i.e. the tissue-, cellular- and molecular-level dynamics.

### *1.1.3 Understanding and quantifying tumour ecology*

#### 1.1.3.1 Computational pathology-enabled spatial analysis

From the perspective of the cancer cell, its ecology can be generally described by *hazards*, which are defined as everything that can kill the cell, and *resources* (8-11). The spatial heterogeneity of hazards and resources can dictate how a cell population will evolve with important impact to the patient prognosis.

Apart from the therapy itself, one of the main hazards a cancer cell faces is the immune system. Tumour immune infiltration has been associated with favourable patient prognosis in many studies (6, 12-21). The spatial distribution of lymphocytes, including their co-localisation with cancer cells have also been shown to have independent prognostic value (22, 23).

Some of the resources, such as nutrients and growth and survival signals, may be provided by the stroma after cancer cells have invaded the basement membrane (24-29). Cancer-associated fibroblasts are associated with poor prognosis (30) as they promote tumour progression, immune evasion and resistance to treatment (31-33). As with immune infiltration, ecological theory predicts that spatial heterogeneity of resources may be important and should promote invasion and metastasis, as cells tend to attempt to escape regions of sparse resources for more rich ones.

In the era of digital pathology, computer vision techniques allow the automatic detection and classification of every cell component. *Yuan and colleagues* have successfully applied the hazard-resources ecological concept in high-grade serous ovarian carcinoma by demonstrating that the quantification of the local ecological microenvironment (“local EcoScore”) was independently prognostic of overall survival (34). Furthermore, a study in the multi-region TRACKing Cancer Evolution through Therapy (Rx) (TRACERx) cohort of 100 patients (35) demonstrated that “immune-cold” regions and regions with extensive and irregular cancer-stroma cell interface in lung adenocarcinoma tumours are strongly associated with poor prognosis (36).

#### 1.1.3.2 Limitations of the approach: a role for functional imaging?

Diagnostic molecular pathology relies on surgical biopsies to obtain information about the histological, molecular and genetic properties of the tumour. However, biopsy samples represent only a “snapshot” of the tumour at a given time or location. Monitoring tumour evolution would require multiple repeated biopsies, which is limited by potential complications and is often not feasible, especially in the paediatric population. Furthermore, the highly heterogeneous nature of many tumour types, results in the biopsy sample not being representative of the entire tumour. This was further demonstrated in the multi-region TRACERx cohort, where the existence of “immune-cold” regions was independently prognostic for survival despite other regions of the tumour being “immune-hot” (36).

The non-invasive nature of anatomical and functional imaging allows the longitudinal interrogation and characterisation of the tumour microenvironment. Although the length scales of these methodologies do not allow the imaging of single cells, they can characterise the regions in which they reside, potentially revealing the spatial variations in the density of different cell populations and inform on tumour *resources* via metrics sensitive to oxygen consumption, tumour metabolism and blood flow.

A useful example of how this can be achieved comes from ecology. Hypothetically, after ground inspection, landscape ecologists have identified that among other species, an area is inhabited by grey and fox squirrels. Grey squirrels are efficient in foraging for the necessary nutrients, but are not equally competent in avoiding predators. Thus, they predominantly live in densely wooded areas. On the contrary, fox squirrels can effectively defend themselves against predators, but are less potent in finding nutrients. As a result, they inhabit sparsely wooded areas on the outskirts of cities. Therefore, by defining the habitat characteristics from a satellite image, ecologists can reliably estimate the spatial distribution of these two species (3). In an analogous way, by identifying and validating the distinct habitats visible on MRI, their molecular properties and dominant cell populations observable on pathology can be reliably predicted.

As a result, functional imaging has a potential role to play in informing on the temporal changes in tumour phenotype both at a regional and global level (3D) and complement the high-resolution “snapshot” provided by computational pathology analysis of surgical biopsies.

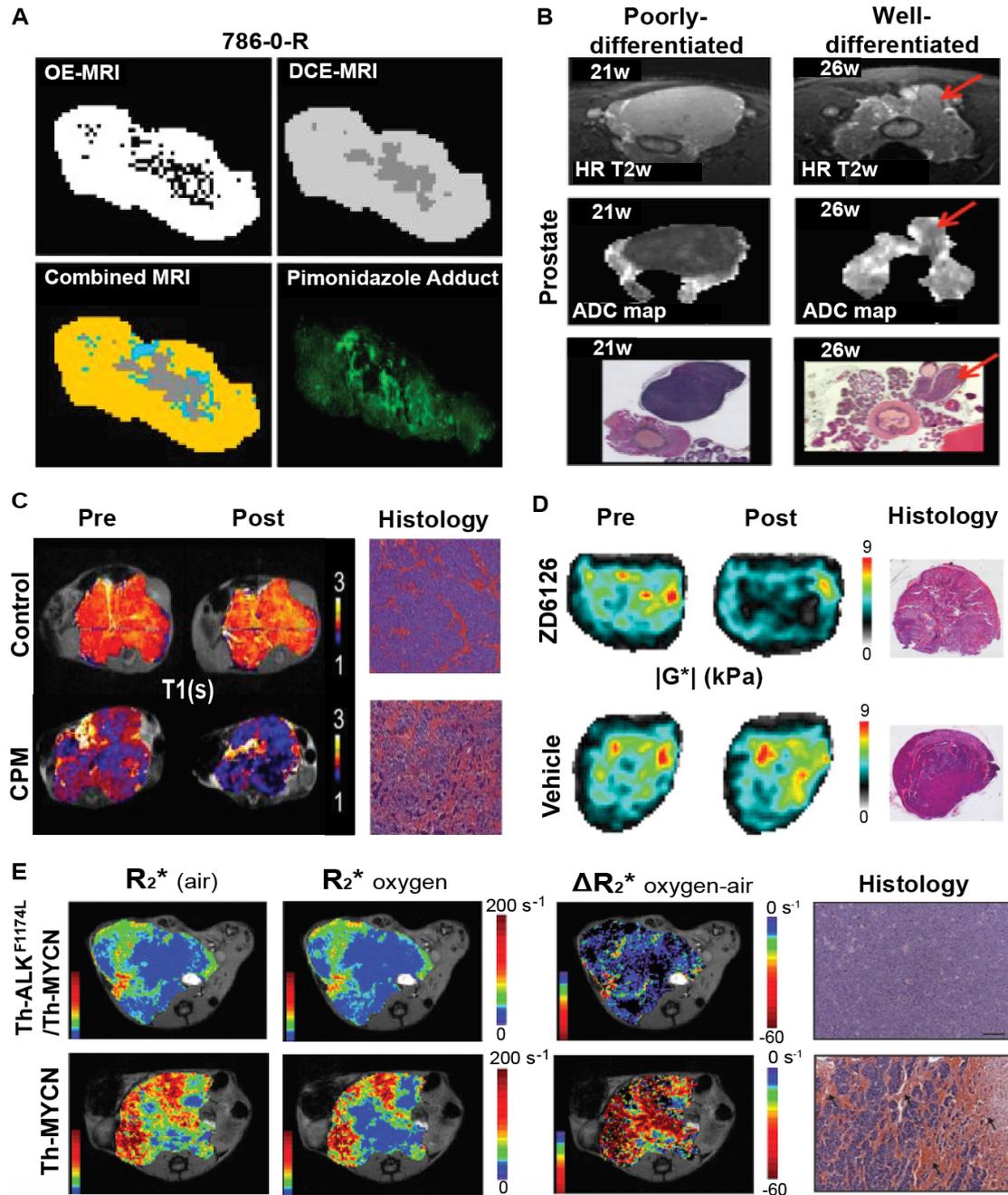
## **1.2 Functional imaging biomarkers, radiomics and habitat imaging for the delivery of precision medicine.**

### *1.2.1 MRI imaging biomarkers to characterise the tumour microenvironment*

Advanced anatomical, structural, functional and molecular MRI strategies not only provide information about tumour shape and boundaries, but also non-invasive metrics that inform on both spatial and temporal changes in key hallmarks of cancer, such as tumour metabolism, vascularisation, inflammation, cellular and stromal integrity, as well as tumour chemical and mechanical properties (pH, interstitial fluid pressure, local stiffness) (37) (Table 1.1). MRI, thus, represents an attractive choice to non-invasively and pathologically characterise the tumour phenotype and its evolution (38) (Figure 1.2).

Methodology	Contrast	Parameters	Sensitive to / Measure	Biomarker of
T <sub>1</sub> Relaxometry	Endogenous	T <sub>1</sub> (s)	Ratio bound/free water, Paramagnetic ions (haemoglobin breakdown)	Cell death Vascular injury
DCE MRI	Exogenous contrast agent (Gd-based)	K <sup>trans</sup> (min <sup>-1</sup> ) IAUGC <sub>60</sub> (mM Gd.min) EF(%)	Delivery of gadolinium contrast agent	Perfusion/perm eability
MT-MRI	Endogenous	MTR	Ratio bound/free water	Cell death Fibrosis
IS-MRI	Endogenous	R <sub>2</sub> <sup>*</sup> (s <sup>-1</sup> )	Deoxyhaemoglobin in erythrocytes	Haemodynamic function
BOLD-MRI	Gas challenge with >30% O <sub>2</sub>	$\Delta R_2^*(s^{-1}) =$ $R_{2^*O_2} - R_{2^*air}$ $fBV$	Re-oxygenation of deoxyhaemoglobin in erythrocytes during gas challenge	Vascular perfusion Hypoxia
SC-MRI	Exogenous contrast agent (USPIO)	fBV (%)	Fractional Blood Volume	Perfusion
OE T <sub>1w</sub> MRI	Gas challenge with >30% O <sub>2</sub>	$\Delta T_1^*(s) =$ $T_{1^*O_2} - T_{1^*air}$ Oxy-R fraction	Delivery of oxygen to tissue during gas challenge Fraction of tumour tissue refractory to gas challenge	Perfusion, Hypoxia
DW imaging	Endogenous	ADC (mm <sup>2</sup> .s <sup>-1</sup> )	Diffusion of water molecules	Cell death
IVIM imaging	Endogenous	D*(mm <sup>2</sup> .s <sup>-1</sup> ) F(%)	Water flowing in randomly oriented capillaries	Perfusion
ASL-MRI	Endogenous	Blood flow (mL. 100 g <sup>-1</sup> .min <sup>-1</sup> )	Diffusion of arterial blood water	Perfusion
MR Elastography	Endogenous	Gd (kPa) GI (kPa)	The viscoelastic properties of tissues via the visualisation of the propagation of exogenously applied mechanical waves	Tissue integrity Cell death Fibrosis Microvessel density

**Table 1.1** Non-invasive and clinically available functional MRI biomarkers of the tumour microenvironment



**Figure 1.2** Non-invasive MR biomarkers **A. Tumour Perfused OxyR: a spatial biomarker of hypoxia.** Representative maps showing oxygen-enhancing (OxyE) and non-enhancing (OxyR) voxel and perfused and non-perfused voxels following oxygen breathing (OE-T<sub>1</sub>-weighted MRI and Dynamic contrast enhanced MRI respectively). Combined maps show that perfused OxyR spatially correlate with pimonidazole adduct formation, the gold-standard for the immunohistopathological detection of hypoxia. **B. Tumour apparent diffusion coefficient (ADC): a biomarker of disease progression in the TRAMP GEM model of prostate cancer.** Parametric maps of ADC allow, in the prostate of the TRAMP GEM model, the detection of malignant tissue from well-differentiated, due to increase cellularity, which restricts water diffusion. More aggressive undifferentiated disease, shows a marked lower ADC compared to well-differentiated disease. **C. Native tumour spin-lattice relaxation time T<sub>1</sub>: A biomarker of response to therapy in GEM model of Neuroblastoma.** Treatment of the Th-MYCN model of neuroblastoma with cyclophosphamide, leads to a reduction in native T<sub>1</sub>. **D. The**

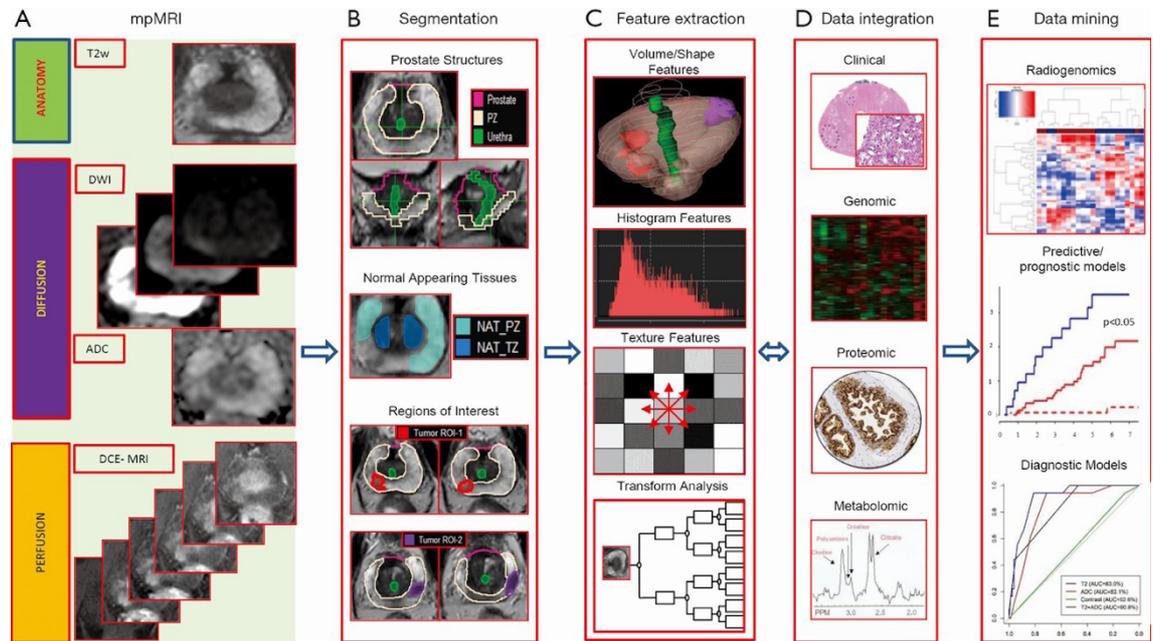
**absolute value of the complex shear modulus  $|G^*|$  measured by MR elastography: a biomarker of tissue integrity.** Parametric maps show a decrease in  $|G^*|$ , 24h following treatment, of a SW620 colon carcinoma tumour xenograft with the vascular-disrupting agent ZD6126, which caused massive central necrosis. **E. Identification of tumours harbouring the  $ALK^{F1174L}$  mutation in  $MYCN$ -driven transgenic mice with Intrinsic Susceptibility (IS)–MRI.** Native  $R_2^*$  maps and  $\Delta R_2^*$  oxygen-air maps obtained following continuous inhalation of 100% oxygen, were able to discriminate between tumours arising in the Th- $ALK^{F1174L}$  /Th- $MYCN$  mouse from tumours arising in the Th- $MYCN$  mouse, due to a differential vascular phenotype characterised by the absence or presence of haemorrhage respectively. [Figure adapted from **A. O'Connor et al** (39), **B. Hill et al** (40), **C. Jamin et al** (41), **D. Li et al** (42) **E. Jamin et al** (43)].

### 1.2.2 Radiomics

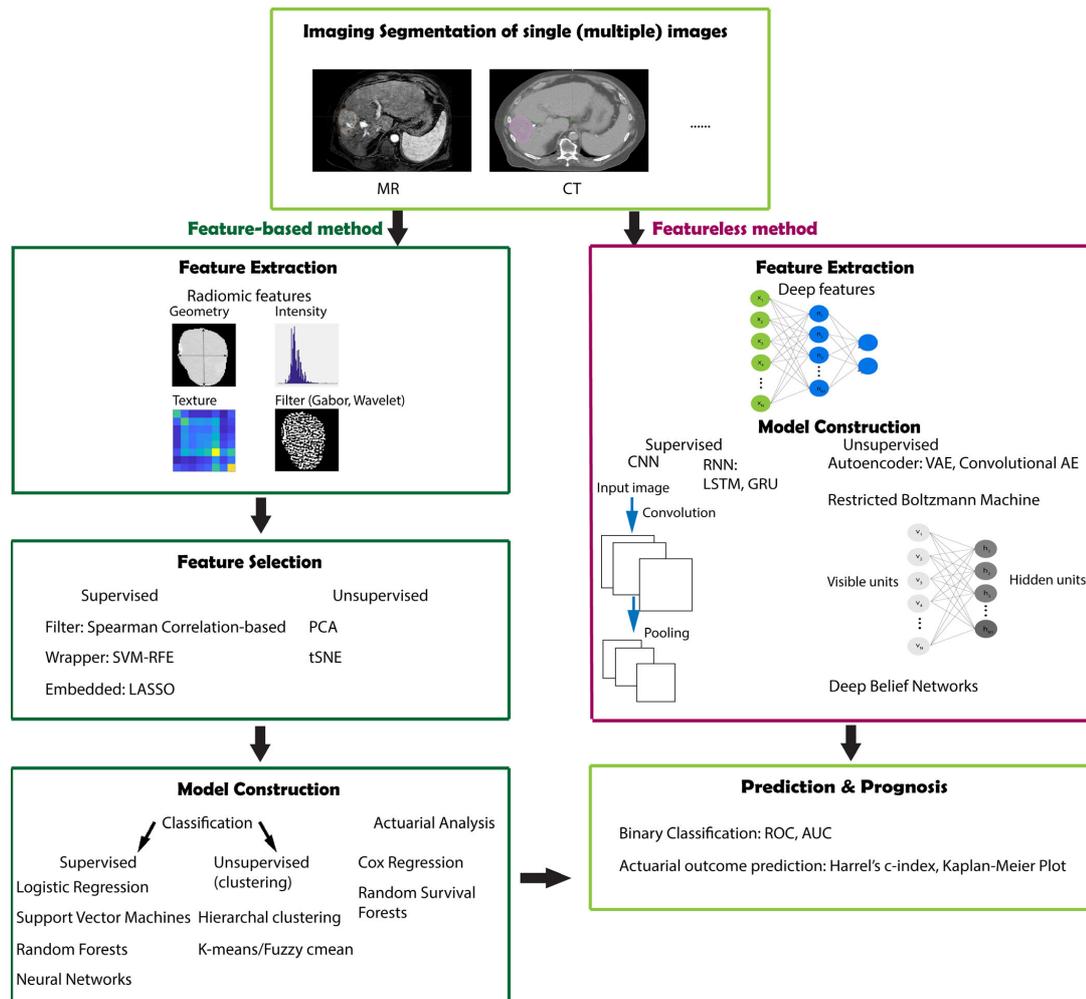
Radiomics refers to the comprehensive quantification of tumour phenotypes by applying a large number of quantitative image features to allow real time data analyses and association of features with prognostic, diagnostic and predictive models (44-48). Image features can also be complemented with other types of data, such as clinical data or genomic information acquired from biopsy samples (“Radiogenomics”) (49-51). Feature extraction from the images can be achieved by using *i*) the so-called “hand-crafted” features, such as first- and second-order statistics and fractal analysis, *ii*) deep learning-based radiomics, where trained deep neural networks automatically extract features and *iii*) hybrid solutions which attempt to exploit the advantages of both approaches (52) (Figure 1.3, Figure 1.4). Radiomics has been shown to be successful in a variety of tasks, such as diagnosis, tumour detection/segmentation/classification, disease staging, survival and recurrence prediction.

However, radiomic features from anatomical images are often abstract and not easily interpretable. Also, deep learning approaches are usually considered as a “black box” which is a major drawback in a clinical setting [although attempts to decipher how neural networks work are being made (53)]. Furthermore, these methods typically fail to satisfactorily generalise and produce the same high-quality results in new independent datasets, especially in data from different centres and scanners. Therefore, all steps of the approach must be easily reproducible, shareable and follow specific standards (44). Finally, the radiology-defined sub-regions (habitats) derived from radiomics are seldom validated with

histopathology, as radiomic techniques typically produce metrics which are not spatially explicit.



**Figure 1.3** A radiomics process pipeline **A.** Acquisition of multi-parametric MR images from standard exams, in this case prostate: T<sub>2</sub>-weighted MRI, diffusion weighted imaging (DWI) and the calculated apparent diffusion coefficient (ADC) maps, dynamic contrast enhanced (DCE)-MRI; **B.** manual or preferably semi-automatic/automatic segmentation of regions of volumes of interest. For the prostate, these volumes are: prostate, peripheral zone (PZ) [and subsequently transition zone (TZ)], urethra, normal appearing tissues in PZ and TZ and tumour regions of interest (ROI); **C.** Extraction and quantification of imaging features, such as intensity volume histogram (first order features), features related to volume/shape, texture features (second order features) and transform analysis features; **D.** radiomic data could then be integrated with clinical, genomic, proteomic and metabolomics data; **E.** the final integrated dataset is mined to develop diagnostic, predictive and prognostic models [Adapted from *Stoyanova et al* (49)].



**Figure 1.4** Workflow for radiomics analysis with “classical” machine learning (left, “hand-crafted” features) and deep learning approaches (right) [Adapted from *Avanzo et al* (54)].

### 1.2.3 Habitat imaging

Habitat imaging approaches typically require the use of computational methods to identify the optimal combination of multiple imaging parameters, all of which are sensitive to different tumour properties, to spatially characterise distinct tumour regions. Subsequently, the habitats are studied to reveal their underlying biology, unique phenotype and genotype and, ultimately, to correlate them with clinical data. Thus, quantitative “habitat” imaging can link the macroscopic spatial heterogeneity with the molecular properties of the tumour and provide new tools for studying the tumour’s ecosystem and inform on clinical decision making, monitor tumour evolution and guide surgical biopsies (2). Examples of various habitat imaging applications are presented in Table 1.2.

Study	Cancer type	Parameters	Habitats	Methods and main findings
Zhou et al (55)	Glioblastoma	MRI: CE-T <sub>1</sub> , T <sub>2</sub> -weighted, FLAIR	Number of habitats: 4 Combinations of "high" and "low" regions Interpretation: Tumour tissue (High CE-T <sub>1</sub> , low FLAIR) Necrosis (Low CE-T <sub>1</sub> , high FLAIR) High cellularity (Low CE-T <sub>1</sub> , low FLAIR)	<ul style="list-style-type: none"> <li>• Each MRI parameter is split into high and low using Otsu thresholding</li> <li>• Radiomic features from the habitats predict patient survival</li> </ul>
Stringfield et al (56)	Glioblastoma multiforme	MRI: CE-T <sub>1</sub> , FLAIR	Number of habitats: 6 Combinations of "high" and "low" regions	<ul style="list-style-type: none"> <li>• CE-T<sub>1</sub> into low, mid, high (Otsu-thresholding), FLAIR higher or lower than white matter (empirical)</li> <li>• Fraction of tumour volume of "habitat 6" (high signal in both CE-T<sub>1</sub> and FLAIR) predictive of survival</li> </ul>
Farhizadeh et al (57, 58)	Soft-tissue sarcoma	MRI: CE-T <sub>1</sub> , T <sub>2</sub>	Number of habitats: 4 Combinations of "high" and "low" regions	<ul style="list-style-type: none"> <li>• Each MRI parameter is split into high and low using Otsu thresholding</li> <li>• Fraction of tumour volume of one habitat (high CE-T<sub>1</sub>, low T<sub>2</sub>) independently predictive of overall and progression-free survival</li> <li>• Observation of evolution during treatment</li> <li>• Habitats could predict the existence of occult metastases</li> </ul>
Carano et al (59)	Human colorectal tumour xenograft mouse model	MRI: ADC, T <sub>2</sub> , Proton density (M <sub>0</sub> )	Number of habitats: 4 Viable tumour, 2 types of necrosis and subcutaneous adipose tissue	<ul style="list-style-type: none"> <li>• K-means clustering with fixed cluster number based on visual histology inspection</li> <li>• Visual histological validation using landmarks.</li> <li>• Observation of evolution during treatment</li> </ul>
Chang et al (60)	Prostate and brain (mouse pre-clinical models), sarcoma and prostate (clinical)	MRI: DCE-MRI	Number of habitats: 3 Well-perfused, hypoxic and non-perfused	<ul style="list-style-type: none"> <li>• Use of simulated data</li> <li>• Statistical learning based on PCA to determine the number of habitats</li> <li>• Spatial deconvolution of habitats from signal vs time curves</li> <li>• Visual histological validation</li> </ul>
Lam et al (61)	human prostate adenocarcinoma tumour xenografts	Saturation transfer MRI, T <sub>1</sub> , T <sub>2</sub>	Number of habitats: 5 Tumour, necrosis/apoptosis, muscle, muscle/connective tissue, blood/oedema	<ul style="list-style-type: none"> <li>• Combination of ICA and GMM to determine the habitats</li> <li>• Visual histological validation</li> <li>• Fraction of necrotic voxels matched scoring from histopathological arrays</li> </ul>
Weigelt et al (62)	One patient of high-grade serous ovarian cancer	MRI: Diffusion coefficient (D), perfusion coefficient (f), DCE-MRI (k-trans) + FDG-PET (SUV)	Number of habitats: 3 (59)	<ul style="list-style-type: none"> <li>• K-means clustering with fixed cluster number</li> <li>• Use of 3D personalised moulds and multi-modal imaging</li> <li>• Identification of distinct regional genomic signatures (growth patterns, hypoxia-related markers, genetic alterations)</li> <li>• Visual histological validation by pathologists (H&amp;E and IHC)</li> </ul>

Wu et al (63)	Breast	DCE-MRI	Number of habitats: 4 3 tumour regions (interpreted as poorly, moderately and highly-perfused) and parenchyma	<ul style="list-style-type: none"> <li>• Spectral clustering</li> <li>• Features from habitats prognostic to recurrence free survival (same features from whole tumour not predictive)</li> <li>• Evidence of hypoxia presence by the interaction of poorly-perfused subregions and parenchyma</li> </ul>
Parra et al (64)	Prostate	DCE-MRI	Number of habitats: 7 Perfusion habitats	<ul style="list-style-type: none"> <li>• Extracting features from the DCE time-activity curves.</li> <li>• Features related to clinical significance disease</li> </ul>
Jardim-Perassi et al (65)	Breast cancer mouse model	MRI: T <sub>2</sub> , T <sub>2</sub> <sup>*</sup> , ADC, DCE-MRI	Number of habitats: 4 viable-normoxic, viable-hypoxic, nonviable hypoxic, nonviable-normoxic	<ul style="list-style-type: none"> <li>• GMM clustering</li> <li>• Quantitative validation from co-registered histology (H&amp;E and IHC)</li> </ul>
O'Connor et al (66, 67)	pre-clinical cancer models, xenografts, and non-small lung cancer patients	MRI: OE-MRI, DCE-MRI	Number of habitats: 3 perfused oxygen-enhanced, perfused-oxygen refractory (hypoxia), non-perfused (necrosis)	<ul style="list-style-type: none"> <li>• Changes in R<sub>1</sub> after oxygen-challenge to directly measure hypoxia</li> </ul>
Panagiotaki et al (68)	Prostate (primarily) + colorectal, breast, brain	VERDICT-MRI	3 compartment model: Characterises water diffusion in the vascular, extracellular-extravascular space (EES) and intracellular (IC) tumour compartments	<ul style="list-style-type: none"> <li>• VERDICT couples DW-MRI to a mathematical model of tumour tissue to access features such as cell size, vascular volume fraction, intra- and extracellular volume fractions, and pseudo-diffusivity associated with blood flow</li> <li>• Histological validation using 3D personalised moulds</li> </ul>
Blackledge et al (69)	Soft-tissue sarcoma	MRI: ADC, Dixon (fat-fraction), pre- and post Gd T1w	Number of habitats: 4 fatty tissue, necrotic tissue, cellular vascularised tumour, poorly vascularised tissue, (class 5: none of this)	<ul style="list-style-type: none"> <li>• Machine learning supervised classification of habitats</li> <li>• Training data provided by radiologists by drawing ROIs on the images</li> <li>• Monitor changes over radiotherapy treatment</li> </ul>

**Table 1.2** Applications of habitat imaging

## **1.3 Functional imaging-guided delivery of precision medicine for children with neuroblastoma.**

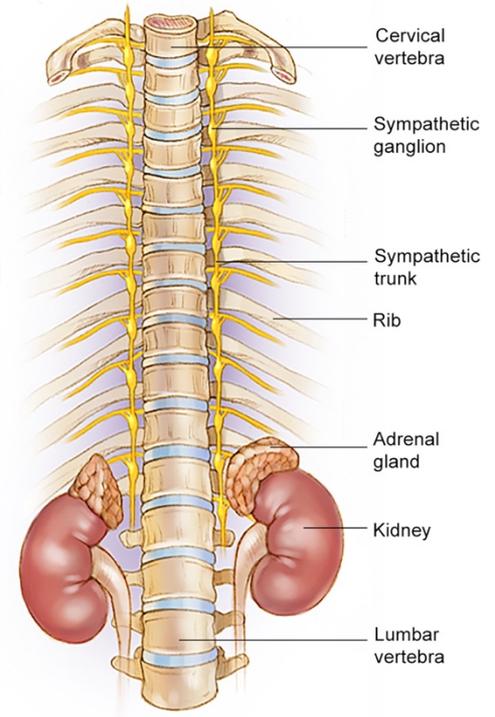
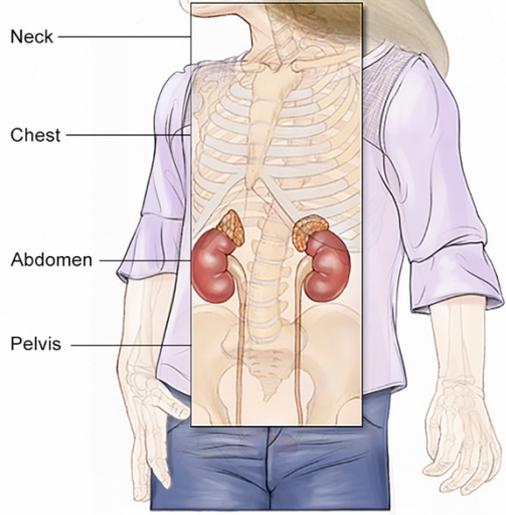
### *1.3.1 Neuroblastoma*

Peripheral neuroblastic tumours are the most common extracranial solid tumours in childhood and the most diagnosed malignancy during infancy (< 1 year old). These tumours arise from anomalies in the embryonic sympatho-adrenal lineage of the neural crest in young children and, thus, can develop anywhere along the sympathetic nervous system and often present as large asymptomatic mass arising from the adrenal medulla (30%), periadrenal and paraspinal sympathetic ganglia (70) (Figure 1.5). Bone and bone marrow metastatic disease is present in 50% of cases (Figure 1.6). Neuroblastic tumours are histologically classified as neuroblastoma (NB), ganglioneuroma (GN) and ganglioneuroblastoma (GNB) (71).

Risk stratification is based on age (1-year-old cut-off), radiological and histological presentation as well as certain molecular features (e.g *MYCN* amplification) (Figure 1.5, Figure 1.6, Figure 1.7).

NB still accounts for 13% of all cancers-related deaths in children. NB is a heterogeneous and pathology-rich disease with a very wide range of clinical outcomes, including the highest rate of spontaneous regression (Stage IVs). Yet half of the children diagnosed with NB, are at a high-risk to develop uncontrollable refractory or relapsing disease, despite the intensity of multimodal therapy, which includes intensive and toxic chemotherapy followed by surgical resection, myeloablation and autologous stem cell rescue, radiation, and intensive biologic/immunotherapy. For the 60% of children with relapsing/refractory disease, the 10-year survival rate remains below 10%. All neuroblastoma survivors are also at high-risk to suffer from long-term disabilities and life-threatening conditions caused by the current anti-cancer therapies (72). There is a crucial need for more efficacious and safer therapies for children with NB.

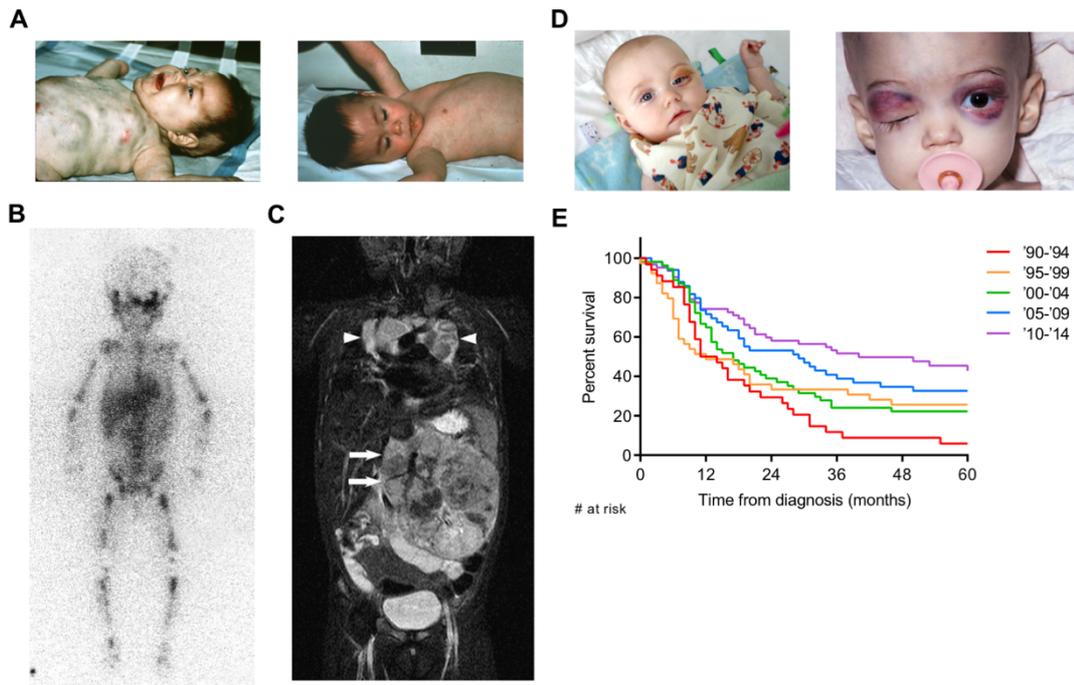
**Primary distribution of neuroblastomas in children**



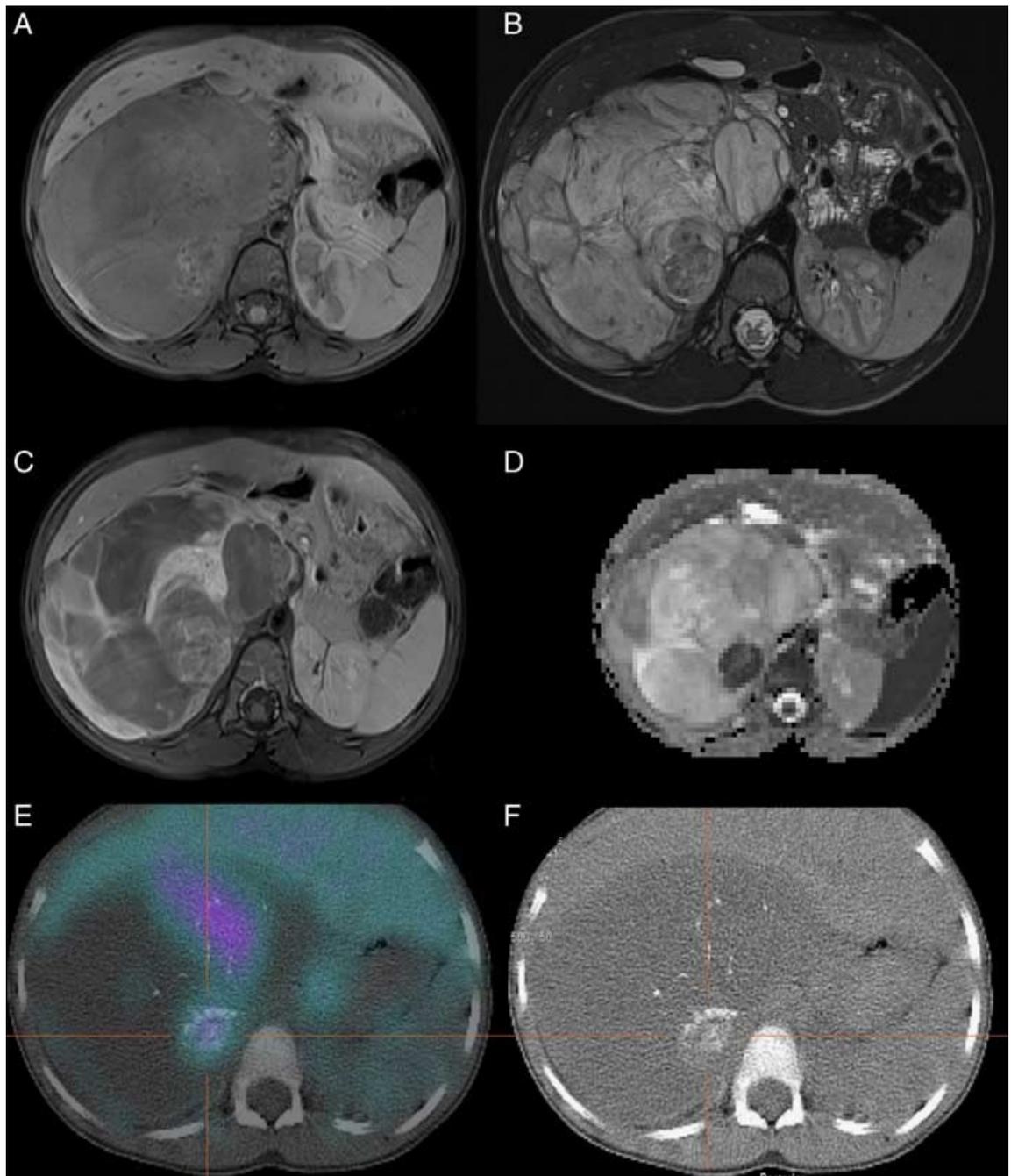
Robert Morreale/Visual Explanations, LLC

© 2005 American Society of Clinical Oncology

**Figure 1.5** Primary distribution of neuroblastoma in children, a tumour of the sympathetic nervous system. <https://www.cancer.net/cancer-types/neuroblastoma-childhood/medical-illustrations>.



**Figure 1.6 Childhood neuroblastoma** **A.** picture illustrates a child <1 year old with stage IVs neuroblastoma with extensive metastatic disease, which spontaneously regressed over a period of 6 months without any therapeutic intervention. **B.**  $^{123}\text{I}$ -meta-iodobenzylguanidine [ $^{123}\text{I}$ -MIBG] scan showing extensive bone and bone marrow metastasis. **C.** Whole-body MRI showing a large primary abdominal tumour distorting the abdominal aorta and a secondary tumour in the upper thorax (73). **D.** Children with neuroblastoma presenting with typical orbital metastasis **E.** Overall survival rates of children with high-risk neuroblastoma from the time of diagnosis (74).



**Figure 1.7** Multi-modal imaging management of children with neuroblastoma. As neuroblastoma is often asymptomatic children usually develop large abdominal tumours detectable by physical palpation usually confirmed by ultrasonography before the child is transferred to a specialised oncology centre. Example of Magnetic resonance imaging (MRI) **A.** axial T<sub>1</sub>-weighted **B.** axial T<sub>2</sub>-weighted showing large abdominal mass **C.** axial T<sub>1</sub>-weighted post-contrast images showing heterogeneous tumour enhancement **D.** apparent diffusion coefficient (ADC) map, **E.** fused <sup>123</sup>I-MIBG scan with SPECT-CT and to **F.** area of calcification on computerized tomography scan. This child presented with a rare case of ganglioneuroblastoma nodular characterised by a 1cm “circular” neuroblastoma nodule in the context of a large ganglioneuroma mass. Note that the neuroblastic lesion showed area of restricted diffusion, MIBG positivity and calcification consistent with the phenotype of undifferentiated neuroblastoma [Adapted from Jain et al (75)].

### 1.3.1.1 Histological characteristics of peripheral neuroblastic tumours

The three types of peripheral neuroblastic tumours are defined by the level of cellular and extracellular maturation. The primordial neural crest cells may remain undifferentiated (neuroblasts) or they may mature (ganglion and Schwann cells). Neuroblasts are small, rounded in contour, show little cytoplasm, and have a darker nuclei and smaller indistinct nucleoli (Figure 1.8), while ganglion cells are fully mature cells with abundant cytoplasm, rounded contour, and large nuclei with distinct and prominent nucleoli. NB is a tumour with malignant potential which primarily comprises by neuroblasts, while GN is composed entirely by mature ganglion cells and other mature tissue and is considered benign. GNB consists of both immature and mature cell types and has intermediate malignant potential (76, 77).

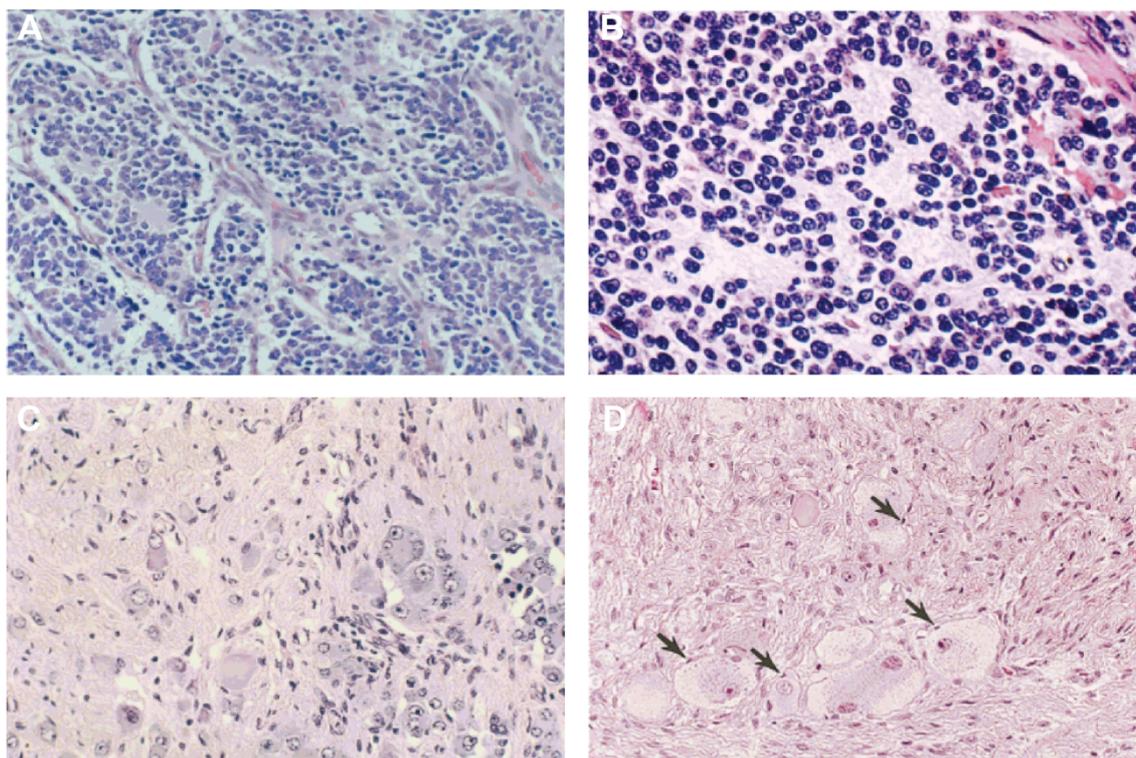
Tumours are classified into favourable or unfavourable prognosis based on their level of cellular differentiation. GN presents with completely (100%) differentiated cellular (and stromal) components, GNB is intermediate (greater than 50% differentiated cells), and NB contains less than 50% differentiated elements. Neuroblasts can present with characteristics of differentiation into ganglion cells, such as nuclear enlargement, nucleoli, cytoplasmic eosinophilia and enlargement, distinct cytoplasmic border, and cell processes (78). Thus, NB can be further sub-classified into undifferentiated (<5% differentiating neuroblasts), poorly differentiated, and differentiating.

Stroma is the tissue surrounding the neuroblasts or/and the ganglion cells. It may consist of Schwann cells, fibrovascular septa, or both (Figure 1.8). Tumours are stratified into stroma-rich and stroma-poor with more mature tumours being stroma-rich. The stroma-rich group is further sub-classified as well-differentiated if it has only mature tissue, intermixed if mature tissue contains random clusters of neuroblasts, and nodular if showing one or a few clearly defined stroma-poor areas trapped in a mature stroma (77).

Other histologic features of interest are the neuropil, which is the fine pattern of cellular (neuritic) processes extending from neuroblasts, formation of Homer

Wright rosettes, which are circular or ovoid columns of tumour cells arranged around a central core of neuropil, as well as necrosis, mitosis, haemorrhage, fibrosis, calcification, lymphocytic infiltrate, and karyorrhexis. Karyorrhexis is the fragmentation of a nucleus into scattered pieces within the cytoplasm, which is an event that occurs during cell death. The mitosis-to-karyorrhexis (MKI) index is defined as the ratio of cells in mitosis and karyorrhexis (apoptosis/5000 neuroblastic cells). A high MKI is associated with unfavourable histology and prognosis (76, 79).

Age of diagnosis is also an important factor in risk stratification. A summary of favourable and unfavourable histological characteristics (Shimada classification system) is presented in Table 1.3.



**Figure 1.8 Histological characteristics of neuroblastic tumours. A-B.** Immature neuroblastic tissue. Neuroblastoma (Schwannian stroma-poor), poorly differentiated subtype, composed of undifferentiated neuroblastic cells with clearly recognisable neuropil. (**A.** original magnification x100; **B.** x200). **C.** Ganglioneuroma (Schwannian stroma-dominant), maturing subtype (original magnification x100). Note that there are some neuroblastic cells that have still not reached full maturation to ganglion cells. Those neuroblastic and ganglion cells are individually scattered and have not formed nests. **D.** Ganglioneuroma (Schwannian stroma-dominant), mature subtype (original magnification x200). Fully mature ganglion cells (arrows) with satellite cells are individually embedded in Schwannian stroma. [Adapted from *Shimada et al* (80)].

Type	Favourable histology	Unfavourable histology
<b>Stroma-rich</b>	Well-differentiated Intermixed	Nodular
<b>Stroma-poor (Age &lt; 18 months)</b>	MKI < 200/5000	MKI > 200/5000
<b>Stroma-poor (Age 18-60 months)</b>	MKI < 100/5000 and differentiation	MKI > 100/5000 or undifferentiated
<b>Age &gt; 5 years</b>	None	All

**Table 1.3** Summary of favourable and unfavourable histological characteristics (Shimada classification system) [Adapted from *Shimada et al* (77)] .

### 1.3.2 Targeted therapies for high-risk refractory and relapse neuroblastoma

Large scale genomics studies have begun to map the landscape of high-risk neuroblastoma, identifying at least four distinct molecularly-defined sub-cohorts (MYCN-TERT, ATRX-ALT, RAS-TP53, ALK) (81-86). Stratified precision medicine approaches for refractory/relapsed neuroblastoma aim to integrate both genomic and biological information to incorporate both novel targeted drugs and immunotherapeutics into treatment schedules.

#### 1.3.2.1 Targeting MYCN

Amplification of the proto-oncogene *MYCN*, is the most established marker of poor prognosis and found in approximately 40% of high-risk disease. *MYCN* is a transcription factor that plays a central role in the biology of high-risk neuroblastoma and as such represents a major, yet elusive, therapeutic target (87). Indirect targeting of this transcription factor through either destabilization of the protein via the selective targeting of Aurora A kinase or targeting of mTOR activity which regulates the major remodelling of metabolic, angiogenic and translational pathways in order to cope with *MYCN* addiction, have shown promising results in preclinical trials and are currently being evaluated in early-phase paediatric clinical trials.

#### 1.3.2.2 The BEACON Neuroblastoma trial

The BEACON-Neuroblastoma Trial is a randomised phase IIb trial of bevacizumab added to temozolomide  $\pm$  irinotecan for children with refractory/relapsed neuroblastoma that started in 2013. One of the aims of the trial design was to evaluate Temozolomide as a chemotherapy backbone for the introduction of novel targeted therapeutics in clinical trials in children with relapse and refractory neuroblastoma. Neuroblastoma tumours are generally hypervascular and vascular-targeted treatments have shown promising results in preclinical studies. Having established its safety profile with manageable toxicity (side effects) in adult cancer (a requirement for a drug to enter early phase trial in children), in addition to showing a survival advantage, the BEACON

Neuroblastoma trial aimed to evaluate the efficacy of Bevacizumab, a monoclonal antibody against vascular endothelial growth factor (VEGF).

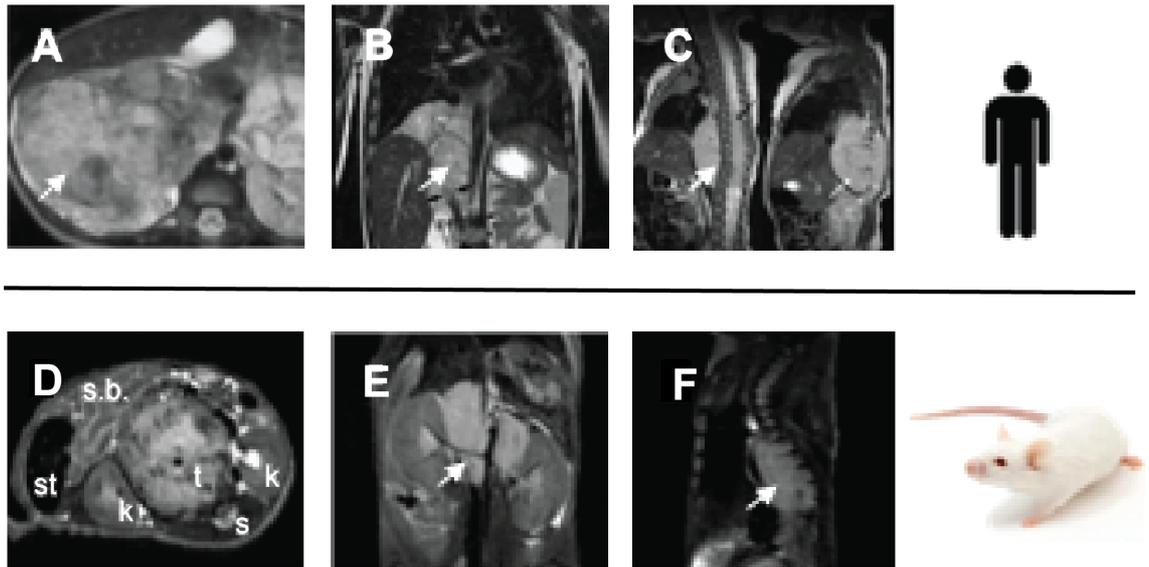
#### 1.3.2.3 The mouse hospital and co-clinical trial concept.

The application of the mouse hospital and co-clinical trial concept, where *in vivo* preclinical and early clinical studies are closely aligned, represents a clear paradigm shift in neuroblastoma translational research (88, 89). This approach integrates more advanced mouse modelling, including genetically-engineered mouse (GEM) models, such as the Th-*MYCN* mouse (90), to accelerate the discovery and evaluation of novel therapeutic strategies, and helps shape the clinical trial pipeline priorities for children with high-risk disease.

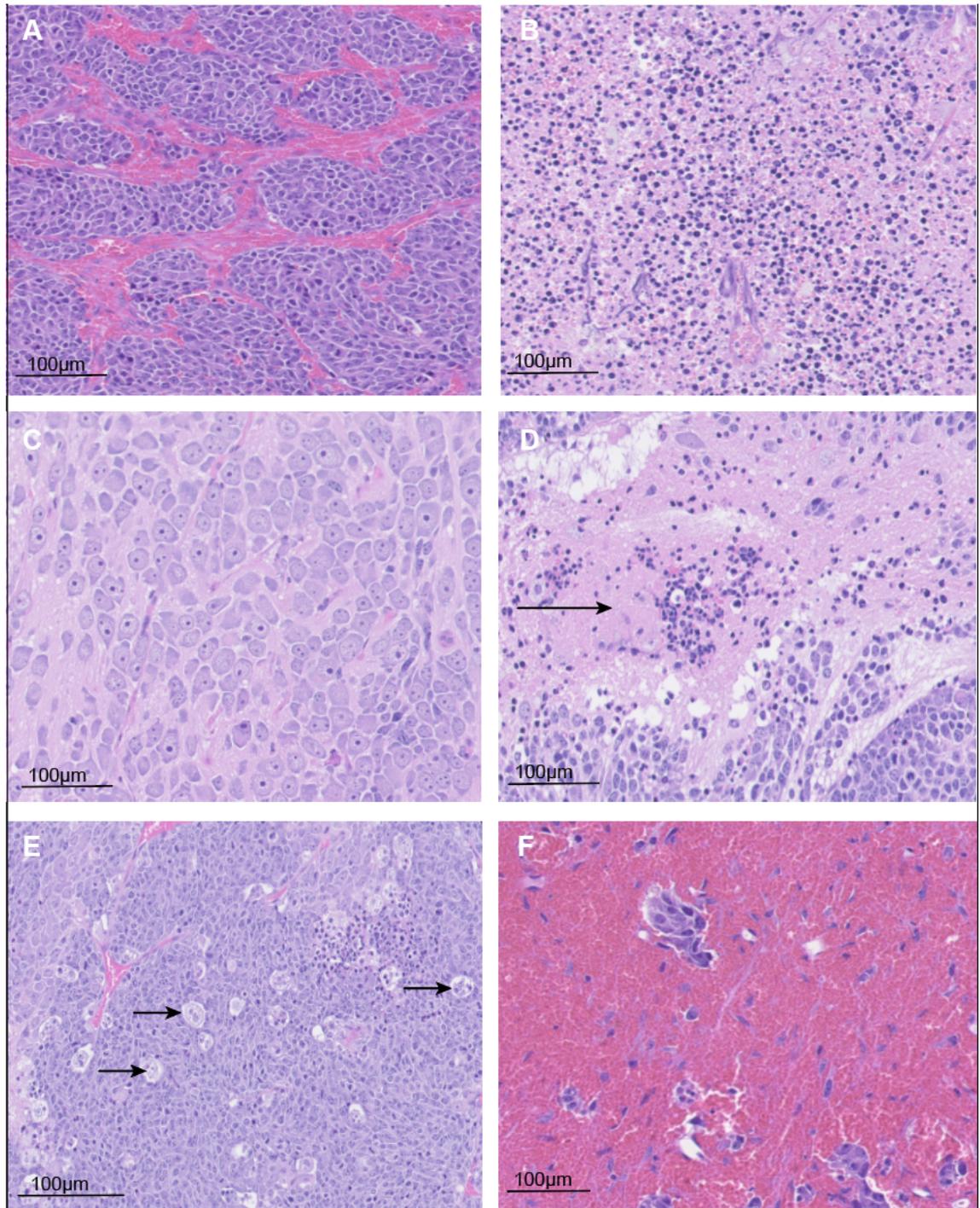
##### 1.3.2.3.1 Genetically engineered murine (GEM) models of neuroblastoma: an information-rich experimental platform

As for most paediatric malignancies, NB is only driven by a small number of genetic aberrations, which makes it amenable to genetically engineered mouse (GEM) modelling approaches. The Th-*MYCN* mouse, the most established GEM model of NB which was generated by targeting *MYCN* expression to the neural crest of mice, is an immunocompetent mouse, which develops spontaneous abdominal tumours, which mirror the major pathophysiological, genetic and radiological features of high-risk *MYCN*-amplified childhood NB (91) (Figure 1.9, Figure 1.10). The Th-*MYCN* mice predominantly present with stroma-poor, undifferentiated or poorly differentiated, “immune cold” tumours with high MKI and high proliferation.

Mutations in the anaplastic lymphoma kinase (ALK) gene are other genomic aberrations present in 10% of children with neuroblastoma that lead to marked poorer prognosis (92). In contrast to *MYCN*, promising ALK inhibitors exist and are currently under evaluation in the neuroblastoma clinic (93-95). Co-expression of *ALK*<sup>F1174</sup> mutation that lead to the constitutive activation of ALK kinase activity with *MYCN* lead to tumours arising early, a different phenotype and clinically relevant resistance to first-generation ALK inhibitor Crizotinib in the Th-*ALK*<sup>F1174L</sup>/*MYCN* GEM models.



**Figure 1.9 Childhood neuroblastoma.** Anatomical **A.** transverse and **B.** coronal T<sub>2</sub>-weighted MRI image from children with abdominal neuroblastoma. **C.** Sagittal T<sub>2</sub>-weighted MRI showing a case of thoracic neuroblastoma growing around the spine with local invasion. **D, E, F** show corresponding representative images of tumours arising in the Th-MYCN mouse. Tumours cause displacement of major abdominal organs such as the kidneys (k), small-bowel (s.b.) and the displacement of the abdominal aorta away from the spine (s) [Clinical images adapted from *Elsayes et al* (96), *Goo et al* (73), and *Nour-Eldin et al* (97)].



**Figure 1.10 Histological Profile of Tumours from *MYCN* Transgenic Mice.** Haematoxylin and eosin (H&E) staining of tumour arising from the Th-*MYCN* genetically engineered model of neuroblastoma. Overall, tumours resembled high-risk childhood neuroblastoma **A**. Islands of undifferentiated neuroblastoma cancer cells separated by fibrous septa **B**. Apoptotic cells distinguishable by the fragmentation of nuclear material and condensation of chromatin within them. **C**. Islands of larger differentiated neuroblastoma cells **D**. Neuropil (arrowed) among undifferentiated tumour **E**. Tingible body macrophages (arrows) **F**. Region of high haemorrhage in Th-*MYCN* GEM model

### 1.3.3 *Biomarkers-guided delivery of precision medicine*

#### 1.3.3.1 Need for non-invasive biomarkers

Biomarkers, defined as a “defined characteristic that is measured as an indicator of normal biological processes or responses to an exposure or intervention, including therapeutic interventions” (98, 99) have contributed to the success of stratified precision medicine in adult cancer, especially by accelerating the clinical development of novel therapeutics. Conventional pharmacodynamic biomarkers necessitate access to post-therapy surgical biopsies, which is often not feasible in children with refractory/relapsed disease. Therefore, non-invasive alternatives, such as imaging biomarkers or liquid biopsies must be pursued (100, 101).

#### 1.3.3.2 Current radiological management of children with neuroblastoma

Evaluation of response to treatment in children with neuroblastoma is based on Response Evaluation Criteria in Solid Tumours (RECIST) using non-invasive anatomical imaging such as computed tomography (CT) or magnetic resonance imaging (MRI). The revised International Neuroblastoma Response Criteria (INRC) guidelines now also include sensitive nuclear medicine-based functional imaging approaches (such as <sup>123</sup>I-meta-iodobenzylguanidine [<sup>123</sup>I-MIBG] scans and [<sup>18</sup>F]-2-fluoro-2-deoxyglucose positron emission tomography/CT [FDG PET/CT]) for the assessment of bone and bone marrow metastatic disease, present in 50% of cases (102). <sup>123</sup>I-meta-iodobenzylguanidine [<sup>123</sup>I-MIBG] scans are one a few examples of success for the field of targeted molecular imaging/theranostics. MIBG is a substrate for the neuro-epinephrine (NET) transporter, which is selectively expressed on the membrane of neuroblastoma cells in 90% of patients. [<sup>131</sup>I]-MIBG is successfully used to target radiation to neuroblastoma cells and is an integral part of frontline therapy against the disease.

#### 1.3.3.3 MRI-based functional imaging for neuroblastoma

MRI is becoming the preferred clinical imaging technique for the management of children with neuroblastoma because of its exquisite soft tissue contrast. MRI provides excellent anatomical information at diagnosis and follow up while sparing children of the exposure to ionising radiation associated with CT (103). Advanced MRI-based functional imaging techniques could potentially be used to define quantitative imaging biomarkers that inform on biologically relevant structure-function relationships (see Section 1.2.1) in neuroblastoma and its microenvironment.

Diffusion MRI is an example of advanced MRI scans that is increasingly being adopted in the neuroblastoma diagnostic clinic with the calculated apparent diffusion coefficient (ADC) being able to distinguish stroma-rich GN and GNB from stroma-poor neuroblastoma (104). The BEACON Neuroblastoma trial also has a small functional imaging arm aiming to evaluate the feasibility of diffusion-weighted imaging (DWI) and dynamic contrast enhanced(DCE-) MRI to monitor treatment response to Bevacuzimab.

#### 1.3.3.4 Validation of imaging biomarkers

As emphasised in the CR-UK and EORTC imaging biomarker roadmap (99), imaging biomarkers must undergo stringent validation before they can be deployed in the clinic. Early imaging biomarker development demands close imaging-pathology correlation, to understand the biological processes underpinning the imaging measurement, which in the paediatric population can only be meaningfully studied using animal models (105, 106).

#### **1.4 The mouse hospital as a robust platform for the validation of imaging biomarkers of neuroblastoma.**

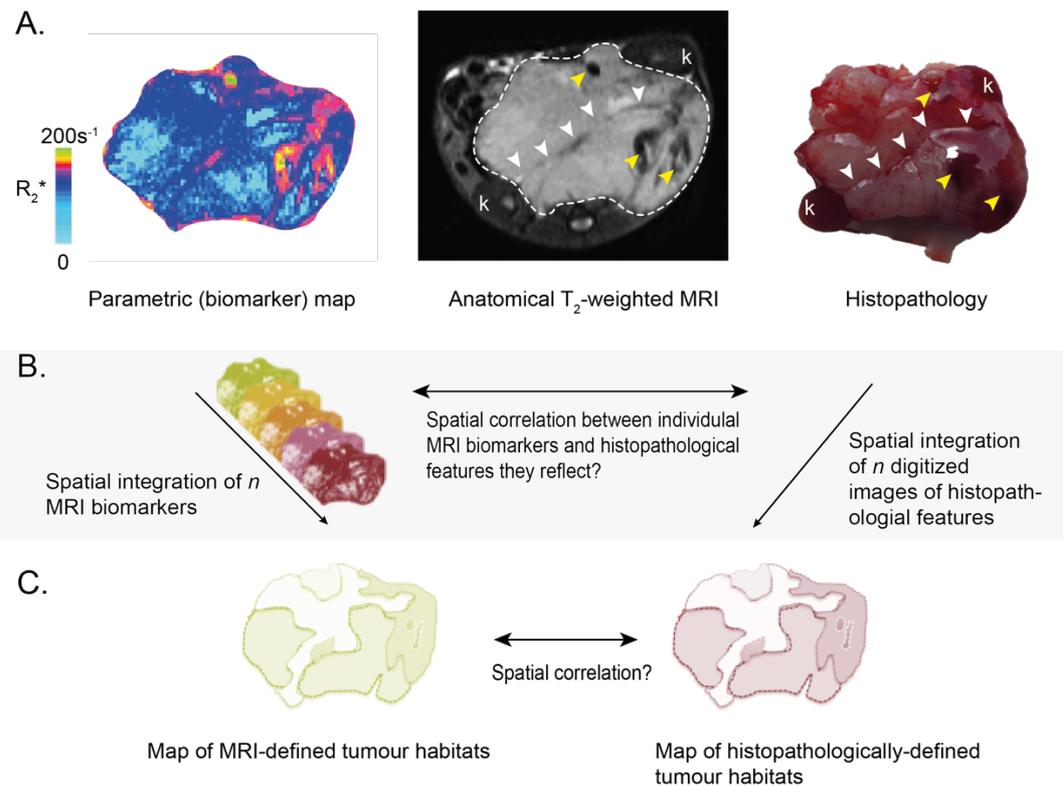
The overarching hypothesis of this project is that the mouse hospital concept provides a unique opportunity to evaluate predictive and prognostic imaging biomarkers of response in neuroblastoma and to perform the close imaging-pathology correlation necessary to understand the biological processes underpinning the imaging measurement and provide the stringent validation needed before they can be deployed clinically.

#### **1.5 Aim of the project**

The project aims to develop computational methodologies for the robust spatial comparison of functional MRI techniques and biomarkers with histopathology to non-invasively map the histopathological hallmarks of high-risk childhood neuroblastoma, and their evolution during treatment, in transgenic *MYCN*-driven mouse models of the disease.

The specific objectives in this project are to develop imaging analysis and spatial statistics tools (Figure 1.11):

1. To characterise of neuroblastoma and its microenvironment in whole-slide digital histology images at both regional (tissue) and cellular level.
2. To evaluate the spatial correlation between the parametric maps of each selected MRI biomarker and their associated histopathological correlates.
3. To spatially integrate the parametric MRI maps to identify radiologically-defined sub-regions (habitats) and evaluate their spatial correlation with histopathologically-defined sub-regions.



**Figure 1.11** Developing radiologic pathology for the non-invasive characterisation of tumour heterogeneity: overview of the methodology. **A.** The primary abdominal tumours in the GEM model of neuroblastoma present with numerous anatomical landmarks (arrows, k, kidney) that are detectable by conventional anatomical MRI, which can be used to ensure precise registration between the MRI-derived parametric maps (example:  $R_2^*$ , sensitive to deoxyhaemoglobin concentration and thus a proxy for haemodynamic vasculature) and the histopathology slices, which is crucial for validation and calibration of the methodology. Imaging analysis and spatial statistics tools will be developed to **B.** evaluate the spatial correlation between the parametric maps of each selected MRI biomarker and the associated histopathological features and **C.** to integrate the different MRI biomarker maps into a spatial map of sub-regions (habitats) and evaluate their spatial correlation with the distribution of sub-regions defined by histopathology.

## 1.6 Structure of the thesis

**Chapter 2** describes the development of deep learning and classical machine learning methodologies and strategies for the characterisation of tumour and its microenvironment in whole-slide digital histology images at both regional (tissue) and cellular level.

**Chapter 3** describes the implementation of a pipeline for the registration of MRI-derived and histology-derived parametric maps, essential for the robust spatial histopathological validation of MRI biomarkers.

The MRI-pathology pipeline was incorporated in *in vivo* studies in the Th-*MYCN* mouse model to:

- a. Demonstrate that Susceptibility-MRI, and the  $R_2^*$  and fractional blood volume (*fBV*) biomarkers, can quantify and map the characteristic haemorrhagic and hypervascular nature of high-risk neuroblastoma and predict response to promising anti-angiogenic treatment (**Chapter 4**).
- b. Demonstrate that  $T_1$ -mapping is sensitive to neuroblastoma histological heterogeneity and can be used to monitor active disease and detect treatment response to two promising *MYCN*-targeted therapeutic strategies (**Chapter 5**).
- c. Explore the pathological determinants of the apparent diffusion coefficient (ADC) measured by diffusion-weighted MRI and its ability to provide information on treatment response in high-risk undifferentiated or poorly differentiated neuroblastoma (**Chapter 6**).
- d. Validate the combined use of the MRI biomarkers validated in the previous chapters to fully characterise radiologically-defined sub-regions (habitats) constituting the rich pathology of neuroblastoma (**Chapter 7**).

## **1.7 Contributions to the work presented in this thesis.**

All *in vivo* MRI studies presented in this thesis were performed by Dr. Yann Jamin as a part of his Children with Cancer Research UK fellowship. Dr. Henrik Failmezger contributed to the development of the superpixel-based conditional random field (SuperCRF) presented in Chapter 2, by implementing the graphical model part of the method.

