A quantitative study of the metastatic process through mathematical modeling

Etienne Baratchart

Etienne Baratchart. A quantitative study of the metastatic process through mathematical modeling. Computation [stat.CO]. Université de Bordeaux, 2016. English. <NNT : 2016BORD0023>. <tel-01314117>

HAL Id: tel-01314117
https://tel.archives-ouvertes.fr/tel-01314117
Submitted on 10 May 2016

Thèse

présentée pour obtenir le grade de

DOCTEUR DE
L'UNIVERSITÉ DE BORDEAUX

École doctorale de Mathématiques et Informatique

Spécialité :

MATHÉMATIQUES APPLIQUÉES ET CALCUL SCIENTIFIQUE

Présentée par

ETIENNE BARATCHART

Sous la direction de SÉBASTIEN BENZEKRY, THIERRY COLIN ET OLIVIER SAUT

Titre :

ÉTUDE QUANTITATIVE DES ASPECTS DYNAMIQUES ET SPATIAUX DU DÉVELOPPEMENT MÉTASTATIQUE À L'AIDE DE MODÈLES MATHÉMATIQUES

Soutenue le 05 Février 2016

JURY

Sébastien Benzekry	INRIA	Encadrant
Andreas Bikfalvi	Université de Bordeaux/INSERM U1029	Invité
Thierry Colin	Institut polytechnique de Bordeaux	Directeur de thèse
Marie Doumic Jauffret	INRIA	Rapportrice
Anne Gégout-Petit	Université de Lorraine	Rapportrice (présidente)
Olivier Saut	CNRS	Directeur de thèse
Angélique Stéphanou	CNRS	Examinatrice

Résumé : L'objet de cette thèse est l'étude du processus métastatique par la confrontation de données *in vivo* chez la souris avec des modèles mathématiques. Plus précisément, des données longitudinales sur la masse métastatique totale combinées à des données IRM fournissant des informations sur le nombre et la taille des macrométastases ont été confrontées à un modèle décrivant l'évolution de la distribution en tailles des métastases par une équation aux dérivées partielles de populations structurées. La théorie sous-jacente au modèle, décrivant le processus métastatique par des métastases initiées par quelques cellules et croissant indépendamment les unes des autres, s'est révélée incapable de décrire les distributions de tailles métastatiques observées à l'IRM, suggérant la présence de phénomènes non pris en compte dans la théorie "standard" du développement métastatique. Ces résultats nous ont conduit à proposer des hypothèses expliquant les différences de distributions métastatiques entre le modèle et les données. Ces hypothèses ont été étudiées expérimentalement par nos collaborateurs biologistes mais également *in silico* à l'aide de modèles d'équations aux dérivées partielles décrivant la croissance de plusieurs métastases pouvant interagir spatialement. Les résultats obtenus à l'aide de notre approche de modélisation suggèrent des interactions jouant un rôle important dans la dynamique métastatique, comme l'agrégation de germes métastatiques ou l'attraction de cellules métastatiques par des foyers métastatiques déjà existants. Une partie de cette thèse est également dédiée à l'analyse mathématique et numérique du nouveau modèle spatial introduit pour l'étude quantitative précédemment évoquée. Ce modèle mécanique décrit notamment l'effet de la pression sur la prolifération des cellules tumorales. Des résultats de convergence de la méthode numérique utilisée sont présentés, ainsi qu'une confrontation du modèle à des données de croissance de métastases pulmonaires. Enfin, une partie traitant des interactions métastases-microenvironnement est également présentée. Des études récentes ont en effet montré que certaines cellules progénitrices de la lignée hématopoïétique ou encore certaines cellules immunitaires pourraient jouer un rôle important dans le développement métastatique. Au cours de cette thèse, ce phénomène appelé niche prémétastatique a été étudié dans la littérature biologique puis modélisé mathématiquement afin de mieux comprendre le rôle de cette niche dans la dynamique métastatique.

Abstract: In this thesis, a quantitative study of the metastatic process in the mouse has been performed thanks to mathematical modeling. Precisely, longitudinal data of the total metastatic burden and MRI data on the macrometastatic size distribution are confronted to a mathematical model describing the metastatic process by the independent growths of metastatic foci starting from one or few cells. This "standard" theory, able to describe the dynamics of the total metastatic burden, is on the other hand unable to describe the observed metastatic size distributions. Indeed, this model predicts many small metastases, whereas the observed metastases are much larger and fewer. In order to explain these differences, we proposed two hypotheses that were not taken into account in the initial theory. In the first one, metastases that are growing in close vicinity could merge, resulting in one larger metastasis. In the second one, metastatic foci could attract arriving circulating tumor cells, resulting also in fewer foci but much larger ones. These hypotheses have been tested experimentally by our biologists collaborators, and *in silico* thanks to a spatial model of tumor growth. The results of this study show that the previously suggested phenomena could have a substantial impact on the number and the sizes of the metastatic foci during metastatic development. Another part of this thesis is devoted to the numerical and mathematical analysis of the previous spatial model. This model takes into account the effect of the pressure on the proliferation of tumor cells. Numerical convergence of the numerical method that has been used and data assimilation on imaging data of pulmonary metastases are presented. Finally, a last part deals with the interactions between metastasis and its supportive stroma. Recent studies shed light on the implication of hematopoietic progenitors in the formation of a permissive soil in the future metastatic site, a phenomenon so-called premetastatic niche. In this thesis, a mathematical model describing the premetastatic and metastatic dynamics is proposed to study quantitative aspects of this phenomenon.

Remerciements

Je tiens en premier lieu à remercier mes directeurs de thèse et encadrant Thierry Colin, Olivier Saut et Sébastien Benzekry pour leurs qualités humaines et scientifiques qui m'ont permis d'apprendre en permanence et d'apprécier les trois années passées au sein de l'équipe INRIA MONC. Thierry Colin pour sa bonne humeur quotidienne contribuant largement à l'ambiance très agréable régnant dans l'équipe, ainsi que pour son approche des sciences très stimulante et ponctuée d'humour, comme en témoignent ces sessions de travail assaisonnées de blagues vaseuses mais souvent très drôles. Olivier Saut pour sa gentillesse, son humour plus feutré, ses compétences et sa patience lorsque je venais l'assaillir de problèmes de code. Mention spéciale pour Sébastien Benzekry, un ami et brillant scientifique qui m'a transmis le goût de la recherche et sa passion pour les sciences appliquées, et à qui je dois beaucoup, tant humainement que scientifiquement.

Je remercie également Anne Gégout-Petit ainsi que Marie Doumic-Jauffret pour avoir pris le temps de lire ma thèse, pour leurs remarques fort utiles qui m'ont permis d'améliorer la qualité du manuscrit ainsi que pour leurs commentaires positifs et encourageants.
Merci également à Angélique Stéphanou, dont j'ai apprécié le contact en conférence et qui m'honore par sa présence dans mon jury de thèse.

Merci bien sûr à Andreas Bikfalvi d'avoir supervisé le projet de ma thèse et d'en avoir permis le financement.

Mes remerciements vont également à Lin et Wiwi pour les mini-cours de biologie improvisés en toute détente devant le tableau, ainsi que pour leur travail expérimental considérable et indispensable à la pertinence scientifique de cette thèse.

Je tiens aussi remercier Emeline Ribot pour sa gentillesse, pour le temps qu'elle a pris à nous expliquer le fonctionnement des séquences IRM et ce qu'on voyait aux images, pour m'avoir permis toutes ces sessions contourage pour la collecte de données, et surtout pour sa contribution à l'obtention de données imagerie ayant joué un rôle très important pour les résultats de cette thèse.
Merci également à Sylvain Miraux pour avoir contribué à l'obtention des données IRM, ainsi que pour m'avoir permis de participer à un stage IRM fort intéressant et enrichissant.

Je tiens également à remercier les autres permanents de l'équipe : Clair Poignard pour sa gentillesse, sa disponibilité, et pour les pistes qu'ils m'a données sur la partie mathématique et numérique de cette thèse, ainsi qu'Annabelle Collin qui n'est arrivée que récemment mais qui m'a donné de précieux conseils.

Mes remerciements vont également à tous les doctorants, postdoc et ingénieurs du couloir, qui m'ont fait passer trois années très sympa, que ce soit pour les discussions boulot ou les sorties tardives et bien arrosées en ville, et désolé par avance si j'oublie quelques personnes : Florian, Marco, Thibaut (Morray), Olivier, les Guillaume, Perrine, Benjamin, Federico, Andrea, Thomas, Cynthia, Marie, Alice, Vivien, Hervé, Agathe, Manon, ...

Merci également aux assistantes d'équipe inria Anne-Laure, Flavie et Sylvie pour leur gentillesse, leur disponibilité et leur aide.

Je n'oublie pas mes amis, dont le contact a été précieux au cours de ces trois années : Laurent, Marc, Julien, Arnaud, JB, Greg, Romain, Mathieu, Félix, Francesco, Jordan, Marie-Laetitia, Antoine et à tous ceux que je ne cite pas mais à qui je pense également.

Merci également à toute ma famille, mon grand-père, mes cousins, oncles et tantes.

Comment ne pas remercier mes parents qui m'ont toujours apporté leur soutien dans mes études et mes projets personnels, et à qui je dois énormément, ainsi qu'à ma soeur à qui je pense et que je ne vois que trop peu.

Enfin, un énorme merci à toi Gwen, pour ce que tu es et tout ce que tu m'as apporté pendant ces huit années.

Contents

Chapter 1

Introduction

Le cancer est la deuxième cause de mortalité au monde derrière les maladies cardiovasculaires [2]. Si le cancer du sein est le plus fréquent chez la femme et le cancer de la prostate le plus fréquent chez l'homme, le cancer causant le plus grand nombre de décès est le cancer du poumon [3]. Le premier stade du cancer consiste en une tumeur localisée dans un organe/tissu. Une tumeur, ou neoplasie (nouveau tissu) se développe dans un organe lorsque certaines cellules de cet organe échappent aux régulations de l'organisme et dérèglent leurs mécanismes de régulation intracellulaire pour proliférer anormalement. Cette prolifération anormale résulte de mutations acquises par les cellules, affectant des gènes ayant un rôle critique dans le contrôle de la prolifération et du comportement des cellules [4]. Ces cellules se multiplient alors de manière incontrôlée, ayant pour conséquence la formation d'une masse de cellules appelée tumeur. Fort heureusement, la très grande majorité des tumeurs demeurent bénignes et se limitent au stade de pseudotumeur. Ces pseudotumeurs restent très localisées au sein de l'organe et sont rapidement éliminées par le système immunitaire. Cependant, dans de rares cas, lorsque certaines mutations s'accumulent chez les cellules, elles induisent un comportement malin de la tumeur. Une tumeur maligne peut devenir invasive, c'est-à-dire envahir les tissus voisins et *in fine* relâcher des cellules dans le système vasculaire pouvant former des tumeurs secondaires dans des organes distants. Ces tumeurs secondaires sont appelées métastases et sont la cause majoritaire (90% des cas) de décès chez les patients atteints d'un cancer [5].

De nombreux traitements ont vu le jour pour lutter contre le cancer et sont utilisés dans différentes situations : chirurgie, chimiothérapies, thérapies ciblées, radiothérapie, etc. Certains de ces traitements ont prouvé leur efficacité (chirurgie pour certaines tumeurs localisées comme la prostate, imatinib pour la leucémie myeloide chronique, etc) en clinique, permettant aujourd'hui de soigner un cancer sur deux en France [6]. La compréhension de la biologie du cancer a permis la mise en place de traitements ciblant par exemple directement les cellules tumorales en prolifération (radiothérapies et chimiothérapies par exemple) ou encore d'autres thérapies ciblant certains mécanismes (protéines extracellulaires, mécanismes intracellulaires, expression de certains gènes) biologiques permettant le phénotype malin des cellules tumorales. Cependant, dans bien des cas, les traitements disponibles restent insuffisants. En effet, pour certains types de cancer (foie, pancréas, poumon), le taux de survie à 10 ans reste inférieur à 10% [7]. Une meilleure compréhension de la biologie du cancer est nécessaire pour améliorer ces traitements. L'intérêt croissant pour les aspects dynamiques de la maladie et les progrès expérimentaux permettant le développement de données quantitatives font de la modélisation mathématique un outil potentiellement utile pour la compréhension de la biologie du cancer ou pour le choix de stratégies thérapeutiques grâce à des prédictions quantitatives dans les cadres précliniques et cliniques.

Comme nous l'avons mentionné, les métastases, tumeurs secondaires se formant dans des organes distants de la tumeur primaire, représentent la première cause de mortalité chez les patients. Paradoxalement, jusque dans les années 2000, relativement peu d'efforts de recherche ont été réalisés pour comprendre la biologie du processus métastatique, les recherches s'étant focalisées sur la compréhension de la tumorigénèse [5]. Le développement métastatique est donc un phénomène encore mal compris, bien qu'il fasse l'objet de recherches de plus en plus nombreuses, dans le but d'une part d'approfondir nos connaissances de cet aspect du cancer afin d'améliorer le diagnostic et de mieux anticiper l'évolution de la maladie chez le patient, et d'autre part de développer des thérapies efficaces contre la dissémination et la croissance des métastases. Lors de la première étape du processus métastatique, des cellules tumorales parviennent à se détacher de la tumeur primaire, notamment par le biais d'un mécanisme appelé "epithelial to mesenchymal transition" (EMT). Au cours de l'EMT, les cellules changent de phénotype, passant du stade de cellules spécialisées et peu mobiles à un stade moins différencié leur conférant une mobilité accrue et la capacité d'exprimer des facteurs leur permettant d'interagir avec le microenvironment. Ces cellules moins différenciées sont parfois capables d'intravaser, c'est-à-dire de pénétrer dans un vaisseau sanguin, et une fois piégées dans les capillaires d'un organe, d'extravaser et d'établir des tumeurs secondaires dans ce nouvel organe. Pour cela, les cellules doivent s'adapter à leur nouvel environnement et établir une niche viable afin d'assurer leur survie et leur prolifération [5]. L'étape comprenant le détachement des cellules de la tumeur primaire, le transport dans le réseau vasculaire et l'arrivée dans l'organe distant, aussi appelée dissémination métastatique, commence à être relativement bien comprise, notamment grâce aux travaux de recherches sur l'EMT [8]. En revanche, l'étape d'adaptation des cellules tumorales au nouveau "sol", à savoir le microenvironnement de l'organe distant, permettant la survie et la prolifération des cellules et *in fine* la colonisation métastatique de l'organe, est encore très mal comprise [9]. L'émergence de théories biologiques traitant des aspects dynamiques du processus (niche prémétastatique [10], *self-seeding* [11]), ainsi que les techniques expérimentales (bioluminescence, marquage aux protéines fluorescentes, techniques d'imagerie) permettant l'obtention de données quantitatives sur la masse métastatique ou encore le nombre et la taille des lésions rendent intéressante l'utilisation de modèles mathématiques pour l'analyse de la dynamique du processus métastatique. Plusieurs aspects de ce processus biologique ont été traités au cours de cet thèse à l'aide de différentes approches et techniques de modélisation.

Le plan de cette thèse est le suivant.

La partie I est consacrée à un état de l'art non exhaustif de la biologie du cancer et de l'usage des modèles mathématiques dans la recherche contre le cancer. Elle aborde également un certain nombre de rappels sur l'assimilation de données et l'estimation statistique de paramètres.

La partie II traite de modèles mathématiques pour la croissance tumorale. Dans un premier temps, nous présentons des modèles dits phénoménologiques et modélisant la croissance tumorale à l'aide d'équations différentielles ordinaires (EDO) du type:

$$\frac{\mathrm{d}V}{\mathrm{d}t} = g(V),$$

où $g(V)$ désigne la loi de croissance pouvant être linéaire en V (modèle exponentiel), logistique, Gompertz, etc. Ces modèles sont confrontés à des données de croissance de tumeur du rein chez la souris issues de deux techniques expérimentales différentes, fournissant ainsi des informations différentes, à savoir le volume tumoral d'une part et le nombre de cellules tumorales (tracées à l'aide d'une protéine fluorescente) d'autre part. Les résultats semblent indiquer que les dynamiques du volume tumoral total et du nombre de cellules tumorales sont différentes. Le stroma, tissu de soutien de la tumeur et composé de cellules non tumorales, pourrait avoir sa propre dynamique, en interaction avec les cellules tumorales, induisant des différences de dynamique entre cellules tumorales et volume tumoral.
Cette partie aborde ensuite des modèles plus mécanistiques et décrivant l'expansion spatiale d'une tumeur. Dans ce type de modèle 2D décrivant les tissus comme une mixture fluide, la dynamique spatio-temporelle des densités de tissu tumoral P et sain S est gouvernée par des équations de conservation, avec une hypothèse de saturation du milieu (densité totale de tissu constante) et une vitesse

de Darcy (gradient de pression) $\mathrm{v} = -k\nabla\Pi$ où Π est la pression du milieu:

$$\begin{aligned}
&\frac{\partial P(t,x)}{\partial t} + \nabla \cdot (\mathrm{v}(t,x)P(t,x)) = \gamma(\Pi)P(t,x),\\
&\frac{\partial S(t,x)}{\partial t} + \nabla \cdot (\mathrm{v}(t,x)S(t,x)) = 0,\\
&\nabla \cdot \mathrm{v}(t,x) = \gamma(\Pi)P(t,x),\\
&\mathrm{v} = -k\nabla\Pi.
\end{aligned} \tag{1.1}$$

Ce type de modèle, introduit entre autres par Preziosi et Ambrosi [12], a récemment été utilisé pour prédire la croissance de métastases pulmonaires chez le patient [13]. Afin de décrire la croissance tumorale de métastases pulmonaires chez la souris, nous avons proposé une loi $\gamma(\Pi)$ où la prolifération des cellules tumorales dépend de la pression de la manière suivante :

$$\gamma(\Pi) = \gamma_0 \exp\left(-\frac{\Pi}{\Pi_c}\right).$$

La prolifération des cellules tumorales est ainsi inhibée lorsque la pression exercée sur elles augmente, une propriété déjà observée dans des études expérimentales [14, 15, 16]. L'introduction de cette loi permet d'obtenir un modèle minimalement paramétré (deux paramètres), proposant une explication mécanique au fait qu'une tumeur croît de moins en moins vite à mesure que son volume augmente, et capable de décrire la croissance de lésions métastatiques pulmonaires chez la souris. Cette loi de prolifération induit dans le modèle une équation elliptique non linéaire sur la pression:

$$\begin{aligned}
-\,k\Delta\Pi &= \gamma_0 \exp\left(-\frac{\Pi}{\Pi_c}\right)P,\\
\Pi_{|\partial\Omega} &= 0.
\end{aligned} \tag{1.2}$$

Après une analyse mathématique et numérique, ce modèle est confronté à des données spatiales (à partir d'imagerie par résonance magnétique) de croissance de métastases pulmonaires afin d'étudier les capacités descriptives du modèle en terme de dynamique du volume et de la forme des lésions.
Pour ce qui est de l'analyse du modèle, une preuve d'existence et d'unicité de solution à l'équation (1.2) pour un domaine Ω Lipschitz et un champ de densité tumorale dans $L^\infty(\Omega)$, ainsi qu'une certaine régularité de cette solution sont présentées. Un algorithme de point fixe est proposé pour la résolution numérique, et sa convergence géométrique dans $H^{3/2}(\Omega)$ est démontrée. Des tests de convergence numérique du système couplé (1.1) ont également été réalisés, dans un premier temps pour le cas simple où la loi de prolifération γ est constante, permettant d'exhiber une solution analytique en géométrie sphérique. Ce cas test a permis de comparer les performances de plusieurs schémas de transport et de choisir le plus efficace en terme de ratio précision sur temps de calcul. Des tests de convergence ont ensuite été réalisés dans le cas où le taux de prolifération dépend de la pression. Enfin, une section présentant une méthodologie et des résultats d'assimilation de données par le modèle est présentée. Il s'agit de données de croissance de métastases pulmonaires. Une méthodologie permettant de borner l'espace des paramètres à partir des données est d'abord présentée, suivie de la calibration du modèle sur les données de dynamique du volume des métastases. Le modèle, capable de décrire la croissance de ces métastases, a ensuite été utilisé pour simuler l'évolution de la forme de ces métastases sous les hypothèses du modèle. Si certaines formes étaient correctement prédites, d'autres prédisaient une forme ne prenant pas en compte les anisotropies visibles dans les données d'imagerie, suggérant que les caractéristiques du milieu que peuvent fournir certaines techniques d'imagerie s'avèreraient utiles pour des prédictions spatiales et quantitatives de l'expansion tumorale.

La partie III, articulée en trois chapitres, contient le coeur de la thèse et concerne les aspects dynamiques du processus métastatique.

Le premier chapitre de cette partie est une étude quantitative des aspects dynamiques du processus métastatique à partir de données de cancer du rein chez la souris et à l'aide d'un modèle mathématique décrivant le processus métastatique à l'échelle de l'organisme. Disposant à la fois de données longitudinales en temps sur la masse métastatique totale (marquage des cellules tumorales avec une protéine fluorescente) dans les poumons et sur la distribution en taille des lésions métastatiques visibles (données d'imagerie par résonance magnétique) dans les poumons, nous avons dans un premier temps confronté la théorie classique du processus métastatique à ces données. Plus précisément, nous avons voulu vérifier si une théorie standard basée sur la dissémination de cellules métastatiques par la tumeur primaire et sur la prolifération de ces cellules donnant naissance à des lésions métastatiques dans l'organe distant était en accord ou non avec les deux types de données citées précédemment. Pour ce faire, nous avons utilisé une formalisation mathématique de cette théorie standard qui avait été effectuée dans un premier temps par Iwata et al. [17]. Ce modèle décrit une distribution continue de tailles métastatiques à l'aide d'une densité $\rho(t,v)$ de métastases structurée en volume de métastases v. Chaque métastase croît selon une loi de croissance du type:

$$\frac{\mathrm{d}v}{\mathrm{d}t} = g(t,v),$$

où $g(t,v)$ est une loi de croissance (de type Gompertz par exemple). La densité métastatique, quant à elle, est décrite par une équation de conservation:

$$\begin{cases} \partial_t\rho(t,v) + \partial_v(\rho(t,v)g(t,v)) = 0 & t \in]0,+\infty[,\ v \in]V_0,+\infty[, \\ g(V_0)\rho(t,V_0) = d(V_p(t)) & t \in]0,+\infty[, \\ \rho(0,v) = 0 & v \in]V_0,+\infty[, \end{cases} \tag{1.3}$$

où V_0 est le volume initial de chaque nouvelle métastase, en général une ou quelques cellules, d'après des observations expérimentales précédentes [18, 19]. $V_p(t)$ est le volume de la tumeur primaire et $d(V_p(t))$ est la loi d'émission de nouvelles métastases, dépendant ici du volume de la tumeur primaire. Ce modèle traduit la théorie classique de la dynamique métastatique, selon laquelle chaque nouvelle métastase naît d'une cellule métastatique survivant dans l'organe distant et où les métastases croissent indépendamment les unes des autres [20, 21, 22, 23, 24, 5]. Le modèle a été validé dans deux études récentes pour décrire la dynamique de la masse métastatique totale sur des données de biolominescence par Hartung et al. [25] sur des données de cancer du sein, et par Benzekry et al. sur des données de cancer du sein et du rein [26]. Cependant, ce modèle n'avait pas encore été confronté à des données sur la distribution en taille métastatique (à l'exception d'un patient dans le papier d'Iwata et al. [17]). Pour ce qui est de nos données (cancer du rein métastatique RENCA), si le modèle s'est révélé capable de décrire la dynamique de la masse métastatique totale (en accord avec les études précédemment mentionnées), il s'est en revanche avéré incapable de décrire le nombre et la taille des lésions métastatiques visibles. En effet, cette théorie standard prédit de trop nombreuses et trop petites lésions, comparé aux observations. Ceci nous a conduit à suggérer des hypothèses comme la fusion de foyers métastatiques spatialement proches, ou encore l'attraction de cellules métastatiques par des foyers existants, deux types d'interactions conduisant à des métastases moins nombreuses et plus massives et donc permettant possiblement d'obtenir des distributions en tailles métastatiques plus réalistes. Nous avons alors expérimentalement mis en évidence les deux phénomènes suggérés, qui se sont révélés être effectivement observés, l'un par IRM, l'autre à l'aide d'un protocole expérimental que nous avons suggéré à nos collègues biologistes, ce qui tend à montrer que ces phénomènes pourraient expliquer les nombres et les tailles métastatiques observés.

Nous nous sommes ensuite intéressés de plus près au phénomène de fusion de métastases évoqué précédemment. Nous nous sommes demandés si les interactions mécaniques entre métastases qui fusionnent pourraient avoir un effet sur leurs croissances respectives. Pour ce faire, nous avons utilisé le modèle spatial de croissance tumoral introduit précédemment (équations (1.1)-(1.2)), prenant en compte l'inhibition de la prolifération par la pression mécanique. Dans un premier temps, nous avons calibré ce modèle sur des données IRM de croissance de métastases pulmonaires (cancer du rein, mêmes données que dans le paragraphe précédent) chez la souris afin d'obtenir un espace des paramètres réaliste pour les simulations qui nous intéressaient. A l'aide de ce modèle calibré, nous avons simulé la fusion de métastases et avons étudié la dynamique de croissance résultant de ces interactions mécaniques entre métastases. Nous avons notamment observé une inhibition importante de la croissance due à ces interactions.

Le dernier chapitre de cette thèse est dédié à la modélisation d'un phénomène appelé niche prémétastatique. Des études récentes ont montré que la colonisation d'un organe distant par des cellules métastatiques pouvait être précédée de la formation d'une niche prémétastatique [10]. La formation de cette niche consiste en une pré-colonisation de l'organe distant par des cellules non tumorales issues de la moelle osseuse. Lors de la phase prémétastatique du poumon par exemple, certains facteurs émis par la tumeur primaire (située dans le sein par exemple) activent d'une part des cellules locales du poumon comme les fibroblastes et d'autre part des cellules progénitrices de la lignée myéloïde. Ces cellules progénitrices, originaires pour une part de la moelle osseuse, vont, grâce aux facteurs émis par les cellules activées du poumon, précoloniser cet organe avant l'arrivée des cellules métastatiques afin de rendre ce site perméable et favorable à la dissémination et à la colonisation métastatique. Ce phénomène est résumé par la figure 1.1.
A partir d'une étude de la littérature, nous avons construit un modèle mathématique décrivant la dynamique de cette niche prémétastatique. Ce modèle décrit la dynamique des différentes espèces impliquées comme les progéniteurs myéloïdes et les fibroblastes. Un couplage a ensuite été fait entre ce modèle et le modèle d'Iwata décrivant la dissémination et la croissance métastatique à l'aide des équation (1.3). Une partie des paramètres du modèle a été calibrée à partir de mesures expérimentales extraites de la littérature. Les paramètres restant ont été calibrés sur les données de Kaplan et al. (voir [10]) et le modèle semble capable de reproduire les dynamiques prémétastatiques et métastatiques. Les perspectives futures en terme de questions quantitatives et cliniquement pertinentes (la précocité de la formation de la niche prémétastatique chez le patient, les voies moléculaires à cibler, etc) sont ensuite discutées.

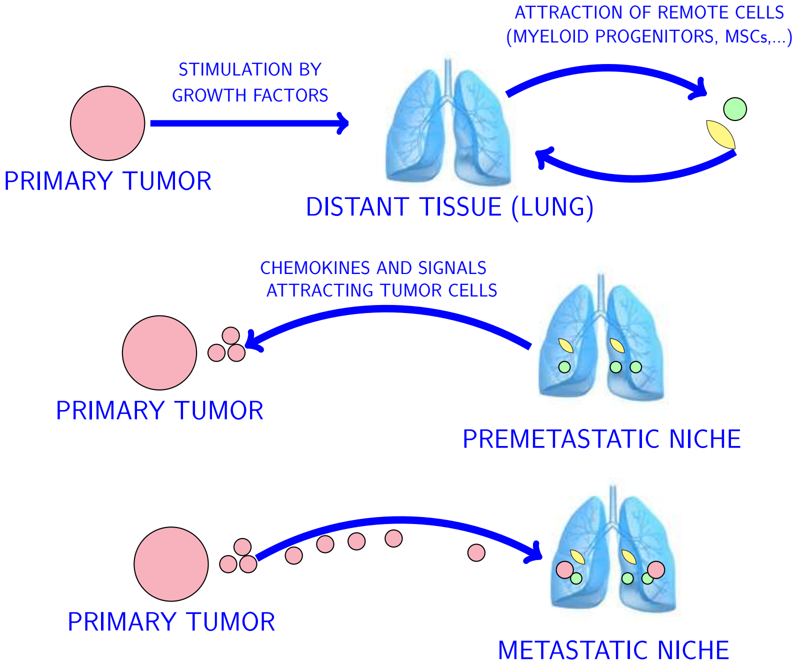


Figure 1.1: Schéma représentant les différentes étapes de la formation d'une niche prémétastatique

Part I

Quantitative biology of cancer

Chapter 2

A short overview on cancer

2.1 Generalities of clinical oncology

Cancer is the second most deadly disease of the world after cardiovascular diseases, with 14.9 million new cases and 8.2 million deaths estimated in 2013 [2]. Whereas breast cancer is the most common one in women and prostate cancer the most common in men, the lung cancer is the most common global cause of cancer-related death [3]. The first stage of cancer is usually a tumor that is initiated by cells that proliferate abnormally in an organ. Tumors of epithelial tissues are called carcinomas, whereas tumor of conjonctive tissues are called sarcomas. One of the first stages of tumor development is the local stage, which is called, in the case of an epithelial tissue, *in situ* carcinoma. The vast majority of tumors remain local and are benign tumors, eliminated by the immune system. However, in rare occasions, tumors can become invasive, spreading through the neighboring tissues. Ultimately, when the tumor is malignant enough, tumor cells detach from the primary, enter into the blood stream and arrest into a distant organ to form secondary tumors, or metastases. Fig 2.1 shows histological slices of a primary tumor and a metastasis in the lung. Clinically, 90% of the cancer-related deaths are due to metastases [5].

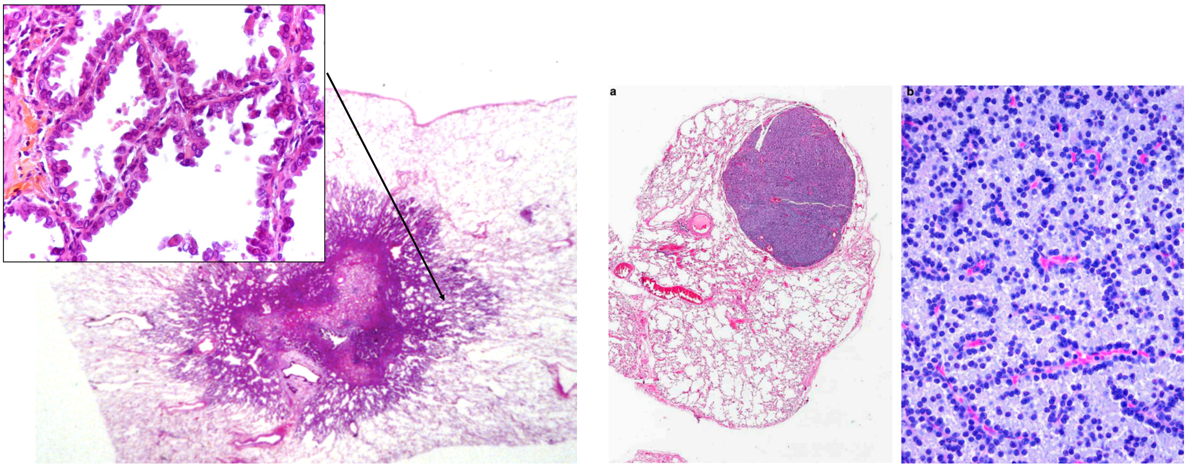


Figure 2.1: Left: Histological slice of a lung primary tumor; Right: Histological slice of a lung metastasis from a thyroid cancer; Images from Institut Bergonié, Bordeaux

Many kinds of treatments have been developed in order to stop the disease: surgery, chemotherapy, targeted therapies, radiotherapy, etc. These treatments have shown substantial effects on the disease. Indeed, The 5-year relative survival rate for all cancers diagnosed between 2004 and 2010 is 68%, up from 49% in 1975 – 1977 [2, page 2], which reflects earlier diagnoses and improvements of the treatments. However, the survival rate of some cancers remain very low: 7% for the pancreas for example. Unfortunately, in many cases, the disease ends up escaping to the treatments. When the cancer is generalized, it is often impossible to perform a surgery to resect all the metastases. Moreover, chemotherapies usually have side-effects on the patient and they are not enough to control the generalized disease, especially because the cancer cells develop resistances. Targeted therapies do not kill tumor cells directly but are directed against specific molecular pathways in order to neutralize the malignant phenotype of the cells. Animal models are often used in order to test these therapies. However, cancer is such a complex cellular and molecular system that effects that are observed on mice are not always observed on patients. The cancer biology has to be better understood, in order to perform more efficient treatments.

2.2 The hallmarks of tumorigenesis and tumor growth

Fundamentally, cancer cells are cells that managed to escape the control of the organism. In a physiological/normal state, many different kinds of cells ensure homeostasis of the tissues. One of the characteristic phenotypes of cancer cells is to proliferate abnormally. But what biological features characterize a cancer cell? Hanahan and Weinberg [4] put forward the hallmarks of cancer. The hallmarks characterizing cancer cell proliferation and tumor growth are first presented in this section, which summarizes the biology that defines a tumor. The hallmarks related to invasion, metastasis and crosstalk with the microenvironment are exposed in the next section.

2.2.1 Key factors of tumorigenesis

Sustaining proliferative signals

Normal cells are able to ensure tissue homeostasis. They can indeed control and limit other cells and their own proliferations. There are two principal ways to do that: 1) extracellular pathways and 2) intracellular pathways. By deregulating these mechanisms, cancer cells can maintain the cell division cycle activated and in this way survive, proliferate and grow pathologically. Such a chronic proliferation needs to sustain proliferative signals [4]. It requires to deregulate extracellular or intracellular pathways by the following ways: 1) secretion of molecules to stimulate other cells or themselves in a cell surface receptor activation fashion. These activated cells secrete growth factors in response which stimulate proliferation of cancer cells by activation of cell surface receptor (e.g EGF, epidermal growth factor) and 2) mutations allow cancer cells to sustain activation of intracellular proteases to maintain a cascade of proteolysis which is necessary to keep the cell division cycle activated. Normally, these intracellular signals are down-regulated when the cell proliferates too much. Mutations of cancer cells allow then to disrupt these regulations [4]. For example, the Ras gene codes for a protein that stimulates proliferation in presence of other signals. When this gene is muted, it codes a protein that stimulates proliferation independently of other signals [27].
However, high level of oncoproteins (Myc, Ras, Raf) expression can induce apoptosis or senescence (a non proliferative stable cellular state) [4]. That is why cancer cells have to develop another capacity: evading growth suppressor.

Evading growth suppressors

Some genes can regulate negatively the cell growth and proliferation. Cancer cells usually manage to evade these signals [4]. Here some of these genes are introduced.

Two major proteins
Two major proteins have been identified as important key molecules of growth inhibition. The first one is RB, or retinoblastoma-associated protein. This protein transduces extracellular growth inhibitory signals to decide whether or not the cell should continue its cell growth and division cycles. Cancer cells with defects in RB pathway function can evade some external growth inhibitory signals [4]. The second one is TP53 protein, which receives inputs from intracellular signals like genome damages, levels of nucleotide pools, growth-promoting signals, glucose, or oxygenation. If these levels

are suboptimal, TP53 can stop the cell division cycle and ultimately lead the cell to apoptosis [4]. TP53 gene is the most frequently mutated gene ($> 50\%$) in human cancer [28].

Contact inhibition
Contact inhibition is another growth inhibition mechanism. It involves NF2 gene, responsible for the production of the cytoplasmic protein Merlin which can lead to coupling cell-surface adhesion molecules (E-cadherin) to transmembrane receptor tyrosine kinases (EGFR, that is to say EGF receptor). Merlin can therefore facilitate adhesion of the cells and disrupt their stimulation by growth factors [4].

Transforming growth factor β (TGF-β)
TGF-β is a protein that has an anti-proliferative effect. It is necessary for cancer cells to evade its effects in order to follow pathological proliferation. TGF-β also facilitates the epithelial to mesenchymal transition (EMT), which is necessary for invasion and metastasis [4].

Resisting cell death

There are several ways for a cell to die but the two main ones are programmed cell death (apoptosis) and necrosis. One hallmark of cancer is the ability of cancer cells to escape from apoptosis [4]. That is why the most common way for a cancer cell to die is by necrosis. The two processes are illustrated in Fig 2.2.

Apoptosis
Apoptosis, or programmed cell death, is triggered in response to physiological stresses like oncogenic signals, DNA damage or hyperproliferation. This is an important research field to design therapies against cancer. Relatively recent studies revealed that apoptosis is attenuated within malignant tumors [4]. There are two major circuits for triggering apoptosis: 1) the extrinsic program, which consists in receiving extracellular death-inducing signals and 2) the intrinsic program consisting in receiving death-inducing signals of intracellular origin. The two types of signals result in activation of proteases (normally latent) which proceed to initiate a cascade of proteolysis for the execution of the apoptotic phase. The cell is progressively disassembled and consumed by other cells. The intrinsic program is more implicated as a barrier to cancer [4]. Balance between pro and anti apoptotic intracellular signals control apoptosis triggering. TP53 induces apoptosis in response to high levels of DNA breaks and chromosomal abnormalities. The most common strategy to escape apoptosis is the loss of TP53 tumor suppressor function [27]. Other factors that trigger apoptosis are insufficient survival factors and hyperactive signaling by oncoproteins (e.g Myc). In order to escape from apoptosis, cancer cells can increase expression of antiapoptotic factors, downregulate proapoptotic signals, increase survival signals and short-circuit extrinsic ligand-induced death pathway [4]. A possible anti-cancer therapy consists in restoring programmed cell death.

Necrosis
Necrosis is a phenomenon involving cell exploding. It is not a breakdown of tissue structure but a genetic program. Necrotic cells release proinflammatory cytokines that stimulate inflammatory cells to clean the place and repair the tissues. However, these inflammatory cells can have an unexpected effect in promoting proliferation, angiogenesis and invasiveness [29, 30]. Moreover, necrotic cells release IL-2 which stimulate other cells to proliferate [4].

Enabling replicative immortality

Replicative immortality barriers
Replicative immortality barriers have to be broken for a cell to acquire a cancerous phenotype. It is commonly accepted that normal cells are not able to undergo an unlimited number of cell cycles [4]. This limitation can occur in two different ways, which are 1) senescence, an irreversible non

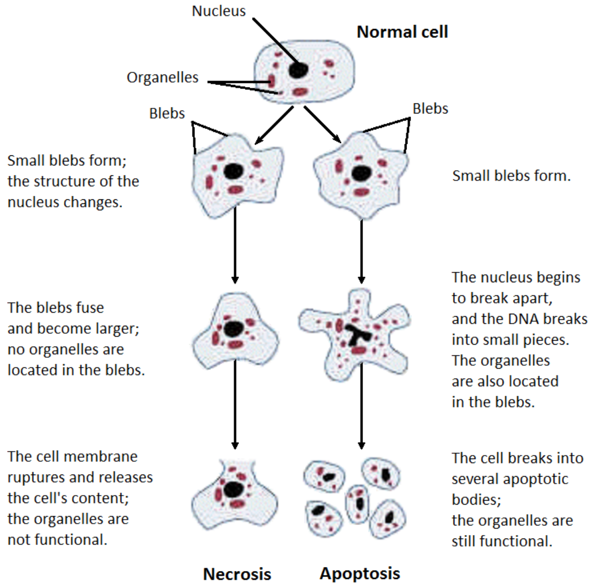


Figure 2.2: Apoptosis and necrosis processes, picture retrieved from [31]

proliferative but stable state, and 2) crisis state that leads to cell death (apoptosis). It seems that such limited replicative potential is linked to the degradation of telomeres [4]. A telomere is a region on a chromosome, which protects the end of the chromosome from deterioration or from fusion with neighboring chromosomes.

During cell cycle division, telomeres become shorter and progressively loose their ability to protect the ends of chromosomal DNAs from end-to-end fusions. Such fusions provoke genomic instabilities [4]. Before this, senescence is triggered in normal cells, but some cells succeed to circumvent this barrier and continue to replicate with genomic damages, resulting in a state of crisis then cell death. This degradation of telomeres can also be considered as the main proliferative potential limitation [4].

The telomerase function

The telomerase, an enzyme that can stimulate telomeres recovering, is not much expressed in normal cells but more expressed in the vast majority of cancer cell lines. This mechanism allows certain cells to acquire a replicative immortality [4].

Genome instability and mutation

Genome instability and tumorigenesis

Hallmarks capabilities are mostly acquired by impairing genetic materials of the cancer cells. For example, mutation or over-expression of the oncogene c-Myc leads to an excessive proliferation and ultimately tumorigenesis [27]. Another example is the tumor suppressor gene TP53. When it is mutated or inhibited, it induces inactivation of apoptosis and ultimately tumorigenesis. There are many other ways than mutations to induce the over/down expression of a gene or of a protein.

Genomic instabilities can accelerate tumorigenesis [4]. The absence of expression of the gene TP53, previously cited, allows the cells to sustain cell cycle division with eroded telomeres. Such configurations enhance the mutability of the genome, accelerating the acquisition of oncogenes or the inhibition of tumor-suppressor genes. It seems therefore understandable that combination of both TP53 and lack of telomerase enhanced tumorigenesis in mice models [4]. On human tumors, for example breast cancer, analyses highlighted important differences between premalignant and malignant tissues:

- Premalignant tissues: non expression of high levels of telomerase, telomere shortening and chromosomal abberations

- Malignant tissues: high levels of telomerase expression, aberrant karyotypes that outcome after telomeres failure but before telomerase acquisition

It seems that the optimal strategy for tumorigenesis is telomere failure to induce genomic mutations more easily, and then telomerase acquisition to get an immortal replicative capacity [4]. The first step requires to evade the main apoptotic signals.

DNA (Deoxyribonucleic acid) replication and transcription

DNA is a molecule with a double helix structure, which is a sequence of nucleotide bases. During the cell-division cycle, DNA has to be replicated to be transmitted to the daughter cell. During this process, several enzymes come into play. The replication fork is created by helicases, which break the hydrogen bonds holding the two DNA strands together. Then, DNA polymerases are responsible of DNA synthesis from parent DNA matrix (see Fig 2.3). The transcription is the first step of proteins synthesis. During this process, DNA is copied in RNA by the RNA polymerase in an analogous fashion.

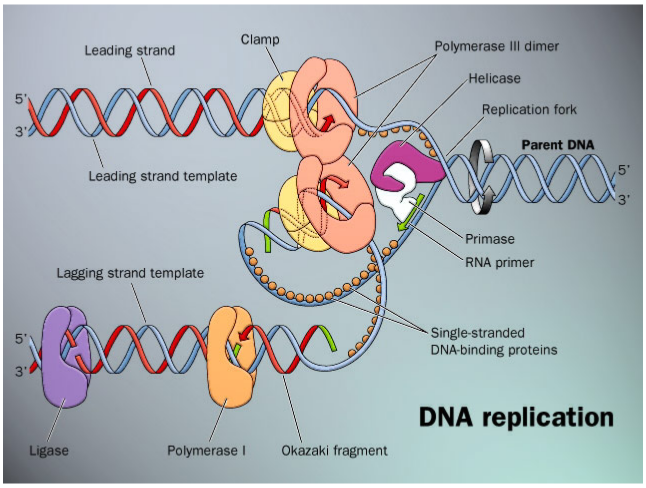


Figure 2.3: DNA replication, retrieved from [32]

Genetic and epigenetic alterations

Genetic and epigenetic alterations are crucial for tumorigenesis. During the DNA replication, DNA polymerase can do replication mistakes, which result in genome modifications, for example one nucleotide basis instead another. Some DNA polymerases have the capacity to correct replicative errors. In highly proliferative cells like cancer cells, the replication mechanisms are highly solicited by the intracellular signals and the DNA replication has to be fast. It enhances the probabilities of mutations during DNA replication [4]. Such mutations result in non well coded proteins, which can't assure their physiological functions, even acquire pathological functions .

Because of the multitude of enzymes (polymerases) or proteins, like histones, which outcome during DNA replication and during the transcription phase, there are many other ways than mutations to impair replication, transcription, as well as genes expression. These are epigenetic alterations [4], see Fig 2.4.

Enhancing the mutability allows to accelerate tumorigenesis process. Some enzymes and proteins have a role of maintenance of the genome integrity, resolving DNA defects. That is why in normal conditions, the rate of spontaneous mutations is usually very low during each cell generation. Thus, cancer cells manage to enhance their mutability in increasing sensitivity to mutagenic agents or compromising genomic maintenance system and surveillance system that eliminate defective materials (TP53 gene) [4].

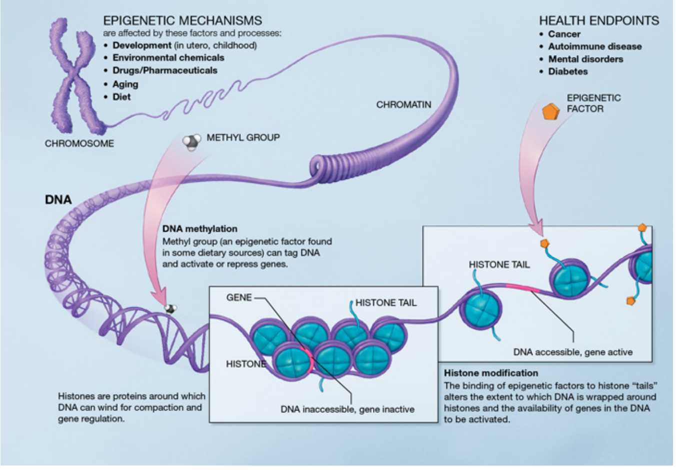


Figure 2.4: Epigenetics altertions, picture retrieved from [33]

2.2.2 Inducing angiogenesis

In order to survive and proliferate, cells need oxygen and nutrients that diffuse from blood vessels [34]. That is why they are generally located at the diffusion distance (between 100 and 200 μm) from the blood vessels. But for a tumor or an organism to grow beyond few cubic millimeters, they need to form new blood vessels by vasculogenesis and angiogenesis [35, 34]. In short, angiogenesis consists in new blood vessels formation and sprouting from existing ones toward the emission source (the tumor for instance) of angiogenic factors.

Angiogenesis is a physiologic transcient process which takes place normally in wound healing and woman cycle [4]. But it is also a critical step of the tumor growth. Indeed, tumor growth needs nutrients and oxygen sustain and evacuation of metaboli wastes which are the functions of blood vessels. That is why the angiogenic switch is almost always activated during tumor growth, resulting in continued sprouting of new blood vessels to sustain neoplastic progression [36]. In the most common paradigm, it is admitted that angiogenesis plays a role in the macroscopic tumor growth only, but it could be an early event in tumorigenesis and play a role in microscopic premalignant neoplastic progression [4]. The neovascularization patterns can be diverse among different types of tumors. Some tumors are hypovascularized with avascular and antiangiogenic zones, whereas other types are densely vascularized. In some tumors, specific oncogenes induce production of angiogenic factors, whereas in other cases, these signals are produced by inflammatory cells [4, 29, 30].

Proangiogenic and antiangiogenic factors

Activating the angiogenic switch depends on the balance between proangiogenic and antiangiogenic factors [37, 38].

One of the main pro-angiogenic factors is the vascular endothelial growth factor (VEGF), which has three receptors. Physiologically, it plays roles in embryonic (vasculogenesis) and postnatal vascular development, and homeostatic survival of endothelial cells. It increases the number of capillaries (sprouting) in a given network. One of its receptors, VEGFR2, activates cascade of signals in the endothelial cells which induce production of factors that increase vessels permeability [39, 40], endothelial cells proliferation and survival-basic fibroblast growth factor (bFGF)-,migration-vascular cell adhesion protein (VCAM) and matrix metalloproteinases (MMPs)-and finally differentiation into mature blood vessels. Another pro-angiogenic factor is the fibroblast growth factor (FGF). FGF-1 stimulates the proliferation and differentiation of all cell types necessary for building an arterial vessel, including endothelial cells and smooth muscle cells. This fact distinguishes FGF-1 from other pro-angiogenic growth factors, such as vascular endothelial growth factor (VEGF), which primarily drives the formation of new capillaries. FGF is also an important factor in the wound-healing process. The platelet-derived growth factor (PDGF) is a mitogenic factor for mesenchymal cells, and it plays a role in angiogenesis because recruitment of new pericytes/smooth muscle cells is necessary to stabilize new vessels.

One of the most common antiangiogenic factors is thrombospondin-1 (TSP). It can bind transmembrane receptors of endothelial cells to activate suppressive signals that counterbalance proangiogenic stimuli [4]. Other endogenous antiangiogenic factors are angiostatin, which is derived from proteolytic cleavage of plasmin [41] and endostatin, which is derived from proteolytic cleavage of 18 collagen [42]. An unbalanced mix of proangiogenic factors leads to aberrant vasculature : excessive vessel branching, distorted vessels, erratic blood flow, hemorrhages [4]. Many targeted therapies have been designed to impair tumor angiogenesis by blocking the effect of angiogenic factors or in using effect of antioangiogenic factors [43].

Role of the microenvironment in angiogenesis

The tumor microenvironment plays also a critical role in angiogenesis. Pericytes and smooth muscle cells are cells of mesenchymal origin, located on the surface of blood vessels and play a role in the blood pressure thus in the blood flow. They have physiological and mechanical roles for the endothelial cells support. For instance they are in part responsible for maintaining endothelial quiescence [39]. They must be detached from the vessels in order that neovessels can be formed. The role of the microenvironment for the formation of a stable vessel is summarized in Fig 2.5 retrived from [40]. As in part mentioned in the legend of the figure, angiogenic factors like VEGF-A destabilize blood vessels and induce $(ang) - 2$ expression by endothelial cells, which functions in an autocrine manner and leads to pericytes dissociation from blood vessels, enhancing the vessels permeability [39]. Destabilized endothelial cells can then proliferate and migrate in the presence of angiogenic factors like VEGF-A [39] to form blood sprouts. New endothelial cells do not originate from pre-existing blood vessels only but are also recruited from the bone marrow as endothelial precursor [43]. Inflammatory cells like macrophages, neutrophils and bone marrow-derived myeloid progenitors are known to play roles in initiation and sustaining of angiogenesis [29, 30]. Some of them can facilitate invasion and protect the vasculature from drugs [4].

2.2.3 Other key factors for ensuring tumor progression

Tumor-promoting inflammation

It is known that tumors are densely infiltrated by imune cells of the innate and adaptative immune systems [29, 44]. This inflammatory response is initially an antitumor response from the immune system that tries to eradicate the tumor. However, inflammation has the unexpected effect to promotes tumor progression in helping to acquire hallmarks capabilities [44, 29, 4]. Indeed, inflammation promote tumor growth by releasing growth factors, survival signals and proangiogenic factors [29]. Inflammation is also associated with releasing of extracellular matrix-modifying enzymes that facilitate angiogenesis, invasion, and metastasis, and inductive signals that lead to activation of epithelial-to-mesenchymal transition (EMT) and other hallmark-facilitating programs [29, 44, 30].

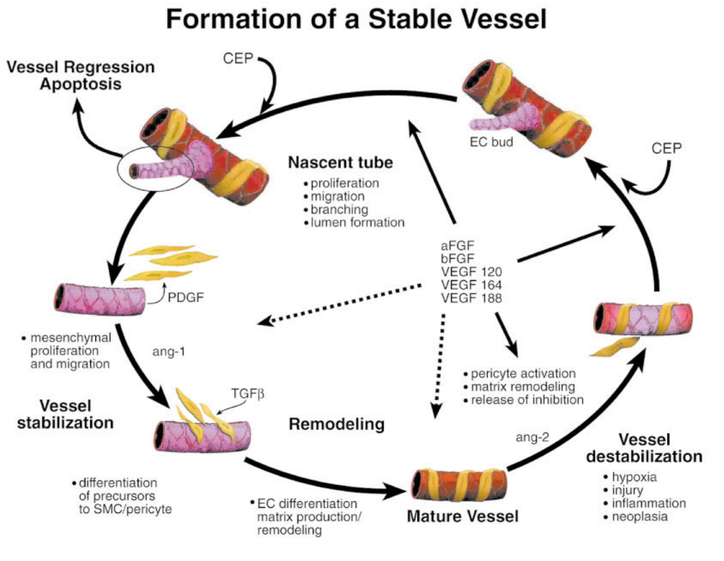
Reprogramming energy metabolism

Metabolism

To proliferate, cancer cells need to transform resources like oxygen into energy, which is contained in the chemical bonds of ATP (adenosine triphosphate) [4]. The first step is the glycolysis, in which glucose is transformed by the cell into pyruvate. During this cycle, 2 ATP units are produced, as explained in Fig 2.6.

During the second step, pyruvate is consumed, as well as oxygen, to produce 32 ATP units, which allow the cell to divide or to ensure other functions. This step is also illustrated in Fig 2.6.

Reprogramming



Current Controlled Trials in Cardiovascular Medicine

Figure 2.5: Assembly of a stable vessel. Local increases in angiogenic factors such as fibroblast growth factor (FGF) and vascular endothelial growth factor (VEGF) during new vessel formation destabilize a portion of an existing vessel (usually a venule). Destabilization is associated with increased angiopoietin $(ang) - 2$ expression and with pericyte activation, matrix remodeling, and induction of pericyte and endothelial cell (EC) migration and proliferation. Newly formed vessels may be dependent on exogenous factors for their survival until they have been remodeled to mature structures. Remodeling involves EC recruitment of pericyte/smooth muscle cell (SMC) precursors via endothelial-derived platelet-derived growth factor (PDGF). Once the mural cell precursor makes contact with the vessel, transforming growth factor (TGF)-β is activated, which in turn suppresses the proliferation and migration and induces the differentiation into SMC/pericytes. In addition to TGF-β, ang-1 produced by the SMC/pericytes is also involved in the stabilization and maintenance of the stable mature vessel. aFGF, acidic fibroblast growth factor; bFGF, basic fibroblast growth factor; CEP, circulating endothelial precursor. Picture and legend are retrieved from [40]

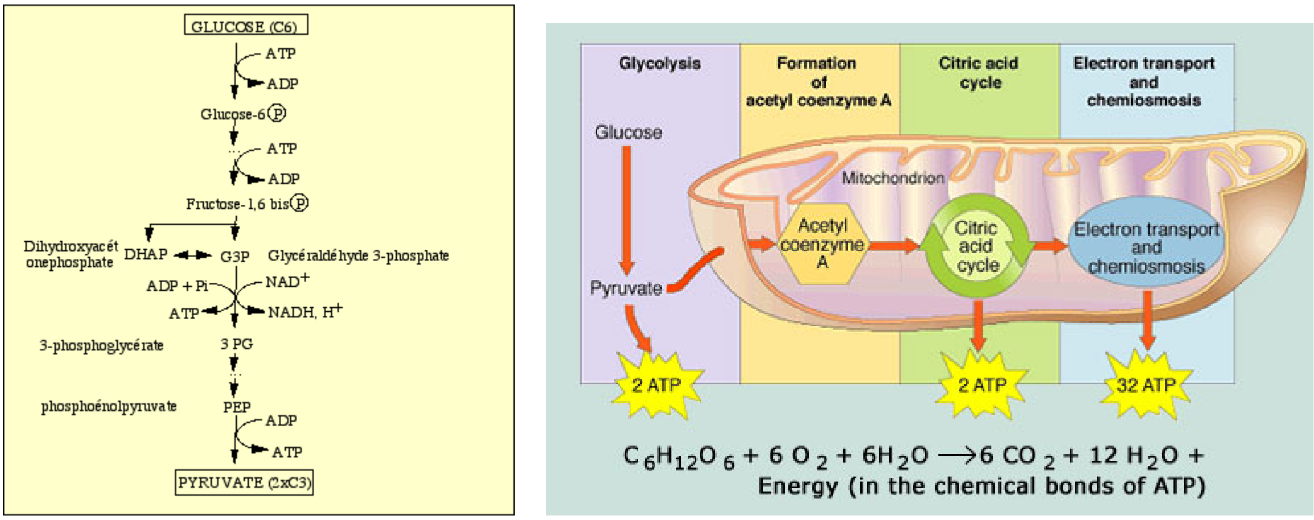


Figure 2.6: Left: Glycolisis: glucose⟶pyruvate+2ATP, picture retrieved from [45]; Right: The second step: pyruvate+oxygen⟶32ATP

However, when some cells lack of oxygen, they can survive with the glycolisis induced energy only. Some cancer cells, which are permanently hypoxic, become specialized in glycolisis and make this process very efficient [4]. When they consume the 2 ATP units resulting from glycolisis, they produce lactate, which make the environment acid and favor the expansion of the tumor cells around [46]. Thus, a tumor is constituted of an heterogeneous population of cells with different metabolisms, which grow in symbiosis [4].

Evading immune destruction

Experimental evidences show that innate and adaptive immune systems have strong anti tumoral potential. Indeed, in many experiments, mice with deficiencies in CD8 + cytotoxic T lymphocytes (CTLs), CD4 + T h 1 helper T cells, or natural killer (NK) cells develop significantly more tumors than control ones [4]. These immune cells have a surveillance function and a tumor cell killing function [44]. It means that tumors that managed to progress until macroscopic sizes evaded from the immune surveillance and managed to limit immunological killing. Moreover, tumor cells that arose in immunodepressed mice are less effecient to proliferate and to induce metastases than tumor cells that arose in immunocompetent hosts [4]. It probably means that immunogenic tumor clones can expand in an immunodepressed host, whereas they are eliminated in an immunocompetent host, giving the way to cells that can evade immune destruction. In patients, immune surveillance has been also highlighted to play an antitumorigenesis role [4]. In particular, lack of both CD8 + cytotoxic T lymphocytes and NK cells seems to lead to a more permissive environment for tumorigenesis. The main question is how cancer cells evade immune destruction? For example, cancer cells may paralyze infiltrating CTLs and NK cells, by secreting TGF-β or other immunosuppressive factors [4]. Cancer cells could also recruit inflammatory cells that are immunosupressive like regulatory T cells (Tregs) and myeloid-derived suppressor cells [44, 29]. Both can suppress the actions of cytotoxic lymphocytes.

2.3 Elements of the metastatic process biology

Metastasis development is a multiple step process in which tumor cells detach from a primary tumor to colonize distant organs [4]. During the last century, many research efforts have been devoted to understand the main aspects of carcinogenesis and more particularly genetic alterations leading to neoplasia, but the metastatic process received relatively less attention [5]. However, the biology of metastasis is more and more studied, especially because secondary tumors generally cause the death of the patient [5]. The metastatic process is summarized by Fig 2.7, retrieved from [20].

2.3.1 About the state of metastasis research in 2000

Metastasis biology is a still poorly understood field, although active research is conducted, especially since 2000 [4]. By then, even less was understood. Metastasis is usually thought as the last stage of cancer progression and is often linked to the phenomenon of local invasion. In [20], Talmadge and Fidler present a historical perspective of advances that have been made on the biology of metastasis.

First discoveries

It has been early observed that certain types of tumors metastasize in particular organs, which has been explained by both mechanical and molecular factors [20]. Steven Paget's “seed and soil” hypothesis, which defends biological compatibility between tumor cells (seed) and the specific organ (soil) is discussed further in this manuscript in chapter 8. It has been early observed that tumor cells can colonize lymphatic nodes by lymphatic channels before to disseminate by hematogenous way, and a distinction between lymphatic and hematogenous metastases has been early made [20]. The importance of adhesion and motility and the role of proteolytic enzymes for tumor cells invasion and metastasis have also been raised before 2000 [20]. Talmadge and Fidler also mentioned discoveries on the critical roles of immune cells like macrophages and NK cells in metastasis and the fact that

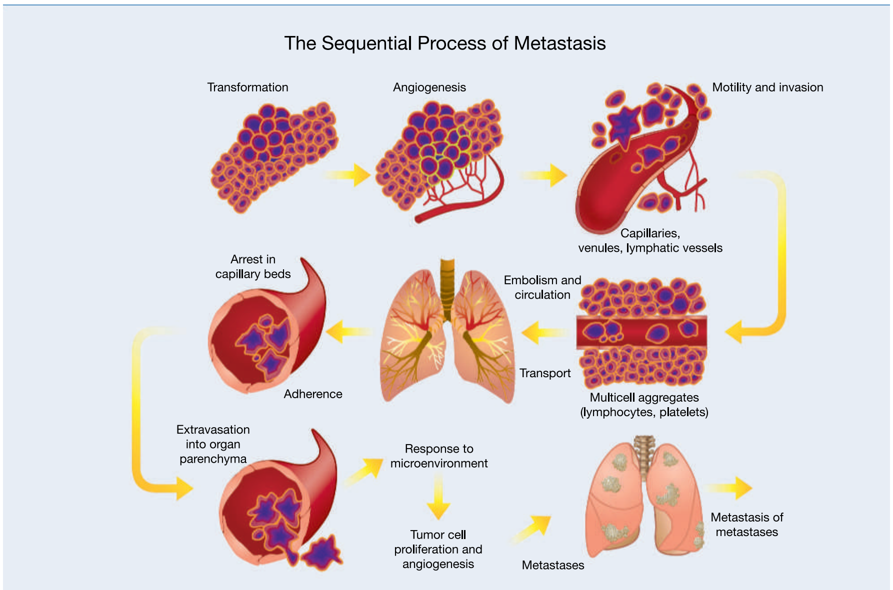


Figure 2.7: Metastatic cascade, steps of the metastatic process, picture retrieved from [20]

metastases are clonal and originate from one surviving cell [22, 20]. Another observation has been made about the link between angiogenesis and metastasis: new blood vessels due to angiogenesis can provide a channel for circulating tumor cells from the primary tumor to the distant site [47].

Cadherins

One of the most characterized feature was loss of E-cadherin, a transmembrane protein involved in cell-cell adhesion [4]. Conversely, other adhesion proteins such as N-cadherin are involved in cell migration (for instance during embryogenesis or inflammation). These are upregulated in many cancers [4].

2.3.2 New discoveries

Epithelial to mesenchymal transition (EMT)

Transformed epithelial cells can undergo a transition to a mesenchymal state that offers them abilities to invade, survive and disseminate. In this, cancer development shares common features with embryogenesis and wound healing. EMT is regulated by transcriptional factors, can be transient or stable and is associated with: change in morphology from polygonal to spindly (like fibroblasts), expression of matrix-degrading enzymes such as matrix metalloproteinases (MMPs), increased motility, heightened resistance to apoptosis [4].
Cells that underwent EMT are often located at the periphery of the tumor, suggesting different microenvironmental stimuli depending on location [4].
It is thought - although not proven - that transcriptional factors regulating EMT are responsible for most of the aggressive features of invasive carcinoma cells.
Involvement of the microenvironment suggests that passage through the multiple steps of the metastatic cascade could happen without further clonogenic modifications of the cancer cells than required in the initial stages of cancer progression. This remain to be tested [4].

Interactions with stromal microenvironment (at primary site)

There is a crosstalk between the cancer cells and surrounding environment. Macrophages and other stromal cells are involved as they can supply matrix-degrading enzymes in order to help invasion and they can also facilitate intravasation [30, 44].

Mesenchymal to epithelial transition (MET)

If EMT is driven by environmental factors, then the reverse transition could occur when the colonizing cell(s) arrive at the distant site. Additionally, it could be simplistic to think in terms of epithelial or mesenchymal as a continuum could exist between both states [4].

Other forms of invasion

- Collective invasion is a - possibly different from EMT - process that often leads to invasion without distant metastasis [30, 4].
- Other possibility is amoeboid (morphological plasticity) invasion, in which cells do not clear a path for themselves by degrading the matrix but rather deform their shape to go through the obstacles. It is still open to determine whether this last form has common biological mechanisms with EMT or not [4].
- Yet another possibility is recruitment of inflammatory cells that produce MMPs, avoiding the need of EMT to produce them [30, 44].

Colonization

The metastatic process consists in two major phases: dissemination and colonization, the latter happening at the distant site [4]. The former could mostly result from EMT while transition of micro to macro metastasis could be differently driven. Colonization is not strictly coupled with dissemination, as proven by the numerous data of large number of micro (and possibly occult) metastatic lesions in patients, without further cancer development [4].

Dormancy
Colonization can be marked by a (possibly long) dormancy phase, during which tumors remain small and non-expanding. Dormancy can be due to intrinsic lack of angiogenesis-stimulating capacity from the metastatic tumor [5]. Nutrient starvation, anti-growth signals and immune pressure are other possibilities [4]. Escape from this dormant state can happen a very long time after dissemination. This dormancy can be provoked by distant action of the primary tumor, for instance through systemic inhibition of angiogenesis [4, 48]. The question of metastasis dormancy is clinically very important because dormant metastases are generally not visible on a patient and the possibility of blowing up of the metastases has to be anticipated by the physician [47].

Metastatic microenvironment
Importance of extracellular matrix components like fibronectin or expression of chemokines for metastasis development [44]. Moreover, the distant site has possibly been *a priori* prepared to form a hospitable soil, possibly by signals emitted from the primary site, a process involving bone marrow derived cells (BMDCs) that are recruited to form a premetastatic niche [10].

Time of dissemination
For a lot of cancers, metastasis is the last step of the multistep process of cancer progression. This is proven by genetic studies showing clonal evolution from ductal adenocarcinoma to metastasis (in prostate) [4]. On the other hand, it has been evidenced in preclinical studies that cells can disseminate pretty early, although clinical relevance of early dissemination remains to be assessed [4]. About this topic, Groom and Chambers mentioned that it is often the case that many patients have already disseminated cells or invisible metastases when diagnosed [47]. It highlights the clinical importance of the early dissemination. The theory of early dissemination and parallel growth of metastases, opposed to the theory of linear progression where lymph nodes are invaded first and tumor cells ultimately disseminate by the hematogenous way, is defended by Klein on the basis of quantitative considerations on clinical data [24]. It remains open to know whether capability to develop macro

metastases is acquired at the primary or secondary site, the latter resulting from selective pressure at the distant (possibly hostile) site [4].

Self-seeding
Metastatic cells can disseminate further from the distant site, in particular back to the primary tumor [49, 50]. Indeed, Kim et al. showed evidences of cells exchange between two tumors implanted in a same mouse [11]. This could give an alternative explanation to the observed cells with metastatic capability at the primary site. Moreover, based on the kinetics argument that tumor doubling times increase with the size of the tumor, Norton notices that multiple seeds grow faster than one large tumor and he proposes that malignant tumors that seed a lot grow fast because of the self-seeding [50].

Genes that mediate metastatic development
In 2000, six genes have been identified, regulating survival and growth of metastatic cells in the secondary sites [47]. Many efforts are currently devoted to find genes that characterize the metastatic program [51].

Chapter 3

Mathematical models in cancer biology

As we could see in the previous chapter, cancer, from tumorigenesis to metastatic colonization, is a very complex biological process. Understanding this complexity is motivated by designing new drugs and finding new treatment schedules. Studies have been so far focused rather on qualitative aspects of the cancer biology. However, measurement techniques progress, genetics development, and growing interest on dynamical aspects of the biological processes are leading researchers to focus more and more on the quantitative aspects. In this context, mathematical models could be powerful tools for data analysis and quantitative predictions.

3.1 A crosstalk between experiments, theoretical biology and mathematical models

3.1.1 Mathematical models applications

Models in sciences

In physics and biology, observing a natural phenomenon on the real system studied is not always possible and use of a miniature model is required for understanding it. These models are not only built for complexity reasons but also for material reasons: in fluid mechanics for example, it is sometimes technically impossible to study the real flow in the conditions that we are interested in. A solution is to study the flow on a miniature model with a reduced scale. In biology, for ethical and health reasons, there are few ways to study the biology of a disease on patients. That is why studying the disease on animals like mice or rats is necessary to understand the biology and to test treatments.
In order to understand natural phenomena, we need to build conceptual models that capture the essential elements of a phenomenon. For example, the idea of the rotation of the earth around the sun, viewed as a fix center, is a model. Indeed, because of the inertia of the sun which is much more massive than the earth, we neglect the fact that the sun is also attracted by the earth that exerts on the sun the same force exerted by the sun on the earth. To understand natural phenomena, as explained in [52] by Wheldon, it is often not necessary to take into account all the phenomena that theoretically could have a causal effect on what we are interested in. Due to the limitations of our brain to understand the interactions of all the elements of the universe, such a simplification is necessary to essentialize the phenomenon. A model consists in picking out the main elements of the phenomenon that have a causal influence on what we perceive and observe. In sciences, theories or conceptual models are proposed to explain natural phenomena. When the theory is rejected by observations, another theory has to be proposed, including new elements that were not taken into account in the former one.

Mathematical models

In physics, mathematics have proven to be useful and even necessary. Mathematical models for physics have been used to make the knowledge progress in confronting theories with experiments. When a theory does not match with the experimental results, it has to be rejected. When the theory matches, it can be considered as valid until another theory is able to explain observations better. For example, circular trajectories of the planets around the suns could relatively well explain the planets trajectories until Kepler proposed elliptic trajectories.
Mathematical models of physical phenomena are also used for their predictive power in industrial fields like aeronautics, aerospace, and many other fields or for scientific fields like astronomy to predict planet trajectories.
Whereas the role of mathematics is essential in physics, it is less obvious in biology. In physics, lots of phenomena can be explained in a quantitative way and lots of theories can be expressed in a quantitative way. A theory explains a phenomenon by the separate effects of few essential elements. In physics, it is often relevant to express the objects of interest (mass, magnetic field, etc) in a quantitative way. Theories where the different effects on a same phenomenon can be expressed quantitatively are adapted to be formalized in mathematical language. A mathematical model, which is the mathematical formalization of a precise quantitative theory, is a relevant tool to deduce implications of theories dealing with quantitative objects. In physics, these models are often formulated in time-dependent differential equations, leading to a solution that describes the time course of an observable quantity. In biology, phenomena are most of the time understood in a qualitative way, which limits the relevance of using mathematical models. However, the same approach as in physics could be made in biology fields where quantitative aspects would be of biological interest. The growing interest in dynamical aspects [53] of the cancer disease and the amount of quantitative data seem to make cancer research a fertile field for mathematical models applications.

A mathematical model for what purpose?

The two principal uses of mathematical models are devoted to: 1) clinical purposes and 2) biological purposes.

Clinical purposes
A first goal of the mathematical models for cancer consists in predicting the evolution of the disease for a patient in order to help the physician in his therapeutic decisions. For example, from three CT-scan images, we could try to predict the fourth six months after to help the doctor to anticipate the evolution of the disease. Another purpose of mathematics can be to simulate the effect of a therapeutic protocol in order to perform scheduling optimization.

Biological purposes
Another goal of the mathematical models for cancer is to improve our understanding of the cancer biology. As mentioned before, cancer biology is very complex and not well understood yet. Understanding it better would help to design new drugs or therapeutic protocols in order to improve the way of how patients are treated. In one hand, a mathematical model can be used to infer insights that are not obvious in the untreated experimental data. A mathematical model formalizes a biological theory that could explain a phenomenon. The goal of the modeling approach is to validate or reject theories in confronting models with observations.

In this thesis, the purpose was of the second type: cancer biology understanding and more particularly the biology of metastasis. We had to keep in mind that the modeling approach can differ in the two previous cases. In a clinical context, a model has to be predictive, in order to infer useful information to the physician. Such a model does not need to be based on a solid and realistic biological framework. A totally phenomenological model that does not describe any biology can be relevant in this context if it is predictive. For biological purposes, it is different because the goal of the model is often to infer reliable biological informations from the parameters or the simulation results. It does not mean that such a model has to contain all the biology of the studied phenomenon. First because it is not possible, as we are always far from understanding the complete biology of the phenomenon, second because the number of parameters that we can identify is limited by the amount of available data, and third because the power of a model is exactly to be able to essentialize the phenomenon in few major components.

3.1.2 Using mathematical models for theoretical cancer biology

Rejecting and validating theories

In the first chapter of his book “Mathematical models in cancer research” [52], Wheldon mentioned, based on Karl Popper’s work: “a theory cannot be proved correct experimentally but only refuted if it makes incorrect predictions about some natural phenomena”. He also mentions that the predictions of a theory are not always simple to know and that mathematical models are sometimes needed to know these predictions. A theory that makes valid predictions only gains in plausibility but is never validated as the truth, the absolute theory that reflect the reality. Indeed, whenever another theory shows predictions valid as well, it is not possible to discriminate one or the other theory. Although of lesser importance than in physics, mathematical models could play a similar role in theoretical biology thanks to a similar approach where observations and data are continuously confronted.

Data-driven modeling

The previous approach calls for a continuous crosstalk between models and experiments. A theory is systematically built on what is perceived and understood of first observations (e.g. the idea of the gravitational force from the observations of falling objects). When a quantitative theory is non-ambiguously formulated, it can be formalized into a mathematical model. The predictions of this model can then be compared to observations. If these predictions are not in accordance with data, the theory has to be changed or enriched.

Biology is an experimental science that has known improvements without the help of mathematical models. A theory can be formulated as a question like: "Does this molecule have a pro or an anti-tumoral effect?" As in physics, such a question is highlighted by previous observations and new experiments have to be made in order to answer to this question or reject/validate some theory. The role of mathematics in biology could be to highlight new questions that focus more on quantitative aspects than in the classical view of the biology.

For instance, a quantitative question could be: "Is the doubling time of a tumor constant during the growth?". Such a question could be answered using a mathematical model, and this is precisely the purpose of the next paragraph. Moreover, as we just mentioned, mathematical models could help in giving rise to new questions, in the sense of suggesting new hypotheses that could be tested experimentally. The scheme of Fig 3.1 summarizes the data-driven modeling approach for theoretical biology.

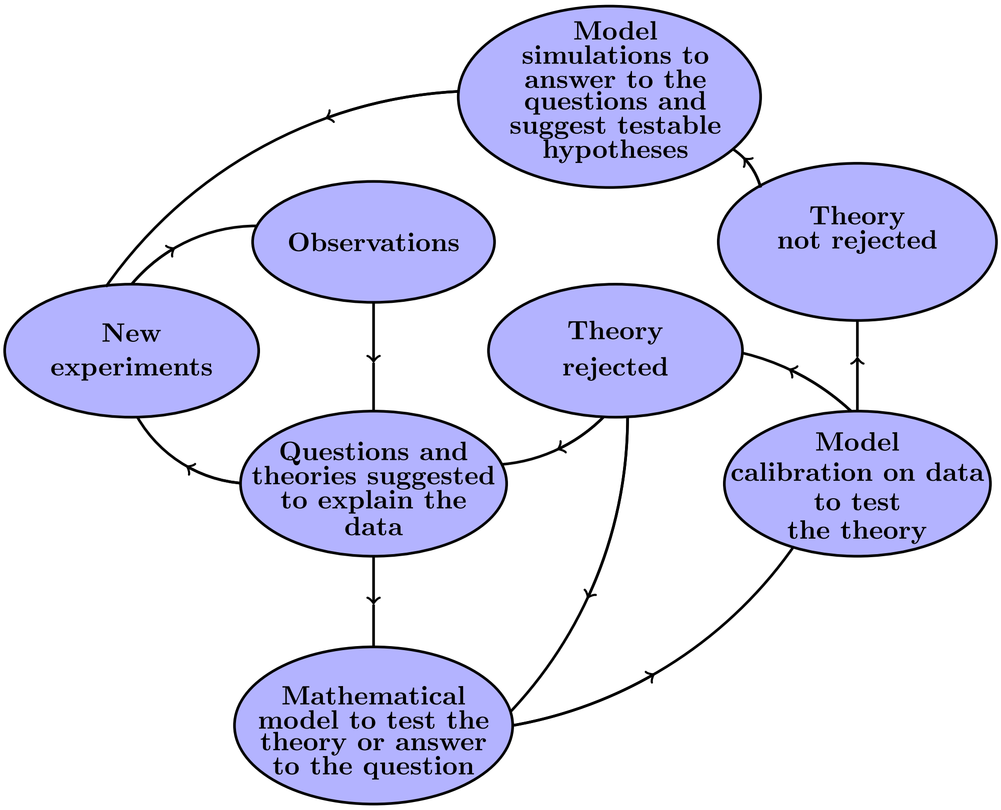


Figure 3.1: Data-driven modeling in cancer theoretical biology

Naturally, this modeling approach is not the only one possible. In this thesis (except in the last chapter), we essentially used this approach. However, other approaches, more theoretical, are possible and sometimes very useful as they can provide general insights on the studied biological process [54].

How complex a mathematical model has to be

As previously mentioned, a mathematical model formalizes a quantitative theory of what are the causal effects of the biological entities of the system on each other. The description depth of a phenomenon can be limited to the macroscopic level like the tissue scale, but can also reach much smaller scales like the cell scale. The level of biological description in the model can be very flexible to describe a same global phenomenon. For example, to describe tumor growth dynamics, a scalar, the tumor volume, and an ordinary differential equation describing the increase of this volume can be sufficient to assess useful quantitative predictions for a patient. However, it would be possible to include in the model a variable representing the vascular density within the tumor and to model its effect on the tumor growth and the reciprocal effect of the tumor on vascular density by angiogenesis. It could also be possible to include the effect of the immune system into the model. Such refined models can be useful as they provide predictions on not necessarily observable biological quantities such as the vascular density or immune infiltration. A good example of such a contribution is the modeling work of Michor et al. on chronic myeloid leukemia, which allowed to infer information on the dynamics of subtypes of cells that have different states of differentiation (whereas only the total number of cells was measurable) [55]. Their analysis generated informative clues on the particular cells that were resistant to treatment (the stem cells compartment). An other use of adding biological aspect into a model is to simulate effects of specific inputs. For example, including angiogenesis into a model of tumor growth can be useful to quantify the effect of an anti-angiogenic therapy [56]. Starting from these observations, it could be tempting to include all the biological aspects that seem important in the studied process: angiogenesis, immune system, key genes, extracellular matrix, growth factors for example, in order for the model to be as realistic as possible. However, as very well explained by Wheldon in [52], biological entities like cells or tissues can hardly be defined as having fundamental and invariant properties because the properties of a cell, for example, are strongly mediated by the other cells present and surrounding factors. In physics, depending on the scale, an object can be defined as having a set of fundamental properties like its mass for example. This is why, according to Wheldon, mathematical biology has for now to deal with relatively simple problems likely to benefit from mathematical models as a quantitative tool. Finding relevant quantitative problems for which progress can be made thanks to mathematical modeling remains one of the main difficulties. However, there are several possible approaches of cancer modeling. The mechanistic approach consists in describing phenomena at a fine scale like the cell scale. It introduces lots of biological elements in the theory, in order to deduce information at a more macroscopic scale and can be called a bottom-up approach. The second approach, more phenomenological, consists in essentializing the phenomenon in few main elements to exhibit general laws/theories that can be confronted to data in order to deduce biological informations. This kind of approach can be called top-down approach.

Simulations and mathematical analysis

Solving the mathematical models is not always simple, depending on the complexity of the model. For some ordinary differential equations-based models, it is possible to exhibit an analytical solution. In these cases, the solving part is not a big deal. However, some ordinary differential equations (ODE) systems are not as simple to solve and a numerical approach is required. Mathematical analysis of the models allows sometimes to exhibit qualitative properties of the solutions and to understand the behavior of the model. For complex systems like some partial differential equations, mathematical analysis is also necessary to show the well-posedness of the problem.

3.2 Example of a top-down modeling approach

The next paragraph presents a simple example of top-down approach that illustrates the modeling philosophy of this thesis. We consider here a simple situation where mathematical models can be used to test theories against data and to infer information that are not obvious in the data. In this example, the goal is to provide the most parsimonious growth laws that could explain data of tumor growth in different situations.

3.2.1 First situation: *in vitro* cell proliferation

Here we deal with kinetics of renal cell carcinoma (RENCA) tumor cells that are growing *in vitro*. RENCA is a murine cell line of kidney cancer. Fig 3.2 shows the data of cell kinetics obtained from [57].