## **Chapter 2 : Development of computational pathology methodologies in whole-slide digital histology images**

### **2.1 The new era of digital pathology**

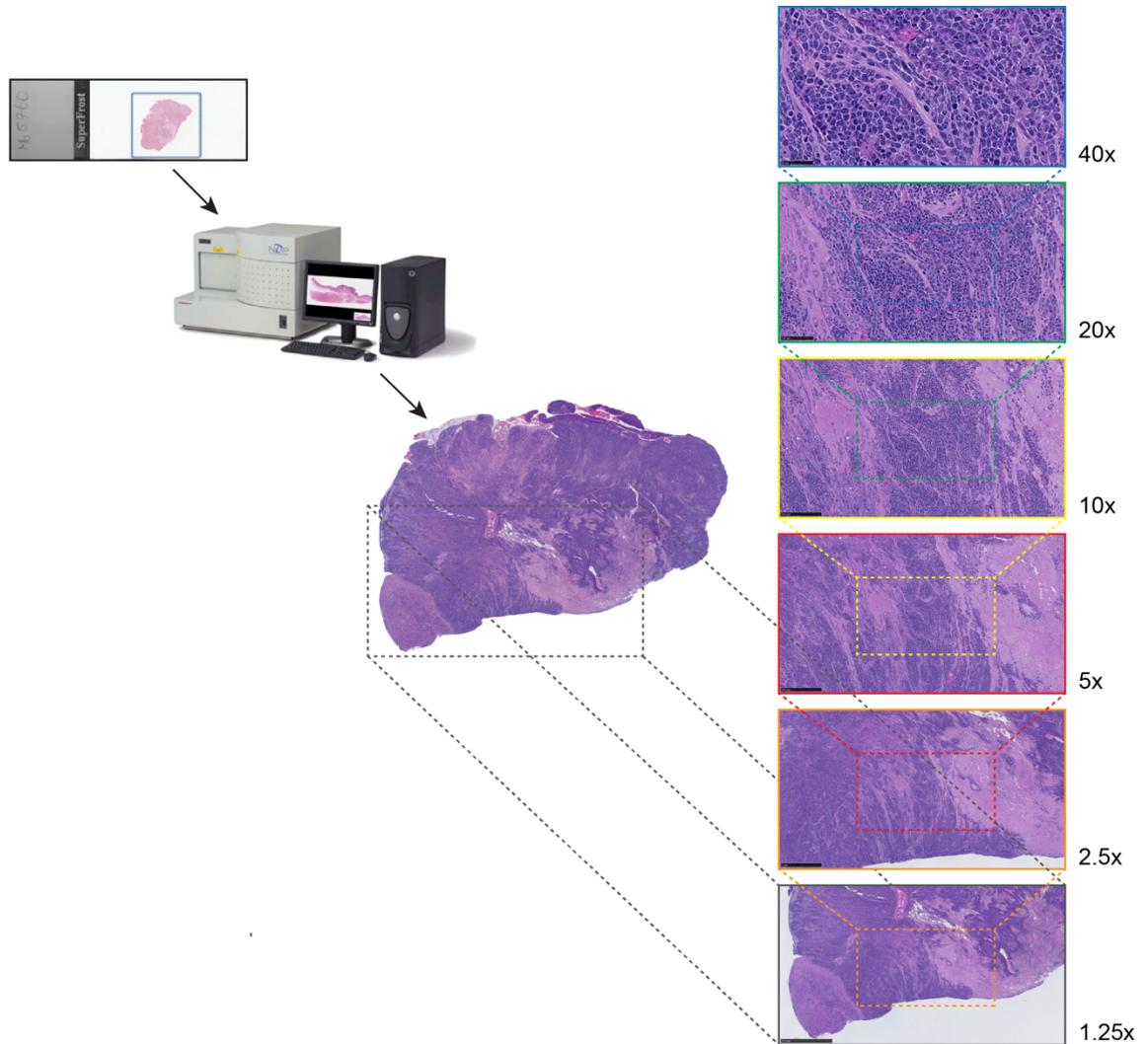
#### *2.1.1 Computer-aided diagnosis*

Cancer is a highly complex, non-autonomous disease. The interactions between micro-environmental selective pressures and cancer cells dictate how cancer progresses and evolves. Accurate and spatially explicit characterization of the tumour micro-environmental landscape, including how cancer cells interact with the extra-cellular matrix and other cellular players such as stromal cells and immune cells within the tumoural niche, is needed to understand the context in which cancer evolves, and may also provide robust predictor of cancer behaviour for risk-stratification (107).

The analysis of histopathological images of surgical tissue specimens, under the microscope, from pathologists is the “gold-standard” for the routine management of patients with cancer and the evaluation of new therapeutic strategies in clinical trials (107-111). Histology images can be stained with various dyes, such as haematoxylin and eosin (H&E) to highlight the cell nuclei from the cytoplasm and reveal the tissue sample’s structure, and immunohistochemistry (IHC) markers to unveil the tissue’s molecular properties. The most widely used (and low-cost) staining method for diagnosis remains the H&E staining. Pathologists, however, can only conduct a qualitative or semi-quantitative analysis to the samples, which is also a time-consuming and cumbersome process. Furthermore, factors such as inter- and intra-observer variability, or fatigue, can affect diagnosis accuracy and therefore treatment planning (112-115).

The relatively recent development of digital scanners has allowed for the glass slides to be digitised at different magnification levels and stored. These whole-slide images (WSI) can then be viewed on any computer (Figure 2.1). In this new era of digital pathology, advanced computational image analysis techniques are

revolutionizing the field of histopathology by providing objective, robust and reproducible quantification of tumour components, thereby assisting pathologists in tasks such as tumour identification and tumour grading (116, 117).



**Figure 2.1** Digitisation of a whole-slide image. The glass-slide is scanned by a digital slide scanner and the images are stored in a pyramid-like structure containing different magnification levels. At the “cell-level” (40x-20x) information about cell distribution and morphology are derived, while at the “tissue-level” (10x-1.25x) the different tissue structures are visible.

### 2.1.2 Machine learning

Image analysis and computer vision algorithms used in digital pathology are primarily focused around machine learning, which is the process of “training” a computer program to make data-driven decisions on its own. The main aim is to

create a model that recognises patterns and makes accurate predictions in the training data and then applies this learned “expertise” successfully in new, unseen data.

#### 2.1.2.1 Machine learning methods

There are different methods of training machine learning algorithms based on how the training data are utilised (118-120). The most common approach, called *supervised learning*, is when each piece of input data has a corresponding label, and the algorithm is trained to predict the label of a new, previously unseen sample. Classification, regression and segmentation tasks typically demand a supervised learning approach. In *unsupervised learning*, the computer is learning to discover patterns in unlabelled input data, clustering and dimensionality reduction being two characteristic examples. Sub-categories include *semi-supervised learning*, which combines a few labelled and many unlabelled data, and *multiple-instance learning*, in which the labels are “weaker” as individual examples are unlabelled and “bags” or groups of instances are labelled instead. Another method which is becoming increasingly popular and has recently shown immense potential (121) is *reinforcement learning*, where “agents” operate with their environment and learn through trial and error and a reward system.

#### 2.1.2.2 Machine learning approaches

For many years, the typical or “classical” way of designing a digital pathology machine learning pipeline was to use mathematical models to extract *features* from an image, such as first- and second-order statistics (eg. texture), then use a *classifier* in the multi-dimensional feature space to separate them into the desired classes. These features are called “hand-crafted”, since they are manually designed.

Some of the most popular classical machine learning algorithms that are based on supervised learning and are widely used in medical image analysis are support vector machines (SVM) (122, 123), which try to maximise the classification margin (distance) in the feature space using the so-called support vectors, and

classification and regression decision trees (CART), which are powerful predictive models while also being easily interpretable. Other advanced ensemble methods are based on decision trees, such as the random forest algorithm, which was created to achieve better generalisation in unseen data. Popular unsupervised learning algorithms for data clustering include the k-means algorithm, which categorises samples based on their distance from the cluster with the nearest mean and the Gaussian mixture modelling (GMM), which is a probabilistic model to identify sub-populations by fitting Gaussian distributions in the overall data. One of the most popular approaches for dimensionality reduction remains the principal component analysis (PCA), which applies a linear transformation to extract features that captures the maximum variance of the data.

#### 2.1.2.3 Deep learning

Deep learning (124, 125) is a branch of machine learning that has recently received much attention as in most tasks deep neural network approaches surpassed the previous state-of-the-art methods by a wide margin, while reaching or even surpassing human performance in certain cases. One of the most characteristic examples is the computer-vision challenge ImageNet, where machine learning algorithms are given natural images and are asked to classify them (e.g. “car”, “dog” etc). In 2012, the deep neural network “AlexNet” won the challenge with a 10% margin from the second place, which improved by another 10% the next year, leading to today’s performance of over 98% accuracy which is higher than human performance (~95-96%) in the same dataset (120, 126, 127).

Artificial neural networks (ANN) are actually one of the first and most well-studied machine learning algorithms, introduced in the 1950s (128) and are inspired by biological neurons. Neural networks consist of a number of connected computational units, the *neurons* or *perceptrons*, arranged in layers. Each neuron is comprised by a set of *weights* which represent its relative importance and a *bias* term. The first layer where the data are fed in the network is called the *input* layer, followed by one or more *hidden* layers, which transform the data as they

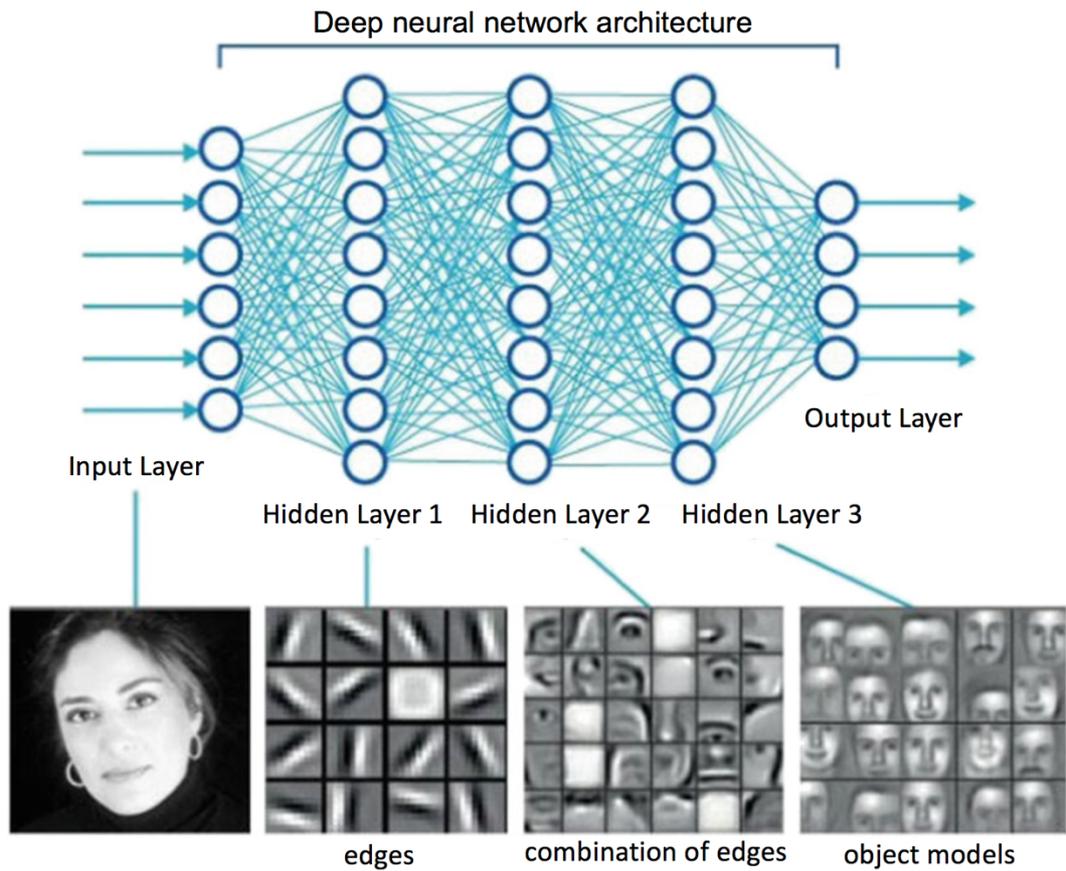
pass through, and ending with the *output* layer that gives the network's predictions. An activation function is attached to each layer, which transforms the output of the previous layer. The aim is to train a model to learn useful representations of the data by optimising an objective function and updating its weights in each iteration. Typically, the networks were "shallower" consisting of an input layer, a hidden layer and an output layer. However, ANN were largely abandoned, due to the high computational cost associated with the increased complexity of network architectures, as well as technical issues which made their training challenging (e.g. the "vanishing gradient" problem, when the amount of error or "gradient" being distributed to all nodes is too small, resulting in the network not being able to converge) (120).

The key idea of deep learning is to include a greater number of layers, making them "deeper" to exploit the *context* of the data in small steps, much like the human visual system. The idea existed since the 1990s (129), however it was the recent advances in computer hardware and graphical processing units (GPUs) that allowed the successful implementation of deep architectures. One of the most important type of networks in medical imaging is the *convolutional neural network* (CNN), which learns and preserves spatial relations in the data and thus is excellent for image analysis (Figure 2.2). The main difference with classical machine learning is that in deep learning, feature learning and performing a task have merged into one step, as deep neural networks extract features and learn useful non-linear representations automatically, directly from the raw data (118-120, 130). Common types of constituent layers, along with strategies incorporated in many modern CNNs to enhance their performance by improving their generalisation to unseen data, regularize them and accelerate the training process are described in Table 2.1.

CNNs alone or in combination with other approaches can be used for almost any supervised learning task, such as classification, regression and segmentation. Other popular network architectures for supervised learning are the *U-Nets* (131), which are inspired by CNNs and reconstruct the image using upscaling and skip connections and are primarily used for segmentation, and the *fully connected networks* (FCN), which can both easily be applied to many different tasks and

handle mixed data types. The most popular type of network for unsupervised learning, typically for dimensionality reduction, is the *autoencoder*, which finds the best non-linear representation of the data in lower dimensions by trying to reconstruct the original data.

Despite the immense success of deep learning, classical machine learning approaches are still widely used especially in cases with fewer data samples or in combination with deep neural networks.



**Figure 2.2** A deep convolutional neural network (CNN) trained for face recognition. Visualisation of the first hidden layer reveals the detected edge features. In the next hidden layer where the context is larger, the network captures shapes and face parts. The next hidden layer where the context is even larger, the network captures bigger part of the faces. The output layer predicts (classifies) the probability the input image belongs into a pre-defined set of categories (classes) (120). Adapted from *Artificial Intelligence in Medical Imaging: Opportunities, Applications and Risks*, page 28, figure 3.4.

<b>Common types of layers in CNN</b>	
<b>Convolutional</b>	A $k \times k$ filter (where $k < \text{image size}$ ) is convolved with the entire image in strides (like a “sliding window”) to extract features. Each region has the same shared <i>weights</i> , allowing for the training of deeper networks with fewer parameters. For image-based tasks, it is quite common to use several layers of convolutions at the input.
<b>Activation</b>	The feature maps from a convolutional layer are passed through nonlinear activation functions. The most popular activation function is the rectified linear unit or <i>ReLU</i> , which outputs a “0” for any negative input and outputs the input if the input is positive, and variations of it such as the <i>leaky ReLU</i> .
<b>Pooling</b>	Pooling is a down-sampling method, where one or more convolution layers are followed by a “pooling layer,” in which the outputs of adjacent convolutions are combined into a single output. This output is usually computed using the maximum value ( <i>max pooling</i> ) or the mean value ( <i>average pooling</i> ). Summarising the information of the previous convolution layer, makes the features more translation-invariant. Instead of pooling, down-sampling can also be achieved by convolutions with increased stride lengths.
<b>Fully Connected</b>	In this layer, every input is multiplied to an activation function and is passed to an output. The fully connected layers are usually applied towards the end of the CNN.
<b>Output</b>	The final activation function is related to the task and the objective function used for training the network. For classification, the most common function is the <i>softmax</i> which converts the values of the previous layer into an output prediction vector of size equal to the number of classes and sum equal to “1”. For regression tasks, a <i>linear</i> activation is appropriate.
<b>Strategies to enhance network performance</b>	
<b>Dropout</b>	It is the process of randomly excluding different neurons during each iteration of training, not allowing this way the weights to become co-dependent (132).
<b>Batch normalization</b>	This approach normalises the activation maps by subtracting the mean and dividing by the standard deviation for each training batch to solve the “vanishing-gradient” problem. It also makes the network less dependent to initialising parameters (133).

**Table 2.1** Common elements of convolutional neural networks (CNN)

### *2.1.3 Aim of the chapter*

This chapter reports on the development of computational methodologies for both cell and tissue classification, their application to the study of tumour histological heterogeneity and hence their potential value to enhance the validation of non-invasive functional imaging of tumour architecture. This chapter is divided in four autonomous sections:

- i. Section **2.3** reports on the development of a cell classification pipeline based on cell morphology and applied for the classification of cells in transgenic model of neuroblastoma.
- ii. Section **2.4** demonstrates how capturing global spatial context can improve cell classification.
- iii. Section **2.5** reports on the development of SuperCRF: an enhanced deep learning approach for incorporating global and local context in whole-slide image analysis.
- iv. Section **2.6** reports on the development of SuperHistopath: a deep learning pipeline for mapping the tumour heterogeneity on low-resolution whole-slide digital histology images.

## **2.2 Cell classification using morphological features**

### *2.2.1 Introduction*

Cells morphology and outlines can be distinguished on histopathological images at 20x magnification, whereas 40x magnification affords the clearer observation of morphological details of the cell nuclei. Many cell types share distinct morphological characteristics, such as their shape, size and nuclei distribution. For example, malignant cells usually have large (>10  $\mu\text{m}$ ) abnormal nuclei of variable texture and shape, while stromal cells (fibroblasts and endothelial cells) share a more spindle-shaped nuclei and lymphocytes have a small (<8  $\mu\text{m}$ ) dark nuclei with reduced cytoplasm (134). Computational pathology algorithms are aiming to quantify those differences to discriminate cells.

In classical machine learning, cell classification requires two independent tasks, feature extraction followed by feature classification.

#### 2.2.1.1 Aim of the section

In this section of Chapter 2, I implemented an image processing pipeline for the classification of cells in the Th-*MYCN* transgenic mouse model of neuroblastoma using a supervised learning approach. Firstly, cell nuclei were segmented, and then a series of well-established hand-crafted features were extracted. The extracted features were subsequently classified using a SVM, which is trained with ground-truth annotations from a pathologist.

### *2.2.2 Materials and Methods*

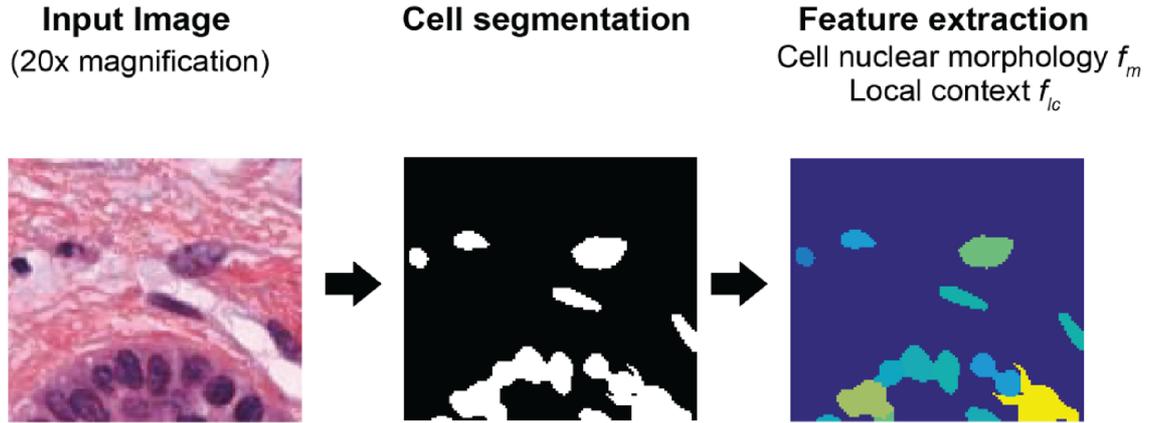
#### 2.2.2.1 Dataset

In total, 7 H&E-stained whole-slide formalin-fixed and paraffin-embedded (FFPE) images from Th-*MYCN* transgenic mouse models of neuroblastoma were used. All images were digitized at 20x magnification using a Hamamatsu NanoZoomer XR scanner (0.46 $\mu\text{m}$  resolution, Hamamatsu, Japan).

Ground-truth annotations were provided by a trainee pathologist (Dr. Matt Clarke, ICR) under the guidance of a senior paediatric neuropathologist (Dr. Tom Jacques, UCL Great Ormond Street Institute of Child Health) for a total of 16320 cells over 5 categories: 4608 undifferentiated neuroblasts, 2057 differentiating neuroblasts, 5252 apoptotic cells, 2246 lymphocytes and 2157 stromal cells (fibroblasts and endothelial cells).

#### 2.2.2.2 Cell segmentation and classification

Image processing was carried out using a pipeline based on CRImage (134). First, cell nuclei were extracted from H&E-staining by Otsu thresholding (135). Then, noisy image structures were deleted using morphological opening and the clustered nuclei were separated by the Watershed algorithm (136). For every nucleus, 91 morphological (134), three local-context and 46 cell-cytoplasm features were extracted (Figure 2.3). The morphological features included *i*) topological properties, such as the size of the nucleus, perimeter, eccentricity, acircularity, (137) *ii*) texture (Haralick) features, which describe the texture of the nuclei and are calculated from the co-occurrence grey level matrix, (138) *iii*) image moments which characterise the two dimensional density distribution function of the grey values of an image (139) and *iv*) Zernike moments which are shape parameters invariant to translation and rotation based on Zernike polynomials (140). A SVM with a radial basis function (RBF,  $\gamma=1/\text{number\_of\_features}$ ) kernel was trained with these features using the R-package *e1071* (141) and classified the cells into 5 categories: undifferentiated neuroblasts, differentiating neuroblasts, apoptotic cells, lymphocytes, stromal cells (Figure 2.4).



**Figure 2.3** The cell segmentation and subsequent feature extraction process from H&E images using the CRImage pipeline.

### 2.2.2.3 Evaluation metrics for classification

Some evaluation metrics for classification tasks are *accuracy*, *precision* and *sensitivity (recall)*, described as following:

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (2.1)$$

$$Precision = \frac{TP}{TP+FP} \quad (2.2)$$

$$Sensitivity (Recall) = \frac{TP}{TP+FN} \quad (2.3)$$

where  $TP$  is true positives,  $FP$  is false positives,  $TN$  is true negatives, and  $FN$  is false negatives.

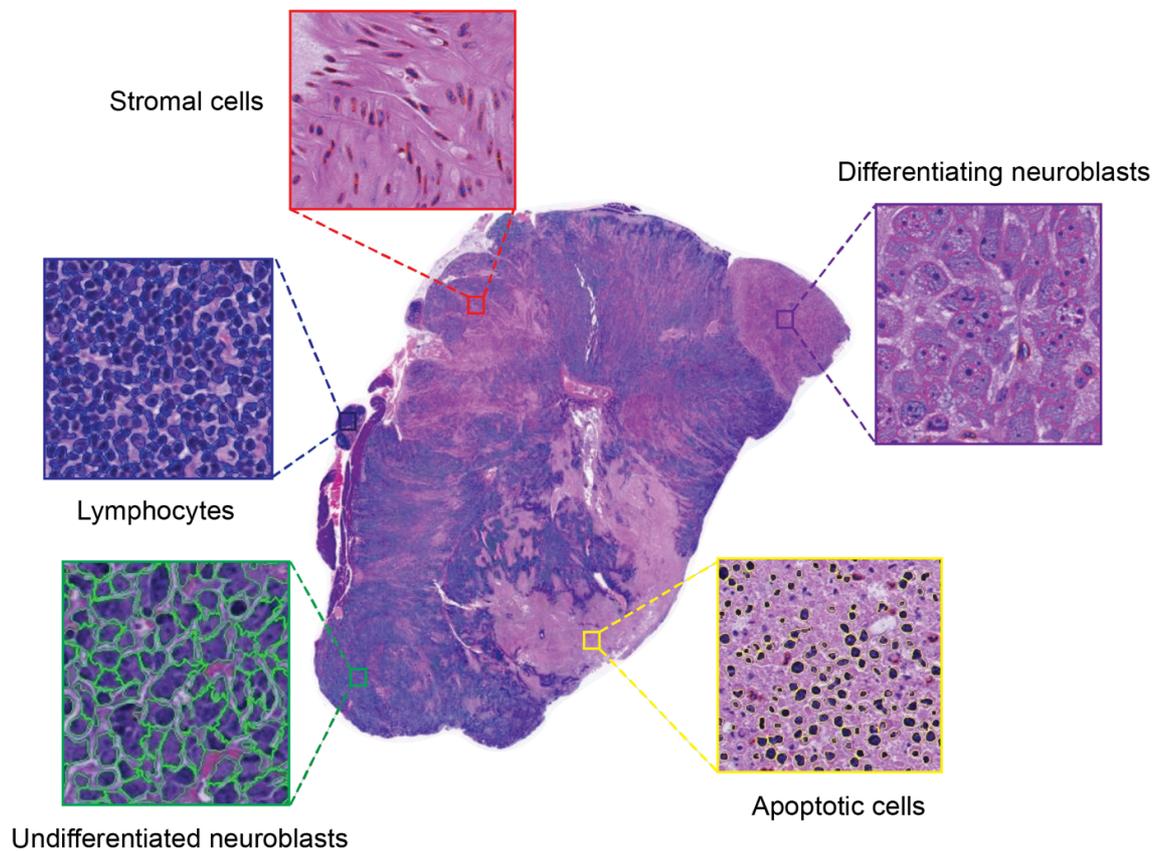
### 2.2.3 Results

Cells were segmented and classified into 5 categories with an overall accuracy of 94.3%, average precision of 94.3%, and average recall of 94.5%, after 50-fold cross validation (Table 2.2, Figure 2.4 and Figure 2.5): undifferentiated neuroblasts (Precision = 98.6%, Recall = 90%), differentiating neuroblasts (Precision = 96.8%, Recall = 98.1%), apoptotic cells (Precision = 95.4%, Recall = 88.9%), lymphocytes (Precision = 96.2%, Recall = 99.9%), stromal cells

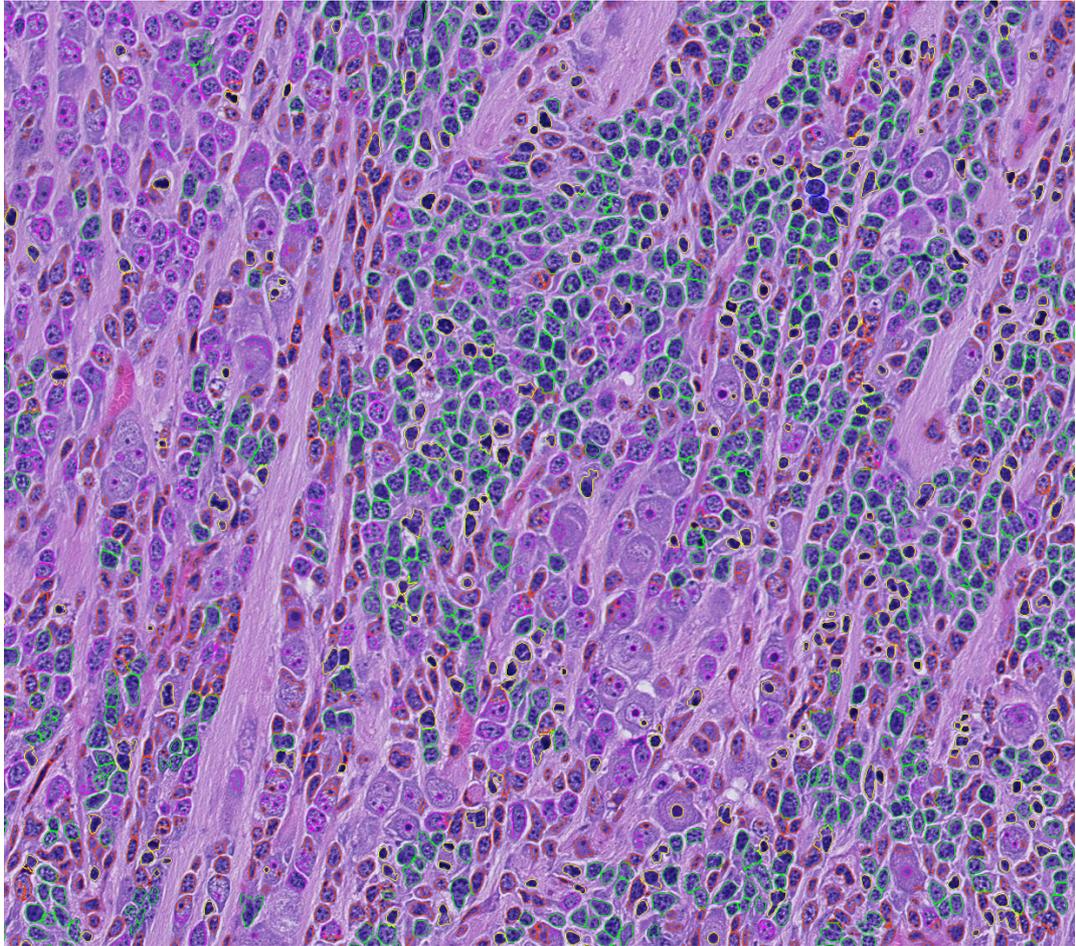
(Precision = 84.5%, Recall = 95.5%). The quality of the results was also visually evaluated by our pathologist.

	Apoptotic cells %	Undifferentiated neuroblasts %	Differentiated neuroblasts %	Lymphocytes %	Stromal cells%
Apoptotic cells %	95.41	2.83	0.12	0.06	1.58
Undifferentiated neuroblasts%	0.63	98.61	0	0	0.76
Differentiated neuroblasts %	0.41	1.29	96.79	0	1.51
Lymphocytes%	3.7	0	0	96.15	0.15
Stromal cells%	7.23	6.47	1.76	0	84.54

**Table 2.2** Confusion matrix of the cell classification results in the Th-MYCN mouse model of neuroblastoma



**Figure 2.4** Computational analysis of a digitized whole-slide histological image of a Th-MYCN neuroblastoma. Cells were segmented and classified into 5 categories with an overall accuracy of 94.3%: undifferentiated neuroblasts (98.61% accuracy, green), differentiating neuroblasts (96.79%, purple), apoptotic cells (95.41%, yellow), lymphocytes (96.15%, blue), stromal cells (84.54%, red).



**Figure 2.5** Representative results of the computational analysis of a digitized whole-slide histological image of a Th-MYCN neuroblastoma; Cell outlines: green – undifferentiated neuroblasts, blue – lymphocytes, red – stromal cells, yellow – apoptotic/dying cells, purple – differentiated neuroblasts.

#### *2.2.4 Discussion*

In Section 2.2, a pipeline for cell segmentation and classification based on cell nuclei and cytoplasm morphology was implemented. The classifier achieved an overall accuracy of 94.3% in the neuroblastoma dataset over 5 different cell categories.

This approach suffers from some known limitations associated with many classical machine learning methods. Despite the quality of the results obtained in this dataset, application of the pipeline to new datasets was always very challenging without extensive optimisations. Cell detection with CRImage was

considered satisfactory as validated by our pathologist in this dataset, yet inferior when compared to the level of detection achieved in other datasets with state-of-the-art approaches. Cell segmentation was particularly sub-optimal when cells are overlapping, a known limitation of the Watershed algorithm.

#### *2.2.5 Conclusion*

This approach provided satisfactory results for the classification of cells in the transgenic mouse models of neuroblastoma and was used to investigate the sensitivity of MRI to the underlying histology of neuroblastoma, as described in Chapters 5-7. However, more sophisticated methods including deep learning approaches should be further explored to achieve high-quality results in more complex tasks.

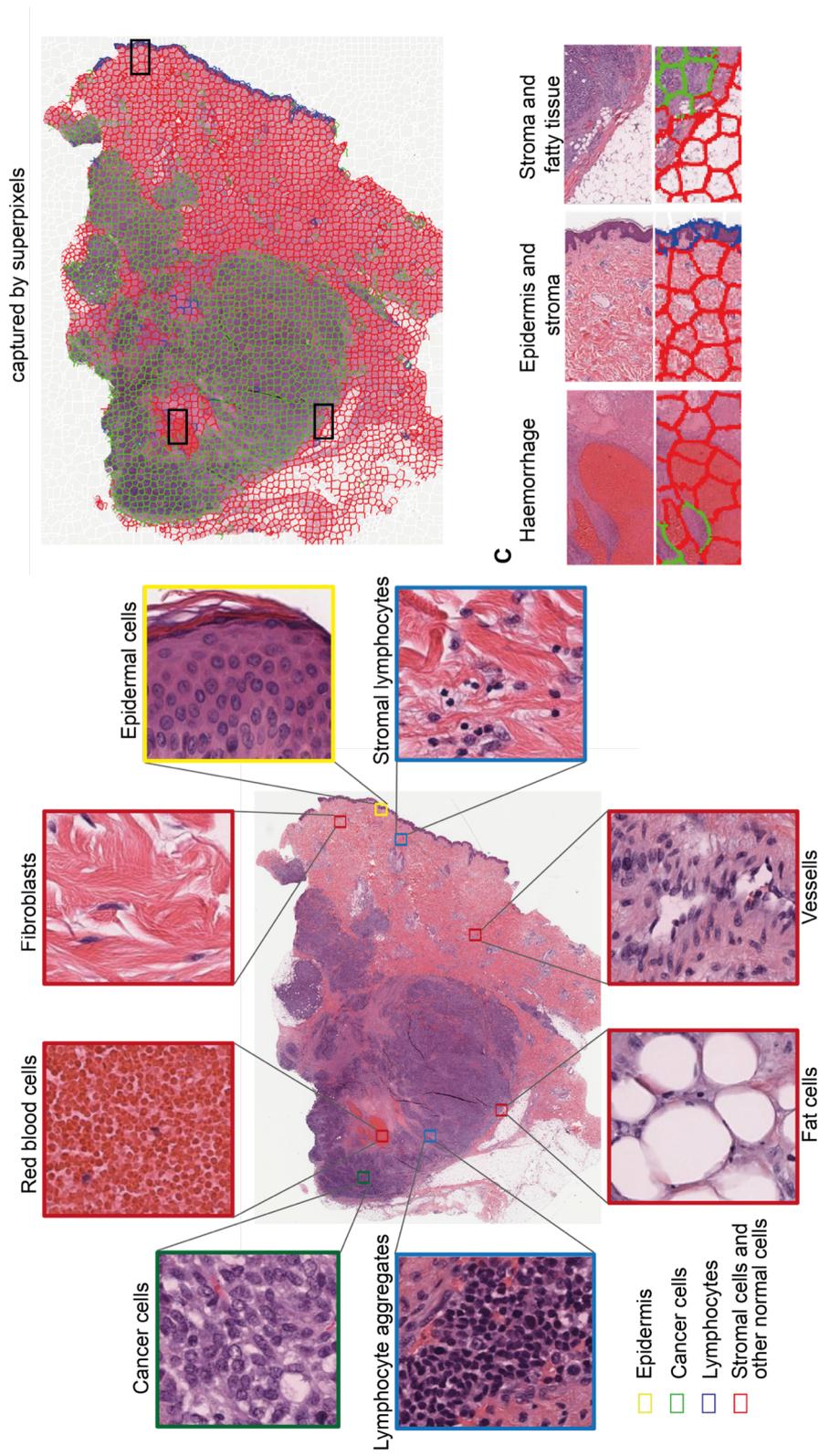
## **2.3 Capturing global spatial context to improve cell classification**

### *2.3.1 Introduction*

Deep learning has revolutionized the field of cell classification, with deep neural networks achieving higher-quality results than the state-of-the-art classical machine learning approaches (142-145). However, even state-of-the-art deep learning algorithms can underperform especially in cases where different cell types appear morphologically similar. Current computed pathology tools focus on individual cell nuclei morphology with limited abstract local context features, because they operate on image patches, which only provided a limited scope of the WSI. Pathologists on the other hand incorporate regional tissue architecture (in practice, by zooming in/out), together with cell morphological features to accurately classify cells. Thus, contextual information can be the key to further improve cell classification algorithms.

#### 2.3.1.1 Aim of the section

In this chapter section, I demonstrate the effectiveness of including global tumour spatial context into single-cell classification. I propose a multi-resolution hierarchical framework, which captures the spatial global context at low magnification, by classifying superpixels into biologically meaningful regions (tumour area, stromal, normal epidermis and lumen/white space, Figure 2.6) and combining them with cell nuclei morphological features at high resolution to improve single cell classification (Figure 2.7). The algorithm was applied on H&E-stained WSIs of clinical cutaneous melanoma.



**Figure 2.6** The complex nature of melanoma architecture. **A.** Heterogeneous tumour stroma makes accurate cell classification a difficult task. **B.** Superpixels captures tumour global architecture by delineating the boundaries among heterogeneous tumour components, including haemorrhage area, fatty tissues, stromal regions, epidermis and cancer nests. **C.** Current classification scheme assigned these components accurately to their respective superpixel classes. Top: example images; Bottom: segmentation and classification using SLIC superpixels.

## 2.3.2 Methodology and Results

### 2.3.2.1 Data

In total, 58 full-face, H&E stained section images from FFPE diagnostic blocks of clinical cutaneous melanoma skin cancer from The Cancer Genome Atlas (TCGA) were used. All digitized (Aperio ImageScope) histology images were scaled to 20x and 1.25x magnification with pixel resolution 0.504 and 8.064 $\mu$ m, respectively, using Bio-Formats (<https://www.openmicroscopy.org/bio-formats/>). To set the ground truth for regional classification at 1.25x magnification, an expert pathologist with over 20 years of experience (Dr. Ioannis Roxanis) provided annotations on the slides for 4 different regions: tumour area, stroma, normal epidermis and lumen/white space. We randomly selected 21 images for training and reserved the remaining 37 images as an independent test set.

For single cell classification, 7 WSIs at 20x were split into sub-images (tiles) of 2000x2000 pixels each, for computational efficiency. 3 WSIs were used for training and 4 for testing. Based on the pathologist's input, we used 3863 cell nuclei (1320 cancer cells, 1100 epidermal cells, 751 lymphocytes, 692 stromal cells) from 82 sub-images for training and 2405 cell nuclei (876 cancer cells, 602 epidermal cells, 417 lymphocytes, 510 stromal cells) from 224 sub-images as an independent test set (Figure 2.8A).

### 2.3.2.2 Superpixel-based tumour region classification

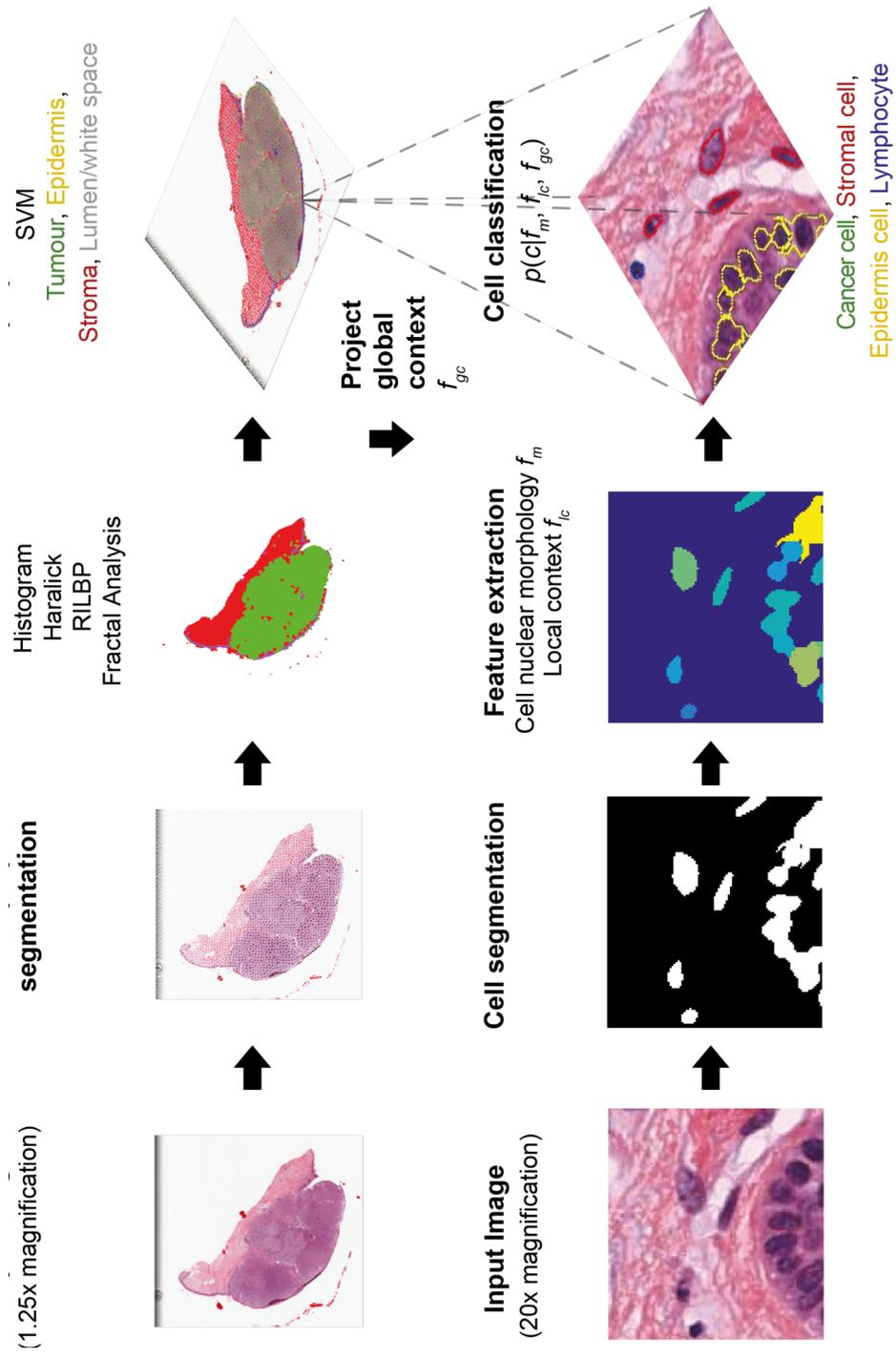
First, each dataset was pre-processed using the Reinhard stain normalization (146) to account for stain variabilities that could affect classification. Then, all images were segmented using the simple linear iterative clustering (SLIC) superpixels algorithm, which groups together similar neighbouring pixels (147). With the pathologist's input, the optimal number of superpixels was selected by visually identifying a superpixel size that capture only homogeneous areas and adhere to image boundaries. This is a critical step for ensuring accurate tissue segmentation, and therefore, classification (Figure 2.6BC). The number of

superpixels for each image was automatically set based on the image size according to Equation 2.1.

$$N_i = \text{ceiling}\left(\frac{S_i}{U}\right) \quad (2.4)$$

The SLIC algorithm inherently provides a roughly uniform superpixel size. Setting  $U = 1250$ , Equation 2.4 gave a mean superpixels size of  $35 \times 35$  pixels, equivalent to an area of approximately  $280 \times 280 \mu\text{m}^2$ . The SLIC superpixels algorithm was proven to be computationally efficient, requiring only 3s to segment a single downscaled image of  $2500 \times 2500$  pixels using a 2.9GHz Intel core i7 processor.

Region annotations were provided by the pathologist (Dr. Ioannis Roxanis). I identified 15477 superpixels belonging in tumour areas, 6989 in stroma areas, 141 in epidermis and 691 in lumen/white space for training based on their isocenter location within the annotated regions.



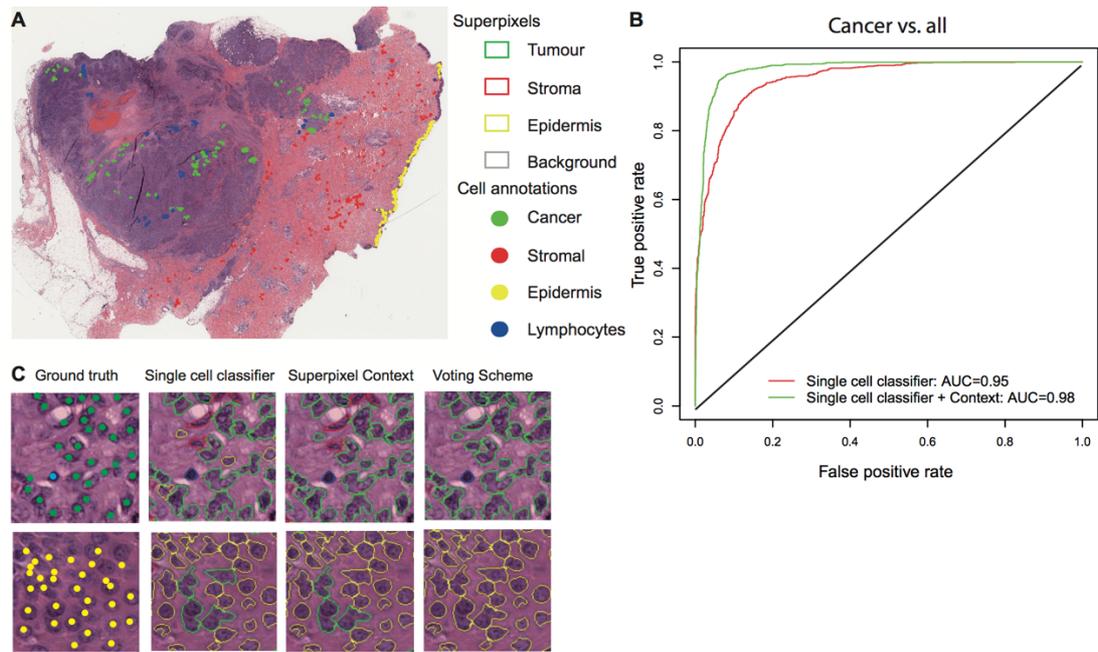
**Figure 2.7** Proposed hierarchical framework to project tumour global context onto single cell classification by integrating superpixel segmentation and classification.

Next, four types of well-established classical machine learning features were extracted, 85 in total, from each superpixel, including seven histogram features (mean values of hue, saturation, and brightness, sum of intensities, contrast, standard deviation, and entropy), and well-established texture features (12 Haralick features (138)), 59 rotation-invariant local binary patterns (RILBP), 7 segmentation-based fractal texture analysis (SFTA) features (148). Features were standardized into z-scores. The mean values and standard deviation of the features from the training set were used for the normalization of the test set. A SVM with a radial basis function (RBF,  $\gamma = 1/\text{number\_of\_features}$ ) was trained with these features to classify superpixels into the different biologically meaningful categories, cancer, stroma, normal epidermis and lumen/white space. To address the class imbalance problem for training, I randomly selected a subset of 5000 cancer and stroma superpixels and increased the penalty in the cost function for the epidermis and lumen/white space classes by a factor of 10.

Performance of classification using individual and various combinations of feature sets was tested (Table 2.3). The use of all 85 features, yielded the highest accuracy (97.7% in the training set using 10-fold cross validation and 95.7% in 2997 superpixels annotated in the 37 images of the independent test set).

Features	Accuracy (%)	Feature combinations	Accuracy
Hist.	95.9 %	Hist. + Haralick	97.3 %
Haralick	91.4 %	Hist. + RILBP	96.3 %
RILBP	88.8 %	Hist. + SFTA	96.9 %
SFTA	85.2 %	Hist. + Haralick + RILBP	97.1 %
		Hist. + Haralick + SFTA	97.1 %
		Hist. + Haralick + RILBP + SFTA	<b>97.7 %</b>

**Table 2.3** Accuracy matrix of the classification of superpixels for single sets of features (left) and various combinations (right).



**Figure 2.8** Superpixels provide global context for single cell classification. **A.** Representative superpixel classification of tumour regions overlaid with ground-truth cell annotations. **B.** ROC curves (cancer vs. all) illustrate the improvement in classification accuracy by adding superpixel context as additional features. **C.** Representative images comparing ground truth annotation and single cell classifiers with and without superpixels.

### 2.3.2.3 Cell Classification based on nuclear features

Image processing was carried out using a pipeline based on CRImage described in Section 2.2. For every nucleus, 91 morphological features ( $f_m$ ) were extracted (134). Three local features  $f_{lc}$  were added: the number of nuclei neighbours in a distance of  $25\mu\text{m}$ , the density at the particular cell position, and the size of the surrounding cytoplasm, calculated by thresholding the image's red channel after excluding the nuclei pixels.

For single cell classification, a SVM with a RBF ( $\gamma=1/\text{number\_of\_features}$ ) kernel was trained with these features and achieved 86.4% accuracy (Table 2.5) in the test dataset. As expected, the classifier underperformed in distinguishing between epidermal and cancer cells (Table 2.4, Figure 2.8), due to their similar morphology and related local context (both exist in crowded environments).

#### 2.3.2.4 Incorporating global context in cell classification

Two different schemes were used in order to integrate regional classification with cell classification. First, the type of area a single cell belonged to, provided by the superpixel classification, was added to the morphological feature set as an extra feature, the global context feature ( $f_{gc}$ ). This resulted in a considerable reduction of misclassification between epidermis and cancer cells (Table 2.4) and led to a much higher accuracy (91.6%, Table 2.5, Figure 2.8) compared to the cell-morphology based classifier.

Secondly, global context given by superpixels served as biological *a priori* knowledge to correct single cell classifications. E.g. stromal cells seldom exist in non-stromal regions, while, cancer cells should only exist in tumour regions and epidermal cells should be found only in epidermal regions. Lymphocytes, however, can infiltrate into both tumour regions and stroma, but are rarely found in epidermis (149). The regional context of a cell is thus of great importance for its annotation. Following these constrains an iterative voting scheme for cells in stromal and tumour regions was implemented (Equation 2.5):

$$c_i = \begin{cases} \textit{epidermis} & \textit{if} & s = \textit{epidermis} \\ t & \textit{ifelse} & t = s \vee t = \textit{lymphocyte} \\ k & \textit{else} & k = s \vee k = \textit{lymphocyte} \end{cases} \quad (2.5)$$

Where  $c_i$  is the cell at position  $i$ ,  $t \in \{\textit{cancer}, \textit{stromal}, \textit{epidermis}, \textit{lymphocyte}\}$  is the most probable annotation of the cell in the SVM,  $s \in \{\textit{cancer}, \textit{stroma}, \textit{epidermis}\}$  is the annotation of the cell's superpixel,  $k \in \{\textit{cancer}, \textit{stromal}, \textit{epidermis}, \textit{lymphocyte}\}$  is the annotation with the next highest unchecked probability in the SVM. This voting scheme (Eq. 2.5) was applied for all cell nuclei and resulted in 92.8% accuracy in the test dataset (Figure 2.8C, Table 2.5).

		Morphology				Morphology + global context				Morphology + Voting scheme			
		C	E	L	S	C	E	L	S	C	E	L	S
Classes	cancer	715	27	20	14	786	27	23	40	844	0	30	2
	epidermis	96	496	5	5	44	550	2	6	48	515	17	22
	lymphocyte	7	12	397	1	12	5	399	1	14	1	401	1
	stromal	18	4	17	471	11	12	18	469	10	4	23	473

**Table 2.4** Confusion matrix of the Single Cell Classifiers, based on nuclei morphology (left), adding of the global context as extra feature in the feature set (middle) and implementation of the voting scheme (right). C: cancer, E: epidermis, L: lymphocyte, S: stromal. Note the confusion between cancer cells and epidermis cells highlighted in red.

Method	Accuracy	Precision	Recall
Single Cell	86.4 %	87.4%	87.9 %
Single Cell + Context	91.6 %	91.5 %	92.2 %
Voting Scheme	92.8 %	92.8 %	92.7 %

**Table 2.5** Classification performance for the three classifiers. Both approaches of incorporating the global context led to markedly better results, with the voting scheme achieving the highest score.

### 2.3.3 Discussion and Conclusion

In this section of Chapter 2, a hierarchical framework was implemented to mirror the way pathologists perceive tumour architecture. This proof-of-concept study shows that the incorporation of global spatial context, in a biologically interpretable way, can overcome limitations associated with cell classification algorithms strictly based on nuclear morphology or abstract local contextual features. The proposed framework can be easily adapted to improve any nuclear morphology-based single-cell classification algorithm and study other cancer types.

Here, the incorporation of global context was done following two simple approaches, by either adding the spatial context as an extra feature or exploit it to constrain the algorithm with biological rules. While both resulted in a marked increase in the classification accuracy (from 86.4% to 91.6% and 92.8% respectively), these hard rules may also introduce some bias in the system. The

next step of this approach would be a method able to model the spatial associations of a cell to its microenvironment in an autonomous way. Also, more sophisticated approaches for incorporating the global context should be explored in combination with state-of-the-art deep learning methods (143).

## **2.4 SuperCRF: An enhanced deep learning approach for incorporating global and local context in whole-slide image analysis**

### *2.4.1 Introduction*

Robust tumour regional classification from lower resolution images can provide the contextual information that is key to further improve single-cell classification algorithms. The aim of this chapter section is to introduce dependencies of a single-cell to its global tissue context and cell neighbourhood and enhance learning results for cell classification from deep CNNs. Probabilistic graphical models have successfully been applied to improve cell classification in time-lapse imaging by considering the temporal context of a cell (150-155). Probabilistic graphical models have also been used successfully in histopathology images for detection and segmentation (156-159), disease and tissue staging (160, 161), and nuclei segmentation (162). In this study, instead of time dependency, graphical models are applied to introduce the spatial context of a cell as additional information to improve single-cell classification.

#### 2.4.1.1 Aim of the section

A multi-resolution hierarchical framework is proposed to mirror the way pathologists perceive tumour architecture, and applied to H&E-stained WSIs of clinical cutaneous melanoma skin cancer (Figure 2.9A) to understand its ability to characterise the immune-stroma-tumour interface and provide improved predictive/prognostic markers.

### *2.4.2 Materials and Methods*

#### 2.4.2.1 Datasets

In total, 105 full-face, H&E stained section images from FFPE diagnostic blocks of melanoma skin cancer from The Cancer Genome Atlas (TCGA) were used. All digitized (Aperio ImageScope) histology images are scaled to 20, 5, and 1.25x magnification with pixel resolution 0.504, 2.016, and 8.064 $\mu$ m, respectively, using Bio-Formats (<https://www.openmicroscopy.org/bio-formats/>). WSIs at 20x

magnification (representative size: 30,000x30,000 pixels), were split into sub-images (tiles) of 2,000x2,000 pixels each, for computational efficiency.

For the purpose of training and testing the different parts of our system the dataset was divided into sub-datasets, namely single-cell classification dataset, 5x sub-dataset, 1.25x sub-dataset and discovery sub-dataset (Table 2.6).

Name	Number of WSIs		Purpose
Single-cell classification sub-dataset	<b>Total</b>	<b>8</b>	Single-cell classification into 4 categories: cancer cells, lymphocytes, stromal cells, epidermal cells
	Training SC-CNN	3 (348 tiles)	
	Training SuperCRF	2 (84 tiles)	
	Testing	3 (290 tiles)	
5x sub-dataset	<b>Total</b>	<b>16</b>	Region classification into 5 categories: tumour, stroma, lymphocyte cluster, normal epidermis, lumen/white space
	Training	10	
	Testing	6	
1.25x sub-dataset	<b>Total</b>	<b>58</b>	Region classification into 4 categories: tumour, stroma, normal epidermis, lumen/white space
	Training	21	
	Testing	37	
Entire dataset	<b>Total</b>	<b>98</b>	Study of the tumour-stroma interface. To accelerate the analysis, 50 tiles (2000x2000 pixels) containing tumours were randomly sampled from every whole-slide image (WSI)

**Table 2.6** Summary of the data used to train and test the different parts of the SuperCRF system, as well as to study the cancer-immune-stroma interface.

#### 2.4.2.2 Single-Cell Classification Using a Spatially Constrained Convolutional Neural Network

A Spatially Constrained Convolutional Neural Network (SC-CNN) (143) was used for single cell classification (Figure 2.9E). SC-CNN uses spatial regression in order to predict the probability of a pixel being the centre of the nucleus. The nucleus is classified by a neighbouring ensemble predictor (NEP) in conjunction with a standard softmax CNN. The network's layers were randomly initialised as we have found that to perform better than transfer learning from real-world datasets in our experiments with pathological samples.

#### 2.4.2.3 Superpixel-Based Tumour Region Classification

The machine learning superpixel-based framework implemented in chapter section 2.3 was applied to classify tumour tissue regions in low resolution (5x and 1.25x) images. With our pathologist's input (Dr. Ioannis Roxanis) the optimal

superpixel size was set using  $U = 1250$  in equation 2.5 for both 5x and 1.25x images.

Region annotations were also provided by our senior pathologist. Overall, for the 1.25x training sub-dataset, I identified 15,477 superpixels belonging in tumour areas, 6,989 in stroma areas, 141 in epidermis and 691 in lumen/white space, while for the 5x training sub-dataset I identified 1,193 superpixels belonging in tumour areas, 1,324 in stroma areas, 360 in epidermis, 506 in lymphocyte clusters and 830 in lumen/white space based on location of their isocentre within the annotated regions.

For the 5x sub-dataset, superpixels were classified into five categories: tumour area, stroma, normal epidermis, lymphocytes cluster, and lumen/white space. The penalty in the cost function was increased for the epidermis and lumen/white space classes by a factor of 10 when training the SVM, to account for class imbalance. For the 1.25x sub-dataset superpixels classification consisted of four categories: tumour area, stroma, normal epidermis, and lumen/white space. A subset of 5,000 cancer and stroma superpixels was randomly selected and the penalty in the cost function was increased for the epidermis and lumen/white space classes by a factor of 10, again to account for class imbalance (Figure 2.9D).

#### 2.4.2.4 SuperCRF

Single-cell based classification approaches often assign a class label based on the morphology of individual cells, regardless of their neighbouring cells. However, these spatial relationships provide important information that is used by pathologists. Conditional random fields (CRF) are undirected graphical models that represent efficient ways to model dependences, by factorizing the probability density into a specific set of conditional dependence (163). Therefore, the tumour microenvironment can be modelled by a CRF by introducing nodes for cells and superpixels, as well as edges whenever there is a spatial relationship between nodes.

Lymphocytes were excluded from the CRF assumption that neighbouring cells have a higher probability to share the same class labels, since they infiltrate, in an inconsistent manner ranging from sparse to highly dense, in tumour as well as stromal tissue. Therefore, lymphocytes kept their label as assigned by the SC-CNN.

Let  $n$  be the total number of cells (besides lymphocytes) in the image and  $c_i \in \{stromal, cancer, epidermis\}$ ,  $i = 1, 2, \dots, n$  the input labels of the cells as assigned by the SC-CNN. Let  $s_i$ , be the corresponding superpixel for a cell  $c_i$  with  $s_i \in \{stromal, cancer, epidermis, white\ space\}$  for 1.25x superpixels and  $s_i \in \{stromal, cancer, epidermis, lymphocyte, white\ space\}$  for 5x superpixels.  $\mathbf{x} \in \{\mathbf{c}, \mathbf{s}\}$  comprises the nodes of the CRF. The CRF assigns output labels  $y_i \in \{stromal, cancer, epidermis, lymphocyte, white\ space\}$  based on the input data. The joint probability distribution over input data and output labels,  $p(y_1, y_2, \dots, y_n | x_1, x_2, \dots, x_n)$  can be modelled by factorizing the probability density into a specific set of conditional dependence relationships (Figure 2.9C).

$$p(y | x) = \frac{1}{Z} \prod_{i=1}^n p(y_i | x_i) = \frac{1}{Z} \exp(\sum E(x_i, y_i, xN_i, yN_i)) \quad (2.6)$$

where  $Z$  is a normalizing constant,  $w$  is a weight vector and

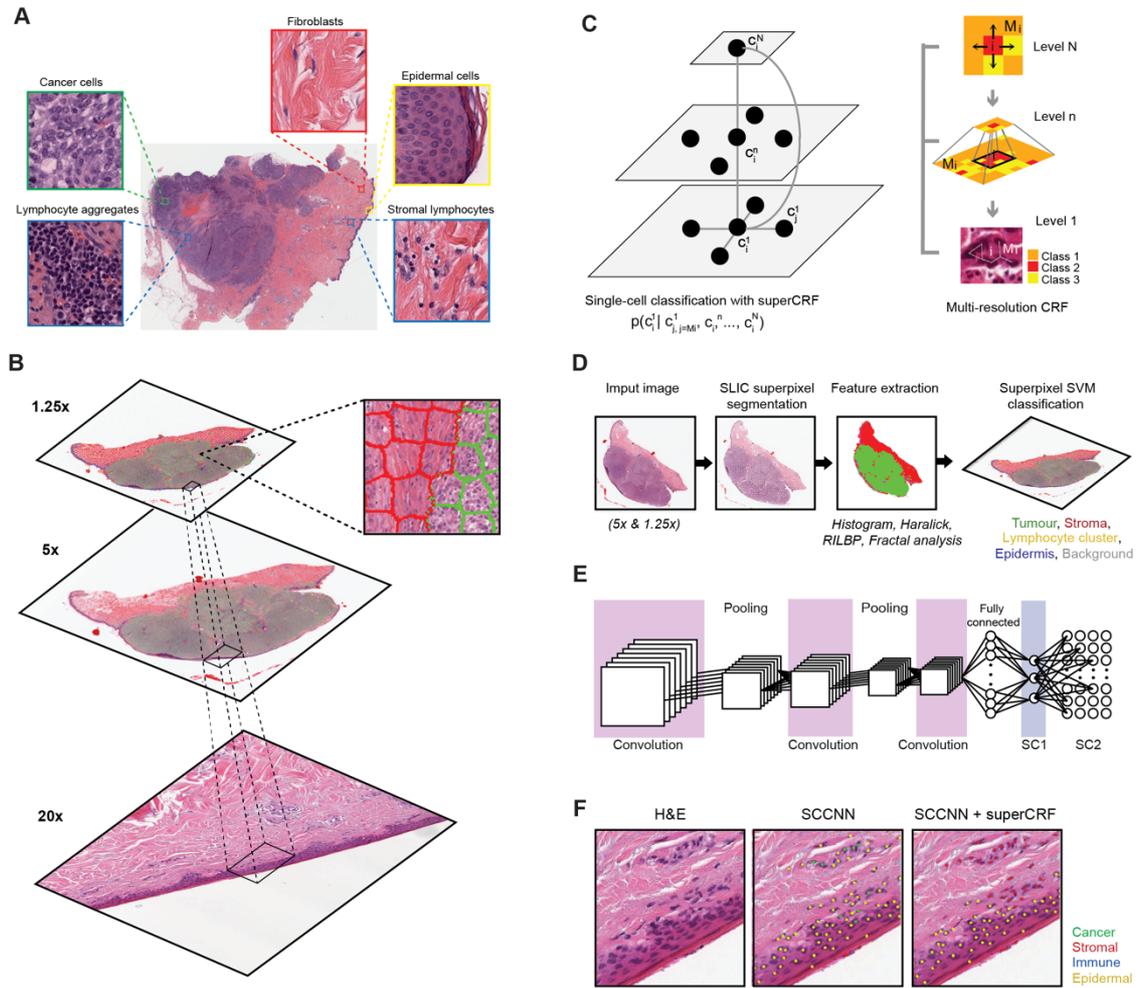
$$E(x_i, y_i, xN_i, yN_i) = \sum \Phi(x_i, y_i) + \sum \psi_c(xN_i, yN_i) \quad (2.7)$$

defines the energy function of the CRF.

The node potentials  $\Phi(x_i, y_i)$  represent the evidence that a cell  $i$ , with the input label  $x_i$  takes the class label  $y_i$ . The node potential can be defined as  $\Phi(x_i, y_i) = f(x_i, y_i) + b$ , with  $f(x_i, y_i) = \begin{cases} 1 & \text{if } y_i = x_i \\ 0 & \text{else} \end{cases}$  and  $b$  representing the bias.

The edge potentials  $\psi_c(xN_i, yN_i)$  model the probability that neighbouring cells take a similar cell label.  $N_i$  is the neighbourhood of cell  $i$ , defined as all the cells that can be found in a defined distance. The edge potentials are defined as:  $\psi_c(x_i, y_i, xN_i, yN_i) = f(x_i, y_i) * f(xN_i, yN_i) + b$ .

The CRF was trained with stochastic gradient descent and the decoding was applied using loopy belief propagation. The toolbox of M. Schmidt was used to train and decode the CRF (164).



**Figure 2.9** Overview of the SuperCRF framework for analysing H&E-stained pathological images of melanoma. **A.** Major histological features of melanoma architecture. **B.** Projection of regional classification results using superpixels from various scales to the 20x magnification for the improvement of single-cell classification. **C.** Graphical representation of node dependencies (cells and superpixels) across different scales. **D.** Region classification scheme using a superpixel based machine-learning method in whole-slide images (5x and 1.25x magnification). **E.** Single-cell classification using a state-of-the-art spatially constrained-convolution neural network (SC-CNN) classifier **F.** representative results of the SC-CNN cell classifier alone and combined with our SuperCRF system. Note the misclassification of various stromal cells by the SC-CNN, which are corrected by our model.

#### 2.4.2.5 Survival analysis

The prognosis value of the abundance of stromal cells and location of lymphocytes was evaluated in the discovery sub-dataset. The ratio of stromal cells to all cells, the ratio of lymphocytes in cancer areas to all lymphocytes, and the ratio of lymphocytes in stroma areas to all lymphocytes were calculated for each patient. Patients were divided into high and low ratio groups, by split at the median value of all scores to ensure that the groups were of similar sizes. Patients with a ratio of lymphocytes being high inside the tumour area and low in the stroma were categorized as the “immune infiltration” group whereas patients with a ratio of lymphocytes being low in the tumour area and high in the stroma were categorized as “immune excluded,” based on the recent classification of the main immune phenotypes of anticancer immunity that predict response to immunotherapy (165). The number of patients belonging to neither of these two groups (high/high  $n = 6$  and low/low  $n = 5$ ) was too small to perform the survival analysis. Non-parametric Kaplan-Meier estimation was used to analyse overall survival in 94 patients. Differences between survival estimates were assessed with the log-rank test. Finally, Cox regression models were adjusted, testing for the independent prognostic relevance of our risk scores. To test if Breslow-thickness (the distance between the upper layer of the epidermis and the deepest point of tumour penetration) was contributing to a high ratio of stromal cells, a multivariate model was created containing both stromal cells ratio and Breslow-thickness, as well as two univariate models containing the covariates separately. Pearson’s correlation was used to test for linear relation between the two variables.

### 2.4.3 *Results*

#### 2.4.3.1 SuperCRF improves accuracy of cell classification

First, the state-of-the-art deep learning method, spatially-constrained CNN (SC-CNN) algorithm was trained to detect and classify cells in high resolution (20x) WSI into four categories: cancer cells, stroma cells, lymphocytes, and epidermis cells. The SC-CNN network yielded an accuracy of 84.63% over 4,059 cells in the independent test set (Table 2.7, Table 2.8). Visual inspection revealed that

the majority of false positives were misclassification of stromal and cancer cells as epidermis, which confirmed the initial motive for the incorporation of regional and spatial information to improve classification.

Subsequently, a conditional random field (CRF) was trained by combining the cellular neighbourhood with tumour region classification (cancer area, stroma, normal epidermis, lymphocyte cluster, and lumen/white space) from low resolution images (5x and 1.25x, Figure 2.9B), given by the superpixel-based machine-learning framework. The superpixel-based region classification was applied on the two datasets of 1.25x and 5x magnification (Figure 2.10A) and achieved high accuracy in regional classification (1.25x sub-dataset: Overall accuracy 97.7% in the training set using 10-fold cross validation and 95.7% in 2997 superpixels annotated in the 37 images of the independent test set; 5x sub-dataset: Overall accuracy 97.1% in the training set using 10-fold cross validation and 95.2% in 1798 superpixels annotated in the six images of the independent test set).

To train SuperCRF, firstly dependencies of single-cells on their cell neighbourhood were introduced. Cells were considered neighbours in the CRF, if they were located in a spatial proximity of 15 $\mu$ m (or 30 pixels), which resulted in an average of 1.3 neighbours per cell. Subsequently, this local neighbourhood was integrated with global context by connecting the CRF single-cell nodes to the regional classification results from superpixels. To determine the best configuration, four different CRFs were trained and their performance was compared in terms of single-cell classification on a test set, including three samples, 290 tiles and 4059 single-cell annotations (1527 cancer cells, 676 lymphocytes, 837 normal epidermis cells, 1019 stromal cells).

In detail, for the first CRF no global context was used, just cell neighbourhood dependencies, i.e., the only edges of the CRF were between neighbouring cells (singleCellCRF). For the second and third CRF superpixel nodes were introduced. Now, single-cell nodes are not only connected to neighbouring cells but every single-cell node is also connected to a superpixel node. A CRF was trained for 5x superpixel classification (CRF5x) and 1.25x superpixel

classification (CRF1.25x). Furthermore, a CRF which included all above dependencies was trained with every single-cell node connected to two superpixel nodes in 5 and 1.25x resolution (SuperCRF). Already the singleCellCRF (Accuracy: 87.6%, Precision: 89.7%, Recall: 89.5%, Table 2.7, Table 2.9) improves the classification accuracy compared to the SC-CNN (84.6%, Precision: 87.6%, Recall: 88.1%, Table 2.7, Table 2.8). However, the use of contextual information by the introduction of superpixel nodes, markedly improves the classification metrics (Accuracy 90.8%, Precision: 92.5%, Recall: 91.1%, Table 2.7, Table 2.10) for CRF1.25x and (Accuracy 91.7%, Precision: 93%, Recall: 91.3%, Table 2.7, Table 2.11) for the CRF5x. The SuperCRF, using nodes from superpixels in both 5 and 1.25x resolution images, as well as the neighbouring cells, resulted in the highest classification outcome (Accuracy 96.5%, Precision: 96.4%, Recall: 96.3%, Table 2.7, Table 2.12, Figure 2.9F, Figure 2.10B).

Method	Accuracy (%)	Precision	Recall
SC-CNN	84.63	0.8756	0.8808
singleCellCRF	87.61	0.8973	0.8946
CRF1.25x	90.79	0.9248	0.9110
CRF5x	91.70	0.9298	0.9126
SuperCRF	<b>96.48</b>	<b>0.9644</b>	<b>0.9629</b>

**Table 2.7** Evaluation of different conditional random fields (CRF) versions and a state-of-the-art spatially constrained-convolution neural network (SC-CNN) deep learning cell-classifier.

		SC-CNN			
		C	E	L	S
Classes (Cells)	Cancer	1149	2	0	11
	Epidermis	348	834	16	204
	Lymphocytes	24	1	660	12
	Stromal	6	0	0	792

**Table 2.8** Confusion matrix of the classified cells from the spatially constrained-convolution neural network (SC-CNN) deep learning cell-classifier. C: cancer cells, E: epidermis cells, L: lymphocytes, S: Stromal cells.

		singleCellCRF			
		C	E	L	S
Classes (Cells)	Cancer	1263	0	1	17
	Epidermis	236	836	12	197
	Lymphocytes	24	1	660	12
	Stromal	4	0	3	793

**Table 2.9** Confusion matrix of the classified cells from singleCellCRF architecture, where the trained conditional random field (CRF) combines the classified cells from the spatially constrained-convolution neural network (SC-CNN), with local context information (neighbouring classified cells). C: cancer cells, E: epidermis cells, L: lymphocytes, S: Stromal cells.