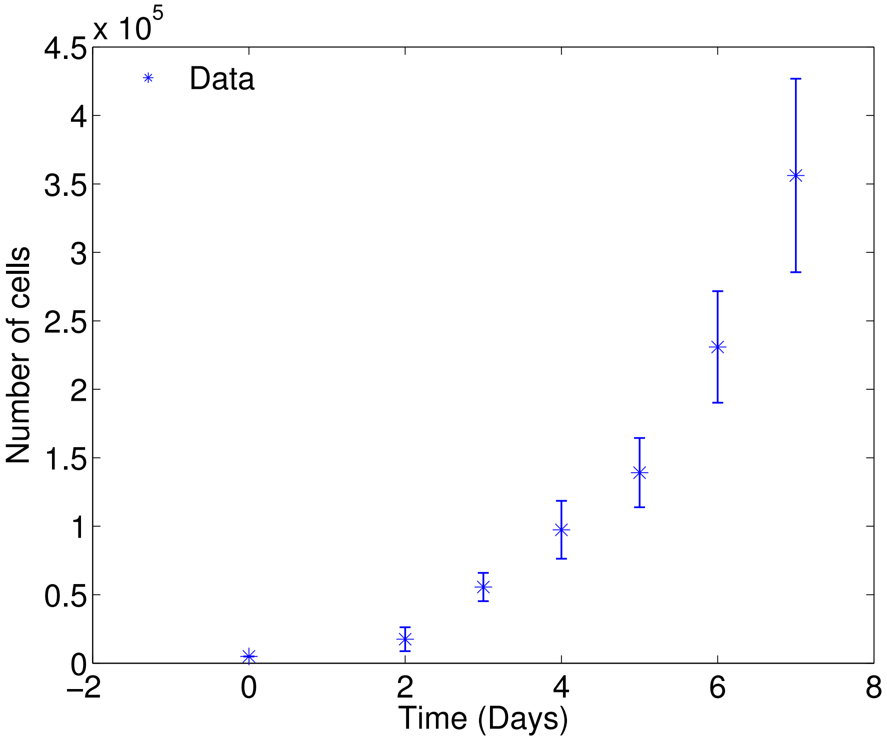


Figure 3.2: *In vitro* proliferation assay of RENCA cells retrieved from [57]. The points of the curve represent the number of cells time course. Mean values ± standard deviation.

We look for a model able to describe these data that would be in accordance with the biological knowledge, with parameters as reliable as possible with experiments, and as simple as possible, according to the parsimony principle. A simple theory of growth that could explain these data would be that a constant fraction of the cancer cells proliferate with a constant doubling time. How to mathematically formalize this theory? The exponential law is the mathematical law that describes such dynamics. The *in vitro* assay being in 2D, we define S the surface occupied by the cells and write the exponential law followed by this surface:

$$\begin{aligned}&\frac{\mathrm{d}S}{\mathrm{d}t} = aS,\\ &S(t=0) = S_0,\end{aligned}$$

where a is the growth rate of the cell line. This parameter is equal to log(2) multiplied by the ratio between the cell fraction in cycle and cell cycle duration. Can this theory explain the data? To answer this question, we calibrated the parameter of the model to fit the data. To do so, we minimized the following weighted least square criterion:

$$LS(a) = \sum_{i=1}^{n} \frac{(S(t_i, a) - y_i)^2}{\Sigma_i^2},$$

where n is the number of data points, $(y_1, .., y_n)$ are the observations at times $(t_1, ..., t_n)$, and $\Sigma_1, ..., \Sigma_n$ are the standard deviations of the observations at times $(t_1, ..., t_n)$, represented by the error bars on Fig 3.2 and retrieved from [57].

We observed that this theory was able to describe the data, as shows Fig 3.3. Indeed, except in one point, the normalized residuals $\frac{S(t_i,a)-y_i}{\Sigma_i}$ were all lesser than 1, meaning that the error of the model is lesser than the variability of the data. Moreover, a statistical test did not reject the null hypothesis of a standard normal distribution of the normalized residuals ($p =$ by the $z-$test). Particularly, it implies that the residuals are distributed around 0, a result expected from a good model.

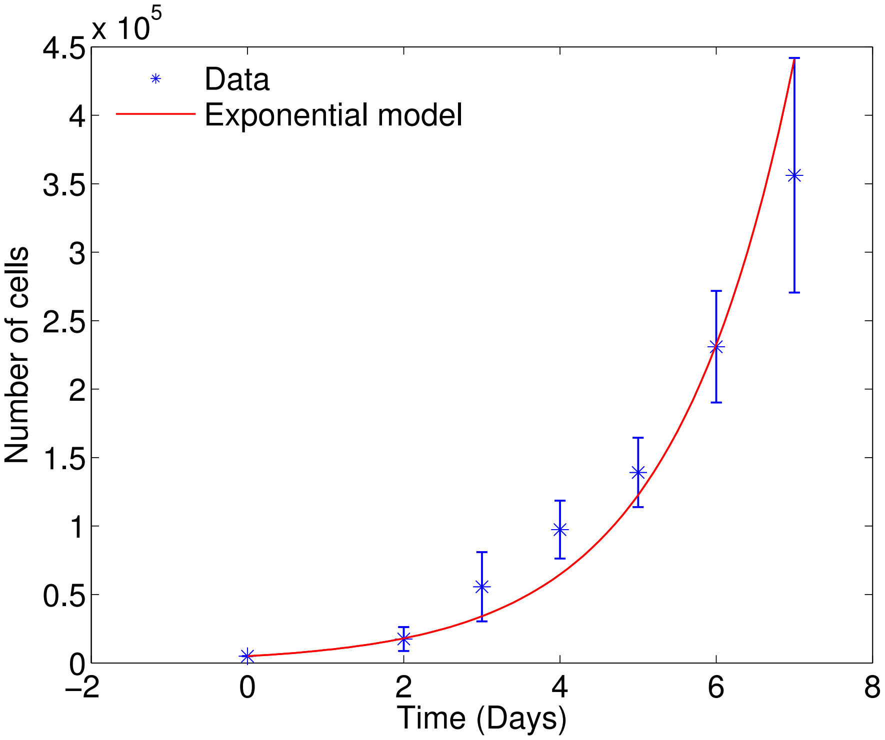


Figure 3.3: *In vitro* number of RENCA cells time course. The points represent the data and the curve the fitted exponential model.

What have we learned?

The theory “a constant fraction of the cancer cells proliferate at a constant doubling time” is able to describe *in vitro growth* of RENCA cells. It does not mean that this theory describes exactly the reality. We could not discriminate other theories that would fit equally well. We also inferred a biological from the data: the calibrated value of the parameter a provides us the *in vitro* doubling time of the RENCA cell line.

3.2.2 Second situation: *in vivo* tumor growth

The previous model is able to describe *in vitro* tumor growth but what about *in vivo* tumor growth? This time, the model is written:

$$\begin{aligned}\frac{\mathrm{d}V}{\mathrm{d}t} &= aV,\\ V(t=0) &= V_0,\end{aligned}$$

where V is the total volume of tumor cells. We consider data of the same cell line (RENCA) but injected in the renal subcapsule of a mouse. Fig 3.4 presents this data. These data are the same as used in section 5.1.2 of the chapter 5 to study tumor cells kinetics in the kidney. In this experiment, tumor cells are tagged with green fluorescent protein (GFP) and each data point corresponds to one distinct animal. A statistical model of the interindividual error is further presented in 5.1.2, resulting

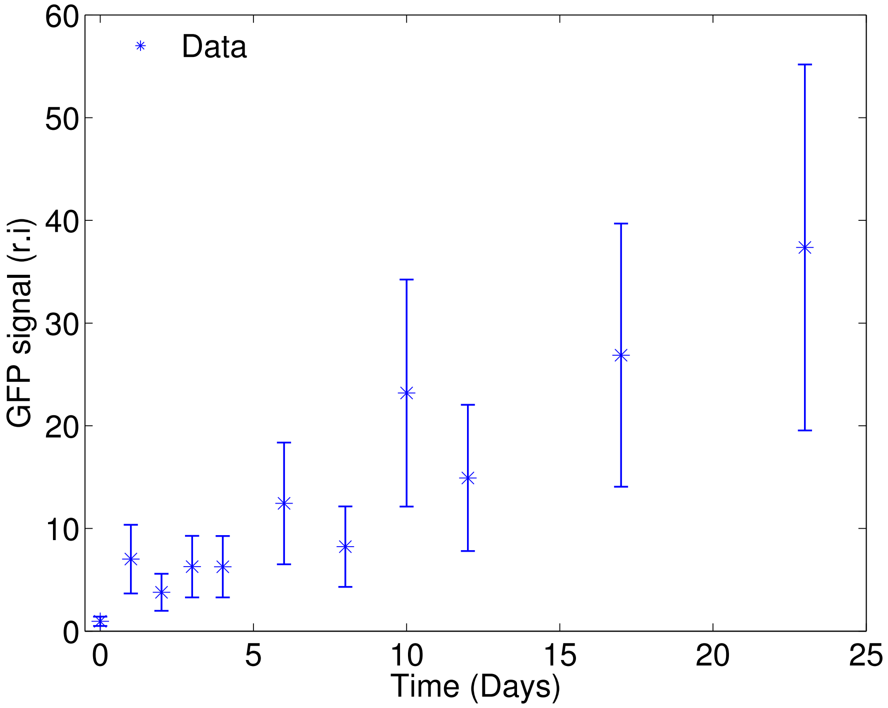


Figure 3.4: *In vivo* proliferation of RENCA cells in the kidney. The points of the curve represent the number of cells time course. Mean values ± standard deviation.

in error bars presented in Fig 3.4 and in the use of the following weighted least square criterion to minimize for the model calibration:

$$LS(a) = \sum_{i=1}^{n} \frac{(V(t_i, a) - y_i)^2}{\Sigma_i^2},$$

where Σ_i represents the model of standard deviation that has been established in 5.1.2.
When we tried to fit the exponential model to this *in vivo* data, we observed that this theory can not describe the data, as shown in Fig 3.5. Indeed, two normalized residuals only are lesser than 1. Moreover, a normal distribution of the residuals is rejected by the $z-$test ($p < 0.01$). These results show that the residuals are biased and not randomly distributed around 0. Indeed, in Fig 3.5, we can see that the model curve is under all the points except the last one.
This means that *in vivo* growth can not be explained by a constant growth rate, which is maybe a too simple theory to describe *in vivo* growth. Indeed, it is often accepted that *in vivo* growth exhibits increasing doubling times [58, 59, 60]. To take this phenomenon into account, we propose a slightly more complex model:

$$\frac{\mathrm{d}V}{\mathrm{d}t} = \alpha V - \beta V \log(V),$$
$$V(t=0) = V_0.$$

This model is called “Gompertz model” because it was introduced by Benjamin Gompertz in 1825 (but not to describe tumor growth). This model describes a growth with a relative growth rate that decreases exponentially. In trying to fit it on the *in vivo* data presented in Fig 3.4, we concluded that this model was able to describe RENCA *in vivo* growth in the kidney, as shown in Fig 3.6. Indeed, 9 normalized residuals on 10 are lesser than 1 and the $z-$test does not reject the null hypothesis of a standard normal distribution of the normalized residuals ($p = 0.2$).

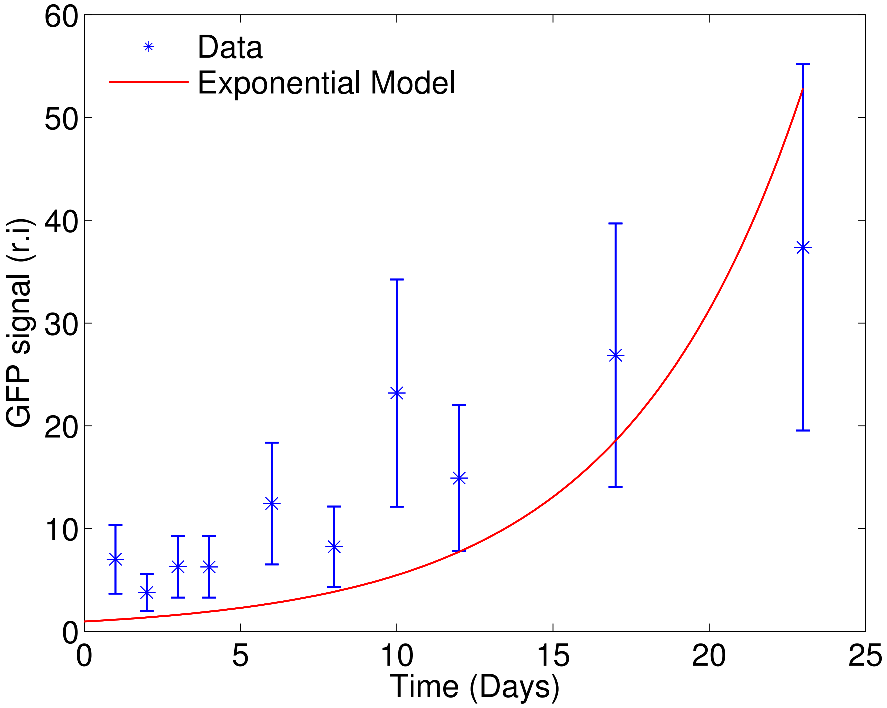


Figure 3.5: *In vivo* number of RENCA cells time course in the kidney. The points represent the data and the curve the fitted exponential model.

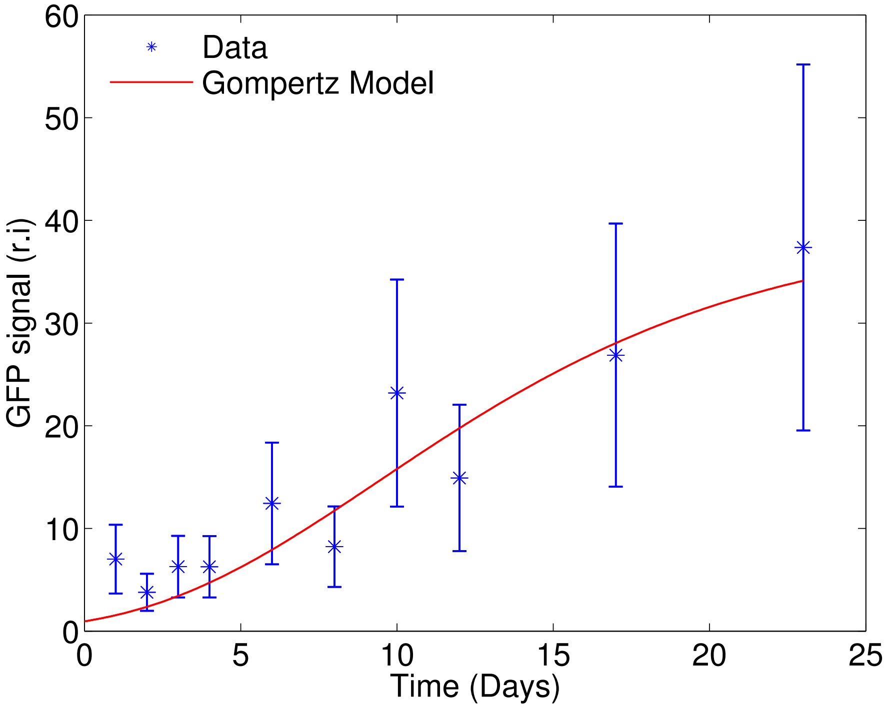


Figure 3.6: *In vivo* number of RENCA cells time course in the kidney. The points represent the data and the curve the fitted Gompertz model.

What have we learned?

The theory "a constant fraction of the cancer cells proliferate at a constant doubling time" is not able to describe *in vivo* growth of RENCA cells. Paradoxically, the fact that this model was not able to fit the data is precious because it shows that this theory has to be rejected and allows to infer a biological insight: the *in vivo* doubling times are not constant. We then observed that the Gompertz was able to describe these data, which means that a growth theory where the growth rate decreases is able to describe *in vivo* growth of a RENCA tumor. But another theory fitting equally well would be as well as valuable. We have to keep in mind that theoretical biology approaches using mathematical model allow to infer biological insights when rejecting theories. It then leads to suggest other or supplementary hypotheses that are tested experimentally. In this thesis, we used a data-driven modeling approach in order to challenge standard theories and suggest other ones that could be tested experimentally.

3.3 A review of mathematical models for cancer research

3.3.1 Mathematical models for tumor growth

As explained by Araujo and McElwain in a review (2004) [61], one of the first uses of mathematics to understand tumor growth had started in 1932 by Mayneord [62]. The author first observed a linear tumor growth of rat sarcoma submitted to X-radiations. He second observed that proliferation after radiation was limited to a thin rim at the tumor surface. Using a mathematical model of tumor growth, he then observed that exponential growth is obtained when a constant fraction of the all cells of the tumor are proliferating. He then deduced that a linear growth could be explained by a proliferating rim becoming more and more thin at the tumor periphery. Another substantial review has been published by Byrne, where a time line of mathematical models for tumor growth is presented [63]. The purpose of the following short review is not to be exhaustive, but to present some of the main mathematical models of tumor growth. The review is divided into different classes of mathematical models of tumor growth. Some examples of modeling studies are presented for each class.

Tumor volume and tumor cell population kinetics models

Ordinary differential equations based models like Gompertz model have known some success to predict tumor volume dynamics in mouse data [64, 65]. Concerning the clinics, Norton used the Gompertz model to describe breast cancer dynamics, with good results showing a potential of this kind of model to be used to forecast clinical tumor growth [58]. Michor et al. also illustrated the potential of mathematical models for clinical purposes in proposing a model of the chronic myeloid leukemia dynamics, describing the dynamics of several populations of cells: stem cells, progenitor cells, differentiated cells and terminally differentiated cells [55]. After calibration of the model on clinical data, simulations suggested that resistance of the treatments (imatinib) are exhibited by the stem cells subpopulation, illustrating how potentially powerful mathematical models could be in inferring information that is not contained in the data.

Mechanical models of tumor spatial expansion

In the other hand, partial derivative equations have been used to describe spatial expansion of tumors like gliomas for instance [66]. One of these models, developed by Swanson et al., has been used to assess predictions on glioma spatial invasion in the brain [67]. This model, based on a diffusion equation with a renewal term, summarizes glioma spatial expansion in two major phenomena that are proliferation of cells and migration of cells. It allows to take in account heterogeneities of the brain media (white matter and grey matter) resulting in heterogeneous migration of the spatial expansion. In another work, Preziosi and Ambrosi proposed a mathematical model of tumor spatial expansion where tissues are considered as multiphasic fluids. In this model, the expansion is not based on cell diffusion but described by a Darcy velocity, that is a gradient of pressure, the pressure resulting from the proliferation of cells [12]. Similar modeling asumptions (conservation laws, Darcy velocity) have been used by Jouganous et al. to assess predictions on clinical pulmonary metastatic lesions [13]. Bresch, Colin et al. used a similar framework to model avascular tumor growth, incorporating a viscoelastic behavior of the tumor tissue [68]. In another recent work, Perthame et al. studied, thanks to an asymptotic analysis, the passage to the stiff limit in the pressure law of state from a model describing tumor cell number density dynamics thanks to a conservation law (Darcy velocity) and a growth term. In this model, the proliferation of tumor cells depends on the medium pressure [69].

Avascular tumor growth models

In the avascular stage, a tumor grows thanks to nutrients diffusion from blood vessels and this kind of growth ends to a dormant steady state where the tumor is not growing anymore due to balance between death rate and growth rate within the tumor cell population [70, 71]. One of the first modeling attempt for describing the mechanisms of growth by diffusion has been introduced by Greespan (1972) [70]. In this theoretical study, simple mathematical models are used to investigate if dormant state in avascular tumor growth is due to metabolic wastes or necrotic core disintegration.

Spatial models of avascular growth have also been developed to describe tumor spheroid growth. Continuous and discrete approaches both demonstrated descriptive ability for tumor spheroid dynamics data [63]. Indeed, the results obtained by the continuous model from [72] and those obtained by the discrete model described in [73] both exhibited a good ability to describe the exponential then linear growth dynamics observed on multicellular spheroids.

For a much more complete review of tumor avascular growth mathematical models, the reader is referred to read the review from Roose, Chapman and Maini [74, 75].

Vascular tumor gowth, angiogenesis modeling

Angiogenesis being a critical step in tumor progression, many mathematical models have been developed to describe biological mechanisms of angiogenesis and vascular tumor growth. In 1996, an early spatial model of vascular tumor growth has been proposed by Chaplain [76]. In this paper, Chaplain introduce models of different cancer stage: avascular tumor growth, angiogenesis and vascular tumor growth, invasion and metastasis. In the model of angiogenesis, angiogenic factors-mediated chemotaxis and proliferation of endothelial cells are modeled, leading to simulation results in accordance with experimental data. In 1998, Anderson and Chaplain proposed two spatial models of vascular tumor growth [77]. The first is a continuous model where endothelial cells, stimulated by tumor angiogenic factors, migrate toward chemotactic gradients and also interact with the ECM (fibronectin). After calibrating parameter values on experimental data, simulation results are compared with *in vivo* experiments. A discrete version of this model is also proposed, allowing to perform simulation of vessel sprouting. These *in silico* neovascular networks are compared with *in vivo* vacular networks. More recently, Billy et al. proposed a spatial model of vascular tumor growth including the maturation state of endothelial cells in blood vessels, in order to assess efficiency of therapies targeting the microenvironment [78].

Other studies focusing more on the effect of endogenous factors of angiogenesis inhibition have been performed by Hahnfeldt et al. using a more phenomenological modeling approach [56]. Results of the studies shed light on promising perspectives of antiangiogenic therapies and of quantitative modeling approaches to analyze systemic dynamics in cancer.

Tumor microenvironment

Some modeling studies focused more on the effects of particular elements of the microenvironment like extracellular matrix (ECM) or nutrients concentration. Anderson et al. investigated the effect of differental microenvironmental conditions (heterogeneity of the ECM, hypoxia, etc) on phenotype of cancer cells using a hybrid agent-based model. [79]. The study revealed that harsh microenvironmental conditions exert a strong selective pressure on the cancer cells, resulting in a more invasive tumor phenotype.

3.3.2 Mathematical models of the metastatic process

Whereas many mathematical models have been built to describe tumor growth dynamics for preclinical and clinical purposes, relatively few efforts have been devoted to metastatic development modeling, although metastases are responsible of 90% of cancer related deaths. Because the metastatic cascade is a very complex process, mathematical models for metastasis often focus on a precise step of the cascade (acquiring an invasive phenotype, growth of established metastases, etc), as mentioned by Scott et al. in a recent review [80]. The goal of this short review is not at all to be exhaustive, but rather to present some of the main models focusing on different steps of the metastatic cascade and that have been used to infer different kinds of biological or clinical insights. Because, the dynamical aspect is very important in the metastatic disease, deterministic models principally based ODE systems have been built to describe the metastatic development. But on the other hand, because of the stochastic nature of cell scale events like mutation development or cell intravasation, several stochastic models have also been used to study the metastatic process.

Modeling the metastatic dissemination to predict the number of metastatic foci

In one of the first modeling works for describing the metastatic cascade, Saidel, Liotta and Kleinerman modeled the main steps of the process: primary tumor growth, angiogenesis and dynamics of the vascular surface of the primary, intravasation, circulating tumor cells flow, extravasation and formation of new metastatic colonies [81]. This model does not describe the growth of metastatic foci but only the birth time of each macrometastasis. In this work, experiments of mouse fibrosarcoma-induced pulmonary metastases allowed to calibrate the model and to assess the descriptive power of the model for each step of the metastasis cascade. The model, able to describe observations in the unperturbed system, was also able to reproduce experimental observations when perturbations like tumor trauma and tumor amputation occur. Predictions of the model are also studied in the case of other perturbations like lung vessel damages and inhibition of vascularization.
Liotta et al. also proposed a stochastic model of metastasis formation based on Markov chains [82]. In this model, metastatic foci do not only birth from single tumor cells but also from circulating cell clumps that arrest in capillary beds. In this work, the authors assessed good agreement between model and data in terms of number of metastasis and the time evolution of the probability to have no metastasis. The study goes further beyond with simulation of surgical removal of the tumor in the cases where a treatment is applied or not. The treatment consists in dislocating the tumor cell clumps. The results suggest that dislocating tumor cell clumps into single cells does not substantially improve the cure but reduce significantly the number of metastatic foci.

Modeling the metastatic dissemination as an evolution process

Other stochastic models have been built to study the metastatic process. One of them has been developed by Michor et al. in [54] to simulate Darwinian competition between tumor cells that acquired metastatic potential or not, thanks to a stochastic process called Moran process [83]. In this work, it is assumed that one mutation is responsible for acquiring the metastatic potential and that this mutation can confer to the tumor cells either advantageous or disadvantageous feature within the tumor. In this study, Michor et al. were wondering if metastatic potential is the property of cells of the main tumor or of a small subset of cells. One of their findings was that under their assumptions (one mutation responsible for the metastatic potential, constant number of cells in the tumor), cells that acquired metastatic potential are likely to have advantageous features within the tumor (high proliferation rate, low death rate), leading ultimately to tumors essentially composed of malignant cells.

Models that describe the metastatic size distribution

Other kinds of models describe the complete metastasis size distribution dynamics. The output of these models can be represented by an histogram of the type of the one presented in Fig 3.7 As explained by Scott et al. in [80], when a patient exhibits a visible metastatic lesion, it is difficult to know if other metastatic lesions are present under the visibility threshold of the used imagery technique.

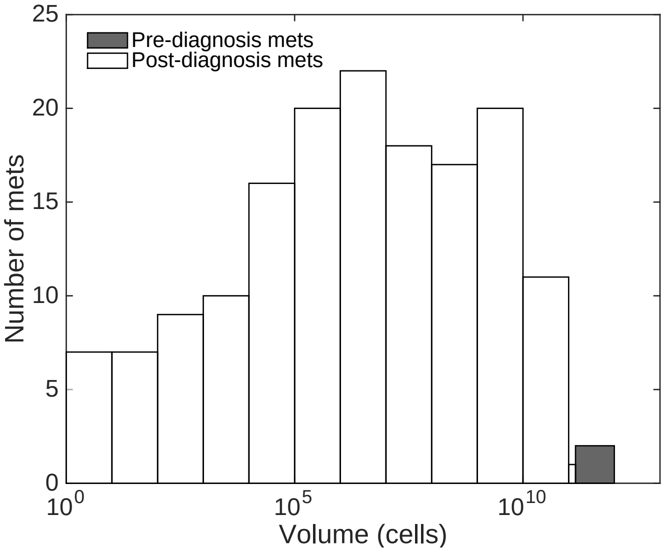


Figure 3.7: Histogram of a metastatic size distribution

In 2000, Iwata et al. built a continuous model describing a density of metastatic tumors distributed with respect to the metastatic sizes [17]. The output is hence a histogram of the same type as in Fig 3.7. In this model, the density of metastases follows a conservation law describing the transport of the density by a velocity field which is the growth law of the metastatic foci. This model has been used on a patient data to estimate the invisible metastatic burden from the visible metastatic lesions in the liver. This model has been recently validated to describe total metastatic burden dynamics from bioluminescence data of xenograft mouse breast cancer [25, 26] and syngeneic kidney cancer mouse model [26]. Similar studies have been performed using stochastic models by Bartoszyinski et al. [84] and Hanin et al. [85], where the authors validated their models against a single patient's data. Stochastic modeling techniques have also been used by Michor and Haeno to predict survival of patients with metastatic pancreatic cancer [86]. This model exhibited excellent abilities to predict the survival of a cohort of patients and has been used to assess treatment strategies. The study suggested a better efficiency of chemotherapy alone as compared to surgery of the primary alone.

Modeling the metastasis dormancy

Other studies have also been devoted to modeling the metastatic dormancy, more precisely the surgery-mediated dormancy interruption [87, 88] or systemic inhibition of angiogenesis causing a global metastatic dormancy [48]. In the last mentioned study, the Hahnfeldt model [56] is used for the growth of the metastases and the primary to model angiogenesis inhibition that a tumor exerts on all other ones. It illustrates that metastatic dynamics could result from these interactions and that latent stages of the disease could result from such global dormancies.

Modeling the blood network to assess the likelihood of organs as target sites

Other studies taking into account anatomical parameters of the patient have been performed. In [89, 90], Scott et al. proposed a model of the vascular network, where circulating tumor cells cross the different organs that have respective filtration capacities. For a specific patient, if the different filtration capacities could be measured, the model could be parametrized and assess the likelihood of future metastasis development in each organ.

Modeling tumor self-seeding

In another study [91], Scott et al. used a similar approach combined with simple growth laws to challenge the theory of Larry Norton where primary tumor growth is accelerated by reseeding of circulating tumor cells that leave and go back to the primary in crossing all the blood network or by emission from the metastatic tumors to the primary [50]. In their study, Scott et al. showed that secondary reseeding (from metastatic tumors) is more likely to accelerate primary tumor growth than primary reseeding (from the primary and back to the primary through the blood flow).

Modeling the metastatic microenvironment

Other studies have been focused on the role of the microenvironment for metastasis establishment. In [92], Araujo et al. studied the role of osteoblasts and osteclasts in the development of prostate cancer bone metastases. The model, able to reproduce the physiological stage where homeostasis in the osteoblastic and osteoclastic activities allow bone remodeling, has also been used to simulate the effect of a prostate tumor cell on this ecosystem. Simulations, in accordance with experimental data, showed how one tumor cell can deregulate this homeostasis and lead to a vicious cycle where tumor cells stimulate osteogenesis.

3.4 Conclusions

In this part, we discussed about the purpose of using mathematical models in cancer research, and more particularly how data can drive mathematical modeling, which in turn drive new experiments. As it can be noticed, a huge number of mathematical models for cancer have already been proposed and confronted with different kinds of data for different purposes like biology, tumor growth forecasting for patients or simulations of treatments scheduling. In this thesis, we adopted a data-driven modeling approach in order to infer biological insights on the biology of metastasis. Given that many kinds of data can be provided by experimentalists, a methodology to treat these data is needed. This is the purpose of the next chapter.

Chapter 4

Data assimilation in cancer modeling

Many different kinds of experimental techniques can be used to follow tumor dynamics in living systems. It gives rise to different kinds of data that require different kinds of statistical treatments to be analyzed. In this chapter, we will first deal with the different measurement techniques to follow tumor growth dynamics and second with the data assimilation techniques that are adapted for these different types of data.

4.1 Different types of data

4.1.1 Different biological systems to study the cancer biology

Tumor growth can be studied in different contexts: *in vitro*, *in vivo* and in patients. Each system has its advantages and drawbacks. Fig 4.1 illustrates the advantages and drawbacks of the different experimental systems:

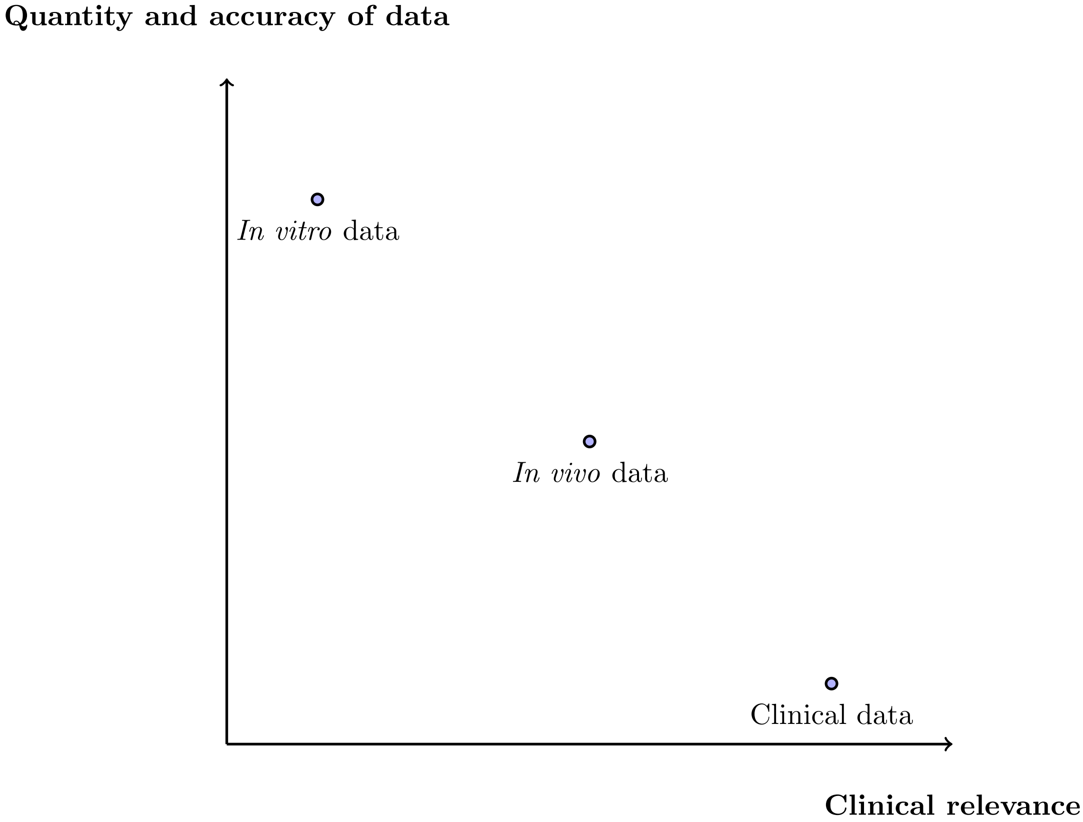


Figure 4.1: Advantages and drawbacks of the different kinds of data

In vitro systems

In vitro systems like 2D monolayers and 3D spheroids allow to get large amounts of observable quantities with a good accuracy. For instance, it is possible to observe the individual proliferating cells and quiescent cells. Moreover, *in vitro* experiments are useful to study the individual effect of a particular factor on a system. However, *in vitro* systems are far from *in vivo* situations because many factors are lacking, especially the interactions with the microenvironment such as other cells, extracellular matrix or endocrine factors. Such factors could have a critical influence on the system and in some cases totally change the output of the experiment. That is why *in vitro* experiments are useful but *in vivo* experiments are needed to study complete biological phenomena.

In vivo animal models

In vivo animal models are very useful because they allow to observe many variables in biological systems relatively close to a human system. However, observing variables *in vivo* is more difficult that *in vitro* and sophisticate protocols (imaging, GFP tracking, etc) are needed, inducing more error measurements. Furthermore, even if *in vivo* systems allow to study cancer that approach clinical cancers much better than *in vitro* ones, many differences are observed between mice and patients, limiting sometimes the relevance of animal models.

Clinical data

Clinical data are the most relevant data to confront to mathematical models. However, the amount of clinical data for a given patient is often limited, due to the limited number of examinations for a patient. Indeed, examinations are sometimes irradiant (CT scan, PET scan, etc), invasive (biopsies), painful for the patients, or expensive and can not be performed frequently enough to provide data that can be exploited by a model. Moreover, clinical data are often provided for treated patients. It is relevant when the goal is to simulate the effect of the treatment and propose therapeutic strategies but not to study the biology of the growth process, which requires to observe the tumor development without the bias introduced by the treatment.

4.1.2 Different experimental techniques to observe tumor growth dynamics

Several experimental systems can be used to follow the growth of a tumor. Some of the main systems are presented here:

Caliper

This method is often used to measure the volume of subcutaneous tumors in animal models. The diameters of the tumor are measured with a caliper. The volume is then approximated, considering the tumor as an ellipsoide.

Advantages
One of the main advantages of this technique is be quick and easy, allowing multiple measurements and then estimation of a measurement error. Moreover, this technique is not as expensive as can be more sophisticated techniques.

Drawbacks
However, measurements by caliper are not very accurate and are only feasible for subcutaneous or easily accessible macroscopic tumors.

Green fluorescent protein (GFP) tracking

GFP tracking consists in manipulating cells to make them express a green fluorescent protein. In order to measure the GFP expression in a tissue sample, which is in theory proportional to the number of GFP-expressing cells in the sample, a quantification of GFP expression by quantitative real-time polymerase chain reaction is performed.

Advantages
This technique has many advantages as it can be performed in many organs for many kinds of cancers, at a small scale (few cells) with a good accuracy. Moreover, it gives information on the tracked cells only. For instance, for a tumor, it can give information on the tumor cells only, which can be biologically more relevant than the volume of the tumor.

Drawbacks
Unfortunately, because the tissue or the organ has to be extracted for the quantitative real-time polymerase chain reaction, it is impossible to have longitudinal data on a same individual with such a technique. In GFP-tracking longitudinal data, one point corresponds to one individual. Another drawback is that the GFP signal does not give direct informations on the volume of the tumor but

on the tumor cells. Moreover, such a signal can be relative and is sometimes difficult to rely to a number of cells.

Bioluminescence tracking

The bioluminescence technique consists in manipulating cells so that they express luciferin to make them emit photons, in order to measure a signal that is in theory proportional to the number of luciferin-expressing cells.

Advantages
Like GFP tracking, bioluminescence can be used to track tumors in many organs for many kinds of cancers and gives information on the tracked cells only, possibly at a small scale. Moreover, unlike GFP tracking, bioluminescence allows to follow the time course of tumor cells on a same individual, and the signal is more easily reliable to a number of cells.

Drawbacks
The prize to pay for these advantages is principally large measurement errors. Furthermore, to track a metastatic burden thanks to bioluminescence, one has sometimes to remove the primary tumor because its emission eclipses the metastatic signal. Moreover, like GFP tracking, the measured signal does not give direct informations on the volume of the tumor but on the tumor cells.

Imaging

There are many kinds of imaging techniques (MRI, CT scan, PET-scan, etc). We present here only magnetic resonance imaging (MRI) and Computed tomography scan (CT scan) data. MRI consists in submitting tissues to a magnetic field in order to get a contrast due to the different relaxation times of the diverse molecules. CT scan consists in submitting the tissues to X-rays in order to measure differences in tissue density. Although these two techniques are different in many aspects, they both allow to generate 3D images, informing on the spatial distribution of the lesions.

Advantages
Imaging techniques have the advantage to give information on distinct lesions. For instance, it is possible to count the metastatic foci and obtain their individual sizes, resulting in much more information than the total metastatic burden only obtained by GFP-tracking or bioluminescence. Moreover, it gives spatial information like the shape of the lesion and its anatomical location. All these information can be monitored during the time on a same individual (patient, mouse, etc).

Drawbacks
The accuracy and detection threshold are limited to the device resolution ($\simeq 1$ mm^3/voxel for clinical images, $\simeq 0.01$ mm^3/voxel in experimental images). Moreover, it takes a long time to delineate the images. That is why automatic segmentation tools are needed.

4.2 A statistical framework for data analysis

Statistics is the science that deals with data. When data is obtained from experiments, we need to know what tendencies are brought out from the data, in short what the data “say” to go further beyond. That is why preliminary analysis has often to be performed before the modeling work and then confront these models to the data. Moreover, statistical methods have also to be used to compare simulation results with data, especially when simulations have to be confronted to data of a population of patients or animals.

4.2.1 Testing if a sample comes from a theoretical distribution

Let a sample of N data points $(x_1, .., x_N)$ that are realisations of N random variables $(X_1, .., X_N)$ assumed to be independent and identically distributed. We *a priori* do not know the distribution of $(X_1, .., X_N)$ but we want to test if the sample $(x_1, .., x_N)$ is likely to come from a theoretical distribution with a mean value μ and a standard deviation σ. If it is true, considering the random variable $\hat{\mu} = \frac{1}{N} \sum_{i=1}^{N} X_i$, we know, thanks to the central limit theorem, that $\sqrt{N}(\hat{\mu} - \mu)$ converges in distribution to a standard normally distributed variable as N tends to infinity:

$$\sqrt{N}\hat{\mu} \xrightarrow{d} \mathcal{N}\left(\mu, \sigma^2\right),$$

where $\mathcal{N}(\mu, \sigma^2)$ is the normal distribution of mean value μ and standard deviation σ. If the sample is large enough (if N is high enough) we then have:

$$\hat{\mu} \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{N}\right),$$

and then:

$$z = \frac{\hat{\mu} - \mu}{\left(\frac{\sigma}{\sqrt{N}}\right)} \sim \mathcal{N}(0, 1).$$

To test if the sample $(x_1, .., x_N)$ is likely to come from the theoretical distribution with the mean value μ and the standard deviation σ, we first compute z and then the cumulative distribution function of a standard normal distribution in the value $|z|$ that we called normcdf($|z|$), as in Matlab, and defined by:

$$\mathrm{normcdf}(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{s^2}{2}} \mathrm{d}s \quad \forall x \in \mathbb{R},$$

which is the probability that a random variable from a standard normal distribution takes a value less than x.

The value $p = 2(1 - \mathrm{normcdf}(|z|)$ is called $p-$value for the null hypothesis that z comes from a standard normal distribution, and then for the null hypothesis that the sample $(y_1, .., y_N)$ has been generated by a distribution with mean μ and standard deviation σ. The null hypothesis is generally rejected if $p \leq 0.05$.

Remark 4.2.1. *In the reality, it is very rare to know the theoretical variance σ and it is replaced by its estimator $\hat{\sigma} = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N} (X_i - \hat{\sigma_X})^2}$. Thanks to the Slutsky theorem, the convergence of $\hat{\mu}$ to the normal distribution is still true.*

Remark 4.2.2. *The previous test is only based on comparisons (between the sample and the theoretical distribution) of mean values. Other tests allow to test theoretical cumulative distributions or theoretical probability distributions against the data. The one-sample Kolmogorov-Smirnov test [93] for instance, challenges the null hypothesis that the population cumulative distribution function of the data is equal to the cumulative distribution function of a specific distribution (normal distribution, exponential distribution, etc). The χ^2 goodness of fit test [94] allows to test if the sample comes from a theoretical probability distribution (normal distribution, exponential distribution, etc) in comparing the empirical and the theoretical probability distributions.*

4.2.2 Testing if two samples come from the same distribution

It is often necessary to test if two samples come from the same statistical distribution. For example, to test if a number of cancer cells is higher in the case where mice are treated than in the case where they are not, a statistical test is necessary. The output of this kind of test is a result for the null hypothesis that the two samples come from the same distributions. Many different tests exist and are adapted for different kinds of situations (sample sizes and hypotheses on the distributions) but we just present here a particular example based on the parametric z-test.

Let two data samples $(x_1, .., x_K)$ and $(y_1, .., y_N)$ that are realisations from respectively K and N random variables $(X_1, .., X_K)$ and $(Y_1, .., Y_N)$ assumed to be independent and identically distributed, with respective standard deviations σ_X and σ_Y and mean values μ_X and μ_Y. We want to test if the sample $(y_i)_i$ and $(x_i)_i$ are likely to have been generated by the same distribution. We introduce the random variable $D = \hat{X} - \hat{Y}$. The variance of D is equal to:

$$VAR(D) = VAR(\hat{X}) + VAR(\hat{Y}) + 2COV(X, Y), \tag{4.1}$$

$$VAR(D) = VAR(\hat{X}) + VAR(\hat{Y}), \tag{4.2}$$

because X and Y are assumed to be independent. If the two samples are large enough, then we have, thanks to the central limit theorem:

$$\hat{X} \sim \mathcal{N}(\mu_X, \frac{\sigma_X^2}{K})$$

$$\hat{Y} \sim \mathcal{N}(\mu_Y, \frac{\sigma_Y^2}{N}),$$

and then, thanks to (4.2):

$$D \sim \mathcal{N}(\mu_X - \mu_Y, \frac{\sigma_X^2}{K} + \frac{\sigma_Y^2}{N}).$$

Known variances

If σ_X and σ_Y are known, then we define:

$$z = \frac{D}{\sqrt{VAR(D)}} \sim \mathcal{N}(\mu_X - \mu_Y, 1).$$

If the samples come from distributions of equal mean values, then:

$$z \sim \mathcal{N}(0, 1).$$

To test if the samples $(y_1, .., y_N)$ and $(x_1, .., x_K)$ are likely to come from the same distribution, we compute z and the cumulative distribution function of a standard normal distribution in $|z|$. The value $p = 2(1 - \text{normcdf}(|z|)$ gives then the $p-$value for the the null hypothesis that z comes from a standard normal distribution, and then for the null hypothesis that the samples $(y_1, .., y_N)$ and $(x_1, .., x_K)$ have been generated by distributions of same mean values.

Unknown variances

If σ_X and σ_Y are unknown and the two samples are large enough, we replace σ_X and σ_Y by $\widehat{\sigma_X} = \sqrt{\frac{1}{N-1}\sum_{i=1}^N (X_i - \widehat{\sigma_X})^2}$ and $\widehat{\sigma_Y} = \sqrt{\frac{1}{N-1}\sum_{i=1}^N (Y_i - \widehat{\sigma_Y})^2}$, which are good estimators of σ_X and σ_Y if the samples are large. In replacing σ_X and σ_Y by the estimators, the parameter z defined above does not follow a a normal distribution but Student distribution. That is why a Student's t-test can be used in this case. If the sample is small, X and Y have to be normally distributed but this is often difficult to verify. That is why non parametric tests such as the two-sample Kolmogorov-Smirnov test are used more often in this case.

Remark 4.2.3. *The previous test challenges the null hypothesis that the two samples come from the same distributions only on the basis of their mean values. Other tests like two-sample Kolmogorov-Smirnov test [95] compute the difference between the cumulative distribution functions of the distributions of the two sample data.*

Remark 4.2.4. *This kind of test is not only used to compare two data samples. It could also be used to compare a distribution coming from model simulations in order to compare outputs of a model with data samples.*

Remark 4.2.5. *There are also tests that are adapted for small sample but specific for normal distributions. These tests are not based on a test of normality but on a Student's t-test.*

4.3 Data assimilation techniques and model parameter estimation

4.3.1 Calibration of a mathematical model

Calibrating the parameters of a model consists in finding the parameters that make the model best fit with the data. It can be seen as an optimization problem where the functional to minimize is a function of the parameters $\theta \in \mathbb{R}^p$ of the model and is expressed as an error between the data and the output of the model.

An optimization problem

Let $(y_1, .., y_N)$ be a sample of N data points that are measured at different time points $(t_1, .., t_N)$. For all i, the data y_i is measured at the time t_i. Let f the function defined in $\mathbb{R}^N \times \mathbb{R}^p$ such that $f(t_i, \theta)$ is the output of a model at the time t_i that we want to compare with the data y_i, p is the number of parameters that have to be calibrated and the parameter space is not necessarily the entire space $\mathbb{R}^p$ but can be a subset Ω of $\mathbb{R}^p$. We now have to define an error criterion to minimize between the model and the data. A classical criterion is the least squares criterion:

$$J(\theta) = \sum_{i=1}^{N} (y_i - f(t_i, \theta))^2,$$

which is often used to fit a model on data. It can be written as the L^2 norm of the difference between the vectors $Y = (y_i)_{1 \leq i \leq N}$ and $F = (f(t_i, \theta))_{1 \leq i \leq N}$:

$$J(\theta) = ||Y - F||_{L^2}^2.$$

This function has to be minimized with respect to θ in order to minimize the error between the model and the data. The optimization problem can be written as follows: we search for a parameter set θ_{opt} such that:

$$\theta_{opt} = \text{argmin}\{J(\theta); \theta \in \Omega\}.$$

Examples of optimization algorithms

Such a function has most of the time no explicit expression depending on the parameters θ, which does not make easy to find a minimum. To solve this, many optimization algorithms have been developed, with two classes of algorithms: deterministic ones and stochastic ones. The first ones are generally fast to find a minimum but can be trapped in local minima. Stochastic models allow to search for a global minimum on the parameter space, but they are generally slower to converge. We present here an example of each kind of algorithm.

Monte Carlo algorithms
The Monte Carlo algorithm consists in random samples in the parameter space and to chose the parameter set that minimizes the criterion. It requires to define the parameter space Ω, a maximal iterations number Maxiter, and an accuracy criterion ε. One can also define an initial condition θ_0 for the parameter set. The algorithm can then be written as follows:

- Initializing: $n = 1$ $\theta_{opt} = \theta_0 \in \Omega$
- While ($n <$ Maxiter and $J(\theta_{opt}) > \varepsilon$) do
  - $\theta =$RandomSample(Ω)
  - If ($J(\theta) < J(\theta_{opt})$ then $\theta_{opt} = \theta$.
  - End if
  - $n = n + 1$

```
• End while
```

This kind of algorithm is slow to converge but it allows to cross the entire space to search for a global minimum. Moreover, it is very easy to implement and parallelize. More elaborated stochastic algorithms exist but they will not be presented here.

Gradient descent algorithms
There are several gradient methods. One of them is the method of steepest descent. It allows, from an initial point of the parameter space, to converge to a local minimum and needs the function J to be differentiable on the parameter space. It consists in moving by steps that go into the opposite direction of the gradient of J at each step. At each step, from a point θ_k, the next point θ_{k+1} is defined as follows:

$$\theta_{k+1} = \theta_k - \rho \nabla J(\theta_{k+1})$$

Where ρ is a positive step. The algorithm can therefore be written as follows:

```
• Initializing: $n = 1$, choosing $\theta \in \Omega$, initialize $D > \varepsilon$
• While ($n <$ Maxiter and $D > \varepsilon$) do
    – Compute $\nabla J(\theta)$
    – $J_{old} = J(\theta)$
    – $\theta = \theta - \rho \nabla J(\theta)$
    – $D = |J(\theta) - J_{old}|$
    – $n = n + 1$
• End while
```

It is often not possible to have an explicit expression of J with respect to the parameters and then it is often not possible to compute the gradient of J explicitly. In this case, partial derivatives of J can be computed by finite differences. More elaborated deterministic methods exist. In the previous method, the step ρ is not necessarily fixed. At each step, the criterion J can be derived with respect to the step, which leads to the optimal step gradient descent.

Remark 4.3.1. *Gauss-Newton method is another differential method that converges faster than the gradient descents when the initial condition is not far from the optimum. This method is specific for least squares minimization problems and requires that the number of parameters p is smaller than the number of data N. The Levenberg-Marquardt algorithm is an hybrid method between Gauss-Newton algorithm and gradient steepest descent [96, 97].*

4.3.2 Statistical framework for data assimilation

Individual data

Observations
Let $(t_i, y_i)_{1 \le i \le N}$ be N couples of time and observations of a quantity on a same individual or patient. y_i is the observation at time t_i. Typically, the observations can represent the time-course of the tumor volume in a mouse, or a number of metastases for instance.

Structural model
The structural model is a function that takes a time $t \in \mathbb{R}$ and a vector of parameter $\theta \in \mathbb{R}^p$ as input and gives an output that can be confronted to the data, a tumor volume in our case.
In theory, we want a structural model that describes the reality. However, a model never describes the reality but rather tries to essentialize it in a theory in order to approach this reality. Moreover, the data are not the reality because they contain measurements error due to the observation techniques

and to the experimentalist. These measurement errors are generally assumed to result from random events [98]. The observations can therefore be considered as random variables related to the structural model that defines a theory that could explain the data:

$$Y_i = f(t_i, \theta^*) + \varepsilon_i \quad \forall i = 1..N, \tag{4.3}$$

where θ^* is the parameter set that we are looking for and ε_i is the error term, comprising the measurement errors [98, 99]. The errors $(\varepsilon_i)_i$ are assumed to be independent and identically distributed. These measurements error, due to random events, are often assumed to be normally distributed: $\varepsilon_i \sim \mathcal{N}(0, \sigma_i^2)$ [98, 99].

A possible error model could be to consider the standard deviation σ_i to be constant: $\sigma_i = \sigma$. We can then define a probability density $p(Y|\theta^*)$, with $Y = (Y_1, \dots, Y_N)$ the random observations. If we consider this probability density as a function of θ with y a fixed sample which is a realization of Y, we define the likelihood $L(\theta) = p(y|\theta)$. The goal is to estimate θ^*, the parameter set that maximizes the likelihood of the data y under the theory f.

The random variables $(Y_i)_i$ being independent, we can write [98]:

$$L(\theta) = p(y|\theta) = \prod_{i=1}^{N} p(y_i|\theta).$$

Therefore, in the we have:

$$L(\theta) = p(y|\theta) = \prod_{i=1}^{N} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(f(t_i,\theta)-y_i)^2}{2\sigma^2}} = \frac{1}{\sigma\sqrt{2\pi}} e^{-\sum_{i=0}^{N} \frac{(f(t_i,\theta)-y_i)^2}{2\sigma^2}}.$$

We see that maximizing the likelihood $L(\theta)$ is equivalent to minimize the classical least squares criterion:

$$\sum_{i=0}^{N} (f(t_i, \theta) - y_i)^2.$$

Note that in the more general case of a non constant variance error model (heteroscedasticity), such as for instance a proportional model $\sigma_i = \sigma f(t_i, \theta^*)$, the likelihood writes:

$$L(\theta) = p(y|\theta) = \prod_{i=1}^{N} \frac{1}{\sigma_i\sqrt{2\pi}} e^{-\frac{(f(t_i,\theta)-y_i)^2}{2\sigma_i^2}} = \frac{1}{\prod_{i=1}^{N} \sigma_i\sqrt{2\pi}} e^{-\sum_{i=0}^{N} \frac{(f(t_i,\theta)-y_i)^2}{2\sigma_i^2}}.$$

Therefore, two possible criteria to minimize are:

$$LS(y_i, \theta) = \sum_{i=0}^{N} \frac{(f(t_i, \theta) - y_i)^2}{2\sigma_i^2},$$

$$LM(y_i, \theta) = \sum_{i=0}^{N} \frac{(f(t_i, \theta) - y_i)^2}{2\sigma_i^2} + \sum_{i=1}^{N} \log\left(f(t_i, \theta)\sqrt{2\pi}\right).$$

Practically, when the goal is to treat individual data, an estimation of the error measurement distribution has to be made if possible. In order to establish a statistical model of this error, a law for the variance can be proposed to be tested on the data. For instance, if we want to test a constant variance model, we have to test statistically (using a χ^2 goodness-of-fit test [94] for instance) if the measurement errors $(\varepsilon_i)_i$ are normally distributed. If we want to test a model of variance that is proportional to the measured data, we have to test if the normalized residuals $(\frac{\varepsilon_i}{f(t_i,\theta)})_i$ are normally distributed. In the case of tumor volume measurement for example, the first model would mean that the error measurement does not depend on the volume of the tumor, whereas in the second model, errors would be proportional to the volume. More elaborated statistical models have been used in [65] to study descriptive and predictive power of classical models of tumor growth.

Parameters identifiability and confidence interval
We consider $\hat{\theta}$ the random variable such that:

$$\hat{\theta} = \operatorname{argmax} L(\theta) = \operatorname{argmax} p(Y|\theta),$$

where Y is seen as a random variable. $\hat{\theta}$ is called the likelihood maximization estimator. Its variance gives information on the identifiability of the model [98]. This variance depends on the sparsity of the data and the sensitivity of f to the parameters. Knowing this variance allows to compute a 95% confidence interval representing a set in which a parameter value is likely to be. Sparse data or poorly sensitivity of a model to a parameter lead to large confidence interval, which means a strong uncertainty on the estimated value of the parameter. In theoretical biology, it is often well-suited to use a model endowed with a good identifiability because the goal can be to derive insights from the parameters estimated values. Identifiability can also be important in dynamics prediction of biological phenomena in order to provide a confidence interval [100, 101, 65].

If no measurement error statistical estimation is possible
Unfortunately, it not always possible to have informations on the measurement error. In this case, classical approaches can be to use a constant error model or a proportional error model. In the first case, the criterion to minimize is the following one:

$$J(\theta) = \sum_{i=0}^{N} \frac{(f(t_i, \theta) - y_i)^2}{\sigma^2},$$

where σ is constant. We can see that the minimization of J is independent of σ, which leads to a classical least squares minimization. Not knowing the value of σ penalizes us only for the goodness of fit estimation. Indeed, when in the cases where the variance of the error is known, the values of the normalized residual $\frac{(f(t_i,\theta)-y_i)^2}{\sigma^2}$ allows to compare the error of the model with the measurement error, which is a relevant way to evaluate the descriptive power of a model.
In the second case where we consider the error proportional to the data, the criterion to minimize is the following one:

$$J(\theta) = \sum_{i=0}^{N} \frac{(f(t_i, \theta) - y_i)^2}{\sigma^2 f(t_i, \theta)^2}. \tag{4.4}$$

This minimization is also independent of the value of σ and corresponds to a minimization of weighted least squares.

Population data

Observations
Again, we follow the time-course of a quantity but this time, each data point corresponds to one distinct individual. It is the case for example for GFP-tracking data when each mouse has to be sacrificed in order to perform the GFP analysis in specific organs. In this case, the time points are also denoted $(t_i)_{1 \leq i \leq N}$ where N is the number of time points, and the measure at the time point i on the individual j is noted y_i^j. For exampe, if individual are mice and that three mice are sacrificed per time point, we would have, for the time i, three measures y_i^1, y_i^2, y_i^3.
In a modeling perspective, this kind of data is not simple to treat because modeling the tumor kinetics of an individual makes sense but modeling the growth of data of several individuals is not as simple to define. Indeed, even if they give informations on tumor kinetics, these data do not reflect the growth of one tumor. Therefore it is necessary to define well the framework of data assimilation in order to calibrate a model and test theories.

Fitting on the mean values

A first approach can be to fit the model on mean values (with respect to the time points) of the data. We first write the mean value $\overline{y_i}$ at the time t_i:

$$\overline{y_i} = \frac{1}{N_i} \sum_{j=1}^{N_i} y_i^j,$$

where N_i is the number of individuals at the time t_i. The function J_{mean} to minimize is then defined by:

$$J_{mean}(\theta) = \sum_{i=1}^{N} \frac{(f(t_i, \theta) - \overline{y_i})^2}{\sigma_i^2},$$

where σ_i is the standard deviation of the error $f(t_i, \theta) - \overline{y_i}$ following a model that has to be defined.

Fitting on all the points
A second approach can be to fit on all the data points in defining the following objective function:

$$J_{all}(\theta) = \sum_{i=1}^{N} \sum_{j=1}^{N_i} \frac{(f(t_i, \theta) - y_i^j)^2}{\sigma_{ij}^2}.$$

This time, σ_{ij} represents the standard deviation of the error $f(t_i, \theta) - y_i^j$ following a model that has to be defined. However, we can see that if σ_{ij} depends on the individual j considered, the data of a same time point won't have the same weight during the fit. For instance, if the error model is proportional to the data, such that: $\sigma_{ij} = \sigma y_i^j$, the small values of a same time point will have a heavier weight than the large values.

Proposition 4.3.1. *If the error model is constant and the individual number is the same for all time points, fitting on the mean values and fitting on all the points are equivalent.*

Proof. Let us denote N_{ind} the total number of individuals and σ the constant standard deviation of the error $(f(t_i, \theta) - y_i^j)$. The objective function used to fit on all the points can be written:

$$J_{all}(\theta) = \frac{1}{\sigma^2} \sum_{i=1}^{N} \sum_{j=1}^{N_{ind}} (f(t_i, \theta) - y_i^j)^2,$$

$$J_{all}(\theta) = \frac{1}{\sigma^2} \sum_{i=1}^{N} \sum_{j=1}^{N_{ind}} f(t_i, \theta)^2 - 2f(t_i, \theta) y_i^j + {y_i^j}^2,$$

$$J_{all}(\theta) = \frac{N_{ind}}{\sigma^2} \sum_{i=1}^{N} f(t_i, \theta)^2 - 2f(t_i, \theta)\overline{y_i} + \frac{1}{\sigma^2} \sum_{i=0}^{N} \sum_{j=1}^{N_{ind}} {y_i^j}^2,$$

so that the term that depends on the parameters θ, in other words the term which is minimized, is the following:

$$\frac{N_{ind}}{\sigma^2} \sum_{i=1}^{N} f(t_i, \theta)^2 - 2f(t_i, \theta)\overline{y_i}.$$

In a similar fashion, computing the objective function used to fit on the mean values leads to:

$$J_{mean}(\theta) = \frac{1}{\sigma^2} \sum_{i=1}^{N} (f(t_i, \theta) - \overline{y_i})^2,$$

$$J_{mean}(\theta) = \frac{1}{\sigma^2} \sum_{i=1}^{N} f(t_i, \theta)^2 - 2f(t_i, \theta)\overline{y_i} + \overline{y_i}^2,$$

so that the term that depends on the parameters θ, in other words the term which is minimized, is the following:

$$\frac{1}{\sigma^2}\sum_{i=1}^{N} f(t_i,\theta)^2 - 2f(t_i,\theta)\overline{y_i},$$

which is the same term divided by the factor N_{ind} as the one for the fit on all the points. □

The population approach and nonlinear mixed-effects models

The population approach is similar to the individual approach but instead to calibrate one parameter set to describe the data of one individual, it consists in calibrating a parameter distribution reflecting the variabilities observed among the population of individuals.

Let a population of N individuals on which we observe longitudinal data. For each individual j, we consider N_j observations and we denote by y_i^j the observation on the individual j at time t_i^j.

As in (4.3), the observations are considered as random variables, but this time the parameter sets θ^j of the individual j also is a random variables:

$$Y_i^j = f(t_i^j,\theta^j) + \varepsilon_i^j \quad \forall j = 1..N, \quad \forall i = 1..N_j,$$

where ε_i^j is the error term, comprising the measurement errors. The random variables $(\theta^j)_j$ are assumed to be independent and identically distributed. This distribution is often assumed to be log-normal:

$$\theta^j \sim \mathcal{LN}(\theta_\mu,\theta_\omega) \quad \forall j = 1..N,$$

where $\theta_\mu \in \mathbb{R}^p$ is the expectation of the distribution of $\theta_\mu \in \mathbb{R}^p \times \mathbb{R}^p$ is the covariance matrix of the distribution. $(\varepsilon_i^j)_{i,j}$ are still assumed to be independent and normally distributed:

$$\varepsilon_i^j \sim \mathcal{N}(0,\sigma_{ij}^2) \quad \forall j = 1..N, \quad \forall i = 1..N_j.$$

For instance, if we consider a model of standard deviation proportional to the observation we have:

$$\sigma_{ij} = \sigma f(t_i^j,\theta^{j*}) \quad \forall j = 1..N, \quad \forall i = 1..N_j.$$

For a sample $y = (y_1^1, y_2^1..y_{N_1}^1, y_1^2,..y_{N_N}^N)$, we can then define a probability density $p(y|\mathcal{LN}(\theta_\mu,\theta_\omega))$ of the population data assuming the parameter distribution $\mathcal{LN}(\theta_\mu,\theta_\omega)$. If we consider this probability density as a function of θ_μ and θ_ω with y constant, we define the likelihood $L(\theta_\mu,\theta_\omega) = p(y|\theta_\mu,\theta_\omega)$. The goal is to estimate θ_μ and θ_ω, the expectation and the covariance matrix of the parameter distribution set that maximize the likelihood of the data y under the theory f. This non linear mixed-effect framework is explained in more details in [102]. Practically, a numerical algorithm has to be used in order to maximize the likelihood that we defined previously. A relatively fast algorithm is well-suited because of the number of model running in order to evaluate the likelihood. The stochastic approximation of expectation maximization algorithm (SAEM) is an algorithm that allows to approach the likelihood and that maximize it using a stochastic algorithm [103].

Goodness of fit

Let $(t_i, y_i)_{1\leq i\leq N}$ be N couples of time and observations of a quantity during the time. Again the output of the model at the time t_i for the parameter set θ is noted $f(t_i,\theta)$. Once the model is calibrated on the data, it is important to evaluate how good the model describes the data. To do so, one can simply observe how a model curve fits the data curve. However, in order to compare descriptive abilities of different models, a quantitative criterion that is a good indicator of the goodness of fit is sometimes required.

Least squares criterion

As we minimize the weighted least squares criterion J defined by the equation (4.4) in order to calibrate the model, a natural goodness of fit criterion is the value of J. However, differences of J to compare the descriptive capabilities of different models are not easy to interpret.

Coefficient of determination
The coefficient of determination, denoted R^2 is a measure of goodness of fit and is defined as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{N} (y_i - f(t_i, \theta))^2}{\sum_{i=0}^{N} (y_i - \bar{y})^2} \tag{4.5}$$

This criterion quantifies how much better the model describes the data than the mean value of the data. If $R^2 = 1$, it means that the model fits perfectly well, whereas if $R^2 \leq 0$ it means that the mean value of the data better describes them than the model does. This criterion is often used to evaluate the goodness of fit of a regression in statistics. It can be seen as the fraction of unexplained variance, as the second term is the ratio between the variance of the error of the model and the total variance of the data.

Distribution of the residuals
Let us recall that the calibration of the model is performed by minimizing the following criterion

$$J(\theta) = \sum_{i=0}^{N} \frac{(f(t_i, \theta) - y_i)^2}{\sigma_i^2},$$

where σ_i represents the standard deviation of the error, often modeled by a normal distribution. Once the model is calibrated on the data, the distribution of the normalized residuals $\frac{y_i - f(t_i, \theta)}{\Sigma_i}$ can be examined and a statistical test of normality can be made for this distribution. A normal distribution of the residuals is a good indicator for the goodness of fit. Indeed such a test eliminates models that exhibit biases like “the model output is always higher/lesser than the data points”. Moreover, if the normalized errors shows a standard normal distribution, it means that the residuals $y_i - f(t_i, \theta)$ are comprised in the observation errors and that a better fit can not be obtained

Remark 4.3.2. *Other goodness of fit criteria like the Akaike information criterion take into account the number of parameters of the model but we won’t introduce them in this thesis.*

4.4 Conclusion

In order to calibrate a mathematical model on specific data, a statistical framework has to be defined for the estimation of the model parameters in order to know what does the calibrated parameter value mean, what confidence can be expected from this value and what confidence can be expected from the predictions. The framework can also depend on what is expected from the model: prediction on one patient, comparison of parameter distribution between two groups, etc. In this chapter, we presented the frameworks that have been used in this thesis to treat the different kinds of data we obtained from different experimental techniques (imaging and GFP-tracking).

Part II

Mathematical modeling of tumor growth

Chapter 5

A pressure-mediated spatial tumor growth model: mathematical analysis, numerical validation and data assimilation

To be submitted

This chapter deals with tumor growth modeling. We first shortly introduce some classical phenomenological models of tumor growth based on ordinary differential equations. These models, adapted to treat longitudinal data of tumor growth, are confronted to different types of data of kidney tumor growth. We then introduce a spatial model of tumor growth describing the spatial expansion of the tumor tissue thanks to conservation equations. The general model is first analyzed in a simple case which allows to compare the precision of different numerical schemes used to integrate the partial differential equations of the model. We then focus on a pressure-mediated proliferation law incorporated in this model, leading to a minimally parameterized model, able to describe the growth of individual metastatic lesions. Indeed, in a recent study, this model has been confronted to mouse data of metastatic tumors from a renal primary disease with encouraging results [1]. A part of this chapter is devoted to the mathematical and numerical analyses of the model. This new proliferation law induces a nonlinear elliptic equation on the pressure. A proof of existence and uniqueness of a weak solution of the nonlinear elliptic equation is presented here. A fixed point method has been used for the resolution, and the theoretical convergence has also been studied, as well as the numerical convergence of the complete system. Finally, the model is confronted to growth data of lung metastases in an animal kidney cancer and exhibits good descriptive power of the volume dynamics and gives some clues on the potential of spatial modeling to well describe the shape time-evolution of *in vivo* tumors.

The imaging data that have been used in this part were provided by Emeline Ribot and Sylvain Miraux from the center of magnetic resonance of biological systems (Bordeaux), that perfomed the MRI sequence used to visualize the mouse lesions. In this work, my contribution was first to propose a novel spatial mathematical model of tumor growth, describing the pressure-mediated inhibition of the tumor cells proliferation. I then performed a short benchmark of the transport solvers implemented in the C++ Cadmos library by Olivier Saut, in order to compare the efficiency of the different schemes in term of ratio accuracy/time computing in a simple case where an analytical solution can be found. I then implemented a numerical method for the resolution of the equations describing the tumor spatial equations, proved the theoretical convergence of the method, and confronted the model to the MRI data of metastatic lesions.

5.1 Longitudinal data of tumor volume: ODE models

In this section, we introduce some phenomenological mathematical models of tumor growth. They are phenomenological in the sense that they focus on simple general laws to describe tumor growth. For instance these models do not care about describing the tumor heterogeneity, cell cycle, molecular pathways, effects of immune system on tumor growth, necrosis, etc. This part deals first with ordinary differential equations-based (ODE) models that allow to simulate tumor volume kinetics. We first introduce the models and the different theories that they formalize and then we present some abilities of these models to describe tumor volume dynamics in the kidney.

Studying the tumor growth kinetics often implies to study the tumor volume kinetics. As we saw previously, many experimental protocols allow to follow the time-course of the tumor volume. GFP-tracking, caliper-using measurements and bioluminescence are part of these techniques. In this context, we need mathematical models adapted to be confronted to longitudinal data of tumor volume. Ordinary differential equations-based models are well adapted to describe this kind of data. In this section, we present some of the classical ODE-based models of tumor growth and their respective abilities to describe tumor growth dynamics. All of these models describe the dynamics of the tumor volume V starting from an intial volume V_0.

5.1.1 Some classical ODE models of tumor growth

The exponential model or Malthusian growth

Theory

An exponential growth expresses the fact that a constant fraction of cells proliferate at a constant growth rate. This kind of growth has been observed *in vitro* for 2D monolayers cell cultures [57], but also for pulmonary metastases [104, 105], although other studies shown different patterns of growth for pulmonary metastases [13]. This simple model of tumor growth is due to Malthus (1766-1834) and assumes that in a given population (in our case a cell population) the number of newborns and the number of deaths are proportional to the number of individuals in the population.

Model and parameters

Under an exponential growth, the dynamics of the volume V is governed by the following ordinary differential equation:

$$\begin{aligned} \frac{\mathrm{d}V}{\mathrm{d}t} &= aV, \\ V(t=0) &= V_0. \end{aligned}$$

The two parameters of the model are the initial volume V_0 and a, the effective growth rate. Indeed, from a population dynamics point of view, the Malthus model assumes that the rate of newborns and the death rate are proportional to the population size, which leads to:

$$\begin{aligned} \frac{\mathrm{d}V}{\mathrm{d}t} &= bV - dV, \\ V(t=0) &= V_0, \end{aligned}$$

where b is the number of newborns per individual per unit time and d the proportion of the total population of dying individuals per unit of time. The effective growth rate is $a = b - d$, which is positive in a growing population. The doubling time T of the cells is directly related to the parameter a: $T = \frac{\log(2)}{a}$. The second parameter V_0 can be fixed when the number of injected cells in a mouse is known, although a fraction of the injected cells often die or do not proliferate. In a patient on the other hand, this parameter corresponds to the first examination. Generally, we do not know when the tumor was initiated, and the initial number of cells of a tumor is still debated. Indeed, monoclonal origin of tumors is generally accepted [106], but some studies tends to show that it is not always the case [107]. As reported in [65], this model has a limited descriptive power *in vivo* (breast and lung cell lines). Taking the parameter V_0 free for the calibration makes the descriptive power much improved but the predictive power still remains limited. In the other models that are presented further, the growth exhibits an increasing doubling time, which means that the growth decelerates. As said previously, the doubling times observed on *in vivo* systems are often increasing (see [58, 108, 60] for clinical evidences).

The logistic model

Theory

Nutrients supply is necessary for the proliferation of cells. In an *in vivo* system, nutrients resources are not unlimited and the larger the tumor is, the more cells are needing nutrients, and the more these nutrients are consumed and limited. To describe that limited resources limit the growth, Verhulst proposed in 1938 a model where birth rate and death rate are respectively decreasing and increasing affine functions of the population size, the tumor volume in our case [109].

Model and parameters

The model formalizing the previous assumption is the following one:

$$\begin{aligned} \frac{\mathrm{d}V}{\mathrm{d}t} &= aV\left(1-\frac{V}{K}\right), \\ V(t=0) &= V_0, \end{aligned}$$

where the parameter K is a carrying capacity traducing the maximal capacity of the environment. The first term aV corresponds to the growth term, whereas the second one $a\frac{V^2}{K}$ expresses that some cells of the population compete with each other for nutrients or living space. The dynamics is essentially exponential for small volumes and exhibits a growth saturation when V tends to K. Analysis of the model shows the existence of an unstable equilibrium point in $V = 0$ and a stable one in $V = K$. For a positive solution, the analytical solution of this nonlinear differential equation is given by:

$$V(t) = K\frac{1}{1+(\frac{K}{V_0}-1)e^{-at}}.$$

This model is often used in ecology, to describe the dynamics of bacteria population for instance, but has also been employed to describe tumor growth [65]. However, this model exhibits limited descriptive and predictive powers for *in vivo* tumor growth, espacially in comparison with other growth laws [65].

The power law model

Theory

In vivo tumors need angiogenesis to grow beyond a diameter of 2-3 millimeters [35]. This process allows them to form their own vasculature ensuring nutrients supplying. Within the tumor, the vascular network is of fractal dimension [110, 111]. If we assume that the proliferating cells inside a tumor are the cells that are in close vicinity of the vasculature, it means that the subset of proliferating cells is of the same fractal dimension as the vasculature.

Model and parameters

A minimal model expressing this asumption was introduced by Dethlefsen et al in [112] and describes the tumor volume dynamics with the following power law:

$$\begin{aligned} \frac{\mathrm{d}V}{\mathrm{d}t} &= aV^b, \\ V(t=0) &= V_0. \end{aligned}$$

This model traduces the fact that the subset of cycling tumor cells is proportional to a power b of the volume, having therefore a smaller dimension than the tumor itself. If $b = \frac{2}{3}$ for instance, the tumor exhibits a proliferative rim at its surface, whereas if $b = 1$, the growth is exponential and the number of proliferative cells is proportional to the tumor volume. If V is expressed in number of cells, a represents the growth rate at one cell. This model has many advantages. First it gives a mechanistic description of the vascular growth in remaining simple (2 parameters). Moreover, it exhibited good descriptive and predictive abilities in a study made on *in vivo* data [65].

The Gompertz model

Theory
The Gompertz model was initially introduced by Benjamin Gompertz in 1825 to describe the probability of death, or human mortality rate, in respect with the age. Used to describe the tumor growth, the model assumes that the volume of a tumor grows with a growth rate that decreases exponentially [113].

Model and parameters
The Gompertz model is written as follows:

$$\begin{aligned} \frac{\mathrm{d}V}{\mathrm{d}t} &= a \log\left(\frac{K}{V}\right) V, \\ V(t=0) &= V_0. \end{aligned} \tag{5.1}$$

K is the maximal size that the tumor can reach, in short the carrying capacity. As the logistic model, the Gompertz model exhibits a saturation of the growth. Indeed, during the growth, the doubling time is continuously increasing and the volume asymptotically tends to K. Integrating this equation leads to the following solution:

$$V(t) = K e^{\log\left(\frac{V_0}{K}\right) e^{-at}}.$$

It can then be noticed that the equation (5.1) can also be written:

$$\begin{aligned} \frac{\mathrm{d}V(t)}{\mathrm{d}t} &= \alpha \exp(-\delta t) V(t), \\ V(t=0) &= V_0, \end{aligned} \tag{5.2}$$

with $\alpha = aK \log\left(\frac{K}{V_0}\right)$ and $\delta = a$. The formulation (5.2) formalizes the hypothesis of a growth rate that decreases exponentially. This model has been proven able to describe *in vivo* tumor growth in numerous animal experimental systems [113, 65, 114] as well as human data [58]. In [65], the model demonstrated also good predictive abilities. This model can also be written in the following form:

$$\begin{aligned} \frac{\mathrm{d}V(t)}{\mathrm{d}t} &= \alpha V(t) - \beta V(t) \log(V(t)) \\ V(t=0) &= V_0, \end{aligned} \tag{5.3}$$

where $\beta = a$ and $\alpha = a \log(K)$. A linear correlation has been shown between α and β in [115] on murine B16 melanoma and rat mammary carcinoma the correlation coefficient depending on the considered specie. The same correlation has been shown by Brunton and Wheldon for human IgG multiple myeloma [116] many different species (mouse, rat, hamster, man) [117].

5.1.2 Tumor dynamics description

This part briefly illustrates the descriptive power of the previous models for data kidney tumor growth on mice. The goal is not to highlight a best model for tumor growth, nor to evaluate the predictive abilities of these models. This would require more data and tumor types. The purpose is rather to illustrate how the growth theories we introduced previously can or can not explain two kinds of growths: tumor volume dynamics or tumor cells dynamics.

These data come from a mouse model of kidney tumor growth. RENCA cells are injected in the renal subcapsule. Two ways to follow the tumor dynamics are used here: GFP-tracking and MRI monitoring (see materials and methods in the appendices for experimental details). The first one gives information on the tumor cells only after quantifying the level of GFP expression in the kidney. The second one gives information on the tumor volume. In the two kinds, the experimental conditions are the same: tumor cell line (renal cell carcinoma), number of tumor cells injected, mouse type. Only the monitoring techniques are different, giving two distinct kinds of information. It is then interesting to know if tumor volume and tumor cells follow similar growth dynamics.

Data of renal tumor cells dynamics

GFP-tracking is used here to follow tumor dynamics in the kidney. It gives information on the tumor cells only after quantifying the level of GFP expression in the kidney.

Data

The data is presented Fig 5.1. Each data point corresponds to a distinct animal ($n = 31$ animals in total). We assume here that these data points result from a same phenomenon that is impaired by random events traducing the variability between different animals. Therefore we consider the tumor burden as a random variable which depends on time as follows :

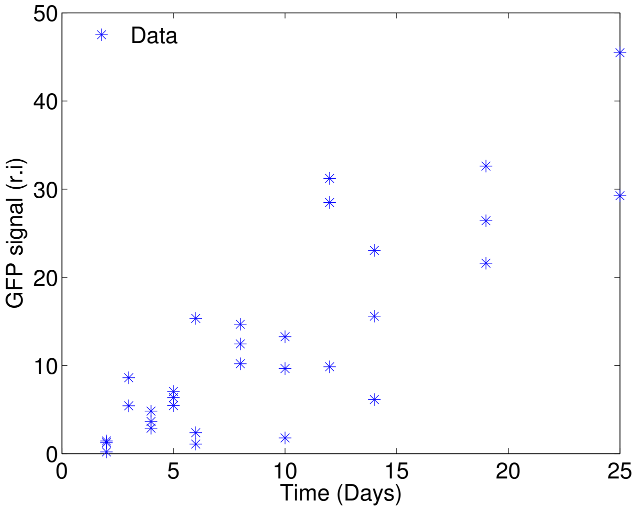


Figure 5.1: GFP data of RENCA tumor growth in the kidney. The GFP signal is proportional to the number of tumor cells

$$\begin{aligned} Y_i &= f(t_i) + \sigma_i \varepsilon_i \quad \varepsilon_i \sim \mathcal{N}(0,1) \\ E[Y_i] &= f(t_i), \end{aligned}$$

where $f(t_i)$, the structural model, represents the theoretical output of the biological phenomenon without any random perturbation and σ_i is the standard deviation of Y_i due to random events.

Error variance model

For each time point, since we have several mice per time point we approach $f(t_i)$ by the empirical mean value on the time point $\overline{y_i} = \frac{1}{n}\sum_{j=1}^{n} y_i^j$ where y_i^j is the data of the individual j at time t_i. For instance, if three data points (three mice) y_i^1, y_i^2 and y_i^3 are available for a time point i, three errors are computed, which are the difference of each data with the mean value: $E_{ij} = \overline{y_i} - y_i^j$. The

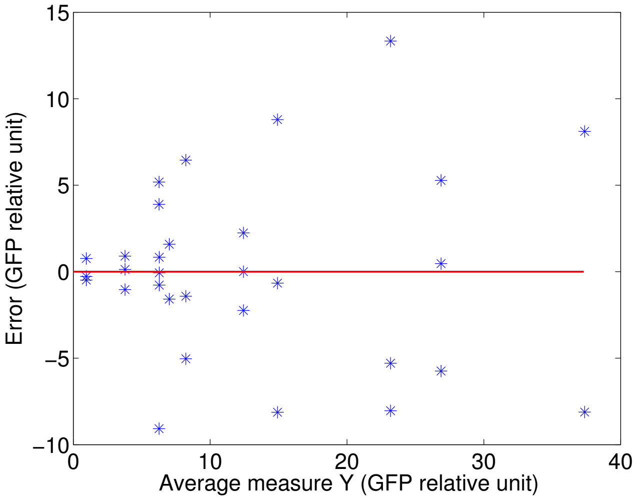


Figure 5.2: Distribution of the interindividual error against the mean (on the individuals) data

distribution of this error against the measurements (the mean values on each time point) is plotted in Fig 5.2.

Random effects are often modeled by normally distributed random variables. Under this assumption, the differences observed between mice at each time point would be normally distributed. Therefore, we tested the null hypothesis of a gaussian distribution for the error sample that are realizations of the random variability $Y_i - f(t_i)$. Under this asumption, the standard deviations σ_i would be constant: $\sigma_i = \sigma$. The χ^2 and the one-sample Kolmogorov-Smirnov tests rejects the gaussian distribution hypothesis (respectively $p = 0.001$ and $p = 0.04$). Therefore the hypothesis of a constant standard deviation may not be reasonable. In Fig 5.2, the variability seems to increase with the value of the mean. Thus, we tested the hypothesis of gaussian distribution with a mean dependant standard deviation σ_i :

$$\begin{aligned} &Y_i = f(t_i) + \sigma_i \varepsilon_i \quad \varepsilon_i \sim \mathcal{N}(0,1) \\ &\sigma_i = \sigma f(t_i) \\ &E[Y_i] = f(t_i) \end{aligned} \tag{5.4}$$

The χ^2 and the one-sample Kolmogorov-Smirnov tests did not reject the hypothesis (respectively $p = 0.15$ and $p = 0.45$). We consequently used the previous model for the variance of the error between the model and the observations.