		CRF1.25x			
		C	E	L	S
Classes (Cells)	Cancer	1295	1	2	28
	Epidermis	203	829	2	78
	Lymphocytes	24	1	660	12
	Stromal	5	6	12	901

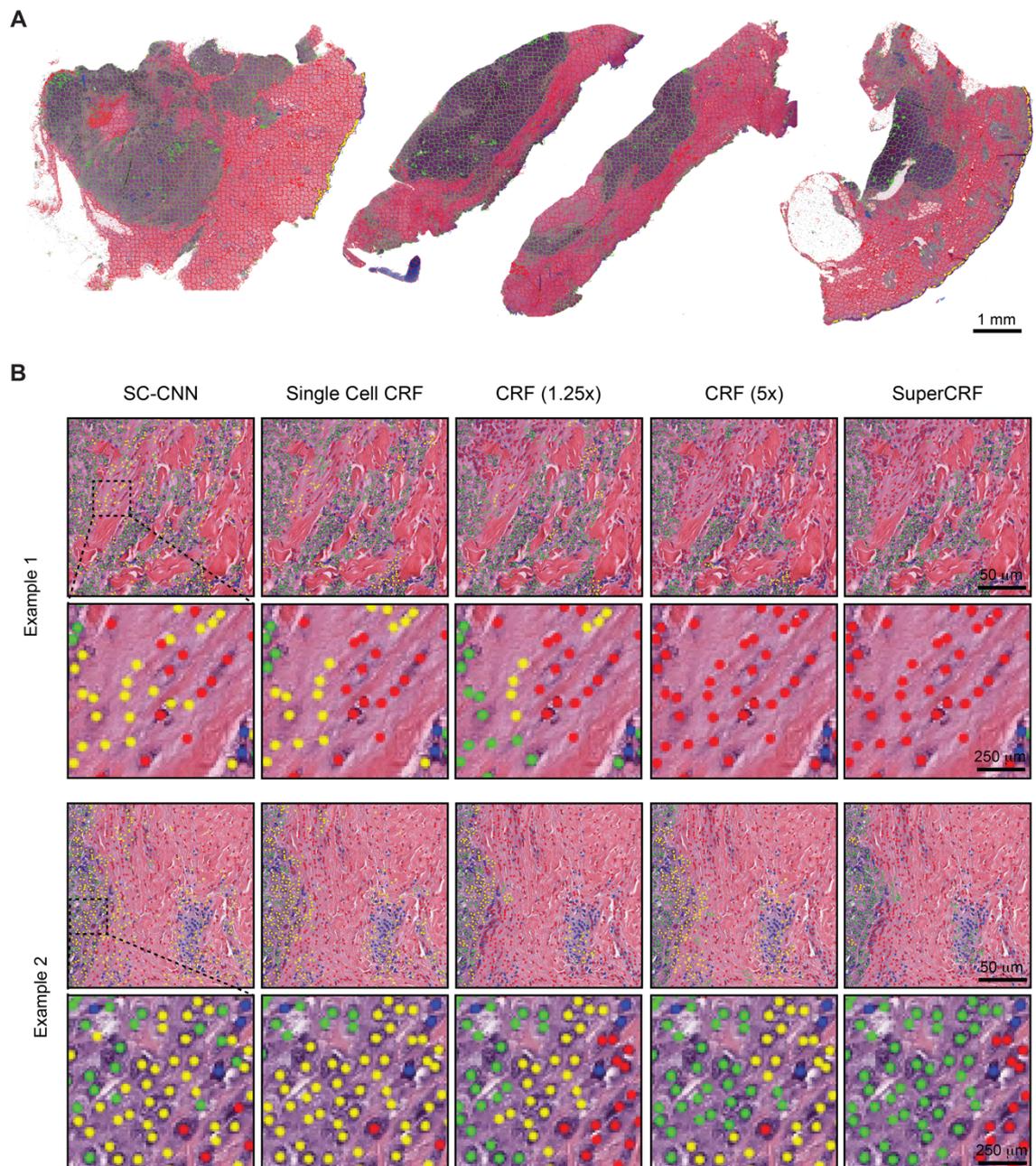
**Table 2.10** Confusion matrix of the classified cells from CRF1.25x architecture, where the trained conditional random field (CRF) combines the classified cells from the spatially constrained-convolution neural network (SC-CNN), with the region classification information from the 1.25x magnification whole-slide images. C: cancer cells, E: epidermis cells, L: lymphocytes, S: Stromal cells.

		CRF5x			
		C	E	L	S
Classes (Cells)	Cancer	1311	2	2	6
	Epidermis	177	778	1	28
	Lymphocytes	24	1	660	12
	Stromal	15	56	13	973

**Table 2.11** Confusion matrix of the classified cells from CRF5x architecture, where the trained conditional random field (CRF) combines the classified cells from the spatially constrained-convolution neural network (SC-CNN), with the region classification information from the 5x magnification whole-slide images. C: cancer cells, E: epidermis cells, L: lymphocytes, S: Stromal cells.

		SuperCRF			
		C	E	L	S
Classes (Cells)	Cancer	1482	24	2	20
	Epidermis	17	795	1	8
	Lymphocytes	24	1	660	12
	Stromal	4	17	13	979

**Table 2.12** Confusion matrix of the classified cells from SuperCRF architecture, where the trained conditional random field (CRF) combines the classified cells from the spatially constrained-convolution neural network (SC-CNN), with the region classification information from the 1.25x and 5x magnification whole-slide images and local context information (neighbouring classified cells). C: cancer cells, E: epidermis cells, L: lymphocytes, S: Stromal cells.



**Figure 2.10** Representative examples of both superpixel and single-cell classification with or without SuperCRF. **A.** Superpixels-based regional classification on representative whole slide images (5x magnification) of melanoma. Green: tumour area, Red: stroma area, Blue: normal epidermis, Yellow: lymphocyte cluster. **B.** Representative images showing cell classification using a state-of-the-art spatially constrained-convolution neural network (SC-CNN) and four conditional random fields (CRF) models. Note the mislabelling of many cancer and stromal cells as epidermis cells when using the SC-CNN and the gradual increase in classification accuracy with the best accuracy achieved with the SuperCRF. Green, cancer cells; Red, stromal cells; Blue, lymphocytes; Yellow, epidermis cells.

#### 2.4.3.2 SuperCRF's increased accuracy of cell classification improves confidence in stromal cell ratio as a predictive feature of survival in melanoma

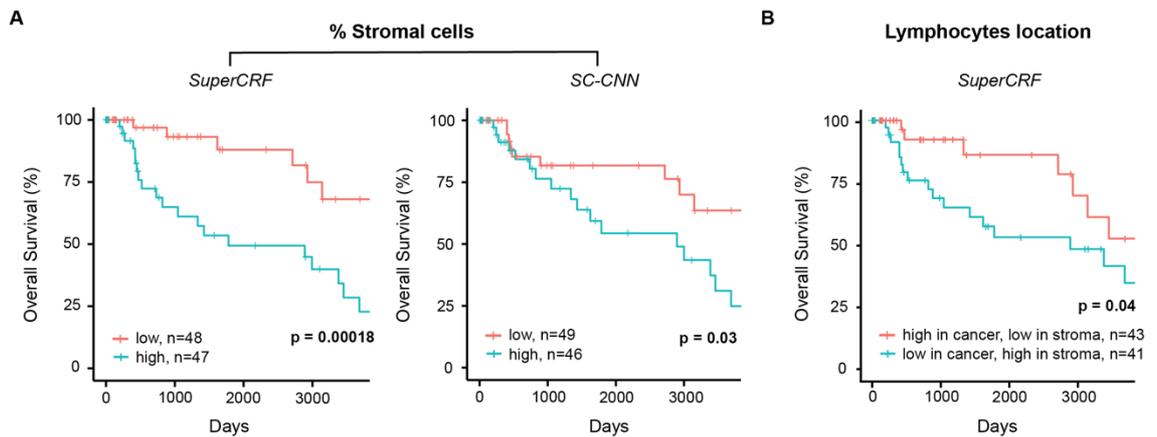
The crosstalk between cancer cells and stromal cells play an active role in tumour invasion and metastasis, and controlling immune infiltration and is increasingly recognized as a hallmark of cancer (166). Tumour-stromal cell ratio has been shown to hold prognostic and predictive information in patient with solid tumours (165, 167, 168). Here, a high stromal cell ratio is also a predictor of poor prognosis in melanoma using both values derived from the multivariate models of SC-CNN and SuperCRF in the discovery sub-dataset. Yet SuperCRF yields a significantly higher confidence in the predictive value of the stromal cell ratio (SuperCRF:  $P < 0.0001$ , Coxph-Regression (discretized by median): HR = 4.1,  $P = 0.006$ ; SC-CNN:  $P = 0.039$ , Coxph-Regression (discretized by median): HR = 2.4,  $P = 0.05$ , Figure 2.11A).

Similar regression coefficients for both stromal cells ratio and Breslow-thickness covariates were observed between the multivariate and the two univariate survival models (1.404 and 0.171, respectively, for the multivariate model and 1.633 and 0.179 for the univariate models) of the SuperCRF. Pearson's correlation showed no correlation between stromal cells ratio and Breslow-thickness ( $r = -0.05$ ), overall indicating that stromal cells ratio is independent to Breslow-thickness.

#### 2.4.3.3 Combining cell and region classification: Location of the immune infiltrate is predictive of survival in melanoma

There is increasing evidence of the value of immune infiltration to provide prognostic information and predictors of response in patient with melanoma [recently reviewed in (169)]. The spatial compartmentalization of immune cells afforded by our SuperCRF (by the cell and region classification results) was used to define the recently-described main immune phenotypes of anticancer immunity that predict response to immunotherapy (165). Patients with a classified "immune excluded" phenotype, defined by a low lymphocyte ratio inside the tumour area and high inside the stroma area, was associated with a significantly worse prognosis compared to "inflamed" tumours characterized by a high ratio of

lymphocytes inside the tumour and a low ratio inside the stroma ( $P = 0.026$ , Cox PH –regression:  $HR = 2.57$ ,  $P = 0.032$ , Figure 2.11B). Taken together, the data is consistent with the model of a stroma-mediated immune suppressive microenvironment that exclude T cells from the vicinity of cancer cells.



**Figure 2.11** Associations between survival outcomes and SuperCRF-define risk groups in the Cancer Genome Atlas (TCGA) cohorts of patients with melanoma. **A.** Kaplan-Meier Survival curves for patients in the high-risk group (blue) and low risk group classified by stromal cells ratio derived from SuperCRF (left) and using only the SC-CNN classifier. Note the difference in the p-value using the two methods. **B.** Kaplan-Meier Survival curves for patients in the high-risk group (blue) and low risk group classified by immune phenotype based on spatial distribution of lymphocytes in different tumour compartments derived from SuperCRF.

#### 2.4.4 Discussion

In this section of Chapter 2, a framework which fuses traditional machine learning with deep learning to model the way pathologists incorporate large-scale tissue architecture and context across spatial scales was implemented, to improve single-cell classification in large whole-section slide images. Using this approach, a marked 11.85% overall improvement in the accuracy of the state-of-the-art deep learning SC-CNN cell classifier was demonstrated. Also, the similar values of both precision and recall and their simultaneous increase in every step show the unbiased nature of the approach.

Computational pathology algorithms, typically exploit the inter-cell phenotypic differences for cell classification, yet even state-of-the-art deep learning algorithms tend to underperform in this task, mainly due to the disproportional numbers of cells sharing similar nuclear morphological features, or due to intra-class diversity, seen for example in tumour stroma (fatty tissue, necrosis, vessels, muscle, fibroblasts, and associated collagen). Whilst computers can quantify morphological differences in a considerably more complex way, pathologists still generally outperform computers in cell classification. An essential reason is that pathologists incorporate key contextual information such as heterogeneous tissue architecture, together with cell morphological features.

The idea that a cancer cell is dependent on its neighbouring cells and global context is comparable to the fundamental concept in landscape ecology that a living population depends on the existing habitats and is not equally spread on the terrain. A particular habitat could favour the development of specific organisms. In practice, landscape ecologists denote the habitats from satellite images and then “ground-truth” them by the detailed small-scale sampling of the habitats of interest (3). This inspired the design of the framework by introducing CRF dependencies between *i*) the cells and their neighbours and *ii*) the cells and to the global context (i.e., habitats from low resolution captured by the classified superpixels).

The proposed framework connects deep learning and classical image processing using probabilistic graphical models. All the information was combined using a CRF graphical model, which have been widely applied in image analysis for pathological images, yet mainly for semantic segmentation (156, 157, 170). Here, we *i)* introduce a new way to capture high-level spatial context using superpixels, *ii)* propose a new CRF model that introduces dependences over space and across different spatial scales, thereby modelling multiple cells and their associated superpixels simultaneously for more accurate classification, *iii)* introduce the concept of context-specific CRF modelling, given that the strength of dependence can be variable according to tumour compartments. There is an increasing interest in combining deep learning with different strategies, or “umbrella approaches,” such as the use of traditional machine learning to spatially explicit context used in this study, with the aim to, not only refine and improve the overall existing deep learning network (157, 171-173), but also facilitate biological interpretation compared to the “black-box”-like approach of deep-learning-only methods. However, optimizing and inventing new and refined deep learning networks is of equal importance, as during experimentation we observed that the better we made our single-cell classifier baseline, the more effective our SuperCRF approach became.

It is also shown that combining cell classification with the global context given by the region classification (both inherent parts of the SuperCRF architecture) can open new avenues to study the cancer microenvironment from histopathological slides. For example, the spectacular response observed in clinical trials of immunotherapy in patients with incurable melanoma calls for a better understanding of the tumour microenvironment and in particular the cancer-immune-stroma interface. Here, the ability of this approach to look at lymphocytes within their cellular and global context can predict melanoma patient survival and potentially provide biomarker stratification for immunotherapeutic approaches, by identifying the three main types of tumour immune-phenotypes including *i)* inflamed tumours which are characterized by infiltrated T Cells within the tumour, and associated with a generally good prognosis *ii)* immune-excluded tumours, in which T cells are present but prevented to infiltrate the tumour due to stromal interaction, and associated with worse prognosis (and obviously *iii)* immune

desert tumours). This could also potentially be extended to provide quantitative biomarkers to characterize the immune infiltrating response to immunotherapy. In accordance with the immune-excluded phenotype it is also demonstrated that, tumours rich in stromal cells had a marked poorer prognosis in patients with melanoma. With p-value lower by two orders of magnitude, this method provides stronger predictive power than by using deep-learning only method for cell classification.

In the future, the framework will be expanded to include an upward optimization step for the superpixels which may include additional classes for cells, regions and structures in order to provide a complete characterization of the tumour microenvironment. This may include deriving further classes from higher resolution images such as the lymphocyte clusters in this study, which were difficult to visualize in 1.25x resolution images. Additional deep learning methods will be explored to perfect the classification of superpixels.

The primary aim of this study was to demonstrate proof-of-principle that the introduction of global and local context as cell dependencies using a probabilistic graphical model as a post-processing step, like an “umbrella,” can significantly improve the performance of deep learning or classical machine learning cell classifiers based only on cell-morphology and abstract local context information. The SC-CNN architecture was chosen as the primary cell classification step due to its state-of-the-art performance in cell detection and classification compared to other well-established deep learning and classical machine learning approaches (143). Alternatively, other promising deep learning networks could potentially be used including InceptionV3 (174), InceptionV4 (175), or a VGG architecture (176).

#### *2.4.5 Conclusion*

The novel general framework SuperCRF improves cell classification by introducing global and local context-based information much like pathologists do. SuperCRF can be implemented in combination with any single-cell classifier and represent valuable tools to study the cancer-stroma-immune interface, which we used to identify predictors of survival in melanoma patients from conventional H&E stained histopathology.

## **2.5 SuperHistopath: A Deep Learning Pipeline for Mapping Tumour Heterogeneity on Low-Resolution Whole-Slide Digital Histopathology Images**

### *2.5.1 Introduction*

The relatively new ability to map the spatial context of each single cell has opened new avenues for the study of the tumour micro-environment, which is key to guide the delivery of precision medicine including immunotherapy. However, computational pathology is still not widely adopted in the oncology clinic. One of the challenges lies in the gigabyte sizes of high-resolution WSIs, which result in computationally expensive approaches. WSIs need to be divided into images patches (typical size: 256x256) before being processed by deep networks such as convolutional neural networks (CNNs) (177). Secondly, single-cell approaches provide markers that are often hard-to-be-evaluated or even interpreted by the pathologists and can be prone to the generalization errors when applied in new unseen dataset. As a result, many promising markers eventually fail to reach the clinic due to a lack of cross-validation in new independent datasets. On the other hand, tissue classification approaches, which target multicellular assemblies and paucicellular areas where individual cells are incorporated into the region segmentation, would be accessible for visual validation by pathologists. Such algorithms would enable the characterization of the distribution and interrelationship of global features that are currently detectable by human perception but not quantifiable without artificial intelligence- (AI)-assisted numerical expression.

Current computed pathology tools primarily focus on individual cell analysis at high-resolution (40x/20x magnification) with limited local context features, whereas pathologists frequently employ collateral information, taking into account the overall tissue micro-architecture. Many established clinical markers are actually identified at low or intermediate magnifications, including tumour architecture-based grading systems (178, 179), stroma-tumour ratio (180, 181), infiltrating lymphocytes (TILs) (169, 182) and necrosis (183-185). This has not been yet fully emulated by computational pathology methodologies. However, some methods for the classification of tissue components have been suggested

either using image patch classification typically with a CNN or pixel-level classification/segmentation typically with a U-Net-like architecture (131), mainly for tasks such as the dichotomized classification of tissue (e.g. cancerous vs non-cancerous) (186, 187), the segmentation of a feature of interest (e.g. glands) (188, 189) or multi-type tissue classification (117, 190-194). For segmentation purposes, U-Net-like architectures are usually preferred over CNNs, which have established limitations in conforming to object contours. Yet, CNNs have also resulted in promising segmentation approaches (157, 195, 196) with the enhanced capability of classifying a large number of categories (126). Multi-scale approaches incorporating information from various image resolutions have also been proposed (197-200). Different approaches have been explored for the classification of epithelium or stroma using superpixels-based segmentation of image patches with either hand-crafted or deep learning features (201, 202). *Bejnordi and colleagues* used a similar method for their multi-scale approach for the classification of tissue or non-tissue components on low resolution images and stroma and background regions from intermediate and high resolution images (203). However, these methods are typically performed on high-magnifications image patches (20-40x and more rarely 10x) and are associated with a high computational cost.

#### 2.5.1.1 Aim of this section

In this section of Chapter 2, I propose a framework (SuperHistopath), which can map most of the global context features that contribute to the rich tumour morphological heterogeneity visible to pathologists at low resolution and used for clinical decision making in a computationally efficient manner. First the SLIC superpixels algorithm (147) is applied directly on the WSI at low resolution (5x magnification) and subsequently the superpixels are classified into different tumour region categories based on pathologists' annotations. SuperHistopath particularly capitalises on:

- i. the use of superpixels which provide visually homogeneous areas of similar size respecting the region boundaries and avoid the potential

- degradation of classification performance associated with image patches, (no matter how small) spanning over multiple tissue categories.
- ii. the use of CNN necessary to accurately classify and map the multiple tissue categories that constitute the rich and complex histological intra-tumoural heterogeneity.
  - iii. the computational efficiency, faster processing speed and lower memory requirements associated with processing the WSI at low resolution without breaking them down to image patches.

SuperHistopath is applied to H&E-stained images from three different cancer types: clinical cutaneous melanoma, triple-negative breast cancer and tumours arising in genetically-engineered mouse models of high-risk childhood neuroblastoma.

## 2.5.2 *Materials and Methods*

### 2.5.2.1 Ethical statement

Clinical datasets generated from diagnostic H&E images were provided anonymised to the researchers. The neuroblastoma preclinical dataset was built from H&E images collected during previous in vivo studies approved by The Institute of Cancer Research Animal Welfare and Ethical Review Body and performed in accordance with the UK Home Office Animals (Scientific Procedures) Act 1986.

### 2.5.2.2 Datasets

All digitized whole-slide images (WSI) used in this study were H&E-stained, FFPE sections, and scaled to 5x magnification as presented in Table 2.13. The framework is applied to clinical patient samples of cutaneous melanoma and triple-negative breast cancer, in addition to tumour samples from transgenic mouse models of childhood neuroblastoma. Both the Th-*MYCN* and Th-*ALK<sup>F1174L</sup>/MYCN* mouse models have been shown to spontaneously develop

abdominal tumours, which mirror the major histopathological characteristics of childhood high-risk disease (43, 204).

Cancer type	Number of WSIs	Digital scanner	Pixel resolution (5x magnification)	Dataset
Cutaneous melanoma	127	Aperio ImageScope	2.016 $\mu\text{m}$	The Cancer Genome Atlas (TCGA)
Triple-negative breast cancer	23	NanoZoomer XR	2.3 $\mu\text{m}$	Internal dataset, Collaboration with The Serbian Institute of Oncology
High-risk neuroblastoma (mouse models)	73	NanoZoomer XR	2.3 $\mu\text{m}$	Internal dataset Tumours samples coming from established Th-MYCN and Th-ALK <sup>F1174L</sup> /MYCN transgenic mouse colonies (205, 206) and processed by a clinical histopathological core facility

**Table 2.13** Summary of the datasets used

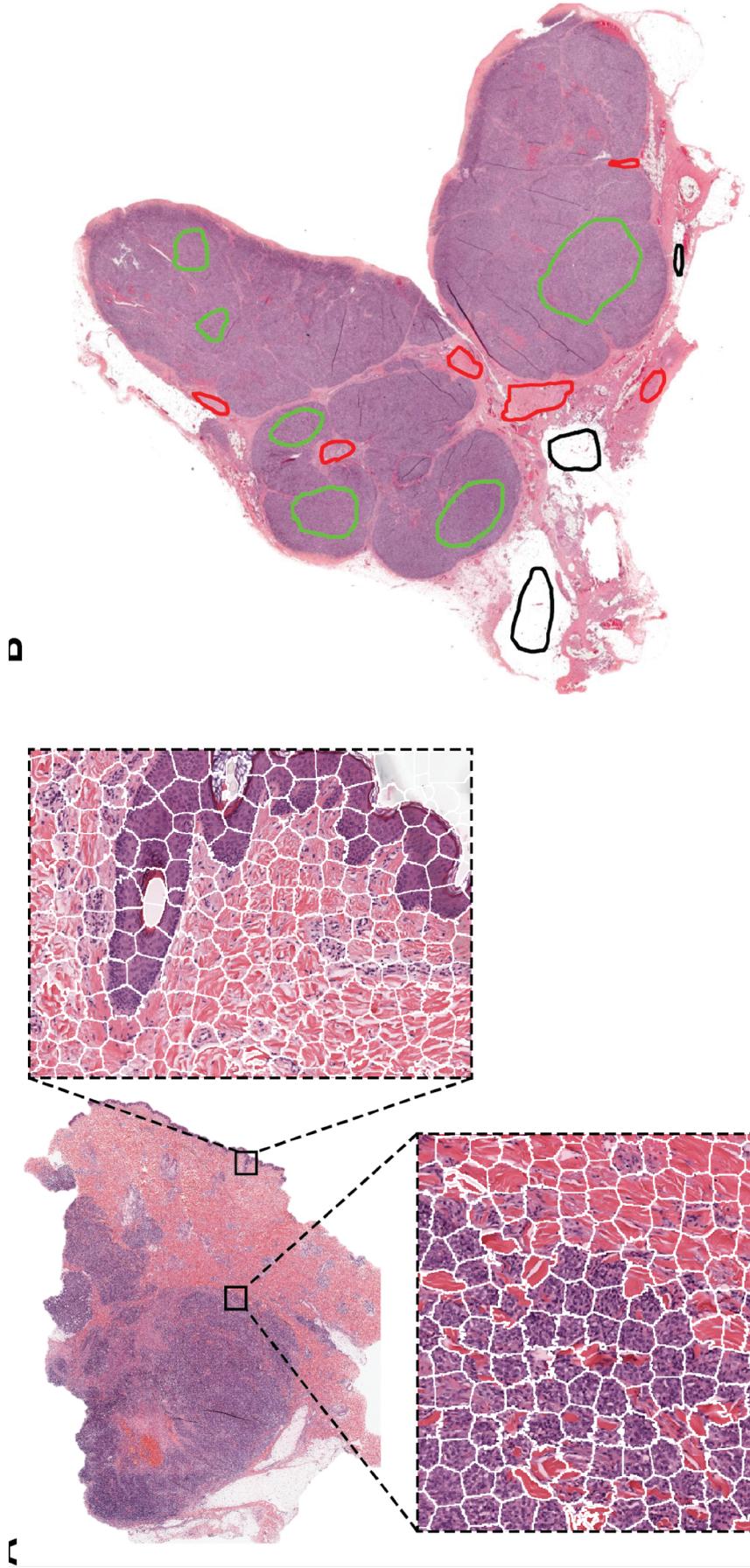
### 2.5.2.3 Region classification

First, each dataset was pre-processed using the Reinhard stain normalization (146) to account for stain variabilities that could affect classification. Then, all images were segmented using the SLIC superpixels algorithm as described in Section 2.3.2.2. Setting  $U = 1500$  on Equation 2.5 to ensure superpixels capture only homogeneous areas (Figure 2.12A), gave a mean superpixels size of 51x51 pixels, equivalent to an area of approximately 117x117  $\mu\text{m}^2$ . Bilinear interpolation was subsequently use to resize each superpixel to a fixed size of 56x56 or 75x75 pixels (the minimum input size for inception-like network architectures).

Region annotations were provided by a senior pathologist with over 20 years of experience for the melanoma and breast cancer clinical datasets (Dr. Ioannis Roxanis), and a senior paediatric neuropathologist with over 20 years of

experience (Prof. Rosa Noguera, University of Valencia) for the neuroblastoma mouse datasets. For training and testing, superpixels were assigned to each category based on their isocentre location within the annotated regions. Note that region annotations for our algorithm do not need to delineate boundaries as illustrated in Figure 2.12B.

The numbers of clinically relevant tissue categories, number of WSIs and superpixels used for training and testing are summarized for each tumour types in Table 2.14. Standard image augmentations, such as rotations ( $90^\circ$ ,  $-90^\circ$ ,  $180^\circ$ ), flips (horizontal and vertical), and contrast (histogram equalization) were performed in each case to capture more variation and even out the training dataset imbalances.



**Figure 2.12** Representative examples of the SLIC superpixels segmentation and ground-truth annotations in TCGA melanoma samples **A**. Whole-slide image segmentation using the SLIC superpixels algorithm. Note how the superpixels adhere to the boundaries of the different components of the tumor with each superpixel containing a single type of tissue **B**. Ground-truth annotations are provided by the pathologists by marking samples of the region components (the different colors represent different regions) without the need for delineating the boundaries of the tumor components.

Cancer type	Number of WSIs used for network training		Regional classification	
Cutaneous melanoma	Total	27	<b>6 categories</b>	<b>Superpixels for training</b>
	Training	22	Tumour tissue	21940
	Testing	5	Stroma	12419
			Normal epidermis	1646
			Lymphocytes cluster	2367
			Fat	15484
			Empty/white space	3412
Triple-negative breast cancer	Total	23	<b>6 categories</b>	<b>Superpixels for training</b>
	Training	18	Tumour tissue	18873
	Testing	5	Stroma	24220
			Necrosis	15102
			Lymphocytes cluster	3472
			Fat	10044
			Empty/white space	16473
High-risk neuroblastoma (mouse model)	Total	60	<b>8 categories</b>	<b>Superpixels for training</b>
	Training	44	Region of undifferentiated neuroblasts	20512
	Testing	16	Tissue damage (necrosis/apoptosis)	17645
			Differentiation region	5740
			Lymphocytes cluster	4009
			Haemorrhage (blood)	6124
			Muscle	6415
			Kidney	14976
		Empty/white space	21470	

**Table 2.14** Summary of the datasets used for training and testing the convolutional neural network. Note that the testing datasets consisted of whole-slide images from different patients from the training dataset.

#### 2.5.2.4 Training of the convolutional neural networks

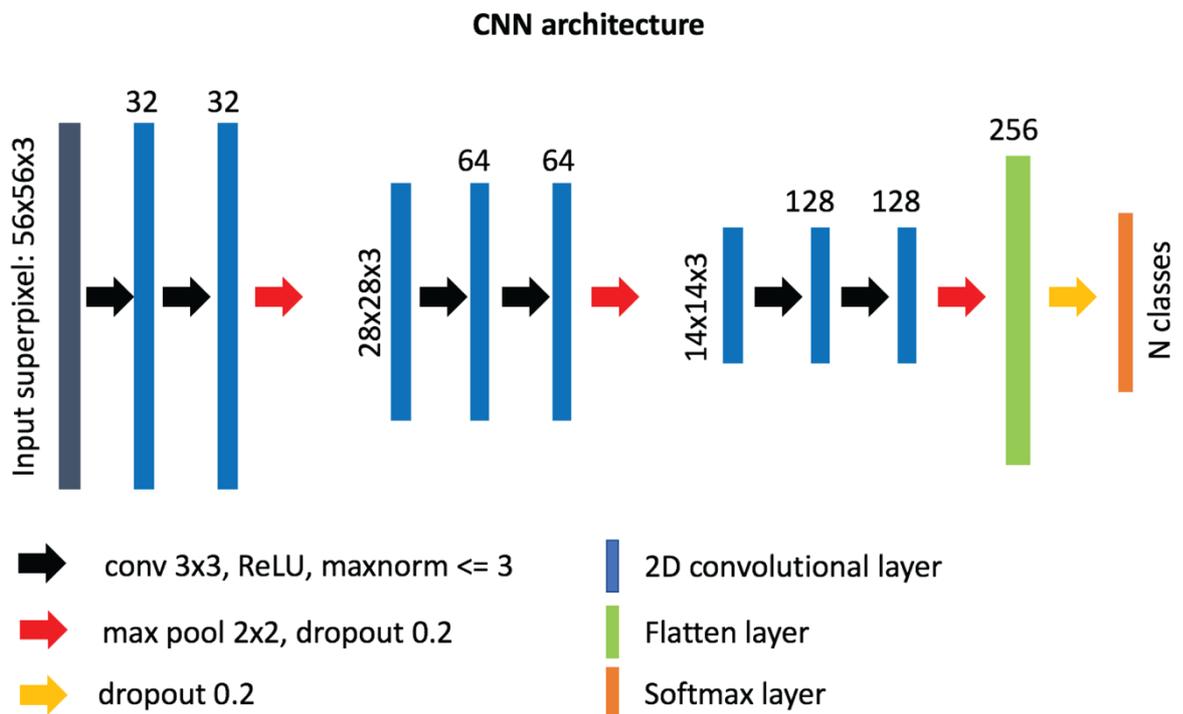
Our custom-designed CNN (Figure 2.13) for superpixel classification consists of 6 convolutional layers (32, 32, 64, 64, 128, 128 neurons respectively) of 3x3 filter size and 3 max-pooling layers, followed by a “flatten” layer and a dense layer of 256 neurons (Table 2.12). A superpixel RGB image (post-interpolation) was used as input into the network and normalised from range 0-255 to range 0-1 using the maximum value. The output of the network was a label assigned to each superpixel based on which region category it belonged to. After empirical experimentation, a ReLU activation function was used in all layers except for the last layer where standard softmax was used for classification. The weights incident to each hidden unit were constrained to have a norm value less than or equal to 3 and a dropout unit of 0.2 was used before every max-pooling operation to avoid overfitting (132). The weights of the layers were randomly initialized using “Glorot uniform” initialization (207), and the network was optimized using the Adam method (208) with a learning rate of  $10^{-3}$  and a categorical cross-entropy cost function. The number of trainable parameters for our custom-made network is ~1.9M. The network was implemented in python (v. 3.6.5) using the Keras/Tensorflow libraries (v. 2.2.4/1.12.0 respectively).

To choose the best network for our framework, we tested other known neural network architectures as implemented in the Keras framework, including InceptionV3 (174), Xception (209), InceptionResNetV2 (175), and ResNet (210). We initialised the weights using the pre-trained ImageNet weights. To optimize each network, we excluded the final classification layer, and added three additional layers, *i*) a global average pooling layer, *ii*) a dense layer of 256

neurons with ReLU activation, constrained to have a norm value less than or equal to 3, and *iii*) a dense layer tailored to the number of classes of each cancer type using the softmax function for classification.

For inception-like architectures (Inception v3, InceptionResNetV2, Xception) only superpixels of size 75x75 were used. We trained all the networks for 50 epochs using batch sizes of 150 and 256 for superpixels of sizes 75x75 and 56x56 respectively, and kept the models with the highest validation accuracy.

The Xception and custom-made networks were re-trained from the beginning for each cancer type, without applying any further changes.



**Figure 2.13** Architecture of our custom-designed convolutional neural network for the classification of superpixels into different tissue-level categories.

## 2.5.2.5 Application of SuperHistopath for the quantification of clinical features of interest

### 2.5.2.5.1 Melanoma

In the melanoma dataset, the number of pixels belonging to each classified category was calculated. For each patient, *i*) the ratio of pixels classified as stroma region to all pixels in tumour compartments, and *ii*) the ratio of pixels classified as clusters of lymphocytes to all pixels in tumour compartments were derived; the prognostic value of these quantitative indices was evaluated using survival analysis. Patients were divided into high- and low-risk groups based on split at the median value of all scores to ensure both groups were of similar size. Kaplan-Meier estimation was used to compare overall survival in the 127 patients. Differences between survival estimates were assessed with the log-rank test and hazard ratios were calculated using Cox's proportional-hazard regression.

### 2.5.2.5.2 Neuroblastoma

In the neuroblastoma dataset, the differences in phenotype between the Th-*ALK<sup>F1174L</sup>/MYCN* (n=7) and Th-*MYCN* tumours (n=6) were evaluated by quantifying the proportion of pixels classified by the SuperHistopath as regions rich in undifferentiated neuroblasts, differentiating neuroblasts, tissue damage (necrosis/apoptosis) haemorrhage and clusters of lymphocytes. Note that *i*) stroma in these tumours was not quantified as they faithfully mirror the stroma-poor phenotype which define high-risk disease *ii*) clusters of lymphocytes universally correspond to encapsulation of lymph node by the tumour, rather than tumour infiltrates, consistent with the "cold" immune phenotype of high-risk disease. The focus was on identifying any significant difference in the ratio of differentiation or the ratio of haemorrhagic regions to all tumour compartments between the two tumour types using the Mann-Whitney U test, with a 5% level of significance.

## 2.5.3 Results

### 2.5.3.1 SuperHistopath can accurately map the complex histological heterogeneity of tumours

#### 2.5.3.1.1 Melanoma

We first developed and evaluated our framework on the H&E-stained, FFPE sections of clinical specimen of cutaneous melanoma scaled to 5x magnification. Figure 2.12 shows the results of the segmentation using the simple linear iterative clustering (SLIC) superpixels algorithm, which groups together similar neighbouring pixels.

The optimized Xception network achieved the highest score and classified the melanoma sample regions into 6 predefined tissue categories of interest: tumour tissue, stroma, cluster of lymphocytes, normal epidermis, fat, and empty/white space with an overall accuracy of 98.8%, an average precision of 96.9%, and an average recall of 98.5% over 14092 superpixels in a separate test set of 5 images (Table 2.15, Table 2.16). Our custom CNN also achieved comparable performance to the state-of-the-art networks with an overall accuracy of 96.7%, an average precision of 93.6%, and an average recall of 93.6% (Figure 2.13, Table 2.17). Figure 2.14 shows qualitative results of our approach's regional classification in representative melanoma WSIs using the optimized Xception network.

Network	Accuracy (%)	Precision (%)	Recall (%)	Parameters (in millions)
InceptionV3	97.5	94.2	96.7	~22.4
InceptionResNetV2	97.7	94.1	97.3	~54.8
ResNet50	93.8	92.2	88.9	~24.2
Xception	98.8	96.9	98.5	~21.4
Our custom-made CNN	96.7	93.6	93.6	~1.9

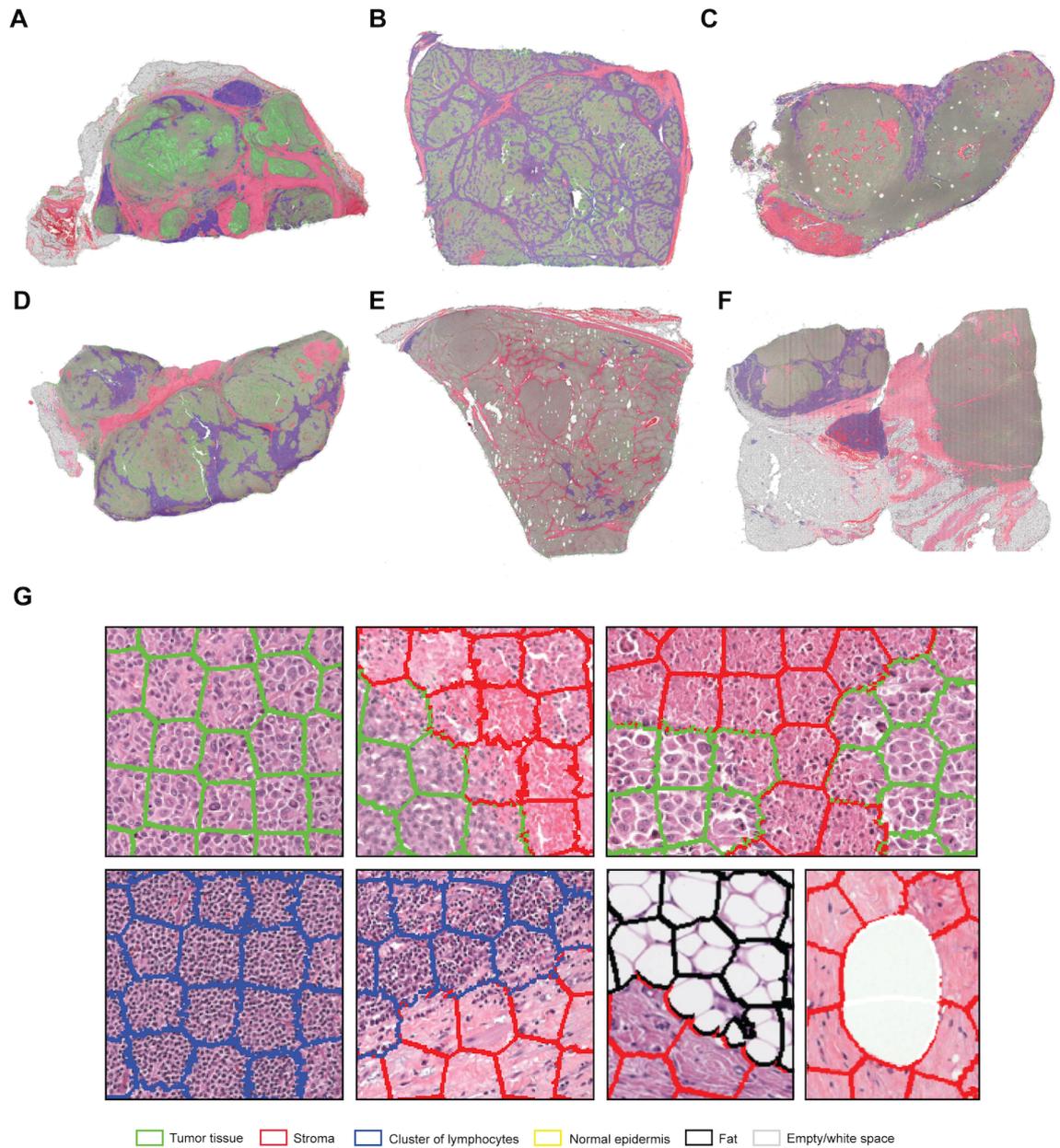
**Table 2.15** Evaluation metrics of the different neural network architectures in the TCGA melanoma test dataset.

	Tumour	Stroma	Epidermis	Lym	Fat	Empty space
Tumour	5286	10	7	8	0	0
Stroma	9	986	0	0	2	0
Epidermis	22	0	545	0	1	0
Lym	0	0	1	821	0	0
Fat	0	9	0	0	5603	3
Empty space	0	0	0	0	98	681

**Table 2.16** Confusion matrix of the classification of superpixels using the optimized Xception network in melanoma patients in 6 categories: tumour, stroma, normal epidermis, cluster of lymphocytes (Lym), fat and empty/white space (separate test set of 5 whole-slide images). Overall accuracy = 98.8%, average precision = 96.9%, average recall = 98.5%.

	Tumour	Stroma	Epidermis	Lym	Fat	Empty space
Tumour	5223	24	27	36	1	0
Stroma	24	937	33	2	1	0
Epidermis	95	0	473	0	0	0
Lym	8	0	1	812	0	0
Fat	0	20	2	0	5481	112
Empty space	0	1	0	0	83	695

**Table 2.17** Confusion matrix of the classification of superpixels using the custom-made CNN in melanoma patients in 6 categories: tumour, stroma, normal epidermis, cluster of lymphocytes (Lym), fat and empty/white space (independent test set of 5 whole-slide images). Overall accuracy = 96.7%, average precision = 93.6%, average recall = 93.6%.



**Figure 2.14 A-F.** Representative examples of the results obtained from the application of the SuperHistopath pipeline in whole-slide images of tumours (5x) of the Cancer Genome Atlas (TCGA) melanoma dataset (**G**, Magnified regions of interest). Note the important clinically-relevant phenotypes characterized by clusters of lymphocytes infiltrating the tumour in samples **B** and **D** or the majority of clusters of lymphocytes residing just outside the tumour area (left and central part) with only a few clusters infiltrating the tumour (right part) in sample **C**.

### 2.5.3.1.2 *Breast cancer*

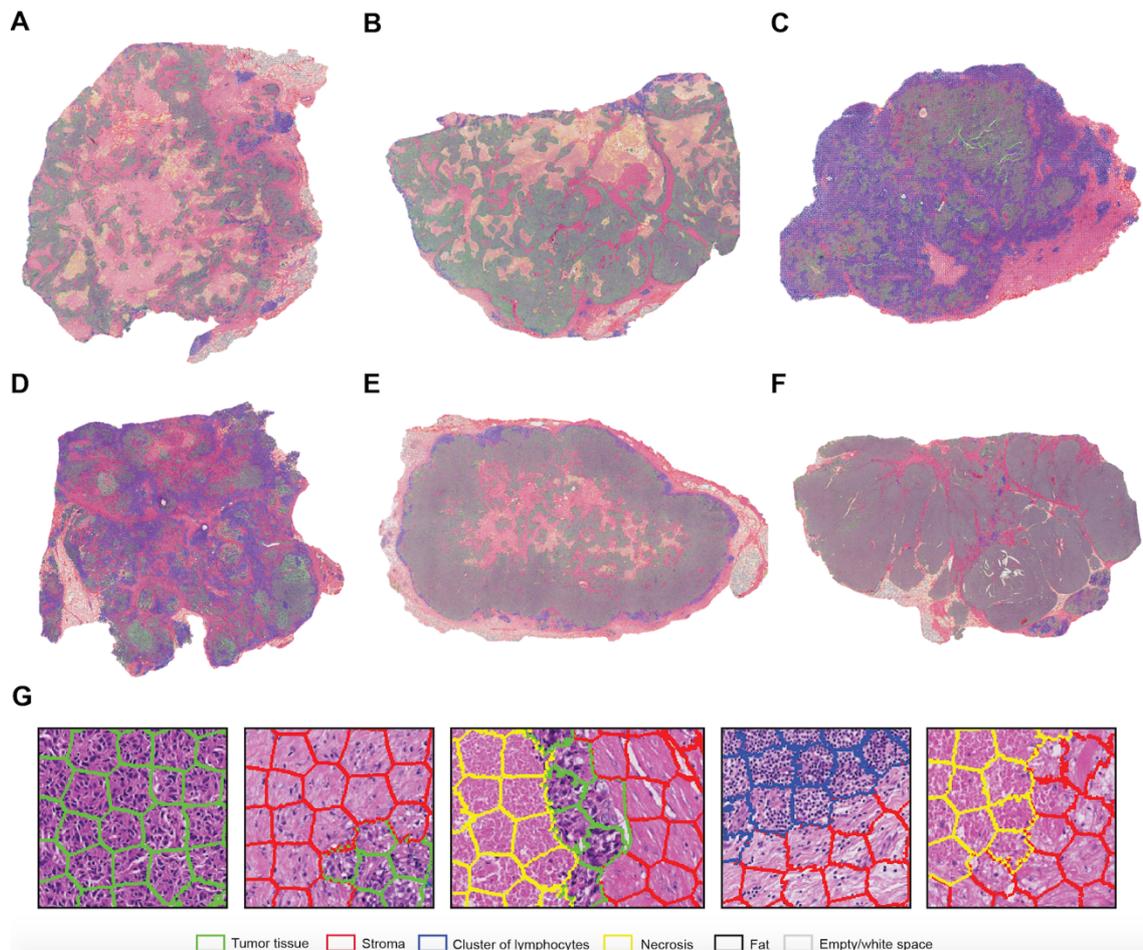
SuperHistopath classified sample regions into 6 predefined tissue categories of interest in triple-negative breast cancer, which is an aggressive disease that is negative for oestrogen receptors, progesterone receptors, and excess HER2 protein: tumour, necrosis, stroma, cluster of lymphocytes, fat, and lumen/empty space with an overall accuracy of 93.1%, an average precision of 93.9%, and an average recall of 93.6% using Xception and 91.7%, 92.5%, 91.8% respectively using our custom-made CNN over 10349 superpixels in the independent test set of 5 images (Table 2.18, Table 2.19). Figure 2.15 shows qualitative results our approach's regional classification in representative triple-negative breast cancer WSIs.

	Tumour	Necrosis	Lym	Stroma	Fat	Empty space
Tumour	1830	13	15	42	0	0
Necrosis	50	1446	2	320	0	0
Lym	4	2	705	10	0	0
Stroma	42	120	20	3836	0	1
Fat	0	0	0	0	562	5
Empty space	0	0	0	0	67	1257

**Table 2.18** Confusion matrix of the classification of superpixels using the optimized Xception network in triple-negative breast cancer patients in 6 categories: tumour, necrosis, cluster of lymphocytes (Lym), stroma, fat and lumen/empty space (separate test set of 5 whole-slide images). Overall accuracy = 93.1%, average precision = 93.9%, average recall = 93.6%.

	Tumour	Necrosis	Lym	Stroma	Fat	Empty space
Tumour	1828	35	3	34	0	0
Necrosis	89	1365	10	354	0	0
Lym	5	0	701	15	0	0
Stroma	60	92	23	3843	0	1
Fat	0	2	4	18	538	5
Empty space	0	1	0	6	71	1246

**Table 2.19** Confusion matrix of the classification of superpixels using the custom-made CNN in triple-negative breast cancer patients in 6 categories: tumour, necrosis, cluster of lymphocytes (Lym), stroma, fat and lumen/empty space (independent test set of 5 whole-slide images). Overall accuracy = 92%, average precision = 92.7%, average recall = 92.2%.



**Figure 2.15 A-F.** Representative examples of the results obtained from the application of the SuperHistopath pipeline in whole-slide images of tumours (5x) of the triple-negative breast cancer (**G**. Magnified region of interest). Note the important clinically-relevant features, such as the amount of tumour necrosis inside tumours **A** and **B**, lymphocytes which, are infiltrating the tumour in large number in samples **C** and **D**, but are surrounding the stroma barrier without infiltrating the tumour in samples **A**, **B**, **E**, **F**.

### 2.5.3.1.3 *Neuroblastoma*

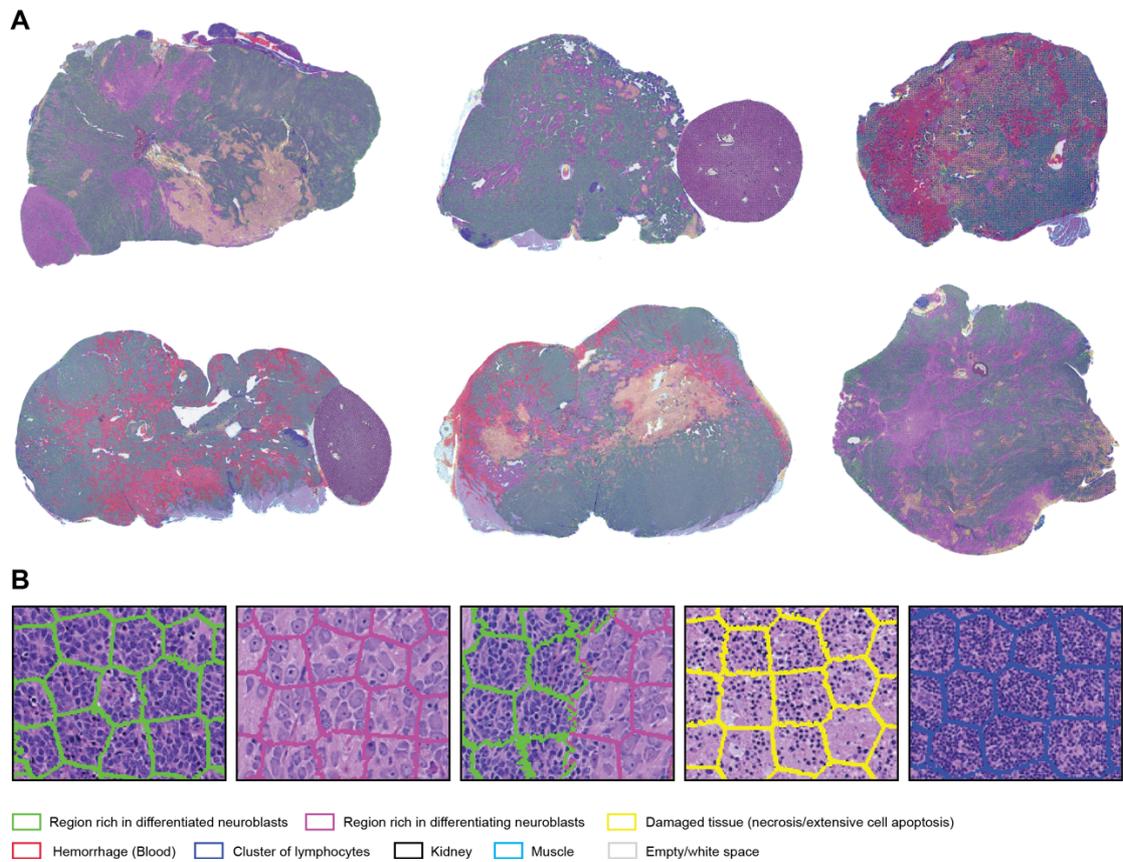
SuperHistopath classified the tumour regions into 8 predefined tissue categories of interest: undifferentiated neuroblasts, tissue damage (necrosis/apoptosis), areas of differentiation, cluster of lymphocytes, haemorrhage, muscle, kidney, and empty/white space with an overall accuracy of 98.3%, an average precision of 98.5%, and an average recall of 98.4% using Xception and 96.8%, 97.1%, 97.2% respectively using our custom-made CNN over 9868 superpixels in the independent test set of 16 images (Table 2.20, Table 2.21). Figure 2.16 shows qualitative results of our approach's regional classification in representative WSIs of neuroblastoma arising in the Th-*MYCN* mouse model.

	Undifferentiated region	Necrosis	Lym	Differentiation	Blood	Empty space	Muscle	Kidney
Undifferentiated region	1403	3	0	14	1	0	0	0
Necrosis	13	1642	1	26	49	2	5	18
Lym	6	5	1150	0	0	0	0	3
Differentiation	0	0	0	1261	0	0	0	0
Blood	1	7	0	0	1327	0	9	0
Empty space	0	2	0	0	0	560	3	2
Muscle	0	2	0	0	1	0	1176	0
Kidney	0	0	0	0	0	0	0	1176

**Table 2.20** Confusion matrix of the classification of superpixels using the optimized Xception network in the Th-*MYCN* and Th-*ALK<sup>F1174L</sup>/MYCN* mouse models in 8 categories: region of undifferentiated neuroblasts, necrosis, cluster of lymphocytes (Lym), haemorrhage (blood), empty/white space, muscle tissue and kidney (separate test set of 16 whole-slide images). Overall accuracy = 98.3%, average precision = 98.5%, average recall = 98.4%.

	Undifferentiated region	Necrosis	Lym	Differentiation	Blood	Empty space	Muscle	Kidney
Undifferentiated region	1419	1	0	1	0	0	0	0
Necrosis	33	1566	2	76	60	1	3	15
Lym	46	4	1114	0	0	0	0	0
Differentiation	18	0	0	1240	1	0	0	2
Blood	0	5	0	7	1330	0	2	0
Empty space	0	4	0	0	0	545	15	3
Muscle	0	0	0	9	1	0	1166	3
Kidney	0	1	0	0	0	0	0	1175

**Table 2.21** Confusion matrix of the classification of superpixels using the custom-made CNN in the Th-*MYCN* and Th-*ALK<sup>F1174L</sup>/MYCN* mouse models in 8 categories: region of undifferentiated neuroblasts, necrosis, cluster of lymphocytes (Lym), haemorrhage (blood), empty/white space, muscle tissue and kidney (independent test set of 16 whole-slide images). Overall accuracy = 96.8%, average precision = 97.2%, average recall = 97.1%.



**Figure 2.16** Representative examples of the results obtained from the application of the SuperHistopath pipeline in whole-slide images of tumours (5x) arising in genetically-engineered mouse models of high-risk neuroblastoma (**B**. Magnified region of interest).

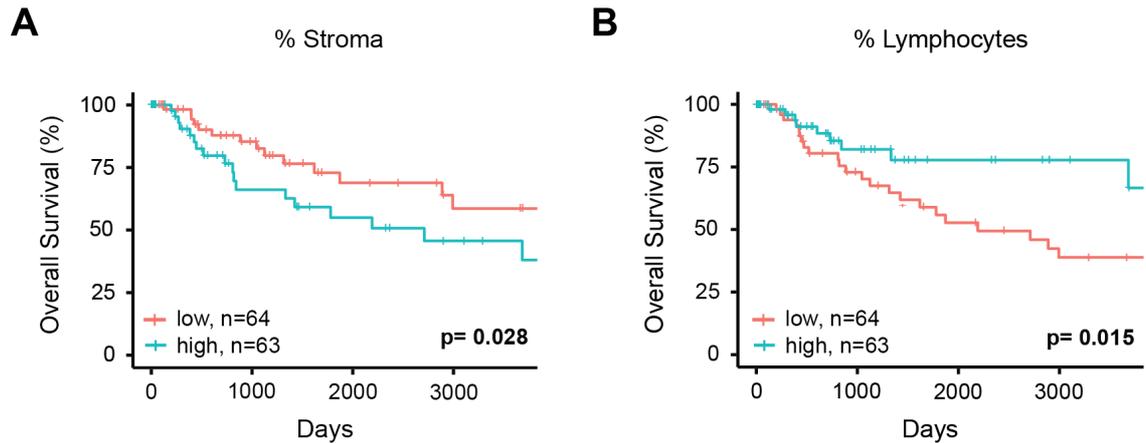
### 2.5.3.2 SuperHistopath pipeline for the analysis of low-resolution WSI affords significant speed advantages

The average time for the SLIC superpixels algorithm to segment a WSI in 5x magnification was < 2 min using a 3.5 GHz Intel core i7 processor. The average time for both the Xception and our custom-made CNN network to classify every superpixel in the images was 1-2min using the same processor. A quick convergence of the networks (around epoch 30) was observed in all cases, which needed ~3 hours for Xception and only ~30 mins for our custom-made CNN using a Tesla P100-PCI-E-16GB GPU card, and therefore the latter was used for experimenting.

### 2.5.3.3 SuperHistopath can provide robust quantification of clinically relevant features

#### 2.5.3.3.1 *Stroma-to-tumour ratio and clusters of lymphocytes abundance as predictive markers of survival in melanoma.*

Firstly, SuperHistopath is used to quantify both the stroma-to-tumour ratio and the immune infiltrate, which have both shown to provide prognostic and predictive information in patient with solid tumours, including melanoma (169, 180, 181). The important role of immune hotspots has been established based on density analysis of single cell classification of lymphocytes in high-resolution images (22, 23). Here, in the melanoma dataset of 127 WSIs i) a high stromal ratio as identified in low resolution WSIs is a predictor of poor prognosis (SuperHistopath:  $P = 0.028$ , Coxph-Regression [discretized by median]:  $HR = 2.1$ ,  $P = 0.0315$ ; Figure 2.17A) and ii) clusters of lymphocytes hold predictive information, with a high lymphocyte ratio being an indicator of favourable prognosis (SuperHistopath:  $P = 0.015$ , Coxph-Regression [discretized by median]:  $HR = 0.4$ ,  $P = 0.018$ ; Figure 2.17B). Pearson's correlation showed no significant correlation between stromal ratio and clusters of lymphocytes ratio ( $r = -0.13$ ,  $P = 0.13$ ), and between absolute sizes of stroma and clusters of lymphocytes ( $r = 0.13$ ,  $P = 0.11$ ). Taken together, the data captured from low resolution (5x) WSIs, are consistent with those extracted from single-cell analysis in high-resolution WSIs in section 2.4.3 (211).



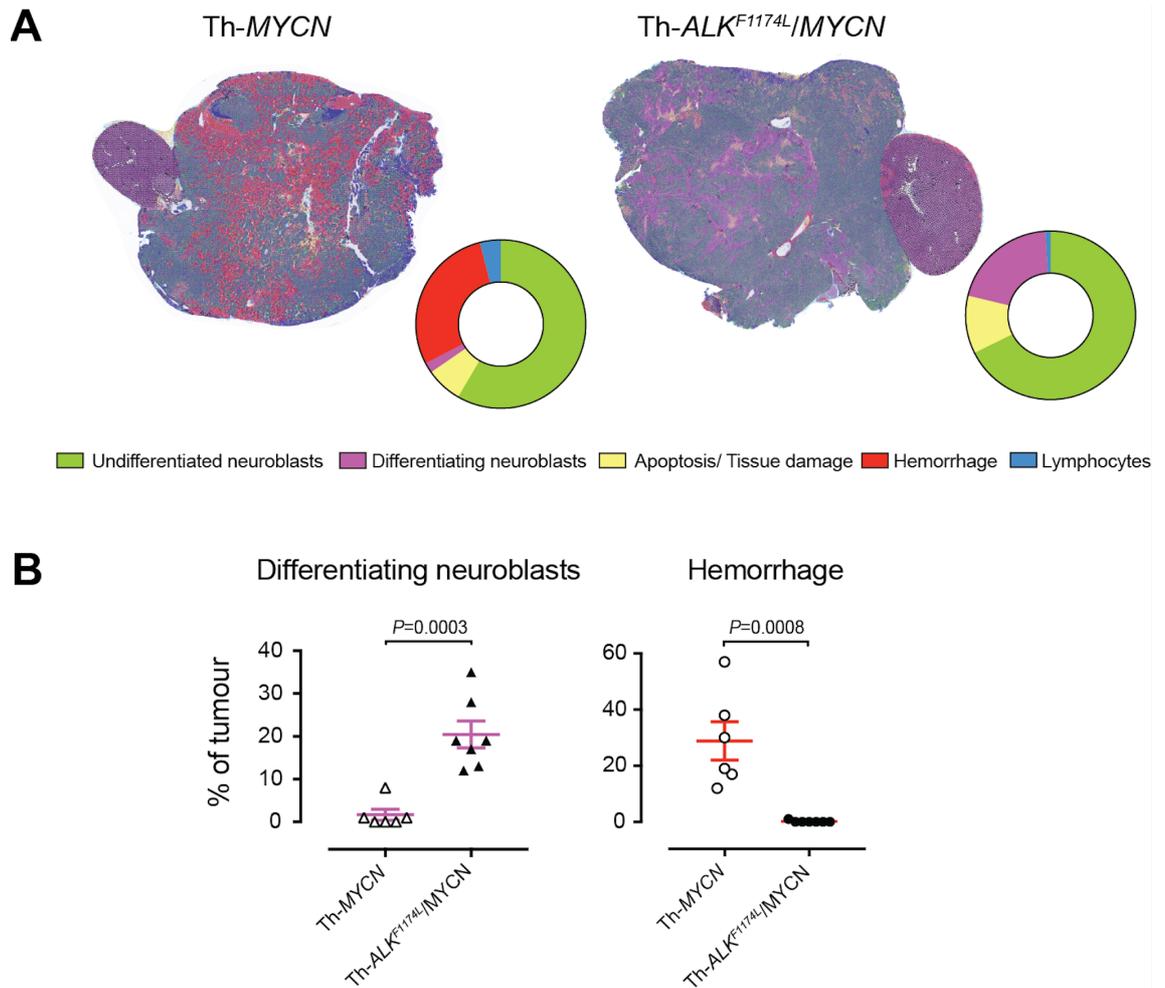
**Figure 2.17** Associations between survival outcomes and SuperHistopath-defined risk groups in the Cancer Genome Atlas (TCGA) cohorts of patients with melanoma. **A.** Kaplan-Meier Survival curves for patients in the high-risk group (blue) and low risk group (red) classified by stromal cells ratio derived from SuperHistopath and **B.** Kaplan-Meier Survival curves for patients in the high-risk group (blue) and low risk group (red) classified by immune infiltrate based on lymphocytes cluster ratio derived from SuperHistopath.

#### 2.5.3.3.2 *Necrosis quantification.*

SuperHistopath is used to quantify tumour necrosis in the breast cancer and childhood neuroblastoma preclinical datasets. Tumour necrosis, defined as confluent cell death or large area of tissue damage, holds predictive and prognostic information, both at diagnosis and after chemotherapy, in many solid tumours including breast cancer and childhood malignancies (183-185, 212, 213). Whilst visible at 5x objective lens magnification, its quantification can often be a challenging task even for experienced pathologists. Here, SuperHistopath provides satisfactory quantification of necrosis in clinical breast cancer samples by distinguishing it from stroma with high specificity (91.2%) and satisfactory precision (79.5%), and in the high-risk neuroblastoma mouse models with both high precision and specificity (93.5% and 98.9% respectively).

#### 2.5.3.3.3 Quantification of neuroblastoma differentiation.

SuperHistopath is used to quantify the phenotype of *MYCN*-driven transgenic mouse models of high-risk stroma-poor neuroblastoma. SuperHistopath can identify areas of differentiation, a critical feature for the stratification of children neuroblastoma, with both high precision and specificity (100% and 96.9% respectively). SuperHistopath also showed that expression of *ALK<sup>F1174L</sup>* mutation significantly shift the *MYCN*-driven phenotype from poorly-differentiated and haemorrhagic phenotype (Th-*MYCN*:  $1.8 \pm 1.3\%$  differentiating area and  $29.2 \pm 6.7\%$  haemorrhage, Figure 2.18A) into a differentiating phenotype also characterised by the almost complete abrogation of the haemorrhagic phenotype (Th-*ALK<sup>F1174L</sup>/MYCN*:  $20.3 \pm 3.1\%$  differentiating area and  $0.2 \pm 0.1\%$  haemorrhage,  $P = 0.0003$  and  $P = 0.0008$  respectively, Figure 2.18B) as previously demonstrated (43, 214).



**Figure 2.18** SuperHistopath-based quantification of tumour phenotype in genetically-engineered mouse model of high-risk neuroblastoma. **A.** Representative SuperHistopath-segmented whole-slide images (5x) and pie chart showing the SuperHistopath quantified mean composition of the tumours arising in Th-MYCN (n=6) and Th- $ALK^{F1174L}/MYCN$  (n=7) mouse models of high-risk neuroblastoma. Note the marked difference of phenotype induced by the expression of the  $ALK^{F1174L}$  mutation characterized by **B.** a significantly increased neuroblastoma differentiating neuroblasts and the total abrogation of the characteristic haemorrhagic phenotype of Th-MYCN tumours.

#### *2.5.4 Discussion*

In this section of chapter 2, I implemented SuperHistopath: a digital pathology pipeline for the classification of tumour regions and the mapping of tumour heterogeneity from low-resolution H&E-stained WSIs, which was demonstrated to be highly accurate in three types of cancer. Combining the application of the SLIC superpixels algorithm directly on low magnification WSIs (5x) with a CNN architecture for the classification of superpixel contributes to SuperHistopath computational efficiency allowing for fast processing, whilst affording the quantification of robust and easily interpretable clinically-relevant markers.

Applying this computational approach on low-resolution images leads to markedly increased processing speed, for both the classification of new samples and network training. Here, the (5x) magnification was chosen as a compromise between tumour structures visibility and computational cost. Specific metrics such as stroma-to-tumour ratio could potentially be derived from images at even lower magnifications (eg. 1.25x) as shown in chapter sections 2.3 and 2.4 (211). Digital histology images are conventionally processed at 40x (or 20x) magnification where cell morphology is most visible. At those resolutions, WSIs are large (representative size at 20x: 60000x60000 pixels), requiring of a lot of memory and images to be divided into patches (tiles) for processing. Under these conditions, the training of new networks for cell segmentation and classification typically requires days and the application to new WSI samples can take hours prior to code optimization. In contrast, the training our neural network until acceptable convergence needed ~30mins and application on new samples ~5mins (for both superpixel segmentation and classification) in this study. High-resolution images are essential when studying cell-to-cell interactions, however we show that the processing of low resolution images is appropriate for the extraction of specific global context features.

Furthermore, SuperHistopath combines the main advantages of regional classification and segmentation approaches. On one hand, classification approaches applied on smaller patches resulting from splitting WSIs allow the

use of CNN for the robust classification of many categories necessary to capture intra-tumour heterogeneity (126), yet at the expense of higher risk of misclassification, especially close to regional boundaries where an image patch, regardless of its size, may contain multiple tumour components. Overlapping (sliding) window approaches can improve the issue, yet at an increased computational cost. On the other hand, segmentation approaches such as U-Net-like architectures can resolve the regional boundaries issue but appear to work better for few classes, typically two. SuperHistopath efficiently combines the use of a segmentation approach using superpixels to adhere to region boundaries with CNN classification to cover the rich tumour histological heterogeneity (here 6-8 region categories depending on the cancer type).

This method also markedly simplifies and accelerates the process of preparing ground-truth (annotations) datasets as *i)* the use of superpixels alleviate the need for careful boundary delineation of the tumour components of interest (Figure 2.12B), a cumbersome and time-consuming process necessary for using U-Net-like architectures and *ii)* each annotated region contains large numbers of superpixels facilitating the collection of the large datasets traditionally required by deep learning methods.

Many promising computational pathology-derived biomarkers ultimately fail to translate in the clinic due to their inherent complexity and the difficulty for pathologists to evaluate them in new datasets. In this proof-of-concept study, I showed that SuperHistopath can quantify well-understood features/markers already used, albeit only qualitatively or semi-quantitatively, by pathologists, including the stroma-to-tumour ratio, lymphocyte infiltration, tumour necrosis, and neuroblastoma differentiation. SuperHistopath-derived results also corroborated those obtained from single-cell analysis on high-resolution samples in chapter section 2.4 (211). The computational efficiency of SuperHistopath, combined with the simple superpixels-enabled data collection, could facilitate its adoption in the clinic to accelerate pathologist workflow, could assist in intra-operative pathological diagnosis and should facilitate working with large datasets in clinical research.

Moving forward, I plan to expand the types of global context features extractable from SuperHistopath in more cancer types. The accuracy of SuperHistopath will also be evaluated on digitized frozen tissue sections to demonstrate its potential to assist in the rapid intra-operative pathological diagnostic. The previous framework (SuperCRF) which incorporates region classification information to improve cell classification will be updated by using SuperHistopath in the place of the classical machine learning framework. Together both SuperHistopath and SuperCRF would provide invaluable tools to study spatial interactions across length scales to provide a deeper understanding of the cancer-immune-stroma interface, key to further unlock the potential of cancer immunotherapy (177).

In this proof-of-concept study, the proposed method is applied to three cancer types with disparate histology without any changes (just retraining). While the approach could thus be virtually extended to any type of cancer, improvements could be made tailored to a specific global feature, cancer type or dataset and could include further exploring *i)* the use of SVM to combine the CNN-extracted features with handcrafted ones, *ii)* the use of other image colour spaces which has been shown to improve classification in certain cases (215), and *iii)* alternative superpixel algorithms such as the efficient topology preserving segmentation (ETPS) algorithm (216). Additionally, further improvement of this proof-of-concept framework could be sought via experimentation with hyperparameter tuning, or the use of other custom and well-established architectures (175, 217). Since superpixels only capture small homogeneous areas, combination with other approaches such as classification of larger image patches with a deepCNN or U-net-like architectures might be more appropriate for the single purpose of segmenting some large and multi-component tumour structures, e.g. certain types of glands (188).

### *2.5.5 Conclusion*

To conclude, the novel pipeline, SuperHistopath can accurately classify and map the complex tumour heterogeneity from low-resolution H&E-stained histology images. The resulting enhanced speed for both training and application (~5mins for classifying a WSI and ~30 mins for network training) and the efficient and simple collection of ground-truth datasets make SuperHistopath particularly attractive for research in rich datasets and would facilitate its adoption in the clinic to accelerate pathologist workflow in the quantification of predictive/prognosis markers derived from global features of interest.

## **Chapter 3 : Development of a MRI-histopathology cross-validation pipeline**

### **3.1 Introduction**

Imaging biomarkers must undergo stringent validation before they can be deployed clinically, (99) a process that can often only be meaningfully done in animal models of cancer. However, histological validation is not a trivial task. Registering the 3D MRI volume to the 2D histology slides is difficult, due to low out-of-plane resolutions, histology tissue deformations, unknown orientations, and lack of details or landmarks in the MRI images (218).

An elegant solution to this problem is to create 3D patient-specific moulds and use a microtome to cut the tumours relatively to the MRI slices (219). This procedure is especially popular with the prostate (220, 221), while successful attempts have been made with renal tumours (222) and breast cancer mouse models (65).

The primary abdominal tumours in the Th-*MYCN* mice, however, present with multiple para- and intra-tumoural anatomical landmarks detectable by conventional MRI. Thus, a 2D approach of manually selecting the best corresponding and most closely-aligned MRI slice to the histology was preferred (223, 224).

#### *3.1.1 Aim of the chapter*

This methodology chapter describes the development of a framework to accurately compare multi-parametric MRI with corresponding histology in the GEM models of neuroblastoma which exploit the numerous abdominal peri- and intra-tumoural landmarks visible on conventional T<sub>2</sub>-weighted MRI.

## 3.2 MRI-histology alignment

### 3.2.1 *Materials and Methods*

#### 3.2.1.1 Animals.

All experiments were performed in accordance with the local ethical review panel, the UK Home Office Animals (Scientific Procedures) Act 1986, the United Kingdom National Cancer Research Institute guidelines for the welfare of animals in cancer research (225).

All MRI studies were performed on a 7T Bruker horizontal bore MicroImaging system (Bruker Instruments, Ettlingen, Germany) using a 3cm birdcage volume coil. Anesthesia was induced by an intraperitoneal 5ml/kg injection of a combination of fentanyl citrate (0.315mg/ml) plus fluanisone (10mg/ml) (Hypnorm, Janssen Pharmaceutical, Oxford, UK) and midazolam (5mg/ml) (Roche, Welwyn Garden City, UK) and water (1:1:2). Core temperature was maintained at ~37°C with warm air blown through the magnet bore.

#### 3.2.1.2 T<sub>2</sub>-weighted MRI.

For all the studies reported in this thesis, anatomical T<sub>2</sub>-weighted transverse and coronal images were acquired from twenty contiguous 1 mm-thick slices through the mouse abdomen, using a rapid acquisition with refocused echoes (RARE) sequence with 4 averages of 128 phase encoding steps over a 3×3 cm field of view, an echo time (TE) of 36ms, a repetition time (TR) of 4.5s and a RARE factor of 8.

#### 3.2.1.3 T<sub>2</sub>-weighted MRI guided tumour excision and cutting.

These images were subsequently used to determine tumour volumes, planning the subsequent functional MRI measurements and tumour excision. At the end of the MRI protocol, the mice were sacrificed and prepared for dissection. After removal of the digestive system, coronal T<sub>2</sub>-weighted images were used to guide the positioning of a microtome blade at the imaging plane and the tumour was

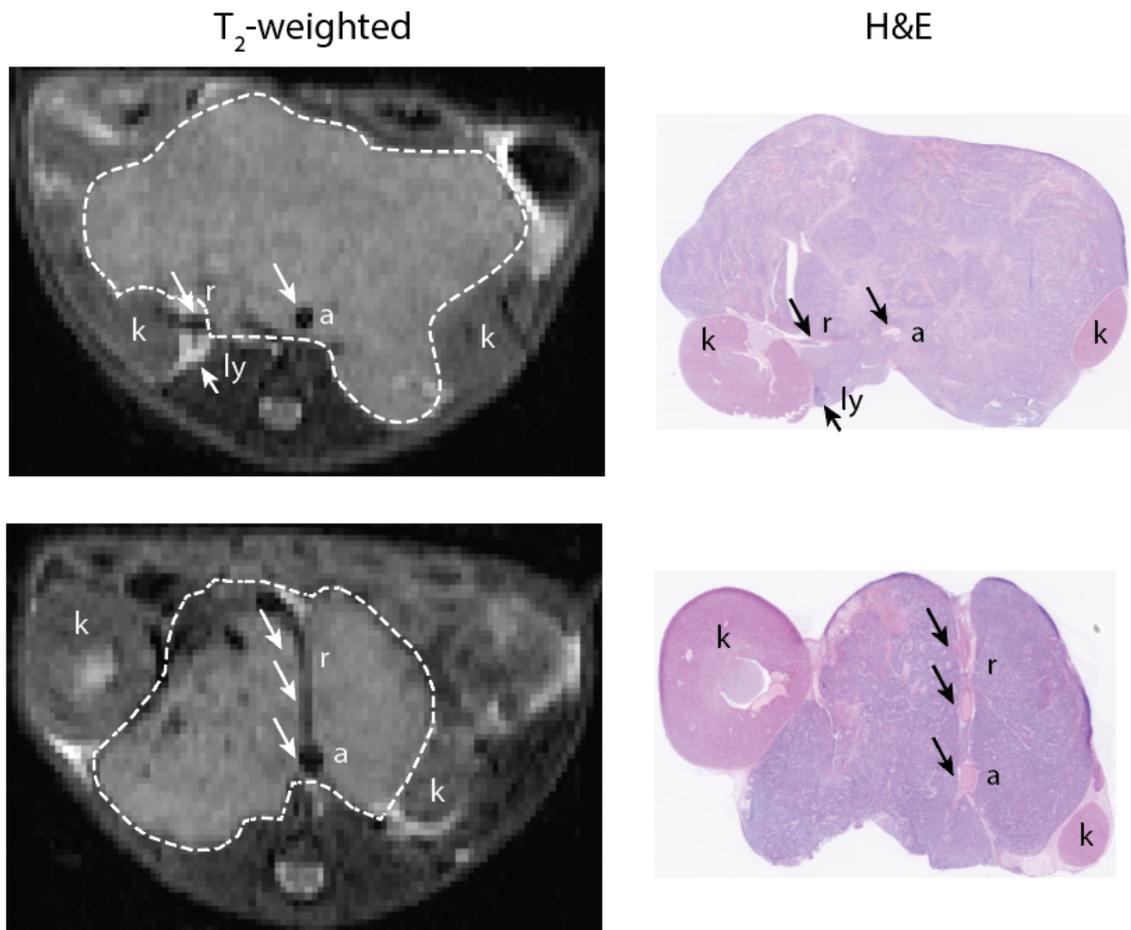
sectioned in half *in situ*. Tumours were then dissected further and freed from their spinal attachment. Tumour halves were then positioned full-face down in separate histological cassettes and placed in 10% formalin (Sigma-Aldrich, Gillingham, UK) for 24 hours.

#### 3.2.1.4 Sample processing and slide digitization.

Formalin-fixed and paraffin-embedded tumours (FFPE) were cut in 3µm sections full face from the cut edge and were stained with H&E and/or an IHC marker (on consecutive slices) at the ICR Breast Cancer Now histopathology core facility. Whole-slide images were digitized using a Hamamatsu NanoZoomer XR scanner (20x magnification, 0.46µm resolution, Hamamatsu, Japan).

#### 3.2.1.5 Supervised visual correspondence and identification of slices of interest.

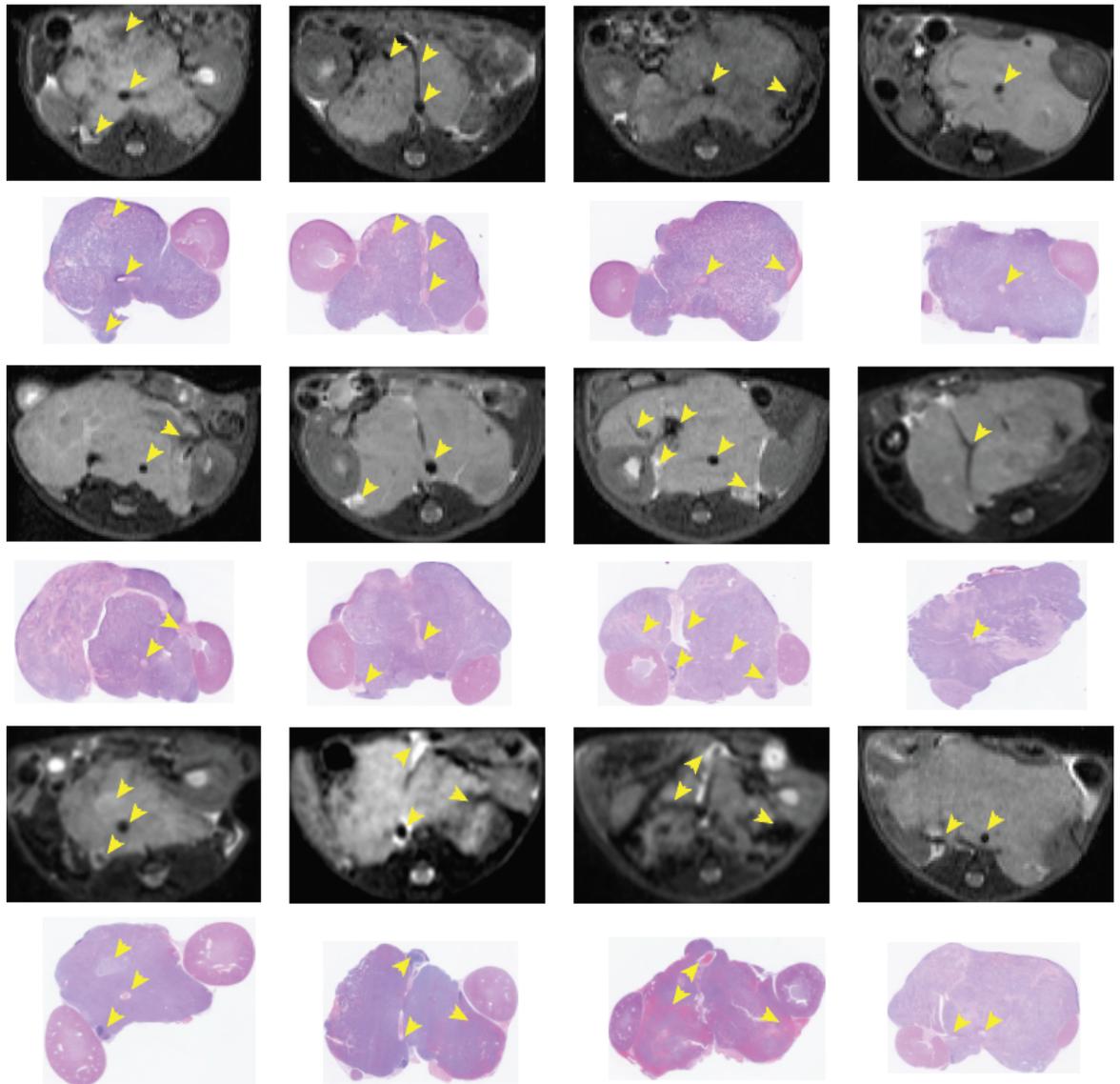
The MRI slice(s) of interest was identified and visually aligned with digitized whole-slide H&E stained images using anatomical landmarks from T<sub>2</sub>-weighted images, including the shape of the tumour, the position of the kidneys, the position and/or orientation of tumour- displaced, and circumferentially surrounded abdominal aorta and renal aorta/vein and the presence of lymph nodes (Figure 3.1).



**Figure 3.1** Representative  $T_2$ -weighted images through the abdomen of two neuroblastoma-bearing Th-MYCN mice showing the numerous anatomical landmarks detected by conventional anatomical  $T_2$ -weighted MRI (abdominal aorta: a, renal artery: r, lymph node: ly, kidney: k) which can be used in addition to tumour shape for the precise alignment of MRI with H&E- stained whole-slide high-resolution histology.

### 3.2.2 Results

The  $T_2$ -weighted MRI-guided excision method led to satisfactory spatial correspondence histopathological images and  $T_2$ -weighted images in **75 %** of the mice (n=53). Representative examples are shown in Figure 3.2.



**Figure 3.2** Representative examples of satisfactory correspondence between T<sub>2</sub>-weighted MRIs with their corresponding histopathological slices from tumours arising in Th-MYCN and Th-ALK<sup>F1174L</sup>/ Th-MYCN transgenic mouse models. Some of the intratumoural features under investigation (e.g. the artery, haemorrhage, lymph nodes, necrotic areas) are marked (yellow arrowheads).

### 3.2.3 Challenges for MRI-histology comparison

The position of the MRI slices relatively to the kidney at necropsy and the numerous anatomical landmarks visible on H&E images, enable the MRI-pathology correspondence necessary to improve biomarker validation through the interrogation of regional variation. However, a direct and pixel-to-pixel MRI-

pathology correlation would be difficult (226) due to a number of technical problems, including:

- *Unmatched spatial resolution:* The MRI “voxels” used here have a resolution of  $0.2 \times 0.2 \text{ mm}^2$  with a slice thickness of 1mm, while the histology images have pixel resolution of  $0.46 \times 0.46 \mu\text{m}^2$  and were cut in  $3 \mu\text{m}$  sections. Thus, tissue structure heterogeneity present within the MRI slice might not be reflected in the histological data.
- *Tissue deformations:* a large amount of deformation (including a universal shrinkage) is associated with tissue fixation and processing for histology.
- *Unmatched tissue planes/image features:* Imprecise cutting angle at excision or during slide cutting, can lead to a portion of the image features captured in the MRI not being present in the histology image. As a result, in certain cases only a qualitative correlation of MRI and histopathology data was feasible.

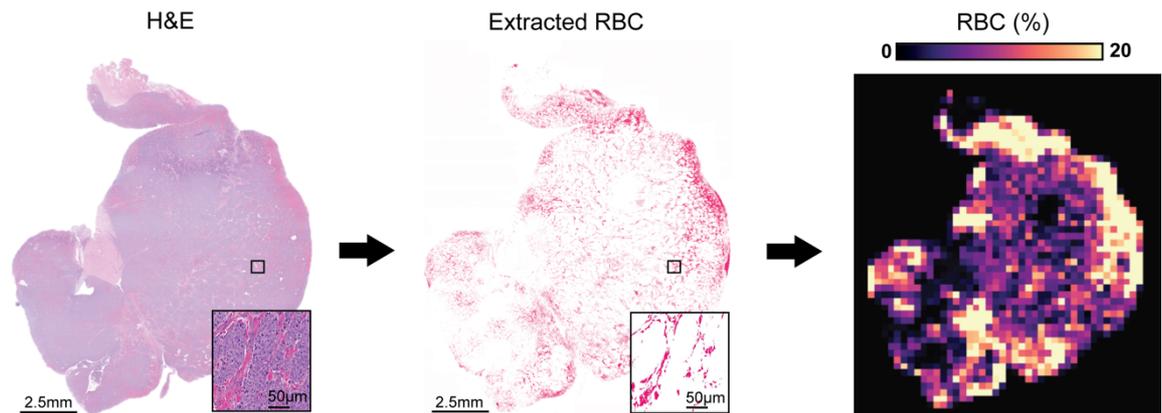
All these limitations potentially impact on our ability to use spatial statistics.

### **3.3 Matching the resolution of histological parametric maps to the MRI resolution**

Features of interest from WSI were derived from IHC or H&E, including the density of cell types derived from the cell classification algorithms described in Chapter 2. In order to ensure the quantitative and meaningful comparison between MRI and histology-derived parametric, it is essential that the resolution of the histology-derived parametric maps is equal to the resolution of MRI parametric maps.

For the Hamamatsu NanoZoomer XR, each pixel in 20x magnification has a squared resolution of  $0.46 \mu\text{m}$  ( $0.000452 \text{ mm}$ ). The MRI resolution used is  $0.234375 \text{ mm}$ , so each pixel in the MRI corresponds to  $0.234375 / 0.000452 = 518.53 = 518$  pixels approximately in the histopathological image. Simply

downscaling the histology image would cause most of its information to be lost. Instead, the extracted feature in the histology image is processed into a binary form (for example the centre of each cell nuclei would be marked as '1' and the rest as '0'). The fraction of pixels occupied by the parameter of interest was calculated within 518x518 pixel-regions. This fraction then represents a single pixel in the final image. An example of this approach is presented in Figure 3.3.

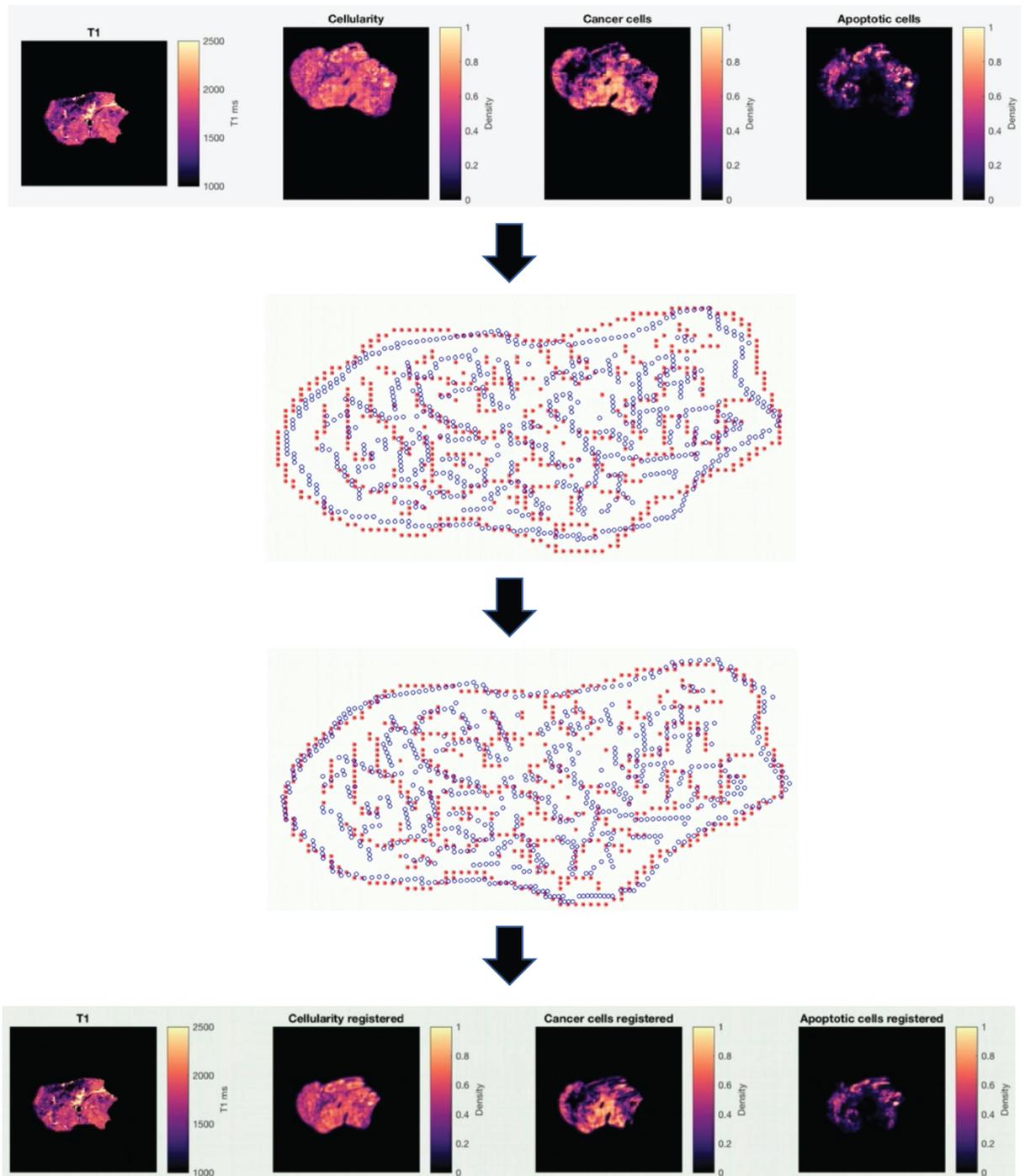


**Figure 3.3** Automatic extraction of red blood cells (RBC) from H&E images in Th-MYCN tumours. Whole-slide images of RBC were converted into binary and processed to match MRI resolution (234x234µm), with the fraction of pixels occupied by RBC within 518x518 pixel-regions representing a single pixel in the final RBC map.

### 3.4 MRI-histology registration

Subsequently, a MRI-histology registration step was incorporated into the framework to account for the numerous deformations. Amongst the many 2D registration approaches that could be applied for that purpose, the automatic point-set registration coherent point drift (CPD) algorithm was chosen as it preserves topologic structures due to the coherent motion of the point sets during registration (227). To avoid introducing any bias, the edges of each maps were extracted with a simple Canny edge detector, to use them as point sets (features). The algorithm can be applied either rigidly, allowing only for rotation, scaling and translation or non-rigidly. The types of deformations occurring in this case, primarily demanded the use of non-rigid registration. An example of the

application of the CPD algorithm, which takes only a few seconds to perform, is presented in Figure 3.4.



**Figure 3.4** Representative example of the application of the coherent point drift (CPD) algorithm. Firstly, derived-from-histology density maps of all the segmented cells (“Cellularity”) were non-rigidly registered to the MRI  $T_1$  images based on features extracted by a Canny edge detector. The same transformation is subsequently applied to the density maps of each classified cell category.

### 3.5 Discussion and conclusion

This chapter demonstrates that the use of anatomical landmarks for the excision and sectioning of tumours provide sufficient spatial correspondence for the comparison of MRI with histopathology in the GEM models of neuroblastoma. The addition of both in-plane registration and resolution matching of the histological maps to the MRI maps further enhance our ability to robustly compare the regional distribution of MRI- and histology-derived parameters. This chapter defines the ‘core’ pipeline applied in the studies described in Chapters 4-7 to evaluate and validate imaging biomarkers and habitat imaging approaches sensitive to neuroblastoma histological hallmarks and their modulation in response to targeted therapy.

The registration achieved with our pipeline will not only enhance the information provided by conventional statistical comparison of representative tumour values, (median, mean tumour) but more importantly, allow the use of more advanced spatial analysis to reflect the tumour heterogeneity and validate habitat imaging.

This is explored more specifically in the following Chapters 4, 5 and 6 where the “core” pipeline was complemented with more case-specific steps, depending on the nature and complexity of the MRI biomarker, including *i)* hotspot analysis such as kernel density estimation (KDE) algorithm to assist visual similarity comparison between MRI and histological maps or quantitative spatial comparisons approach using either *ii)* ecology-inspired algorithms, such as the non-parametric Mantel test (228), or *iii)* sub-regional analysis. In Chapter 7, the co-registered histology-derived maps were used to combine the different MRI parameters and train machine learning algorithms to characterise the heterogeneous neuroblastoma tumours.

In conclusion, the MRI-histopathology cross-validation pipeline presented in this chapter provides the close imaging-pathology correlation necessary to understand the biological processes underpinning imaging measurements in the

clinically relevant *MYCN*-driven models of high-risk neuroblastoma, and provides the stringent validation needed to support the deployment of novel MRI-based functional imaging methodologies including habitat imaging in the neuroblastoma clinic.

## **Chapter 4 : Validation of $R_2^*$ and fractional Blood Volume ( $fBV$ ) biomarkers for the imaging of the hemodynamic vasculature of neuroblastoma**

### **4.1 Introduction**

#### *4.1.1 Vascular-targeted therapies for high risk neuroblastoma*

In neuroblastoma, a high vascular index is associated with increased disseminated disease, amplification of *MYCN*, unfavourable histology, and overall poor prognosis (229). Increased microvascular proliferation and other specific vascular morphological patterns are associated with even poorer prognosis, highlighting the pivotal role of angiogenesis in determining the clinical behaviour of neuroblastoma (230, 231). The VEGF family is a key regulator of angiogenesis in neuroblastoma and high VEGF expression at the time of diagnosis is associated with poor outcome (232). Numerous antiangiogenic therapies are being evaluated in early phase paediatric clinical trials in children with solid tumours, including the anti-VEGF monoclonal antibody bevacizumab in combination with temozolomide (BEACON, NCT02308527), and the tyrosine kinase inhibitors regorafenib (NCT02085148) and pazopanib as a single agent (NCT01956669) or in combination with metronomic oral topotecan (TOPAZ), as well as the panVEGFR inhibitor cediranib currently being evaluated in children with metastatic alveolar soft part sarcoma (NCT00942877).

#### *4.1.2 Non-invasive functional imaging of tumour vasculature*

Perfusion MRI (pMRI) methodologies such as dynamic contrast enhanced (DCE) MRI using low molecular weight gadolinium chelates are often used to evaluate vascular response to VEGF signalling inhibitors in adult oncology clinical trials. However, these techniques suffer from marked measurement variability and are challenging to perform in young children (233, 234). Arterial spin labelling (ASL) MRI is an emerging and attractive contrast-free approach for the evaluation of cerebral perfusion, yet many challenges need to be overcome before its utility to assess vascular perfusion in extracranial tumours can be evaluated. Two complementary magnetic susceptibility-based MRI approaches are being actively

exploited to assess tumour vascular function and response. Tumour vasculature has been studied using (i) intrinsic susceptibility (IS) MRI, which measures the intrinsic transverse relaxation ( $R_2^*$ ) contrast produced by paramagnetic deoxyhaemoglobin within tumour blood vessels, and (ii) susceptibility contrast (SC) MRI, which relies on the intravenous administration of ultra-small superparamagnetic iron oxide (USPIO) particles. Distribution of USPIO particles within tumours causes regional increases in tumour  $R_2^*$  ( $\Delta R_2^*$ ), from which the fractional blood volume ( $fBV$ , %) can be derived (235, 236). The long intravascular half-life of USPIO particles enables steady-state acquisition and high-resolution mapping of regional variations in tumour perfusion. Both  $R_2^*$  and  $fBV$  has been shown to be sensitive imaging biomarkers of response to vascular-targeted therapies *in vivo* in the preclinical setting (41, 237, 238).

#### 4.1.3 Aim of this chapter

In this chapter, the utility of SC- and IS-MRI to inform on the functional tumour vasculature, and its response to the potent pan-VEGFR inhibitor cediranib was evaluated in tumours arising in the Th-*MYCN* GEM model of neuroblastoma. Comparison of the regional distribution of tumour  $fBV$  and  $R_2^*$  with aligned whole-slide digitized pathology stained for vessel density and red blood cell (RBC) aggregation respectively was performed via hotspot mapping and spatial statistics. This approach demonstrated how susceptibility-MRI can robustly characterize the vascular phenotype of neuroblastoma *in vivo* and help identifying that both baseline  $fBV$  and  $R_2^*$  are predictive imaging biomarkers of tumour response to cediranib. The pathologic attributes associated with the differential sensitivity to cediranib treatment in the Th-*MYCN* model are discussed and positioned against available clinical MRI findings of high-risk neuroblastoma, including initial experience with IS-MRI.

## 4.2 Materials and Methods

### 4.2.1 Animals

All experiments were approved by The Institute of Cancer Research Animal Welfare and Ethical Review Body and performed in accordance with the UK Home Office Animals (Scientific Procedures) Act 1986, the United Kingdom National Cancer Research Institute guidelines for the welfare of animals in cancer research (225) and the ARRIVE (animal research: reporting in vivo experiments) guidelines (239).

### 4.2.2 In vivo study design

A total of 68 mice were enrolled with a median tumour volume of  $801 \pm 63 \text{ mm}^3$  (median  $\pm 1$  s.e.m., ranging from 143 to  $2,055 \text{ mm}^3$ ).

Study 1. First the sensitivity of MRI to cediranib-induced acute modulation of neuroblastoma vasculature was evaluated. IS- and SC-MRI were performed prior to (day 0) and 24 hours after treatment started (day 2). Mice were treated on day 1 with 6 mg/kg of cediranib orally (obtained under material transfer agreement with AstraZeneca, n=10) or vehicle (n=8).

Study 2. In an additional cohort, (n=12) the effect of sustained daily treatment was further evaluated, with IS- and SC-MRI performed prior to, 24 hours and 7 days after daily treatment with 6 mg/kg cediranib. MRI data were not collected from two mice at the 24-hour timepoint.

Study 3. Guided by the results of Study 2, IS-MRI data from additional Th-MYCN mice (n=25) were subsequently acquired prior to daily treatment with cediranib for 7 days (bringing the total number of mice from which pre-treatment  $R_2^*$  data were acquired to 37). The volumetric response to cediranib over 7 days treatment

was monitored by T<sub>2</sub>-weighted MRI only, and compared with that from mice treated daily with vehicle (n=12).

A more detailed description of the animal experiments can be found in the Appendix.

#### 4.2.3 MRI

For all the mice, contiguous anatomical T<sub>2</sub>-weighted transverse images were acquired through the mouse abdomen for the quantification of tumour volume. The transverse relaxation rate R<sub>2</sub><sup>\*</sup> (second<sup>-1</sup>), which is sensitive to the concentration of paramagnetic species, principally deoxyhaemoglobin, was quantified prior to and following the administration of USPIO using a multiple gradient-recall echo (MGE) sequence with eight averages, and an acquisition time of 3 minutes 20 seconds. Tumour R<sub>2</sub><sup>\*</sup> maps were calculated by fitting a single exponential to the signal intensity echo time curve on a voxel-by-voxel basis using a robust Bayesian approach using in-house software (ImageView, developed in IDL; ITT Visual Information Systems). Parametric maps of tumour *fBV* (%) were subsequently calculated using the USPIO-induced change in R<sub>2</sub><sup>\*</sup> ( $\Delta R_2^*$ ).

A more detailed description of the MRI methods can be found in the Appendix.

#### 4.2.4 Computational pathology/digital histology

*Digitized histology.* Histology FFPE sections were acquired as described in Chapter 3 and stained with H&E. I digitized the whole-slide H&E images using a Hamamatsu NanoZoomer XR scanner (20x magnification, 0.46 $\mu$ m resolution, Hamamatsu, Japan).

*MRI-histology alignment.* For each tumour, the MRI slice of interest was visually aligned with the digitized whole-slide H&E stained image using anatomical landmarks as described in Chapter 3.

*RBC extraction from H&E-stained sections.* A macro was written in Fiji to extract RBC from each image tile (2000x2000 pixels) by applying colour deconvolution to extract the eosin colour channel followed by applying colour deconvolution to extract the eosin colour channel followed by the application of Otsu's automatic threshold detection method, both using ImageJ/Fiji plugins (with Java 8) (135, 240, 241). The algorithm's accuracy was tested using independent annotation of 561 RBC and 591 non-RBC points in nine samples.

*Endomucin extraction from endomucin-stained IHC sections.*

A macro was written in Fiji to extract endomucin from each tile by applying colour deconvolution to extract the brown colour channel followed by the application of maximum entropy threshold detection method, both using ImageJ/Fiji plugins (Java 8). The algorithm's accuracy was tested using independent annotation of stained/nonstained points (688/734) on nine different samples.

*Generation of RBC and endomucin parametric maps.* Whole-slide images of RBC and endomucin staining were binarised and processed to match MRI resolution (234x234 mm), with the fraction of pixels occupied by RBC or endomucin staining within 518x518 pixel-regions representing a single pixel in the final RBC map.

*MRI- and histology-derived parametric map registration.* This was performed using the CPD algorithm as described in Chapter 3. Both rigid and non-rigid registration was applied to corresponding pairs of images ( $R_2^*$  or *fBV* maps and computed maps of RBC and endomucin, respectively) in Matlab. Independently, a manual registration was also performed by rotation and scaling to match.

*Spatial comparison between MRI and histology.* Comparison of  $R_2^*$  versus RBC and *fBV* versus endomucin parametric maps was performed visually following the application the kernel density estimation (KDE) to display hotspots of high, above the 85th percentile for each tumour sample,  $R_2^*$  and RBC (%), *fBV* (%), and endomucin staining (%) values using the "MASS" package in R (22, 23).

To provide further evidence of spatial correlation, the Mantel test, a nonparametric analysis of associations between corresponding positions of two

distance matrices, was applied on registered pairs of MRI maps and computed maps of histopathologic features using the "ade4" library in R. Distance matrices were calculated using the "dist" function.

*Statistical analysis.* Statistical analysis was performed with GraphPad Prism 6 (GraphPad Software Inc.). The mean of median values for all the quantitative MRI parameters, the mean values for tumour volume, RBC, and endomucin were used for statistical analysis with a 5% level of significance. Any significant differences in tumour volume and quantitative histopathologic parameters were identified using the Student two-tailed unpaired t-test, with a 5% level of significance. Significant correlations between the mean values. All linear correlations were determined by Pearson's correlation method. Level of significance was 5%.

#### *4.2.5 Clinical investigations*

Based on the results of the preclinical investigations presented above, we aimed to *i)* evaluate the clinical relevance of the vascular phenotype observed in the Th-*MYCN* tumours by reviewing available diagnostic MRI data from patients with *MYCN*-amplified neuroblastoma, and *ii)* evaluate the feasibility of acquiring IS-MRI in children with neuroblastoma. Informed written consent was obtained from all parents or patients following IRB approval of the prospective study with IS-MRI at the Royal Marsden Hospital. For the retrospective review of routine diagnostic imaging at Great Ormond Street Hospital (GOSH), anonymized data were provided to the researchers via a collaboration with Dr. Kieran McHugh.

*Retrospective analysis of MRI scans performed at initial diagnosis.* Diagnostic imaging was acquired for patients (n=19, 13 male/6 female) with confirmed *MYCN*-amplified disease (ages 1 day to 3.6 years old) on a 1.5 T Magnetom Avanto system (Siemens Healthcare) including T<sub>2</sub>-weighted, T<sub>1</sub>-weighted imaging, prior and after the injection of Gd chelate-based contrast agent, and diffusion-weighted MRI.

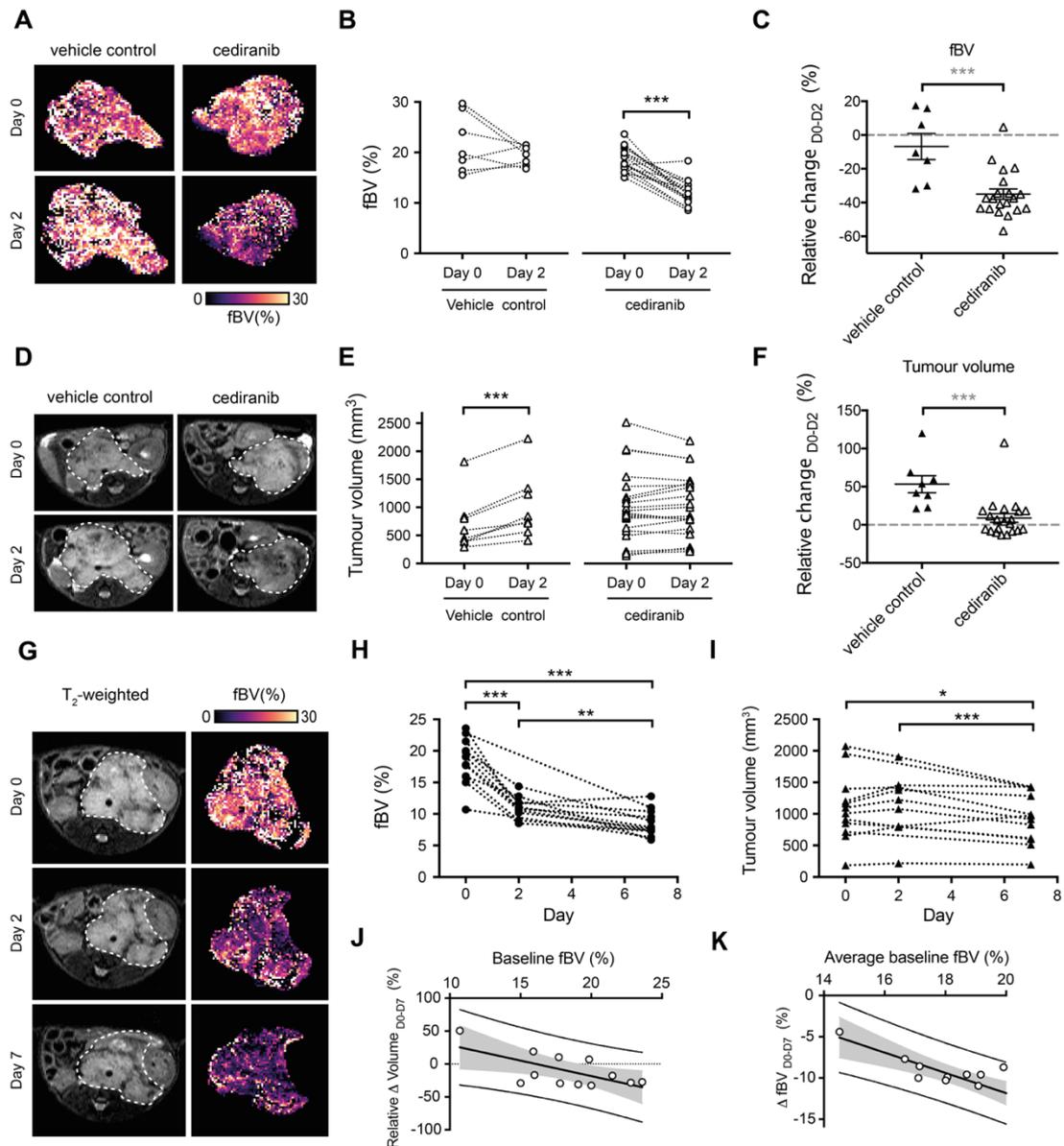
*Prospective study of the feasibility of IS-MRI in patients with refractory/relapsing neuroblastoma (Royal Marsden Hospital).* This was performed as a collaboration with Dr Neil Jerome, Dr Fernando Carceller and Prof Dow-Mu Koh. The three patients (n=2 male/1 female) with refractory/relapsing disease included in this study were scanned on a 1.5 T Magnetom Avanto system and included T<sub>2</sub>-weighted sequence, T<sub>1</sub>-weighted, prior and after (last scan of the session) the injection of Gd chelate-based contrast agent, and a diffusion-weighted MRI sequence. IS-MRI was performed in 2.5 minutes using a 2D MGE sequence. Images were analysed and R<sub>2</sub>\* quantified in Matlab.

A more detailed description of the protocols and MRI methods can be found in the Appendix.

### **4.3 Results**

#### *4.3.1 Baseline tumour fBV predicts response to cediranib in the Th-MYCN GEM model of neuroblastoma*

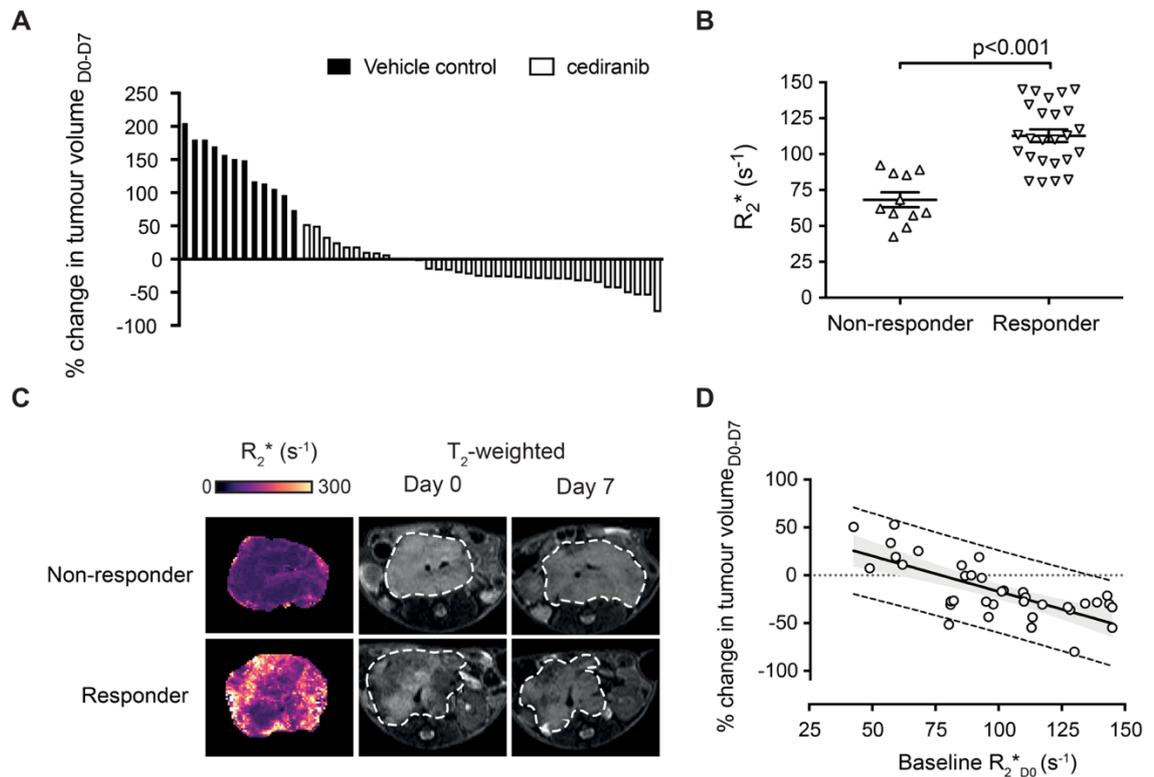
Changes in vascular function in Th-MYCN mice were first assessed following treatment with cediranib. At the time of enrolment, the mean tumour *fBV* quantified using SC-MRI in the Th- MYCN mice (n=19) was  $19.6 \pm 1\%$ , consistent with the high vascular index of neuroblastoma. Treatment with cediranib caused an acute and highly significant reduction at 24 hours in tumour *fBV* ( $\Delta fBV_{\text{cediranib}_{24\text{h}}} = -28 \pm 4\%$  vs.  $\Delta fBV_{\text{vehicle}_{24\text{h}}} = -6 \pm 8\%$ ,  $P = 0.0008$ ), which was accompanied by a highly significant cytotoxic response ( $\Delta \text{Volume}_{\text{cediranib}_{D0-D2}} = 18 \pm 8\%$  vs.  $\Delta \text{Volume}_{\text{vehicle}_{D0-D2}} = 53 \pm 11\%$ ,  $P = 0.0006$ ; Figure 4.1A–F). Sustained daily treatment over 7 days caused a further significant reduction in *fBV* and antitumour activity (Figure 4.1G–I). The average baseline *fBV* strongly correlated with the cediranib-induced reduction in *fBV* ( $r = -0.83$ ,  $P = 0.0008$ ; Figure 4.1J), and the value of *fBV* at baseline correlated with the subsequent cediranib-induced tumour volumetric response at day 7 ( $r = -0.65$ ,  $P = 0.02$ ; Figure 4.1K).



**Figure 4.1** Susceptibility-contrast MRI-derived *fBV* predicts response to VEGFR (panVEGFR) inhibitor cediranib in the Th-MYCN model of neuroblastoma. **A**. Representative tumour parametric *fBV* maps in the Th-MYCN mice prior to (day 0) and 24 hours (day 2) after treatment with 6 mg/kg cediranib or vehicle. **B** and **C**. Changes (**B**) and relative changes (**C**) in tumour median *fBV* prior to (day 0) and 24 hours (day 2) after treatment with 6 mg/kg cediranib or vehicle. **D**. Representative T<sub>2</sub>-weighted anatomical MRI taken through the abdomen of tumour-bearing Th-MYCN mice prior to (day 0) and 24 hours (day 2) after treatment with 6 mg/kg cediranib or vehicle. **E** and **F**. Changes (**E**) and relative changes (**F**) in tumour volume prior to (day 0) and 24 hours (day 2) after treatment with 6 mg/kg cediranib or vehicle. Data are the individual median value for each tumour and the cohort mean  $\pm$  1 s.e.m. (\*\*\*,  $P < 0.005$ , two-tailed paired Student t test; \*\*,  $P < 0.01$ ; \*\*\*  $P < 0.005$ , two-tailed unpaired Student t test, 5% level of significance). **G**. Representative T<sub>2</sub>-weighted anatomical MRI and tumour *fBV* parametric map of a Th-MYCN mouse prior, 24 hours, and 7 days after daily treatment with 6 mg/kg cediranib. **H** and **I**. Changes in tumour *fBV* (**H**) and tumour volume (**I**) in Th-MYCN mice prior, 24 hours, and 7 days after daily treatment with 6 mg/kg cediranib. **J**. Average tumour median *fBV* correlated with change in *fBV* over the 7 days of the trial ( $r = -0.83$ ;  $P = 0.0008$ ). **K**. Tumour median *fBV* correlated with change in tumour volume over the 7 days of the trial ( $r = -0.65$ ;  $P = 0.02$ ).

#### 4.3.2 Baseline tumour $R_2^*$ predicts response to cediranib in the Th-MYCN GEM model of neuroblastoma

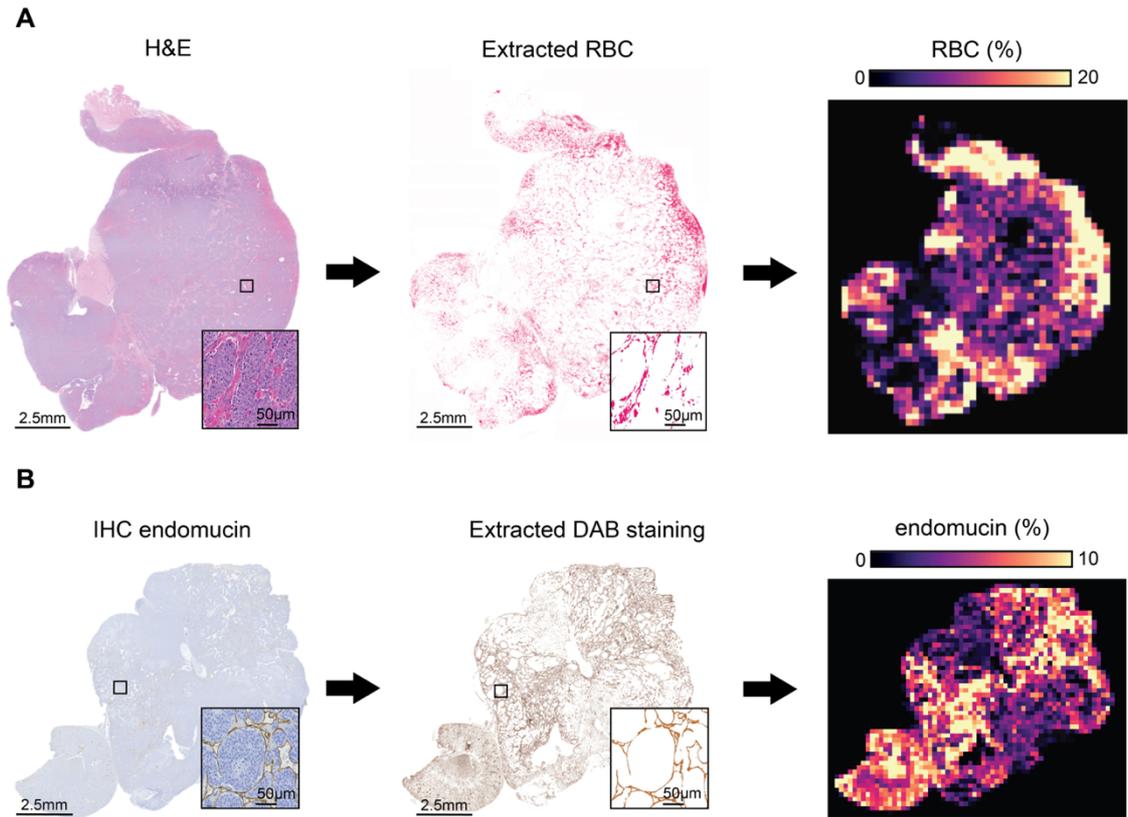
The analysis of the cohort of mice in which IS-MRI was routinely performed pre-treatment confirmed that cediranib effectively suppress the aggressive tumour growth that is typical of this model ( $DVolume_{cediranib\_D0-D7} = 16 \pm 5\%$  vs.  $DVolume_{vehicle\_D0-D7} = 142 \pm 11\%$ ,  $P < 0.0001$ ,  $n = 37$  and  $12$ ). However, a range of volumetric response, from progressive disease to partial response, was observed (Figure 4.2A). Examination of our pre-treatment  $T_2$ -weighted anatomical images and native  $R_2^*$  parametric maps revealed that responsive tumours exhibited a characteristically heterogeneous appearance, with areas of hypointense  $T_2$  signal and relatively fast  $R_2^*$  ( $R_2^* = 113 \pm 4 \text{ second}^{-1}$ ; Figure 4.2B and C), whereas progressive tumours typically demonstrated a more homogeneous, isointense appearance on  $T_2$  and significantly slower native  $R_2^*$  values ( $R_2^* = 68 \pm 5 \text{ second}^{-1}$ ,  $P < 0.0001$ ). Cediranib-induced changes in tumour volume over 7 days treatment correlated with native tumour  $R_2^*$  measured before treatment ( $r = -0.72$ ,  $P < 0.0001$ ; Figure 4.2D).



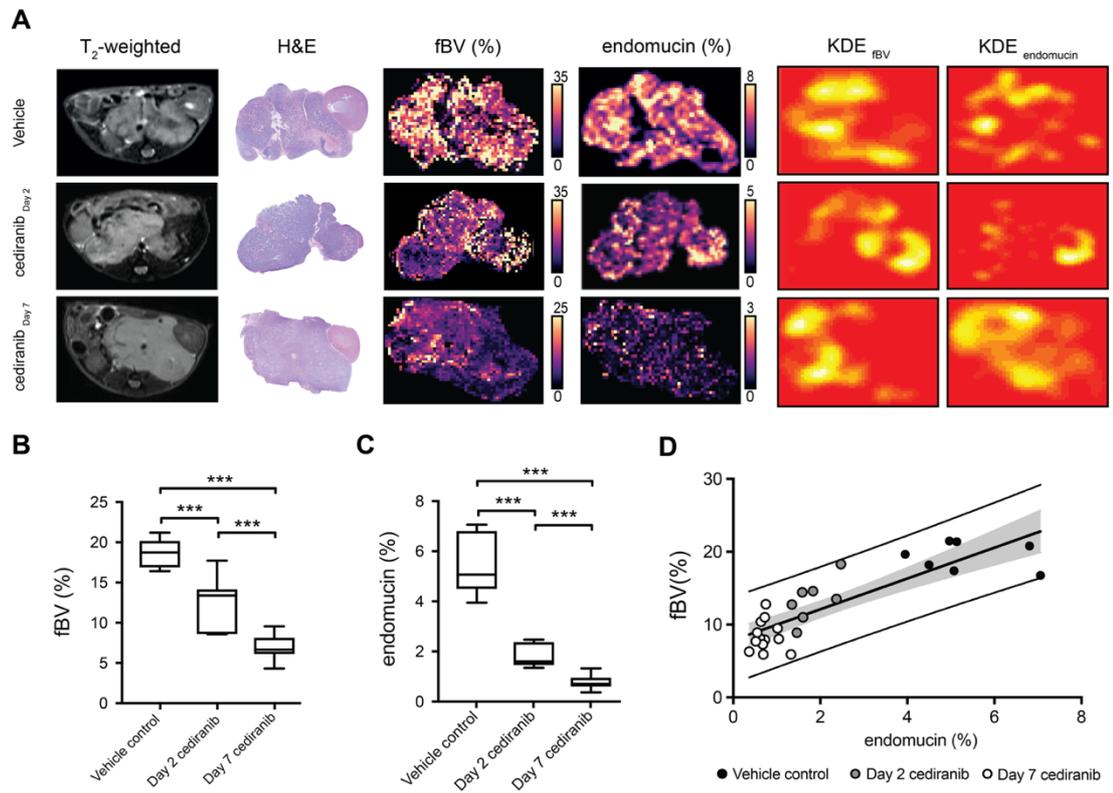
**Figure 4.2** Intrinsic susceptibility MRI-derived transverse relaxation rate  $R_2^*$  predicts response to VEGFR (panVEGFR) inhibitor cediranib in the Th-MYCN model of neuroblastoma. **A.** Waterfall plot documenting relative changes in tumour volume in the Th-MYCN mouse model of neuroblastoma following 7-day treatment with daily dose of 6 mg/kg cediranib or vehicle ( $DVolume_{cediranib\_D0-D7} = -16 \pm 5\%$  vs.  $DVolume_{vehicle\_D0-D7} = 142 \pm 11\%$ ;  $P < 0.0001$ , two-tailed unpaired Student t test with a 5% level of significance). **B.** Baseline  $R_2^*$  value (day 0) of responsive and nonresponsive tumours. Data are the individual median value for each tumour and the cohort mean  $\pm$  1 SEM ( $P < 0.0001$ , two-tailed unpaired Student t test with a 5% level of significance). **C.** Representative baseline native  $R_2^*$  maps for responsive and progressive tumours in Th-MYCN transgenic mice (assessed by anatomical MRI) following daily treatment with 6 mg/kg cediranib for 7 days. **D.** Native tumour median  $R_2^*$  (day 0) correlated with relative change in tumour volume following daily treatment with 6 mg/kg cediranib for 7 days ( $r = -0.72$ ,  $P < 0.001$ ).

#### 4.3.3 Spatial heterogeneity in *fBV* reflects regional variations in the microvasculature of neuroblastoma

The numerous anatomical landmarks (kidney, spleen, abdominal arteries, lymph nodes) clearly evident on T<sub>2</sub>-weighted MRI were used to guide the careful excision and subsequent orientation of digitized pathology with SC- and IS-MRI-derived parametric maps. Intra-tumoural regional differences in *fBV* visually reflected spatial variations in computed maps of vascular endothelial endomucin staining, automatically extracted from high-resolution IHC images with an accuracy of 97%, after automatic registration using the CPD algorithm (Figure 4.3A and B). This was corroborated by the location, size, shape, and orientation of hotspots identified on KDE maps of high values (above the 85th percentile of each tumour sample) of *fBV* and endomucin (Figure 4.4A). Application of the Mantel statistical test revealed significant ( $P < 0.05$ ) correlation ( $0.10 < r < 0.58$ ; Table 4.1) between the distance matrices of *fBV* and endomucin in 16 of 23 tumours. Endomucin staining fraction corroborated the decrease in *fBV* measured longitudinally over 7 days, with significantly lower tumour values determined in the cohort treated with cediranib for 7 days than in the cohort of animals treated with a single dose of cediranib; both cohorts demonstrated lower values of *fBV* compared with the vehicle control group (Figure 4.4B and C). Calculated tumour median values of *fBV* strongly correlated with median values of endomucin fraction area across the cohorts ( $r = 0.85$ ,  $P < 0.0001$ ; Figure 4.4D).



**Figure 4.3** Pipeline for the generation of computed maps of RBC and endomucin staining from histopathologic images in Th-*MYCN* tumours. **A.** RBCs were automatically extracted from H&E-stained whole section slides, with an accuracy of 99.74%. **B.** Endomucin staining was extracted from chromogenic immunohistochemistry whole section slides, with an accuracy of 97%. Note that the contrast of the whole slide extracted endomucin was enhanced, solely for printing visibility. Whole-section images were subsequently processed to match the MRI resolution, with the fraction of pixels occupied by RBCs (**A**) and endomucin (**B**) within 518x518 pixels of the original images, representing a single pixel in the final maps.



**Figure 4.4** Susceptibility-contrast MRI-derived *fBV* correlates with endothelial cell marker endomucin fraction in tumours of the Th-MYCN model of neuroblastoma. **A**. Representative T<sub>2</sub>-weighted anatomical MRI images and their corresponding H&E-stained histology, parametric *fBV* maps, and computed areas of vascular endothelial cell marker endomucin staining, KDE hotspot maps of high values (above the 85th percentile of each tumour sample) of *fBV* and endomucin fraction in the Th-MYCN model of neuroblastoma, 24 hours and 7 days after daily treatment with cediranib or vehicle (24 hours). Note the different scale for *fBV* and endomucin maps across the groups to better capture the intra-tumoural heterogeneity. **B** and **C**. Median tumour *fBV* (**B**) and median tumour endomucin fraction area (**C**) in the Th-MYCN GEM model 24 hours and 7 days after treatment with cediranib or vehicle (\*\*\*,  $P < 0.001$ , two-tailed unpaired Student t test). **D**. Tumour median *fBV* correlated with endomucin fraction area ( $r = 0.85$ ,  $P < 0.0001$ ).

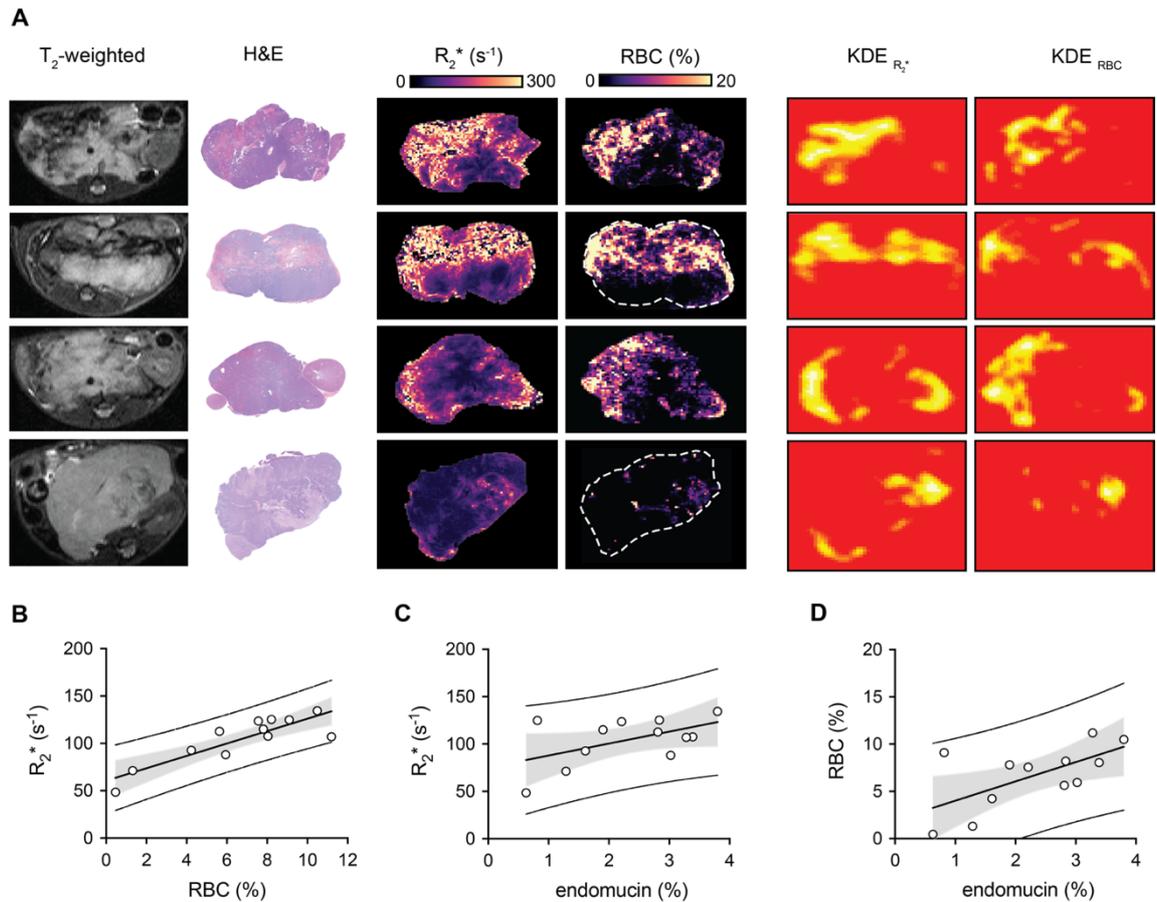
Mouse number	r	p-value	Registration type
M104252835	0.15	0.08*	Rigid
M125555	0.58	0.004	non-rigid
M125559	0.1	0.04	non-rigid
M125689	0.023	0.0007	Manual
M126140	0.33	0.003	non-rigid
M126134	0.3	0.007	Rigid
M126143	0.35	0.0001	non-rigid
M126835	-	-	
M126882	-	-	
M128379	-	-	
M128380	-	-	
M128381	-	-	
M104252874	0.14	0.002	Rigid
M104252879	-	-	
M104252883	-	-	
M104252898	-	-	
M104252901	-	-	
M104252904	0.27	0.008	non-rigid
M104252920	-	-	
M104252921	0.35	0.0001	Rigid
M104252923	0.51	0.0001	Manual
M104252941	0.22	0.0375	Rigid
M104252984	0.31	0.0012	Rigid

- r < 0.1, \*Not significant

**Table 4.1** Summary of Mantel test results showing the highest correlation of distance matrices obtained between MRI fractional blood volume *fBV* and computed area of vascular endothelial cell marker endomucin staining maps following manual, non-rigid and rigid registration.

#### *4.3.4 Regional heterogeneity in $R_2^*$ reflects variations in RBC distribution in neuroblastoma*

Regional differences in parametric  $R_2^*$  maps visually reflected spatial variations in extravasated RBC aggregation or RBC-filled necrotic areas (defined as large areas of cell damage) in the digitized maps, which was corroborated by the location, size, shape, and orientation of hotspots on KDE maps of high values (above the 85th percentile of each tumour sample) of  $R_2^*$  and RBC, automatically extracted from high-resolution H&E-stained images with an accuracy of 99.7%, after automatic registration using the CPD algorithm (Figure 4.5A). The Mantel statistical test revealed significant ( $P < 0.05$ ) spatial correlation ( $0.10 < r < 0.76$ ; Table 4.2) between the distance matrices of  $R_2^*$  and RBC in 9 of 12 tumours. Calculated tumour median values of  $R_2^*$  strongly correlated with the mean values of RBC fraction area ( $r = 0.87$ ,  $P < 0.0003$ ). Tumour median  $R_2^*$  and mean value of RBC fraction also correlated with endomucin staining ( $r = 0.53$ ,  $P = 0.08$  and  $r = 0.64$ ,  $P = 0.02$ , respectively; Figure 4.5B–D).



**Figure 4.5** Intrinsic susceptibility MRI-derived transverse relaxation rate  $R_2^*$  correlates with tumour RBC content. **A**. Representative  $T_2$ -weighted MRI images and their corresponding H&E histology, maps of the transverse relaxation rate  $R_2^*$  and computed area of RBC detected from H&E-stained images, KDE hotspot maps of high values (above the 85th percentile of each tumour sample) of  $R_2^*$ , and RBC fraction in tumours arising in control Th-*MYCN* transgenic mouse. **B**. Tumour median  $R_2^*$  positively correlated with the mean value of RBC fraction detected from H&E-stained tumours arising in the Th-*MYCN* transgenic mouse model ( $r = 0.87$ ;  $P = 0.0003$ ).  $R_2^*$  values (**C**) and mean value of RBC fraction (**D**) also correlated with endomucin staining ( $r = 0.53$ ,  $P = 0.08$  and  $r = 0.64$ ,  $P = 0.02$ ).

Mouse number	r	p-value	Registration type
M104252849	0.32	0.0002	non-rigid
M116446	0.76	0.0001	non-rigid
M116490	0.48	0.0001	Manual
M117093	0.4	0.0006	non-rigid
M117718	0.11	0.055 <sup>x</sup>	non-rigid
M117723	0.46	0.0005	non-rigid
M117729	0.69	0.0001	non-rigid
M118312	0.43	0.0002	non-rigid
M118359	0.51	0.0001	non-rigid
M118756	-	-	
M118765	-	-	
M65760	0.27	0.0006	Manual

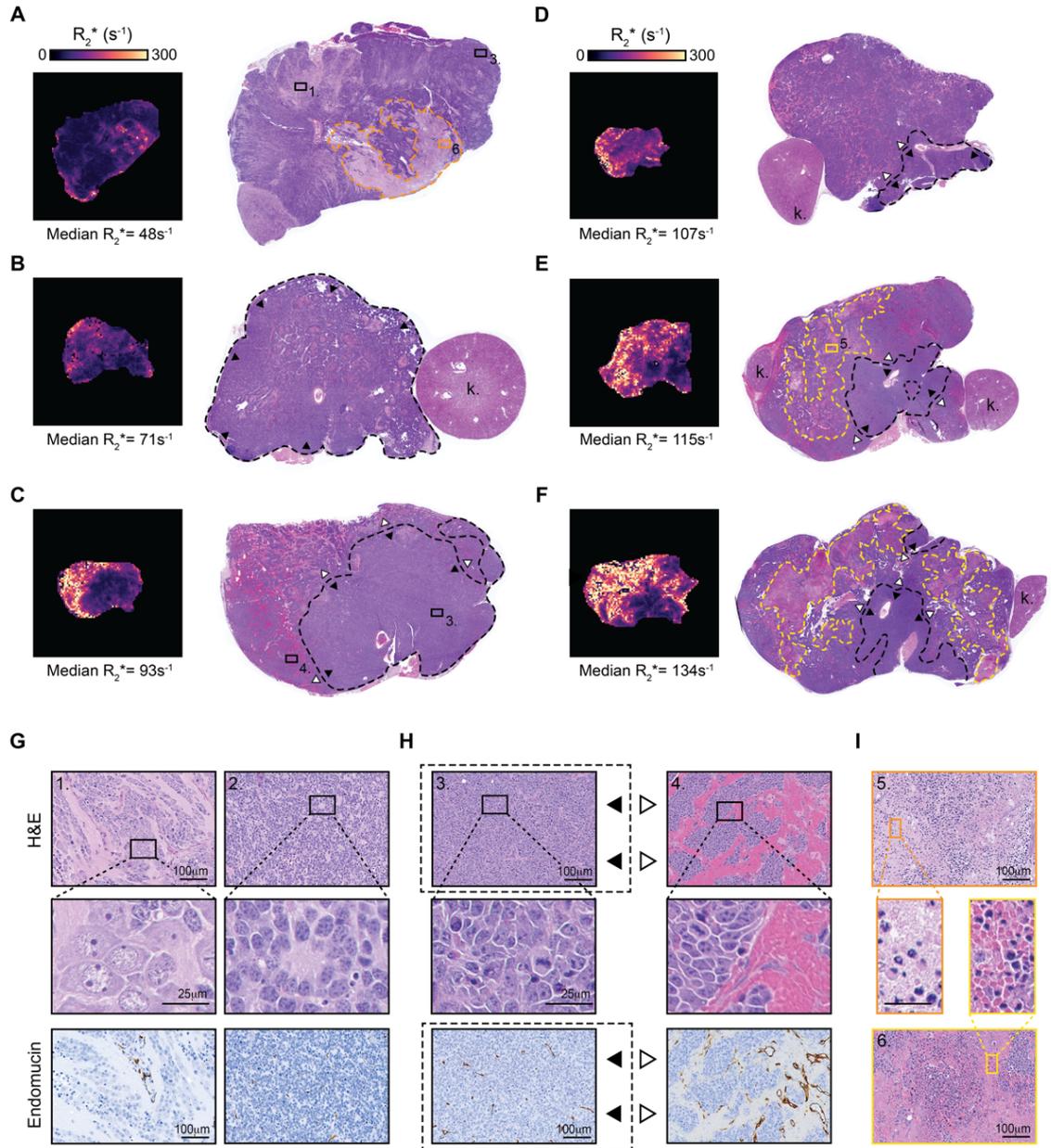
- r < 0.1, <sup>x</sup> Not significant

**Table 4.2** Summary of Mantel test results showing the highest correlation of distance matrices obtained between MRI transverse relaxation  $R_2^*$  and computed red blood cell maps following manual, non-rigid and rigid registration.

#### 4.3.5 Innate resistance to cediranib is associated with a differentiating phenotype

We then performed a semi-quantitative analysis to calibrate the tumour  $R_2^*$  phenotype against pathologic evaluation of aligned H&E stained sections (Figure 4.6A–I; Table 4.3). For tumours with a median  $R_2^*$  between ~80 and ~110  $\text{second}^{-1}$ , increasing tumour median  $R_2^*$  reflected increasing area/proportion of the tumour presenting with a haemorrhagic and hypervascular phenotype, characterized by large sinusoidal-shaped vessels. In comparison, non-haemorrhagic regions remained well-vascularized but with markedly reduced haemorrhage and smaller capillary-like shape vessels. In these tumours, the extravasated RBCs conserved their biconcave shape and appeared intact. Tumours with a median  $R_2^*$  > 110  $\text{second}^{-1}$  additionally presented with large necrotic regions filled with both intact and damaged RBCs. Tumours exhibiting a median  $R_2^*$  < 70  $\text{second}^{-1}$  (comparable to tumours that progressed while on cediranib), presented a very different phenotype, characterized by the absence of haemorrhage (including in necrotic areas), reduced microvessel density with

smaller more regular capillaries-like vessels, and large regions of differentiating neuroblasts (yet with only very few mature ganglion cells) arranged in islands, separated by a large amount of neuropil.