Fitting criterion

As shown in part 4, maximizing the likelihood of the data under the model f is equivalent to minimize the least square criterion weighted with the variances:

$$LS(\theta) = \sum_{i=1}^{n} \frac{(f(t_i,\theta) - y_i)^2}{(\sigma f(t_i,\theta))^2},$$

where y_i is the mean data at day t_i and θ is the parameter vector of the model. Hence, we define the normalized residuals as following : $NE_i = \frac{(f(t_i,\theta)-y_i)}{\sigma f(t_i,\theta)}$. We can then also write :

$$LS(\theta) = \sum_{i=1}^{n} NE_i^2.$$

Fitting results

Several theories have been tested against these data:

- The exponential model
- The logistic model
- The power law model
- The Gompertz model

Fig 5.3 shows the curves of the four fitted models. In taking a look at the curves, we can see that the exponential model does not seem able to describe these data. The logistic model, although better than the exponential model, has a lower descriptive power than the Gompertz and the power law models.

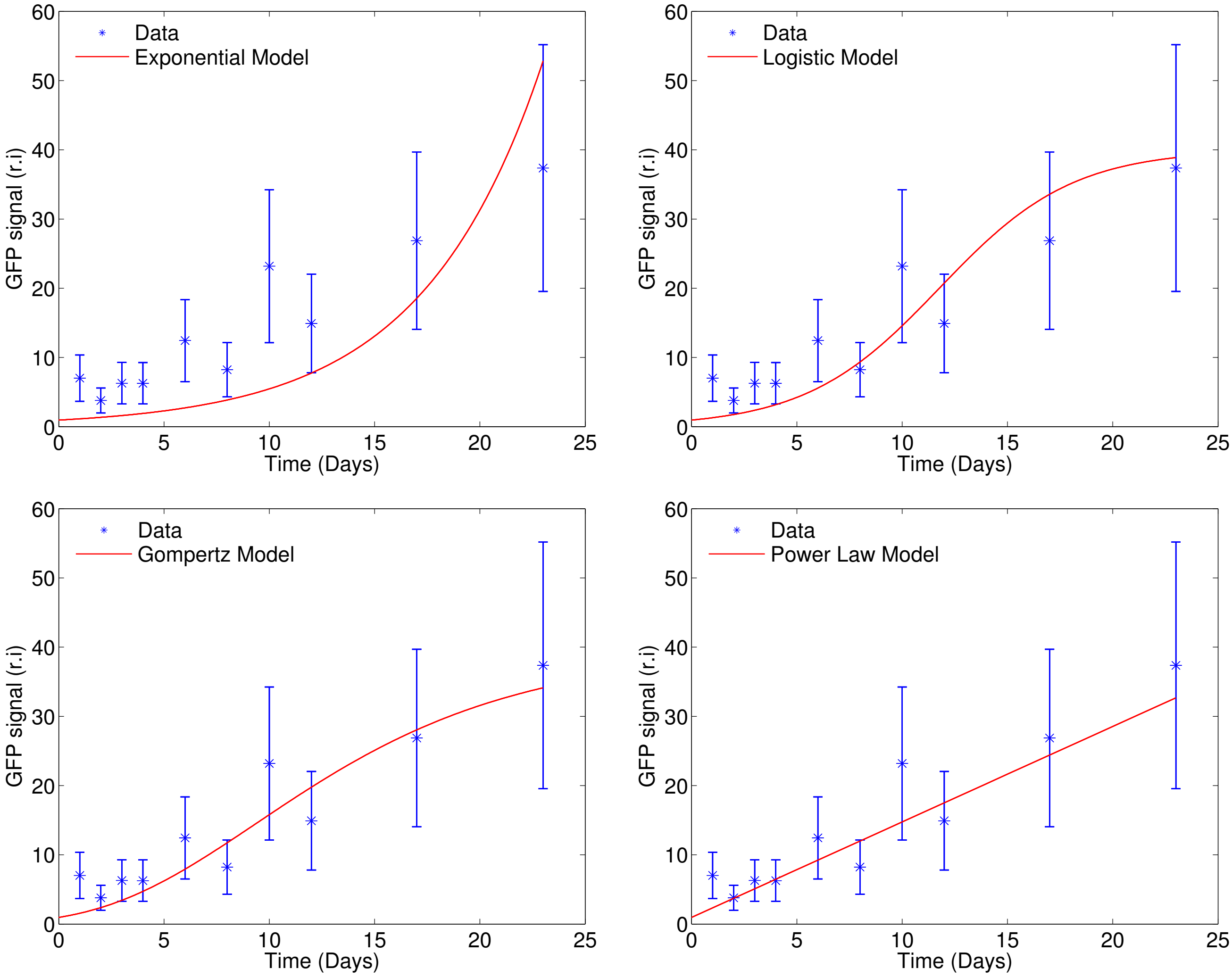


Figure 5.3: Renal tumor cells dynamics in the kidney: fitting results of the different models. Mean values ± std

Table 5.1 shows the fitting results with the value on the criteria LS, R^2, the number of normalized residuals that are higher than 1, which means that the error between data and model is higher than the standard deviation of the error model, and the result of a statistical test for the null hypothesis that the normalized residuals come from a standard normal distribution. The $p-$value for the $z-$test is precised, but the same results have been obtained using other statistical tests (one-sample Kolmogorov-Smirnov test, χ^2 goodness of fit test, $t-$test).

The results of the table seem to confirm the tendencies of the curves. The exponential model and the logistic model exhibit repsectively only 2 and 5 errors on 10 less than interanimal standard

	LS	R^2	Number of normalized residuals ≤ 1	Normal distribution of normalized residuals
Exponential model	18.14	0.21	2	No ($p < 0.01$)
Logistic model	9.93	0.76	5	Yes ($p = 0.09$)
Gompertz model	6.80	0.85	9	Yes ($p = 0.2$)
Power law model	4.16	0.82	9	Yes ($p = 0.5$)

Table 5.1: Fitting results of the four models for RENCA tumor cells growth data

deviation. Moreover the value of R^2 is much smaller for the exponential model than for the other ones. The Gompertz model exhibits the best R^2 value but the power law model shows a lower weigthed least squares criterion. These two models are those showing the best abilities to describe the data. Moreover, the exponential model is the only one to fail in the test of normal residuals, although the logistic model exhibits a poorly significant $p-$value (0.09). Taking a look at the parameter values is also interesting. In particular, for the power law model, we observe that the optimal value for the power parameter b is 0, which means that the proliferative part of the cancer cells is of null fractal dimension, which leads to a linear growth. Indeed, if we consider the growth of the tumor volume as is done in the next paragraph, it makes sense to consider that the proliferating cells have the same fractal dimension than the vasculature. But here, we consider the kinetics of the tumor cells only without taking into account the other cell types like stromal cells for instance. It could be possible that the proportion of tumor cells in respect with the tumor volume is not constant during the growth. In this case it would not make sense to consider that the vasculature is of fractal dimension with respect to the set of tumor cells.

Remark 5.1.1. *Due to the nature of the data, a mixed-effect statistical framework could be used to perform a population fit. This could be done in a future work to reinforce these results and tendencies.*

Conclusions

A theory where tumor cells are proliferating with a constant doubling time is not in accordance with dynamics of RENCA cells proliferating in the kidney. The power law theory of growth is able to describe the dynamics of tumor cells with a linear pattern of growth. It means that the proliferative subset of tumor cells is neither a constant proportion of them, neither a constant power of the set of tumor cells like the surface (proliferative rim), but rather a constant number of tumor cells. It could make sense with the theory of cancer stem cells suggesting that only a small subset of cancer cells have the infinite replication ability [118, 119, 120].

It should be interesting to test on the same animal model (RENCA cells injected in the renal subcapsule of mice) if the tumor volume follows the same growth pattern as the tumor cells. It would give information on the dynamics of the tumor stroma, which is the non tumor part of the total tumor volume and probably plays an essential role in the tumor development. This is the object of the following paragraph.

Data on kidney tumor volume dynamics

The same experimental protocol was used here. These kidney tumors were followed by MRI in the kidney, as illustrated by Fig 5.4. We delineated two kidney tumors in several time points in order to analyze their dynamics and compare them to the dynamics of GFP+ tumor cells of the previous experiment. This time, we can therefore follow the tumor time course in a same mouse. Because the segmentation was difficult and took a long time to perform, we could not do several segmentations of a same lesion to estimate a statistical model of error. We hence chose a very arbitrary error bar of 10% of the tumor volume on each data points, which is considered rather to give a relative idea of the precision of the models, not to strictly evaluate the precision of the models with respect to a precise error measurement. The fitting results are presented for the two tumors in the figure 5.5.

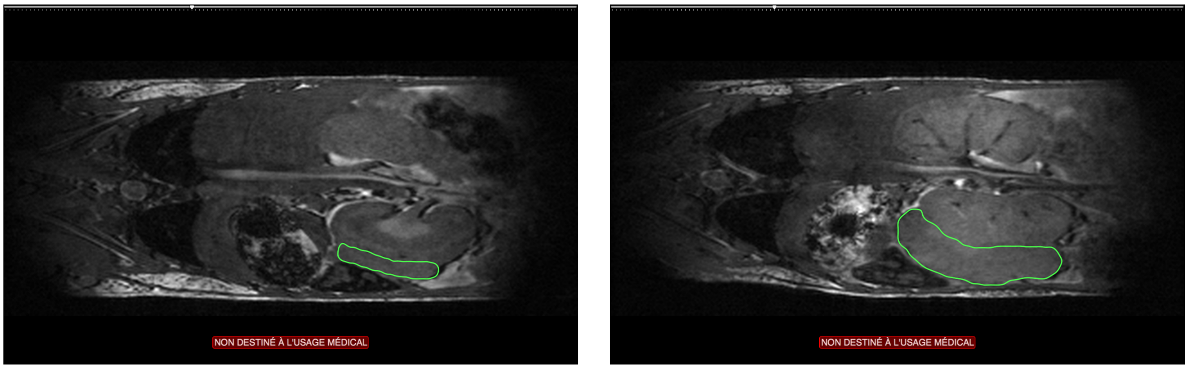


Figure 5.4: Two successive MRI images of the same RENCA tumor in the kidney. Left: day 10 after injection. Right: day 17 after injection

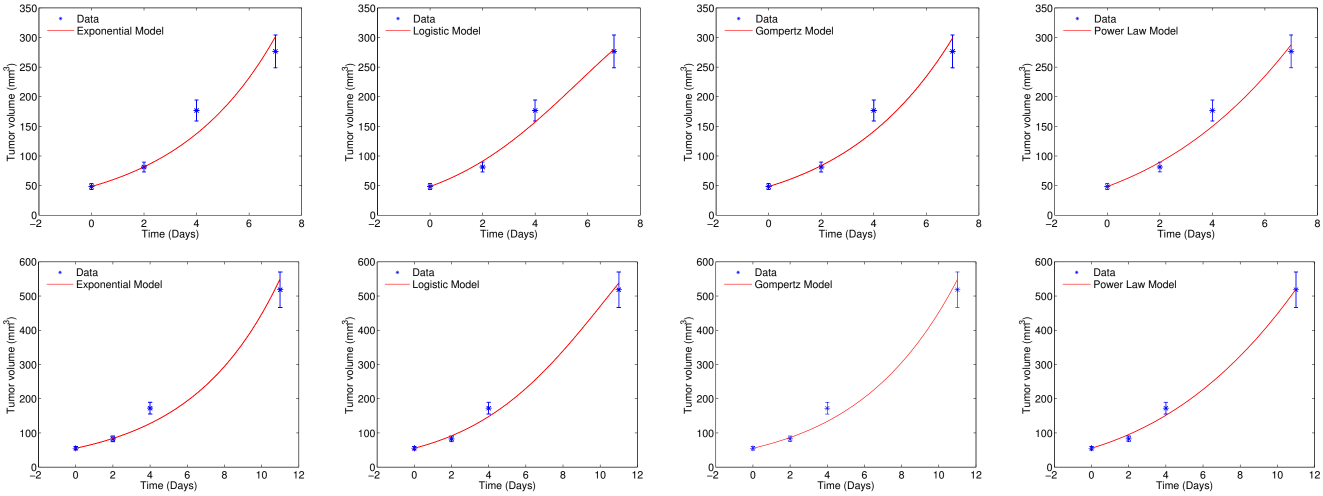


Figure 5.5: Renal tumor volume dynamics in the kidney observed by MRI: fitting results of the different models. Top: first mouse. Bottom: second mouse. From left to right: exponential model, logistic model, Gompertz model, power law model.

Table 5.2 presents the scores of the four models. This time, no statistical test has been made for the distribution of normalized residuals. Indeed, such a test would not make sense as only three points were available per mouse. Globally the exponential model is less efficient to describe *in vivo* tumor growth in the kidney. This model is nevertheless better for data on the tumor volume than it was for tumor cells dynamics. The three other models, the logistic growth law, the Gompertz growth and the power growth law are equivalent to describe this dynamics. Interestingly, the two values of the power parameter b are respectively equal to 0.68 and 0.70. These two values are very different from the null value estimated for tumor cells dynamics data. It means that the tumor volume did not follow a linear dynamics as did the tumor cells only.

Conclusions
Although describing the tumor volume dynamics better than the turmor cell dynamics, the exponential model seems to have a limited descriptive power of kidney tumor growth. The three other tested models, describing the growth with a doubling time that increases with the size of the tumor, are more able to describe kidney tumor growth, although the logistic model has less ability to describe tumor cells dynamics (GFP data) than the Gompertz and the power law models. Of course, we did not use enough data in this short study to properly establish the descriptive abilities of these models for kidney tumor growth. More mice would be necessary.

In the context of biology understanding, which is a different problematic than predictive purposes,

	First mouse			Second mouse		
	LS	R^2	Number of normalized errors ≤ 1	*LS*	R^2	Number of normalized errors ≤ 1
Exponential model	7.18	0.91	1	5.70	0.66	2
Logistic model	3.31	0.97	1	2.73	0.92	1
Gompertz model	5.56	0.93	2	4.72	0.72	2
Power law model	3.79	0.98	1	3.43	0.86	2

Table 5.2: Fitting results of the four models for kidney tumor volume growth data

models with biologically reliable parameters like the power law model are very useful. In our data, this model revealed that growth of tumor cells and growth of tumor volume are governed by different dynamics. Indeed, the calibrated value of the parameter b was equal to 0 for the tumor cells dynamics, leading to a linear growth, and equal to 0.68 and 0.70 for the two tumor volumes. Several remarks can be made on these values. First, the two values obtained with the tumor volumes are very close, indicating that this parameter possibly characterize the RENCA growth in the kidney, although it has to be confirmed by more data. Second, these values are comprised between $\frac{2}{3}$ and 1, meaning that the proliferative part of the tumor is neither the surface nor the entire tumor but a an infiltrating subset of tumor cells not limited at the surface but also not consisting in an entire core within the tumor. It makes sense with results obtained for the value of this parameter with data of *in vivo* lung and breast cancer [65]. Third, and more importantly for our point, we can therefore observe that values of b to obtain tumor volume growth are much different from those obtained to describe tumor cells dynamics for the same cell line (RENCA) and the same experimental protocol. It means that dynamics of tumor volume and tumor cells are quite different. The tumor volume growth is governed by the fractal dimension of its vasculature whereas the tumor cells dynamics seems linear. The tumor volume is constituted of the tumor cells and the tumor stroma. These results show first that RENCA tumor in the kidney can not be considered as a mass of tumor cells, as the stroma probably constitute a substantial part of the tumor. Moreover, as tumor cells and tumor volume dynamics are quite different, the stroma has probably its own dynamics and it would be interesting to investigate on this dynamics that could be dependant on the tumor cells dynamics. Further work could be devoted to study this, combining at the same time data on the tumor volume and the tumor cells dynamics.

5.2 Spatial data of tumor expansion: spatial PDE models

This section is devoted to models of spatial tumor growth based on partial differential equations. We first present a general model structure based on conservation equations to describe tumor spatial expansion. We secondly different growth laws that can be introduced in the model. It is sometimes useful to study spatial aspects of tumor dynamics. Indeed, in cancer biology, some phenomena are spatial by nature (tumor heterogeneity, migration of cells, spatial interactions between metastases, etc.). Studying this kind of phenomena requires data giving spatial information like image-based data. Moreover, studying these spatial phenomena with a mathematical modeling approach leads to incorporate a spatial dimension in the model. For example, a spatial problematic in tumor growth could be to simulate not the volume kinetics only but also the shape evolution of the tumor.

5.2.1 A spatial tumor growth mathematical model based on conservation equations

We are interested here in the spatial dynamics of tumor growth at the tissue scale. One of the main assumptions here is to consider that at this scale, the medium can be seen as a continuous porous medium. The 2D model we used for the spatial growth describes a saturated flow in a porous medium with two species: the tumor tissue and the healthy tissue, denoted respectively by P and S. The third variable of our problem is a pressure field Π. The model describes, on a domain Ω, the passive motion of the tissues due to growth of volume caused by proliferation. Under the mass-balance assumption, it gives the following equations:

$$\frac{\partial P(t,x)}{\partial t} + \nabla \cdot (\mathrm{v}(t,x)P(t,x)) = \gamma(t,x,P,S,\Pi)P(t,x), \tag{5.5}$$

$$\frac{\partial S(t,x)}{\partial t} + \nabla \cdot (\mathrm{v}(t,x)S(t,x)) = 0, \tag{5.6}$$

where γ is the growth rate of the proliferating cells. We also make the assumption of a saturated flow, that is $P + S = 1$ at each point of the domain. This hypothesis leads to the following condition on the velocity of the flow :

$$\nabla \cdot \mathrm{v} = \gamma P. \tag{5.7}$$

As we are going to show in the next part, this saturation assumption means that the tumor volume changes are due to proliferation.
We consider the lungs as a porous medium with a porosity $k(t,x)$ and that the velocity v is due to pressure gradients within the tumor. This velocity field is hence modeled with a Darcy law, as in [12]:

$$\mathrm{v} = -k\nabla\Pi, \tag{5.8}$$

where Π is the pressure field. Therefore, the velocity v is the passive motion velocity, due to the pressure exerted by the proliferative tissue on the surrounding tissues. These tissues are in this way ”pushed out” and move from the high pressure areas to the low pressure ones, as illustrated by Fig 5.6.

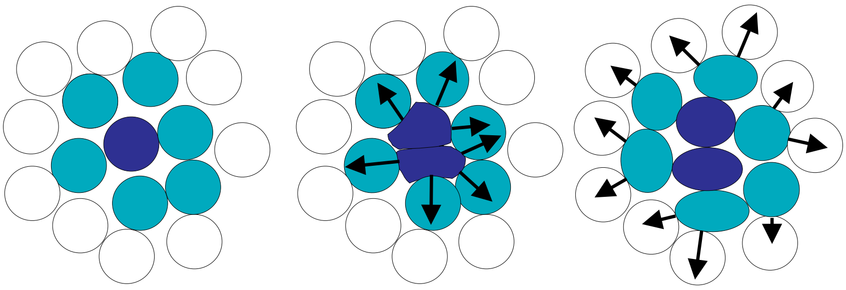


Figure 5.6: Scheme of cells in cycle that are pushing each other when they proliferate, picture from [121]

Considering metastatic growth in the lungs, the size of the computational domain Ω is fixed at the same order as mouse lungs ($\simeq 1\ \mathrm{cm}^3$). Assuming that no mechanical interactions occur with the organ boundaries (for instance, the possible deformations of the organ due to the growth are neglected here), we suppose that the domain Ω is large enough to consider the pressure on the boundaries as equal to the homeostatic pressure of the body. We model this setup by a Dirichlet boundary condition on the pressure:

$$\Pi = \Pi_{eq} \text{ on } \partial\Omega.$$

Collecting (5.7) and (5.8), and considering the porosity k constant, we obtain that Π satisfies a Poisson's equation with Dirichlet boundary conditions:

$$\begin{cases} -k\Delta\Pi = \gamma P, \\ \Pi_{|\partial\Omega} = \Pi_{eq}. \end{cases} \tag{5.9}$$

Similar models using more cellular species were used by Bresch, Colin and colleagues in [122] to describe avascular tumor growth, by Billy and colleagues in [78] to describe angiogenic tumor growth, and by Saut, Lagaert and colleagues in [123] to assess the effect of antiangiogenic drugs on glioblastoma. Here the goal is not to study hypoxia, angiogenesis nor effect of drugs on the growth, but more to find a simple way to model tumor spatial expansion allowing to capture essential mechanisms of macroscopic growth. That is why we consider here one species only, P, which is the density of tumor cells, with a proliferation rate γ, which is distributed in space, reflecting the heterogeneity of proliferation in the tumor.

Qualitative study: the characteristics

We introduce the characteristic curves $X(t, y)$ associated to velocity field v, that is the solution to:

$$\begin{aligned} &\frac{\partial X(t,y)}{\partial t} = \mathrm{v}(t, X(t,y)), \\ &X(0,y) = y. \end{aligned}$$

If we switch to Lagrangian coordinates, that is writing $\tilde{P}(t,y) = P(t, x(t,y))$, one gets the following equation on $\tilde{P}$:

$$\partial_t \tilde{P}(t,y) = \partial_t P(t, X(t)) + \partial_t X(t,y) \cdot \nabla P(t, X(t,y)), \tag{5.10}$$

$$\partial_t \tilde{P}(t,y) = -\nabla \cdot (\mathrm{v}P)(t, X(t,y)) + (\gamma P)(t, X(t,y)) + \mathrm{v}(t, X(t,y)) \cdot \nabla P(t, X(t,y)), \tag{5.11}$$

$$\partial_t \tilde{P}(t,y) = -\nabla \cdot \mathrm{v}(t, X(t,y)) \tilde{P}(t,y) + (\gamma \tilde{P})(t,y), \tag{5.12}$$

$$\partial_t \tilde{P}(t,y) = -\gamma \tilde{P}^2(t,y) + \gamma \tilde{P}(t,y), \tag{5.13}$$

$$\partial_t \tilde{P}(t,y) = \gamma \tilde{P}(1 - \tilde{P})(t,y). \tag{5.14}$$

It therefore follows that $\tilde{P}$ satisfies a logistic growth along the characteristics. This is due to the saturation hypothesis ($\nabla \cdot \mathrm{v} = \gamma P$). As a consequence, a discontinuous initial condition for P, for example the density 1 inside the tumor and 0 outside, is conserved during the growth.
One of the important things we are interested in is the kinetics law on the global tumor burden that derives from this spatial model.

Burden and volume dynamics

Integrating (5.5) on the whole domain gives rise the following equations on the tumor burden $\int_\Omega P \mathrm{d}x$:

$$\begin{aligned} &\int_\Omega \frac{\partial P}{\partial t} \mathrm{d}x + \int_\Omega \nabla \cdot (\mathrm{v}P) \mathrm{d}x = \int_\Omega \gamma P \mathrm{d}x, \\ &\frac{\mathrm{d}}{\mathrm{dt}} \int_\Omega P \mathrm{d}x + \int_{\partial\Omega} P \mathrm{v} \cdot n_\sigma \mathrm{d}\sigma = \int_\Omega \gamma P \mathrm{d}x, \end{aligned}$$

where $\sigma = \partial\Omega$ and n_σ is the outgoing normal on σ. Considering that the tumor does not reach the boundaries of the domain, the term $\int_{\partial\Omega} P \mathrm{v}.n_\sigma \mathrm{d}\sigma$ is vanishing. So the burden equation can be written as follows:

$$\frac{\mathrm{d}}{\mathrm{dt}} \int_\Omega P \mathrm{d}x = \int_\Omega \gamma P \mathrm{d}x \tag{5.15}$$

At the initial time one has $P(0,x) \geq 0$ within an open set $\Omega_0 \subset \Omega$ and $P(0,x) = 0$ for all $x \in \Omega \setminus \Omega_0$. Then for all time $t > 0$, $P(t,x) \geq 0$ within an open set $\Omega_t \subset \Omega$ and $P(0,x) = 0$ for $x \in \Omega \setminus \Omega_t$. Therefore, using the characteristics $X(t,x)$ associated to v, one has $\Omega_t = X(t, \Omega_0)$. Introducing the Jacobian transformation $J(t,y) = \det(D_y X(t,y)$, one has $\partial_t J(t,y) = \nabla \cdot \mathrm{v}(t, X(t,y)) J(t,y)$ [124].

Therefore the derivative of the tumor volume with respect to time can then be written as:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\int_{\Omega_t}\mathrm{d}x &= \frac{\mathrm{d}}{\mathrm{d}t}\int_{\Omega_0} J(t,y)\mathrm{d}y\\ &= \int_{\Omega_0}\frac{\partial J(t,y)}{\partial t}\mathrm{d}y\\ &= \int_{\Omega_0}\nabla\cdot \mathrm{v}(t,X(t,y))J(t,y)\mathrm{d}y\\ &= \int_{\Omega_t}\nabla\cdot \mathrm{v}(t,x)\mathrm{d}x\\ &= \int_{\Omega_t}\gamma P\mathrm{d}x\end{aligned}$$

Because $P(t,x) \geq 0\ \forall x \in \Omega_t$ and $P(t,x) = 0$ for $x \in \Omega \setminus \Omega_t$ we have $\int_{\Omega_t}\gamma P\mathrm{d}x = \int_{\Omega}\gamma P\mathrm{d}x$ and then $\int_{\Omega_t}\gamma P\mathrm{d}x = \frac{\mathrm{d}}{\mathrm{d}t}\int_{\Omega}P\mathrm{d}x$ due to (5.15).
As $\frac{\mathrm{d}}{\mathrm{d}t}\int_{\Omega}P\mathrm{d}x = \frac{\mathrm{d}}{\mathrm{d}t}\int_{\Omega_t}P\mathrm{d}x$, we finally have:

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\Omega_t}\mathrm{d}x = \frac{\mathrm{d}}{\mathrm{d}t}\int_{\Omega_t}P\mathrm{d}x \tag{5.16}$$

We denote by V_T the volume of the tumor, which embraces both non tumoral and tumoral cells inside the tumor, and by V_P the volume occupied by the tumor cells only (we called it the tumor burden previously). Equation (5.16) can be written

$$\frac{\mathrm{d}V_T}{\mathrm{d}t} = \frac{\mathrm{d}V_P}{\mathrm{d}t}.$$

It describes that the total volume of the tumor and the volume occupied by the tumor cells only are increasing at the same speed. This modeling asumption comes from the saturation hypothesis. In this paradigm of tumor growth, tumor cells burden and tumor volume have the same growth dynamics. In the context of *in vivo* model of tumor growth, such an asumption would be right if GFP+ (green fluorescent protein) or Luc+ (luciferase) tumor cells had the same dynamics than the tumor volume. This is not true for *in vivo* RENCA tumor growth in the kidney, as we showed in section 5.1.The validity of such an assumption probably depends on the cell line, the location (organ) and the nature of the lesion (primary tumor, metastasis).

5.2.2 Some phenomenological growth laws to model the proliferation of tumor cells

Exponential growth

If we consider a constant growth rate, $\gamma = \gamma_0$ is taken constant. In such a case, the equation (5.15) on the tumor burden becomes :

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\Omega}P\mathrm{d}x = \gamma_0\int_{\Omega}P\mathrm{d}x.$$

In such a case, the tumor burden follows an exponential growth with the growth rate γ_0. If we consider the growth all along the characteristics, the equation (5.23) becomes $\frac{\partial\tilde{P}(t,y)}{\partial t} = \gamma\tilde{P}(t,y)(1-\tilde{P}(t,y))$ and can be exactly solved :

$$\tilde{P}(t,y) = \frac{1}{1+(\frac{1}{\tilde{P}(0,y)}-1)\exp(-\gamma_0 t)}.$$

In this way, a continuous initial condition asymptotically tends to a discontinuous one, the density being 1 inside the tumor and 0 outside.

However, as previously mentioned, the *in vivo* doubling time can rarely be considered as constant. That is why we are interested to use a model with a doubling that is evolving with the environment conditions (nutrient supply, competition for nutrients, pressure conditions, etc.).

Logistic growth law

To formalize the doubling time increasing during *in vivo* tumor growth, we first assume a simple model of growth rate, which depends explicitly on the tumoral density, for example a logistic growth law : $\gamma(P) = \gamma(1-P)$. Such a growth rate depicts the competition between tumor cells for the nutrients. Under such a growth assumption, the model writes:

$$\frac{\partial P(t,x)}{\partial t} + \nabla \cdot (\mathrm{v}(t,x)P(t,x)) = \gamma(1-P)P(t,x),$$
$$\frac{\partial S(t,x)}{\partial t} + \nabla \cdot (\mathrm{v}(t,x)S(t,x)) = 0.$$

Or:

$$\frac{\partial P(t,x)}{\partial t} + \nabla \cdot (\mathrm{v}(t,x)P(t,x)) = \gamma S(t,x)P(t,x), \tag{5.17}$$
$$\frac{\partial S(t,x)}{\partial t} + \nabla \cdot (\mathrm{v}(t,x)S(t,x)) = 0. \tag{5.18}$$

The growth rate is written here as γS, which expresses that nutrients and vasculature are contained in the healthy tissue or that the growth needs a supportive stroma that is composed of healthy cells. However, under such a model, the growth is strongly dependent on the initial distribution. For instance, if the distribution is 1 inside the tumor and 0 outside, which could corresponds to a compact metastasis, there is no growth. That is why we need to enrich the model and add some biology in the proliferation process, considering for example pressure or vasculature aspects.

Gompertz-like growth

As we previously mentioned, the Gompertz model has a good ability to describe *in vivo* tumor growth for many cancer types. That is why we seek a spatial model that exhibits a Gompertz growth when integrated. In order to perform this, we introduce a resource variable M which contains nutrients, growth factors and all other factors needed for the growth. The resources are considered as spatially homogeneous and follow the following decreasing dynamics:

$$\frac{\mathrm{d}M}{\mathrm{d}t} = -\delta M,$$
$$M(0) = M_0,$$

so that $M(t) = M_0 \exp(-\delta t)$. Taken as simple as possible, the growth law is considered as a linear function of the resources:

$$\gamma(M) = \gamma_0 M.$$

The complete system is written as follows:

$$\frac{\partial P(t,x)}{\partial t} + \nabla \cdot (\mathrm{v}(t,x)P(t,x)) = \gamma_0 M P(t,x), \tag{5.19}$$
$$\frac{\partial S(t,x)}{\partial t} + \nabla \cdot (\mathrm{v}(t,x)S(t,x)) = 0, \tag{5.20}$$
$$\frac{\mathrm{d}M(t)}{\mathrm{d}t} = -\delta M(t). \tag{5.21}$$

When integrating the equation verified by the tumor density P and assuming that the tumor never reaches the domain boundary, we get:

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\Omega} P(t,x)\mathrm{d}x = \gamma_0 M_0 \exp(-\delta t)\int_{\Omega} P(t,x)\mathrm{d}x.$$

We can see that the tumor burden $B(t) = \int_{\Omega} P(t,x)\mathrm{d}x$ follows the Gompertz growth law because the growth rate is is decreasing exponentially, as in (5.2).

Hypoxia-mediated proliferation

It is commonly accepted that lack of nutrients is one of the limiting factors for tumor growth. That is why we introduce here a vasculature variable M which represents the nutrients concentration on the domain, and varies between 0 and 1. The growth rate is considered depending on this concentration in a linear fashion: $\gamma(M) = \gamma_0 M$. Furthermore, the nutrients are consumed by the tumor cells in the following fashion: $\frac{\partial}{\partial t}M = -\delta PM$

$$\begin{aligned}
&\frac{\partial P(t,x)}{\partial t} + \nabla \cdot (\mathrm{v}(t,x)P(t,x)) = \gamma M(t,x)P(t,x),\\
&\frac{\partial S(t,x)}{\partial t} + \nabla \cdot (\mathrm{v}(t,x)S(t,x)) = 0,\\
&\frac{\partial}{\partial t}M(t,x) = -\delta P(t,x)M(t,x).
\end{aligned}$$

This model principally contains three parameters that are γ, the maximal proliferation rate, $M(0)$ the initial vasculature (if we consider that the vasculature is homogeneous initially) and δ, the consumption rate of nutrients by the tumor cells.

Pressure-mediated proliferation

Hanahan and Weinberg put forward contact inhibition between cells in [4] to ensure tissue homeostasis. They highlight the capacity of cancer cells to ignore these inhibition signals as a hallmark of cancer. In [15], Stylianopoulos et al showed that tumor cells, in proliferating uncontrollably, induce mechanical stresses in surrounding microenvironnement of murine and human tumors. Furthermore, Stylianopoulos showed in [16] that exerted pressure impairs the proliferation in two ways: the direct effect of the pressure on the cancer cells growth and the collapsing effect of the pressure on the blood vessels, resulting in a lack of nutrients for the cancer cells. Based on these considerations, it seems relevant to model a proliferation law depending on the pressure.

In the previously presented model, the tissues motion is directed along the pressure gradients. It means that cancer cells proliferate and that the exerted pressure pushes out the neighboring tissues. This pressure is not solely due to mechanical constraints (solid stresses, interstitial fluid pressure, etc.) exerted by the neighbouring cells on each other, but also represents a more phenomenological pressure, that reflects the basic assumption of our modelling strategy for the tumour tissue being constituted by a fluid mixture in a porous medium. Other studies based on numerical simulation also have been performed to investigate the effect of the pressure on proliferation. In [125], Montel et al discussed the fact that cells proliferate faster on the surface than in the bulk of a tumor spheroid. A classical reason is that nutrients do not penetrate deeply in the spheroid. But Montel suggests a mechanical effect due to the necessity for a cell to deform its environment in order to proliferate. Drasdo and Höhme have performed a study on 2D monolayers and 3D spheroids tending to show that pressure conditions have a higher impact on doubling time than nutrients lacking [126]. Based on this framework, a reasonable assumption is to consider that the doubling time of the tumor cells increases with the pressure. Moreover, in [125], Montel et al performed experiments where tumour cells were submitted to different pressure constraints and observed a decrease in proliferation when pressure was applied. In their study, simulation results that were compared to experimental ones showed an exponential decreasing of proliferation with pressure. However, the bulk and surface division rate were not affected equally by stresses. In our model, we use a pressure-mediated proliferation law expressing direct effects of mechanical stresses on proliferation as well as indirect effects of proliferation on the micro environment (collapsing of blood vessels leading to lack of nutrients):

$$\gamma(\Pi) = \gamma_0 \exp\left(-\frac{\Pi}{\Pi_c}\right), \tag{5.22}$$

where Π represents the pressure field, γ_0 the maximal proliferation rate, and Π_c a characteristic pressure of proliferation inhibition.
Under these modeling assumptions, high pressures imply proliferation decreasing, but not apoptosis, consistently with Montel et al. [125].

5.2.3 Confronting spatial tumor growth models with tumor growth data

As ODE-based models, PDE-based spatial models allow to simulate tumor volume kinetics but they also make it possible to simulate the spatial expansion, including the time evolution of the tumor shape, which is not possible with ODE-based models. However, the prize to pay in terms of computing cost and mathematical complexity are sometimes very high. Before performing any data assimilation, such a model has to be analyzed mathematically to be sure of the well-posedness of the equations, and numerical methods, sometimes very sophisticated, have to be developed in order to simulate the model. Moreover, optimization algorithms need sometimes a lot of model runs to find an optimum and the computational costs of such PDE-based models limits sometimes dramatically the possibility to fit the model on data.

5.3 A simple case for numerical validation: constant growth rate

The goal of this section is to chose the most adapted numerical scheme to solve the equations (5.5)-(5.9). To do so, we considered a simple case for which an analytical solution can be exhibited, allowing to evaluate the accuracy of the numerical schemes. The ratio of accuracy over computing time of different transport schemes have been compared. It appeared that the WENO5 scheme was the more efficient. Convergence tests have also been made but the goal was only to test the convergence of the schemes and not to reach high orders of convergence. Indeed, reaching a second order of convergence would require to recompute the velocity field v at each step of the splitting scheme (5.24), requiring to solve a Laplacian on the domain at each step. It would be possible but very costly in computing time, whereas we do not need here to capture local phenomena like shocks requiring high-order-accurate-in-space schemes. This is why we decided to stay at a first order of convergence.

We first consider a constant growth rate, that is γ is taken as a constant: $\gamma = \gamma_0$. The equation (5.15) on the tumor burden becomes :

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega P \mathrm{d}x = \gamma_0 \int_\Omega P \mathrm{d}x.$$

In such a case, the tumor burden follows an exponential growth with the growth rate γ_0. If we consider the growth all along the characteristics, the equation (5.23) becomes $\frac{\partial \tilde{P}(t,y)}{\partial t} = \gamma \tilde{P}(t,y)(1-\tilde{P}(t,y))$ and can be exactly solved :

$$\tilde{P}(t,y) = \frac{1}{1+\left(\frac{1}{\tilde{P}(0,y)}-1\right)\exp(-\gamma t)}. \tag{5.23}$$

Therefore a non zero continuous initial condition tends to the constant solution equal to 1 in every point of Ω. Moreover, as said previously, a discontinuous solution with a density equal to 1 inside the tumor and 0 outside is transported during the growth: The final density is still equal to 1 inside the tumor and 0 outside.

5.3.1 Numerical scheme

A well adapted numerical scheme has to be chosen to solve the system (5.5)-(5.8). We consider a square domain and we define a cartesian spatial grid on this domain. A time discretization is also defined: $t^n = n\Delta t^n$ where Δt is the discretization time step. We define the sequences $(P^n)_n$, $(\mathrm{v}^n)_n$ and $(\Pi^n)_n$ by:

$$\begin{aligned} P^n(x) &= P(t^n, x) \ \ \forall n, \\ \Pi^n(x) &= \Pi(t^n, x) \ \ \forall n, \\ \mathrm{v}^n(x) &= \mathrm{v}(t^n, x) \ \ \forall n. \end{aligned}$$

Global scheme

At a time t^n, for a given distribution P^n, we can compute the remaining variables of the model: first we determine Π^n, then v^n and finally P^{n+1} with an explicit method in time.
The first step consists in discretizing the Laplacian operator on the spatial grid in order to solve the problem (5.9).
It provides us a second-order in space approximation of Π^n. From $\mathrm{v}^n = \nabla \cdot \Pi^n$, an approximation of the velocity field v^n is then computed at the second order too with a centered finite differences method. Now that the velocity field is determined, the equation on P can be solved. We first write it as:

$$\begin{aligned} \frac{\partial P(t,x)}{\partial t} + \mathrm{v}(t,x)\cdot\nabla P(t,x) &= \gamma_0 P(t,x) - \nabla.\mathrm{v}(t,x)P(t,x), \\ \frac{\partial P(t,x)}{\partial t} + \mathrm{v}(t,x)\cdot\nabla P(t,x) &= \gamma_0 P(t,x)(1 - P(t,x)). \end{aligned}$$

This equation is solved by a splitting method, as it has been done in similar models [127, 122]. As in [127], a Strang spitting is chosen here, which writes, from the P^n at the time t^n:

$$\begin{aligned} &\begin{cases} \frac{\partial \tilde{P}(t,x)}{\partial t} = \gamma_0 \tilde{P}(t,x)(1 - \tilde{P}(t,x)) \ \ on\,]0; \frac{\Delta t}{2}[, \\ \tilde{P}(0,x) = P^n, \end{cases} \\ &\begin{cases} \frac{\partial P^*(t,x)}{\partial t} + \mathrm{v}(t,x).\nabla P^*(t,x) = 0 \ \ on\,]0; \Delta t[, \\ P^*(0,x) = \tilde{P}\left(\frac{\Delta t}{2}, x\right), \end{cases} \\ &\begin{cases} \frac{\partial P(t,x)}{\partial t} = \gamma_0 P(t,x)(1 - P(t,x)) \ \ on\,]0; \frac{\Delta t}{2}[, \\ P(0,x) = P^*(\Delta t, x). \end{cases} \end{aligned} \tag{5.24}$$

The first and third logistic equations can be solved exactly. Indeed, these equations have respectively for solutions $\tilde{P}(t,x) = \frac{P^n(x)}{P^n(x) + (1-P^n)\exp(-\gamma_0 t)}$ and $P(t,x) = \frac{P^*(\Delta t, x)}{P^*(\Delta t,x) + (1-P^*(\Delta t,x))\exp(-\gamma_0 t)}$.