**Figure 4.6** Calibration of intrinsic susceptibility MRI-derived tumour  $R_2^*$  phenotype with histopathologic phenotype in the Th-MYCN genetically engineered murine model of neuroblastoma. **A–F.** Susceptibility-weighted MRI-derived maps of the tumour transverse relaxation  $R_2^*$  and their corresponding H&E-stained whole-section slides (k., kidney), arranged by increasing median tumour  $R_2^*$  value. Tumours with very low median value of  $R_2^*$  ( $< 60 \text{ second}^{-1}$ ; **A**) present with a differentiating phenotype characterized by large areas of differentiating neuroblasts (**G**), with enlarged cytoplasm and more prominent nucleoli, surrounded by large amounts of neuropil (1.), alongside dense regions of undifferentiated neuroblasts (2.). Note that haemorrhage is absent from these tumours and the vasculature is characterized by sparse capillaries-like blood vessels as revealed by IHC staining for the vascular endothelial cells marker endomucin. For a median value of  $R_2^*$  of  $\sim 70 \text{ second}^{-1}$  and above, tumours present with dense area of poorly differentiated neuroblasts presenting with numerous mitotic features (**H**), yet tumour median  $R_2^*$  is determined by the increasing ratio of region with large area of haemorrhage and high vascular density (4., ---,  $\triangleright$ ) over region with lower density of capillary like vessel (3., ---,  $\blacktriangleright$ ). Note that the extravasated RBC appear intact.

Additionally, tumours with median  $R_2^*$  over  $110 \text{ second}^{-1}$  show vast area of cellular damage (necrosis) filled with intact and damaged RBCs (I; ---, 5.). Necrotic area with low RBC content (---, 6.) only present with small regional increase in  $R_2^*$  (as seen in **A**).

Mouse number	$R_2^*$ ( $\text{s}^{-1}$ )	Histology <sup>‡</sup>	Haemorrhage	Necrosis (%)	Specific Characteristics
M116446	134	Poorly differentiated	+++	35	Large central area of necrosis with RBC
M118359	125	Poorly differentiated	+++	6	-
M117729	125	Poorly differentiated	+++	19	Large central area of necrosis with RBC surrounded by numerous foci of differentiating neuroblasts
M117093	124	Poorly differentiated	+++	N.D.	-
M118756	115	Poorly differentiated	+++	14	Large area of necrosis with RBC
M104252849	113	Poorly differentiated	+++	15	Large area of necrosis with RBC
M118765	108	Poorly differentiated	+++	N.D.	-
M116490	107	Poorly differentiated	+++	18	Large are of necrosis with RBC
M117723	93	Poorly differentiated	++	2	Haemorrhage only present on one half of tumour
M118312	88	Poorly differentiated	++	>1	10% tumour present with abundant neuropil
M117718	71	Poorly differentiated	+	7	Numerous small foci of necrosis with or without RBC or not, with or without differentiating neuroblasts
M65760	48	Poorly differentiated	N.D.	17	Large area including a nodule of differentiating neuroblasts separated by large amount of neuropil. No RBC in necrotic area.

<sup>‡</sup> international neuroblastoma pathology classification

**Table 4.3** Transverse relaxation rate  $R_2^*$  and pathological characteristics of the Th-*MYCN* tumours.

#### 4.3.6 *MYCN-amplified childhood neuroblastomas are haemorrhagic*

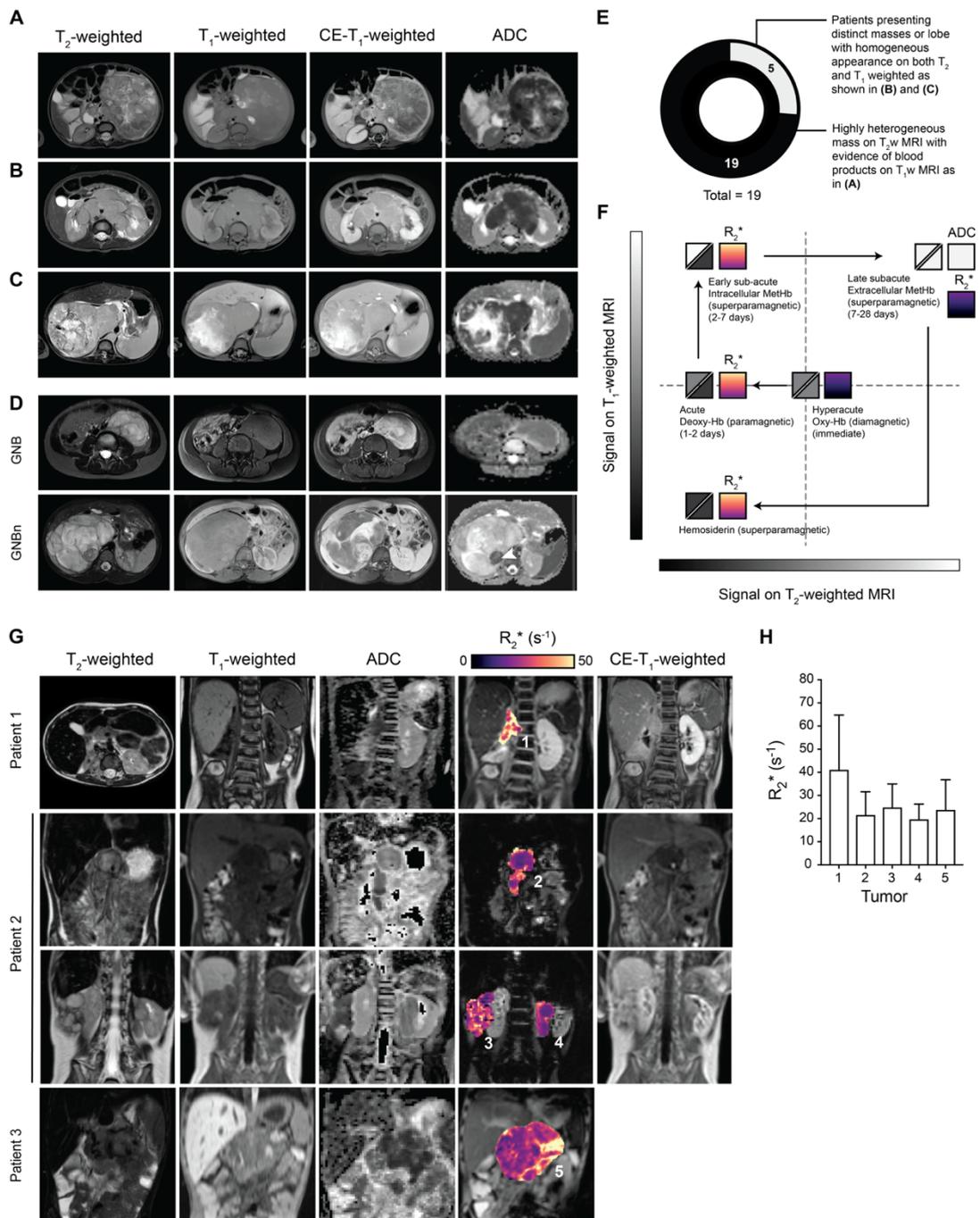
To understand the relevance of the tumour MRI phenotypes observed in the Th-*MYCN* GEM mouse to *MYCN*-driven childhood neuroblastoma, we reviewed anatomical T<sub>2</sub>-weighted, T<sub>1</sub>-weighted and contrast enhanced (CE-) T<sub>1</sub>-weighted images, as well as diffusion-weighted (DW) MRI-derived apparent diffusion coefficient (ADC) maps, acquired from 19 patients with *MYCN*-amplified neuroblastoma at the time of diagnosis. All patients presented with abdominal tumours showing a high degree of heterogeneity (both hyper- and hypointensity) on T<sub>2</sub>-weighted MRI and regional hyperintensity on T<sub>1</sub>-weighted MRI (Figure 4.7A–E). Based on the well-established appearance of ageing blood on both T<sub>2</sub>- and T<sub>1</sub>-weighted MRI in hematoma (Figure 4.7F; and (242)), this appearance suggests the presence of blood products at different stage of haemoglobin degradation. Most tumours presented a central region, which did not enhance on CE-MRI and showed elevated ADC values, suggesting the presence of large areas of nonviable tissue. Additionally, a few tumours presented with cystic components, with/without blood. Interestingly, 5 of the 19 patients also presented with contiguous or satellite masses showing homogeneous signal on both T<sub>1</sub> (isointense) and T<sub>2</sub> (hyperintense) MRI, homogeneous contrast enhancement and low ADC values, indicative of water restriction (Figure 4.7B and C). This MRI phenotype suggests a non-haemorrhagic and cellular-dense tumour phenotype within these lobes.

We also compared the radiology of two cases of ganglioneuroblastoma: a ganglioneuroblastoma intermixed (GNB) and ganglioneuroblastoma nodular (GNBn), two entities that are uncommonly *MYCN*-amplified. Yet GNB nodular are characterized by the presence of macroscopic undifferentiated neuroblastoma nodule(s) of Schwannian stroma-poor components coexisting with ganglioneuroblastoma-intermixed of stroma-rich component or ganglioneuroma (fully differentiated) of stroma-dominant component. The neuroblastoma nodule in the GNBn was easily identifiable with low ADC values, consistent with a stroma-poor neuroblastoma phenotype, and demonstrated marked hypointensity on T<sub>2</sub>-weighted images, hyperintensity on T<sub>1</sub>-weighted images and high

gadolinium avidity, overall suggesting that the neuroblastic nodule presented a highly vascular and haemorrhagic phenotype.

#### *4.3.7 Initial experience with IS-MRI in children with neuroblastoma*

Based on the initial results obtained in the Th-*MYCN* model, we have introduced IS-MRI into clinical trials in patients with relapsed/refractory neuroblastoma as a proof of concept. Our initial experience based on three patients scanned at 1.5T (Figure 4.7G) showed that tumours (n = 5) presented with a heterogeneous regional distribution of  $R_2^*$ , with tumour median  $R_2^*$  values ranging from 20 to 41  $\text{second}^{-1}$  (Figure 4.7H), a range similar to the one measured in mice when corrected for the difference of magnetic field strength. Regionally, higher  $R_2^*$  values were associated with hypointensity on  $T_2$ -weighted images, higher gadolinium avidity, and lower ADC values. Combined with the apparent haemorrhagic nature of *MYCN*-amplified neuroblastoma, this pilot study suggests that IS-MRI should be prospectively evaluated in children with neuroblastoma to determine its clinical utility as an imaging biomarker.



**Figure 4.7** Haemorrhage in *MYCN*-amplified childhood neuroblastoma MRI and initial experience with SW-MRI in the neuroblastoma clinic. **A–C.** show abdominal axial fat-suppressed STIR T<sub>2</sub>-weighted-MRI images, fat-suppressed SPAIR T<sub>1</sub>-weighted MRI images, before and after administration of gadolinium (Gd)-based contrast agent, and diffusion-weighted MRI-derived ADC maps of children with neuroblastoma at the time of diagnosis. **A–C.** *MYCN*-amplified high-risk neuroblastoma in a 10-month old boy (**A**), a 2-year old boy (**B**), a 8-month-old boy (**C**), **D.** Comparison between a GNB in a 4-year-old female patient and GNB in a 5-year-old male patient. Note the presence of the neuroblastic nodule in the GNBn easily identified on ADC maps (arrowhead). **E.** Proportion of patients with *MYCN*-amplified neuroblastoma (n = 19) presenting with a MRI phenotype, suggestive of the presence or absence of a hemorrhagic phenotype, based on the well-established knowledge of the appearance of aging blood in hematoma

on conventional MRI (242), as illustrated in **F. G.** Abdominal  $T_2$ -weighted and  $T_1$ -weighted images (not fat-suppressed), ADC maps, intrinsic susceptibility MRI-derived transverse relaxation  $R_2^*$  map of tumours, and post Gd contrast-enhanced  $T_1$ -weighted MRI images in children with refractory/relapsing neuroblastoma. (Patient 1, 7-year-old male; Patient 2, 5-year-old male; Patient 3, 6-year-old female). **H.** Tumour median  $R_2^*$  values for each individual tumour shown in **G** ( $\pm$ SD). Note that  $R_2^*$  values increased monotonically and approximatively linearly with magnetic field strength-dependent and, as such, clinical  $R_2^*$  value at 1.5T was estimated to be four times lower than if they were measured at 7T (243).

#### 4.4 Discussion

There is a current paradigm shift toward a stratified precision medicine approach for children with cancer. Central to this approach is the integration of both genomic and biological information including pharmacodynamic and predictive biomarkers of response to incorporate urgently needed novel therapies into treatment schedules. With increasing evidence for the contribution of angiogenesis in determining and predicting the biological behaviour of neuroblastoma, therapeutic approaches targeting the unique vascular phenotype of neuroblastoma are being accelerated in early phase paediatric clinical trials. However, there are currently no validated biomarkers predicting the clinical response or therapeutic benefits to anti-angiogenic therapy.

Although MRI is becoming the preferred clinical imaging method in the management of children with neuroblastoma particularly at diagnosis and for routine follow-up, the unique ability of advanced functional MRI techniques to non-invasively quantify changes in tumour vascular architecture and function has not yet been fully exploited (244). The MRI-histopathology correlation pipeline presented in Chapter 3, combined with the use of KDE hotspot mapping and additional evidence from the nonparametric Mantel statistical test demonstrated that *fBV*, measured *in vivo* by SC-MRI, spatially correlated with stained area for the endothelial cell marker endomucin in tumours arising in the Th-*MYCN* mouse model of neuroblastoma. This approach thus validates *fBV* as not only a sensitive but also specific imaging biomarker of tumour vascular perfusion, and its therapeutic modulation by the potent VEGFR inhibitor cediranib. Furthermore, we found a negative correlation between tumour *fBV* and cediranib-mediated changes in *fBV*, suggesting baseline *fBV* as a potential predictive biomarker of

vascular response to VEGFR inhibition. Importantly, the data presented in this chapter suggest that both baseline tumour *fBV* and  $R_2^*$  are predictive of the longer-term tumour response to VEGFR-targeted therapies such as cediranib.

Several preclinical studies have demonstrated that quantitation of tumour *fBV* using SC-MRI provides a sensitive biomarker of response to vascular-targeted agents *in vivo* (236-238). The data herein highlight the potential of SC-MRI to assess neuroblastoma vascular response to VEGF signalling inhibitors (245-248), and to *MYCN*-targeted therapeutics whose mechanism of action is predicted to elicit an antiangiogenic effect (43, 249-251). For example, direct silencing of *MYCN* through bromodomain extra terminal (BET) domain inhibition, destabilization of *MYCN* protein via the selective targeting of Aurora A kinase or mTOR/PI3k signalling, and inhibition of anaplastic lymphoma kinase (ALK) and RET kinases [reviewed in (88)] have all been linked with angiogenic blockade in aligned preclinical trials against neuroblastoma or other paediatric solid tumours (250-254).

The high vascular permeability and density present in the Th-*MYCN* tumours and detected by IS- and SC-MRI illustrate the archetypal VEGF-driven vascular remodelling triggered by members of the MYC family oncogenes (250, 255). There is increasing evidence that enhanced angiogenic signalling is part of a wider MYC-driven adaptive program to nutrient deprivation, to which MYC-driven tumours rapidly develop an addiction to maintain their growth (255, 256). As a result, any blockade of this program via MYC deactivation, metabolic inhibition, or vascular blockade triggers a wave of apoptosis, resulting in the unique rapid tumour debulking observed in MYC-driven tumours (255, 256). The correlation between baseline *fBV* and  $R_2^*$  and cediranib-induced tumour debulking established in this chapter illustrates a similar dependence on *MYCN*-driven angiogenic and vascular remodelling in the Th-*MYCN* GEM model. In contrast, tumours at the lower range of baseline  $R_2^*$  and *fBV* values, which progress despite cediranib treatment, presented with a more normalized vasculature. I also identified tumours with large areas of differentiating neuroblasts, a phenotype similar to that in *MYCN*-driven GEM models of neuroblastoma harbouring constitutive activation of ALK, and in which reduced VEGFR signalling has been

reported (249, 257, 258). Clinically, differentiated neuroblastoma displays low vascularity (229, 230), partly due to the release of strong antiangiogenic factors that directly antagonize the effect of VEGF (259, 260). Additionally, in differentiating, Schwannian stroma-poor neuroblastoma, VEGF-mediated cross-talk between differentiating neuroblasts and endothelial cells has been identified, which mimics normal vascular maturation in the developing nervous system (261). Overall in this chapter, reduced tumour sensitivity to cediranib was associated with a more stable vascular phenotype, suggesting a reduced dependence of these nonresponsive tumours on VEGF signalling to sustain growth.

The high tumour *fBV* values of ~20% reported in the Th-*MYCN* mouse models used here are consistent with the unique hypervascular phenotype of childhood neuroblastoma. The median endomucin staining area in the Th-*MYCN* tumours (2.3%) is consistent with the mean stained area of endothelial cell marker CD31 (1.7%) in a cohort of 458 primary childhood neuroblastomas (231). Structurally, the large areas of irregular sinusoidal vessels found in the haemorrhagic regions of Th-*MYCN* tumours emulate the major unstable vascular patterns associated with unfavourable clinical prognosis. In contrast, tumours with lower  $R_2^*$ , differentiated or not, presented a more normalized vascular pattern associated with favourable prognosis factors in primary neuroblastoma. The retrospective radiological analysis of 19 cases of *MYCN*-amplified neuroblastoma revealed that all tumours presented with a MRI phenotype consistent with haemorrhage. Additionally, the initial experience here with IS-MRI in children with refractory/relapse neuroblastoma demonstrated a range of tumour median  $R_2^*$  values, comparable to those measured in the Th-*MYCN* model when corrected for the difference in magnetic field strength (243). Collectively, these data demonstrate that the Th-*MYCN* GEM model of neuroblastoma recapitulates the pathophysiology of childhood neuroblastoma and provide a unique clinically relevant and information-rich platform to identify and validate imaging biomarkers of the neuroblastoma microenvironment.

DCE-MRI and CT-based functional imaging are commonly used to evaluate tumour perfusion in the clinic. Yet their associated biomarkers are difficult to

interpret as they reflect both vascular perfusion and permeability. More importantly, the acquisition of robust DCE-MRI data can be challenging in young children (233) and CT-based functional imaging methodologies should be preferentially avoided due to the higher inherent sensitivity of children to the negative effects of ionizing radiation (262). The data herein provide a strong rationale for the incorporation of both IS- and SC-MRI into functional imaging-embedded clinical trials for new therapeutic strategies that directly or indirectly modulate neuroblastoma vascular phenotype. IS-MRI is a rapid, quantitative, safe and non-invasive  $R_2^*$  relaxometry measurement that has become the reference measurement to assess both liver and cardiac iron content and a primary endpoint in paediatric clinical trials in highly transfused children suffering from haematological diseases who are at risk of liver and cardiac iron overload (263). The multi-gradient recalled echo MRI sequence is universally available on clinical MRI scanners and IS-MRI can be easily incorporated in the neuroblastoma clinic, as demonstrated herein, adding only two and half minutes to the current routine scanning session. IS-MRI has been incorporated in an ancillary functional imaging study of the BEACON Neuroblastoma trial. This chapter also demonstrates that haemorrhage detectable with IS-MRI, or the absence thereof, could hold prognostic/diagnostic value. Clinically, the haemorrhagic nature of undifferentiated neuroblastoma is well described, yet its routine reporting remains only anecdotal (264). Its value for risk-stratification could be explored by introducing non-invasive IS-MRI at the time of diagnosis. As a matter of fact, introducing susceptibility-weighted imaging (SW-MRI), a variant of IS-MRI, which combines the information from both the magnitude and the phase of the MR signal, would quantify haemorrhage with the added value of detecting calcification (opposite phase to haemorrhage). SW-MRI can assist in the differential diagnosis of neuroblastoma (frequently calcified) versus Wilms' tumours (seldom calcified) as this not easily detected on conventional MRI. SC-MRI, which combines  $R_2^*$  mapping with the use of USPIO particles, does not require bolus timing or complex image acquisition schemes, and provides an attractive steady-state imaging alternative to DCE-MRI. Recent clinical studies highlighting the safe off-label use of the USPIO preparation ferumoxytol for MRI in adults and children is a promising step-forward, warranting further evaluation of the approach in clinical trials (265).

## 4.5 Conclusion

IS- and SC-MRI are robust non-invasive imaging techniques to characterize, quantify, and map the unique vascular phenotype of neuroblastoma and its therapeutic modulation. With the central role of angiogenesis in determining and predicting the clinical behaviour of neuroblastoma, baseline  $fBV$  and  $R_2^*$ , have the potential to provide diagnostics and prognosis information at the time of diagnosis and provide currently unavailable predictive and pharmacodynamic biomarkers for antiangiogenic therapy against neuroblastoma.

## Chapter 5 : Exploring the histological correlates of native $T_1$ mapping in neuroblastoma pathology and its response to *MYCN*-targeted therapy *in vivo*

### 5.1 Introduction

#### 5.1.1 *MYCN*-targeted therapies for high-risk *MYCN*-driven neuroblastoma

Amplification of the proto-oncogene *MYCN* is the most common genomic aberration in neuroblastoma, which defines a subgroup of children with a high-risk disease. *MYCN* plays a central role in the biology of high-risk neuroblastoma and as such represents a major therapeutic target. Small molecule inhibitors targeting the stability of *MYCN* protein have shown strong anti-tumour activity in the Th-*MYCN* model and are being evaluated in early-phase paediatric clinical trials (88, 266-268). These include the selective inhibitor of Aurora A kinase, alisertib (MLN8237, NCT01601535), and selective inhibitors of mTOR activity (NCT01331135, NCT01467986, NCT01625351, NCT02343718, NCT02574728, NCT02638428, NCT02813135).

#### 5.1.2 Reduction in the native spin lattice relaxation time $T_1$ as a generic biomarker of response to therapy

*McSheehy and colleagues* have demonstrated that an early reduction in native spin-lattice relaxation time  $T_1$  measured using inversion recovery true fast imaging with steady-state precession (IR-TrueFISP) MRI, correlated with eventual change in tumour volume in a wide range of chemosensitive xenograft models following treatment with five drugs with varied mechanism of action (28). *Jamin and colleagues* demonstrated that a reduction in median tumour  $T_1$ , can provide a sensitive biomarker of response to cyclophosphamide, which is a ubiquitous component of various frontline protocols for neuroblastoma, and to anti-vascular therapies in the Th-*MYCN* model (41). While  $T_1$  is a fundamental parameter in MRI, the pathological determinants underpinning the reduction in native tumour  $T_1$  associated with successful therapy remains unclear, although potentially associated with the release of paramagnetic species during cell death.

### 5.1.3 Aim of this chapter

In this chapter, I demonstrate how native tumour  $T_1$  provides a robust biomarker of response to alisertib and the mTOR inhibitor vistusertib (AZD2014) in the Th-*MYCN* model of neuroblastoma. The comparison of native  $T_1$  maps with those derived from multi-parametric MRI and computational pathology demonstrated that native  $T_1$  mapping (the voxel-wise quantification of  $T_1$ ) is sensitive to the rich histological presentation of neuroblastoma, including regional differences in the density of undifferentiated, differentiating and apoptotic neuroblasts populations. The potential application of  $T_1$  mapping for diagnosis/prognosis, surgical planning and the evaluation of novel therapies for children with neuroblastoma, as well as the mechanism of contrast, are discussed.

## 5.2 Materials and Methods

### 5.2.1 Animals, imaging and drug treatment schedule

A total of 46 mice were enrolled with a median tumour volume of  $861 \pm 86 \text{ mm}^3$  (derived from  $T_2$ -weighted MRI; median  $\pm 1$  s.e.m., ranging from 280 to 2557  $\text{mm}^3$ ). MRI was performed prior to treatment (Day 0). Mice were left to recover for 24h, and then treated (Day 1) with 30 mg/kg p.o. of Alisertib (MLN8237, purchased from Selleckchem, n=11) or vehicle (10% 2-hydroxypropyl  $\beta$ -cyclodextrin, 1%  $\text{NaHCO}_3$ , n=9), or 25 mg/kg p.o. of Vistusertib (AZD2014, obtained under material transfer agreement with AstraZeneca, n=14) or vehicle (5% DMSO, 95% PEG300, n=12). Post-treatment MRI was performed 24h after treatment started (Day 2).

To retrospectively analyse the relationship between  $T_1$  and  $R_2^*$ , a total of 71 mice were included.

A more detailed description of the animal experiments can be found in the Appendix.

### 5.2.2 MRI

For all the mice, contiguous anatomical  $T_2$ -weighted transverse images were acquired through the mouse abdomen for the quantification of tumour volume, optimization of the local field homogeneity using the FASTmap algorithm, and for planning the subsequent multiparametric MRI measurements. In addition to IR-TrueFISP MRI for quantification of the spin-lattice ( $T_1$ ) and spin-spin ( $T_2$ ) relaxation times, these also included measurement of the apparent diffusion coefficient (ADC), the transverse relaxation rate  $R_2^*$  and the magnetization transfer ratio (MTR).

A more detailed description of the MRI methods can be found in the Appendix.

### 5.2.3 Computational pathology/digital histology

*Digitized histology.* Histology FFPE sections were acquired as described in Chapter 3 and stained with H&E. I digitized the whole-slide H&E images using a Hamamatsu NanoZoomer XR scanner (20x magnification, 0.46 $\mu$ m resolution, Hamamatsu, Japan).

*MRI-histology alignment.* For each tumour, the MRI slice of interest was visually aligned with the digitized whole-slide H&E stained image using anatomical landmarks as described in Chapter 3.

*Cell segmentation and classification.* Image processing was carried out using the CRImage-based pipeline, as described in Chapter 2. Cells were segmented and subsequently classified into 5 categories: undifferentiated neuroblasts, differentiating neuroblasts, apoptotic cells, lymphocytes, stromal cells.

*Generation of cellular density and classified cell parametric maps.* Whole-slide images of cells were processed to match the MRI resolution (234 x 234  $\mu$ m), with the number of segmented cells and classified cells within 518x518 pixel-regions representing a single pixel in the final cell density maps. Density maps were normalized to their sample's maximum number of cells/classified cells in order to facilitate the evaluation intra-tumour heterogeneity.

*MRI- and histology-derived parametric map registration.* This was performed using the CPD algorithm as described in Chapter 3. Firstly, density maps of all the segmented cells were non-rigidly registered to the T<sub>1</sub> images based on features extracted by a Canny edge detector. Using the edge features extracted from the density maps of all the segmented cells, the density maps of each classified cell category were subsequently registered to the T<sub>1</sub> maps.

*Spatial quantitative comparison between MRI parametric maps or between MRI- and histology-derived maps.* In order to explore and understand the biological determinants of native T<sub>1</sub>, native T<sub>1</sub> values were compared with maps of the other

MRI parameters (including  $R_2^*$  as a marker for haemorrhage as demonstrated in Chapter 4) and the histology-derived parametric maps. Taking into account the inherent difficulties in quantitatively comparing MRI to histology discussed in Chapter 3, an unbiased region-based approach was chosen over a pixel-to-pixel comparison. The  $T_1$  parametric maps were divided into sub-regions of high and low values using various thresholds summarized in Table 5.3. A binary mask was created for each sub-region and applied to the second parametric map. This analysis was performed in 13 tumours across both vehicle cohorts for which precise MRI-Pathology registration was possible. Statistical comparison of sub-regional median values between the two parametric maps was performed and the process was repeated in reverse.

*Cleaved caspase-3 immunohistochemistry and quantification.* Cleaved caspase-3 (CC3) (Asp175) (Cell Signalling, #9661) antibodies were used. Cleaved caspase 3 expression was quantified as the percentage of stained area on digitized whole-slide histopathological images (10x) by applying colour deconvolution to extract the brown colour channel followed by the application of a manual threshold using Fiji. Computational analysis of Cleaved caspase 3 expression was conducted on digitized whole-slide histopathological images (20x) using the free software QuPath-version 0.1.2 (<https://qupath.github.io/>). CC3 staining for each cell was quantified as the mean optical density of nuclear DAB staining using the *Positive Cell Detection* module (269-273). Results of detected CC3 positive cells were exported and processed to produce maps of equivalent MRI-resolution in Matlab as described previously.

#### 5.2.4 Statistical Analysis

Statistical analysis was performed with GraphPad Prism 7 (GraphPad Software Inc., La Jolla, USA). The mean values for tumour volume, and the mean of the median values for all the quantitative MRI parameters were used. All the absolute and treatment-induced relative changes in MRI parameters were assumed to be normally distributed, which was confirmed using the D'Agostino-Pearson omnibus K2 normality test. Student's two-tailed t-test was used to assess any significant differences in quantitative MRI parameters and tumour volume upon

treatment (paired), and in the magnitude of these changes compared to the control cohort (unpaired), with a 1% level of significance. Further statistical analysis was performed with the Bonferroni correction ( $n=5$ ). Any significant differences between groups identified in the sub-regional analysis were identified using the Wilcoxon signed rank test with a 5% level of significance. Significant correlations were determined using linear regression analysis, confirmed by using the robust regression and outlier removal approach (274).

### 5.3 Results

#### 5.3.1 Alisertib and vistusertib elicit significant anti-tumour activity associated with a decrease in native $T_1$

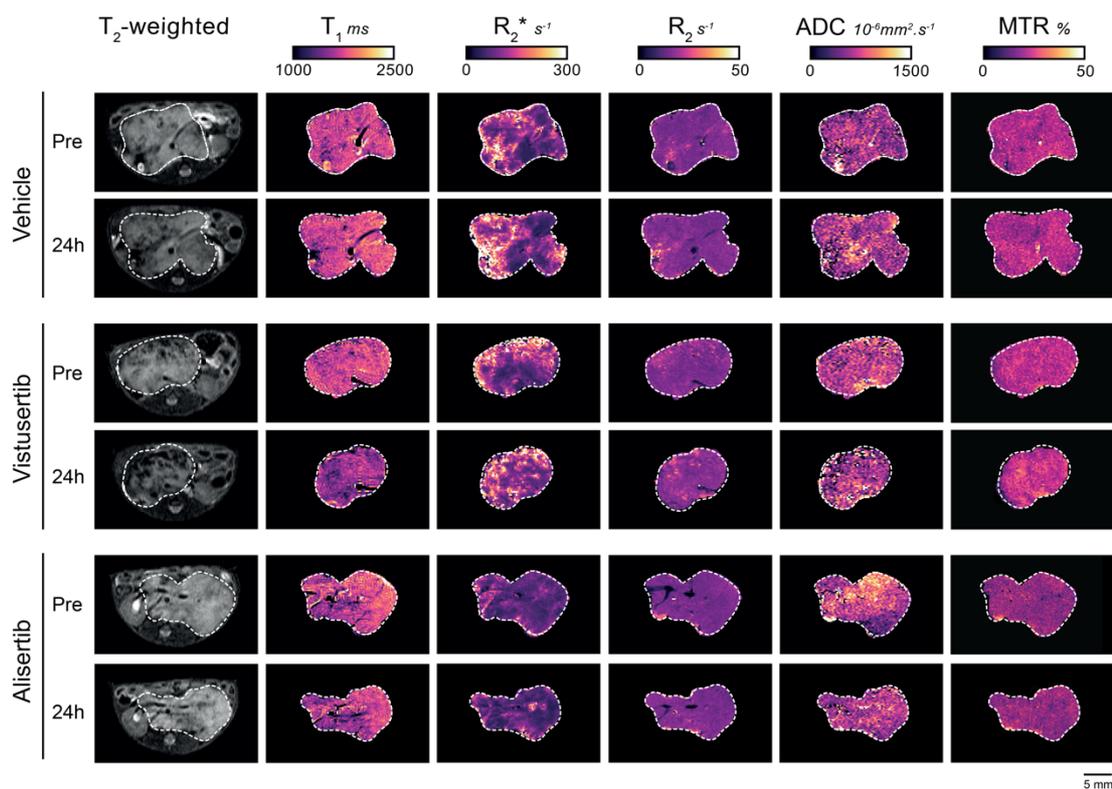
The Th-MYCN GEM model of neuroblastoma recapitulates the aggressiveness of the clinical disease, with an observed average  $31 \pm 4\%$  increase in tumour volume measured over the 48h experimental time course (Table 5.1, Table 5.2, Figure 5.1). Despite this, tumour median values for all the MRI parameters remained stable over 48h in the vehicle treated cohorts (coefficients of variation  $\text{CoV}_{T_1} = 2.4\%$ ,  $\text{CoV}_{T_2} = 5.8\%$ ,  $\text{CoV}_{ADC} = 14.0\%$ ,  $\text{CoV}_{R_2^*} = 11.7\%$ ,  $\text{CoV}_{MTR} = 6.2\%$ ). There was no significant difference in tumour volume between the different treatment cohorts at the time of enrolment (Figure 5.2). Treatment with either alisertib or vistusertib led to a highly significant reduction in native  $T_1$  ( $-9.3 \pm 0.9\%$  and  $-5.4 \pm 1.1\%$ , both  $P < 0.0001$ ) and was associated with a significant reduction in tumour volume with vistusertib ( $-42 \pm 5.1\%$ ,  $P < 0.0001$ ) but not alisertib, although a reduction in tumour volume was seen in 9 out of 11 treated mice (Figure 5.3). Both the alisertib and vistusertib treated groups elicited significant anti-tumour activity when compared to their respective vehicle control cohorts (both  $P < 0.0001$ ). No significant changes in tumour native  $T_2$ , ADC,  $R_2^*$ , or MTR were determined following treatment with either alisertib or vistusertib, nor any treatment-induced relative changes compared with vehicle controls.

	Vehicle control				25mg/kg Vistusertib			
	Pre	24h post	Relative changes	n	Pre	24h post	Relative changes	n
<b>Tumour volume</b>	1185±221mm <sup>3</sup>	1474±245mm <sup>3</sup>	31±5.5%	14	966±127 mm <sup>3</sup>	577 ±93 mm <sup>3</sup> ( <b>&lt;0.0001</b> ) §	-42.0 ±5.1% ( <b>&lt;0.0001</b> ) ‡	12
<b>T<sub>1</sub></b>	1723 ±16ms	1771±14ms	2.9±0.8%	14	1712±25ms	1553±25ms ( <b>0.0001</b> )§	-9.3 ±0.9 % ( <b>&lt;0.0001</b> )‡	12
<b>T<sub>2</sub></b>	62±1ms	60±1ms	-4.0±1.7 %	14	63±2ms	58±2ms	-8.0±2.7 %	12
<b>R<sub>2</sub>*</b>	102±6s <sup>-1</sup>	105 ±7s <sup>-1</sup>	4.5±3.3%	14	99±6s <sup>-1</sup>	113±10s <sup>-1</sup>	13.3 ±7.9%	12
<b>MTR</b>	22.5±0.6%	22.3±0.4%	-0.4 ±2.2%	14	22.2±0.3%	23.0±0.5ms	3.9±2.2%	11
<b>ADC</b>	593 ±26 .10 <sup>-6</sup> mm <sup>2</sup> .s <sup>-1</sup>	569±22 .10 <sup>-6</sup> mm <sup>2</sup> .s <sup>-1</sup>	-2.5±4.7%	14	689±42 .10 <sup>-6</sup> mm <sup>2</sup> .s <sup>-1</sup>	686 ±40 .10 <sup>-6</sup> mm <sup>2</sup> .s <sup>-1</sup>	2.4 ±8.6%	10

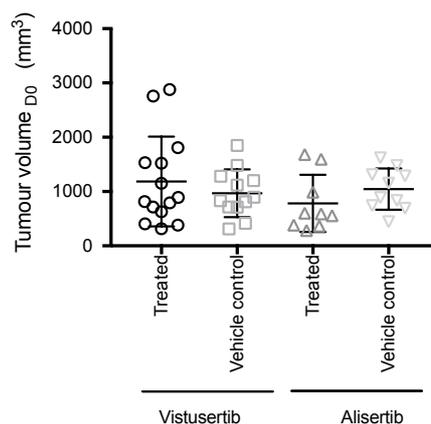
**Table 5.1** Summary of the response of the Th-MYCN transgenic model of neuroblastoma to vistusertib. Data are presented as mean of tumour median value ± 1 s.e.m. § Student's two-tailed paired t-test, ‡ Student's two-tailed unpaired t-test, both incorporating a Bonferroni correction (n=6) and assuming a 1% level of significance. The difference in the number of mice associated with the different parameters reflects that it was not possible to acquire the full protocol in all cases.

	Vehicle control				30mg/kg Alisertib			
	Pre	24h post	Relative changes	n	Pre	24h post	Relative changes	n
<b>Tumour volume</b>	781±176mm <sup>3</sup>	981±198mm <sup>3</sup>	30.5±5.0%	9	1037±109 mm <sup>3</sup>	938 ±129 mm <sup>3</sup>	-11.2 ±4.3% ( <b>&lt;0.0001</b> ) ‡	11
<b>T<sub>1</sub></b>	1754 ±36ms	1750±32ms	-0.2±0.6%	9	1776±26ms	1679±21ms ( <b>0.0008</b> )§	-5.4±1.1 % ( <b>&lt;0.001</b> )†	11
<b>T<sub>2</sub></b>	62±3ms	60±2ms	-3.2±3.6 %	6	62±1ms	63±1ms	1.9±2.9%	7
<b>R<sub>2</sub>*</b>	109±14s <sup>-1</sup>	117 ±20s <sup>-1</sup>	5.0±8.0%	6	114±11s <sup>-1</sup>	108±9s <sup>-1</sup>	-1.2 ±11.7%	7
<b>MTR</b>	23.6±0.7%	22.3±1.0%	-5.7 ±3.8%	6	22.6±0.4%	21.5±0.5ms	-0.37±3.2%	7
<b>ADC</b>	664 ±66 .10 <sup>-6</sup> mm <sup>2</sup> .s <sup>-1</sup>	607 ±17 .10 <sup>-6</sup> mm <sup>2</sup> .s <sup>-1</sup>	-5.6±6.3%	6	615±29 .10 <sup>-6</sup> mm <sup>2</sup> .s <sup>-1</sup>	664 ±33 .10 <sup>-6</sup> mm <sup>2</sup> .s <sup>-1</sup>	8.6 ±6.3%	5

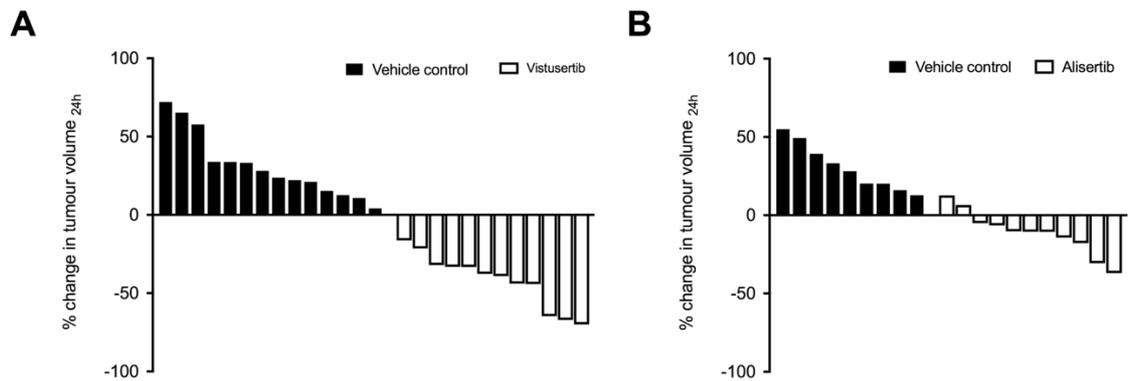
**Table 5.2** Summary of the response of the Th-MYCN transgenic model of neuroblastoma to alisertib. Data are presented as mean of tumour median value ± 1 s.e.m. § Student's two-tailed paired t-test, ‡ Student's two-tailed unpaired t-test, both incorporating a Bonferroni correction (n=6), and assuming a 1% level of significance. The difference in the number of mice associated with the different parameters reflect that it was not possible to acquire the full protocol in all cases



**Figure 5.1** Representative  $T_2$ -weighted anatomical MR images of tumour-bearing Th-MYCN mice and associated parametric maps of the tumour spin-lattice relaxation time  $T_1$ , transverse relaxation rate  $R_2^*$ , spin-spin relaxation rate  $R_2$  ( $=1/T_2$ ), apparent diffusion coefficient (ADC) and magnetization transfer ratio (MTR), prior to and 24 hours following treatment with 25mg/kg vistusertib, 30mg/kg alisertib or vehicle.



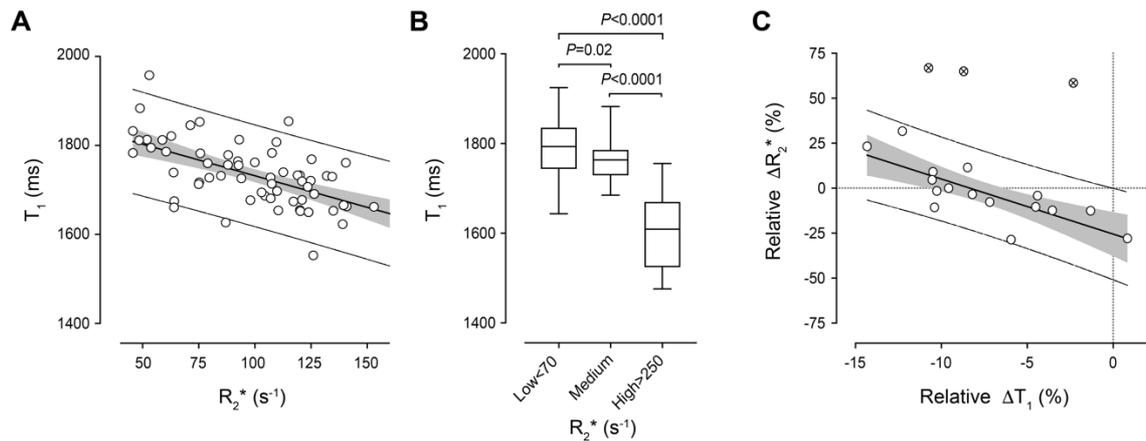
**Figure 5.2** Summary of the tumour volumes quantified using  $T_2$ -weighted MRI at the time of enrolment ( $D_0$ ). There was no significant difference in tumour volume at  $D_0$  between the different treatment cohorts ( $P > 0.01$ , one-way ANOVA with a 5% level of significance).



**Figure 5.3** Waterfall plots documenting relative changes in tumour volume in the Th-*MYCN* mouse model of neuroblastoma 24 hours following treatment with **A.** 25mg/kg vistusertib, **B.** 30mg/kg alisertib, or their respective vehicle.

### 5.3.2 Low native tumour $T_1$ correlates with high tumour red blood cell content.

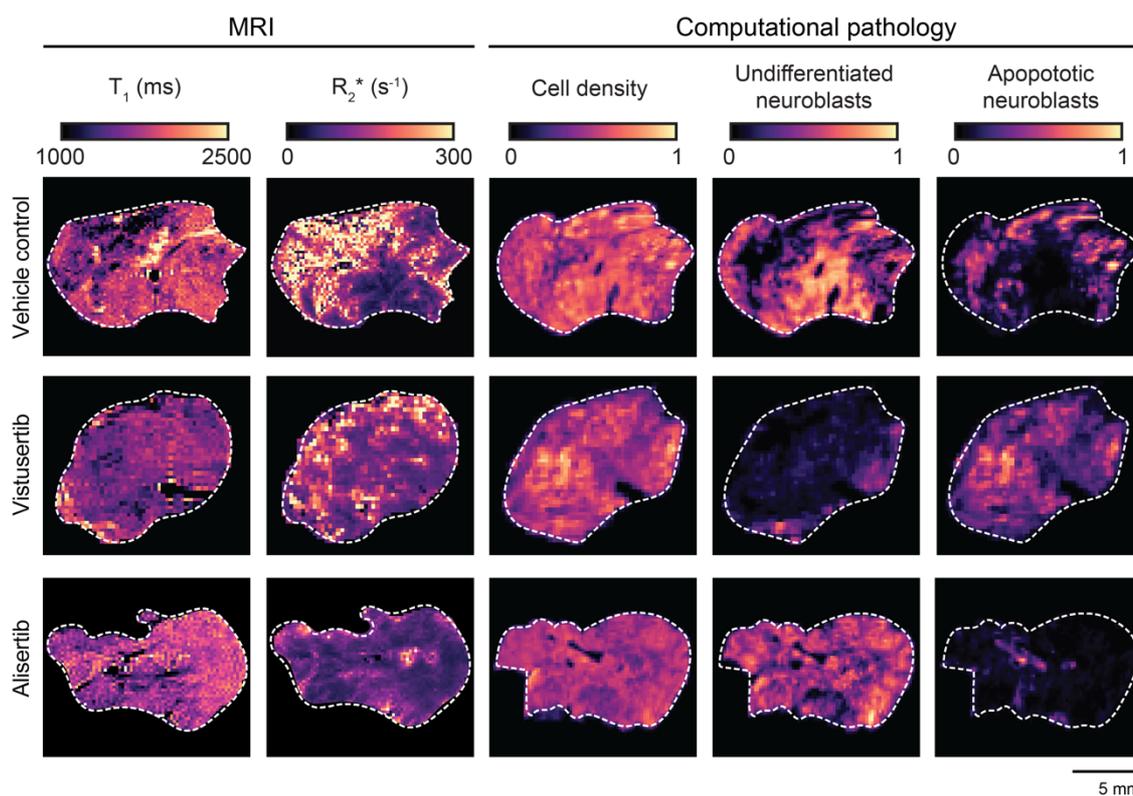
The effect of the characteristic haemorrhagic phenotype on native  $T_1$  was evaluated by comparing  $T_1$  and  $R_2^*$  maps (Figure 5.1) based on the robust relationship between native  $R_2^*$  and RBC aggregates established in Chapter 4. Regions of high  $R_2^*$  co-localized with regions of low native  $T_1$ . Retrospective analysis of measurements made in 71 untreated tumours arising in GEM models of neuroblastoma revealed that the median native  $T_1$  inversely correlated with native median  $R_2^*$  ( $r = -0.59$ ,  $P < 0.0001$ ) (Figure 5.4A). Sub-regional analysis using established empirical  $R_2^*$  threshold values ( $R_2^* < 70\text{s}^{-1}$  as no haemorrhage (43, 275) and  $> 250\text{s}^{-1}$  as purely RBC, and mixed regions of neuroblasts and RBC for the  $R_2^*$  values between) identified significantly different values of  $T_1$  associated with low ( $< 70\text{s}^{-1}$ ), intermediate and high ( $> 250\text{s}^{-1}$ )  $R_2^*$  (Figure 5.4B). Comparison of the relative changes in median  $T_1$  and  $R_2^*$  with treatment revealed a significant negative correlation ( $r = -0.78$ ,  $P = 0.002$  with Bonferroni correction [n=5], Figure 5.4C). Importantly, both positive and negative changes in tumour  $R_2^*$  occurred with treatment, thereby accentuating the sensitivity of native  $T_1$  to RBC deposition, but excludes changes in the content of paramagnetic RBCs or other such species as the main cause of reduction in  $T_1$  upon treatment.



**Figure 5.4 A.** Scatter graph of the native spin lattice relaxation time  $T_1$  against native transverse relaxation rate  $R_2^*$  from 71 untreated tumours arising in genetically-engineered murine models of neuroblastoma. Linear regression analysis and associated 95% confidence and prediction intervals are shown. A highly significant negative correlation was obtained ( $r = -0.59$ ,  $P < 0.0001$  with Bonferroni correction [ $n=5$ ]). **B.** Box-and-whisker plot showing the difference in native  $T_1$  in sub-regions categorized by low ( $< 70s^{-1}$ ), intermediate ( $70s^{-1} < R_2^* < 250s^{-1}$ ) and high ( $> 250s^{-1}$ ) values of  $R_2^*$  measured in Th-MYCN tumours treated with vehicle ( $n= 13$ ). Data are medians and interquartile range. **C.** Scatter graph of relative changes in native tumour  $R_2^*$  ( $\Delta R_2^*$ ) and relative changes in native  $T_1$  ( $\Delta T_1$ ) 24 hours following treatment with either alisertib or vistusertib. Bold lines represent linear regression with crossed dots indicating outliers determined using the robust regression and outlier removal approach. Grey shaded area indicates the 95% confidence intervals while dashed lines indicate 95% prediction confidence. A significant negative correlation was obtained ( $r = -0.78$ ,  $P = 0.002$  with Bonferroni correction [ $n=5$ ]).

### 5.3.3 High native tumour $T_1$ correlates with high density of undifferentiated neuroblasts and with low density of apoptotic neuroblasts

I then focused on the major histological component of these tumours, i.e. the dense cellular network. I generated parametric maps of classified undifferentiated neuroblasts and apoptotic cells density and compared them with spatially-registered native  $T_1$  maps. In vehicle control tumours, regions exhibiting high values of  $T_1$  co-localized with dense regions of undifferentiated neuroblasts (Figure 5.5).

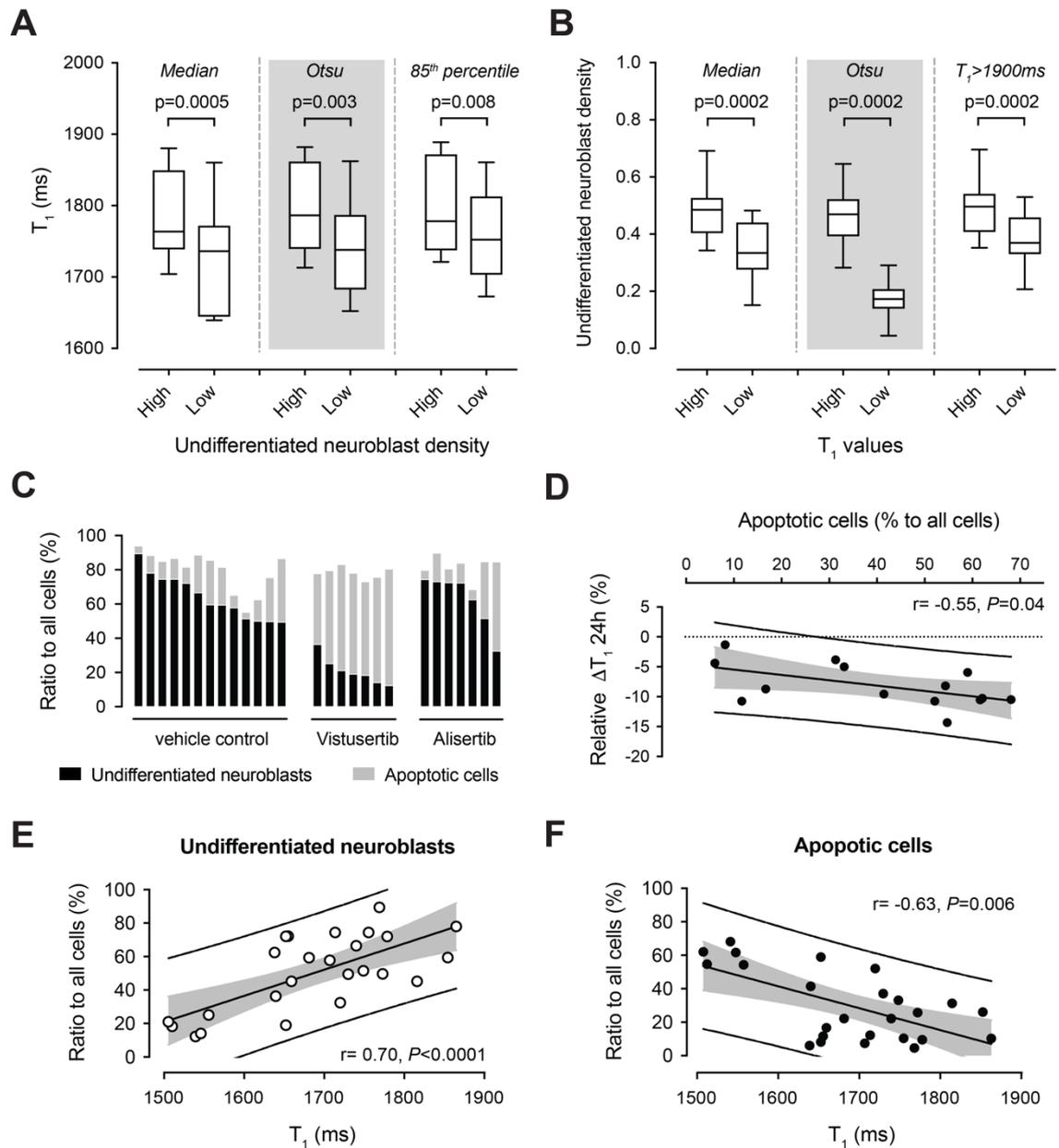


**Figure 5.5** Representative MRI-derived parametric maps of the tumour spin-lattice relaxation time  $T_1$  and transverse relaxation rate  $R_2^*$ , and registered histopathology-derived parametric maps of cell density including undifferentiated and apoptotic neuroblasts in the Th-MYCN model of neuroblastoma, 24 hours following treatment with either vehicle control, 25mg/kg vistusertib or 30mg/kg alisertib.

Threshold-based sub-regional analysis confirmed that regions with higher  $T_1$  values corresponded to areas of increased density of undifferentiated neuroblasts and, reciprocally, regions with higher neuroblast density had higher native  $T_1$  values (Table 5.3 and Figure 5.6A&B). Interestingly, areas dense in apoptotic cells in vehicle control tumours also corresponded to regions of lower native  $T_1$ . The widespread reduction in  $T_1$  seen in the vistusertib-treated tumours was associated with a widespread and significantly higher fraction of apoptotic cells ( $57 \pm 3\%$  compared to  $16 \pm 3\%$  in vehicle control,  $P < 0.0001$ ) and tissue damage, concomitant with a significantly lower fraction of undifferentiated neuroblasts ( $21 \pm 3\%$  compared to  $64 \pm 4\%$  in vehicle control,  $P < 0.0001$ ) (Figure 5.6C).

Thresholds on density values	T <sub>1</sub> values from segmented cells density thresholds			T <sub>1</sub> values from classified undifferentiated neuroblasts density thresholds		
	High	Low	p-value	High	Low	p-value
<b>Median</b>	1753±18	1742±19	<b>0.01</b>	1763±17	1736±20	<b>0.0005</b>
<b>Otsu</b>	1761±17	1721±21	<b>0.0005</b>	1786±17	1737±18	<b>0.003</b>
<b>85<sup>th</sup> percentile</b>	1769±22	1761±16	0.38	1778±18	1752±17	<b>0.008</b>
<b>Thresholds on T<sub>1</sub> values</b>						
Thresholds on T <sub>1</sub> values	Segmented cells density			Classified undifferentiated neuroblasts density		
	High	Low	p-value	High	Low	p-value
<b>Median</b>	0.58±0.02	0.50±0.02	<b>0.0002</b>	0.48±0.03	0.33±0.03	<b>0.0002</b>
<b>Otsu</b>	0.57±0.02	0.33±0.02	<b>0.0002</b>	0.47±0.03	0.17±0.02	<b>0.0002</b>
<b>T<sub>1</sub> &gt; 1900</b>	0.57±0.02	0.52±0.02	<b>0.0002</b>	0.49±0.03	0.37±0.03	<b>0.0002</b>

**Table 5.3** Summary of sub-regional analysis for the comparison of T<sub>1</sub> with segmented cell density and classified undifferentiated neuroblasts density.

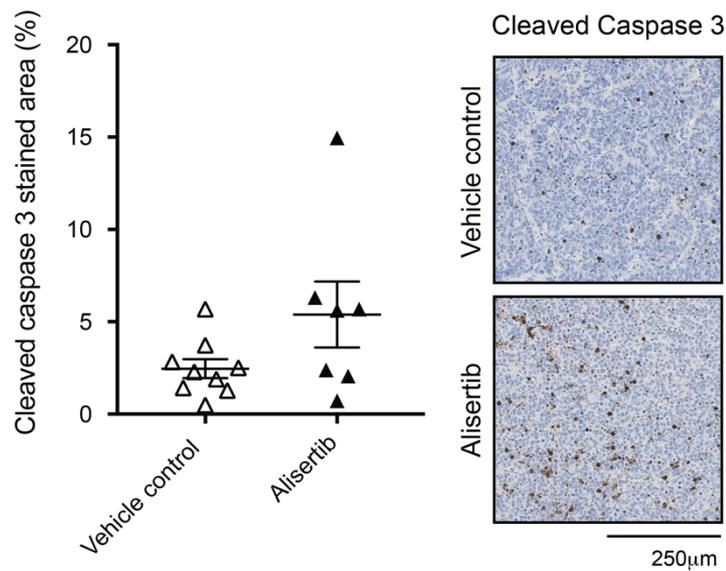


**Figure 5.6 A.** Box-and-whisker plot showing the difference in native  $T_1$  values in sub-regions categorized by low and high density of undifferentiated neuroblasts. Dichotomization was achieved using either median, Otsu or 85<sup>th</sup> percentile thresholds on registered histopathology-derived parametric maps of segmented and classified undifferentiated neuroblasts in vehicle control Th-MYCN tumours ( $n=13$ ). **B.** Box-and-whisker plot showing the difference in undifferentiated neuroblast density in sub-regions categorized by low and high native  $T_1$  values, defined using either median, Otsu or  $T_1 > 1900\text{ms}$  thresholds in vehicle control Th-MYCN tumours. Data are medians and interquartile range. ( $P$ , Wilcoxon signed rank test with a 5% level of significance) **C.** Proportion of undifferentiated and apoptotic neuroblasts relative to all cells derived from cell segmentation and classification from haematoxylin and eosin (H&E) stained histopathology from Th-MYCN tumours 24 hours following treatment with either vehicle control, 25mg/kg vistusertib or 30mg/kg alisertib. **D.** Scatter graph showing that the reduction in native tumour  $T_1$  over 24h treatment with either alisertib or vistusertib correlated with an increased proportion of apoptotic cells present in the tumour at the time of excision ( $r = 0.55$ ,  $P=0.04$ ). **E,F.** Scatter graphs showing that median tumour

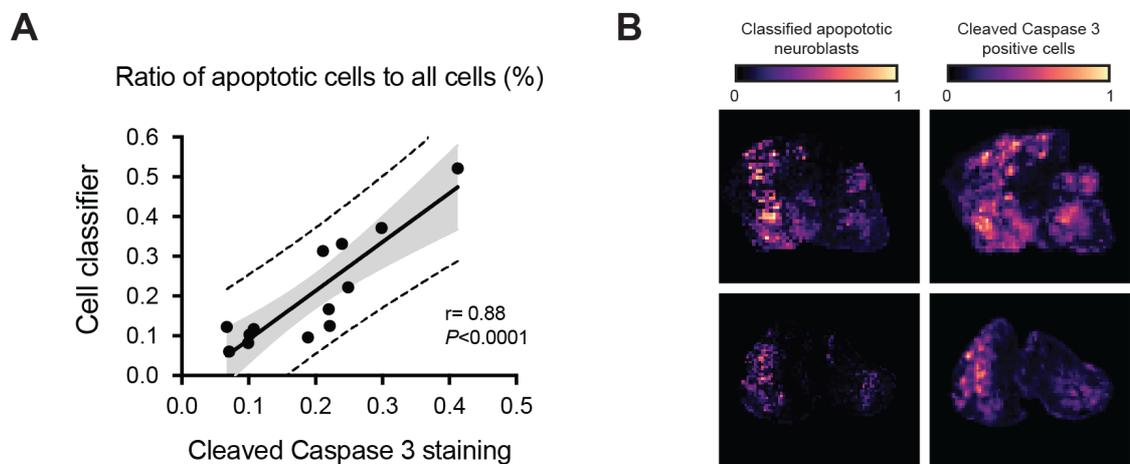
native  $T_1$  in the treated and vehicle control cohorts positively correlated with the proportion of undifferentiated neuroblasts ( $r = 0.70$ ,  $P < 0.0001$ ), and negatively correlated with the proportion of apoptotic neuroblasts ( $r = -0.63$ ,  $P = 0.006$ ). Grey shaded area indicates 95% confidence intervals while dashed lines indicate 95% prediction confidence.

The more modest but widespread reduction in  $T_1$  in the alisertib-treated tumours was not associated with any detectable differences in the fraction of apoptotic or undifferentiated neuroblasts on corresponding H&E staining, as confirmed by cleaved caspase 3 staining (Figure 5.7, Figure 5.8). Note that this response was however associated with the reduction in tumour volume seen in 9 out of the 11 mice treated (Table 5.1), and the absence of any significant difference in tumour  $T_1$  post-treatment between the alisertib and vehicle control cohorts (contrary to that seen with vistusertib,  $P < 0.0001$ ).

Combining the MRI data from vistusertib- and alisertib- treated mice with matched histopathology revealed a significant negative correlation between treatment-induced reduction in  $T_1$  over 24h and the proportion of apoptotic neuroblasts present in the tumour at the study endpoint ( $r = -0.55$ ,  $P = 0.04$ , Figure 5.6D). Combining the MRI data from vehicle control, vistusertib and alisertib-treated tumours with matched histopathology showed a positive correlation between median  $T_1$  and the ratio of undifferentiated neuroblasts ( $r = 0.70$ ,  $P < 0.0001$ , Figure 5.6E) and a negative correlation with the fraction of apoptotic cells ( $r = -0.63$ ,  $P = 0.006$ , Figure 5.6F).



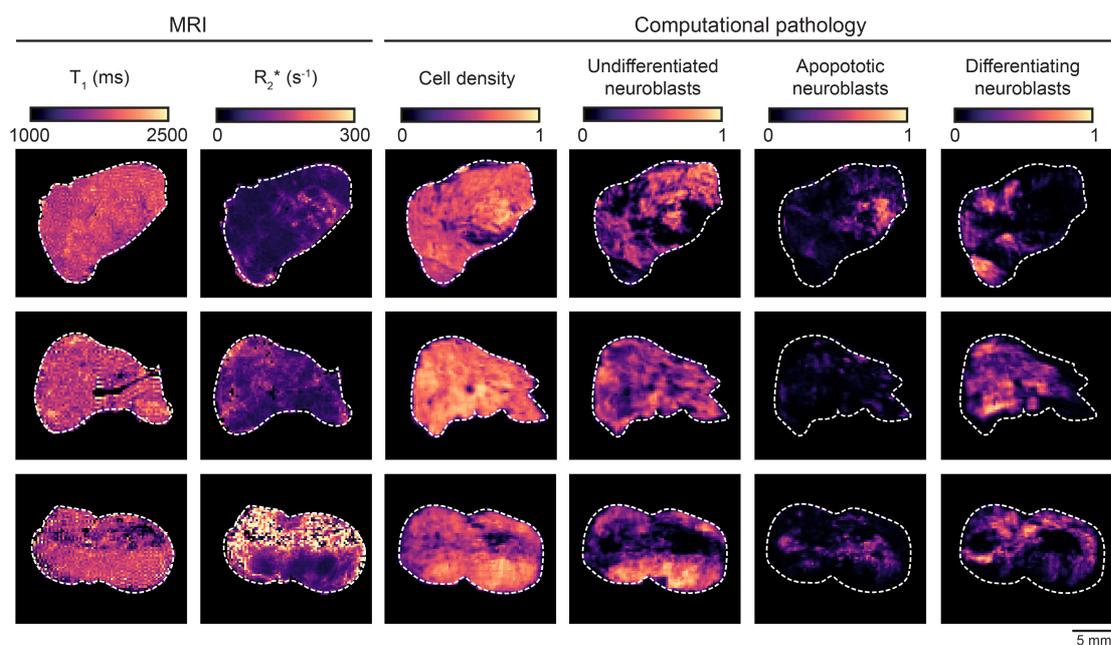
**Figure 5.7** Quantification of cleaved caspase-3 staining in tumours arising in the Th-MYCN mouse model of neuroblastoma 24h after treatment with 30mg/kg alisertib or vehicle.



**Figure 5.8 A.** Scatter graph showing that the proportion of apoptotic cells in tumours arising in the Th-MYCN mouse model of neuroblastoma 24h after treatment with 30mg/kg alisertib or vehicle and quantified using the cell classification algorithm significantly correlated ( $r = 0.88$ ,  $P < 0.0001$ ) with the proportion of cells staining positively for cleaved caspase 3 on adjacent IHC slides. The grey shaded area indicates the 95% confidence intervals while the dashed lines indicate 95% prediction confidence  
**B.** Histopathology-derived parametric maps showing apoptotic cells density determined using the cell classification algorithm and the density of cells staining positively for cleaved caspase 3 in Th-MYCN tumours 24 hours following treatment with 30mg/kg alisertib or vehicle.

### 5.3.4 Regions rich in differentiating neuroblasts are associated with lower $T_1$ values.

Three tumours exhibited a significant amount of differentiating neuroblasts (yet with only very few mature ganglion cells). In these tumours (Figure 5.9), previously shown to have very low levels of haemorrhage (275), regional differences in  $T_1$  visually and spatially corresponded to differences in undifferentiated neuroblast density, with regions of low  $T_1$  and low density undifferentiated neuroblasts corresponding to hotspots of differentiating neuroblasts, arranged in islands separated by a large amount of neuropil or simply interspersed with undifferentiated neuroblasts.



**Figure 5.9** Three cases of differentiating tumours in the Th-MYCN model of neuroblastoma. **A.** Representative MRI-derived parametric maps of the tumour spin-lattice relaxation time  $T_1$  and transverse relaxation rate  $R_2^*$ , and registered representative pathology-derived parametric maps of tumour cell density including undifferentiated, apoptotic and differentiating neuroblasts.

## 5.4 Discussion

This chapter exploited computational pathology methodologies to explore the sensitivity of  $T_1$  mapping to the rich histological presentation of neuroblastoma, and how it can provide a sensitive biomarker of response to two clinically-relevant *MYCN*-targeted therapeutics in the Th-*MYCN* GEM model of neuroblastoma.

### *T<sub>1</sub> mapping of neuroblastoma histopathology and its regional heterogeneity*

Using this approach, four major determinants of the regional heterogeneity observed on native  $T_1$  maps were identified, and confirmed using quantitative sub-regional analysis and included: *i*) regions with high  $T_1$  values which corresponded to hotspots of undifferentiated neuroblasts, characterized by a high level of proliferation, whereas *ii*) regions rich in differentiating neuroblasts exhibited lower  $T_1$  values, and both *iii*) regions with large amounts of extravasated RBCs and *iv*) large areas of cell damage, with or without RBCs, were both associated with very low  $T_1$  values. The association between  $T_1$  and extravasated RBCs was an expected finding consistent with the linear relationship of blood  $T_1$  with haematocrit level and haemorrhage (276).

### *Reduction in tumour native T<sub>1</sub> is associated with a reduction in undifferentiated neuroblast density*

Data with vistusertib treatment indicate that the reduction in native  $T_1$  was associated with a shift in tumour composition characterized by rapid loss of tumour regions with higher  $T_1$  values, a consequence of cell death, with the post-therapy tumour  $T_1$  values determined by dying and remaining haemorrhagic fractions. A similar conclusion can be drawn on the contrast mechanism underpinning the reduction in  $T_1$  upon treatment with alisertib, based on both the known mechanism of response to alisertib through apoptosis in this model and the observed reduction in tumour volume in our study (268). However, this could not be confirmed using endpoint histopathological assessment, potentially due to the high inter-tumour heterogeneity both in terms of the amount of apoptosis

present at the time of enrolment (as shown by the endpoint histopathology in the vehicle cohorts) and in the actual response to alisertib treatment in this model as recently reported (268). The absence of any significant relative change in  $R_2^*$ , a validated biomarker for RBC (277), or  $T_2$ , ADC and MTR, which all relate to tissue water content/binding, strongly suggests that the overall decrease in  $T_1$  is being driven by the loss of the tissue fraction with high  $T_1$  values, i.e. regions with a high density of undifferentiated neuroblasts, rather than a gain of new MRI contrast e.g. that resulting from cell death-mediated release of paramagnetic ions (278, 279).

### *Why is $T_1$ sensitive to neuroblastoma histopathology and its modulation?*

By definition, the spin-lattice  $T_1$  relaxation time refers to the interaction or energy transfer between the excited  $^1\text{H}$  spin and the molecules within the surrounding molecular structure. The  $T_1$  value, i.e. the efficacy of the spin-lattice relaxation, is dependent on molecular tumbling of the molecule in which the proton resides. For MRI applications, this molecule is primarily water, which can be present in three states associated with different  $T_1$  values: *i*) free water (free to move, high  $T_1$ ), *ii*) "structured" water (bound to a macromolecule by a single bond where molecular tumbling is still possible, lower  $T_1$ ), *iii*) "bound" water (found in solids, bound by multiple bonds, high  $T_1$ ). The general consensus is that the reduced tissue  $T_1$  of structured water is a consequence of its interaction with proteins and other macromolecules. Tissue  $T_1$  thus depends on compartmentalization of structured water and the amount of molecular crowding within each different compartment. Cancer cells and tumour tissue typically have elevated  $T_1$  values compared to normal tissues, the original observation that demonstrated the potential of MRI for cancer diagnosis. Elevated tumour  $T_1$  remains attributed to a difference in intracellular water structure and order compared to normal cells (280, 281).  $T_1$  has also been suggested to change during cell cycle and mitosis *in vitro*, a phenomenon also attributed to different levels of water-macromolecule interactions (282, 283). However, very early work in MRI-detectable isolated large cells such as *Xenopus* oocytes and *Aplysia* neurons confirmed that cell nuclei exhibit higher  $T_1$  values than the cytoplasm (1.85 vs 1.2s respectively for *Xenopus* oocytes at 7T), and that degradation/permeabilization of the nuclear

envelope causes an equilibration of  $T_1$  values (284, 285). A more recent study reported anomalously rapid hydration water diffusion dynamics near DNA surfaces, which demonstrates that water interacts differently with DNA compared to protein. More precisely, water behaves like free water near DNA (286), which would explain both the higher nuclear  $T_1$ , and the change in  $T_1$  observed during mitosis when the chromatin is condensed and DNA is less accessible to water molecules and the nuclear membrane completely disappears.

Poorly or undifferentiated neuroblastoma are defined as small round nuclei with stippled chromatin (diffuse open chromatin) with scant eosinophilic cytoplasm and indistinct cell borders. This definition is thus self-explanatory for the higher  $T_1$  values reported here in areas of dense, undifferentiated neuroblasts (dense cells with a high nuclear/cytoplasmic ratio and minimal extracellular compartment). We can also assume that any reduction in undifferentiated cell density, or change in cell phenotype and/or intracellular compartmentalization, in a sufficiently large number of cells would thus result in a reduction in  $T_1$  (77). The reduced native  $T_1$  associated with dense areas of differentiating neuroblasts, characterized by lower nuclear to cytoplasmic ratio, lower cell density, and possibly surrounded by abundant eosinophilic neuropil, supports this hypothesis. Many of the events occurring during apoptosis, including water loss, pyknosis and karyorrhexis, would also align with a reduction in  $T_1$  if happening in a sufficient number of cells (287). Interestingly, both pyknosis and karyorrhexis are steps common to apoptosis, necrosis and senescence, indicating a potential generic sensitivity of  $T_1$  to cell death. As virtually all undifferentiated neuroblasts in this model are positive (and apoptotic cells negative) for the proliferation marker Ki67 (204), this hypothesis corroborates the studies by *McSheehy and colleagues* showing that a reduction in native  $T_1$  positively correlates with Ki67 staining (278, 279).

*Potential further clinical applications in guiding risk stratification and surgical planning and early clinical trials to develop new drugs*

The differential diagnosis and risk-stratification for children with neuroblastoma is based on criteria including histological features such as the grade of tumour differentiation. The sensitivity of T<sub>1</sub> mapping to regions rich in undifferentiated, apoptotic or differentiating neuroblasts seen in the Th-*MYCN* model herein suggests its potential to non-invasively classify tumours by favourable and unfavourable histology. It may also help identify anaplastic lymphoma kinase (ALK) positive regions, mutations associated with poor outcome in neuroblastoma, and for which small molecule inhibitors are currently being developed. Interestingly, ALK mutations have been shown to be associated with a differentiating molecular signature, confirmed at a pathological level in several *MYCN*- and *ALK*-mutated GEM models (206, 214, 249, 258). T<sub>1</sub> mapping may also afford additional prognostic value in confirmed cases of neuroblastoma, in which high cellular density of proliferative cells is associated with poor outcome, whereas a high density of apoptosis suggests a more favourable outcome (288). Finally, T<sub>1</sub> mapping may help identify the nature of tumours following the induction phase of frontline therapy, where it is uncertain if a mass is comprised of undifferentiated neuroblastoma or apoptotic or differentiated disease. In this regard, T<sub>1</sub>-mapping would provide additional and complementary information to semi-quantitative molecular imaging strategies such as MIBG and FDG-PET scans and help confirm the nature and heterogeneity of the disease associated with MIBG avid/non-avid and FDG (positive/negative) disease. This is important as discrepancies exist between the expression of the norepinephrine transporter (NET), responsible for the uptake of MIBG, and the presence of an aggressive cellular phenotype. These include MIBG non-avid disease that presents in ~10% of children, and reduced NET protein expression in high risk *MYCN*-amplified disease (289). Additionally, targeted therapies against *MYCN* or *ALK* can lead to the modulation of vascular perfusion (and hence the delivery of radiolabeled MIBG and FDG), glucose uptake and NET expression, which may potentially lead to a change in MIBG avidity which does not reflect, or makes it difficult to assess, changes in the extent of active disease using the current INRC guidelines. Treatment with the histone deacetylase inhibitor Vorinostat has for example been

shown to be effective against neuroblastoma while increasing NET transporter expression in neuroblastoma (290).  $T_1$  mapping has the potential to help improve the accuracy of detection of active disease for enhance surgical tissue sampling, surgical resection planning and response assessment.

### *Translating $T_1$ mapping into the neuroblastoma clinic*

The voxelwise quantification of  $T_1$  is an essential component of many MRI-based functional and molecular imaging techniques being developed to study the tumour microenvironment and for the evaluation of novel targeted therapies, including immunotherapy (291-296). DCE-MRI, arterial spin labelling (ASL-) MRI and oxygen-enhanced  $T_1$ -MRI are being evaluated clinically to assess tumour vascular perfusion/permeability and hypoxia. However, native tumour  $T_1$  maps acquired in the clinic are often only estimated and seldom reviewed or interpreted (41). In contrast, the clinical adoption of native  $T_1$  mapping has increased the potential for the non-invasive and differential diagnosis of cardiac pathology (297, 298) and the staging of chronic liver disease (299). The cardiac MR experience has shown that  $T_1$  mapping is simple to perform and analyse, minimally subjective, and highly reproducible (~2% CoV for a modified Look-Locker inversion recovery MOLLI sequence over 24 hours (300)). However, there are many acquisition schemes available for  $T_1$  mapping, and the measured  $T_1$  will depend on the precision and reproducibility of each scheme, and how is it affected by motion, flow and off-resonance effects. In our study, one of the advantages of the IR-TrueFISP technique, aside from its high accuracy, is that it is inherently flow compensated in the directions of slice selection and readout, especially at the blood velocity observed in tumours (301), allowing us to exclude changes in vascular flow as a source of reduction in native  $T_1$ . Moving forward, including  $T_1$ -mapping in an ethically-approved clinical study within the standard-of-care frontline chemotherapy would provide the study to rapidly evaluate and validate  $T_1$ -mapping potential for the neuroblastoma clinic. Such a study would also inform on the potential of native  $T_1$ -mapping to help better define bone and bone marrow metastasis and its response to treatment.

### *Beyond neuroblastoma*

The potential value of native T<sub>1</sub> reduction as a generic biomarker of early tumour response to therapy was first demonstrated by *McSheehy and colleagues* (278, 279). By understanding the spatial relationship of T<sub>1</sub> mapping with regional variations in neuroblastoma phenotype, new light is shed into the biology underpinning native T<sub>1</sub> contrast, based on cell anatomy. The data demonstrated in this chapter strongly support the use of T<sub>1</sub> mapping as a generic approach to assess early response to cancer treatment, especially since *i)* the “small-blue-round-cell tumour” phenotype, characterized by monotonous proliferations of small, undifferentiated or poorly differentiated cells with scant cytoplasm, is actually used to refer to the phenotype of a large group of highly aggressive tumours, including many high-risk paediatric malignancies such rhabdomyosarcoma and medulloblastoma (and adult cancers such as certain subtypes of sarcoma, carcinoma, lymphoma, and melanoma) and *ii)* both pyknosis and karyorrhexis are common steps to the major cell death processes. However, it would be important to understand the disease or tissue-specific factors, which may also affect native T<sub>1</sub> including the presence of oedema, fat or melanin.

## 5.5 Conclusion

In summary, Chapter 5 demonstrates that native  $T_1$  mapping can precisely and quantitatively map the rich histopathology of neuroblastoma tumours and its modulation by *MYCN*-targeted therapeutics in the clinically-relevant Th-*MYCN* model of neuroblastoma. By providing strong evidence for the sensitivity of native  $T_1$  to dense areas of undifferentiated neuroblasts, these data suggest further application for diagnosis, risk stratification and surgical planning, and that its potential as a biomarker of successful response to therapy could be extended to larger subsets of aggressive paediatric and adult tumours. Widely available on conventional clinical scanners, our study provides a strong rationale for the incorporation of  $T_1$  mapping both at the time of diagnosis and in early phase clinical trials to guide clinical decision making and the delivery of precision medicine to children with neuroblastoma.

## **Chapter 6 : Evaluating the sensitivity of ADC to childhood neuroblastoma pathology using gaussian mixture models and computational pathology *in vivo*.**

### **6.1 Introduction**

As discussed in Section 1.3.3.3, diffusion-weighted MRI is increasingly being used for assisting neuroblastoma diagnosis (302-304). The more benign forms, ganglioneuroma and ganglioneuroblastoma, are Schwannian stroma-rich tumours characterized by high median ADC value. On the contrary, aggressive undifferentiated or poorly-differentiated neuroblastoma are stroma-poor tumours characterized by monotonous proliferations of small, undifferentiated or poorly differentiated cells with scant cytoplasm, in which water diffusion is highly restricted (low ADC). *Jain and colleagues* recently reported a case study demonstrating the utility of DW-MRI in the diagnostic ganglioneuroblastoma nodular (GNBn) (305). DWI successfully detected and guided biopsy of a poorly differentiated neuroblastoma nodule within the context of a well-differentiated ganglioneuroma (Figure 1.7).

#### *6.1.1 Aim of this chapter*

The utility of ADC in the differential diagnosis in neuroblastoma is based on the discrimination of stroma-rich versus stroma-poor disease. In this chapter I use Gaussian mixture modelling and computational pathology methods in the MRI-histopathology registered datasets to further evaluate the sensitivity of diffusion-weighted imaging to the heterogeneous histopathology of stroma poor, poorly- and un-differentiated neuroblastoma using the Th-*MYCN* model of neuroblastoma. Since DWI is increasingly being used in the neuroblastoma clinic, and ADC is being actively investigated as an imaging biomarker of treatment response in oncology (306, 307), ADC is highly likely to be incorporated to evaluate response in early phase clinical trials designed to evaluate the efficacy of novel targeted therapies in children with high-risk neuroblastoma. Here, GMM analysis will be used to provide an essential understanding of the neuroblastoma-specific histological determinants of median ADC, which may explain why no apparent change in median tumour ADC over 24 hours was detected in the

tumours of the Th-MYCN model upon treatment with vistusertib, despite evidence of marked tissue response revealed by endpoint histological assessment.

## **6.2 Materials and methods**

*In vivo study design and protocol, including MRI acquisition and analysis, computational pathology methodologies for cell segmentation and classification are described in Chapter 5. The MRI-histopathological registration frame used in this chapter is described in Chapter 3.*

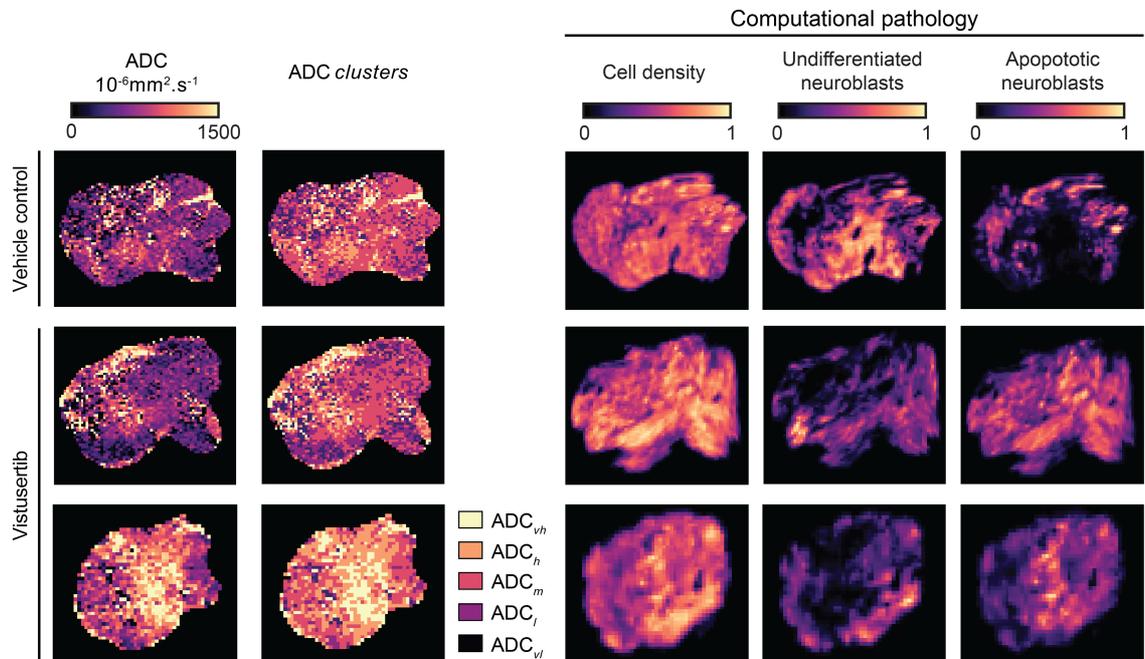
### *6.2.1 Gaussian mixture models for tumour clustering:*

From visual inspection of the registered ADC-H&E images, we identified 5 pathological determinants of regional heterogeneity in ADC maps (Figure 6.1): 1) red blood cells accumulation, 2) large areas of tissue damage with low ADC (restricted diffusion), 3) monotonous sheet of small undifferentiated cells, 4) areas rich in islands of differentiating neuroblasts separated by neuropil and 5) large areas of tissue damage associated with very high ADC (unrestricted diffusion). We used GMMs to cluster data in an unsupervised way. To complement the visual assessment, we computed the Bayesian and Akaike information criteria (BIC, AIC) to determine the appropriate number of clusters (2-10 were tested). GMMs were applied in 6 tumours from the vehicle-control group, containing representative proportions of all the areas of interest. A GMM was applied 1000 times on the ADC data with random initialisations and the median value of the threshold for each cluster was selected.

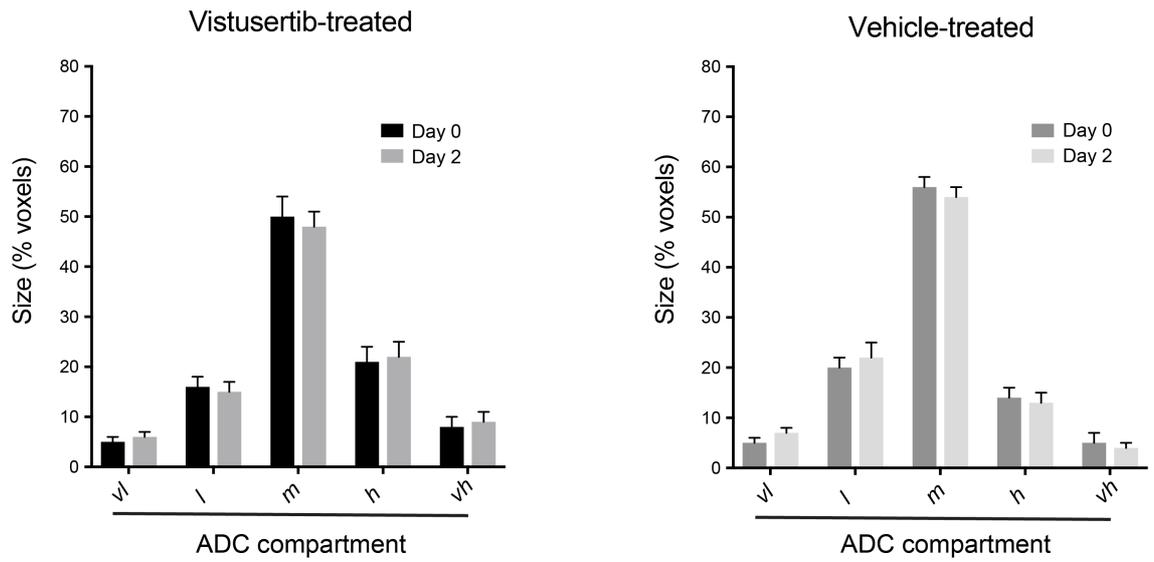
Both vehicle control and vistusertib-treated tumours were then classified using the defined clusters.

### 6.3 Results

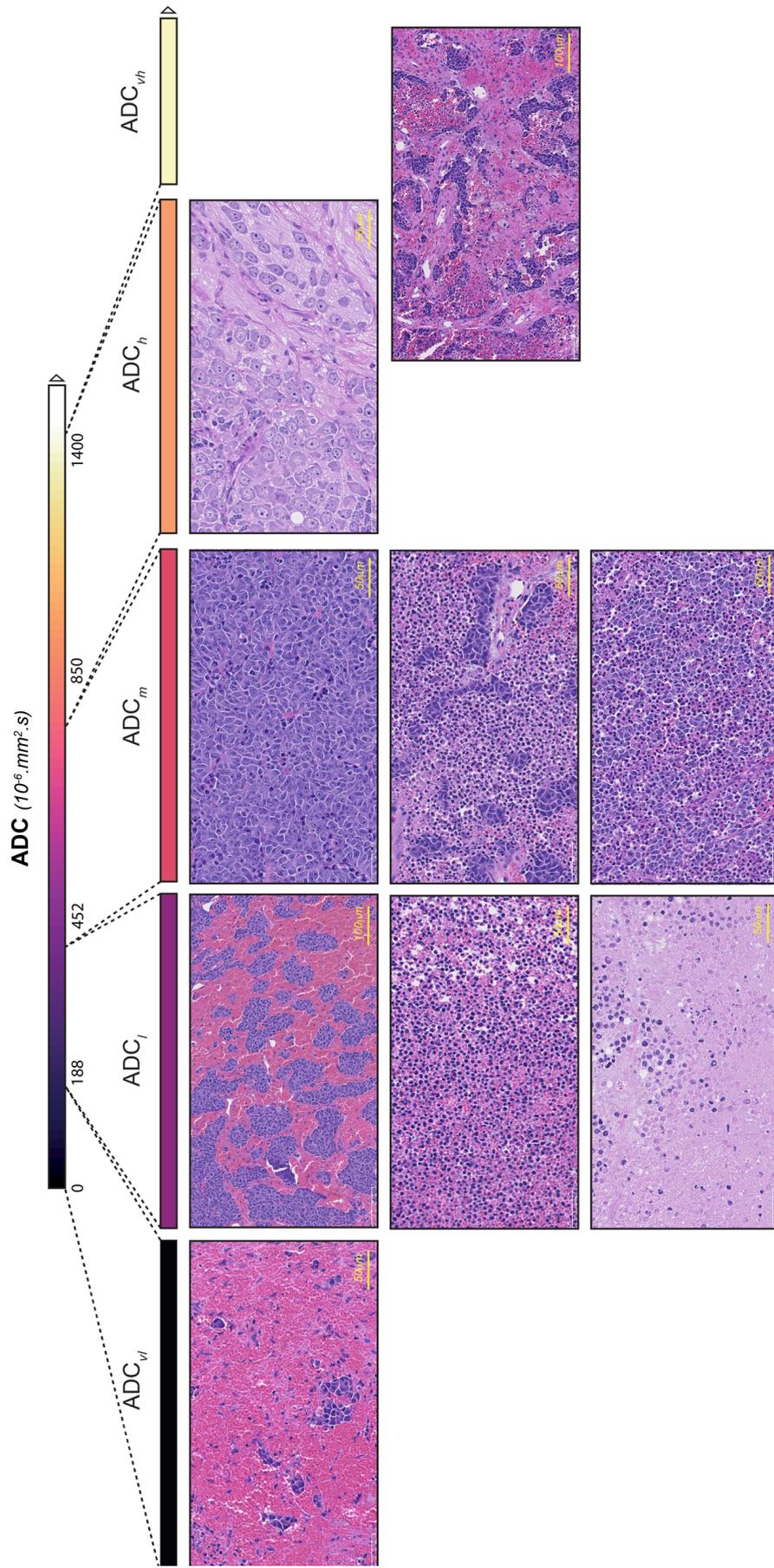
Both BIC and AIC criteria indicated to cluster the data using between 3 to 8 clusters, which we fixed to 5 based on the initial hypothesis. Application of the GMMs resulted in 5 compartments:  $ADC_{vh}$  ( $0-188 \cdot 10^{-6} \text{ mm}^2 \cdot \text{s}^{-1}$ ),  $ADC_l$  ( $188-452 \cdot 10^{-6} \text{ mm}^2 \cdot \text{s}^{-1}$ ),  $ADC_m$  ( $452-850 \cdot 10^{-6} \text{ mm}^2 \cdot \text{s}^{-1}$ ),  $ADC_h$  ( $850-1400 \cdot 10^{-6} \text{ mm}^2 \cdot \text{s}^{-1}$ ) and  $ADC_{vh}$  ( $1400-2000 \cdot 10^{-6} \text{ mm}^2 \cdot \text{s}^{-1}$ ). Figure 6.1 show representative examples of segmented maps according to the ADC compartments and Figure 6.2 plots the mean size the compartment of the different tumours used in this study. Figure 6.3 summarises representative histopathological features found in each of the ADC compartments.



**Figure 6.1** Representative maps of tumour ADC values and ADC compartments derived from Gaussian mixture models with registered histopathology-derived maps of cell density including undifferentiated and apoptotic neuroblasts in the Th-MYCN model of neuroblastoma, 24h following treatment with vehicle control or vistusertib.



**Figure 6.2** Gaussian mixture modelling of the apparent diffusion coefficient distribution in the tumours of the Th-MYCN model of neuroblastoma prior and after treatment with vistusertib or vehicle control.

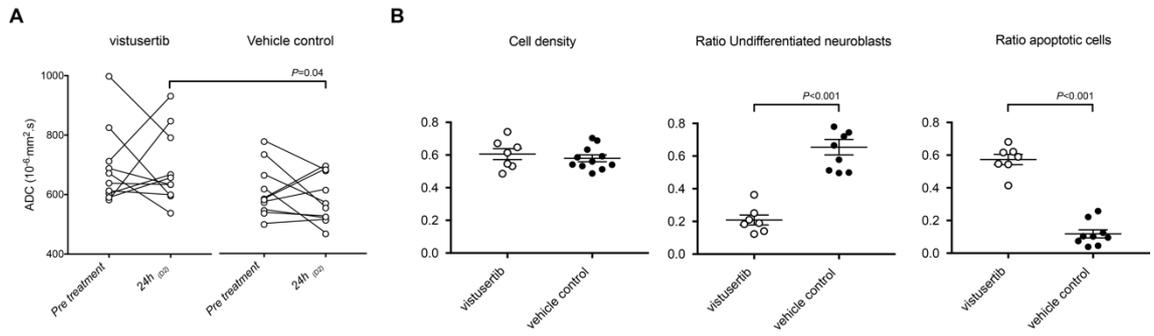


**Figure 6.3** Representative histopathological characteristic of the ADC compartments derived from Gaussian mixture modelling of the apparent diffusion coefficient distribution in the Th-MYCN model of neuroblastoma.

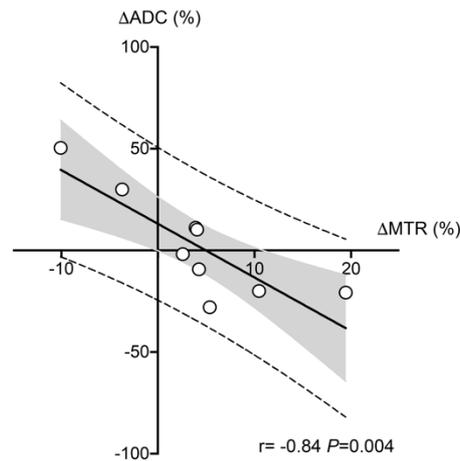
### *6.3.1 Median tumour ADC and ADC compartments are not sensitive to tissue response following treatment with vistusertib.*

As discussed in Chapter 5, treatment with vistusertib was not associated with any significant change in tumour median ADC over 24h (Figure 6.4). There was no correlation between the change in tumour median ADC and the change in native  $T_1$  which is shown to be a robust marker of response to treatment with vistusertib in Chapter 5. Interestingly there was a significant and strong correlation between change in ADC and MTR over treatment with vistusertib (Figure 6.5). However, the absence of significant change in median tumour MTR further corroborated the lack of sensitivity of median tumour ADC to tissue response to vistusertib. Vistusertib treatment was not associated with any change in the size of any of the ADC compartments, despite the significant reduction in tumour burden in all treated mice. Interestingly the median tumour ADC in vistusertib-treated tumours was higher than in the vehicle control cohort at the 24 hours timepoint ( $689 \pm 40 \cdot 10^{-6} \text{ mm}^2 \cdot \text{s}^{-1}$  vs  $584 \pm 26 \cdot 10^{-6} \text{ mm}^2 \cdot \text{s}^{-1}$ ,  $P = 0.04$ ). This was particularly associated with significantly larger  $\text{ADC}_h$  and  $\text{ADC}_{vh}$  compartments in treated tumours ( $22 \pm 3\%$  vs  $13 \pm 3\%$ ,  $P = 0.03$  and  $9 \pm 2\%$  vs  $4 \pm 1\%$ ,  $P = 0.01$  respectively). It is, however, important to stress that there was no significant change in these two compartments over treatment.

Visual inspection of H&E-stained images revealed marked tissue remodelling in tumour treated tumours, which was confirmed quantitatively by the significantly lower fraction of undifferentiated neuroblasts ( $66 \pm 4\%$  vs  $21 \pm 3\%$ ,  $P < 0.0001$ ) and higher fraction in apoptotic cells ( $17 \pm 3\%$  vs  $57 \pm 3\%$ ,  $P < 0.0001$ ) in vistusertib-treated tumours compared to vehicle control tumours at 24h. However, there was no significant difference in the overall cell density between the two groups (Figure 6.5).



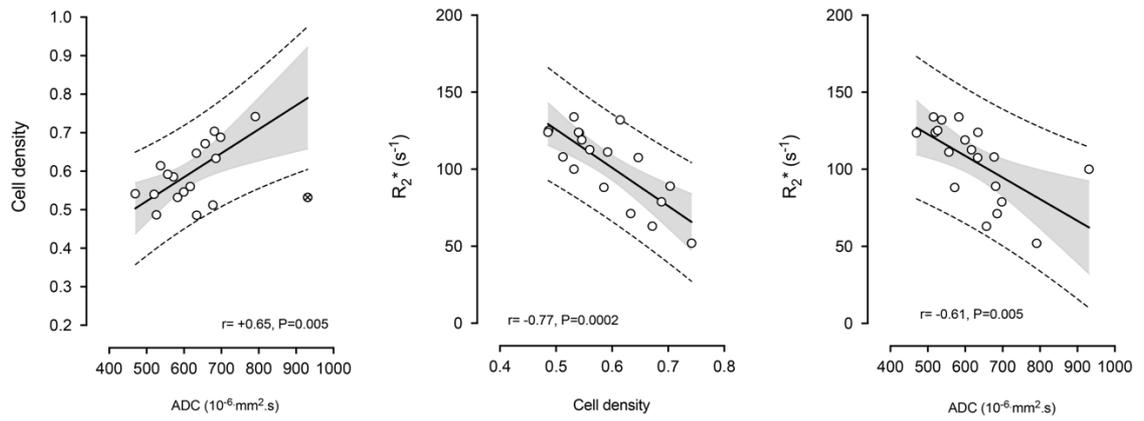
**Figure 6.4 A.** Changes in tumour median apparent diffusion coefficient (ADC) in the Th-MYCN model of neuroblastoma 24h following treatment with vehicle control or vistusertib. Note that there was not significant change in tumour median ADC over treatment. Interestingly there was significant difference in tumour median ADC between treated and vehicle control tumours at the study endpoint (24h). **B.** Difference in tumour median cell density, the ratio of undifferentiated neuroblasts and apoptotic cells to all cells in tumours 24h after treatment with vehicle control or vistusertib.



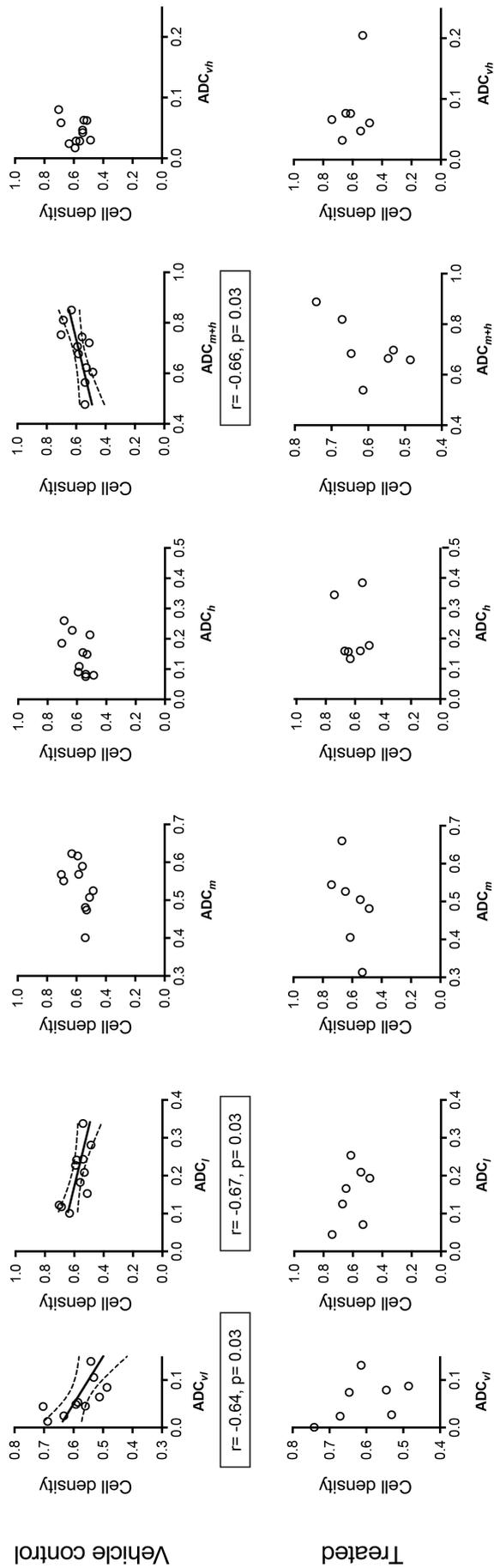
**Figure 6.5** Relationship between the change in median tumour ADC and the change in tumour median magnetisation transfer ratio (MTR) in tumours of the Th-MYCN model of neuroblastoma 24h after treatment with vistusertib.

### 6.3.2 Low native tumour ADC correlates with high red blood cell content

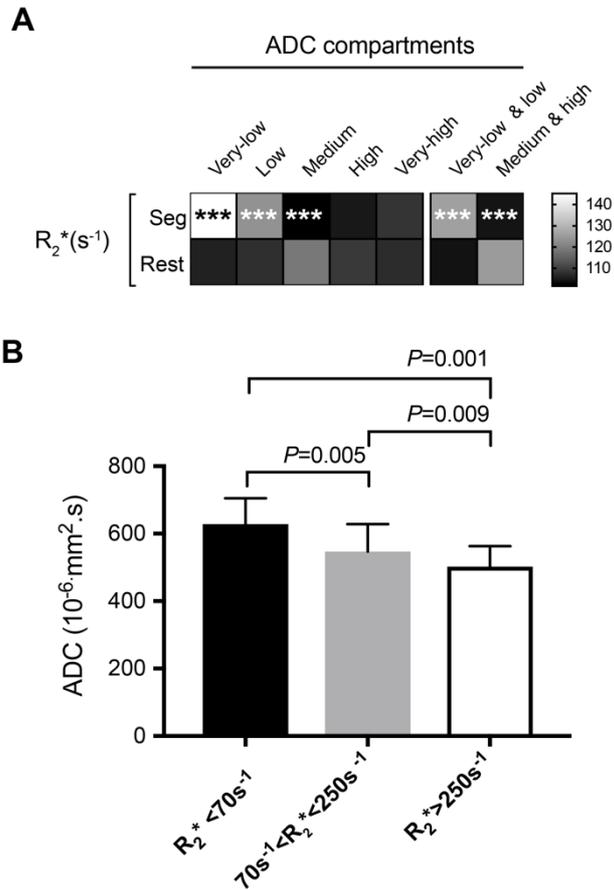
Counterintuitively, median tumour ADC correlated positively with median tumour cell density ( $r = +0.65$ ,  $P = 0.005$ , Figure 6.6). This result was confirmed by the compartment analysis with  $ADC_m + ADC_h$ , containing the majority of the cells, positively correlating with the overall cell density in vehicle control ( $r = 0.66$ ,  $P = 0.03$ , Figure 6.7). A significant negative correlation was also found between both  $ADC_{vl}$  and  $ADC_l$  with cell density in the vehicle control ( $r = -0.64$ ,  $P = 0.03$  and  $r = -0.67$ ,  $P = 0.03$ ). Based on the original hypothesis, we compared ADC maps with  $R_2^*$  maps showing that regions with high  $R_2^*$ , which correspond to areas rich in RBC as demonstrated in Chapter 4, were associated with low ADC values. Sub-regional analysis corroborated this observation by demonstrating first that  $ADC_{vl}$  and  $ADC_l$  tumour compartments, had significantly higher median  $R_2^*$  than the rest of the tumours (Figure 6.8). Reciprocally, sub-regional analysis using the empirical threshold on  $R_2^*$  maps defined in Chapter 4 ( $R_2^* < 70s^{-1}$  as no haemorrhage (308, 309) and  $> 250s^{-1}$  as purely RBC, and mixed regions of neuroblasts and RBC for the  $R_2^*$  values between) shows significantly different median ADC values for each  $R_2^*$  based compartment with ADC decreasing as  $R_2^*$  increased. Comparison of median tumour values showed significant negative correlation between  $R_2^*$  and cell density ( $r = -0.77$ ,  $P = 0.0002$ ) and  $R_2^*$  and ADC ( $r = -0.61$ ,  $P = 0.005$ ), indicating that red blood cell aggregation is a major determinant of low median ADC.



**Figure 6.6** Relationships between median tumour ADC, median tumour  $R_2^*$  and median tumour cell density in neuroblasts in tumours of the Th-MYCN model of neuroblastoma 24h after treatment with vistusertib or vehicle control.



**Figure 6.7** Relationships between ADC compartments derived from Gaussian mixture models and median tumour cell density in neuroblasts in tumours of the Th-MYC/N model of neuroblastoma 24h after treatment with vistusertib or vehicle control.



**Figure 6.8 A.** Sub-regional analysis demonstrates that several ADC compartments have significantly different median  $R_2^*$ . \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.005$ , Wilcoxon signed rank test. **B.** Box-and-whisker plot showing the difference in ADC in sub-regions categorized by low ( $< 70 \text{ s}^{-1}$ ), intermediate ( $70 \text{ s}^{-1} < R_2^* < 250 \text{ s}^{-1}$ ) and high ( $> 250 \text{ s}^{-1}$ ) values of  $R_2^*$  measured in Th-MYCN tumours treated with vehicle ( $n = 13$ ). Data are medians and interquartile range.

### 6.3.3 Increasing cell death is associated with a gradual reduction in ADC.

In treated tumours, the change in the proportion of cells composing the tumours would potentially explain why the relation between the ADC compartments and cell density disappeared. Figure 6.1 illustrates discrepancies in the relation between ADC and the underlying pathology of neuroblastoma, showing *i)* a treated tumour with no apparent difference in ADC or ADC compartment maps compared to a vehicle control despite marked differences in cellular composition (viable vs apoptotic cells) and *ii)* two treated tumours with marked differences in ADC and ADC compartment maps but that appear to have similar phenotype (reduced viable cells and high fraction of apoptotic cells). Sub-regional analysis was thus subsequently carried out to characterise the cellular composition of each compartment in vehicle control and treated tumours. The median  $T_1$  values, which provide insight on the longitudinal changes occurring upon treatment, were also included in the analysis having established the sensitivity of  $T_1$  to the underlying pathology of neuroblastoma, particularly its sensitivity for regional high density of undifferentiated neuroblasts (Chapter 5).

The ADC compartments were associated with differences in  $T_1$ , cell density and undifferentiated neuroblast density values.

In vehicle-control tumours:

- Not surprisingly, the  $ADC_{vl}$  compartment had the significantly lowest median  $T_1$  and cell density, consistent with a high  $R_2^*$  value, which has been shown to be associated with a high content of RBC and very low undifferentiated neuroblast density (Figure 6.9).
- For the other compartments, median  $T_1$  values correlated with the difference in undifferentiated density, reinforcing the undifferentiated cells density as a major determinant of native  $T_1$  in neuroblastoma.  $ADC_m$  and  $ADC_h$  compartments contained higher density of undifferentiated neuroblasts and had significantly higher  $T_1$  value than the rest of the tumour.

- Finally,  $ADC_{vh}$ , which corresponds to  $ADC > 1400 \cdot 10^{-6} \text{ mm}^2 \cdot \text{s}^{-1}$ , associated with un- or less-restricted diffusion, was consistently associated with lower  $T_1$  median value and lower undifferentiated cell density. As a matter a fact, the size of  $ADC_{vh}$  rapidly decreased with the increasing fraction of undifferentiated neuroblasts in vehicle control tumours ( $r = -0.85$ ,  $P = 0.001$ ).

Upon treatment with vistusertib:

- The median  $T_1$  of  $ADC_l$ ,  $ADC_m$ ,  $ADC_h$  and  $ADC_{vh}$  compartments, significantly dropped corroborating the significantly lower density of undifferentiated neuroblasts and concomitant higher density of apoptotic cells found in these compartments treated at 24h compared to vehicle control.
- There was no correlation between the change in median native  $T_1$  and change in ADC upon treatment neither was a correlation between native  $T_1$  and ADC in the treated tumours. Note that  $ADC_l$ ,  $ADC_m$ ,  $ADC_h$ , and  $ADC_{vh}$  compartments have similar density of apoptotic cells in the treated tumours.

In treated tumours, the size of ADC compartments was associated the density of undifferentiated neuroblasts:

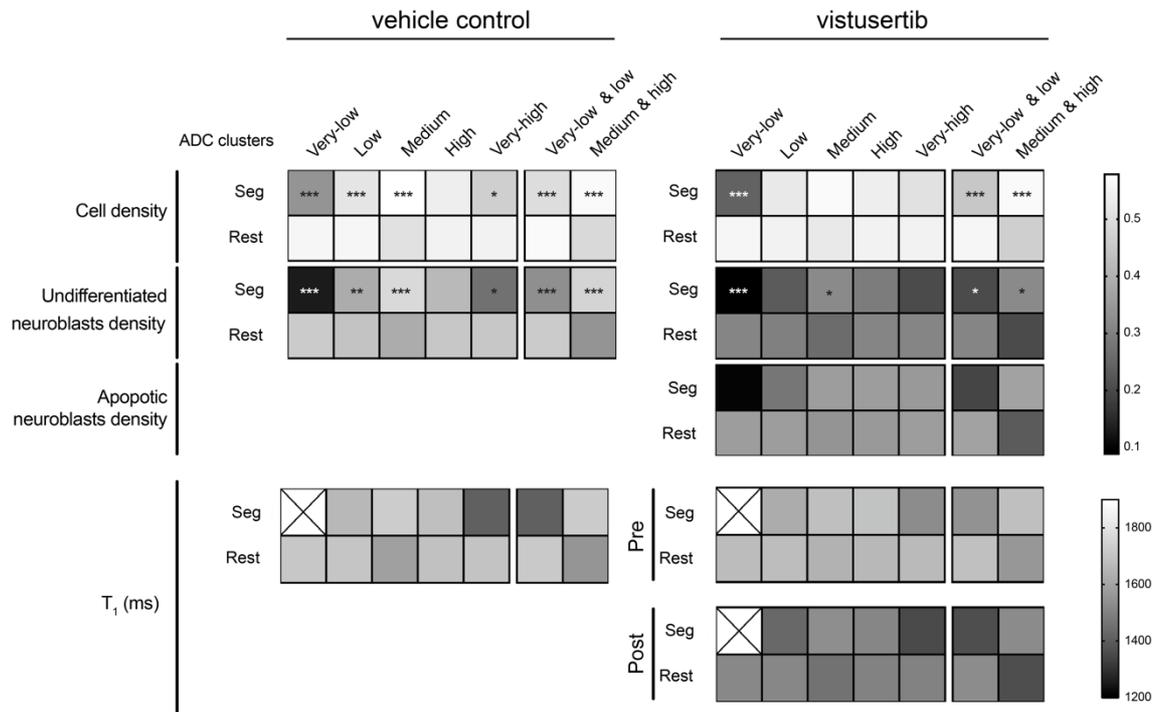
- The size of compartments  $ADC_h$  and  $ADC_{m+h}$  significantly and positively correlated to the ratio of undifferentiated cells ( $r = +0.81$ ,  $P = 0.03$  and  $r = +0.77$ ,  $P = 0.02$ , respectively).
- The size of compartments  $ADC_{vl}$  and  $ADC_l$  inversely correlated to the ratio of undifferentiated neuroblasts ( $r = -0.81$ ,  $P = 0.03$  and  $r = -0.83$ ,  $P = 0.02$ , respectively) (Figure 6.10).

Conversely, these relations were inverted when compared with the ratio of apoptotic cells:

- The size of compartments  $ADC_h$  and  $ADC_{m+h}$  negatively correlated to the ratio of apoptotic cells ( $r = -0.76$ ,  $P = 0.04$  and  $r = -0.80$ ,  $P = 0.03$ , respectively).
- The size of compartments  $ADC_{vl}$  and  $ADC_l$  positively correlated to the ratio of apoptotic cells ( $r = +0.83$ ,  $P = 0.03$  and  $r = -0.84$ ,  $P = 0.02$ , respectively).

Retrospective visual inspection of these regions in the original H&E images showed that large areas with a high amount of apoptotic cells representatively belong to the  $ADC_m$  compartment (which contains the regions with high density in undifferentiated neuroblastoma in control tumours), while regions in which the level of apoptosis increased towards secondary necrosis or large regions of tissue damage are contained in the  $ADC_l$  compartment, regardless of the presence of red blood cells.

Finally, in vehicle control tumours, there was a significant negative correlation between the size of the  $ADC_{vh}$  compartment, associated with unrestricted diffusion, and the density of undifferentiated neuroblasts. This relationship was characterised by a steep slope indicative of the strong effect of the characteristic tumour phenotype of dense small blue cells on that compartment.



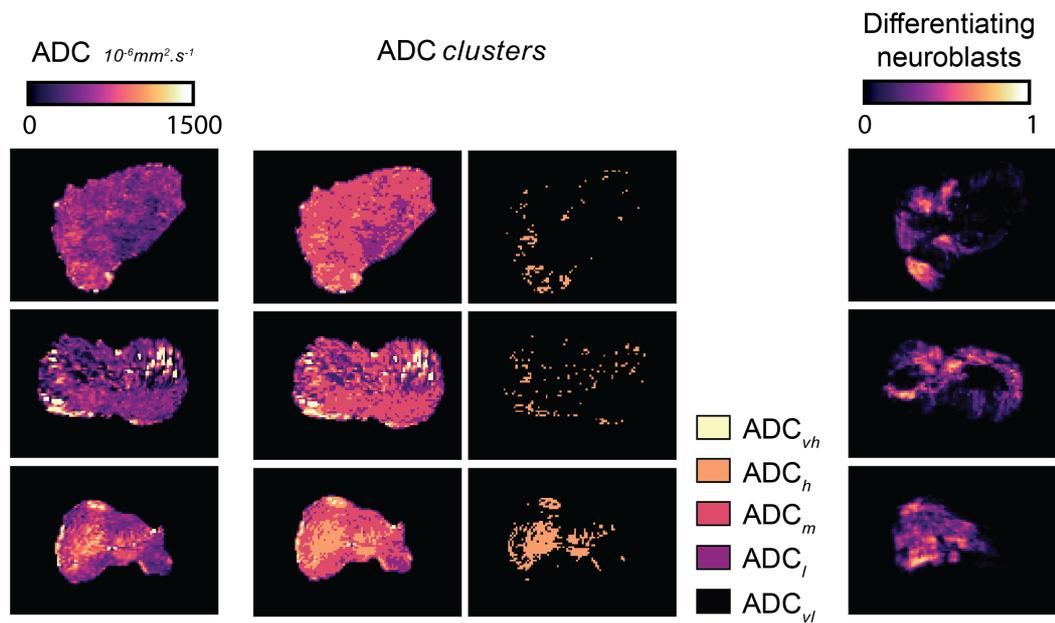
**Figure 6.9** Sub-regional analysis defining the cell density, the fraction of undifferentiated neuroblasts and apoptotic cells, and native T<sub>1</sub> value associated with the ADC compartment \*P<0.05, \*\*P<0.01, \*\*\*P<0.005, Wilcoxon signed rank test.



#### *6.3.4 ADC is sensitive to areas of therapy induced maturation and areas rich in differentiating neuroblasts*

Some of the treated tumours did show higher ADC and consequently larger  $ADC_h$  and  $ADC_{vh}$  regions as shown in Figure 6.1. These regions spatially associated with regions with a high fraction of apoptotic neuroblasts. Yet visual inspection of these regions on the H&E images revealed areas with sparsely distributed mature ganglion (not normal) cells and abundant neuropil in the background, consistent with neuroblastoma maturation towards a ganglioneuroblastoma often seen at the time of surgical resection following induction chemotherapy.

Based on the initial hypothesis, segmented  $ADC_h$  regions spatially associated with regions rich in differentiating neuroblasts (Figure 6.11), usually characterised by islands of larger cells separated by abundant neuropil in the three control tumours characterised by a more differentiating phenotype described in Chapters 4 and 5.



**Figure 6.11**  $ADC_h$  compartment derived from Gaussian mixture modelling spatially associates with area dense in differentiating neuroblasts in 3 cases of differentiating tumours arising in the Th-MYCN model of neuroblastoma.

## 6.4 Discussion

The insights into neuroblastoma histopathology developed in the previous chapters guided the application of GMMs to derive ADC compartments in this chapter. This approach further demonstrated the sensitivity of DWI to tumour microstructure and the rich histopathology of neuroblastoma focusing on the characteristic Schwannian stroma-poor, poorly- or un-differentiated phenotype, which is characteristic of high-risk disease, which is faithfully recapitulated in the tumours of the Th-*MYCN* mouse model. However, this MRI-histopathology cross validation study demonstrates that median/mean tumour ADC cannot be used as a reliable early sensitive pharmacodynamic biomarker of response to targeted treatment in the Th-*MYCN* model. This study particularly demonstrates that ADC is unequivocally insensitive to the morphological cellular changes occurring during the apoptotic process and that the consequences of cell death on tumour architecture are associated with conflicting effects on ADC measurements; in particular extensive regions of tissue damage or convergent cell death, haemorrhagic or not, are associated with reduced ADC values, while tissue maturation and cellular differentiation are associated with high ADC values.

DWI has become a powerful tool in the diagnostic oncology clinic as malignant lesions often have a lower ADC due to restricted water diffusivity (306, 307). This is mainly attributed to elevated tissue cellularity or density compared to normal tissue or more benign tumours which restrict translational water diffusion, which justified the introduction of DWI in the neuroblastoma diagnostic clinic (302-305, 310). However, median/mean ADC can also be influenced by the ratio between free and bound water, the difference in the tortuosity of the extracellular space and the composition of the extracellular matrix. Focal necrosis, glandular structure and fibrosis are among the histopathological features that can also affect mean/median tumour ADC (311). ADC measurements are also affected by  $^1\text{H}$  relaxation ( $T_1$  and  $T_2$ ) and paramagnetic substances. This study demonstrates that the main determinant of mean/median tumour ADC in the Th-*MYCN* mouse, is the extent of red blood cells extravasation associated with reduced ADC, consistent with the presentation of intracranial hematomas in the subacute phase

(during which RBC are intact) (312). As discussed in Chapter 4, clinically, the haemorrhagic nature of undifferentiated neuroblastoma, including nodular neuroblastoma components in the context of GNBn, is well described (313, 314), yet its routine reporting remains only anecdotal. It would be extremely valuable to introduce IS-MRI in the neuroblastoma clinic to understand the contribution of haemorrhage to the lower ADC values of high-risk disease in the clinic and its potential role to enhance the specificity of ADC in the differential diagnosis of neuroblastoma from its benign counterparts.

The potential of ADC for early-detection of treatment response in solid tumours has been actively investigated (306, 307), based on the assumption that conventional chemotherapy or vascular-targeted therapy leads to early alterations in tissue microstructure: cell swelling and cell death are accompanied by a loss of cell membrane integrity and changes to tumour cell density. As these changes often precede a change in tumour volume, the resulting increase in the interstitial space leads to an increase in freedom of water molecules to move and an associated increase of the measured ADC. This was demonstrated in several *in vivo* studies in subcutaneous models of cancer (306, 307). This is true in the context of liquefactive necrosis, in which lysed cells are converted into a fluid phase within hours. Liquefactive necrosis is usually associated with bacterial infection and with ischemic injury in the brain (a poorly understood exception in the body) (315). Yet liquefactive necrosis is common in subcutaneous xenograft models of cancer (usually visible as hypertense regions on T<sub>2</sub>-weighted images). However, preclinical and clinical studies have shown the cyclic nature of the ADC response, due to contradictory effect of cell swelling and cell death on ADC. Others have shown no change in ADC, suggesting an incomplete necrotic process where partially-ruptured cell membranes remain an obstacle to water diffusivity (316), highlighting the crucial need for careful timing necessary to meaningfully interpret ADC measurements. In the Th-MYCN model, the extensive regions of tissue damage or convergent cell death are consistent with the appearance of coagulative necrosis (315). In contrast to liquefactive necrosis, coagulative necrosis, the other major pattern of necrosis, is characterized by the maintenance of normal architecture of necrotic tissue for several days after cell death and is the result of infraction or ischemia, and characteristic of radiation-

induced necrosis. Characterised by homogeneous clusters and sheets of dead and degraded tumour cells that coalesce into an amorphous coagulum, coagulative necrosis has, not surprisingly, the opposite effect on ADC, i.e. it is characterised by extremely low ADC, as shown here in the Th-*MYCN* and previously in high-grade glioma (317, 318). Coagulative necrosis and may also be the major pathological determinant of the strong inverse correlation observed between the relative changes in median ADC and MTR over treatment with vistusertib.

The sensitivity of ADC, and especially an increase in ADC, to therapy-inducing apoptosis is a little more contentious, as illustrated by a study showing the sensitivity of ADC to conventional chemotherapy agent irinotecan but not to the apoptosis-inducing drug birinapant in a xenograft model of human colon carcinoma (319). Also, careful discrimination should be made between the terms “early biomarkers of therapy-induced apoptosis” and “early biomarkers of response to apoptosis-inducing drugs”. The Th-*MYCN* model of neuroblastoma recapitulates the *MYCN* driven, apoptosis-primed state in neuroblastoma, which thereby sensitizes the neuroblastoma cells to drug-induced apoptosis (320, 321). Destabilization of the *MYCN* protein by mTor inhibitors including AZD8055, the analogue of the clinical lead vistusertib, has been shown to lead to the rapid induction of widespread apoptosis in the Th-*MYCN* model (322) and thus represented an appropriate model system to understand the sensitivity of ADC for apoptosis in neuroblastoma. This MRI-cross validation study shows that tumour ADC is not an early biomarker of apoptosis, as supported by the ADC<sub>m</sub> compartment containing both regions of dense undifferentiated and regions dense in apoptotic cells. This is however not surprising as apoptosis is characterised by a series of typical morphological events designed to prevent the release of intracellular content which would lead to an inflammatory response (323). These well-orchestrated events include shrinkage of the cell, the hypothesis for ADC potential sensitivity to apoptosis, and the fragmentation into membrane-bound apoptotic bodies. However, while this leads to the reduction in the overall cell volume, it is associated with no change or can even be associated with an increase in surface area due to the creation of the apoptotic bodies as elegantly described by *Nunez and colleagues* (324). The increase in cell

membrane surface area will cause additional restriction to translational water diffusion. The final stage in the apoptotic process is the phagocytosis of the apoptotic bodies by neighbouring cells including macrophages. If phagocytosis does not happen, apoptotic bodies will be degraded in a process similar to necrosis, defined secondary necrosis. The full process is estimated to take 6-24h *in vivo* and as a result, within the time frame of our study. In subcutaneous xenografts, successful targeted therapy seldom leads to a rapid reduction in tumour burden but often leads to cytostatic/lentic effects over days. As a result, the apoptotic process within a persistent tumour volume will lead to an increase in extracellular space resulting in increased ADC. In this case, ADC could be considered an early biomarker of response to apoptosis-inducing treatment. However, as mentioned above, secondary (liquefactive) necrosis, native extracellular space and existing necrotic (liquefactive) areas, necrosis induced by inhibition of downstream signalling (i.e. angiogenesis) could also be responsible for an intermixed or combined necrosis/apoptosis scenario resulting in an increase in ADC. In the case of the Th-*MYCN* mouse model, the native reduced extracellular space (stroma-poor), the coagulative nature of necrosis and finally the rapid reduction in tumour volume, which emulates the chemosensitivity of *MYCN*-amplified tumours to frontline therapy, can only lead to a reduction in tumour ADC in the Th-*MYCN* tumours.

This is actually contradictory with the small number of preliminary clinical studies that have shown a consistent increase in ADC in neuroblastic tumours especially with those showing a good response following induction chemotherapy (325-328). However, this is after over >70 days of intensive chemotherapy (well beyond the time scale of the cell-death process) and the histopathology of the resected tumours has been shown to mirror the wide range of RECIST responses with the presence necrosis, scarring and areas of neuroblastoma-like, ganglioneuroblastoma-like and ganglioneuroma-like differentiation. Remarkable RECIST responses are typically associated with neuroblastoma maturation, which would thus be consistent with an increase in tumour ADC. In this study, we demonstrate the sensitivity of the higher GMM compartments  $ADC_h$  to regions rich in differentiating neuroblastoma, characterised by islands of larger cells separated with increased neuropil background. Vistusertib-treated tumour

regions belonging to the  $ADC_h$  and  $ADC_{vh}$  compartments were associated with regions that resemble chemotherapy-induced neuroblastoma maturation, demonstrating that as early as 24 hours, histological features with opposite effect on ADC can coexist, thereby rendering tumour ADC inadequate to assess response to treatment, at least on its own. As a matter of fact, the strong correlation between the change in ADC and the change in MTR over treatment also suggests a change in the content of free water, with the presence of positive and negative changes in MTR over treatment potentially reflecting opposite contributions from cell death and tissue maturation. Based on the results obtained in Chapter 5, incorporating  $T_1$ -mapping in the routine clinic would improve specificity of ADC for active disease and cell death in neuroblastoma. A similar approach combining DWI and native  $T_1$ -weighted MRI has been shown to reliably identify and quantify necrosis clinically in Wilm's tumours, the most common paediatric renal tumours (329).

## **6.5 Conclusion**

This study re-iterates the great sensitivity but an apparent lack of specificity of ADC for tumour microstructure and the rich pathology of high-risk neuroblastoma. Based on the results presented in Chapter 4 and 5, Chapter 6 warrants further investigation into a multi-parametric or habitat imaging approach combining DWI, IS-MRI and  $T_1$ -mapping to improve the specificity of MRI-based functional imaging to map the hallmarks of neuroblastoma pathology with the aim to enhance differential diagnosis and early response assessment to therapy.

## Chapter 7 : Habitat imaging of childhood neuroblastoma using multi-parametric MRI and co-registered histopathology

### 7.1 Aim of this chapter

The previous chapters report on the in-depth understanding of the pathological determinants of single MRI biomarkers and their sensitivity to some of neuroblastoma histological hallmarks used clinically for risk-adaptation. As introduced in section 1.2.3, habitat imaging refers to the combination of the different MRI biomarkers to reveal regional differences in phenotype. Most of habitat imaging approaches are based on the clustering of the MRI data, followed by the attempt to histologically validate and explain the resulting clusters. The MRI-histopathology framework described in Chapter 3 potentially allows the use of more direct supervised learning approaches, using the quantitative computational histology data (e.g. cell density) as labels to train machine learning algorithms and predict the underlying histology from the multi-parametric MRI data.

This chapter reports on a proof-of-concept study that explores the potential of combining  $R_2^*$ , native  $T_1$ , and ADC to predict tumour histological properties and non-invasively map neuroblastoma pathology using two types of supervised learning approaches:

- *Classification.* The results from the regional classification enabled by *SuperHistopath* (Chapter 2) are used as output (ground-truth) labels to train well-established machine learning algorithms to classify MRI values as distinct histologically-defined sub-regions (habitats).
- *Regression.* The approach used to match the resolution of histology images to the MRI, results in continuous variables (e.g. density of undifferentiated neuroblasts in a pre-defined area), which are used as outputs for regression-based approaches for training algorithms to directly predict the values of the histological parameter from the MRI data.

## 7.2 Materials and Methods

*In vivo study design and protocol, including MRI acquisition and analysis, computational pathology methodologies for cell segmentation and classification are described in Chapter 5. The MRI-histopathological registration framework used in this chapter is described in Chapter 3.*

For this proof-of-concept study, 11 vehicle-control Th-MYCN mouse tumours were used from the cohort described in Chapters 5 and 6.

### *7.2.1 Classifying tumour regions based on restrictions*

Based on the knowledge acquired from the studies described in Chapters 4, 5 and 6, specific tumour regions (e.g. haemorrhage) can already be classified with high certainty without the need of a machine learning classifier. This section aimed to calculate what percentage of the tumour can be confidently mapped using simple robust restrictions based on our understanding of neuroblastoma pathology. A summary of these assumptions is presented in Table 7.1.

<b>Tumour region</b>	<b>Restriction</b>
<b>Haemorrhage</b>	$R_2^* \geq 250 \text{ s}^{-1}$ <b>OR</b> $T_1 \leq 1000 \text{ ms}$ <b>OR</b> $\text{ADC} \leq 100 \cdot 10^{-6} \text{ mm}^2 \cdot \text{s}^{-1}$
<b>Extensive cell apoptosis/necrosis</b>	$\text{ADC} \leq 350 \cdot 10^{-6} \text{ mm}^2 \cdot \text{s}^{-1}$ <b>AND</b> $T_1 \leq 1500 \text{ ms}$ <b>AND</b> $R_2^* \geq 170 \text{ s}^{-1}$
<b>Undifferentiated neuroblasts without the presence of blood</b>	$1900 < T_1 < 2300 \text{ ms}$

**Table 7.1** A summary of the empirical thresholds used to confidently characterise specific tumour regions in Th-MYCN mouse model of neuroblastoma.

### 7.2.2 Habitat characterisation based on supervised machine learning

The purpose of this section is to classify each pixel within the tumour into one of four distinct types of habitats: 1) region of undifferentiated neuroblasts, 2) extensive cell apoptosis/necrosis (tissue damage), 3) region of differentiating neuroblasts, and 4) haemorrhage. SuperHistopath results will be used as labels to train a machine learning algorithm to classify the MRI data.

#### 7.2.2.1 Spatial registration of MRI and SuperHistopath region classifier

As described in section 3.3, the histology images are divided into non-overlapping square grids of 518x518 pixels and a representative value for each area is needed to process them to MRI resolution. Since each pixel in this area can have different classification values, the representative value of the area is determined by the class with the most occurrences. The results were registered to the native  $T_1$  maps using the CPD algorithm as described in Chapter 3. The method of 'nearest neighbour' was used for interpolation.

#### 7.2.2.2 Training dataset

For training, I selected labels from four tumours (mice 2, 6, 7, 11) with a relatively balanced representation of all four classes and in which good registration was achieved upon visual inspection. This resulted in 5091 labels for regions of undifferentiated neuroblasts, 1211 labels for regions of extensive cell apoptosis/necrosis/tissue maturation, 677 labels for regions of differentiating neuroblasts and 616 labels for haemorrhage. To address the class imbalance problem, 1500/5091 labels for regions of undifferentiated neuroblasts were randomly selected.

The resulting dataset was randomly split into training and validation (80% and 20% respectively).

#### 7.2.2.3 Features

Each MRI pixel was represented by three features:  $R_2^*$ ,  $T_1$  and ADC values. To address the pixel-to-pixel mismatches between MRI and histology, median values of the pixels' square neighbourhood (2 pixels width) for both  $T_1$  and ADC were added. The correlation between  $R_2^*$  values and their neighbourhood values was too high ( $r > 0.9$ ) and therefore the value of the  $R_2^*$  pixel neighbourhood was not used as a feature.

Each feature was normalised to the range of 0 to 1 by dividing it with a maximum value ( $300\text{s}^{-1}$  for  $R_2^*$  and 3000ms for both  $T_1$  and ADC).

#### 7.2.2.4 Machine learning algorithms

A wide variety of machine learning algorithms for classification were incorporated, including models of logistic regression, linear discriminant analysis (LDA), k-nearest neighbour (KNN), decision trees (CART), Gaussian naïve Bayes and support vector machines (SVM). Ensemble methods were also used including Adaboost, gradient boosting machines (GBM), random forest and extra trees

classifiers. The models were initially assessed using 10-fold cross-validation using the accuracy metric.

All code was written in Python (v. 3.6.10) and the models were trained with the scikit-learn library (v. 0.23.2).

### *7.2.3 Prediction of key histological features using a regression approach*

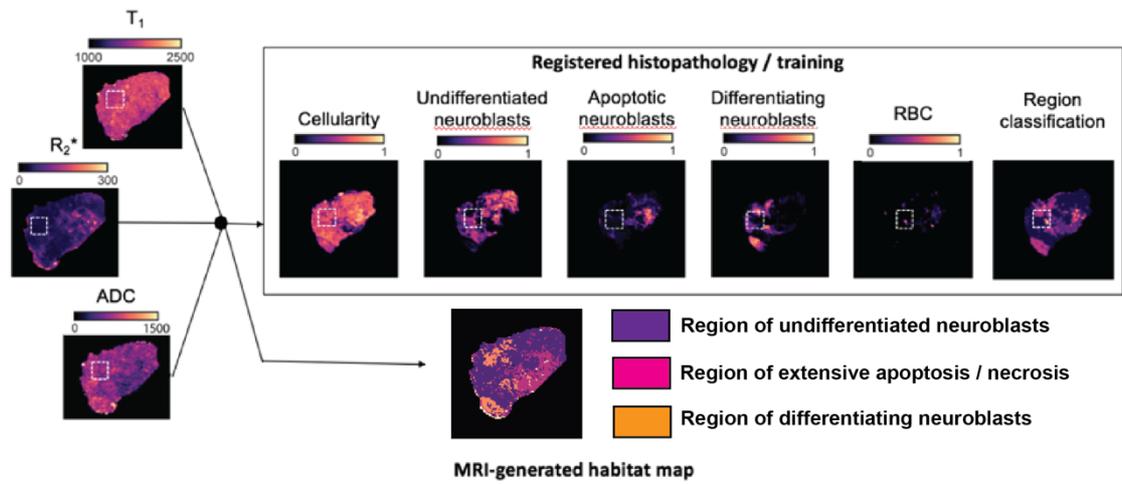
The purpose of this section is to predict and map key histological features, typically calculated from histological image analysis, such as the density of different cell populations, directly from the MRI data. The histology-derived features maps from the cell classifier were used as labels for supervised regression machine learning algorithms. Each MRI pixel is represented by the 5 input features described in Section 7.2.2.3 and 5 histological labels used as outputs (cell density, density of undifferentiated neuroblasts, density of apoptotic cells, density of differentiating cells and density of RBC).

The models were trained to both predict each label separately (single output) and all labels simultaneously (multiple output). The regression models incorporated in this approach included linear regression, lasso, elastic net, k-nearest neighbours, decision trees (CART), support vector regression (SVR), XGBoost (330) and light GBM. Ensemble methods were also used, such as Adaboost, gradient boosting machines, random forests and extra trees regressors. From the above algorithms, only the inherently multiple output algorithms were used to predict multiple labels.

The training dataset was the same as in the classification Section 7.2.2.2 and was randomly split into training and validation (80% and 20% respectively). The models were initially assessed using 10-fold cross-validation and the negative mean squared error metric.

All code was written in Python (v. 3.6.10) and the models were trained with the scikit-learn library (v. 0.23.2).

An overview of the method is presented in Figure 7.1.



**Figure 7.1 An overview of the methodology for habitat imaging.** The supervised machine learning algorithm was trained using the values from registered histology-derived maps as labels to characterise distinct tumour sub-regions (classification) or predict the histology values (regression) from multi-parametric MRI ( $T_1$ ,  $R_2^*$ , ADC) data.

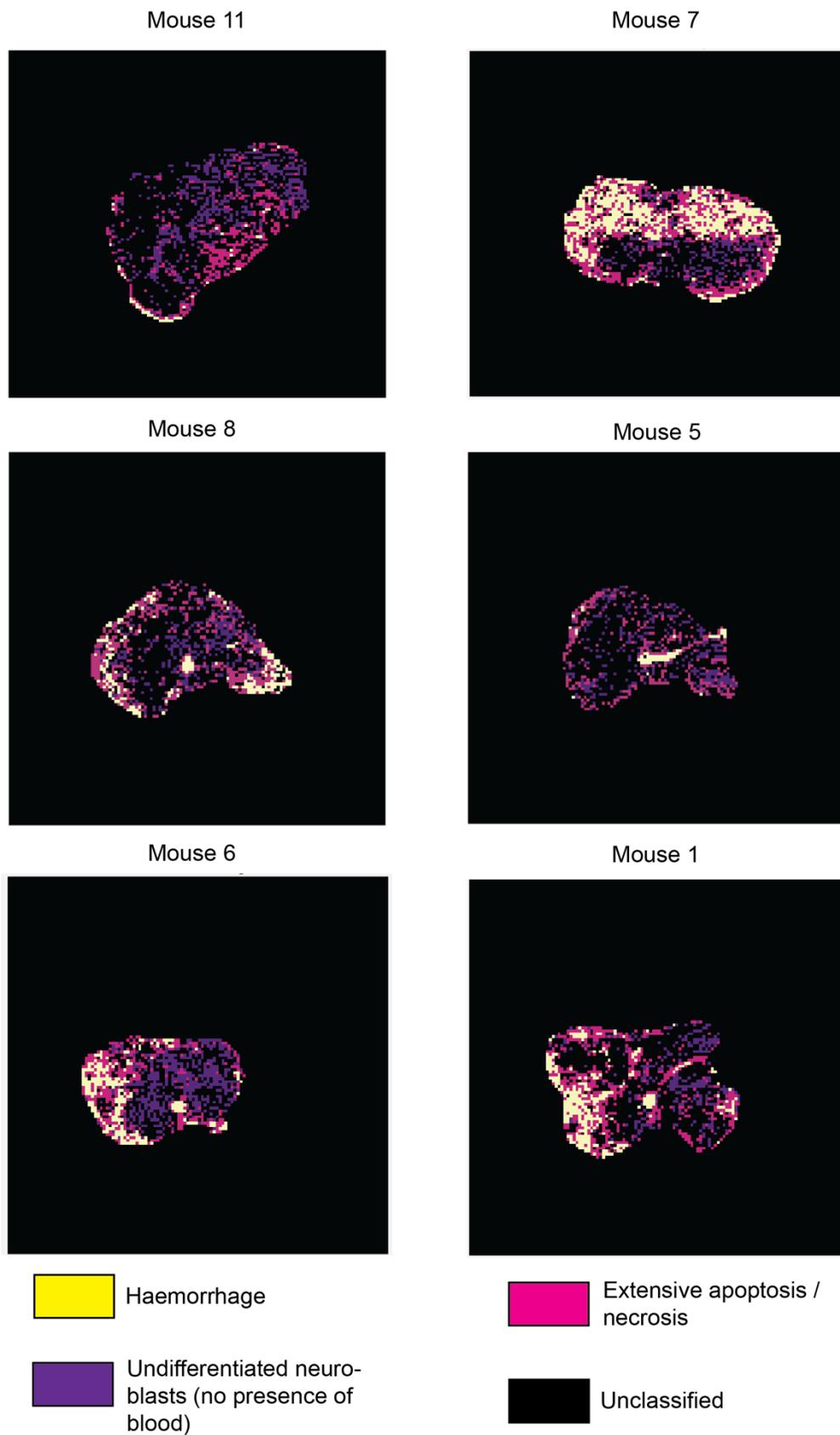
### 7.3 Results

#### 7.3.1 Empirical thresholds can confidently characterise a large proportion of the tumour

On average, 57.2% of the tumour can be confidently characterised using the empirical thresholds (Table 7.2). Representative images of classified tumours using the empirical thresholds are presented in Figure 7.2.

Mouse	Classification percentage (%)
1	49.2
2	71.1
3	55.3
4	68.3
5	39.4
6	64.3
7	70
8	48.3
9	76
10	49.6
11	37.9

**Table 7.2** Percentages of the tumour classified with high certainty after the application of empirical thresholds in all mice.



**Figure 7.2** Representative examples of classified tumours using empirical thresholds for  $R_2^*$ ,  $T_1$  and ADC. Note the unclassified intra-tumour areas in black.