Transport solver

We have now to chose a transport scheme for the second step of the splitting. We tested five schemes of transport with different spatial orders of accuracy:

- A first order upwind scheme
- A second order semi lagrangian scheme [128, 129]
- A second order TVD Lax Wendroff scheme
- A fifth order WENO scheme (with a forward Euler time scheme) [130, 131]
- A fifth order WENO scheme (with a TVD third-order Runge-Kutta scheme) [132]

5.3.2 Test cases: numerical results

Convergence

We tested the convergence of the scheme for a smooth initial condition (gaussian bell) and for non regular one (square pulse). First order of convergence was obtained at every point of the domain for the smooth initial condition and with respect to the L^1 norm for the non smooth initial condition.

A smooth initial condition

Fig 5.7 shows the evolution of the errors in repect with mesh size for the smooth initial condition, which is a gaussian bell. The table 5.3 shows the convergence orders of each scheme.

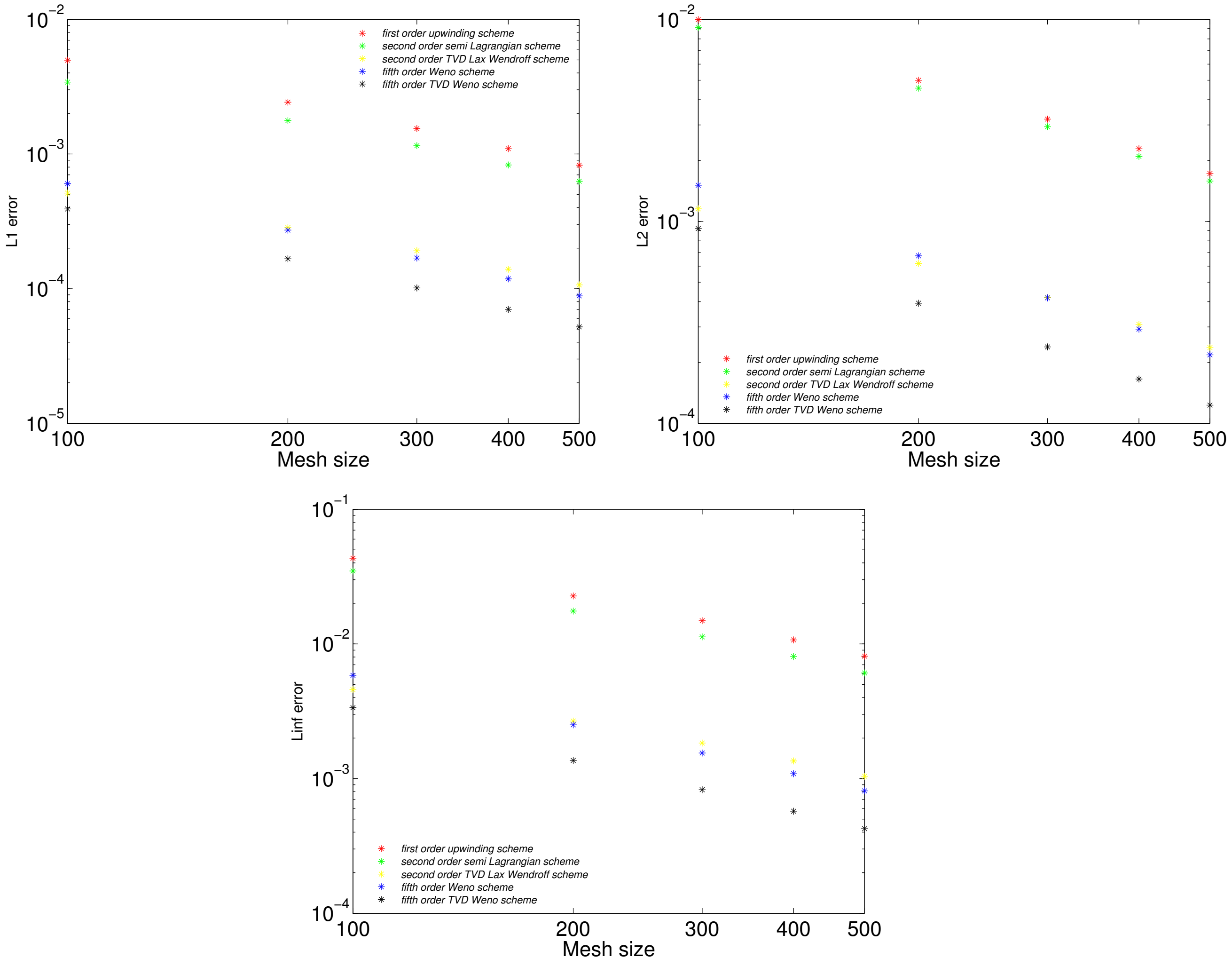


Figure 5.7: L^1, L^2 and L^∞ errors evolution in log-log scale for a Gaussian Bell initial condition and for $\gamma = \gamma_0 = 0.2$ day^{-1}

Transport scheme	L^1 order	L^2 order	L^∞ order
First order upwinding scheme	1.11	1.09	1.03
Second order semi Lagrangian scheme	1.05	1.08	1.08
Second order TVD Lax Wendroff scheme	0.97	0.97	0.91
Fifth order WENO scheme	1.19	1.20	1.22
Fifth order TVD WENO scheme	1.25	1.25	1.28

Table 5.3: L^1, L^2 and L^∞ orders of convergence of the five transport schemes for a Gaussian Bell initial condition and for $\gamma = \gamma_0 = 0.2$ day^{-1}

The Strang splitting ensures a second order of convergence in time if each method in the different steps of the splitting are second order methods. As expected, upwind scheme shows first order of convergence. The WENO scheme shown also first order of convergence because the time scheme is at the first order (forward Euler). For the TVD Lax Wendroff and WENO schemes, we do not obtain better order than first one because using Runge Kutta methods require to evaluate the derivative of the equation (5.24) at multiple time points and then to evaluate the velocity field at these points. Here we only used the velocity field at the previous time step. The issue is similar for the semi-Lagrangian scheme, which requires to compute the velocity field at two time steps.

A non smooth initial condition

Fig 5.8 shows the evolution of the errors in respect with mesh size for the non smooth initial condition, which is a square pulse.

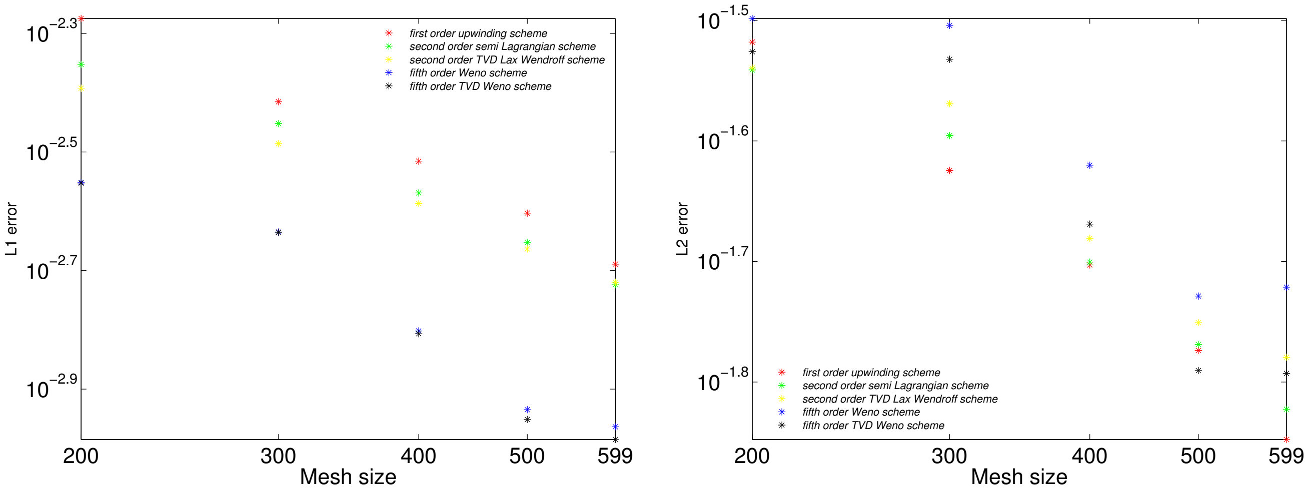


Figure 5.8: L^1, L^2 and L^∞ errors evolution in log-log scale for a Square Pulse initial condition and for $\gamma = \gamma_0 = 0.2$ day^{-1}

The table 5.4 shows the convergence orders of each scheme.

Transport scheme	L^1 order	L^2 order
First order upwinding scheme	0.86	0.68
Second order semi Lagrangian scheme	0.79	0.61
Second order TVD Lax Wendroff scheme	0.70	0.55
Fifht order WENO scheme	0.94	0.55
Fifht order TVD WENO scheme	0.99	0.65

Table 5.4: L^1, L^2 and L^∞ convergence orders of the five transport schemes for a Square Pulse initial condition and for $\gamma = \gamma_0 = 0.2$ day^{-1}

Here we have almost a first order of convergence in L^1 norm but not in L^2 norm. It is because the scheme computes approached derivatives assuming the solution as smooth enough whereas the solution is not continuous at the tumor boundary. For this reason, there is no convergence in L^∞ error.

Accuracy

We tested the convergence of different numerical schemes. Now we want to chose the “best” scheme, in other words the more accurate relatively to the computing time. In order to do this, we first examine a case where it is possible to obtain the exact solution.

Spherical geometry

We consider the system (5.17) in 3D with a spherical geometry, in other words a spherical tumor (radius R_T, center $r = 0$) and spherical boundaries (radius R_B, center $r = 0$). We consider now a Heaviside function for the initial tumor distribution: ($P = 1$ inside the tumor, 0 outside). In $1D$, all along a radius with coordinates θ and φ being fixed, the tumor distribution is illustrated in Fig 5.9.

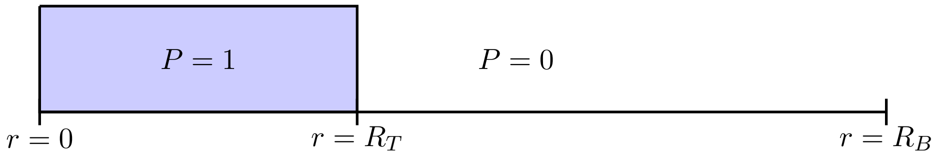


Figure 5.9: Spherical geometry in radial coordinates: a square pusle initial condition on the tumor density

With respect to the growth equation along the characteristics (5.23), a piece of tissue with a density equal to 1 (respectively 0) is transported along the characteristic and keeps its density equal to 1 (respectively equal to 0). Hence, the Heaviside function extends at the velocity field (at the interface between tumor and healthy tissue) speed, as it is shown in Fig 5.10.

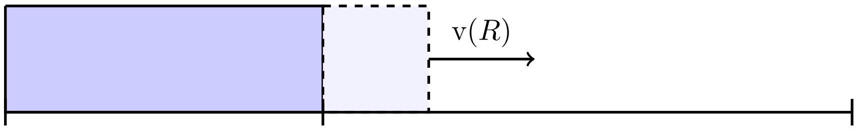


Figure 5.10: Spherical geometry in radial coordinates: transport of a tumor square pusle

This result can be extended in two and three dimensions. In 2D, we have a circular domain Ω and the initial condition is a tumor disk, as illustrated in Fig 5.11

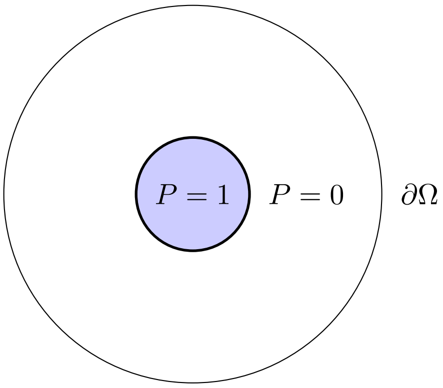


Figure 5.11: Spherical geometry in polar coordinates: the initial condition on the tumor density is a disk.

This problem is invariant by rotation and the velocity field, which is a gradient of pressure can be written:

$$\mathrm{v}(t,r,\theta) = \frac{1}{r}\frac{\partial\Pi(t,r,\theta)}{\partial\theta}e_\theta + \frac{\partial\Pi(t,r,\theta)}{\partial r}e_r,$$

where e_r and e_θ are the vectors of the polar basis. Because of the invariance by rotation, v is then purely radial: $\mathrm{v}(t,r,\theta) = \mathrm{v}(t,r)e_r$. Consequently, for two points $X_1 = (r,\theta_1)$ and $X_2 = (r,\theta_2)$ with the same radius r considered at the time $t > 0$, we have:

$$X_1 = X(t,y_1),$$
$$X_2 = X(t,y_2),$$

with X following the characteristic equation:

$$\frac{\partial X(t,y)}{\partial t} = \mathrm{v}(t,r)e_r,$$
$$X(0,y) = y.$$

Because the velocity is purely radial and does not depend on θ, we have $y_1 = (r_0,\theta_1)$ and $y_2 = (r_0,\theta_2)$. The angles θ_1 and θ_2 are conserved along the characteristics. Furthermore, if the initial tumor shape is circular, we have : $P(0,y_1) = P(0,y_2) = P(0,r_0)$ and then $P(t,X_1) = P(t,X_2) = P(t,r)$. Finally, we observe that an initial circular shape is conserved during time, as illustrated by Fig 5.12.

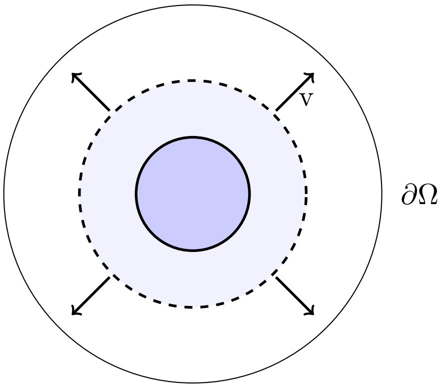


Figure 5.12: Spherical geometry in polar coordinates: transport of a tumor disk

We prove in the next part that while the tumor does not reach the boundaries, the domain size has no impact on the solution. Therefore, assuming that the tumor does not reach the boundaries, the solution with any domain size is the solution with an infinite domain.

Computing the velocity field in spherical geometry
Let us recall the Poisson equation giving the pressure field:

$$\begin{cases} -k\Delta\Pi = \gamma P \\ \Pi_{|\partial\Omega} = 0 \end{cases}$$

If the geometry is spherical, assuming the porosity k as constant with $k = 1$, the equation in radial coordinates becomes:

$$\begin{cases} -\dfrac{\partial^2\Pi}{\partial r^2} - \dfrac{2}{r}\dfrac{\partial\Pi}{\partial r} = \gamma P \\ \Pi(R_B) = 0 \\ \dfrac{\partial\Pi}{\partial r}(0) = 0 \end{cases} \tag{5.25}$$

We consider a Heaviside function for the tumor distribution: $P = 1$ for $0 \leq r \leq R_T$ and $P = 0$ for $R_T \leq r \leq R_B$. As we previously saw, the Heaviside function is transported and for all time t, the tumor distribution at the time t is an Heaviside function too.
Integrating the equation (5.25) leads to the following solution:

$$\begin{cases} \Pi(r) = \frac{\gamma_0}{6}(R_T^2 - r^2) + \frac{\gamma_0 R_T^3}{3}\left(\frac{1}{R_T} - \frac{1}{R_B}\right) & if\ r \in [0; R_T] \\ \Pi(r) = \frac{\gamma_0 R_T^3}{3}\left(\frac{1}{r} - \frac{1}{R_B}\right) & if\ r \in [R_T; R_B] \end{cases}$$

As we can see, Π depends on the domain size R_B. However, assuming the porosity k equal to 1, the velocity field has the following expression:

$$\begin{cases} \mathrm{v}(r) = \frac{\gamma_0 r}{3} & if\ r \in [0; R_T] \\ \mathrm{v}(r) = \frac{\gamma 0 R_T^3}{3r^2} & if\ r \in [R_T; R_B] \end{cases}$$

A first remark can be made on the velocity of the tumor interface $\mathrm{v}(t, R_T)$, which is the tumor expansion speed. We can see that under an exponential growth, this speed is increasing linearly with the radius of the tumor. This means that the larger the tumor is, the faster it expands, which makes sense with the idea of exponential dynamics. More importantly for our purpose, the velocity field does not depend on the domain size. Therefore the solution P does not depend on the domain size. This domain can be therefore extended to $[0; +\infty[$ without changing the solution. These results can be extended in two and three dimensions. In two dimensions also, the solution does not depend on the domain size and the domain can be extended to $\mathbb{R}^2$ without change the solution.

Analytical solution
We previously observed that in the simple case of γ constant considered here, the tumor mass follows an exponential growth. Therefore, for a given initial tumoral distribution with an initial mass $M(0)$, we exactly know the mass at t: $M(t) = M(0)\exp(\gamma_0 t)$. We also know that an initial spherical distribution gives a spherical distribution. In order to determine an analytical solution, we have thus to determine the radius R_T of this distribution, such that:

$$\int_0^R P(t,r) 2\pi r \mathrm{d}r = \int_0^R 2\pi r \mathrm{d}r = M(t).$$

We finally have $R_T = \sqrt{\frac{M(t)}{\pi}}$. The expression of the analytical solution is now known:

$$\begin{aligned} &P(t,r) = 1\ \ if\ r < \sqrt{\frac{M(0)\exp(\gamma_0 t)}{\pi}} \\ &P(t,r) = 0\ else \end{aligned}$$

It allows us to test the consistance and the accuracy of the numerical method we will use to solve the equations in more compicated cases where we do not know an analytical solution.

Numerical results
As we saw before, for a circular geometry, we know the analytical solution. However, we need a scheme that can treat more general types of geometry on a cartesian grid. That is why we consider a square domain, which is more appropriate for cartesian grids. The problem is that the solution with circular boundaries is different than the one on a square domain. We know the exact solution on a circular domain but not on a square one. To practically simulate the circular case in a square domain, we used a penalization method. More precisely, we considered a square domain with particular conditions on the porosity in the area that is in the square but not in the included circle. This area is colored in red in Fig 5.13.

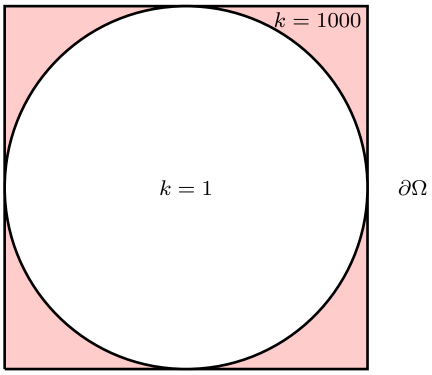


Figure 5.13: Imposing a 1000-fold greater porosity between the boundary and the tumor front than within the tumor allows to simulate the circular geometry.

We imposed a 1000 times higher porosity in the red area than within the circle, keeping a Dirichlet boundary condition on the square. It models a case of circular boundaries with Dirichlet conditions. Indeed, Fig 5.14 shows a comparison between two simulations where the porosity is either constant (equal to 1) on the whole domain or equal to 1 within the included circle and equal to 1000 in the red area, for the same value of γ. It gives the same final burden but two different shapes, the second case giving a circular tumor, whereas in the other case the squared domain changes the shape, as we can see in Fig 5.14.

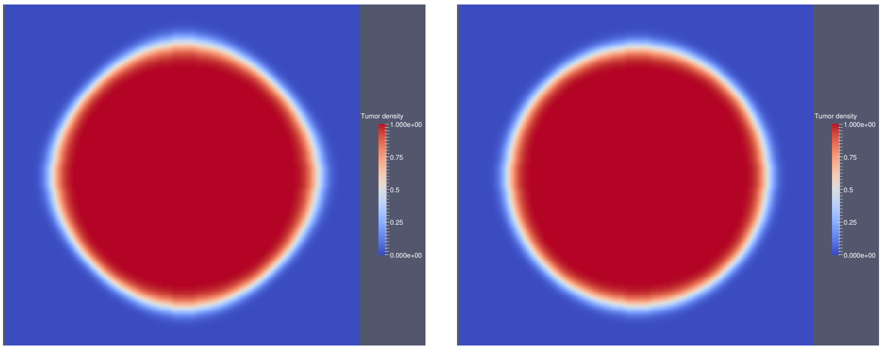


Figure 5.14: Simulation results for the tumor final distribution: final time $T = 1$day, $\gamma_0 = 0.67\text{day}^{-1}$, Mesh size 200×200. Left: $k = 1$, Right: $k = 1$ in the included circle, $k = 1000$ outside.

We considered a circular initial tumor with a given surface and simulated the model with a given parameter γ_0 and a given final time ensuring that the tumor at the final time is “far enough” from the boundaries. Knowing the analytical solution for this geometry, we wanted to compare the accuracy of each scheme relatively to a reasonable computing time (40s). The required mesh sizes for this computing time and the errors relatively to the analytical solution are presented for each scheme in the table 5.5.

Transport scheme	Required mesh size	L^1 relative error	L^2 relative error
First order upwind scheme	350	14.11%	21.10%
Second order semi Lagrangian scheme	290	14.46%	23.07%
Second order TVD Lax Wendroff scheme	320	12.22%	18.94%
Fifht order WENO scheme	300	8.41%	20.44%
Fifht order TVD WENO scheme	245	10.48%	23.10%

Table 5.5: L^1 error and L^2 error for fixed computing time

We see that the WENO scheme has better L^1 and L^2 accuracies than all other schemes except the Lax Wendroff scheme for the L^2 accuracy. We also compared the computing times of the schemes for a given reasonable precision. To perform it, we simulated the growth for each scheme with a mesh refined enough to obtain a L^1 relative error of 10% and compared the computing time of the schemes. The table 5.6 summarizes these results. The simulations with each scheme are presented in the figure A.1 of the appendices.

Transport scheme	Required mesh size	Computing time(s)
First order upwind scheme	750	406
Second order semi Lagrangian scheme	750	715
Second order TVD Lax Wendroff scheme	530	173
Fifth order WENO scheme	250	23
Fifth order TVD WENO scheme	250	45

Table 5.6: Computing times of the five schemes with a fixed relative L^1 error of 10%

We notice that the WENO scheme required the shortest computing time. The results presented here have been obtained with the value $\gamma_0 = 0.67$ day^{-1} (a fast growing tumor) for the proliferation rate but we got similar results with $\gamma_0 = 0.2$ day^{-1}.

Taking γ constant allowed us to test and compare the accuracy of several schemes. However, *in vivo* tumor growth rarely shows constant growth rates. That is why we are interested in a spatial model ensuring that the doubling time of the mass is increasing with time. Such a property can be obtained with the spatial model presented previously in considering a proliferation that depends on the environment conditions (nutrient supplying, competition of the cells for nutrients, pressure conditions, etc.). In this study, we consider a simple model where proliferation rate depends on the environmental pressure Π. This model has already been used and confronted with *in vivo* metastatic growth in mice in [1]. This pressure-mediated growth law induces a nonlinear elliptic equation on the pressure Π. In the next part, we prove the existence of solutions for this nonlinear equation, propose a fixed-point method for the resolution, and prove the theoretical convergence of this algorithm. We then test the numerical convergence of the complete system and finally confront the model to growth data of lung metastasis in mice in order to study the capacity of the model to reproduce the shape evolution of a metastasis during the growth. For these simulations, we chose the WENO transport scheme, which exhibited good accuracy/computational time ratio in the previous test case.

5.4 Pressure-mediated proliferation: mathematical and numerical analysis

We consider the pressure-mediated growth law presented in the equation (5.22) of the section 5.2.2:

$$\gamma(\Pi) = \gamma_0 \exp\left(-\frac{\Pi}{\Pi_c}\right)$$

The major difference between this model and the previous one (with γ constant) is the resolution of the equation for the pressure Π. Let us recall that the domain Ω is Lipschitz (square domain). For a given distribution P of tumor density, the equation on Π is now the following nonlinear equation:

$$\left\{\begin{array}{l} -k\Delta\Pi = \gamma_0 \exp\left(-\frac{\Pi}{\Pi_c}\right) P \\ \Pi_{|\partial\Omega} = 0 \end{array}\right. \tag{5.26}$$

In this section, we suppose that Ω is a Lipschitz domain and $P \in L^\infty(\Omega)$ and non negative. We prove here the existence and uniqueness for the problem (5.26) and propose a fixed-point algorithm for the resolution. Moreover, a result of regularity of the solution is also presented.

5.4.1 Existence, uniqueness and regularity of solution

We first prove uniqueness and existence of a solution of (5.26) in $H_0^1(\Omega)$. In this section, without loss of generality, we fix $\gamma_0 = 1$, $k = 1$ and $\Pi_c = 1$. Moreover, we assume that $P \in L^\infty(\Omega)$.

Proposition 5.4.1. *Let $P \in L^\infty(\Omega)$ and non negative. The system (5.26) has a unique solution in $H_0^1(\Omega)$ and this solution is non negative.*

Proof. The variational form of the previous equation is written as follows:

$$\int_\Omega \nabla\Pi \cdot \nabla\varphi - \int_\Omega \exp(-\Pi) P\varphi = 0 \tag{5.27}$$

for all $\varphi \in H_0^1(\Omega)$.

Uniqueness

Assuming that a weak solution exists, let us prove that such a solution is unique. Let $g \in H_0^1(\Omega)$ and $h \in H_0^1(\Omega)$ such that, for all $\varphi \in H_0^1(\Omega)$:

$$\int_\Omega \nabla g \cdot \nabla\varphi - \int_\Omega \exp(-g) P\varphi = 0$$
$$\int_\Omega \nabla h \cdot \nabla\varphi - \int_\Omega \exp(-h) P\varphi = 0.$$

Writing the difference between the two equations and taking $\varphi = g - h$, it comes:

$$\int_\Omega |\nabla(g-h)|^2 = \int_\Omega (\exp(-g) - \exp(-h))(g-h)P.$$

For all $x \in \Omega$, $(\exp(-g(x)) - \exp(-h(x)))(g(x) - h(x)) \leq 0$. We can then deduce that $g = h$, which proves the uniqueness of the solution.

Positivity

Assuming that a weak solution exists, let us prove that such a solution is non negative. Let Π be a weak solution and Π^- the negative part of Π defined as:

$$\begin{aligned} \Pi^- &= \Pi \quad \text{if } \Pi < 0 \\ \Pi^- &= 0 \quad \text{otherwise.} \end{aligned}$$

Given that $\Pi \in H^1(\Omega)$, $\Pi^- \in H^1(\Omega)$ [133, chap.II, Cor.2.1.8]. In choosing $\varphi = \Pi^-$, (5.27) can be written:

$$\int_\Omega \nabla\Pi \cdot \nabla\Pi^- - \int_\Omega \exp(-\Pi)P\Pi^- = 0.$$

Based on the result from [133] showing that $\nabla\Pi^- = \nabla\Pi$ if $\Pi < 0$ and $\nabla\Pi^- = 0$ otherwise, the previous equation can be written:

$$\int_\Omega |\nabla\Pi^-|^2 = \int_\Omega \exp(-\Pi)P\Pi^-.$$

We can conclude that $\Pi^- = 0$ because $|\nabla\Pi^-|^2 \geq 0$ and $\exp(-\Pi)P\Pi^- \leq 0$. The solution Π is then non negative.

Existence

We define :

$$G(x) = \int_0^x \exp(-s)\mathrm{d}s$$

and $f : H_0^1(\Omega) \longrightarrow \mathbb{R} \cup +\infty$ such that $f : u \longmapsto \int_\Omega G(u(x))P(x)\mathrm{d}x$. We define now a functional $J : H_0^1(\Omega) \longrightarrow \mathbb{R}$ such that:

$$J(u) = \frac{1}{2}\int_\Omega |\nabla u|^2 - f(u).$$

To prove that J is well-defined, we need to show that if $u \in H_0^1(\Omega)$, then $\exp(-u) \in L^1(\Omega)$. This follows from the Trudinger inequality, asserting that whenever $\mathcal{O}$ is a bounded Lipschitz subset of $\mathbb{R}^N$, then

$$\int_{\mathcal{O}} \exp\left(|u|^{\frac{N}{N-1}}\right) < \infty, \quad \forall u \in W^{1,N}(\mathcal{O}), \tag{5.28}$$

see [134, chap.VIII, rem.17]. The assumption there is that $\mathcal{O}$ is C^1-smooth but generalizing to Lipschitz domains is straightforward using the extension theorem [135, chap.VI, thm.5]. In our case, (5.28) becomes:

$$\int_\Omega \exp\left(|u|^2\right) < \infty \;\; \forall u \in H^1(\Omega).$$

Moreover, $\forall a, b \in \mathbb{R}$, we have $2ab \leq a^2 + b^2$. Thus, $\forall u \in H^1(\Omega)$, we have that $u \leq \frac{u^2+1}{2}$. Hence

$$0 \leq \exp(u) \leq \exp\left(\frac{u^2+1}{2}\right) \in L^1(\Omega)$$

by (5.28). We then conclude that $\exp(u) \in L^1(\Omega)$ for all $u \in H^1(\Omega)$. Therefore, the functional J is well-defined. Next, let us prove that J is differentiable and express its differential. This amounts to show that f is differentiable. If this is the case, then

$$DJ(u)\cdot\varphi = \int_\Omega \nabla u \cdot \nabla\varphi - Df(u).\varphi \quad \forall\varphi \in H_0^1(\Omega).$$

The differential of f at $u \in H_0^1(\Omega)$ applied to $\varphi \in H_0^1(\Omega)$ can be computed as follows:

$$\begin{aligned} f(u+\varphi)-f(u) &= \int_\Omega (G((u+\varphi)(x)) - G(u(x)))P(x)\mathrm{d}x \\ f(u+\varphi)-f(u) &= \int_\Omega \int_{u(x)}^{(u+\varphi)(x)} \exp(-s)\mathrm{d}sP(x)\mathrm{d}x \\ f(u+\varphi)-f(u) &= \int_\Omega (-\exp(-(u+\varphi)(x)) + \exp(-u(x)))P(x)\mathrm{d}x. \end{aligned}$$

According to the Lagrange form of Taylor's theorem applied to this right hand side, there exists $\xi(x) \in [u(x),(u+\varphi)(x)]$ such that:

$$f(u+\varphi)-f(u) = \int_\Omega (\varphi(x)\exp(-u(x)) - \frac{1}{2}\exp(-\xi(x))\varphi(x)^2)P(x)\mathrm{d}x,$$

therefore we arrive at the following inequality:

$$\left| f(u+\varphi)-f(u) - \int_\Omega \varphi(x)\exp(-u(x))P(x)\mathrm{d}x \right| \leq \frac{1}{2}\int_\Omega \exp\left(|u(x)| + |(\varphi)(x)|\right)\varphi^2(x)P(x)\mathrm{d}x. \quad (5.29)$$

We now prove that there is a real constant C such that

$$\left| f(u+\varphi)-f(u) - \int_\Omega \varphi(x)\exp(-u(x))P(x)\mathrm{d}x \right| \leq C\|\varphi\|^2_{H_0^1} \quad \text{when} \quad \|\varphi\|_{H_0^1} \longrightarrow 0. \quad (5.30)$$

It has been proved in [136] that there exist 2 real constants c_1 and c_2 depending only on Ω such that $\forall\, 1 \leq p < \infty$ and $\forall u \in H^1(\Omega)$,

$$\|\exp(|u|)\|_{L^p} \leq c_1 \exp\left(c_2 p \|u\|^2_{H^1}\right). \quad (5.31)$$

Let $p_1, p_2 > 1$ such that $\frac{1}{p_1} + \frac{1}{p_2} = \frac{1}{2}$. Thanks to Hölder's inequality, (5.29) implies:

$$\left| f(u+\varphi)-f(u) - \int_\Omega \varphi(x)\exp(-u(x))P(x)\mathrm{d}x \right| \leq \frac{1}{2}\|\exp(|u|)\|_{L^{p_1}}\|\exp(|\varphi|)\|_{L^{p_2}}\left\|\varphi^2\right\|_{L^2}\|P\|_{L^\infty}.$$

Using (5.31), we can write

$$\left| f(u+\varphi)-f(u) - \int_\Omega \varphi(x)\exp(-u(x))P(x)\mathrm{d}x \right| \leq \frac{1}{2}c_1^2\exp\left(c_2p_1\|u\|^2_{H^1}\right)\exp\left(c_2p_2\|\varphi\|^2_{H^1}\right)\left\|\varphi^2\right\|_{L^2}\|P\|_{L^\infty}.$$

In order to prove (5.30), we can assume that $\|\varphi\|_{H_1} \leq 1$. Then

$$\left| f(u+\varphi)-f(u) - \int_\Omega \varphi(x)\exp(-u(x))P(x)\mathrm{d}x \right| \leq \frac{1}{2}c_1^2\exp\left(c_2p_1\|u\|^2_{H^1}\right)\exp\left(c_2p_2\right)\left\|\varphi^2\right\|_{L^2}\|P\|_{L^\infty}.$$

Thus, there exists a constant $C = C(u)$ such that:

$$\left| f(u+\varphi)-f(u) - \int_\Omega \varphi(x)\exp(-u(x))P(x)\mathrm{d}x \right| \leq C \parallel \varphi^2 \parallel_{L^2}, \quad \|\varphi\|_{H_1} \leq 1.$$

Moreover, $\left\|\varphi^2\right\|_{L^2} = \|\varphi\|^2_{L^4}$, hence:

$$\left| f(u+\varphi)-f(u) - \int_\Omega \varphi(x)\exp(-u(x))P(x)\mathrm{d}x \right| \leq C\|\varphi\|^2_{L^4}, \quad \|\varphi\|_{H_1} \leq 1.$$

The injection of $H^1(\Omega)$ into $L^4(\Omega)$ is continuous, then there exists a constant K such that:

$$\left| f(u+\varphi) - f(u) - \int_\Omega \varphi(x)\exp(-u(x))P(x)\mathrm{d}x \right| \leq K\|\varphi\|^2_{H^1_0}$$

where K is a real constant. Therefore, f is differentiable and so is J. Its differential can be written as follows:

$$DJ(u)\cdot\varphi = \int_\Omega \nabla\Pi\cdot\nabla\varphi - \int_\Omega \exp(-u)P\varphi$$

which corresponds to the variational form (5.27). Using the Riesz representation theorem, we write $\langle J'(u),\varphi\rangle_{H^1_0} = DJ(u)\cdot\varphi$. It comes then:

$$\langle J'(u)-J'(v), u-v\rangle_{H^1_0} = \int_\Omega (\nabla u - \nabla v)^2 - \int_\Omega (\exp(-u)-\exp(-v))(u-v)P,$$

and as $\exp(-x)$ is decreasing, we have that $\int_\Omega(\exp(-u)-\exp(-v))(u-v)P \leq 0$ and then:

$$\langle J'(u)-J'(v), u-v\rangle_{H^1_0} \geq \|u-v\|^2_{H^1_0}.$$

Therefore J is α-convex with $\alpha = 1$. J being differentiable and hence continuous, J has a unique minimum on $H^1_0(\Omega)$ [137, chap.IX, Th.9.2.6]. Let Π be the unique function of $H^1_0(\Omega)$ that minimizes J, it verifies: $\langle J'(\Pi),\varphi\rangle = 0$ and therefore

$$\int_\Omega \nabla\Pi\cdot\nabla\varphi - \int_\Omega \exp(-\Pi)P\varphi = 0.$$

Π is then a weak solution of the problem. □

5.4.2 A semi implicit iterative algorithm to solve the problem

In this section, we don't assume $\gamma_0 = 1$, $k = 1$ and $\Pi_c = 1$ anymore. In order to solve the problem (5.26), we define an iterative algorithm. To do so, let us define the following sequence $(\Pi_n)_n$ of $H^1_0(\Omega)$:

$$\Pi_{n+1} = \Pi_n + \rho k\Delta\Pi_{n+1} + \rho\gamma_0\exp\left(-\frac{\Pi_n}{\Pi_c}\right)P, \tag{5.32}$$

with $\Pi_0 \in H^1_0(\Omega)$, $k,\gamma_0,\rho > 0$ and $P \in L^\infty(\Omega)$. This algorithm corresponds to an implicit time discretization of the heat equation with an explicit nonlinear source term. We prove in this section that for well chosen values of the step ρ, the sequence $(\Pi_n)_n$ defined by the algorithm (5.32) converges geometrically in $H^{3/2}(\Omega)$ to the solution of (5.26). It allows to conclude that this solution is in $H^{3/2}(\Omega)$ and thus has a better regularity that initially proved in the previous paragraph.

Proposition 5.4.2. *Let $P \in L^\infty(\Omega)$, P non negative, and $\Pi_0 \in H^1_0(\Omega)$. The algorithm defined, $\forall n \geq 0$, by:*

$$\begin{cases} \Pi_{n+1} = \Pi_n + \rho k\Delta\Pi_{n+1} + \rho\gamma_0\exp\left(-\frac{\Pi_n}{\Pi_c}\right)P, \\ \Pi_{n+1}{}_{|\partial\Omega} = 0 \end{cases} \tag{5.33}$$

is well-posed in $H^1_0(\Omega)$.

Proof. Let $\Pi_n \in H^1_0(\Omega)$, the variational form of the algorithm is written as follows:

$$\int_\Omega \Pi_{n+1}\varphi = \int_\Omega \Pi_n\varphi - \rho k\int_\Omega \nabla\Pi_{n+1}\cdot\nabla\varphi + \rho\gamma_0\int_\Omega \exp\left(-\frac{\Pi_n}{\Pi_c}\right)P\varphi,$$

for all $\varphi \in H^1_0(\Omega)$. We consider the bilinear form a of H^1_0 defined as follows:

$$a(u,v) = \int_\Omega uv + \rho k\int_\Omega \nabla u\cdot\nabla v.$$

This bilinear form is continuous. Indeed, using Cauchy-Schwartz inequality, we have:

$$|a(u,v)| \leq \|u\|_{L^2}\|v\|_{L^2} + \rho k\|u\|_{H_0^1}\|v\|_{H_0^1}.$$

The Poincaré inequality gives then:

$$|a(u,v)| \leq (C^2 + \rho k)\|u\|_{H_0^1}\|v\|_{H_0^1}.$$

Clearly, a is also coercive. Indeed, we can write:

$$a(u,u) = \|u\|^2_{L^2} + \rho k\|u\|^2_{H_0^1},$$

and then we have:

$$a(u,u) \geq \rho k\|u\|^2_{H_0^1}. \tag{5.34}$$

Moreover, the following linear form L:

$$L(v) = \int_\Omega \Pi_n v - \rho\gamma_0 \int_\Omega \exp\left(-\frac{\Pi_n}{\Pi_c}\right) Pv$$

is well defined because $\Pi_n \in L^2(\Omega)$ and $\exp\left(-\frac{\Pi_n}{\Pi_c}\right) \in L^2(\Omega)$ by (5.31). Moreover L is continuous. The Lax-Milgram theorem allows to conclude that there is a unique solution of (5.33) in H_0^1. □

Proposition 5.4.3. *Let $P \in L^\infty(\Omega)$ and non negative. For all $n \geq 0$, if $\Pi_n \geq 0$, then $\Pi_{n+1} \geq 0$ where Π_{n+1} is defined by (5.33).*

Proof. Assuming that $\Pi_n \geq 0$, let us take the L^2 scalar product of (5.32) with $\Pi^-_{n+1} = \min(0, \Pi_{n+1})$:

$$\int_\Omega |\Pi^-_{n+1}|^2 = \int_\Omega \Pi_n\Pi^-_{n+1} - \rho k\int_\Omega |\nabla\Pi^-_{n+1}|^2 + \rho\gamma_0\int_\Omega \exp\left(-\frac{\Pi_n}{\Pi_c}\right) P\Pi^-_{n+1} = 0.$$

Π_n being non negative, the right hand side of the equation is negative. Consequently, $\left\|\Pi^-_{n+1}\right\|_{L^2} \leq 0$, which means that $\Pi^-_{n+1} = 0$ and then $\Pi_{n+1} \geq 0$. Therefore $\forall n \in \mathbb{N}$, if $\Pi_n \geq 0$, then $\Pi_{n+1} \geq 0$. □

Convergence of the algorithm and regularity of the solution

Let $\rho > 0$, $k > 0$, and A be the linear operator from $L^2(\Omega)$ into $H_0^1(\Omega)$ such that $A(\varphi) = u$ with u the unique solution of the following system:

$$\left\{\begin{array}{l} u = \rho k\Delta u + \varphi \\ u_{|\partial\Omega} = 0 \end{array}\right. \tag{5.35}$$

Note that u uniquely exists by using the Lax-Milgram theorem, using an argument similar to that in (5.34). Note also that this argument shows that A is continuous from L^2 to H_0^1.

Proposition 5.4.4. *The operator A is continuous from $L^2(\Omega)$ to $H^{3/2}(\Omega)$: There exists a constant C_A such that*

$$\|A(\varphi)\|_{H^{3/2}} \leq C_A\|\varphi\|_{L^2}.$$

Proof. Let $f \in L^2(\Omega)$ and let us define the following problem:

$$k\Delta u = f \tag{5.36}$$

$$u_{|\partial\Omega} = 0. \tag{5.37}$$

The domain Ω being Lipschitz, it has been proved in [138] that the problem (5.37) has a unique solution in $H^{3/2}(\Omega)$ and that the operator $R : f \mapsto u$ is continuous from L^2 to $H^{3/2}$. For all $\varphi \in L^2(\Omega)$ and $u \in H_0^1(\Omega)$ such that $A(\varphi) = u$, u verifies the equation (5.37) with $f = \frac{u}{\rho} - \frac{\varphi}{\rho}$. Such a function f is

$L^2(\Omega)$, therefore thanks to [138], we know that the solution of the problem (5.35) is in $H^{3/2}(\Omega)$ and that the operator A is continuous from $L^2(\Omega)$ into $H^{3/2}(\Omega)$. So there exists a positive constant C_A such that:

$$\|A(\varphi)\|_{H^{3/2}} \leq C_A \|\varphi\|_{L^2},$$

which achieves the proof. □

Proposition 5.4.5. *Let $\gamma_0, \Pi_c > 0$, $P \in L^\infty(\Omega)$, P non negative and $\Pi_0 \in H_0^1(\Omega)$. For a step ρ such that $0 \leq \rho \leq \frac{\Pi_c}{\gamma_0 \|P\|_{L^\infty}}$, then the sequence $(\Pi_n)_n$ defined by (5.33) converges geometrically in $L^2(\Omega)$.*

Proof. We now consider the L^2 scalar product of (5.35) with u:

$$\|u\|_{L^2}^2 = -\rho k \|u\|_{H_0^1}^2 + \int_\Omega \varphi u.$$

The Cauchy-Schwartz inequality allows to write:

$$\|u\|_{L^2}^2 \leq -\rho k \|u\|_{H_0^1}^2 + \|u\|_{L^2} \|\varphi\|_{L^2}.$$

Moreover, Poincaré inequality leads to:

$$\|u\|_{L^2}^2 \leq -\frac{\rho k}{C} \|u\|_{L^2}^2 + \|\varphi\|_{L^2} \|u\|_{L^2},$$

which can be written:

$$\|u\|_{L^2} \leq \frac{1}{1 + \frac{\rho k}{C}} \|\varphi\|_{L^2}.$$

A is therefore L^2-contracting: there exists a positive constant $K_A < 1$ such that $\forall \varphi \in L^2(\Omega)$, $\|A(\varphi)\|_{L^2} \leq K_A \|\varphi\|_{L^2}$. Moreover, $\forall \varphi \in L^2(\Omega)$ such that $\varphi \geq 0$, then $A(\varphi) \geq 0$ (Same proof as in proposition 5.4.3). We can deduce that

$$A(|\varphi|) = |A(\varphi)| \quad \forall \varphi \in L^2(\Omega), \tag{5.38}$$

and also that:

$$A(\varphi_2) \geq A(\varphi_1) \quad \forall \varphi_1, \varphi_2 \in L^2(\Omega), \varphi_2 \geq \varphi_1. \tag{5.39}$$

Let us now consider two consecutive iterations of the algorithm (5.33):

$$\Pi_n = A(\Pi_{n-1} + \rho\gamma_0 \exp\left(\frac{-\Pi_{n-1}}{\Pi_c}\right) P) \tag{5.40}$$

$$\Pi_{n+1} = A(\Pi_n + \rho\gamma_0 \exp\left(\frac{-\Pi_n}{\Pi_c}\right) P). \tag{5.41}$$

Writing (5.41)-(5.40) leads to:

$$\Pi_{n+1} - \Pi_n = A\left(\left(1 - \frac{\rho\gamma_0 P}{\Pi_c}\right)(\Pi_n - \Pi_{n-1}) + \rho\gamma_0 P\left(F\left(\frac{\Pi_n}{\Pi_c}\right) - F\left(\frac{\Pi_{n-1}}{\Pi_c}\right)\right)\right),$$

where $F(\Pi_n) = F \circ \Pi_n$, $F(\Pi_{n-1}) = F \circ \Pi_{n-1}$, and F is the function from $\mathbb{R}^+$ into $\mathbb{R}^+$ defined by

$$F(x) = \exp(-x) - 1 + x.$$

Using (5.38), we can write:

$$|\Pi_{n+1} - \Pi_n| = A\left(\left|\left(1 - \frac{\rho\gamma_0 P}{\Pi_c}\right)(\Pi_n - \Pi_{n-1}) + \rho\gamma_0 P\left(F\left(\frac{\Pi_n}{\Pi_c}\right) - F\left(\frac{\Pi_{n-1}}{\Pi_c}\right)\right)\right|\right). \tag{5.42}$$

Taking the L^2 norm of (5.42) gives:

$$\|\Pi_{n+1}-\Pi_n\|_{L^2} = \left\| A\left(\left| \left(1-\frac{\rho\gamma_0 P}{\Pi_c}\right)(\Pi_n-\Pi_{n-1}) + \rho\gamma_0 P\left(F\left(\frac{\Pi_n}{\Pi_c}\right) - F\left(\frac{\Pi_{n-1}}{\Pi_c}\right)\right)\right|\right)\right\|_{L^2}.$$

Using (5.39) allows then to write:

$$\|\Pi_{n+1}-\Pi_n\|_{L^2} \leq \left\| A\left(\left| \left(1-\frac{\rho\gamma_0 P}{\Pi_c}\right)(\Pi_n-\Pi_{n-1})\right| + \left|\rho\gamma_0 P\left(F\left(\frac{\Pi_n}{\Pi_c}\right) - F\left(\frac{\Pi_{n-1}}{\Pi_c}\right)\right)\right|\right)\right\|_{L^2}.$$

We now assume that $0 \leq \rho \leq \frac{\Pi_c}{\gamma_0\|P\|_{L^\infty}}$. It implies that $\left|1-\frac{\rho\gamma_0 P}{\Pi_c}\right| = \left(1-\frac{\rho\gamma_0 P}{\Pi_c}\right)$, allowing to write:

$$\|\Pi_{n+1}-\Pi_n\|_{L^2} \leq \left\| A\left(\left(1-\frac{\rho\gamma_0 P}{\Pi_c}\right)|\Pi_n-\Pi_{n-1}| + \rho\gamma_0 P\left|F\left(\frac{\Pi_n}{\Pi_c}\right) - F\left(\frac{\Pi_{n-1}}{\Pi_c}\right)\right|\right)\right\|_{L^2}.$$

Moreover, F is 1-Lipschitz because its derivative on $\mathbb{R}^+$ is:

$$\frac{\mathrm{d}F}{\mathrm{d}x}(x) = -\exp(-x)+1,$$

which implies that $0 \leq \frac{\mathrm{d}F}{\mathrm{d}x} \leq 1$ on $\mathbb{R}^+$. This allows us to write:

$$\|\Pi_{n+1}-\Pi_n\|_{L^2} \leq \|A\left(\left(1-\frac{\rho\gamma_0 P}{\Pi_c}\right)|\Pi_n-\Pi_{n-1}| + \rho\gamma_0 P\left|\frac{\Pi_n}{\Pi_c}-\frac{\Pi_{n-1}}{\Pi_c}\right|\right)\|_{L^2}$$

and then:

$$\|\Pi_{n+1}-\Pi_n\|_{L^2} \leq \|A(|(\Pi_n-\Pi_{n-1})|)\|_{L^2}.$$

A being K_A-contracting, we can write:

$$\|\Pi_{n+1}-\Pi_n\|_{L^2} \leq K_A\|\Pi_n-\Pi_{n-1}\|_{L^2} \leq K_A^n\|\Pi_1-\Pi_0\|_{L^2}$$

It implies a geometrical convergence of $(\Pi_n)_n$ in $L^2(\Omega)$. □

Proposition 5.4.6. *Let $\gamma_0, \Pi_c > 0$, $P \in L^\infty(\Omega)$, P non negative and $\Pi_0 \in H_0^1(\Omega)$. For a step ρ such that $0 \leq \rho \leq \frac{\Pi_c}{\gamma_0\|P\|_{L^\infty}}$, then the sequence $(\Pi_n)_n$ defined in (5.33) converges geometrically in $H^{3/2}(\Omega)$ and then the solution of (5.26) is in $H^{3/2}(\Omega)$.*

Proof. The convergence of $(\Pi_n)_n$ in $L^2(\Omega)$ norm implies the L^2 geometrical convergence of the sequence $(\rho\gamma_0\exp\left(\frac{-\Pi_n}{\Pi_c}\right)P)_n$. Indeed, the exponential function being 1-Lipschitz on $\mathbb{R}^-$, we can write, $\forall n \in \mathbb{N}$:

$$\left\|\rho\gamma_0\exp\left(\frac{-\Pi_n}{\Pi_c}\right)P - \rho\gamma_0\exp\left(\frac{-\Pi_{n-1}}{\Pi_c}\right)P\right\|_{L^2} \leq \rho\gamma_0\|P\|_{L^\infty}\|\Pi_n-\Pi_{n-1}\|_{L^2}$$

$$\left\|\rho\gamma_0\exp\left(\frac{-\Pi_n}{\Pi_c}\right)P - \rho\gamma_0\exp\left(\frac{-\Pi_{n-1}}{\Pi_c}\right)P\right\|_{L^2} \leq \rho\gamma_0\|P\|_{L^\infty}K_A^{n-1}\|\Pi_1-\Pi_0\|_{L^2}.$$

It implies that the sequence $(\Pi_n + \rho\gamma_0\exp\left(\frac{-\Pi_n}{\Pi_c}\right)P)_n$ also converges geometrically in L^2 norm. The proposition 5.4.4 allows then to conclude that $\forall n \geq 0$:

$$\|\Pi_{n+1}-\Pi_n\|_{H^{3/2}} = \left\|A(\Pi_n-\Pi_{n-1}+\rho\gamma_0(\exp\left(\frac{-\Pi_n}{\Pi_c}\right) - \exp\left(\frac{-\Pi_{n-1}}{\Pi_c}\right))P)\right\|_{H^{3/2}}$$

$$\|\Pi_{n+1}-\Pi_n\|_{H^{3/2}} \leq C_A\left\|\Pi_n-\Pi_{n-1}+\rho\gamma_0(\exp\left(\frac{-\Pi_n}{\Pi_c}\right) - \exp\left(\frac{-\Pi_{n-1}}{\Pi_c}\right))P\right\|_{L^2}$$

Because the sequence $\left(\Pi_n + \rho\gamma_0\exp\left(\frac{-\Pi_n}{\Pi_c}\right)P\right)_n$ also converges geometrically in L^2 norm, it therefore exists a constant $0 \leq K \leq 1$ such that:

$$\|\Pi_{n+1}-\Pi_n\|_{H^{3/2}} \leq C_A K^{n-1}\left\|\Pi_1-\Pi_0+\rho\gamma_0(\exp\left(\frac{-\Pi_1}{\Pi_c}\right) - \exp\left(\frac{-\Pi_0}{\Pi_c}\right))P\right\|_{L^2}.$$

Therefore (Π_n) converge geometrically in $H^{3/2}$ norm, and the solution of (5.26) is in $H^{3/2}$. □

Transport scheme	L^1 order	L^2 order
First order upwinding scheme	0.9043	0.7196
Second order semi Lagrangian scheme	0.8770	0.7051
Second order TVD Lax Wendroff scheme	0.8480	0.7060
Fifht order WENO scheme	1.1213	0.7725
Fifht order TVD WENO scheme	1.1259	0.7870

Table 5.7: L^1 and L^2 convergence orders of the five transport schemes for a Square Pulse initial condition and using the pressure mediated proliferation law. Parameter values: $\gamma_0 = 0.78$ day^{-1}; $\Pi_c = 0.0026$ Pa

5.4.3 Implementation and numerical results

Implementation

For the simulations, we chose the implicit algorithm because it was faster than the explicit one. To implement this algorithm, we fixed a tolerance threshold $\varepsilon \geq 0$ for the relative error $\frac{\|\Pi_{n+1}-\Pi_n\|_\infty}{\|\Pi_{n+1}\|_\infty}$ and maximal number of iteration called maxiter. We then implement the algorithm:

- Initializing: $\Pi_c = 0$, $n = 0$
- While ($\frac{\|\Pi_{n+1}-\Pi_n\|_\infty}{\|\Pi_{n+1}\|_\infty} \geq \varepsilon$ and $n \leq$ maxiter) solve $\Pi_{n+1} = \Pi_n + \rho\Delta\Pi_{n+1} + \rho\exp(-\Pi_n)\,P$ and $n = n+1$
- $\Pi = \Pi_{n+1}$

Numerical test of convergence

We tested the convergence of the scheme for non smooth initial condition (square pulse) because in further simulations and data assimilations, tumors are modelled like this. First order of convergence was obtained with respect to the L^1 norm. Fig 5.15 shows the evolution of the errors with respect to mesh size and simulations of the test.

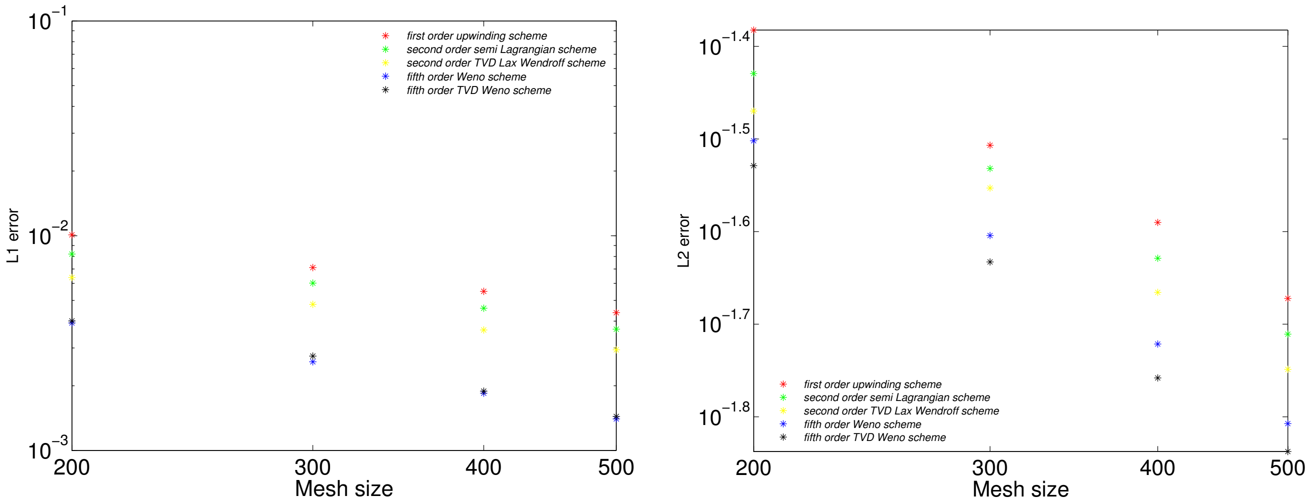


Figure 5.15: L^1 and L^2 errors evolution in log-log scale for a Square Pulse initial condition with the pressure-mediated proliferation model. $\gamma_0 = 0.78$ day^{-1}; $\Pi_c = 0.0026$ r·i (ri: relative unit)

The table 5.7 shows the convergence orders of each scheme.

5.5 Pressure-mediated proliferation: data assimilation and tumor shape dynamics

The fist goal of this part is to check if the pressure-mediated growth model is able to describe *in vivo* tumor growth. The second one is to investigate how a real tumor shape evolves under such a pressure-mediated growth law and if this predicted shape is in accordance with the observed final shape.

5.5.1 Calibration method

Assuming each metastasis as spherical, we fitted the model on volume dynamics data in minimizing the least squares objective function

$$LS(\gamma_0, \Pi_c) = \sum_{i=1}^{n} \frac{(f(t_i, \gamma_0, \Pi_c) - V_i)^2}{f(t_i, \gamma_0, \Pi_c)},$$

where n is the number of time points, $f(t_i, \gamma_0, \Pi_c)$ is the volume computed by the model at the time t_i and V_i is the volume data at the time t_i. In order to recover the parameters γ_0 and Π_c to fit the data, we used a Monte-Carlo method, which was easy to implement and parallelize. Parameters were randomly chosen in the set obtained below and the couple realizing the minimum of the objective function was selected as the best-fit. Boundaries of the parameter space have been established by an analysis of the model and biological considerations for the parameters values.

5.5.2 Bounding the parameter space

We consider data of tumor growth and we assume that we have n time points of data. Let $\{V_i\}_{0\leq i\leq n-1}$ be the volumes of the tumor at the times $\{t_i\}_{0\leq i\leq n-1}$. V_{n-1} is hence the final volume of the tumor. Without loss of generality, we assume here that $V_{i-1} \leq V_i, \ \forall 1 \leq i \leq n-1$. We also assume the tumor as spherical.

Here we assume the tumor distribution as a Heaviside function: $P = 1$ within the tumor, $P = 0$ outside. We consider now the equation of the mass dynamics (5.15) derived from the equation on the proliferative cells (5.5):

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\Omega} P \mathrm{d}x = \int_{\Omega} \gamma P \mathrm{d}x$$

Here, $\gamma = \gamma_0 \exp\left(-\frac{\Pi}{\Pi_c}\right)$ and we have naturally $\gamma_0 \exp\left(-\frac{\Pi}{\Pi_c}\right) \leq \gamma_0$. We have hence:

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\Omega} P \mathrm{d}x \leq \gamma_0 \int_{\Omega} P \mathrm{d}x$$

Thus we can write:

$$V_{n-1} \leq V_0 \exp\left(\gamma_0 t_{n-1}\right)$$

Finally, we have:

$$\gamma_0 \geq \frac{1}{t_{n-1}} \log\left(\frac{V_{n-1}}{V_0}\right)$$

Moreover, γ_0 represents the maximal proliferation rate of the cancer cells, in other words the proliferation rate in optimal conditions of proliferation. In [104], it is established that mammalian cancer cells can not proliferate with a shorter doubling time than 10 hours, which corresponds to a proliferation rate of $\gamma_{0,max} = 1.67$ day^{-1}. We found hence boundaries for the parameter γ_0:

$$1.67 \geq \gamma_0 \geq \frac{1}{t_{n-1}} \log\left(\frac{V_{n-1}}{V_0}\right) \tag{5.43}$$

Now we have to establish boundaries for the second parameter. We write:

$$\gamma = \gamma_0 \exp\left(-\frac{\Pi}{\Pi_c}\right),$$

which implies:

$$\gamma \geq \gamma_0 \exp\left(-\frac{\max(\Pi)}{\Pi_c}\right), \tag{5.44}$$

where $\max(\Pi)$ is the maximal value of the pressure Π during all the growth (from the time t_0 to t_{n-1}). We have now to estimate a value of $\max(\Pi)$. We consider P^{n-1} the final tumor distribution and define Π and $\tilde{\Pi}$ the solutions of the following sytems:

$$\begin{cases} -k\Delta\Pi = \gamma_0 \exp\left(-\frac{\Pi}{\Pi_c}\right) P^{n-1} \\ \Pi_{|\partial\Omega} = 0 \end{cases} \tag{5.45}$$

$$\begin{cases} -k\Delta\tilde{\Pi} = \gamma_0 P^{n-1} \\ \tilde{\Pi}_{|\partial\Omega} = 0 \end{cases} \tag{5.46}$$

Let $\varphi \in H_0^1(\Omega)$, we write (5.45)-(5.46) in variational form:

$$\int_\Omega \nabla(\Pi - \tilde{\Pi}) \cdot \nabla\varphi - \int_\Omega \left(\gamma_0 \exp\left(-\frac{\Pi}{\Pi_c}\right) P^{n-1} - \gamma_0 P^{n-1}\right)\varphi = 0$$

Taking $\varphi = (\Pi - \tilde{\Pi})^+$, we obtain:

$$\int_\Omega |(\nabla(\Pi - \tilde{\Pi}))^+|^2 - \int_\Omega (\gamma_0 \exp\left(-\frac{\Pi}{\Pi_c}\right) P^{n-1} - \gamma_0 P^{n-1})(\Pi - \tilde{\Pi})^+ = 0$$

We have $\gamma_0 \exp\left(-\frac{\Pi}{\Pi_c}\right) P^{n-1} - \gamma_0 P^{n-1} \leq 0$ and $(\Pi - \tilde{\Pi})^+ \geq 0$ so we can conclude that $(\Pi - \tilde{\Pi})^+ = 0$ and then $\Pi \leq \tilde{\Pi}$.

The tumor being spherical, we make also the assumption of a circular domain, which is reasonable if the domain is far enough from the tumor. Such a circular geometry allows us to consider the radial problem (5.25). Let us recall the analytical solution $\tilde{\Pi}$ of this system:

$$\begin{cases} \tilde{\Pi}(r) = \frac{\gamma_0}{6}(R_T^2 - r^2) + \frac{\gamma_0 R_T^3}{3}(\frac{1}{R_T} - \frac{1}{R_B}) & if\ r \in [0; R_T] \\ \tilde{\Pi}(r) = \frac{\gamma_0 R_T^3}{3}(\frac{1}{r} - \frac{1}{R_B}) & if\ r \in [R_T; R_B] \end{cases}$$

Where R_T is the radius of the final tumor (time t_{n-1}). We see that the maximal value of $\tilde{\Pi}$ is $\tilde{\Pi}(0) = \frac{\gamma_0 R_T^2}{6} + \frac{\gamma_0 R_T^3}{3}(\frac{1}{R_T} - \frac{1}{R_B})$.

As $\Pi \leq \tilde{\Pi}$, we can conclude that $\max(\Pi) \leq \tilde{\Pi}(0)$. Based on (5.44), we hence have:

$$\gamma \geq \gamma_0 \exp\left(-\frac{\tilde{\Pi}(0)}{\Pi_c}\right)$$

Again, we consider the equation of the mass dynamics (5.15) derived from the equation on the proliferative cells (5.5):

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega P\mathrm{d}x = \int_\Omega \gamma P\mathrm{d}x \geq \gamma_0 \exp\left(-\frac{\tilde{\Pi}(0)}{\Pi_c}\right)\int_\Omega P\mathrm{d}x$$

We can then write:

$$
\begin{aligned}
&V_{n-1} \geq V_0 \exp\left(\gamma_0 \exp\left(-\frac{\tilde{\Pi}(0)}{\Pi_c}\right) t_{n-1}\right) \\
&\frac{\log\left(\frac{V_{n-1}}{V_0}\right)}{t_{n-1}} \geq \gamma_0 \exp\left(-\frac{\tilde{\Pi}(0)}{\Pi_c}\right) \\
&\log\left(\frac{\log\left(\frac{V_{n-1}}{V_0}\right)}{t_{n-1}\gamma_0}\right) \geq -\frac{\tilde{\Pi}(0)}{\Pi_c} \\
&\frac{\tilde{\Pi}(0)}{\Pi_c} \geq \log\left(\frac{\log\left(\frac{V_{n-1}}{V_0}\right)}{t_{n-1}\gamma_0}\right)
\end{aligned}
$$

We finally obtain:

$$
\Pi_c \leq \frac{\tilde{\Pi}(0)}{\log\left(\frac{t_{n-1}\gamma_0}{\log\left(\frac{V_{n-1}}{V_0}\right)}\right)}
$$

And then we have boundaries on the parameter Π_c:

$$
0 \leq \Pi_c \leq \frac{\frac{\gamma_0 R_T^2}{6} + \frac{\gamma_0 R_T^3}{3}\left(\frac{1}{R_T} - \frac{1}{R_B}\right)}{\log\left(\frac{t_{n-1}\gamma_0}{\log\left(\frac{V_{n-1}}{V_0}\right)}\right)}
$$

Boundaries on γ_0 and Π_c allow to do Monte-Carlo random sampling in a restricted space, which leads to a shorter computational time and a better identifiability of the parameters.

5.5.3 Results

We consider here data on mice that were injected with RENCA cells in the renal subcapsular space. It leads to the formation of a kidney tumor and then metastases in the lungs. Four metastases have been segmented at several days, providing data points on the dynamics of the volume. We fitted these four dynamics with the pressure-mediated growth model thanks to the Monte-Carlo method, assuming each metastasis as spherical. As we can see in Fig 5.16, the model fits well to the data on volume.

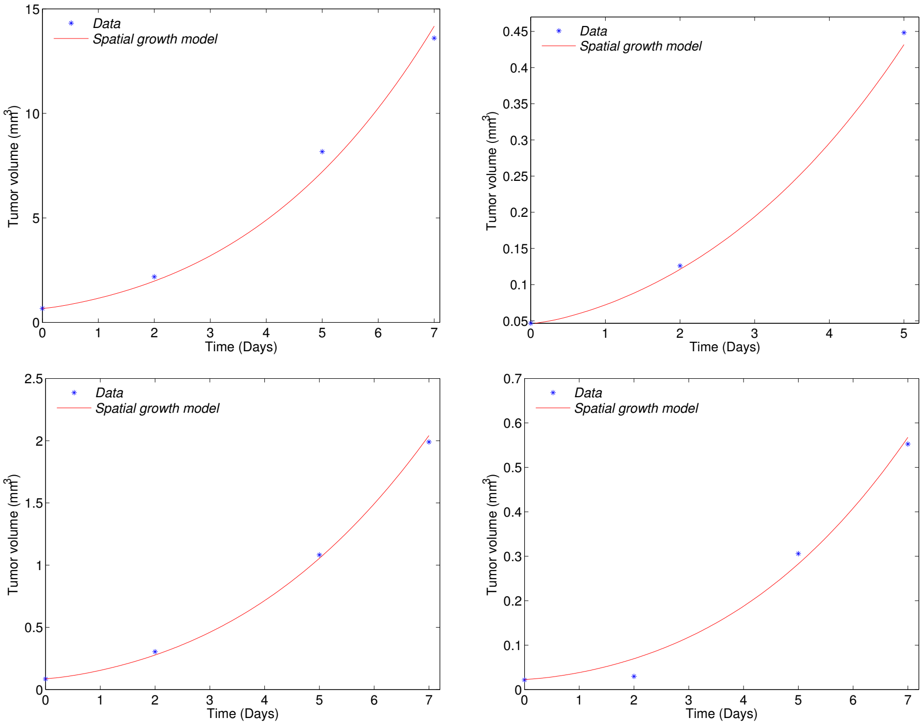


Figure 5.16: Four individual metastatic growths fitted with the spatial model.

With the calibrated model, we then simulated the growth of the four metastases starting from the observed shape and observed how the shape evolved under the model and compared the predicted final shape to the observed one. Fig 5.17 presents one of the predicted shapes incorporated into the dicom image.

The final shape predicted by the model is quite close to the observed one in the case shown in Fig 5.17. The three other metastatic shapes are presented in Fig 5.18.

One of the three other shapes is acceptable, but the two others are not well predicted. Indeed, in the model, the spatial expansion seems isotropic whereas the observed tumors seem to favor particular directions for tumor stretching. Such anisotropies could be induced by the tissue structure and environmental factors, which are not taken into account in the model. Furthermore, if the initial shape was already stretched, anisotropies were probably impacting the early growth. However, the shape dynamics could possibly depend on the proliferation distribution only. Although a deeper study would be necessary to answer to this question, we briefly investigated it on the basis of the four previous metastases. We compared the mass dynamics in the cases of the real shape and a spherical shape with the parameters calibrated on the mass dynamics in each case. The results are presented in Fig 5.19.

Among the four metastases, we obtained 0.6%, 2.7%, 1.1% and 2.9% differences of final burden between the spherical case and the real shape. It suggests that the shape has only a minimal if not negligible impact on the burden dynamics.

5.6 Conclusions

In this chapter, we first introduced some classical phenomenological ODE-based models of tumor growth and challenged their descriptive abilities on renal tumor growth data. Interestingly, we observed that tumor volume and tumor cells in the kidney do not exhibit the same growth dynamics.

We then introduced a model of tumor spatial expansion at the tissue scale, based on conservation equations and a pressure-mediated proliferation law. This model was able to fit the growth of several lung metastatic nodules and simply explains the growth rate decay by a phenomenological pressure

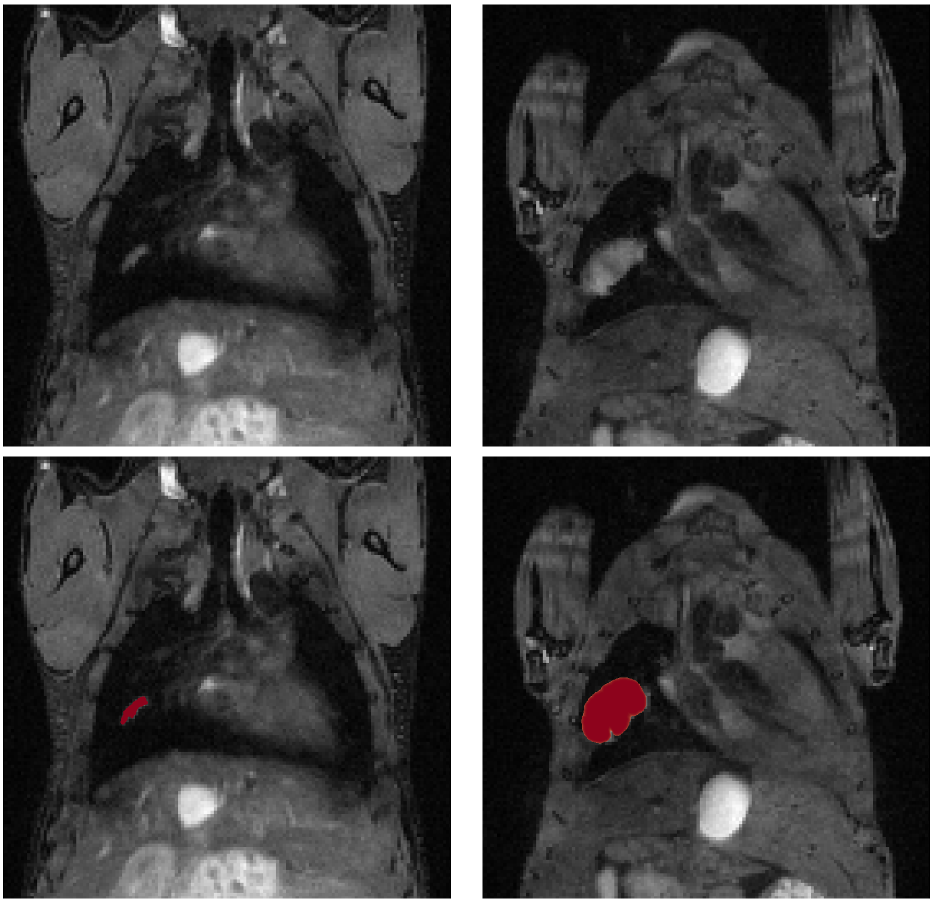


Figure 5.17: Top: Coronal MRI data of the lungs at days 19 and 26. Bottom: simulated growth by the model using the fitted parameters and starting from the real shape of the observed metastasis at day 19 on the coronal MRI slice.

due to proliferation. The model exhibited a good descriptive power for the tumor volume dynamics and the shape time-evolution has been investigated on the four metastatic lesions. To perform the simulations, we chose the fifth order WENO scheme for the transport part, which exhibited the best accuracy/computing cost ratio on test cases with a constant growth rate. The pressure-mediated proliferation law induced a nonlinear elliptic equation on the pressure. We proved existence and uniqueness of a $H^{3/2}$ solution for a given tumor distribution and Lipschitz domain. At each time step, we solved this equation thanks to a fixed-point method. We proved that the algorithm converges geometrically in $H^{3/2}$ norm and tested the numerical convergence of the complete system. A methodology to obtain boundaries for the parameter space depending on the fitted data was also presented. This parameter space restriction allowed to highly reduce the number of iterations in the Monte-Carlo algorithm and to better identify the parameters.

Several perspectives come out from this study. The model could be confronted to more data. Identifiability and predictive power of the model could be studied in a further work. On the numerical side, comparisons have been made between finite differences Eulerian schemes but it could be possible to compare these methods with the level-set method that has been used in [139] for the same conservation equations. On the mathematical side, a proof of existence of the tumor density P in the problem of the conservation laws coupled with the nonlinear equation on the pressure, could be investigated. On the light of this perspective, we tried to obtain the maximal possible regularity on the pressure Π in order to have a regular enough velocity v. Indeed, the velocity being the pressure gradient, a H^s regularity on Π ensures a H^{s-1} regularity on v. We would like to have a velocity field in $W^{1,\infty}$, allowing to define a flow for the conservative transport equations (5.5)-(5.6). Such a regularity on the

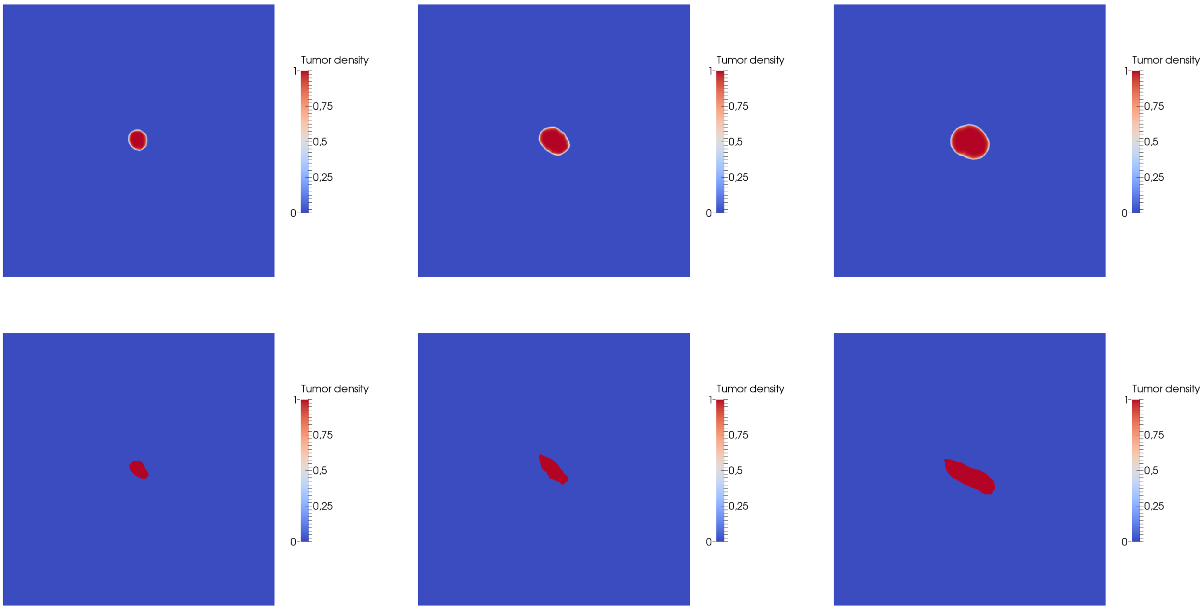


Figure 5.18: Top: final shapes of the three other metastases resulting from the simulated growth by the model calibrated on the volume dynamics of the metastasis. Bottom: The real final shapes of the three metastases, delineated from coronal MRI slices.

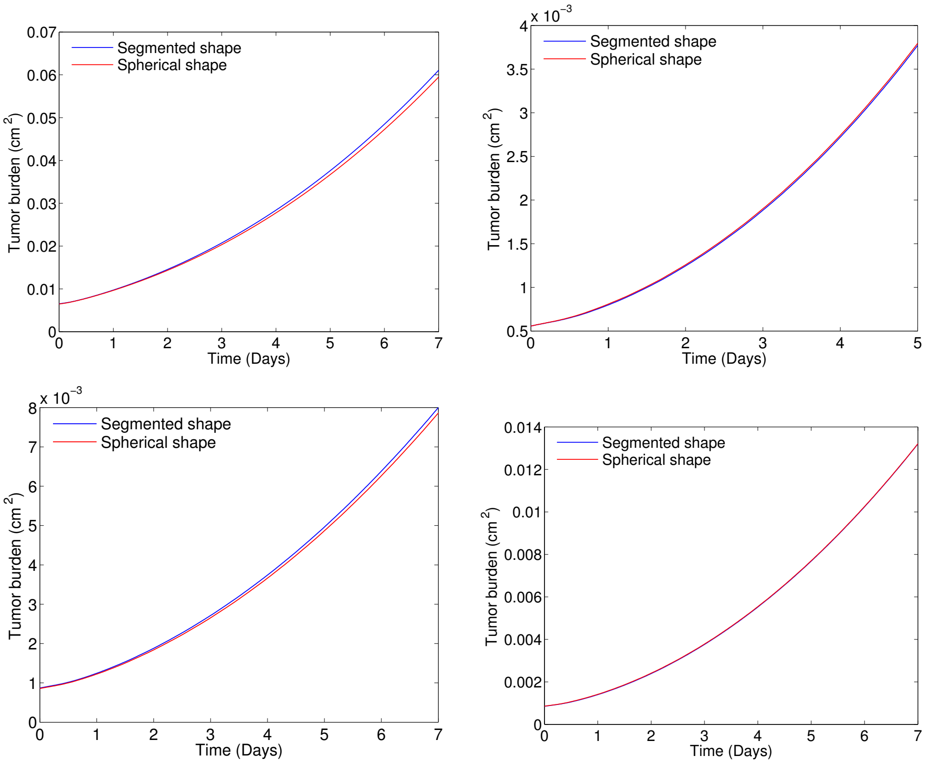


Figure 5.19: Spherical and non-spherical shapes. Volume dynamics of the two simulations using the four parameters sets retrieved from the calibrations. The final relative differences are 0.6%, 2.7%, 1.1% and 2.9%.

velocity field would require a pressure field in $W^{2,\infty}$, which is much more regular than the results we obtained here. Indeed, the square domain Ω being Lipschitz but not C^1, it limits the regularity of the solution of the elliptic equation (5.26). In order to obtain a more regular pressure field, a smooth domain should be considered. If it allows to obtain a regular enough pressure field, existence and uniqueness of the coupled problem between (5.5)-(5.7) and (5.26) could be investigated thanks to techniques used by Michel et al. in [140].

The next part, the main of this thesis, is devoted to modeling the metastatic process. Different kind of mathematical models will be used in these quantitative studies, including the model we previously analyzed.

Part III

Quantitative modeling of the metastatic process

Cette partie, dédiée à la modélisation du processus métastatique, est le coeur principal de cette thèse et est constituée de trois chapitres. Les deux premiers sont fortement connectés puisque une majeur partie de ces deux chapitres a été publiée dans un même article [1]. Dans le premier, on s'intéresse à la dynamique métastatique dans les poumons de souris atteintes de tumeur du rein. On confronte une théorie classique du processus métastatique, basée sur la formation de métastases monoclonales (à partir d'une cellule survivante) croissant indépendamment les unes des autres, à des données longitudinales de la masse métastatique dans le poumon. Cette théorie, formalisée mathématiquement par Iwata en 2000 [17], semble capable de décrire la dynamique de la masse métastatique totale, mais s'avère en revanche incapable de décrire les distributions en taille des métastases observées à l'IRM. En effet, la théorie classique prédit la formation de métastases petites et nombreuses, alors que les métastases observées à l'IRM sont moins nombreuses et plus massives. Ceci indique que certains phénomènes ne sont probablement pas pris en compte dans la théorie classique. Nous avons alors suggéré deux phénomènes pouvant expliquer les distribution en taille observées : la fusion de foyers métastatiques croissant à proximité les uns des autres et l'attraction de cellules tumorales circulantes par des foyers ou des niches métastatiques déjà établis dans le poumon. Le premier phénomène, la fusion de foyers métastatiques, est observé sur nos données IRM. Un modèle basé sur la théorie classique mais incluant également l'espace limité de l'organe et la possibilité pour les métastatses voisines de fusionner (cette partie n'a en revanche pas encore été publiée.) a alors été proposé pour savoir si ce phénomène de fusion pouvait expliquer les distributions en taille métastatique observées à l'IRM. Malgré une faible amélioration des prédictions par rapport à la théorie classique, cette théorie de la fusion ne semble pas suffisante pour expliquer les données. La seconde hypothèse, à savoir l'attraction de cellules tumorales circulantes par des foyers métastatiques existants, a donc été étudiée. Pour ce faire, un protocole expérimental a été proposé à nos collaborateurs biologistes pour mettre en évidence ce phénomène.