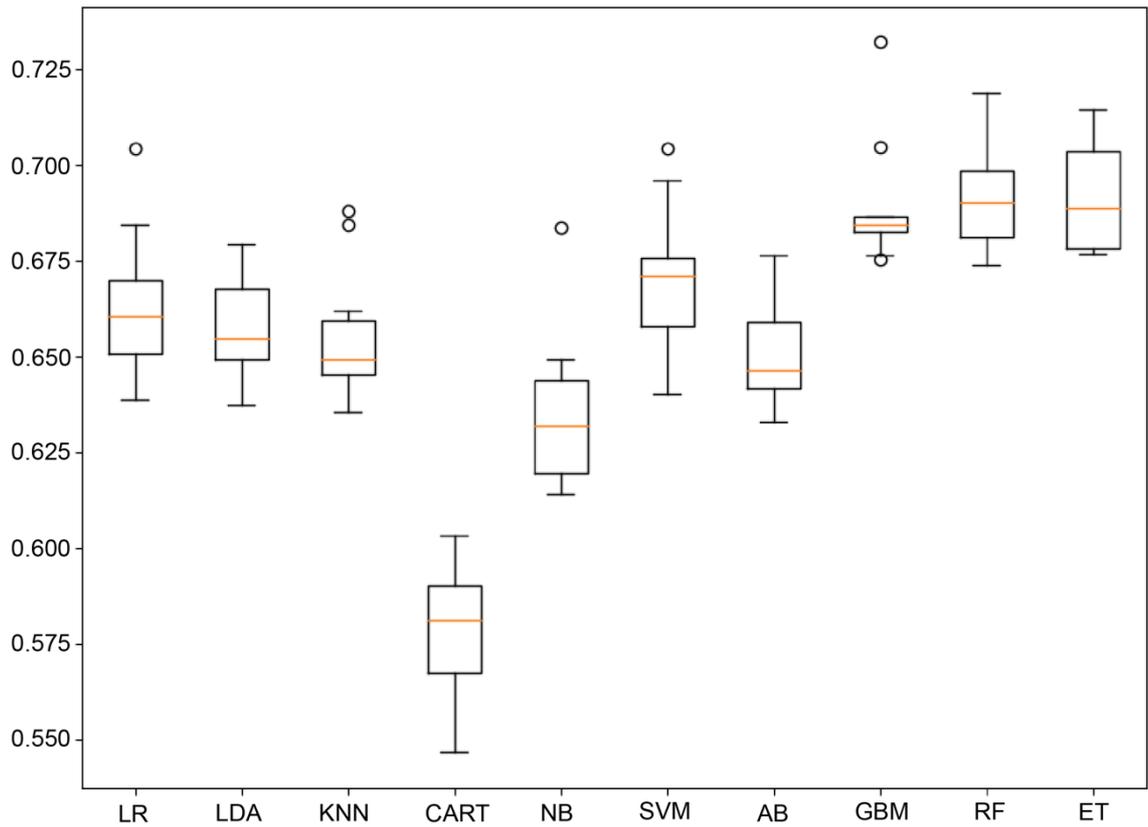
### *7.3.2 Classification of neuroblastoma tumour habitats from multi-parametric MRI*

The random forest classifier consistently achieved the highest classification accuracy after the 10-fold cross-validation during the experimentation process, followed by the extra trees classifier. In addition, random forests are inherently resistant to overfitting and were therefore selected as the final classification model. An overview of the accuracies achieved after the 10-fold cross validation from all the methods is presented in Figure 7.3.

Subsequently, the model was tuned and the final selected parameters were 100 number of estimators (trees) and no maximum depth. The tuned model achieved 69.2% accuracy after 10-fold cross-validation and 69.4% accuracy in the validation set.

Finally, the similarity of the entire MRI-generated habitat maps was compared pixel-to-pixel to the region classification histology maps using the accuracy metric (equal to Jaccard similarity index, since the task is multi-label). The average similarity score achieved is 53.6%. Detailed scores for all tumours are given in Table 7.3.

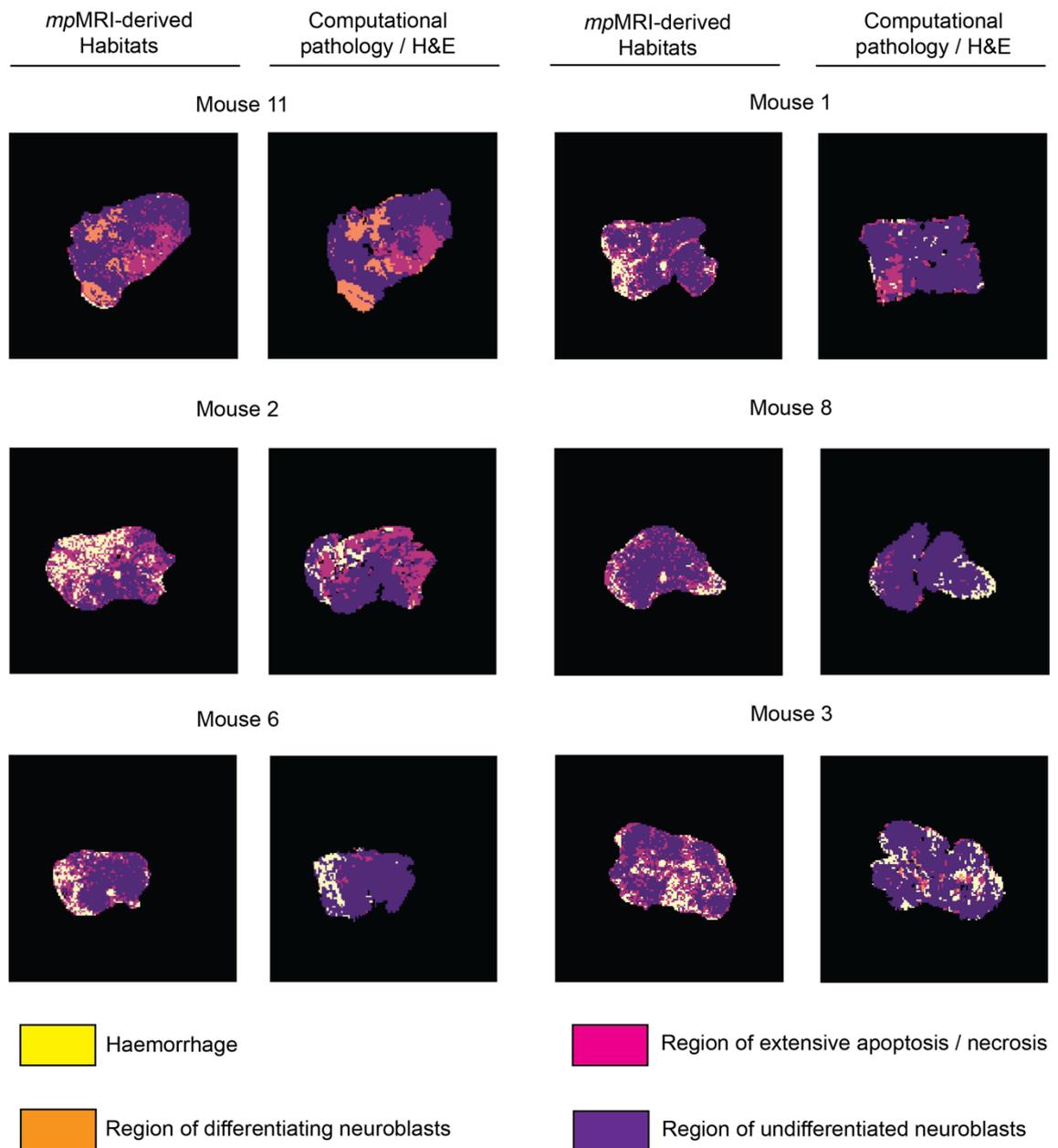
Representative classified tumours are presented in Figure 7.4.



**Figure 7.3** Overview of the 10-fold cross-validation classification accuracies between the MRI-derived and histology-derived data achieved by the linear regression (LR), linear discriminant analysis (LDA), k-nearest neighbour (KNN), decision trees (CART), Gaussian naïve Bayes (NB), support vector machines (SVM), Adaboost (AB), gradient boosting machines (GBM), random forests (RF) and extra trees (ET) algorithms.

Mouse	MRI-Histology similarity (%)
1	59
2	44.3
3	46.9
4	37.3
5	66.3
6	62.4
7	41.2
8	61.6
9	46.6
10	54.1
11	70.1

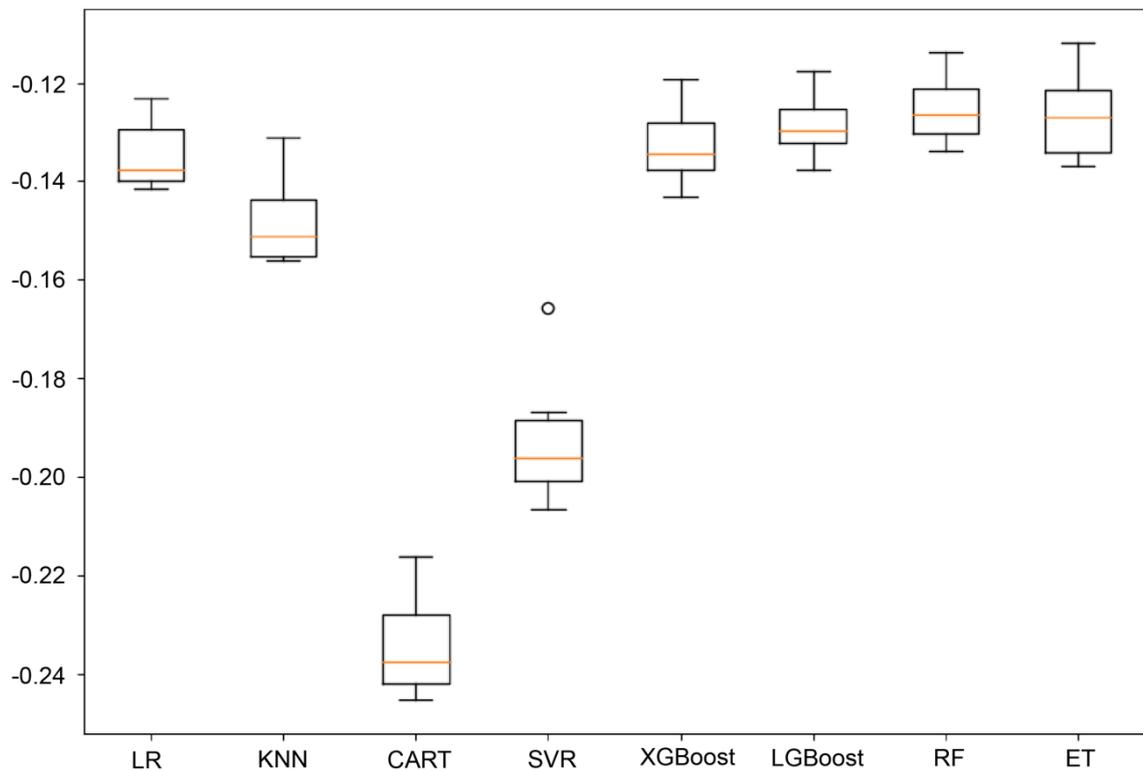
**Table 7.3** Summary of pixel-to-pixel similarity between the MRI-derived habitats and the region classification maps from histology. The accuracy metric was used to measure similarity (note that accuracy equals to the Jaccard similarity index for a multi-label task).



**Figure 7.4** Representative examples of multi-parametric MRI (*mp*MRI)-derived habitat maps and the region classification maps from histology (SuperHistopath). Note that many areas of haemorrhage are classified as tissue damage after the processing into MRI resolution of the SuperHistopath results, since areas of tissue damage very often include extensive haemorrhage.

### 7.3.3 Multi-parametric MRI can map the density of different cell populations

The random forest classifier consistently achieved the lowest mean squared error after the 10-fold cross-validation during the experimentation process and was selected as the final regression model. Extra trees, XGBoost and light gradient boosting machines also performed well. An overview of the mean square error achieved after the 10-fold cross validation from all the methods is presented in Figure 7.5.



**Figure 7.5** Overview of the mean square error between the MRI-derived and histology-derived data achieved after the 10-fold cross-validation by the linear regression (LR), k-nearest neighbour (KNN), decision trees (CART), support vector regression (SVR), XGBoost, light gradient boosting machines (LGBost), random forests (RF) and extra trees (ET) algorithms.

Subsequently, the random forests model was tuned and the final selected parameters were 100 number of estimators (trees) and no maximum depth. The multi-output tuned model achieved 0.019 mean-squared error after 10-fold cross-validation and 0.019 mean-squared error in the validation set.

Finally, the similarity of the entire MRI-generated habitat maps was compared to the region classification histology maps using the mean squared error (MSE) and structural similarity index (SSIM) metrics. The multi-output model (trained with all output labels) achieved an average MSE of 0.0037, 0.0089, 0.0016, 0.0014, and 0.0008 for cellular density, undifferentiated neuroblasts, apoptotic neuroblasts, differentiating neuroblasts and RBC density respectively. The SSIM indexes were 0.90, 0.88, 0.90, 0.88 and 0.92 for cellular density, undifferentiated neuroblasts', apoptotic neuroblasts', differentiating neuroblasts' and RBC density respectively.

The single-output model performed similarly for cellular density and undifferentiated neuroblasts density (MSE: 0.0037, 0.0089 and SSIM: 0.90, 0.87 respectively), but performed slightly worse for apoptotic and differentiating cell densities (MSE: 0.0017, 0.002 and SSIM: 0.89, 0.87 respectively), indicating that the output values may not be entirely independent. Therefore, the multi-output model was selected. Detailed scores for all tumours using the multi-output and single-output models are given in Table 7.4 and Table 7.5 respectively.

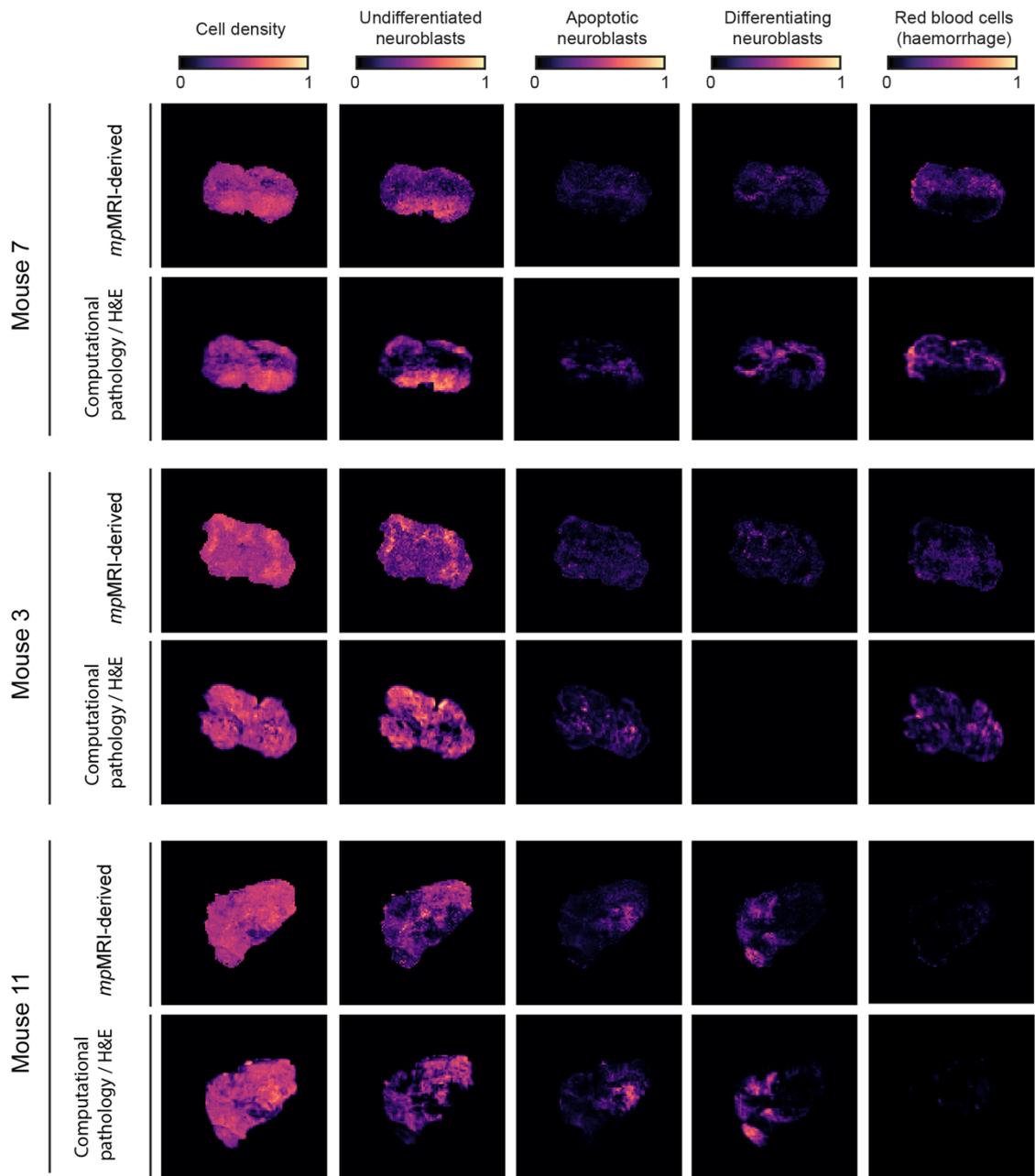
Representative MRI-generated cell density maps using the multi-output RF model are presented in Figure 7.6.

Multi-output Random Forest model					
Mouse	Cell density	Undifferentiated neuroblasts density	Apoptotic cells density	Differentiating neuroblasts density	RBC density
<b>1</b>	MSE: 0.0062 SSIM: 0.86	MSE: 0.013 SSIM: 0.83	MSE: 0.0033 SSIM: 0.86	MSE: 0.001 SSIM: 0.84	MSE: 0.0006 SSIM: 0.91
<b>2</b>	MSE: 0.0038 SSIM: 0.90	MSE: 0.0068 SSIM: 0.89	MSE: 0.0032 SSIM: 0.91	MSE: 0.0004 SSIM: 0.87	MSE: 0.0005 SSIM: 0.94
<b>3</b>	MSE: 0.0039 SSIM: 0.86	MSE: 0.0079 SSIM: 0.83	MSE: 0.0016 SSIM: 0.88	MSE: 0.0016 SSIM: 0.81	MSE: 0.0019 SSIM: 0.86
<b>4</b>	MSE: 0.0029 SSIM: 0.92	MSE: 0.0059 SSIM: 0.90	MSE: 0.0014 SSIM: 0.90	MSE: 0.0004 SSIM: 0.90	MSE: 0.0008 SSIM: 0.90
<b>5</b>	MSE: 0.0044 SSIM: 0.90	MSE: 0.0064 SSIM: 0.89	MSE: 0.0011 SSIM: 0.95	MSE: 0.0077 SSIM: 0.89	MSE: 0.0001 SSIM: 0.96
<b>6</b>	MSE: 0.0022 SSIM: 0.95	MSE: 0.0033 SSIM: 0.94	MSE: 0.0004 SSIM: 0.92	MSE: 0.0002 SSIM: 0.94	MSE: 0.0003 SSIM: 0.97
<b>7</b>	MSE: 0.0047 SSIM: 0.93	MSE: 0.0033 SSIM: 0.90	MSE: 0.0006 SSIM: 0.89	MSE: 0.0014 SSIM: 0.93	MSE: 0.0016 SSIM: 0.91
<b>8</b>	MSE: 0.0042 SSIM: 0.90	MSE: 0.02 SSIM: 0.86	MSE: 0.0013 SSIM: 0.87	MSE: 0.0008 SSIM: 0.87	MSE: 0.0008 SSIM: 0.91
<b>9</b>	MSE: 0.0038 SSIM: 0.88	MSE: 0.013 SSIM: 0.85	MSE: 0.0023 SSIM: 0.89	MSE: 0.0005 SSIM: 0.86	MSE: 0.0013 SSIM: 0.89
<b>10</b>	MSE: 0.0038 SSIM: 0.89	MSE: 0.015 SSIM: 0.85	MSE: 0.0013 SSIM: 0.95	MSE: 0.0008 SSIM: 0.85	MSE: 0.0008 SSIM: 0.91
<b>11</b>	MSE: 0.0024 SSIM: 0.93	MSE: 0.0026 SSIM: 0.90	MSE: 0.0006 SSIM: 0.90	MSE: 0.0009 SSIM: 0.94	MSE: $5 \times 10^{-5}$ SSIM: 0.98
<b>Average</b>	MSE: 0.0037 SSIM: 0.90	MSE: 0.0089 SSIM: 0.88	MSE: 0.0016 SSIM: 0.90	MSE: 0.0014 SSIM: 0.88	MSE: 0.0008 SSIM: 0.92

**Table 7.4** Summary of the mean square error (MSE) and structural similarity index (SSIM) between the MRI-derived maps and the histology-derived maps for all Th-MYCN tumours using the multi-output random forests algorithm.

	<b>Single-output Random Forest model</b>			
<b>Mouse</b>	<b>Cell density</b>	<b>Undifferentiated neuroblasts density</b>	<b>Apoptotic cells density</b>	<b>Differentiating neuroblasts density</b>
<b>1</b>	MSE: 0.0064 SSIM: 0.86	MSE: 0.013 SSIM: 0.83	MSE: 0.0033 SSIM: 0.86	MSE: 0.0019 SSIM: 0.83
<b>2</b>	MSE: 0.0039 SSIM: 0.90	MSE: 0.0069 SSIM: 0.89	MSE: 0.003 SSIM: 0.91	MSE: 0.0016 SSIM: 0.85
<b>3</b>	MSE: 0.0041 SSIM: 0.86	MSE: 0.0083 SSIM: 0.83	MSE: 0.0017 SSIM: 0.87	MSE: 0.0024 SSIM: 0.80
<b>4</b>	MSE: 0.003 SSIM: 0.92	MSE: 0.006 SSIM: 0.90	MSE: 0.0017 SSIM: 0.89	MSE: 0.0008 SSIM: 0.89
<b>5</b>	MSE: 0.0045 SSIM: 0.90	MSE: 0.0064 SSIM: 0.89	MSE: 0.0012 SSIM: 0.90	MSE: 0.0079 SSIM: 0.89
<b>6</b>	MSE: 0.002 SSIM: 0.95	MSE: 0.0033 SSIM: 0.94	MSE: 0.0005 SSIM: 0.94	MSE: 0.0007 SSIM: 0.93
<b>7</b>	MSE: 0.0021 SSIM: 0.93	MSE: 0.0037 SSIM: 0.90	MSE: 0.0008 SSIM: 0.90	MSE: 0.0019 SSIM: 0.91
<b>8</b>	MSE: 0.0048 SSIM: 0.90	MSE: 0.02 SSIM: 0.86	MSE: 0.0014 SSIM: 0.88	MSE: 0.0012 SSIM: 0.86
<b>9</b>	MSE: 0.0044 SSIM: 0.87	MSE: 0.013 SSIM: 0.85	MSE: 0.0025 SSIM: 0.86	MSE: 0.0013 SSIM: 0.86
<b>10</b>	MSE: 0.0039 SSIM: 0.89	MSE: 0.0151 SSIM: 0.84	MSE: 0.0014 SSIM: 0.88	MSE: 0.0012 SSIM: 0.85
<b>11</b>	MSE: 0.0024 SSIM: 0.93	MSE: 0.0028 SSIM: 0.90	MSE: 0.0006 SSIM: 0.94	MSE: 0.0009 SSIM: 0.94
<b>Average</b>	MSE: 0.0037 SSIM: 0.90	MSE: 0.0089 SSIM: 0.87	MSE: 0.0017 SSIM: 0.89	MSE: 0.002 SSIM: 0.87

**Table 7.5** Summary of the mean square error (MSE) and structural similarity index (SSIM) between the MRI-derived maps and the histology-derived maps for all Th-MYCN tumours using the single-output random forests algorithm.



**Figure 7.6** Representative examples of multi-parametric MRI (*mpMRI*)-derived and histology derived maps of different cell population densities.

## 7.4 Discussion

In this chapter, I incorporated multi-parametric MRI and machine learning coupled with the MRI-histology pipeline to demonstrate the potential of multi-parametric MRI to non-invasively characterise distinct tumour sub-regions (habitats) and quantitatively map the density of the different cell populations which constitute both tissue and cellular criteria currently used for risk-adapted treatment for children with neuroblastoma.

The robust validation of the MRI biomarkers and deeper understanding of the pathological determinants of  $R_2^*$ , native  $T_1$  and ADC in neuroblastoma enabled the confident non-invasive characterisation of a significant proportion of the tumour through the use of empirical thresholds. This approach enabled by the incorporation of  $T_1$ -mapping and IS-MRI could further assist radiologists and oncologists in decision making for guiding biopsies in areas of aggressive disease, avoid haemorrhagic areas and adequately assess areas of tissue damage (and potentially improve current response criteria). However, the translation of this approach in the paediatric clinical setting would be limited by the need for standardization of the protocols and equipment, a known limitation to the effective translation of quantitative MRI biomarkers.

Supervised classification approaches are usually not incorporated for this kind of tasks because of MRI-histology mismatches, leading to many labelling mistakes, and making the building of a reliable training set challenging. Here, only four samples where the alignment and registration seemed best upon visual inspection were used to reduce the impact of wrong labels. Pixel-to-pixel comparison of MRI-derived habitats and the histology images was used to estimate the performance of the approach and achieve the moderate, but encouraging (taking into account all the known mismatch issues), average of 53.6%. In future studies, overall tumour multi-resolution approaches based on human perception, such as the method used by *Jardim-Perassi and colleagues* (65) and developed by *Costanza et al* (331) and *Kuhnert et al* (332) could be incorporated to meaningfully compare the similarity between MRI and histology images.

In the Th-*MYCN* model, the inconsistent presence of blood in all histological tumour compartments, which affects the MRI measurements and the regional classification, represents a major obstacle to the characterisation of the tumour habitats (especially the presence of haemorrhage in regions of tissue damage leading to the classification as either category being correct). A potential solution would be the combination of the classification approach with the empirical thresholds, enabling the machine learning algorithm to decide only in the areas of higher uncertainty. (Yet the clinical translation will suffer from the same limitations described above)

Tumours are inherently heterogeneous and providing a unique label assigned to a histological area as large as an MRI pixel ( $\sim 230\mu\text{m} \times 230\mu\text{m}$ ) is actually not optimal especially in regions with intermixing cells and haemorrhage, as mentioned above, or regions of diffused apoptosis, for example. This is why a regression approach with multiple labels to map the density of different cell populations (including red blood cells) can potentially provide a more accurate and informative approach for the non-invasive mapping of neuroblastoma pathology. The MRI-derived histology maps showed good correspondence with the actual histology-derived maps in all tumour samples upon visual inspection, while achieving an average of 0.0047 MSE and  $\sim 0.90$  SSIM scores for all categories in the MRI-histology comparison. Practically, regression also allows for more flexibility in building a machine learning training set. Neighbouring areas normally do not differ dramatically in terms of cell density, whereas a wrong label during training in a classification task is a more serious issue. It also allows for the easier training of under-presented classes, such as the differentiating regions in our dataset, as the algorithm essentially learns through negative examples, meaning that it learns which MRI features lead to small density values of differentiating neuroblasts.

In this small dataset, the random forests algorithm was selected as the final model as it achieved the best score during the 10-fold cross validation and the validation dataset both in classification and regression methods. Random forests are a well-

established and robust approach, generally resistant to overfitting. To further improve the performance in a larger and more representative dataset, combination, through stacked model approaches, with other methods that performed well, such as the extra trees, GBM and SVM in classification and extra trees, XGBoost and light GBM in regression, should be considered. In the future, the method will also be applied to the treated samples of our dataset (chapters 5 and 6), as well as to a newly acquired larger dataset of various neuroblastoma mouse models. To achieve better results, especially in classification, the registration process will also be optimized by experimenting with newer state-of-the-art algorithms, such as the Bayesian CPD (333).

## **7.5 Conclusion**

This proof-of-concept study shows the potential of multi-parametric MRI to fully characterise neuroblastoma histology and provide quantitative data typically acquired from histology image analysis, such as maps of cellular density, densities of the different cell populations and classification of distinct tumour sub-regions (habitats). Confirmation of this approach in a larger and more diverse dataset would demonstrate the full potential of multi-parametric MRI for “non-invasive” or “virtual” biopsies to capture the entire tumour volume and enhance non-invasive diagnostic and treatment monitoring for children with neuroblastoma and pave new ways in studying and monitoring tumour evolution alongside liquid biopsies.

## Chapter 8 : Conclusions

### 8.1 Summary of results

#### *8.1.1 Novel computational approach for enhanced quantitative histopathology*

##### 8.1.1.1 Superpixel-based Conditional Random Fields (SuperCRF): Incorporating global and local context for enhanced deep learning in melanoma histopathology

I have implemented a framework which fuses traditional machine learning with deep learning to model the way pathologists incorporate large-scale tissue architecture and context across spatial scales, to improve single-cell classification in large whole-section slide images. Using this approach, a marked 11.85% overall improvement in the accuracy of the state-of-the-art deep learning SC-CNN cell classifier was demonstrated. SuperCRF can be implemented in combination with any single-cell classifier and represent valuable tools to study the cancer-stroma-immune interface, which I used to identify predictors of survival in melanoma patients from conventional H&E stained histopathology. I particularly demonstrated that in accordance with the immune-excluded phenotype that, tumours rich in stromal cells had a marked poorer prognosis in patients with melanoma. With p-value lower by two orders of magnitude, this method provided stronger predictive power than by using a deep-learning cell-morphology-only method for cell classification.

##### 8.1.1.2 SuperHistopath: A deep learning pipeline for mapping tumour heterogeneity on low-resolution whole-slide digital histopathology images

I also developed and validated the SuperHistopath framework, which combines the application of the SLIC superpixels algorithm directly on low magnification H&E-stained WSIs (5x) with a CNN for the classification of superpixels in distinct tumour regions. SuperHistopath was applied successfully to characterise three cancer types with disparate histology (melanoma, breast and neuroblastoma) without any changes (just retraining), showing that the approach could be virtually extended to any type of cancer. SuperHistopath's results of enhanced speed for both training and application (~5mins for classifying a WSI and ~30 mins for network training) and the efficient and simple collection of ground-truth datasets

make SuperHistopath particularly attractive for research in rich datasets and would facilitate its adoption in the clinic to accelerate pathologist workflow in the quantification of predictive/prognosis markers derived from global features of interest.

### *8.1.2 A robust MRI-histopathology cross-validation pipeline*

The Th-MYCN model of childhood neuroblastoma represents a robust platform not only for the evaluation of novel therapeutics, but also for the evaluation and validation of imaging biomarkers of response. I demonstrated that the use of anatomical landmarks visible on T<sub>2</sub>-weighted images for guiding the excision and sectioning of tumours provide sufficient spatial correspondence for the comparison of MRI with histopathology in the GEM models of neuroblastoma. The addition of both in-plane registration using the CPD algorithm and resolution matching of the histological maps to the MRI maps, further enhance our ability to robustly compare the regional distribution of MRI- and histology-derived parameters. The registration achieved with our pipeline will not only enhance the information provided by conventional statistical comparison of representative tumour values, (median, mean tumour) but more importantly, allow the use of more advanced spatial analysis to reflect the tumour heterogeneity and validate habitat imaging.

### *8.1.3 Imaging the hemodynamic vasculature of neuroblastoma with susceptibility MRI for the non-invasive prediction of response to anti-angiogenic treatment*

The application of the MRI-histopathology cross correlation platform demonstrated that IS- and SC-MRI are robust non-invasive imaging techniques to characterize, quantify, and map the unique vascular phenotype of neuroblastoma and its therapeutic modulation with the panVEGFR small molecule inhibitor cediranib. Importantly, the data presented in Chapter 4 strongly suggest that both baseline tumour *fBV* and  $R_2^*$  are predictive of the longer-term tumour response to VEGFR-targeted therapies. With the central role of angiogenesis in determining and predicting the clinical behaviour of neuroblastoma, baseline *fBV* and  $R_2^*$ , have the potential to provide diagnostic and prognosis information at the time of diagnosis and provide currently

unavailable predictive and pharmacodynamic biomarkers for antiangiogenic therapy against neuroblastoma.

#### *8.1.4 Non-invasive MRI native $T_1$ -mapping detects response to MYCN-targeted therapies in the Th-MYCN model of neuroblastoma*

Chapter 5 demonstrated that native  $T_1$ -mapping can precisely and quantitatively map the rich histopathology of neuroblastoma tumours and that a reduction in native  $T_1$  is a robust biomarker of response to treatment with *MYCN*-targeted therapeutics in the clinically-relevant Th-*MYCN* model of neuroblastoma, including Aurora A kinase inhibitor, alisertib and mTOR inhibitor vistusertib. By providing strong evidence for the sensitivity of native  $T_1$  to dense areas of undifferentiated neuroblasts, the data afforded by the MRI-Histology pipeline suggests further application for diagnosis, risk stratification and surgical planning, and that its potential as a biomarker of successful response to therapy could be extended to larger subsets of aggressive paediatric and adult tumours.

#### *8.1.5 ADC is sensitive to the microstructure of poorly- or un-differentiated neuroblastoma but is not an early biomarker of response to apoptosis-inducing MYCN-targeted therapies.*

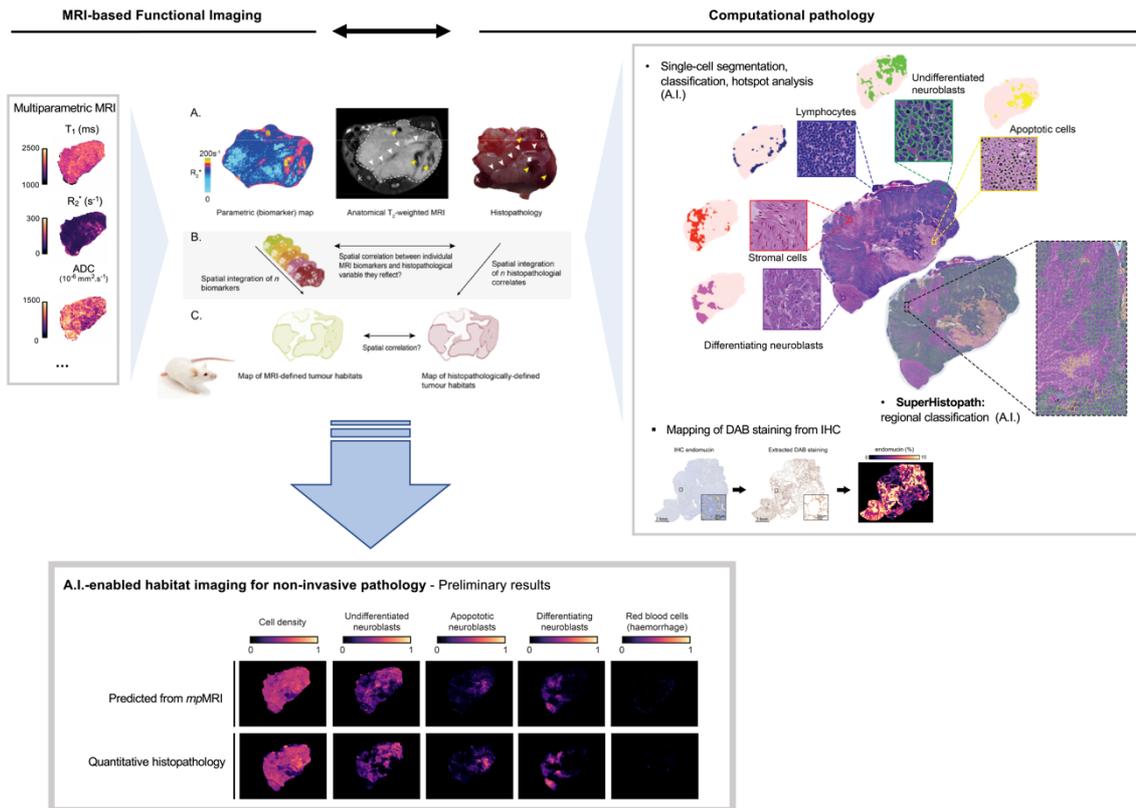
The insights into neuroblastoma histopathology developed in the previous chapters guided the application of GMMs to derive ADC compartments to further demonstrate the sensitivity of DWI to tumour microstructure and the rich histopathology of neuroblastoma focusing on the Schwannian stroma-poor, poorly- or un-differentiated phenotype, which is characteristic of high-risk disease and faithfully recapitulated in the tumours of the Th-*MYCN* mouse model. However, this MRI-histopathology cross-validation study demonstrates that median/mean tumour ADC cannot be used as a reliable early sensitive pharmacodynamic biomarker of response to targeted treatment in the Th-*MYCN* model. This study particularly demonstrated that ADC is unequivocally insensitive to the morphological cellular changes occurring during the apoptotic process and that the consequences of cell death on tumour architecture are associated with conflicting effects on ADC. More specifically, extensive regions of tissue damage

or convergent cell death, haemorrhagic or not, are associated with reduced ADC values, while tissue maturation and cellular differentiation are associated with higher ADC values.

*8.1.6 Habitat imaging: Towards non-invasive pathology for children with neuroblastoma.*

Finally, in Chapter 7, I used supervised machine learning classification and regression approaches to combine the parametric maps of native  $R_2^*$ , native  $T_1$  and ADC and demonstrated the potential of multi-parametric MRI to fully characterise neuroblastoma histology. This proof-of-concept study shows that the approach can non-invasively provide quantitative data typically acquired from histology image analysis, such as maps of cellular density, densities of the different cell populations and classification of distinct tumour sub-regions. Confirmation of this approach in a larger and more diverse dataset would demonstrate the full potential of multiparametric MRI for “non-invasive” or “virtual” biopsies to capture the entire tumour volume and enhance non-invasive diagnostic and treatment monitoring for children with neuroblastoma and pave new ways in studying and monitoring tumour evolution alongside liquid biopsies.

Figure 8.1 summarises the result and progress in this project in relation to the initial hypothesis presented in Figure 1.11.



**Figure 8.1** Towards non-invasive pathology: summary of the results

## 8.2 Future Directions

### 8.2.1 A mouse hospital and co-clinical approach

This thesis provides a strong rationale for the incorporation of  $T_1$ -mapping and IS-MRI alongside diffusion weighted MRI (which is already increasingly used clinically) both at the time of diagnosis and in early phase clinical trials to guide clinical decision making and the delivery of precision medicine to children with neuroblastoma. Widely available on conventional clinical scanners,  $T_1$ -mapping and IS-MRI are rapid, contrast-free scans that have been successfully incorporated across international multi-center and multi-scanner platforms to assess liver function as a part of a multiparametric approach LiverMultiScan (Perspectum Diagnostics, Oxford, UK) to measure fat content and correlates of iron, fibrosis and inflammation in chronic liver disease. Including  $T_1$ -mapping and IS-MRI in an ethically-approved clinical study within the standard-of-care frontline

chemotherapy would allow for the rapid evaluation and validation of the potential of the habitat imaging approach for the neuroblastoma clinic.

#### 8.2.1.1 **MAGNIFY**: comprehensive **MA**ppin**G** of high-risk **Neuroblastoma** heterogeneity to predict and target re**F**ractor**Y** disease

Based on the work of this thesis, Dr Jamin has initiated MAGNIFY: a clinical study within the context of standard-of-care for localised neuroblastoma and a team science effort between Great Ormond Street Hospital, GOS Institute of Child Health, the ICR and the Royal Marsden Hospital. In MAGNIFY, functional imaging ( $T_1$ ,  $R_2^*$ , ADC and contrast-enhanced  $T_1$ -weighted MRI) will be used alongside molecular imaging (MIBG scans) to guide multi-regional tumour biopsies at the time of diagnosis and post-induction treatment, together with interval serial blood (ctDNA), to map the regional heterogeneity of genomic, transcriptomic and microenvironmental (including immuno-proteomic) signatures at diagnosis, and their evolution following induction chemotherapy to discern subclasses of high risk disease, especially signatures associated with resistance to therapy. These signatures will enhance our understanding of the biology of high-risk disease and help us develop new clinical studies of biomarker-driven risk-adapted therapies. Incorporating this information into routine clinical practice is the key to a better understanding of neuroblastoma heterogeneity across length scales i.e. from the tissue level to the genomic level, essential for enhancing clinical decision-making. This study would also provide a unique opportunity to validate the habitat imaging approach clinically.

#### *8.2.2 Incorporating MRI biomarkers of the tumour microenvironment*

Once the habitat imaging approach has been validated in a larger cohort for the accurate detection of active disease, the next step would be to integrate functional and molecular MRI scans to capture the regional heterogeneity of microenvironmental factors within region of active disease including biomechanics (MR-elastography), blood supply (SC-MRI, contrast-enhanced  $T_1$ w-MRI, arterial spin labelling) hypoxia (oxygen-enhanced  $T_1$ w-MRI) and metabolism (Deuterium Metabolic Imaging) and demonstrate the application of

the approach *i*) for the monitoring and guiding of the multi-omics characterization of resistant disease in GEM models of genomic variants (MYCN, ALK, ATRX) and *ii*) in GEM models of chemo-refractory disease as well as *iii*) its potential for monitoring tumour evolution.

### *8.2.3 Refining the MRI-Histopathology pipeline*

#### 8.2.3.1 A larger and more diverse dataset

Moving forward, the methodologies developed here will be optimized and applied in a new independent dataset including GEM models of genomic variants of neuroblastoma (MYCN, ALK, TrP53, ATRX) as well as tumours treated with targeted modulators of these genomic drivers and of the microenvironment, and finally with tumours that acquire resistance to frontline chemotherapy.

#### 8.2.3.2 Improving registration

The current 2D registration CPD algorithm will be compared with new state-of-the-art methods (333-335). Also, 3D registration approaches that involve the precise cutting of the whole tumour, such as those described in Chapter 3, should be considered.

#### 8.2.3.3 Improving cell classification for the enhanced characterisation of the tumour ecosystem

For the cell classification step of the pipeline, the SuperCRF framework will be used, as the classifier based on CRImage, despite being successful in this dataset suffers from the known limitations of classical machine learning algorithms. As a first step in that direction, the state-of-the-art cell detection/classification algorithms, SC-CNN (143) and MapDe (336), have already been applied, but they both severely underperformed in the case of differentiating neuroblasts, resulting in over-detection, likely because of their bigger size and the presence of multiple nucleoli. Due to the morphology of these cells, including their very clear cell boundaries, cell segmentation approaches, such as U-Net or U-Net-like deep learning architectures (131, 188) are more likely to succeed. Moreover, the region classification part of the SuperCRF framework

will be updated with the deep learning SuperHistopath pipeline. I will further explore the spatial interactions across length scales to provide a deeper understanding of the cancer-immune-stroma interface. Ultimately, my aim is to establish a network which will provide a complete characterization of every component of the tumour microenvironment where all the parts will interact with each other like an ecological landscape.

## References

1. McGranahan N, Swanton C. Biological and therapeutic impact of intratumor heterogeneity in cancer evolution. *Cancer Cell*. 2015;27(1):15-26.
2. Gatenby RA, Grove O, Gillies RJ. Quantitative imaging in cancer evolution and ecology. *Radiology*. 2013;269(1):8-14.
3. Kim JY, Gatenby RA. Quantitative clinical imaging methods for monitoring intratumoral evolution. *Cancer Gene Networks: Springer*; 2017. p. 61-81.
4. Kleppe M, Levine RL. Tumor heterogeneity confounds and illuminates: assessing the implications. *Nature medicine*. 2014;20(4):342-4.
5. Pattabiraman DR, Weinberg RA. Tackling the cancer stem cells—what challenges do they pose? *Nature reviews Drug discovery*. 2014;13(7):497.
6. Maley CC, Koelble K, Natrajan R, Aktipis A, Yuan Y. An ecological measure of immune-cancer colocalization as a prognostic factor for breast cancer. *Breast Cancer Research*. 2015;17(1):131.
7. Ibrahim-Hashim A, Robertson-Tessi M, Enriquez-Navas PM, Damaghi M, Balagurunathan Y, Wojtkowiak JW, et al. Defining cancer subpopulations by adaptive strategies rather than molecular properties provides novel insights into intratumoral evolution. *Cancer research*. 2017;77(9):2242-54.
8. Amend SR, Pienta KJ. Ecology meets cancer biology: the cancer swamp promotes the lethal cancer phenotype. *Oncotarget*. 2015;6(12):9669.
9. Amend SR, Roy S, Brown JS, Pienta KJ. Ecological paradigms to understand the dynamics of metastasis. *Cancer letters*. 2016;380(1):237-42.
10. Pienta KJ, Robertson BA, Coffey DS, Taichman RS. The cancer diaspora: Metastasis beyond the seed and soil hypothesis. *Clinical cancer research*. 2013;19(21):5849-55.
11. Yang KR, Mooney SM, Zarif JC, Coffey DS, Taichman RS, Pienta KJ. Niche inheritance: a cooperative pathway to enhance cancer cell fitness through ecosystem engineering. *Journal of cellular biochemistry*. 2014;115(9):1478-85.
12. Lloyd MC, Rejniak KA, Brown JS, Gatenby RA, Minor ES, Bui MM. Pathology to enhance precision medicine in oncology: lessons from landscape ecology. *Advances in anatomic pathology*. 2015;22(4):267.
13. Kirilovsky A, Marliot F, El Sissy C, Haicheur N, Galon J, Pagès F. Rational bases for the use of the Immunoscore in routine clinical settings as a prognostic and predictive biomarker in cancer patients. *International immunology*. 2016;28(8):373-82.
14. Mlecnik B, Bindea G, Kirilovsky A, Angell HK, Obenauf AC, Tosolini M, et al. The tumor microenvironment and Immunoscore are critical determinants of dissemination to distant metastasis. *Science translational medicine*. 2016;8(327):327ra26-ra26.
15. Galon J, Costes A, Sanchez-Cabo F, Kirilovsky A, Mlecnik B, Lagorce-Pagès C, et al. Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. *Science*. 2006;313(5795):1960-4.
16. Sato E, Olson SH, Ahn J, Bundy B, Nishikawa H, Qian F, et al. Intraepithelial CD8+ tumor-infiltrating lymphocytes and a high CD8+/regulatory T cell ratio are associated with favorable prognosis in ovarian cancer. *Proceedings of the National Academy of Sciences*. 2005;102(51):18538-43.

17. Loi S, Sirtaine N, Piette F, Salgado R, Viale G, Van Eenoo F, et al. Prognostic and predictive value of tumor-infiltrating lymphocytes in a phase III randomized adjuvant breast cancer trial in node-positive breast cancer comparing the addition of docetaxel to doxorubicin with doxorubicin-based chemotherapy: BIG 02-98. *J Clin Oncol*. 2013;31(7):860-7.
18. Clemente CG, Mihm Jr MC, Bufalino R, Zurrida S, Collini P, Cascinelli N. Prognostic value of tumor infiltrating lymphocytes in the vertical growth phase of primary cutaneous melanoma. *Cancer: Interdisciplinary International Journal of the American Cancer Society*. 1996;77(7):1303-10.
19. Motz GT, Coukos G. Deciphering and reversing tumor immune suppression. *Immunity*. 2013;39(1):61-73.
20. Galon J, Angell HK, Bedognetti D, Marincola FM. The continuum of cancer immunosurveillance: prognostic, predictive, and mechanistic signatures. *Immunity*. 2013;39(1):11-26.
21. Gentles AJ, Newman AM, Liu CL, Bratman SV, Feng W, Kim D, et al. The prognostic landscape of genes and infiltrating immune cells across human cancers. *Nature medicine*. 2015;21(8):938-45.
22. Heindl A, Sestak I, Naidoo K, Cuzick J, Dowsett M, Yuan Y. Relevance of spatial heterogeneity of immune infiltration for predicting risk of recurrence after endocrine therapy of ER+ breast cancer. *JNCI: Journal of the National Cancer Institute*. 2018;110(2):166-75.
23. Nawaz S, Heindl A, Koelble K, Yuan Y. Beyond immune density: critical role of spatial heterogeneity in estrogen receptor-negative breast cancer. *Modern Pathology*. 2015;28(6):766.
24. Ornitz DM, Itoh N. The fibroblast growth factor signaling pathway. *Wiley Interdisciplinary Reviews: Developmental Biology*. 2015;4(3):215-66.
25. Hanahan D, Coussens LM. Accessories to the crime: functions of cells recruited to the tumor microenvironment. *Cancer cell*. 2012;21(3):309-22.
26. Rattigan YI, Patel BB, Ackerstaff E, Sukenick G, Koutcher JA, Glod JW, et al. Lactate is a mediator of metabolic cooperation between stromal carcinoma associated fibroblasts and glycolytic tumor cells in the tumor microenvironment. *Experimental cell research*. 2012;318(4):326-35.
27. Sotgia F, Martinez-Outschoorn UE, Howell A, Pestell RG, Pavlides S, Lisanti MP. Caveolin-1 and cancer metabolism in the tumor microenvironment: markers, models, and mechanisms. *Annual Review of Pathology: Mechanisms of Disease*. 2012;7:423-67.
28. Martinez-Outschoorn UE, Sotgia F, Lisanti MP. Power surge: supporting cells "fuel" cancer cell mitochondria. *Cell metabolism*. 2012;15(1):4-5.
29. Nieman KM, Kenny HA, Penicka CV, Ladanyi A, Buell-Gutbrod R, Zillhardt MR, et al. Adipocytes promote ovarian cancer metastasis and provide energy for rapid tumor growth. *Nature medicine*. 2011;17(11):1498.
30. Paulsson J, Micke P, editors. Prognostic relevance of cancer-associated fibroblasts in human cancer. *Seminars in cancer biology*; 2014: Elsevier.
31. Östman A, Augsten M. Cancer-associated fibroblasts and tumor growth—bystanders turning into key players. *Current opinion in genetics & development*. 2009;19(1):67-73.

32. Hirata E, Girotti MR, Viros A, Hooper S, Spencer-Dene B, Matsuda M, et al. Intravital imaging reveals how BRAF inhibition generates drug-tolerant microenvironments with high integrin  $\beta$ 1/FAK signaling. *Cancer cell*. 2015;27(4):574-88.
33. Hwang RF, Moore T, Arumugam T, Ramachandran V, Amos KD, Rivera A, et al. Cancer-associated stromal fibroblasts promote pancreatic tumor progression. *Cancer Res*. 2008;68(3):918-26.
34. Nawaz S, Trahearn NA, Heindl A, Banerjee S, Maley CC, Sottoriva A, et al. Analysis of tumour ecological balance reveals resource-dependent adaptive strategies of ovarian cancer. *EBioMedicine*. 2019;48:224-35.
35. Jamal-Hanjani M, Wilson GA, McGranahan N, Birkbak NJ, Watkins TB, Veeriah S, et al. Tracking the evolution of non-small-cell lung cancer. *New England Journal of Medicine*. 2017;376(22):2109-21.
36. AbdulJabbar K, Raza SEA, Rosenthal R, Jamal-Hanjani M, Veeriah S, Akarca A, et al. Geospatial immune variability illuminates differential evolution of lung adenocarcinoma. *Nature Medicine*. 2020:1-9.
37. Narunsky L, Oren R, Bochner F, Neeman M. Imaging aspects of the tumor stroma with therapeutic implications. *Pharmacology & Therapeutics*. 2014;141(2):192-208.
38. Dominiotto M, Rudin M. Could magnetic resonance provide in vivo histology? *Comprehensive Systems Biomedicine*. 2014:7.
39. O'Connor JP, Boulton JK, Jamin Y, Babur M, Finegan KG, Williams KJ, et al. Oxygen-enhanced MRI accurately identifies, quantifies, and maps tumor hypoxia in preclinical cancer models. *Cancer research*. 2016;76(4):787-95.
40. Hill DK, Kim E, Teruel JR, Jamin Y, Widerøe M, Sjøgaard CD, et al. Diffusion-weighted MRI for early detection and characterization of prostate cancer in the transgenic adenocarcinoma of the mouse prostate model. *Journal of Magnetic Resonance Imaging*. 2016;43(5):1207-17.
41. Jamin Y, Tucker ER, Poon E, Popov S, Vaughan L, Boulton JK, et al. Evaluation of clinically translatable MR imaging biomarkers of therapeutic response in the TH-MYCN transgenic mouse model of neuroblastoma. *Radiology*. 2013;266(1):130-40.
42. Li J, Jamin Y, Boulton J, Cummings C, Waterton J, Ulloa J, et al. Tumour biomechanical response to the vascular disrupting agent ZD6126 in vivo assessed by magnetic resonance elastography. *British journal of cancer*. 2014;110(7):1727.
43. Jamin Y, Glass L, Hallsworth A, George R, Koh D-M, Pearson AD, et al. Intrinsic susceptibility MRI identifies tumors with ALK F1174L mutation in genetically-engineered murine models of high-risk neuroblastoma. *PLoS one*. 2014;9(3):e92886.
44. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology*. 2015;278(2):563-77.
45. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, Van Stiphout RG, Granton P, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *European journal of cancer*. 2012;48(4):441-6.
46. Kumar V, Gu Y, Basu S, Berglund A, Eschrich SA, Schabath MB, et al. Radiomics: the process and the challenges. *Magn Reson Imaging*. 2012;30(9):1234-48.
47. Oikonomou A, Khalvati F, Tyrrell PN, Haider MA, Tarique U, Jimenez-Juan L, et al. Radiomics analysis at PET/CT contributes to prognosis of recurrence and survival in lung cancer treated with stereotactic body radiotherapy. *Scientific reports*. 2018;8(1):1-11.

48. Gillies R, Anderson A, Gatenby R, Morse D. The biology underlying molecular imaging in oncology: from genome to anatome and back again. *Clinical radiology*. 2010;65(7):517-21.
49. Stoyanova R, Takhar M, Tschudi Y, Ford JC, Solórzano G, Erho N, et al. Prostate cancer radiomics and the promise of radiogenomics. *Translational Cancer Research*. 2016;5(4):432-47.
50. Vargas HA, Huang EP, Lakhman Y, Ippolito JE, Bhosale P, Mellnick V, et al. Radiogenomics of high-grade serous ovarian cancer: multireader multi-institutional study from the Cancer Genome Atlas Ovarian Cancer Imaging Research Group. *Radiology*. 2017;285(2):482-92.
51. Sala E, Mema E, Himoto Y, Veeraraghavan H, Brenton J, Snyder A, et al. Unravelling tumour heterogeneity using next-generation imaging: radiomics, radiogenomics, and habitat imaging. *Clinical radiology*. 2017;72(1):3-10.
52. Afshar P, Mohammadi A, Plataniotis KN, Oikonomou A, Benali H. From handcrafted to deep-learning-based cancer radiomics: challenges and opportunities. *IEEE Signal Processing Magazine*. 2019;36(4):132-60.
53. Paul R, Schabath M, Balagurunathan Y, Liu Y, Li Q, Gillies R, et al. Explaining deep features using Radiologist-Defined semantic features and traditional quantitative features. *Tomography*. 2019;5(1):192.
54. Avanzo M, Wei L, Stancanella J, Vallières M, Rao A, Morin O, et al. Machine and deep learning methods for radiomics. *Medical Physics*. 2020;47(5):e185-e202.
55. Zhou M, Chaudhury B, Hall LO, Goldgof DB, Gillies RJ, Gatenby RA. Identifying spatial imaging biomarkers of glioblastoma multiforme for survival group prediction. *Journal of Magnetic Resonance Imaging*. 2017;46(1):115-23.
56. Stringfield O, Arrington JA, Johnston SK, Rognin NG, Peeri NC, Balagurunathan Y, et al. Multiparameter MRI Predictors of Long-Term Survival in Glioblastoma Multiforme. *Tomography*. 2019;5(1):135.
57. Farhidzadeh H, Chaudhury B, Scott JG, Goldgof DB, Hall LO, Gatenby RA, et al., editors. Signal intensity analysis of ecological defined habitat in soft tissue sarcomas to predict metastasis development. *Medical Imaging 2016: Computer-Aided Diagnosis; 2016: International Society for Optics and Photonics*.
58. Farhidzadeh H, Chaudhury B, Zhou M, Goldgof DB, Hall LO, Gatenby RA, et al., editors. Prediction of treatment outcome in soft tissue sarcoma based on radiologically defined habitats. *Medical Imaging 2015: Computer-Aided Diagnosis; 2015: International Society for Optics and Photonics*.
59. Carano RA, Ross AL, Ross J, Williams SP, Koeppen H, Schwall RH, et al. Quantification of tumor tissue populations by multispectral analysis. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*. 2004;51(3):542-51.
60. Chang Y-CC, Ackerstaff E, Tschudi Y, Jimenez B, Foltz W, Fisher C, et al. Delineation of tumor habitats based on dynamic contrast enhanced MRI. *Scientific Reports*. 2017;7(1):1-14.
61. Lam WW, Oakden W, Karami E, Koletar MM, Murray L, Liu SK, et al. An Automated Segmentation pipeline for intratumoural Regions in Animal Xenografts Using Machine Learning and Saturation transfer MRI. *Scientific reports*. 2020;10(1):1-14.

62. Weigelt B, Vargas HA, Selenica P, Geyer FC, Mazaheri Y, Blecua P, et al. Radiogenomics analysis of intratumor heterogeneity in a patient with high-grade serous ovarian cancer. *JCO Precision Oncology*. 2019;3:1-9.
63. Wu J, Cao G, Sun X, Lee J, Rubin DL, Napel S, et al. Intratumoral spatial heterogeneity at perfusion MR imaging predicts recurrence-free survival in locally advanced breast cancer treated with neoadjuvant chemotherapy. *Radiology*. 2018;288(1):26-35.
64. Parra NA, Lu H, Choi J, Gage K, Pow-Sang J, Gillies RJ, et al. Habitats in DCE-MRI to predict clinically significant prostate cancers. *Tomography*. 2019;5(1):68.
65. Jardim-Perassi BV, Huang S, Dominguez-Viqueira W, Poleszczuk J, Budzevich MM, Abdalah MA, et al. Multiparametric MRI and Coregistered Histology Identify Tumor Habitats in Breast Cancer Mouse Models. *Cancer research*. 2019;79(15):3952-64.
66. O'Connor JP, Rose CJ, Waterton JC, Carano RA, Parker GJ, Jackson A. Imaging intratumor heterogeneity: role in therapy response, resistance, and clinical outcome. *Clin Cancer Res*. 2015;21(2):249-57.
67. Salem A, Little RA, Latif A, Featherstone AK, Babur M, Peset I, et al. Oxygen-enhanced MRI is feasible, repeatable, and detects radiotherapy-induced change in hypoxia in Xenograft models and in patients with non-small cell lung Cancer. *Clinical Cancer Research*. 2019;25(13):3818-29.
68. Panagiotaki E, Walker-Samuel S, Siow B, Johnson SP, Rajkumar V, Pedley RB, et al. Noninvasive quantification of solid tumor microstructure using VERDICT MRI. *Cancer research*. 2014;74(7):1902-12.
69. Blackledge MD, Winfield JM, Morgan V, Thway K, Strauss D, Collins D, et al. Supervised machine-learning enables segmentation and evaluation of heterogeneous post-treatment changes in multi-parametric MRI of soft-tissue sarcoma. *Frontiers in oncology*. 2019;9:941.
70. Matthay K, Maris J, Schleiermacher G, Nakagawara A, Mackall C, Diller L. Neuroblastoma. vol. 2. *Nat Rev Dis Primers*. 2016:16078.
71. Lonergan GJ, Schwab CM, Suarez ES, Carlson CL. From the archives of the AFIP: neuroblastoma, ganglioneuroblastoma, and ganglioneuroma: radiologic-pathologic correlation. *Radiographics*. 2002;22(4):911-34.
72. Maris JM. Recent advances in neuroblastoma. *N Engl J Med*. 2010;362(23):2202-11.
73. Goo HW. Whole-body MRI of neuroblastoma. *European journal of radiology*. 2010;75(3):306-14.
74. Tas M, Reedijk A, Karim-Kos H, Kremer L, van de Ven C, Dierselhuis M, et al. Neuroblastoma between 1990 and 2014 in the Netherlands: Increased incidence and improved survival of high-risk neuroblastoma. *European Journal of Cancer*. 2020;124:47-55.
75. Jain N, Halbert J, Patel PA, Biassoni L, Anderson J, Sebire N, et al. Importance of Magnetic Resonance Imaging With Diffusion-weighted Imaging in Guiding Biopsy of Nodular Ganglioneuroblastoma: A Case Report. *Journal of Pediatric Hematology/Oncology*. 2019.
76. Joshi VV. Peripheral neuroblastic tumors: pathologic classification based on recommendations of international neuroblastoma pathology committee (Modification of shimada classification). *Pediatric and Developmental Pathology*. 2000;3(2):184-99.

77. Shimada H, Ambros IM, Dehner LP, Hata J, Joshi VV, Roald B, et al. The international neuroblastoma pathology classification (the Shimada system). *Cancer: Interdisciplinary International Journal of the American Cancer Society*. 1999;86(2):364-72.
78. Joshi VV, Cantor AB, Altshuler G, Larkin EW, Neill JS, Shuster JJ, et al. Age-linked prognostic categorization based on a new histologic grading system of neuroblastomas. A clinicopathologic study of 211 cases from the pediatric oncology group. *Cancer*. 1992;69(8):2197-211.
79. Askin FB, Perlman EJ. Neuroblastoma and peripheral neuroectodermal tumors. *American journal of clinical pathology*. 1998;109(4 Suppl 1):S23-30.
80. Shimada H, Ambros IM, Dehner LP, Hata J, Joshi VV, Roald B. Terminology and morphologic criteria of neuroblastic tumors: recommendations by the International Neuroblastoma Pathology Committee. *Cancer*. 1999;86(2):349-63.
81. Ackermann S, Cartolano M, Hero B, Welte A, Kahlert Y, Roderwieser A, et al. A mechanistic classification of clinical phenotypes in neuroblastoma. *Science*. 2018;362(6419):1165-70.
82. Chen Y, Takita J, Choi YL, Kato M, Ohira M, Sanada M, et al. Oncogenic mutations of ALK kinase in neuroblastoma. *Nature*. 2008;455(7215):971.
83. George RE, Sanda T, Hanna M, Fröhling S, Luther II W, Zhang J, et al. Activating mutations in ALK provide a therapeutic target in neuroblastoma. *Nature*. 2008;455(7215):975.
84. Mossé YP, Laudenslager M, Longo L, Cole KA, Wood A, Attiyeh EF, et al. Identification of ALK as a major familial neuroblastoma predisposition gene. *Nature*. 2008;455(7215):930.
85. Peifer M, Hertwig F, Roels F, Dreidax D, Gartlgruber M, Menon R, et al. Telomerase activation by genomic rearrangements in high-risk neuroblastoma. *Nature*. 2015;526(7575):700.
86. Valentijn LJ, Koster J, Zwijnenburg DA, Hasselt NE, van Sluis P, Volckmann R, et al. TERT rearrangements are frequent in neuroblastoma and identify aggressive tumors. *Nature genetics*. 2015;47(12):1411.
87. Brodeur GM. Neuroblastoma: biological insights into a clinical enigma. *Nature Reviews Cancer*. 2003;3(3):203-16.
88. Moreno L, Caron H, Georger B, Eggert A, Schleiermacher G, Brock P, et al. Accelerating drug development for neuroblastoma-new drug development strategy: an innovative therapies for children with cancer, European network for cancer research in children and adolescents and international society of paediatric oncology Europe neuroblastoma project. *Expert opinion on drug discovery*. 2017;12(8):801-11.
89. Clohessy JG, Pandolfi PP. Mouse hospital and co-clinical trial project—from bench to bedside. *Nature reviews Clinical oncology*. 2015;12(8):491.
90. Weiss WA, Aldape K, Mohapatra G, Feuerstein BG, Bishop JM. Targeted expression of MYCN causes neuroblastoma in transgenic mice. *EMBO J*. 1997;16(11):2985-95.
91. Chesler L, Weiss WA, editors. Genetically engineered murine models—contribution to our understanding of the genetics, molecular pathology and therapeutic targeting of neuroblastoma. *Seminars in cancer biology*; 2011: Elsevier.
92. De Brouwer S, De Preter K, Kumps C, Zabrocki P, Porcu M, Westerhout EM, et al. Meta-analysis of Neuroblastomas Reveals a Skewed *ALK* Mutation

- Spectrum in Tumors with *MYCN* Amplification. *Clinical Cancer Research*. 2010;16(17):4353-62.
93. Guan J, Tucker ER, Wan H, Chand D, Danielson LS, Ruuth K, et al. The ALK inhibitor PF-06463922 is effective as a single agent in neuroblastoma driven by expression of ALK and MYCN. *Dis Model Mech*. 2016;9(9):941-52.
  94. Heukamp LC, Thor T, Schramm A, De Preter K, Kumps C, De Wilde B, et al. Targeted expression of mutated ALK induces neuroblastoma in transgenic mice. *Sci Transl Med*. 2012;4(141):141ra91.
  95. Matthay KK, George RE, Yu AL. Promising therapeutic targets in neuroblastoma. *Clin Cancer Res*. 2012;18(10):2740-53.
  96. Elsayes KM, Mukundan G, Narra VR, Lewis Jr JS, Shirkhoda A, Farooki A, et al. Adrenal masses: MR imaging features with pathologic correlation. *Radiographics*. 2004;24(suppl\_1):S73-S86.
  97. Nour-Eldin N-EA, Abdelmonem O, Tawfik AM, Naguib NN, Klingebiel T, Rolle U, et al. Pediatric primary and metastatic neuroblastoma: MRI findings: pictorial review. *Magnetic resonance imaging*. 2012;30(7):893-906.
  98. Biomarkers Definitions Working G. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther*. 2001;69(3):89-95.
  99. O'Connor JP, Aboagye EO, Adams JE, Aerts HJ, Barrington SF, Beer AJ, et al. Imaging biomarker roadmap for cancer studies. *Nat Rev Clin Oncol*. 2017;14(3):169-86.
  100. Daldrup-Link HE, Sammet C, Hernanz-Schulman M, Barsness KA, Cahill AM, Chung E, et al. White paper on P4 concepts for pediatric imaging. *Journal of the American College of Radiology*. 2016;13(5):590-7. e2.
  101. George SL, Izquierdo E, Campbell J, Koutroumanidou E, Proszek P, Jamal S, et al. A tailored molecular profiling programme for children with cancer to identify clinically actionable genetic alterations. *European Journal of Cancer*. 2019;121:224-35.
  102. Park JR, Bagatell R, Cohn SL, Pearson AD, Villablanca JG, Berthold F, et al. Revisions to the international neuroblastoma response criteria: a consensus statement from the National Cancer Institute clinical trials planning meeting. *Journal of Clinical Oncology*. 2017;35(22):2580.
  103. Weiser DA, Kaste SC, Siegel MJ, Adamson PC. Imaging in childhood cancer: a Society for Pediatric Radiology and Children's Oncology Group Joint Task Force report. *Pediatr Blood Cancer*. 2013;60(8):1253-60.
  104. Gahr N, Darge K, Hahn G, Kreher BW, von Buiren M, Uhl M. Diffusion-weighted MRI for differentiation of neuroblastoma and ganglioneuroblastoma/ganglioneuroma. *European journal of radiology*. 2011;79(3):443-6.
  105. Boulton JK, Jamin Y, Jacobs V, Gilmour LD, Walker-Samuel S, Halliday J, et al. False-negative MRI biomarkers of tumour response to targeted cancer therapeutics. *British journal of cancer*. 2012;106(12):1960.
  106. Waterton JC, Pylkkanen L. Qualification of imaging biomarkers for oncology drug development. *European journal of cancer*. 2012;48(4):409-15.
  107. Gurcan MN, Boucheron L, Can A, Madabhushi A, Rajpoot N, Yener B. Histopathological image analysis: A review. *IEEE reviews in biomedical engineering*. 2009;2:147.

108. Tabesh A, Teverovskiy M, Pang H-Y, Kumar VP, Verbel D, Kotsianti A, et al. Multifeature prostate cancer diagnosis and Gleason grading of histological images. *IEEE transactions on medical imaging*. 2007;26(10):1366-78.
109. Kumar R, Srivastava R, Srivastava S. Detection and classification of cancer from microscopic biopsy images using clinically significant and biologically interpretable features. *Journal of medical engineering*. 2015;2015.
110. Veta M, Pluim JP, Van Diest PJ, Viergever MA. Breast cancer histopathology image analysis: A review. *IEEE Transactions on Biomedical Engineering*. 2014;61(5):1400-11.
111. Hajdu SI. Microscopic contributions of pioneer pathologists. *Annals of Clinical & Laboratory Science*. 2011;41(2):201-6.
112. Allard FD, Goldsmith JD, Ayata G, Challies TL, Najarian RM, Nasser IA, et al. Intraobserver and interobserver variability in the assessment of dysplasia in ampullary mucosal biopsies. *The American journal of surgical pathology*. 2018;42(8):1095-100.
113. Gomes DS, Porto SS, Balabram D, Gobbi H. Inter-observer variability between general pathologists and a specialist in breast pathology in the diagnosis of lobular neoplasia, columnar cell lesions, atypical ductal hyperplasia and ductal carcinoma in situ of the breast. *Diagnostic pathology*. 2014;9(1):121.
114. Krupinski EA, Tillack AA, Richter L, Henderson JT, Bhattacharyya AK, Scott KM, et al. Eye-movement study and human performance using telepathology virtual slides. Implications for medical education and differences with experience. *Human pathology*. 2006;37(12):1543-56.
115. Mukhopadhyay S, Feldman MD, Abels E, Ashfaq R, Beltaifa S, Cacciabeve NG, et al. Whole slide imaging versus microscopy for primary diagnosis in surgical pathology: a multicenter blinded randomized noninferiority study of 1992 cases (pivotal study). *The American journal of surgical pathology*. 2018;42(1):39.
116. Kothari S, Phan JH, Stokes TH, Wang MD. Pathology imaging informatics for quantitative analysis of whole-slide images. *J Am Med Inform Assoc*. 2013;20(6):1099-108.
117. Campanella G, Hanna MG, Geneslaw L, Miraflor A, Silva VWK, Busam KJ, et al. Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. *Nature medicine*. 2019;25(8):1301-9.
118. Lundervold AS, Lundervold A. An overview of deep learning in medical imaging focusing on MRI. *Zeitschrift für Medizinische Physik*. 2019;29(2):102-27.
119. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, et al. A survey on deep learning in medical image analysis. *Medical image analysis*. 2017;42:60-88.
120. Ranschaert ER, Morozov S, Algra PR. *Artificial Intelligence in Medical Imaging: Opportunities, Applications and Risks*: Springer; 2019.
121. Silver D, Schrittwieser J, Simonyan K, Antonoglou I, Huang A, Guez A, et al. Mastering the game of go without human knowledge. *Nature*. 2017;550(7676):354-9.
122. Cortes C, Vapnik V. Support-vector networks. *Machine learning*. 1995;20(3):273-97.
123. Vapnik V. *The nature of statistical learning theory*: Springer science & business media; 2013.
124. LeCun Y, Bengio Y, Hinton G. Deep learning. *nature*. 2015;521(7553):436-44.
125. Hinton G. Deep learning—a technology with the potential to transform health care. *Jama*. 2018;320(11):1101-2.

126. Krizhevsky A, Sutskever I, Hinton GE, editors. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*; 2012.
127. Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, et al. Imagenet large scale visual recognition challenge. *International journal of computer vision*. 2015;115(3):211-52.
128. Rosenblatt F. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*. 1958;65(6):386.
129. Shih-Chung BL, Freedman MT, Lin J-S, Mun SK. Automatic lung nodule detection using profile matching and back-propagation neural network techniques. *Journal of Digital Imaging*. 1993;6(1):48-54.
130. Janowczyk A, Madabhushi A. Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases. *Journal of pathology informatics*. 2016;7.
131. Ronneberger O, Fischer P, Brox T, editors. U-net: Convolutional networks for biomedical image segmentation. *International Conference on Medical image computing and computer-assisted intervention*; 2015: Springer.
132. Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research*. 2014;15(1):1929-58.
133. Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv:150203167*. 2015.
134. Yuan Y, Failmezger H, Rueda OM, Ali HR, Gräf S, Chin S-F, et al. Quantitative image analysis of cellular heterogeneity in breast tumors complements genomic profiling. *Science translational medicine*. 2012;4(157):157ra43-ra43.
135. Otsu N. A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics*. 1979;9(1):62-6.
136. Cloppet F, Boucher A, editors. Segmentation of overlapping/aggregating nuclei cells in biological images. *2008 19th International Conference on Pattern Recognition*; 2008: IEEE.
137. Pau G, Fuchs F, Sklyar O, Boutros M, Huber W. EBImage—an R package for image processing with applications to cellular phenotypes. *Bioinformatics*. 2010;26(7):979-81.
138. Haralick RM, Shanmugam K. Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics*. 1973;3(6):610-21.
139. Hu M-K. Visual pattern recognition by moment invariants. *IRE transactions on information theory*. 1962;8(2):179-87.
140. Zernike F. Diffraction theory of the cut procedure and its improved form, the phase contrast method. *Physica*. 1934;1(689-704):56.
141. Dimitriadou E, Hornik K, Leisch F, Meyer D, Weingessel A. Misc functions of the Department of Statistics (e1071), TU Wien. R package. 2008;1:5-24.
142. Chen CL, Mahjoubfar A, Tai L-C, Blaby IK, Huang A, Niazi KR, et al. Deep learning in label-free cell classification. *Scientific reports*. 2016;6:21471.
143. Sirinukunwattana K, Raza SEA, Tsang Y-W, Snead DR, Cree IA, Rajpoot NM. Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images. *IEEE transactions on medical imaging*. 2016;35(5):1196-206.
144. Jacobs JG, Brostow GJ, Freeman A, Alexander DC, Panagiotaki E. Detecting and classifying nuclei on a budget. *Intravascular Imaging and Computer Assisted Stenting*,

- and Large-Scale Annotation of Biomedical Data and Expert Label Synthesis: Springer; 2017. p. 77-86.
145. Shkolyar A, Gefen A, Benayahu D, Greenspan H, editors. Automatic detection of cell divisions (mitosis) in live-imaging microscopy images using convolutional neural networks. 2015 37th annual international conference of the IEEE engineering in medicine and biology society (EMBC); 2015: IEEE.
  146. Reinhard E, Adhikhmin M, Gooch B, Shirley P. Colourtransfer between images. *IEEE Computer graphics and applications*. 2001;21(5):34-41.
  147. Achanta R, Shaji A, Smith K, Lucchi A, Fua P, Ssstrunk S. SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE transactions on pattern analysis and machine intelligence*. 2012;34(11):2274-82.
  148. Imran M, Hashim R, Khalid NEA, editors. Segmentation-based Fractal Texture Analysis and ColourLayout Descriptor for Content Based Image Retrieval. *Intelligent Systems Design and Applications (ISDA), 2014 14th International Conference on*; 2014: IEEE.
  149. Spetz A-L, Strominger J, Groh-Spies V. T cell subsets in normal human epidermis. *The American journal of pathology*. 1996;149(2):665.
  150. Held M, Schmitz MH, Fischer B, Walter T, Neumann B, Olma MH, et al. CellCognition: time-resolved phenotype annotation in high-throughput live cell imaging. *Nature methods*. 2010;7(9):747.
  151. Zhong Q, Busetto AG, Fededa JP, Buhmann JM, Gerlich DW. Unsupervised modeling of cell morphology dynamics for time-lapse microscopy. *nature methods*. 2012;9(7):711.
  152. Failmezger H, Frhlich H, Tresch A. Unsupervised automated high throughput phenotyping of RNAi time-lapse movies. *BMC bioinformatics*. 2013;14(1):292.
  153. Failmezger H, Praveen P, Tresch A, Frhlich H. Learning gene network structure from time laps cell imaging in RNAi Knock downs. *Bioinformatics*. 2013;29(12):1534-40.
  154. Niederberger T, Failmezger H, Uskat D, Poron D, Glauche I, Scherf N, et al. Factor graph analysis of live cell-imaging data reveals mechanisms of cell fate decisions. *Bioinformatics*. 2015;31(11):1816-23.
  155. Failmezger H, Dursun E, Dmcke S, Endeke M, Poron D, Schroeder T, et al. Clustering of samples with a tree-shaped dependence structure, with an application to microscopic time lapse imaging. *Bioinformatics*. 2019;35(13):2291-9.
  156. Karimaghaloo Z, Arnold DL, Arbel T. Adaptive multi-level conditional random fields for detection and segmentation of small enhanced pathology in medical images. *Medical image analysis*. 2016;27:17-30.
  157. Chen L-C, Papandreou G, Kokkinos I, Murphy K, Yuille AL. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE transactions on pattern analysis and machine intelligence*. 2018;40(4):834-48.
  158. Li Y, Ping W. Cancer Metastasis Detection With Neural Conditional Random Field. *arXiv preprint arXiv:180607064*. 2018.
  159. Zanjani FG, Zinger S, editors. Cancer detection in histopathology whole-slide images using conditional random fields on deep embedded spaces. *Medical Imaging 2018: Digital Pathology*; 2018: International Society for Optics and Photonics.

160. Rajapakse JC, Liu S, editors. Staging tissues with conditional random fields. 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society; 2011: IEEE.
161. Li C, Chen H, Xue D, Hu Z, Zhang L, He L, et al., editors. Weakly Supervised Cervical Histopathological Image Classification Using Multilayer Hidden Conditional Random Fields. International Conference on Information Technologies in Biomedicine; 2019: Springer.
162. Paramanandam M, O'Byrne M, Ghosh B, Mammen JJ, Manipadam MT, Thamburaj R, et al. Automated segmentation of nuclei in breast cancer histopathology images. *PloS one*. 2016;11(9):e0162053.
163. Lafferty J, McCallum A, Pereira FC. Conditional random fields: Probabilistic models for segmenting and labeling sequence data. 2001.
164. Schmidt M. UGM: A Matlab toolbox for probabilistic undirected graphical models. 2007.
165. de Kruijf EM, van Nes JG, van de Velde CJ, Putter H, Smit VT, Liefers GJ, et al. Tumor–stroma ratio in the primary tumor is a prognostic factor in early breast cancer patients, especially in triple-negative carcinoma patients. *Breast cancer research and treatment*. 2011;125(3):687-96.
166. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *cell*. 2011;144(5):646-74.
167. Wu J, Liang C, Chen M, Su W. Association between tumor-stroma ratio and prognosis in solid tumor patients: a systematic review and meta-analysis. *Oncotarget*. 2016;7(42):68954.
168. Scheer R, Baidoshvili A, Zoidze S, Elferink MA, Berkel AE, Klaase JM, et al. Tumor-stroma ratio as prognostic factor for survival in rectal adenocarcinoma: a retrospective cohort study. *World journal of gastrointestinal oncology*. 2017;9(12):466.
169. Barnes TA, Amir E. HYPE or HOPE: the prognostic value of infiltrating immune cells in cancer. *British journal of cancer*. 2017;117(4):451-60.
170. Kokkinos I, editor UberNet: Training a Universal Convolutional Neural Network for Low-, Mid-, and High-Level Vision Using Diverse Datasets and Limited Memory. *CVPR*; 2017.
171. Arnab A, Zheng S, Jayasumana S, Romera-Paredes B, Larsson M, Kirillov A, et al. Conditional random fields meet deep neural networks for semantic segmentation. *IEEE Signal Processing Magazine*. 2018;2.
172. Qin Y, Kamnitsas K, Ancha S, Navavati J, Cottrell G, Criminisi A, et al. Autofocus Layer for Semantic Segmentation. *arXiv preprint arXiv:180508403*. 2018.
173. Roy AG, Navab N, Wachinger C. Concurrent Spatial and Channel Squeeze & Excitation in Fully Convolutional Networks. *arXiv preprint arXiv:180302579*. 2018.
174. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z, editors. Rethinking the inception architecture for computer vision. *Proceedings of the IEEE conference on computer vision and pattern recognition*; 2016.
175. Szegedy C, Ioffe S, Vanhoucke V, Alemi AA, editors. Inception-v4, inception-resnet and the impact of residual connections on learning. *Thirty-First AAAI Conference on Artificial Intelligence*; 2017.
176. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:14091556*. 2014.

177. Komura D, Ishikawa S. Machine learning methods for histopathological image analysis. *Computational and structural biotechnology journal*. 2018;16:34-42.
178. Humphrey PA, Moch H, Cubilla AL, Ulbright TM, Reuter VE. The 2016 WHO classification of tumours of the urinary system and male genital organs—part B: prostate and bladder tumours. *European urology*. 2016;70(1):106-19.
179. Rakha EA, Reis-Filho JS, Baehner F, Dabbs DJ, Decker T, Eusebi V, et al. Breast cancer prognostic classification in the molecular era: the role of histological grade. *Breast Cancer Research*. 2010;12(4):207.
180. Ma W, Wang J, Yu L, Zhang X, Wang Z, Tan B, et al. Tumor-stroma ratio is an independent predictor for survival in esophageal squamous cell carcinoma. *Journal of Thoracic Oncology*. 2012;7(9):1457-61.
181. Chen Y, Zhang L, Liu W, Liu X. Prognostic significance of the tumor-stroma ratio in epithelial ovarian cancer. *BioMed research international*. 2015;2015.
182. Ruan M, Tian T, Rao J, Xu X, Yu B, Yang W, et al. Predictive value of tumor-infiltrating lymphocytes to pathological complete response in neoadjuvant treated triple-negative breast cancers. *Diagnostic pathology*. 2018;13(1):66.
183. Renshaw AA, Chevillet JC. Quantitative tumour necrosis is an independent predictor of overall survival in clear cell renal cell carcinoma. *Pathology*. 2015;47(1):34-7.
184. Pichler M, Hutterer GC, Chromecki TF, Jesche J, Kappel-Kettner K, Rehak P, et al. Histologic tumor necrosis is an independent prognostic indicator for clear cell and papillary renal cell carcinoma. *American journal of clinical pathology*. 2012;137(2):283-9.
185. Bredholt G, Mannelqvist M, Stefansson IM, Birkeland E, Bø TH, Øyan AM, et al. Tumor necrosis is an important hallmark of aggressive endometrial cancer and associates with hypoxia, angiogenesis and inflammation responses. *Oncotarget*. 2015;6(37):39676.
186. Wang D, Khosla A, Gargeya R, Irshad H, Beck AH. Deep learning for identifying metastatic breast cancer. *arXiv preprint arXiv:160605718*. 2016.
187. Nahid A-A, Mehrabi MA, Kong Y. Histopathological breast cancer image classification by deep neural network techniques guided by local clustering. *BioMed research international*. 2018;2018.
188. Raza SEA, Cheung L, Shaban M, Graham S, Epstein D, Pelengaris S, et al. Micro-Net: A unified model for segmentation of various objects in microscopy images. *Medical image analysis*. 2019;52:160-73.
189. Bándi P, van de Loo R, Intezar M, Geijs D, Ciompi F, van Ginneken B, et al., editors. Comparison of different methods for tissue segmentation in histopathological whole-slide images. 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017); 2017: IEEE.
190. Bejnordi BE, Zuidhof G, Balkenhol M, Hermsen M, Bult P, van Ginneken B, et al. Context-aware stacked convolutional neural networks for classification of breast carcinomas in whole-slide histopathology images. *Journal of Medical Imaging*. 2017;4(4):044504.
191. Vu QD, Graham S, Kurc T, To MNN, Shaban M, Qaiser T, et al. Methods for segmentation and classification of digital microscopy tissue images. *Frontiers in bioengineering and biotechnology*. 2019;7.

192. Araújo T, Aresta G, Castro E, Rouco J, Aguiar P, Eloy C, et al. Classification of breast cancer histology images using convolutional neural networks. *PLoS one*. 2017;12(6):e0177544.
193. Wetteland R, Engan K, Eftestøl T, Kvikstad V, Janssen EA, editors. Multiclass tissue classification of whole-slide histological images using convolutional neural networks. *Proceedings of the 8th International Conference on Pattern Recognition Applications and Methods*; 2019.
194. Xu Z, Moro CF, Kuznyecov D, Bozóky B, Dong L, Zhang Q, editors. Tissue Region Growing for Histopathology Image Segmentation. *Proceedings of the 2018 3rd International Conference on Biomedical Imaging, Signal Processing*; 2018.
195. Chan L, Hosseini MS, Rowsell C, Plataniotis KN, Damaskinos S, editors. Histosegnet: Semantic segmentation of histological tissue type in whole slide images. *Proceedings of the IEEE International Conference on Computer Vision*; 2019.
196. Xu Y, Jia Z, Wang L-B, Ai Y, Zhang F, Lai M, et al. Large scale tissue histopathology image classification, segmentation, and visualization via deep convolutional activation features. *BMC bioinformatics*. 2017;18(1):281.
197. Song Y, Zhang L, Chen S, Ni D, Lei B, Wang T. Accurate segmentation of cervical cytoplasm and nuclei based on multiscale convolutional network and graph partitioning. *IEEE Transactions on Biomedical Engineering*. 2015;62(10):2421-33.
198. Romo D, García-Arteaga JD, Arbeláez P, Romero E, editors. A discriminant multi-scale histopathology descriptor using dictionary learning. *Medical Imaging 2014: Digital Pathology*; 2014: International Society for Optics and Photonics.
199. Qin P, Chen J, Zeng J, Chai R, Wang L. Large-scale tissue histopathology image segmentation based on feature pyramid. *EURASIP Journal on Image and Video Processing*. 2018;2018(1):75.
200. Xu Z, Zhang Q, editors. Multi-scale context-aware networks for quantitative assessment of colorectal liver metastases. *2018 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)*; 2018: IEEE.
201. Beck AH, Sangoi AR, Leung S, Marinelli RJ, Nielsen TO, Van De Vijver MJ, et al. Systematic analysis of breast cancer morphology uncovers stromal features associated with survival. *Science translational medicine*. 2011;3(108):108ra13-ra13.
202. Xu J, Luo X, Wang G, Gilmore H, Madabhushi A. A deep convolutional neural network for segmenting and classifying epithelial and stromal regions in histopathological images. *Neurocomputing*. 2016;191:214-23.
203. Bejnordi BE, Litjens G, Hermsen M, Karssemeijer N, van der Laak JA, editors. A multi-scale superpixel classification approach to the detection of regions of interest in whole slide histopathology images. *Medical Imaging 2015: Digital Pathology*; 2015: International Society for Optics and Photonics.
204. Moore HC, Wood KM, Jackson MS, Lastowska MA, Hall D, Imrie H, et al. Histological profile of tumours from MYCN transgenic mice. *Journal of clinical pathology*. 2008.
205. Brockmann M, Poon E, Berry T, Carstensen A, Deubzer HE, Rycak L, et al. Small molecule inhibitors of aurora-a induce proteasomal degradation of N-myc in childhood neuroblastoma. *Cancer Cell*. 2013;24(1):75-89.
206. Berry T, Luther W, Bhatnagar N, Jamin Y, Poon E, Sanda T, et al. The ALK(F1174L) mutation potentiates the oncogenic activity of MYCN in neuroblastoma. *Cancer Cell*. 2012;22(1):117-30.

207. Glorot X, Bengio Y, editors. Understanding the difficulty of training deep feedforward neural networks. Proceedings of the thirteenth international conference on artificial intelligence and statistics; 2010.
208. Kingma DP, Ba J. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980. 2014.
209. Chollet F, editor Xception: Deep learning with depthwise separable convolutions. Proceedings of the IEEE conference on computer vision and pattern recognition; 2017.
210. He K, Zhang X, Ren S, Sun J, editors. Deep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition; 2016.
211. Zormpas-Petridis K, Failmezger H, Raza SEA, Roxanis I, Jamin Y, Yuan Y. Superpixel-based Conditional Random Fields (SuperCRF): Incorporating global and local context for enhanced deep learning in melanoma histopathology. *Frontiers in oncology*. 2019;9:1045.
212. Gilchrist KW, Gray R, Fowble B, Tormey DC, Taylor 4th S. Tumor necrosis is a prognostic predictor for early recurrence and death in lymph node-positive breast cancer: a 10-year follow-up study of 728 Eastern Cooperative Oncology Group patients. *Journal of clinical oncology*. 1993;11(10):1929-35.
213. Hanafy E, Al Jabri A, Gadelkarim G, Dasaq A, Nazim F, Al Pakrah M. Tumor histopathological response to neoadjuvant chemotherapy in childhood solid malignancies: is it still impressive? *Journal of Investigative Medicine*. 2018;66(2):289-97.
214. Lambertz I, Kumps C, Claeys S, Lindner S, Beckers A, Janssens E, et al. Upregulation of MAPK negative feedback regulators and RET in mutant ALK neuroblastoma: implications for targeted treatment. *Clinical Cancer Research*. 2015;21(14):3327-39.
215. Gowda SN, Yuan C, editors. ColorNet: Investigating the importance of colourspaces for image classification. Asian Conference on Computer Vision; 2018: Springer.
216. Yao J, Boben M, Fidler S, Urtasun R, editors. Real-time coarse-to-fine topologically preserving segmentation. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2015.
217. Tan M, Le QV. Efficientnet: Rethinking model scaling for convolutional neural networks. arXiv preprint arXiv:1905.11946. 2019.
218. Laura CO, Shekhar R, Wesarg S, Ballester MÁG, Drechsler K, Sato Y, et al. Clinical Image-Based Procedures. *Translational Research in Medical Imaging: 4th International Workshop, CLIP 2015, Held in Conjunction with MICCAI 2015, Munich, Germany, October 5, 2015. Revised Selected Papers: Springer*; 2016.
219. Drew B, Jones EC, Reinsberg S, Yung AC, Goldenberg SL, Kozlowski P. Device for sectioning prostatectomy specimens to facilitate comparison between histology and in vivo MRI. *Journal of Magnetic Resonance Imaging*. 2010;32(4):992-6.
220. Bailey C, Bourne RM, Siow B, Johnston EW, Brizmohun Appayya M, Pye H, et al. VERDICT MRI validation in fresh and fixed prostate specimens using patient-specific moulds for histological and MR alignment. *NMR in Biomedicine*. 2019;32(5):e4073.
221. Wu HH, Priester A, Khoshnoodi P, Zhang Z, Shakeri S, Afshari Mirak S, et al. A system using patient-specific 3D-printed molds to spatially align in vivo MRI with ex vivo MRI and whole-mount histopathology for prostate cancer research. *Journal of Magnetic Resonance Imaging*. 2019;49(1):270-9.

222. Mireia C-O, Gehrung M, Ursprung S, Gill AB, Warren AY, Gallagher FA, et al. 3D-printed moulds of renal tumours for image-guided tissue sampling in the clinical setting. *BioRxiv*. 2019:658831.
223. Kalavagunta C, Zhou X, Schmechel SC, Metzger GJ. Registration of in vivo prostate MRI and pseudo-whole mount histology using Local Affine Transformations guided by Internal Structures (LATIS). *Journal of Magnetic Resonance Imaging*. 2015;41(4):1104-14.
224. Chappelow J, Bloch BN, Rofsky N, Genega E, Lenkinski R, DeWolf W, et al. Elastic registration of multimodal prostate MRI and histology via multiattribute combined mutual information. *Medical Physics*. 2011;38(4):2005-18.
225. Workman P, Aboagye EO, Balkwill F, Balmain A, Bruder G, Chaplin DJ, et al. Guidelines for the welfare and use of animals in cancer research. *Br J Cancer*. 2010;102(11):1555-77.
226. Bourne RM, Bailey C, Johnston EW, Pye H, Heavey S, Whitaker H, et al. Apparatus for histological validation of in vivo and ex vivo magnetic resonance imaging of the human prostate. *Frontiers in oncology*. 2017;7:47.
227. Myronenko A, Song X. Point set registration: Coherent point drift. *IEEE transactions on pattern analysis and machine intelligence*. 2010;32(12):2262-75.
228. Mantel N. The detection of disease clustering and a generalized regression approach. *Cancer research*. 1967;27(2 Part 1):209-20.
229. Meitar D, Crawford SE, Rademaker AW, Cohn SL. Tumor angiogenesis correlates with metastatic disease, N-myc amplification, and poor outcome in human neuroblastoma. *Journal of clinical oncology*. 1996;14(2):405-14.
230. Peddinti R, Zeine R, Luca D, Seshadri R, Chlenski A, Cole K, et al. Prominent microvascular proliferation in clinically aggressive neuroblastoma. *Clinical cancer research*. 2007;13(12):3499-506.
231. Tadeo I, Bueno G, Berbegall AP, Fernández-Carrobles MM, Castel V, García-Rojo M, et al. Vascular patterns provide therapeutic targets in aggressive neuroblastic tumors. *Oncotarget*. 2016;7(15):19935.
232. Jakovljević G, Čulić S, Stepan J, Bonevski A, Seiwert S. Vascular endothelial growth factor in children with neuroblastoma: a retrospective analysis. *Journal of Experimental & Clinical Cancer Research*. 2009;28(1):143.
233. Miyazaki K, Jerome NP, Collins DJ, Orton MR, d'Arcy JA, Wallace T, et al. Demonstration of the reproducibility of free-breathing diffusion-weighted MRI and dynamic contrast enhanced MRI in children with solid tumours: a pilot study. *European radiology*. 2015;25(9):2641-50.
234. O'Connor JP, Jayson GC. Do imaging biomarkers relate to outcome in patients treated with VEGF inhibitors? *Clinical Cancer Research*. 2012;18(24):6588-98.
235. Wu EX, Tang H, Jensen JH. Applications of ultrasmall superparamagnetic iron oxide contrast agents in the MR study of animal models. *NMR Biomed*. 2004;17(7):478-83.
236. Robinson SP, Howe FA, Griffiths JR, Ryan AJ, Waterton JC. Susceptibility contrast magnetic resonance imaging determination of fractional tumor blood volume: a noninvasive imaging biomarker of response to the vascular disrupting agent ZD6126. *International Journal of Radiation Oncology\* Biology\* Physics*. 2007;69(3):872-9.

237. Robinson SP, Boulton JK, Vasudev NS, Reynolds AR. Monitoring the vascular response and resistance to sunitinib in renal cell carcinoma in vivo with susceptibility contrast MRI. *Cancer Research*. 2017;canres. 0248.2017.
238. Vlachogiannis G, Hedayat S, Vatsiou A, Jamin Y, Fernández-Mateos J, Khan K, et al. Patient-derived organoids model treatment response of metastatic gastrointestinal cancers. *Science*. 2018;359(6378):920-6.
239. Kilkenny C, Browne WJ, Cuthill IC, Emerson M, Altman DG. Improving bioscience research reporting: the ARRIVE guidelines for reporting animal research. *PLoS Biol*. 2010;8(6):e1000412.
240. Schindelin J, Rueden CT, Hiner MC, Eliceiri KW. The ImageJ ecosystem: An open platform for biomedical image analysis. *Molecular reproduction and development*. 2015;82(7-8):518-29.
241. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: an open-source platform for biological-image analysis. *Nature methods*. 2012;9(7):676-82.
242. Allen LM, Hasso AN, Handwerker J, Farid H. Sequence-specific MR imaging findings that are useful in dating ischemic stroke. *Radiographics*. 2012;32(5):1285-97.
243. Almeida GS, Panek R, Hallsworth A, Webber H, Papaevangelou E, Boulton JK, et al. Pre-clinical imaging of transgenic mouse models of neuroblastoma using a dedicated 3-element solenoid coil on a clinical 3T platform. *British journal of cancer*. 2017;117(6):791-800.
244. Manias KA, Gill SK, MacPherson L, Foster K, Oates A, Peet AC. Magnetic resonance imaging based functional imaging in paediatric oncology. *European Journal of Cancer*. 2017;72:251-65.
245. Daudigeos-Dubus E, Le Dret L, Bawa O, Opolon P, Vievard A, Villa I, et al. Dual inhibition using cabozantinib overcomes HGF/MET signaling mediated resistance to pan-VEGFR inhibition in orthotopic and metastatic neuroblastoma tumors. *International journal of oncology*. 2017;50(1):203-11.
246. Daudigeos-Dubus E, Le Dret L, Lanvers-Kaminsky C, Bawa O, Opolon P, Vievard A, et al. Regorafenib: antitumor activity upon mono and combination therapy in preclinical pediatric malignancy models. *PloS one*. 2015;10(11).
247. Calero R, Morchon E, Johnsen JI, Serrano R. Sunitinib suppress neuroblastoma growth through degradation of MYCN and inhibition of angiogenesis. *PloS one*. 2014;9(4).
248. Rössler J, Monnet Y, Farace F, Opolon P, Daudigeos-Dubus E, Bourredjem A, et al. The selective VEGFR1-3 inhibitor axitinib (AG-013736) shows antitumor activity in human neuroblastoma xenografts. *International journal of cancer*. 2011;128(11):2748-58.
249. Cazes A, Lopez-Delisle L, Tsarovina K, Pierre-Eugène C, De Preter K, Peuchmaur M, et al. Activated Alk triggers prolonged neurogenesis and Ret upregulation providing a therapeutic target in ALK-mutated neuroblastoma. *Oncotarget*. 2014;5(9):2688.
250. Chanthery YH, Gustafson WC, Itsara M, Persson A, Hackett CS, Grimmer M, et al. Paracrine signaling through MYCN enhances tumor-vascular interactions in neuroblastoma. *Science translational medicine*. 2012;4(115):115ra3-ra3.
251. Di Paolo D, Ambrogio C, Pastorino F, Brignole C, Martinengo C, Carosio R, et al. Selective therapeutic targeting of the anaplastic lymphoma kinase with liposomal siRNA

- induces apoptosis and inhibits angiogenesis in neuroblastoma. *Molecular Therapy*. 2011;19(12):2201-12.
252. Bid HK, Kerk S. BET bromodomain inhibitor (JQ1) and tumor angiogenesis. *Oncoscience*. 2016;3(11-12):316.
253. Romain C, Paul P, Kim KW, Lee S, Qiao J, Chung DH. Targeting Aurora kinase-A downregulates cell proliferation and angiogenesis in neuroblastoma. *Journal of pediatric surgery*. 2014;49(1):159-65.
254. Chen Z, Zhao Y, Yu Y, Pang JC, Woodfield SE, Tao L, et al. Small molecule inhibitor regorafenib inhibits RET signaling in neuroblastoma cells and effectively suppresses tumor growth in vivo. *Oncotarget*. 2017;8(61):104090.
255. Kortlever RM, Sodikin NM, Wilson CH, Burkhart DL, Pellegrinet L, Swigart LB, et al. Myc cooperates with Ras by programming inflammation and immune suppression. *Cell*. 2017;171(6):1301-15. e14.
256. Dejure FR, Eilers M. MYC and tumor metabolism: chicken and egg. *The EMBO journal*. 2017;36(23):3409-20.
257. Berry T, Luther W, Bhatnagar N, Jamin Y, Poon E, Sanda T, et al. The ALKF1174L mutation potentiates the oncogenic activity of MYCN in neuroblastoma. *Cancer cell*. 2012;22(1):117-30.
258. Lopez-Delisle L, Pierre-Eugène C, Louis-Brennetot C, Surdez D, Raynal V, Baulande S, et al. Activated ALK signals through the ERK–ETV5–RET pathway to drive neuroblastoma oncogenesis. *Oncogene*. 2018;37(11):1417-29.
259. Huang D, Rutkowski JL, Brodeur GM, Chou PM, Kwiatkowski JL, Babbo A, et al. Schwann cell-conditioned medium inhibits angiogenesis. *Cancer research*. 2000;60(21):5966-71.
260. Chlenski A, Liu S, Crawford SE, Volpert OV, DeVries GH, Evangelista A, et al. SPARC is a key Schwannian-derived inhibitor controlling neuroblastoma tumor angiogenesis. *Cancer Research*. 2002;62(24):7357-63.
261. Poliani P, Mitola S, Ravanini M, Ferrari-Toninelli G, D'Ippolito C, Notarangelo L, et al. CEACAM1/VEGF cross-talk during neuroblastic tumour differentiation. *The Journal of Pathology: A Journal of the Pathological Society of Great Britain and Ireland*. 2007;211(5):541-9.
262. Brisse HJ, McCarville MB, Granata C, Krug KB, Wootton-Gorges SL, Kanegawa K, et al. Guidelines for imaging and staging of neuroblastic tumors: consensus report from the International Neuroblastoma Risk Group Project. *Radiology*. 2011;261(1):243-57.
263. Verlhac S, Morel M, Bernaudin F, Béchet S, Jung C, Vasile M. Liver iron overload assessment by MRI R2\* relaxometry in highly transfused pediatric patients: an agreement and reproducibility study. *Diagnostic and interventional imaging*. 2015;96(3):259-64.
264. Peuchmaur M, d'Amore ES, Joshi VV, Hata Ji, Roald B, Dehner LP, et al. Revision of the International Neuroblastoma Pathology Classification: confirmation of favorable and unfavorable prognostic subsets in ganglioneuroblastoma, nodular. *Cancer: Interdisciplinary International Journal of the American Cancer Society*. 2003;98(10):2274-81.
265. Toth GB, Varallyay CG, Horvath A, Bashir MR, Choyke PL, Daldrup-Link HE, et al. Current and potential imaging applications of ferumoxytol for magnetic resonance imaging. *Kidney international*. 2017;92(1):47-66.

266. Mosse YP, Fox E, Teachey DT, Reid JM, Safgren SL, Carol H, et al. A phase II study of alisertib in children with recurrent/refractory solid tumors or leukemia: Children's Oncology Group Phase I and pilot Consortium (ADVL0921). *Clinical Cancer Research*. 2019;25(11):3229-38.
267. Vaughan L, Clarke PA, Barker K, Chanthery Y, Gustafson CW, Tucker E, et al. Inhibition of mTOR-kinase destabilizes MYCN and is a potential therapy for MYCN-dependent tumors. *Oncotarget*. 2016;7(36):57525.
268. Brockmann M, Poon E, Berry T, Carstensen A, Deubzer HE, Rycak L, et al. Small molecule inhibitors of aurora-a induce proteasomal degradation of N-myc in childhood neuroblastoma. *Cancer cell*. 2013;24(1):75-89.
269. Caruana D, Wei W, Martinez-Morilla S, Rimm DL, Reisenbichler ES. Association between low estrogen receptor positive breast cancer and staining performance. *NPJ Breast Cancer*. 2020;6(1):1-6.
270. Michael IP, Saghafinia S, Tichet M, Zangger N, Marinoni I, Perren A, et al. ALK7 signaling manifests a homeostatic tissue barrier that is abrogated during tumorigenesis and metastasis. *Developmental cell*. 2019;49(3):409-24. e6.
271. Ribeiro GP, Endringer DC, De Andrade TU, Lenz D. Comparison between two programs for image analysis, machine learning and subsequent classification. *Tissue and Cell*. 2019;58:12-6.
272. Bankhead P, Loughrey MB, Fernández JA, Dombrowski Y, McArt DG, Dunne PD, et al. QuPath: Open source software for digital pathology image analysis. *Scientific reports*. 2017;7(1):16878.
273. Loughrey MB, Bankhead P, Coleman HG, Hagan RS, Craig S, McCorry AM, et al. Validation of the systematic scoring of immunohistochemically stained tumour tissue microarrays using QuPath digital image analysis. *Histopathology*. 2018;73(2):327-38.
274. Motulsky HJ, Brown RE. Detecting outliers when fitting data with nonlinear regression—a new method based on robust nonlinear regression and the false discovery rate. *BMC bioinformatics*. 2006;7(1):123.
275. Zormpas-Petridis K, Jerome NP, Blackledge MD, Carceller F, Poon E, Clarke M, et al. MRI imaging of the hemodynamic vasculature of neuroblastoma predicts response to antiangiogenic treatment. *Cancer research*. 2019;79(11):2978-91.
276. Lu H, Clingman C, Golay X, Van Zijl PC. Determining the longitudinal relaxation time (T1) of blood at 3.0 Tesla. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*. 2004;52(3):679-82.
277. Acerbi I, Cassereau L, Dean I, Shi Q, Au A, Park C, et al. Human breast cancer invasion and aggression correlates with ECM stiffening and immune cell infiltration. *Integrative Biology*. 2015;7(10):1120-34.
278. McSheehy PM, Weidensteiner C, Cannet C, Ferretti S, Laurent D, Ruetz S, et al. Quantified tumor T1 is a generic early-response imaging biomarker for chemotherapy reflecting cell viability. *Clinical Cancer Research*. 2010;16(1):212-25.
279. Weidensteiner C, Allegrini PR, Sticker-Jantscheff M, Romanet V, Ferretti S, McSheehy PM. Tumour T1 changes in vivo are highly predictive of response to chemotherapy and reflect the number of viable tumour cells—a preclinical MR study in mice. *BMC cancer*. 2014;14(1):88.
280. Damadian R. Tumor detection by nuclear magnetic resonance. *Science*. 1971;171(3976):1151-3.

281. Hollis DP, Economou JS, Parks LC, Eggleston JC, Saryan LA, Czeisler JL. Nuclear magnetic resonance studies of several experimental and human malignant tumors. *Cancer Research*. 1973;33(9):2156-60.
282. Beall PT, Asch BB, Chang DC, Medina D, Hazlewood CF. Distinction of normal, preneoplastic, and neoplastic mouse mammary primary cell cultures by water nuclear magnetic resonance relaxation times. *Journal of the National Cancer Institute*. 1980;64(2):335-8.
283. Beall PT, Hazlewood CF, Rao PN. Nuclear magnetic resonance patterns of intracellular water as a function of HeLa cell cycle. *Science*. 1976;192(4242):904-7.
284. Päufer S, Zschunke A, Khuen A, Keller K. Estimation of water content and water mobility in the nucleus and cytoplasm of *Xenopus laevis* oocytes by NMR microscopy. *Magnetic resonance imaging*. 1995;13(2):269-76.
285. Schoeniger J, Aiken N, Hsu E, Blackband S. Relaxation-time and diffusion NMR microscopy of single neurons. *Journal of Magnetic Resonance, Series B*. 1994;103(3):261-73.
286. Franck JM, Ding Y, Stone K, Qin PZ, Han S. Anomalously rapid hydration water diffusion dynamics near DNA surfaces. *Journal of the American Chemical Society*. 2015;137(37):12013-23.
287. Kerr JF, Wyllie AH, Currie AR. Apoptosis: a basic biological phenomenon with wideranging implications in tissue kinetics. *British journal of cancer*. 1972;26(4):239-57.
288. Gestblom C, Hoehner J, Pålman S. Proliferation and apoptosis in neuroblastoma: subdividing the mitosis-karyorrhexis index. *European Journal of Cancer*. 1995;31(4):458-63.
289. DuBois SG, Mody R, Naranjo A, Van Ryn C, Russ D, Oldridge D, et al. MIBG avidity correlates with clinical features, tumor biology, and outcomes in neuroblastoma: a report from the Children's Oncology Group. *Pediatric blood & cancer*. 2017;64(11):e26545.
290. More SS, Itsara M, Yang X, Geier EG, Tadano MK, Seo Y, et al. Vorinostat increases expression of functional norepinephrine transporter in neuroblastoma in vitro and in vivo model systems. *Clinical Cancer Research*. 2011;17(8):2339-49.
291. Little RA, Jamin Y, Boulton JK, Naish JH, Watson Y, Cheung S, et al. Mapping hypoxia in renal carcinoma with oxygen-enhanced MRI: comparison with intrinsic susceptibility MRI and pathology. *Radiology*. 2018;288(3):739-47.
292. Brett Heimlich J, Speed JS, O'Connor PM, Pollock JS, Townes TM, Meiler SE, et al. Endothelin-1 contributes to the progression of renal injury in sickle cell disease via reactive oxygen species. *British journal of pharmacology*. 2016;173(2):386-95.
293. Haller S, Zaharchuk G, Thomas DL, Lovblad K-O, Barkhof F, Golay X. Arterial spin labeling perfusion of the brain: emerging clinical applications. *Radiology*. 2016;281(2):337-56.
294. Messiou C, Orton M, Ang JE, Collins DJ, Morgan VA, Mears D, et al. Advanced solid tumors treated with cediranib: comparison of dynamic contrast-enhanced MR imaging and CT as markers of vascular activity. *Radiology*. 2012;265(2):426-36.
295. Aghighi M, Theruvath AJ, Pareek A, Pisani LL, Alford R, Muehe AM, et al. Magnetic resonance imaging of tumor-associated macrophages: clinical translation. *Clinical Cancer Research*. 2018;24(17):4110-8.
296. Perrin J, Capitao M, Mouglin-Degraef M, Guérard F, Faivre-Chauvet A, Rbah-Vidal L, et al. Cell Tracking in Cancer Immunotherapy. *Frontiers in Medicine*. 2020;7:34.

297. Oishi D, Jeungho S, Al-Wakeel N, Berger F, Kuehne T, Kozerke S, et al. T1 mapping in ischaemic heart disease. *European Heart Journal–Cardiovascular Imaging*. 2014;15(6):597-602.
298. Schelbert EB, Messroghli DR. State of the art: clinical applications of cardiac T1 mapping. *Radiology*. 2016;278(3):658-76.
299. Pavlides M, Banerjee R, Sellwood J, Kelly CJ, Robson MD, Booth JC, et al. Multiparametric magnetic resonance imaging predicts clinical outcomes in patients with chronic liver disease. *Journal of hepatology*. 2016;64(2):308-15.
300. Sado DM, Maestrini V, Piechnik SK, Banyersad SM, White SK, Flett AS, et al. Noncontrast myocardial T1 mapping using cardiovascular magnetic resonance for iron overload. *Journal of magnetic resonance imaging*. 2015;41(6):1505-11.
301. Wu WC, Jain V, Li C, Giannetta M, Hurt H, Wehrli FW, et al. In vivo venous blood T1 measurement using inversion recovery true-FISP in children and adults. *Magnetic resonance in medicine*. 2010;64(4):1140-7.
302. Gahr N, Darge K, Hahn G, Kreher BW, von Buiren M, Uhl M. Diffusion-weighted MRI for differentiation of neuroblastoma and ganglioneuroblastoma/ganglioneuroma. *Eur J Radiol*. 2011;79(3):443-6.
303. Peschmann AL, Beer M, Ammann B, Dreyhaupt J, Kneer K, Beer AJ, et al. Quantitative DWI predicts event-free survival in children with neuroblastic tumours: preliminary findings from a retrospective cohort study. *Eur Radiol Exp*. 2019;3(1):6.
304. Meeus EM, Novak J, Withey SB, Zarinabad N, Dehghani H, Peet AC. Evaluation of intravoxel incoherent motion fitting methods in low-perfused tissue. *J Magn Reson Imaging*. 2017;45(5):1325-34.
305. Jain N, Halbert J, Patel PA, Biassoni L, Anderson J, Sebire N, et al. Importance of Magnetic Resonance Imaging With Diffusion-weighted Imaging in Guiding Biopsy of Nodular Ganglioneuroblastoma: A Case Report. *J Pediatr Hematol Oncol*. 2019.
306. Sinkus R, Van Beers BE, Vilgrain V, DeSouza N, Waterton JC. Apparent diffusion coefficient from magnetic resonance imaging as a biomarker in oncology drug development. *Eur J Cancer*. 2012;48(4):425-31.
307. Padhani AR, Koh DM. Diffusion MR imaging for monitoring of treatment response. *Magn Reson Imaging Clin N Am*. 2011;19(1):181-209.
308. Jamin Y, Glass L, Hallsworth A, George R, Koh DM, Pearson AD, et al. Intrinsic susceptibility MRI identifies tumors with ALKF1174L mutation in genetically-engineered murine models of high-risk neuroblastoma. *PLoS One*. 2014;9(3):e92886.
309. Zormpas-Petridis K, Jerome NP, Blackledge MD, Carceller F, Poon E, Clarke M, et al. MRI Imaging of the Hemodynamic Vasculature of Neuroblastoma Predicts Response to Antiangiogenic Treatment. *Cancer Res*. 2019;79(11):2978-91.
310. Humphries PD, Sebire NJ, Siegel MJ, Olsen OE. Tumors in pediatric patients at diffusion-weighted MR imaging: apparent diffusion coefficient and tumor cellularity. *Radiology*. 2007;245(3):848-54.
311. Hill DK, Heindl A, Zormpas-Petridis K, Collins DJ, Euceda LR, Rodrigues DN, et al. Non-Invasive Prostate Cancer Characterization with Diffusion-Weighted MRI: Insight from In silico Studies of a Transgenic Mouse Model. *Front Oncol*. 2017;7:290.
312. Atlas SW, DuBois P, Singer MB, Lu D. Diffusion measurements in intracranial hematomas: implications for MR imaging of acute stroke. *AJNR Am J Neuroradiol*. 2000;21(7):1190-4.

313. Ambros IM, Hata J, Joshi VV, Roald B, Dehner LP, Tuchler H, et al. Morphologic features of neuroblastoma (Schwannian stroma-poor tumors) in clinically favorable and unfavorable groups. *Cancer*. 2002;94(5):1574-83.
314. Shimada H, Ambros IM, Dehner LP, Hata J, Joshi VV, Roald B, et al. The International Neuroblastoma Pathology Classification (the Shimada system). *Cancer*. 1999;86(2):364-72.
315. Adigun R, Basit H, Murray J. Necrosis, Cell (Liquefactive, Coagulative, Caseous, Fat, Fibrinoid, and Gangrenous). *StatPearls*. Treasure Island (FL)2020.
316. Jardim-Perassi BV, Huang S, Dominguez-Viqueira W, Poleszczuk J, Budzevich MM, Abdalah MA, et al. Multiparametric MRI and Coregistered Histology Identify Tumor Habitats in Breast Cancer Mouse Models. *Cancer Res*. 2019;79(15):3952-64.
317. LaViolette PS, Mickevicius NJ, Cochran EJ, Rand SD, Connelly J, Bovi JA, et al. Precise ex vivo histological validation of heightened cellularity and diffusion-restricted necrosis in regions of dark apparent diffusion coefficient in 7 cases of high-grade glioma. *Neuro Oncol*. 2014;16(12):1599-606.
318. Zakhari N, Taccone MS, Torres C, Chakraborty S, Sinclair J, Woulfe J, et al. Diagnostic Accuracy of Centrally Restricted Diffusion in the Differentiation of Treatment-Related Necrosis from Tumor Recurrence in High-Grade Gliomas. *AJNR Am J Neuroradiol*. 2018;39(2):260-4.
319. Papaevangelou E, Almeida GS, Jamin Y, Robinson SP, deSouza NM. Diffusion-weighted MRI for imaging cell death after cytotoxic or apoptosis-inducing therapy. *Br J Cancer*. 2015;112(9):1471-9.
320. Ham J, Costa C, Sano R, Lochmann TL, Sennott EM, Patel NU, et al. Exploitation of the Apoptosis-Primed State of MYCN-Amplified Neuroblastoma to Develop a Potent and Specific Targeted Therapy Combination. *Cancer Cell*. 2016;29(2):159-72.
321. Fulda S, Lutz W, Schwab M, Debatin KM. MycN sensitizes neuroblastoma cells for drug-induced apoptosis. *Oncogene*. 1999;18(7):1479-86.
322. Vaughan L, Clarke PA, Barker K, Chanthery Y, Gustafson CW, Tucker E, et al. Inhibition of mTOR-kinase destabilizes MYCN and is a potential therapy for MYCN-dependent tumors. *Oncotarget*. 2016;7(36):57525-44.
323. Saraste A, Pulkki K. Morphologic and biochemical hallmarks of apoptosis. *Cardiovasc Res*. 2000;45(3):528-37.
324. Nunez R, Sancho-Martinez SM, Novoa JM, Lopez-Hernandez FJ. Apoptotic volume decrease as a geometric determinant for cell dismantling into apoptotic bodies. *Cell Death Differ*. 2010;17(11):1665-71.
325. Neubauer H, Li M, Muller VR, Pabst T, Beer M. Diagnostic Value of Diffusion-Weighted MRI for Tumor Characterization, Differentiation and Monitoring in Pediatric Patients with Neuroblastic Tumors. *Rofo*. 2017;189(7):640-50.
326. Serin HI, Gorkem SB, Doganay S, Ciraci S, Unal E, Guzel M, et al. Diffusion weighted imaging in differentiating malignant and benign neuroblastic tumors. *Jpn J Radiol*. 2016;34(9):620-4.
327. McDonald K, Sebire NJ, Anderson J, Olsen OE. Patterns of shift in ADC distributions in abdominal tumours during chemotherapy-feasibility study. *Pediatr Radiol*. 2011;41(1):99-106.
328. Demir S, Altinkaya N, Kocer NE, Erbay A, Oguzkurt P. Variations in apparent diffusion coefficient values following chemotherapy in pediatric neuroblastoma. *Diagn Interv Radiol*. 2015;21(2):184-8.

329. Rogers HJ, Verhagen MV, Shelmerdine SC, Clark CA, Hales PW. An alternative approach to contrast-enhanced imaging: diffusion-weighted imaging and T1-weighted imaging identifies and quantifies necrosis in Wilms tumour. *Eur Radiol.* 2019;29(8):4141-9.
330. Chen T, Guestrin C, editors. Xgboost: A scalable tree boosting system. *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*; 2016.
331. Costanza R. Model goodness of fit: a multiple resolution procedure. *Ecological modelling.* 1989;47(3-4):199-215.
332. Kuhnert M, Voinov A, Seppelt R. Comparing raster map comparison algorithms for spatial modeling and analysis. *Photogrammetric Engineering & Remote Sensing.* 2005;71(8):975-84.
333. Hirose O. A Bayesian Formulation of Coherent Point Drift. *IEEE Transactions on Pattern Analysis and Machine Intelligence.* 2020.
334. Simonovsky M, Gutiérrez-Becker B, Mateus D, Navab N, Komodakis N, editors. A deep metric for multimodal registration. *International conference on medical image computing and computer-assisted intervention*; 2016: Springer.
335. Cheng X, Zhang L, Zheng Y. Deep similarity learning for multimodal medical images. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization.* 2018;6(3):248-52.
336. Raza SEA, AbdulJabbar K, Jamal-Hanjani M, Veeriah S, Le Quesne J, Swanton C, et al., editors. Deconvolving Convolutional Neural Network for Cell Detection. *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*; 2019: IEEE.

## Publications arising from the work presented in this thesis and collaborations

### Scientific journals

---

Konstantinos Zormpas-Petridis, Rosa Noguera, Daniela Kolarevic Ivankovic, Ioannis Roxanis, Yann Jamin, and Yinyin Yuan. “**SuperHistopath: A Deep Learning Pipeline for Mapping Tumour Heterogeneity on Low-Resolution Whole-Slide Digital Histology Images**”. Under review at *Frontiers in Oncology*.

Konstantinos Zormpas-Petridis, Evon Poon, Matthew Clarke, Neil Jerome, Jessica Bault, Matthew Blackledge, Fernando Carceller, Alexander Koers, Giuseppe Barone, Andrew Pearson, Lucas Moreno, John Anderson, Neil Sebire, Kieran McHugh, Dow-Mu Koh, Louis Chesler, Yinyin Yuan, Simon Robinson, and Yann Jamin. “**Non-invasive MRI native T1-mapping detects response to MYCN-targeted therapies in the Th-MYCN model of neuroblastoma**”. *Cancer Research*, 2020.

Konstantinos Zormpas-Petridis\*, Henrik Failmezger\*, Shan E Ahmed Raza, Ioannis Roxanis, Yann Jamin and Yinyin Yuan. “**Superpixel-based Conditional Random Fields (SuperCRF): Incorporating global and local context for enhanced deep learning in melanoma histopathology**”. *Frontiers in Oncology*, 2019. \* Joint first author.

Konstantinos Zormpas-Petridis, Neil Jerome, Matthew Blackledge, Fernando Carceller, Evon Poon, Matthew Clarke, Ciara McErlean, Giuseppe Barone, Alexander Koers, Sucheta Vaidya, Lynley Marshall, Andrew Pearson, Lucas Moreno, John Anderson, Neil Sebire, Kieran McHugh, Dow-Mu Koh, Yinyin Yuan, Louis Chesler, Simon Robinson, and Yann Jamin. “**MRI Imaging of the Hemodynamic Vasculature of Neuroblastoma Predicts Response to Anti-angiogenic Treatment**”. *Cancer Research*, 2019.

***From collaborations:***

Konstantinos Zormpas-Petridis, Nina Tunariu, Andra Curcean, Christina Messiou, Sebastian Curcean, David J Collins, Julie Hughes, Yann Jamin, Dow-Mu Koh, and Matthew D. Blackledge. "**quickDWI: Accelerating whole-body diffusion-weighted MRI with Artificial Intelligence.**" Under Revision at *Radiology*.

Jin Li\*, Konstantinos Zormpas-Petridis\*, Jessica K.R. Boulton, Emma Reeves, Andreas Heindl, Maria Vinci, Filipa Lopes, Craig Cummings, Caroline J. Springer, Louis Chesler, Chris Jones, Jeffrey C. Bamber, Yinyin Yuan, Ralph Sinkus, Yann Jamin<sup>†</sup> and Simon P. Robinson<sup>†</sup>. "**Investigating the Contribution of Collagen to the Tumour Biomechanical Phenotype with Non-invasive Magnetic Resonance Elastography**". *Cancer Research*, 2019. \* Joint first author.

D. K. Hill, A. Heindl, K. Zormpas-Petridis, D. J. Collins, L. R. Euceda, D. N. Rodrigues, S. A. Moestue, Y. Jamin, D. M. Koh, Y. Yuan, T. F. Bathen, M. O. Leach, and M. D. Blackledge. "**Non-Invasive Prostate Cancer Characterization with Diffusion-Weighted MRI: Insight from In silico Studies of a Transgenic Mouse Model**". *Frontiers in Oncology*, 2017.

**Conference Papers**

---

Konstantinos Zormpas-Petridis\*, Henrik Failmezger\*, Ioannis Roxanis, Matthew Blackledge, Yann Jamin, Yinyin Yuan. "**Capturing Global Spatial Context for Accurate Cell Classification in Skin Cancer Histology**". Computational Pathology and Ophthalmic Image Analysis, pp52-60, *COMPAY MICCAI 2018*, Granada, Spain. \* Joint first author.

## Awards

---

*Summa Cum Laude* award for the “**Accelerating clinical diffusion-weighted MRI using deep-learning: Potential utility in metastatic prostate cancer and malignant mesothelioma**” abstract, ISMRM 28<sup>th</sup> Virtual Conference & Exhibition, 18<sup>th</sup> – 23rd April, 2020.

*Magna Cum Laude* award for “**Evaluating the sensitivity of ADC to childhood neuroblastoma pathology using gaussian mixture modeling and computational pathology *in vivo***” abstract, ISMRM 28<sup>th</sup> Virtual Conference & Exhibition, 18<sup>th</sup> – 23rd April, 2020.

*Magna Cum Laude* award for the “**T<sub>1</sub> mapping of neuroblastoma pathology: insight from a computational pathology study in the Th-MYCN transgenic mouse model**” abstract, ISMRM 27<sup>th</sup> Annual Meeting & Exhibition, Montreal, 11<sup>th</sup> – 16<sup>th</sup> May, 2019

*Magna Cum Laude* award for the “**Susceptibility contrast-MRI predicts response to the vascular endothelial growth factor receptor inhibitor cediranib in the Th-MYCN model of neuroblastoma**” abstract, ISMRM 26<sup>th</sup> Annual Meeting & Exhibition, Paris, 16<sup>th</sup> – 21<sup>st</sup> June, 2018

### Patent submission (collaboration)

---

**Patent;** Inventors: Matthew Blackledge and Konstantinos Zormpas-Petridis.  
Patent Application No: 1913481.6; Ref: AIMRIAC, Mewburn ref: 7552821.  
“DIFFUSION-WEIGHTED MAGNETIC RESONANCE IMAGING”.

For “quickDWI: Accelerating whole-body diffusion-weighted MRI with Artificial Intelligence”

## Peer-Reviewed Abstracts

---

Konstantinos Zormpas-Petridis, Matthew D. Blackledge, Louis Chesler, Yinyin Yuan, Simon P. Robinson and Yann Jamin. “**Evaluating the sensitivity of ADC to childhood neuroblastoma pathology using gaussian mixture modeling and computational pathology *in vivo***”. Oral presentation and digital poster to the ISMRM 28<sup>th</sup> Virtual Conference & Exhibition, 18<sup>th</sup> – 23<sup>rd</sup> April, 2020

Konstantinos Zormpas-Petridis, Nina Tunariu, Andra Curcean, Christina Messiou, David Collins, Yann Jamin, Dow-Mu Koh, and Matthew D. Blackledge. “**Accelerating clinical diffusion-weighted MRI using deep-learning: Potential utility in metastatic prostate cancer and malignant mesothelioma**”. Oral presentation and digital poster to the ISMRM 28<sup>th</sup> Virtual Conference & Exhibition, 18<sup>th</sup> – 23<sup>rd</sup> April, 2020

Konstantinos Zormpas-Petridis, Nina Tunariu, Dow-Mu Koh, Yann Jamin, and Matthew D. Blackledge. “**Accelerating whole-body diffusion-weighted MRI with Artificial Intelligence**”. Oral presentation at the Radiological Society of North America (RSNA) 2019 Annual Meeting, Chicago, USA.

Konstantinos Zormpas-Petridis, Evon Poon, Matthew Clarke, Yinyin Yuan, Louis Chesler, Simon P. Robinson and Yann Jamin. “**Native T<sub>1</sub> mapping of neuroblastoma pathology – Non-invasive early detection of response to MYCN-targeted therapy in the Th-MYCN model**”. Poster at the 5<sup>th</sup> Neuroblastoma Research Symposium, 11<sup>th</sup> – 12<sup>th</sup> April, Cambridge, 2019

Konstantinos Zormpas-Petridis, Matthew D. Blackledge, Matthew Clarke, Louis Chesler, Yinyin Yuan, Simon P. Robinson and Yann Jamin. “**Native T<sub>1</sub> of neuroblastoma pathology in the Th-MYCN transgenic mouse model: insights from computational pathology**”. Oral presentation at the 27<sup>th</sup> Postgraduate Symposium of the British Chapter of ISMRM (BC-ISMRM), 5<sup>th</sup> April, Birmingham, 2019

Konstantinos Zormpas-Petridis\*, Henrik Failmezger\*, Ioannis Roxanis, Matthew Blackledge, Yann Jamin, Yinyin Yuan. “**Capturing global spatial context for accurate cell classification in skin cancer histology**”. Oral presentation at the internal ICR Informatics conference, 22<sup>nd</sup> March, 2019.

Konstantinos Zormpas-Petridis, Matthew D. Blackledge, Matthew Clarke, Louis Chesler, Yinyin Yuan, Simon P. Robinson and Yann Jamin. “**T<sub>1</sub> mapping of neuroblastoma pathology: insight from a computational pathology study in the Th-MYCN transgenic mouse model**”. Power pitch oral presentation and digital poster at the ISMRM 27<sup>th</sup> Annual Meeting & Exhibition, Montreal, 11<sup>th</sup> – 16<sup>th</sup> May, 2019

Konstantinos Zormpas-Petridis, Matthew D. Blackledge, Louis Chesler, Yinyin Yuan, Simon P. Robinson and Yann Jamin. “**Susceptibility contrast-MRI predicts response to the vascular endothelial growth factor receptor inhibitor cediranib in the Th-MYCN model of neuroblastoma**”. Oral presentation at the ISMRM 26<sup>th</sup> Annual Meeting & Exhibition, Paris, 16<sup>th</sup> – 21<sup>st</sup> June, 2018

Konstantinos Zormpas-Petridis, Matthew D. Blackledge, Louis Chesler, Yinyin Yuan, Simon P. Robinson and Yann Jamin. “**Spatial cross-evaluation of MRI with histology validate R<sub>2</sub>\* as a biomarker of vascular hemodynamics in the Th-MYCN model of neuroblastoma**”. E-poster at the ISMRM 26<sup>th</sup> Annual Meeting & Exhibition, Paris, 16<sup>th</sup> – 21<sup>st</sup> June, 2018

Jin Li, Konstantinos Zormpas-Petridis, Andreas Heindl, Jessica K.R. Boulton, Craig Cummings, Jeffrey C. Bamber, Yinyin Yuan, Ralph Sinkus, Yann Jamin, Simon P. Robinson. “**Collagen is a major determinant of the viscoelastic properties of stromal-dense tumours: insights from pre-clinical MRE**”. Oral presentation at the ISMRM 26<sup>th</sup> Annual Meeting & Exhibition, Paris, 16<sup>th</sup> – 21<sup>st</sup> June, 2018

Konstantinos Zormpas-Petridis, Matthew D. Blackledge, Louis Chesler, Yinyin Yuan, Simon P. Robinson and Yann Jamin. “**Susceptibility-weighted MRI provides predictive biomarkers of response to vascular endothelial growth factor receptor inhibition in the Th-MYCN model of neuroblastoma**”. Oral presentation at the 26<sup>th</sup> Postgraduate Symposium of the British Chapter of ISMRM (BC-ISMRM), May 18<sup>th</sup> London, 2018.

Konstantinos Zormpas-Petridis, Matthew D. Blackledge, Evon Poon, Matthew Clarke, Neil P. Jerome, Alexander Koers, Chris Jones, Dow-Mu Koh, Andrew D.J. Pearson, Lucas Moreno, Yinyin Yuan, Louis Chesler, Simon P. Robinson and Yann Jamin. “**Susceptibility-weighted MRI predicts response to vascular endothelial growth factor receptor inhibition in the Th-MYCN transgenic model of neuroblastoma**”. Poster at the ANR 2018 conference, San Francisco, 2018.

K. Zormpas-Petridis, D.K. Hill, A. Heindl, Y. Yuan, D.J Collins, D.M. Koh, T.F. Bathen, M.O. Leach, Y. Jamin, M.D. Blackledge. “**Luminal space correlates with ADC in a transgenic model of prostate cancer**”. Oral presentation at the 25<sup>th</sup> Postgraduate Symposium of the British Chapter of ISMRM, London 2017.

## Appendix

### Additional Materials and Methods of Chapter 4

#### *Animals*

Transgenic Th-*MYCN* mice were genotyped to detect the presence of the human *MYCN* transgene (18). The study was performed using both male and female hemizygous mice, which developed palpable tumours at 50 to 130 days with a 25% penetrance. Tumour development was monitored weekly by palpation by an experienced animal technician. Mice with palpable tumours were then enrolled (day 0) and their tumour volume was subsequently monitored by MRI. A total of 68 mice were enrolled with a median tumour volume of  $801 \pm 63$  mm<sup>3</sup> (median  $\pm$  1 s.e.m., ranging from 143 to 2055 mm<sup>3</sup>). Mice were housed in specific pathogen-free rooms in autoclaved, aseptic microisolator cages (maximum of four mice per cage).

#### *Preclinical study design*

**Study 1.** We first evaluated the sensitivity of MRI to cediranib-induced acute modulation of neuroblastoma vasculature. IS- and SC-MRI were performed prior to (day 0) and 24 hours after treatment started (day 2). Mice were treated on day 1 with 6 mg/kg of cediranib orally (obtained under material transfer agreement with AstraZeneca, n = 10) or vehicle (n = 8). One mouse was excluded due to failed remote contrast injection. The 24-hour imaging timepoint was chosen based on *i*) preliminary evidence that cediranib does not elicit any significant volume reduction at this timepoint in the highly chemosensitive Th-*MYCN* model (12) and *ii*) cediranib caused significant reductions in DCE-MRI parameters at this timepoint in the adult phase I clinical trial (19).

**Study 2.** In an additional cohort (n = 12) we further evaluated the effect of sustained daily treatment, with IS- and SC-MRI performed prior to, 24 hours and

7 days after daily treatment with 6 mg/kg cediranib. MRI data were not collected from two mice at the 24-hour timepoint.

**Study 3.** Guided by the results of Study 2, we subsequently acquired IS-MRI data from additional Th-*MYCN* mice (n = 25) prior to daily treatment with cediranib for 7 days (bringing the total number of mice from which pre-treatment  $R_2^*$  data were acquired to 37). The volumetric response to cediranib over 7 days treatment was monitored by  $T_2$ -weighted MRI only, and compared with that from mice treated daily with vehicle (n = 12).

### *MRI*

All MRI studies were performed on a 7T Bruker horizontal bore MicroImaging system (Bruker Instruments) using a 3 cm birdcage volume coil. Anaesthesia was induced by an intraperitoneal 5 mL/kg injection of a combination of fentanyl citrate (0.315 mg/mL) plus fluanisone (10 mg/mL; Hypnorm, Janssen Pharmaceutical) and midazolam (5 mg/mL; Roche) and water (1:1:2). A lateral tail vein was cannulated with a 27G butterfly catheter (Hospira) for remote administration of USPIO particles. Core temperature was maintained at approximately 37 °C with warm air blown through the magnet bore.

For all the mice, anatomical  $T_2$ -weighted transverse images were acquired from 20 contiguous 1-mm-thick slices through the mouse abdomen, using a rapid acquisition with refocused echoes (RARE) sequence with four averages of 128 phase encoding steps over a 3x3 cm field of view, an echo time (TE) of 36 milliseconds, a repetition time (TR) of 4.5 seconds and a RARE factor of 8. These images were used to determine tumour volumes, and for planning the subsequent functional MRI measurements, which included optimization of the local field homogeneity using FASTmap algorithm and the measurement of the baseline transverse relaxation rate  $R_2^*$  ( $\text{second}^{-1}$ ), which is sensitive to the concentration of paramagnetic species, principally deoxyhaemoglobin.  $R_2^*$  was quantified using a multiple gradient-recall echo (MGE) sequence with eight averages, and an acquisition time of 3 minutes 20 seconds. Images were

acquired using eight echoes spaced 3 milliseconds apart, an initial TE of 6 milliseconds, a flip angle  $\alpha = 45^\circ$  and a TR of 200 milliseconds. A dose of 150  $\mu\text{mol Fe/kg}$  of the USPIO particle preparation P904 (overall particle size  $\sim 25\text{--}30$  nm diameter; Guerbet) was then administered intravenously. After 3 minutes to allow for equilibration, a second set of identical MGRE images acquired.

All the MRI data were acquired with a matrix size of  $128 \times 128$  over a  $3 \times 3$  cm field of view. Tumour volumes were determined using segmentation from regions of interest drawn on  $T_2$ -weighted images for each tumour-containing slice using OsiriX. Tumour  $R_2^*$  maps were calculated from regions of interest drawn for each tumour-containing slice from the MGRE images acquired prior to and following administration of USPIO particles by fitting a single exponential to the signal intensity echo time curve on a voxel-by-voxel basis using a robust Bayesian approach using in-house software (ImageView, developed in IDL; ITT Visual Information Systems). Parametric maps of tumour  $fBV$  (%) were subsequently calculated using the USPIO-induced change in  $R_2^*$  ( $\Delta R_2^*$ ), as described previously (20).

## **Additional Materials and Methods of Chapters 5-7**

### *Animals, imaging, and drug treatment schedule*

Transgenic Th-*MYCN* mice were genotyped to detect the presence of the human *MYCN* transgene (7). The study was performed using both male and female homozygous mice, which developed a single palpable abdominal tumour at 40–80 days old with 100% penetrance. Tumour development was monitored weekly by palpation by an experienced animal technician. A total of 46 mice were enrolled with a median tumour volume of  $861 \pm 86 \text{ mm}^3$  (derived from  $T_2$ -weighted MRI; median  $\pm 1$  s.e.m., ranging from 280 to  $2557 \text{ mm}^3$ ). MRI was performed prior to treatment (day 0). Mice were left to recover for 24 hours, and then treated (Day 1) with 30 mg/kg orally of alisertib (MLN8237, purchased from Selleckchem,  $n = 11$ ) or vehicle (10% 2-hydroxypropyl  $\beta$ -cyclodextrin, 1%  $\text{NaHCO}_3$ ,  $n = 9$ ), or 25 mg/kg orally of vistusertib (AZD2014, obtained under material transfer

agreement with AstraZeneca, n = 14) or vehicle (5% DMSO, 95% PEG300, n = 12). Post-treatment MRI was performed 24 hours after treatment started (Day 2). Mice were housed in specific pathogen-free rooms in autoclaved, aseptic microisolator cages (maximum of 4 mice per cage) and allowed access to sterile food and water *ad libitum*.

## *MRI*

All MRI studies were performed on a 7T Bruker horizontal bore MicroImaging system (Bruker Instruments) using a 3 cm birdcage volume coil. Anaesthesia was induced by an intraperitoneal 5 mL/kg injection of a combination of fentanyl citrate (0.315 mg/mL) plus fluanisone (10 mg/mL; Hypnorm, Janssen Pharmaceutical) and midazolam (5 mg/mL; Roche) and water (1:1:2). Core temperature was maintained at approximately 37 °C with warm air blown through the magnet bore.

For all the mice, contiguous anatomical T<sub>2</sub>-weighted transverse images were acquired through the mouse abdomen for the quantification of tumour volume, optimization of the local field homogeneity using the FASTmap algorithm, and for planning the subsequent multiparametric MRI measurements. In addition to IR-TrueFISP MRI for quantification of the spin-lattice (T<sub>1</sub>) and spin-spin (T<sub>2</sub>) relaxation times, these also included measurement of the apparent diffusion coefficient (ADC), the transverse relaxation rate R<sub>2</sub><sup>\*</sup> and the magnetization transfer ratio (MTR) using the MRI sequences and parameters listed in Appendix Table A1.

Tumour volumes were determined using segmentation from regions of interest drawn on each tumour-containing T<sub>2</sub>-weighted MRI slice using OsiriX. All the multiparametric MRI data were fitted voxelwise using in-house software (ImageView, working under IDL, ITT) with a robust Bayesian approach that provided estimates of T<sub>1</sub>, T<sub>2</sub>, ADC and R<sub>2</sub><sup>\*</sup>. MTR (%) was calculated as  $MTR = (1 - M_{25ppm} / M_{100ppm}) \times 100$  and fitted voxelwise using in-house code written in Matlab (The Mathworks).

Sequence	Parameters	TE (ms)	TR (ms)	NS	FOV (cm <sup>2</sup> )	Matrix size	Slices	Slice thickness	Sequence specific parameters
<b>T<sub>2</sub>-weighted RARE</b>	Tumour volume (mm <sup>3</sup> )	36	4500	4	3x3	128x128	20	1 mm	RARE factor = 8, TT=3min40s
<b>IR-TrueFISP</b>	T <sub>1</sub> (ms), T <sub>2</sub> (ms)	1.2	2.5	8	3x3	128x128	1	1 mm	8 segments, scan TR= 10s, 50 TI = 28-1930ms), TT=10min 40s
<b>Diffusion-weighted MRI</b>	ADC (10 <sup>-6</sup> mm <sup>2</sup> .s <sup>-1</sup> )	32	1500	4	3x3	128x128	3	1 mm	EPI readout 5b-values = 200 to 1000 s.mm <sup>-2</sup> , TT= 4min
<b>MGRE</b>	R <sub>2</sub> * (s <sup>-1</sup> )	6	200	4	3x3	128x128	3	1 mm	8 echoes, 3ms apart, TT= 3min 20s
<b>MT-RARE</b>	MTR (%)	7.5	1600	4	3x3	128x128	1	1 mm	RARE factor=8, Saturation pulse length= 1.3s, strength B <sub>1</sub> = 8 μT, offset frequency= +25ppm (MT effects "on") and +100ppm (MT effects "off"), TT= 1min 36s

**Table A.1** Summary of the MRI sequences and parameters used in this study