Dans le deuxième chapitre de cette partie, nous nous intéressons aux interactions mécaniques entre métastases qui fusionnent. Pour ce faire nous modélisons la croissance spatiale de métastases et simulons à l'aide de ce modèle la croissance de foyers métastatiques proches les uns des autres, afin de quantifier l'impact de ces interactions mécaniques sur la dynamique globale. Le modèle utilisé est le modèle spatial présenté dans la partie précédente, décrivant l'expansion spatiale d'une tumeur par un gradient de pression et prenant en compte l'effet inhibiteur de la pression sur la prolifération des cellules cancéreuses. Les simulations effectuées semblent indiquer que l'impact des interactions mécaniques entre deux métastases sur leurs croissances respectives est potentiellement très important. En effet, à partir du modèle calibré sur la croissance de lésions métastatiques aux poumons, nous observons une réduction d'environ 30% de la masse métastatique totale due aux interactions mécaniques lorsque deux métastases fusionnent. Ces résultats pourraient avoir des implications en terme de dynamique métastatique globale.

Le troisième et dernier chapitre présente une étude de modélisation de la dynamique du microenvironnement métastatique. Plus précisément, le phénomène de niche prémétastatique est étudié et modélisé. Des études expérimentales récentes ont mis en évidence le rôle de cellules myéloïdes issues de la moelle osseuse dans la formation d'un environnement permissif au sein du futur site métastatique pour la formation et le développement des métastases [10, 141, 142]. Cette niche prémétastatique pourrait être cruciale pour le recrutement et la survie des cellules métastatiques au sein d'un organe *a priori* hostile, puisqu'empêcher la formation de cette niche empêche également la formation des métastases dans les études expérimentales citées précédemment. Après avoir étudié la littérature biologique de ce phénomène, nous avons construit un modèle mathématique basé sur cette littérature. Les valeurs des paramètres de ce modèle ont pour une part été fixées à partir de mesures expérimentales tirées de la littérature, tandis que les autres paramètres ont été calibrés sur les données de l'article de Kaplan et al. [10]. Le modèle semble capable de reproduire les dynamiques des cellules myeloïdes ainsi que de la masse métastatique. Les perspectives futures sont ensuite discutées, notamment les questions quantitatives pouvant être traitées par le modèle, comme l'histoire naturelle de la formation de cette niche prémétastatique ou encore les voies moléculaires et cellulaires pertinentes à cibler pour limiter au maximum le développement métastatique chez un patient.

This part is devoted to modeling the metastatic process. The first two chapters are strongly connected because both studies have been driven by the same data on metastasis dynamics in murine renal cell carcinoma. Parts of these two chapters have been recently accepted for publication in a same paper [1]. These two chapters deal with describing metastatic development and dynamics. The first one focuses on dynamics at the organism scale and the metastatic size distribution, whereas the second one is devoted to spatial interactions between metastatic foci.

The third chapter deals with the impact of the microenvironment on the metastasis development. More precisely, the dynamical aspects of the premetastatic niche formation are studied with a mechanistic model describing the cellular dynamics in the premetastatic and metastatic phases.

Except for the data from Kaplan et al [10], all the experimental data were provided by Andreas Bikfalvi, Lin Cooley, Wilfried Souleyreau and Raphael Pineau for the GFP analysis part, and Emeline Ribot and Sylvain Miraux for the imaging part.

Concerning the GFP analysis, the experiments were designed by Andreas Bikfalvi, Lin Cooley and Wilfried Souleyreau and were performed by Lin Cooley, Wilfried Souleyreau and Raphael Pineau. My contribution was first to analyse these data in using the modeling framework of Iwata et al. and Benzekry et al., and in using the computing library coded by Sébastien Benzekry in Matlab for the simulations of the metastatic distributions. Then, discussions with our biologist colleagues and a study of the literature led me to propose biological hypotheses and an experimental protocol with red-tagged and green-tagged cells to test them. Preliminary experiments were then performed by the biologist collaborators.

As in the previous part, the imaging data that have been used in this part were provided by Emeline Ribot and Sylvain Miraux from the center of magnetic resonance of biological systems (Bordeaux), that perfomed the MRI sequence used to visualize the mouse lesions. To simulate tumor-tumor mechanical interactions, I then used the spatial model I implemented (see the previous part).

Concerning the premetastatic niche, I first studied the biological literature and discussed with our biologist colleagues to better understand the phenomenon. Then I tried to find relevant quantitative questions to address, proposed a mathematical model describing the phenomenon, implemented it in C++, and confronted it to the data of Kaplan et al [10].

Chapter 6

Challenging the standard theory of metastatic colonization using a data-driven modeling approach

Published in Plos Computational Biology [1]

The biology of the metastatic colonization process remains a poorly understood phenomenon. To improve our knowledge of its dynamics, we conducted a modeling study based on multi-modal data from an orthotopic murine experimental system of metastatic renal cell carcinoma. The standard theory of metastatic colonization usually assumes that secondary tumors, once established at a distant site, grow independently from each other and from the primary tumor. Using a mathematical model that translates this assumption into equations, we challenged this theory against our data that included: 1) dynamics of primary tumor cells in the kidney and metastatic cells in the lungs, retrieved by green fluorescent protein tracking, and 2) magnetic resonance images (MRI) informing on the number and size of macroscopic lesions. Critically, when calibrated on the growth of the primary tumor and total metastatic burden, the predicted size distributions were not in agreement with the MRI observations. Moreover, tumor expansion only based on proliferation was not able to explain the volume increase of the metastatic lesions. These findings strongly suggested rejection of the standard theory, demonstrating that the time development of the size distribution of metastases could not be explained by independent growth of metastatic foci. The model predicted an increase in the number of lesions, but of smaller size when compared to the data. This led us to revise the standard theory and to propose two hypotheses in order to explain the observations: 1) small metastatic foci merge into larger ones and/or 2) circulating tumor cells may join already established tumors.

Introduction

Metastasis, the spread of cancer cells from a primary tumor to secondary location(s) in the body, is the ultimate cause of death for the majority of cancer patients [118, 5]. Although studied for more than 180 years [20], increasing efforts in recent years contributed to a better understanding of this aspect of tumor development [5, 4] with exciting new discoveries [41, 10, 11, 143] that potentially have important clinical implications. The metastatic process can be coarsely divided into two major phases: 1) dissemination of detaching cells from the primary tumor to a secondary site and 2) colonization of this distant organ [118, 47]. While the former has been relatively well elucidated, in particular due to recent advances about the epithelial-to-mesenchymal transition [21] and advances on our understanding of molecular and genetic determinants [8], the latter remains not fully understood, especially during the colonization phase [118, 9]. This is due, in part, to experimental limitations that hinder our ability to observe colonization of organs by tumor cells and the development of tumor lesions.

In this context, mathematical models provide powerful tools to potentiate data analysis, infer hidden information, test biological hypotheses against the empirical data and simulate a range of conditions that may be confronted to the biological reality. In recent years, several models for tumor growth have been developed (see [63, 61] for historical reviews), based on multiple modeling techniques from non-spatial models based on ordinary differential equations (see [65] for a benchmark of these against experimental *in vivo* data) to discrete agent-based models [126, 79, 144] and continuous partial differential equations based on tissue mechanics laws [12, 145]. However, despite a large body of literature for modeling tumor growth, relatively little effort has been devoted to the development and validation of mathematical models describing the biology of the metastatic process (see [82, 81] for an early and notable exception, [86, 91] for more recent studies and [80] for a recent review). In 2000, Iwata and colleagues proposed a simple mathematical model for the growth of a population of metastatic colonies [17], which was recently shown able to fit experimental data describing the increase in total metastatic burden [25, 26]. In this mathematical description, each metastasis grows independently from the others and from the primary tumor. We report herein a theoretical study to test this hypothesis using in vivo data derived from a metastatic renal carcinoma model in mice. We first show that the standard theory of metastatic initiation in which distinct foci grow independently from each other (as assumed in [82]) predicted an unrealistically large number of metastases, while the tumors sizes were too small.

In a space-limited organ (such as the lungs), where two neighbouring metastatic foci are growing in close vicinity, they might enter in contact and interactions occur, ultimately leading to the merging of the metastatic foci. This phenomenon is not taken into account in a classical description of metastasis development, although it can lead to important differences in the number and sizes of the colonies. Moreover, mechanical interactions could occur during metastases merging, possibly impacting the global dynamics. Therefore, we next conducted a simulation study to quantify the effect of mechanical interactions between two neighbouring tumors. Based on mechanical laws for tissue growth, we derived a minimally parameterized model (2 parameters). This second spatial model, based on a pressure-mediated growth law, once fitted to magnetic resonance imaging data of individual metastatic tumor growths, offered an adapted framework to perform simulations of spatially interacting tumors. These revealed significant impact of the interactions on the global growth and allowed to test if merging by passive motion could explain the data that are not in accordance with the classical model. To our knowledge, this is the first time that data on size distribution of metastasis at this resolution (with such a small visibility threshold, of the order of 0.05 mm^3) is reported and analysed in lights of a theoretical model.

As an initial step, we studied the growth rates of individual metastatic tumors. Then, we calibrated a more elaborated mathematical model of tumor growth and metastatic dissemination using quantitative data derived from green fluorescent protein (GFP)-tracking of primary and metastatic tumors (see Fig 6.2, $n = 31$ mice). Finally, we used the model to investigate predictions of the standard theory with regard to number and sizes of metastatic lesions and compared them to Magnetic Resonance

Imaging (MRI) data (see Fig 6.4, $n = 6$ mice).

6.1 Growth rates of individual metastatic tumors

RENCA cells were injected orthotopically in the sub-capsular space of the kidney of Balb/c mice. The first metastatic cells were observed in the lungs at day 14 and the first macro-metastases at days 18-19, as shown on (Fig 6.1). No metastasis was observed in other organs.

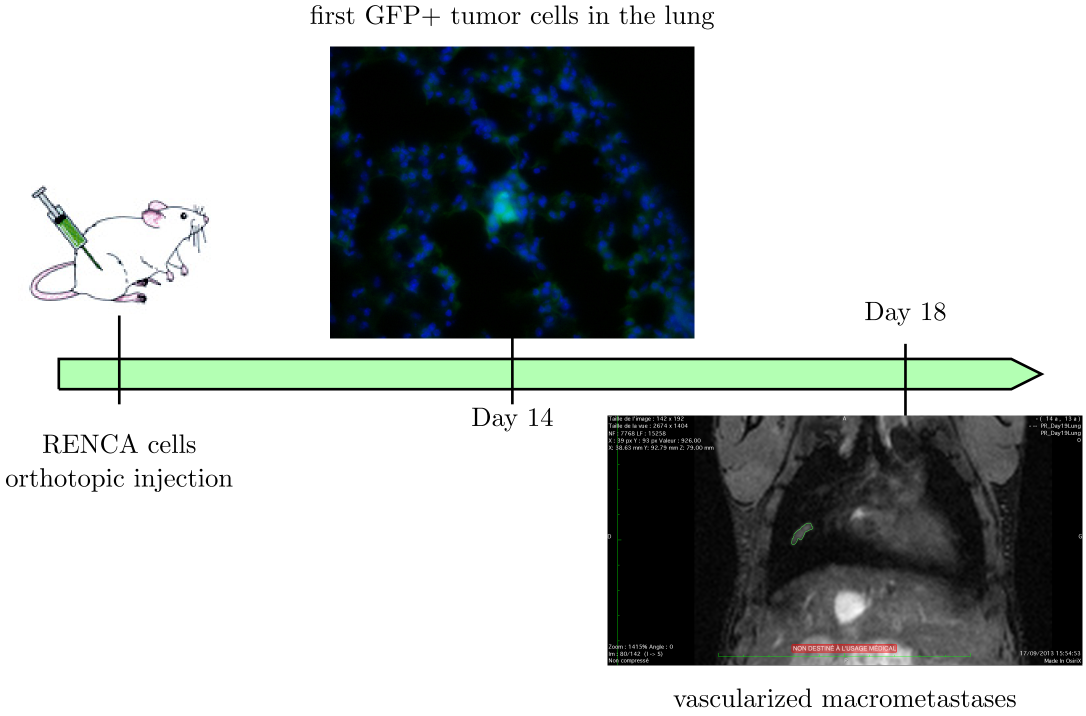


Figure 6.1: At day 14 after GFP+ RENCA cells injection, the first tumor cells were observed in the lungs. At days 18-19, the first macro-metastases were observed by MRI.

Assuming in a theoretical model that each metastasis originates from one surviving cell would imply that some metastases grow from the volume of one cell (10^{-6} mm^3, according to the well-established conversion rule 1 mm$^3 \simeq 10^6$ cells [146]), to a volume of few mm^3 (between 0.022 and 12 mm^3) in five days at most. This would give tumor doubling times comprised between 5 and 8 hours, which represent less than one third of the doubling time observed in vitro (24.5 hours [57]). Even if considering that the metastases arose from few cells (2-50) instead of one [18, 19], this would imply doubling times between 5.5 and 13.5 hours. These doubling times would also have to remain constant during 5 days. Such a fast growth is highly improbable since no mammalian cell has a cell cycle length smaller than 10 hours [104]. Moreover, the doubling time has been reported to be non-constant and to increase during *in vivo* growth [65]. Hence, growth at initiation would have to be even faster, in order to fit the data. Therefore, the theory consisting in describing each metastasis with a tumor expansion only based on cell proliferation seems unlikely.

6.2 Primary kidney tumor and the dynamics of lung metastasis:

The standard theory of metastatic development assumes that secondary tumors are seeded from the primary tumor and that, once established at the distant sites (the lungs in our case), they grow independently from each other and from the rest of the organism [20, 21, 22, 23, 24, 5], as distinct foci initiated by single or few cells. We tested this theory against the following data: a) dynamics of the increase of GFP+ tumor cells in the kidney, b) longitudinal quantification of GFP+ tumor cells in the lungs and c) MRI images of lung metastasis. To this aim, we formalized the standard theory into a mathematical model. The model was then fitted to the data a) and b) and predictions were compared to the data c). For data a) and b) (Fig 6.2), quantification of GFP expression by quantitative real-time polymerase chain reaction (see Materials and methods in Appendices) required sacrifice of the animals. Therefore, each data point corresponds to a distinct animal ($n = 31$ animals in total). Since tumor cells were not detected in the lungs before day 14, measured GFP signals in the lungs in the first days were considered as noise. Including very early time points for the fit would therefore result in a strong bias because the model would be fitted on GFP values that do not reflect the presence of tumor cells. Thus, we considered the data points only starting from day 10 for the initial point for comparison and ignored the previous data points. At day 25, the GFP signal in the lungs of the two mice was within the noise level. Considering as highly unlikely the event of no metastases at such an advanced time, particularly when observing that imaged mice at day 24 either exhibited lots of metastases in the lungs or were dead due to metastatic disease, we concluded to a technical issue and removed these two data points from the analysis.

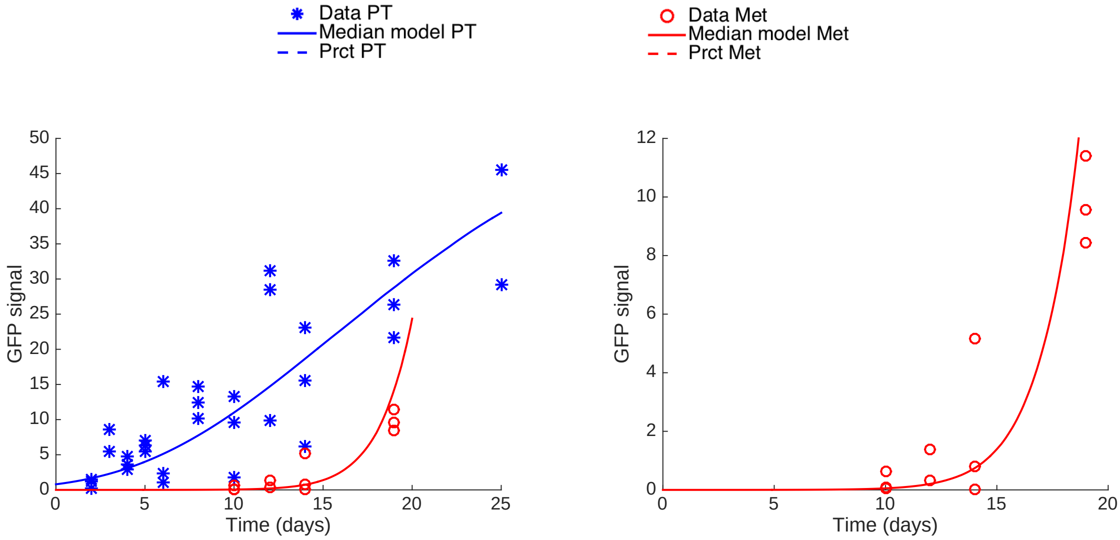


Figure 6.2: (A) Fits of the primary tumor and metastatic burden dynamics, under a mathematical model assuming independent growth of each secondary tumor and using mixed-effects modeling for statistical representation of the population distribution of the parameters and measurement error. (B) Fit on the metastatic burden. In panels (A) and (B), each data point corresponds to one distinct mouse (n=31 animals in total). Simulations were obtained using equation (6.1) for the primary tumor growth and equation (6.3) for the metastatic burden, endowed with a lognormal distribution of the parameters with the following values (median ± standard deviation): $\alpha = 0.417 \pm 171$ day^{-1}, $\beta = 0.106 \pm 0.0478$ day^{-1} and $\mu = 9.72 \times 10^{-6} \pm 0.428 \times 10^{-6}$ cell day^{-1}. PT= Primary tumor. Met = Metastatic burden. Prct=10% and 90% percentiles

6.2.1 Mathematical model:

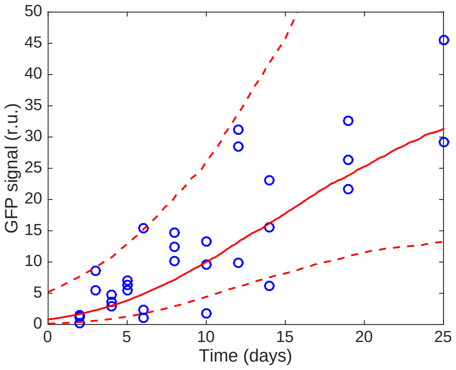
The standard theory as described above has been mathematically formalized by Iwata et al. [17]. This model was reported to provide a valid description of the dynamics of total metastatic burden in two animal models, including ours (human breast carcinoma xenograft [25, 26] and syngeneic renal cell carcinoma [26]).

The model was adapted here as follows. We assumed that the GFP signal was proportional to the number of cancer cells, itself proportional to the tumor volume observed by MRI. The primary tumor volume at time t was denoted $V_p(t)$ and its growth rate $g_p(V_p)$. The primary tumor disseminates metastatic cells into the lungs according to a volume-dependent dissemination law $d(V_p(t)$. The metastatic colonies then grow into the lungs according to the same growth law as the primary tumor. The model describes the size distribution of the metastatic lesions at the distant site by means of a size-dependent density $\rho(t, v)$ of metastatic colonies of size v at time t, i.e $\rho(v, t)\mathrm{d}v$ is the number of metastatic colonies with a size comprised between v and $v + \mathrm{d}v$. Secondary emission of metastases (i.e., metastases from metastases) was neglected here.

Tumor growth was modeled by means of the Gomp-Exp model [52], which is characterized by two phases: first an exponential phase (with a growth rate given by the in vitro proliferation rate), then a Gompertz phase (i.e., exponentially decreasing growth rate). For the primary tumor growth, the model writes

$$\begin{cases} \dfrac{\mathrm{d}V_p(t)}{\mathrm{d}t} = g(V_p(t)), \\ V_p(0) = V_{inj}, \\ g_p(V_p) = \min\left(\lambda V_p, (\alpha - \beta \ln(V_p))\, V_p\right), \end{cases} \tag{6.1}$$

where λ is the in vitro proliferation rate of RENCA cells (retrieved from [57]), α corresponds to the specific growth rate at the size of one relative unit of GFP signal, β is the rate of exponential decrease of the specific growth rate and V_{inj} is the amount of injected cells. The Gompertz model has been proven able to describe in vivo tumor growth in numerous animal experimental systems [65, 113, 64] as well as human data [58]. Adjunction of the initial exponential phase was considered here because the Gompertz model exhibits an infinite relative growth rate for arbitrary small cell numbers, a feature that was not considered relevant, especially for the metastases that start from one cell. At this point we needed to retrieve the GFP signal associated to one cell. To achieve this, we performed a preliminary fit of the primary tumor GFP signals with the parameter V_{inj} subject to optimization, using population mixed-effects statistical modeling (see Fig A.1). We then assumed proportionality between the number of cells injected (100 000) and the estimated signal at day 0. The GFP signal associated to one cell was then derived and is denoted by V_0. We found $V_0 = 7.96.10^{-6}$ relative units of GFP signal.



Par.	Unit	Median value	NSE (%)
α	[day^{-1}]	0.384 (20.07)	25.4
β	[day^{-1}]	0.1 (14.36)	39.1
V_0	[GFP(r.u)]	0.796 (695.8)	81.6

Figure 6.3: Population fit of the primary tumor dynamics. The initial volume is calibrated during the fit. Right panel: the points represent the data, the curve represents the median dynamics, and the dashed curves the percentiles. Left panel: values of the parameters resulting from the population fit of the primary tumor dynamics. NSE: normalized standard error.

Growth of secondary tumors was assumed to follow the same law (Gomp-Exp) and parameter values. More complex modeling including different growth parameters for the primary tumor and metastasis were tested but did not substantially improve the fits, while increasing the uncertainty in parameter estimation due to increased number of degrees of freedom. At present time, there is no detailed study on the shape that the function d should have. However, it is often assumed to follow the law [17, 25, 26]:

$$d(V_p) = \mu V_p^{\gamma},$$

where parameter μ is related to an intrinsic (genetic) probability of the metastatic potential of the tumor cells, combined to the probability of successfully passing all the steps required for establishment of a metastasis (detachment from the primary tumor, intravasation into the blood circulation, survival in transit, arrest and extravasation at the distant site, establishment of a new colony [20, 9]). Precise identification of the value of parameter γ was not possible on our data set, and in the following analysis, we arbitrarily fixed $\gamma = 1$, corresponding to the simplest assumption (emission proportional to the tumor volume). Overall, the model writes as a transport equation on ρ, endowed with suitable boundary and initial conditions [17]:

$$\begin{cases} \partial_t \rho(t,v) + \partial_v(\rho(t,v)g(v)) = 0 & t \in]0,+\infty[, \ v \in]V_0,+\infty[, \\ g(V_0)\rho(t,V_0) = d(V_p(t)) & t \in]0,+\infty[, \\ \rho(0,v) = 0 & v \in]V_0,+\infty[. \end{cases} \tag{6.2}$$

From the solution of this problem, the main quantity of interest for our purpose was the total metastatic burden, defined by

$$M(t) = \int_{V_0}^{+\infty} v\rho(t,v)dv. \tag{6.3}$$

The model output was fitted to the GFP expression data from the lungs (Fig 6.2).

6.2.2 Results

Due to the large inter-animal variability, we used a nonlinear mixed-effects statistical framework for fitting the model to the data and resulting estimation of the parameters [102] (see also [25, 26]) for applications to bioluminescence data of metastatic burden), which is particularly well suited for sparse longitudinal data in an animal population. Briefly, this framework considers estimation of a (parametric) distribution of the parameters within the population. This allows pooling all the data points together, thus leading to an increase of the robustness of the estimation and of the descriptive power for inter-animal variability. For the maximization of likelihood associated to nonlinear mixed-effects modeling, we used a version of a stochastic expectation maximization algorithm implemented in the Matlab function nlmefitsa [147]. To simulate the model, a well-adapted method using an integral formulation for $M(t)$ and the fast Fourier transform algorithm was employed [148]. This ensured reduction of the computational cost of simulations, which was necessary due to the very high number of runs required by nlmefitsa. Data from the primary tumor and the metastatic burden were fitted together, and the model demonstrated satisfactorily descriptive power for the total metastatic burden 6.2, in accordance with other studies [25, 26].

The calibrated model was further used to predict the distribution of macro-metastases visible in the MRI images, and to confront this prediction to the observations. Among the MRI data, images of only one mouse (over 6) were eligible for reliable assessment of the complete size distribution of macro-metastases, which was performed by manual segmentation of metastatic lesions in each of the 142 coronal slices of the MRI (resolution 156 μm × 155 μm × 155 μm), for each time point. In the other mice, the images had no sufficiently defined contours to properly establish a complete size distribution of the metastases (see Fig A). However, segmentation of the largest metastasis for each mouse at day 19 could be performed. In the mouse where number and size of the lesions could properly be assessed, the smallest detectable metastasis had a volume of 0.05 mm^3, which was taken as the minimal visibility threshold. We defined a macro-metastasis as a metastasis having a size larger than this value. Results of the model simulation for the metastatic size distribution are reported in Fig 6.4A, together with the experimental data. Inter-animal variability was simulated using population distribution of the parameters (lognormal distribution and coefficients of variation reported in Table 7.1), retrieved from the population mixed-effects fit. The maximal volumes predicted by the model/standard theory were considerably smaller than those observed by MRI. For example, at $T = 19$ days, while the total metastatic burden was similar in the data and in the model (Fig 6.2), the macro-metastatic burden was three-fold larger in the data than in the model's average prediction (6.4B), and the largest metastasis five-fold larger. At $T = 26$ days, although macro-metastatic burdens were similar in the

Par.	**Unit**	**Median value (CV %)**	**95% CI**
α	day^{-1}	0.417 (41)	(0.329–0.557)
β	day^{-1}	0.106 (45.1)	(0.0372–0.145)
μ	cell^{-1} · day^{-1}	9.721×10^{-6} (21.7)	$(1.81\text{–}52.21) \times 10^{-6}$

Table 6.1: CI = confidence interval. CV = Coefficient of Variation, in per cent = $\frac{\text{std}}{\text{est}} \times 100$, with est and std respectively the median value and standard deviation of the estimated lognormal population distribution of the parameters resulting from the nonlinear mixed-effects statistical estimation procedure.

data and in the model, the standard theory predicted that the largest tumor would have a volume of only 1.14 mm^3 in average (std = 0.755 mm^3), while the largest observed metastasis had a volume more than 10 fold larger (13.6 mm^3). This was compensated by a considerably larger number of metastatic lesions in the model (95.4 ± 47 versus 11 in the data). For each of these quantities, the $p-$value (probability to obtain the data value – or larger – under the null hypothesis that the data would have been generated by the model, evaluated numerically) was statistically significant ($p < 10^{-5}$).

These conclusions are limited by the fact that the entire time course of metastatic size distribution of only one mouse was available for reliable comparison with the model. However, in all the 6 mice, the size of the largest metastasis at day 19 could be measured and ranged (0.45–12) mm^3, which was significantly larger than the model predictions (6.4C, $p < 10^{-5}$ by the z-test). To give an idea, the largest metastases predicted by the model ranged $(9.5 \times 10^{-4}\text{–}0.3)$ mm^3. This strongly suggested that the standard theory was not able to describe the volumes of individual foci. Moreover, even without statistical comparison of the model's predictions to the empirical data, the numbers predicted by the model (in particular the number of macro-metastatic lesions at day 26) seem highly unrealistic. To assess the robustness of our results regarding several assumptions of the model, we investigated varying several parameters. First, metastases might initiate from a size larger than from one cell [18, 19]. We therefore performed the entire analysis for different values of V_0 (see discussion and Fig A.3), and found similar inconsistencies with the data in terms of largest metastases. Data-consistent and biologically plausible results in terms of number of metastases would require initial sizes larger than 100 cells, which is biologically unrealistic in view of the size of capillaries and experimental works that demonstrated that tumor cell clumps comprise less than 10 cells [18]. Similar results demonstrating inconsistency of the standard theory were also obtained when re-performing the analysis for variable values of the parameter γ (Fig A.4).

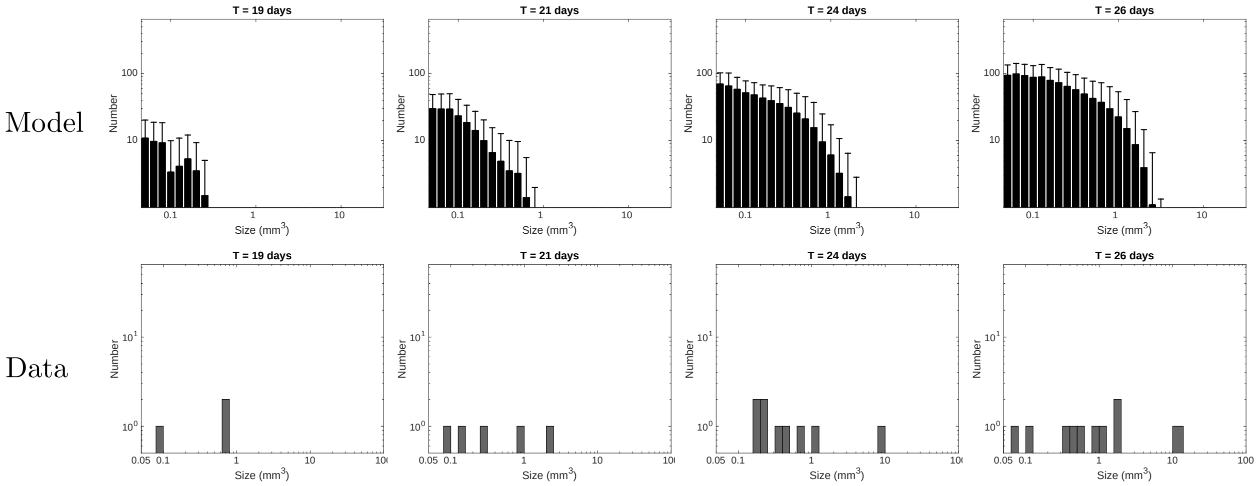
These results strongly indicate that the standard theory of metastatic progression as described by the model employed here (i.e., dissemination and independent growth), when calibrated to data of total metastatic burden, was in contradiction with the experimental observations with regard to the number of metastatic foci and their size distributions. It is beyond the scope of the present work to elaborate (and validate against the data) a unified model able to recapitulate the behaviour of metastatic tumors during the colonization process. However, as a first step toward this objective, we put forward two assumptions to correct the inconsistency of the standard theory:

- non-trivial interactions between metastases
- interactions between the metastatic foci and the circulating tumor cells (cells attraction and aggregation)

6.3 The two hypotheses: merging and attraction

6.3.1 Merging

We indeed observed merging of two metastases in our data (between days 21 and 24, see Fig 6.5) and therefore decided to investigate this further.



T	19		21		24		26	
	Data	Model	Data	Model	Data	Model	Data	Model
Number	4	4.8 (2.7)	5	17.6 (9.52)	9	52.9 (27.5)	11	95.4 (47)
Macro-burden (mm^3)	1.58	0.524 (0.361)	3.55	2.47 (1.8)	11.5	12.2 (9.94)	21.2	28.6 (24.4)
Size of largest met (mm^3)	0.78	0.143 (0.089)	2.19	0.326 (0.221)	8.17	0.766 (0.535)	13.6	1.14 (0.755)

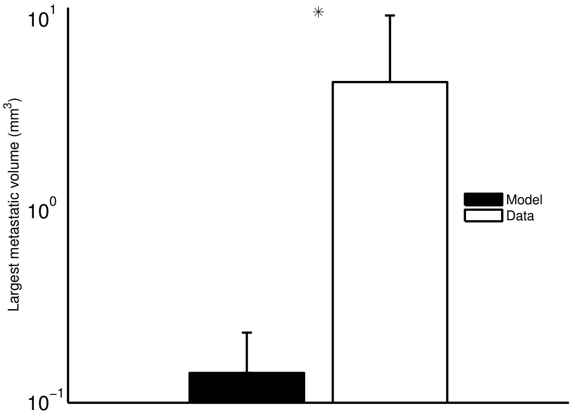
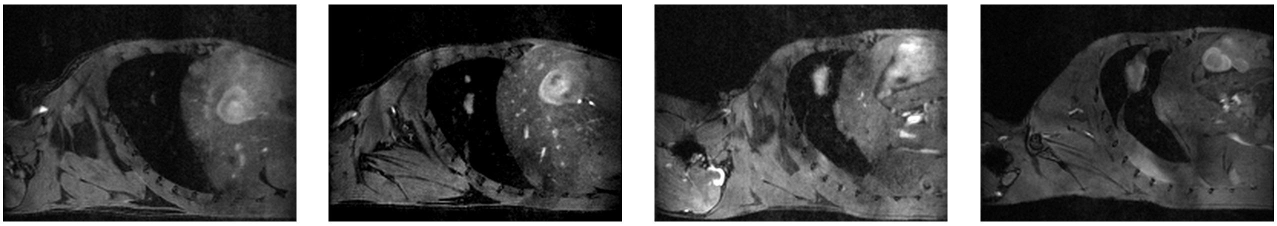


Figure 6.4: (A) Top row: Simulation of the mathematical formalism of the standard theory (i.e. dissemination and independent growth of the resulting tumor foci), using the parameter values inferred from the data of the total metastatic burden (total GFP signal in the lungs). Only tumors larger than the visible threshold at MRI (0.05 mm^3) are plotted. Simulations were obtained using equations (6.1)-(6.2) for the time evolution of the density of secondary tumors, endowed with a lognormal distribution of the parameters for inter-animal variability, with the following values (retrieved from the population mixed-effects fit, median ± standard deviation): $\alpha = 0.417 \pm 0.171$ day^{-1}, $\beta = 0.106 \pm 0.0478$ day^{-1} and $\mu = 9.72 \times 10^{-6} \pm 0.428 \times 10^{-6}$ cell^{-1} · day^{-1}. Shown are the results of 1000 simulations, mean ± standard deviation. Bottom row: Observations of macro-metastases numbers and sizes in one mouse on MRI data. (B) Comparison of several metrics derived from the metastatic size distributions. For the model, numbers are represented as mean value and standard deviation in parenthesis. The data corresponds to the mouse presented in the upper histogram. (C) Comparison of the largest metastatic size at day 19 between model ($n = 1000$ simulated animals) and observations ($n = 6$ animals), log scale. The observed largest metastases are significantly larger than simulated ones ($p < 10^{-5}$ by the z-test).



Day 19 Day 21 Day 24 Day 26

Figure 6.5: From left to right: Sagittal slide of the lungs from Day 19 to Day 26 of the same mouse: Two tumors are growing close to each other and they merge between Day 21 and 24

If metastases confined in a limited space can merge when by passive motion due to proliferation when they are growing in close vicinity, it could result in a different metastatic distribution, with less but larger metastases than described in the standard theory. Would it be sufficient to obtain such different distributions as in the data? To answer this question, we proposed a little extension of the standard theory described by Iwata's model. It consists in the same model describing birth of metastases growing following a given parametrized growth law (Gomp-exp in our case), but this time, metastases appear in a space limited organ, the lungs in our case. They appear randomly in the lungs with an equal probability for each point of the lungs. This assumption is wrong because it is known that metastatic cells arrive in particular sites of the organ like bronchioles and distal alveoli [10], but it makes the model simpler. To make the metastatic cells appear randomly into the lungs, we meshed mice 3D lungs segmented from mouse MRI data, as shown in Fig 6.6.

Figure 6.6: 3D mesh of mice lungs retrieved from MRI segmentation

Each metastasis, starting from one cell, was considered as spherical and was located by the coordinates of its center. For each couple of metastases, they were considered to have merged if the sum of their radii (recovered from scalar simulation of the volume using the growth law) was greater than the distance between their centers. A new distribution of metastases can be computed in gathering all the aggregates of merging metastatic tumors. An other hypothesis is that metastases that enter in contact do not impair their respective growth, which is probably also a little simplistic. Contact interactions between merging metastases are considered in the next chapter. In this new model, the total burden is the same as in the former model, but because the location of each metastasis is random, the size distribution is stochastic as well. In order to test if this simple model of merging was able to better describe the observed metastatic size distributions at days 19, 21, 24 and 26, we performed 10 simulation (not more because computing time of the algorithm was long) of this new model with random locations of the metastatic seeds. The parameters α, β and μ have been fixed to the median value retrieved from the population fit (see the table 7.1). The results of these simulations, that is mean and standard deviation (on the 10 simulations) of the number of macrometastases, the macro-burden and the size of the largest met are presented in the table 6.3.1.

T	19			21		
	Data	Standard theory	Merging theory	Data	Standard theory	Merging theory
Number	4	5	5 (0.309)	5	23	23 (0.707)
Macro-burden (mm^3)	1.580	0.520	0.528 (0.0212)	3.550	2.538	2.539 (0.00250)
Largest met (mm^3)	0.780	0.146	0.146 (1.15×10^{-4})	2.190	0.308	0.344 (0.0973)

T	24			26		
	Data	Standard theory	Merging theory	Data	Standard theory	Merging theory
Number	9	55	51 (2.345)	11	102	95 (3)
Macro-burden (mm^3)	11.5	10.420	10.511 (0.0487)	21.2	23.366	23.594 (0.0786)
Largest met (mm^3)	8.17	0.693	0.797 (0.0866)	13.6	1.0606	1.386 (0.179)

Table 6.2: Results of the model including merging of metastases are compared to the observations and to the results of the standard theory. Results of the merging theory, represented as mean value and standard deviation in parenthesis, were retrieved from 10 simulations with random locations of the metastatic foci in the lungs. The parameters for standard and merging theories are the following ones: $\alpha = 0.417$ day^{-1}, $\beta = 0.106$ day^{-1} and $\mu = 9.72 \times 10^{-6}$ cell$^{-1} \cdot$ day^{-1}.

As we can see, despite slight increases of the largest metastases and the macrometastatic burden, as well as a reduced number of macrometastases, the merging theory was far to recover the observations, which suggests that other phenomena may occur. We remark that substantial differences between the standard theory and the merging theory in metastasis number, macro-burden and metastatic size, occured rather late, that is at days 24-26.

6.3.2 Attraction

Evidences of cells detaching from metastasis and then attracted to the primary site have been experimentally shown, a phenomenon called self-seeding [11]. Chemokine-mediated attraction of tumor cells to premetastatic sites are mentioned in [10]. Joyce and Pollard also mentioned that chemokines mediate tissue tropism by tumor cells attraction to specific sites [44]. The attraction of tumor cells to specific metastatic sites could result in a different metastatic size distribution than described by the standard theory. Indeed, a tumor cell entering into the lungs could join an existing metastatic niche instead of forming a new metastasis, resulting potentially in less but larger metastases.

In order to assess if such a phenomenon could occur, two experimental protocols have been proposed. In the first one, red-tagged and green-tagged RENCA cells are injected with equal proportions into the renal subcapsular space of mice. The goal is to observe if mixed-color metastases are present in the lungs, which would show that metastases are not only clonal from a single surviving cell (standard theory). Such an experimental protocol has already been used in [19] for breast cancer cell lines. In this study, such mixed-color metastases are due to circulating tumor cell clusters establishing new metastases. However, in order to show evidences of metastatic cells attraction, we proposed a second experiment. In this one, a first injection of red-tagged RENCA cells into the renal subcapsular was performed, followed by intravenous injection of green-tagged RENCA cells in the tail vein a few days after. The cells injected into the tail vein going directly into the lungs, we could observe if the cells injected into the tail vein join existing red-tagged metastatic foci or establish independent foci. Some of the first results are presented in Fig 6.7.

In these first results, we could observe mixed-color metastatic foci into the lungs in the two experimental protocols. Moreover, in these lung slices some metastatic foci were constituted of a one-colored (red in Fig 6.7) core surrounded by cells of the other color, which could be explained by a recruitment of intravenously injected green cells by the established red metastatic foci.

These preliminary results have to be confirmed and validated by further experiment to properly establish that established metastatic foci can attract or recruit new arriving metastatic cells.

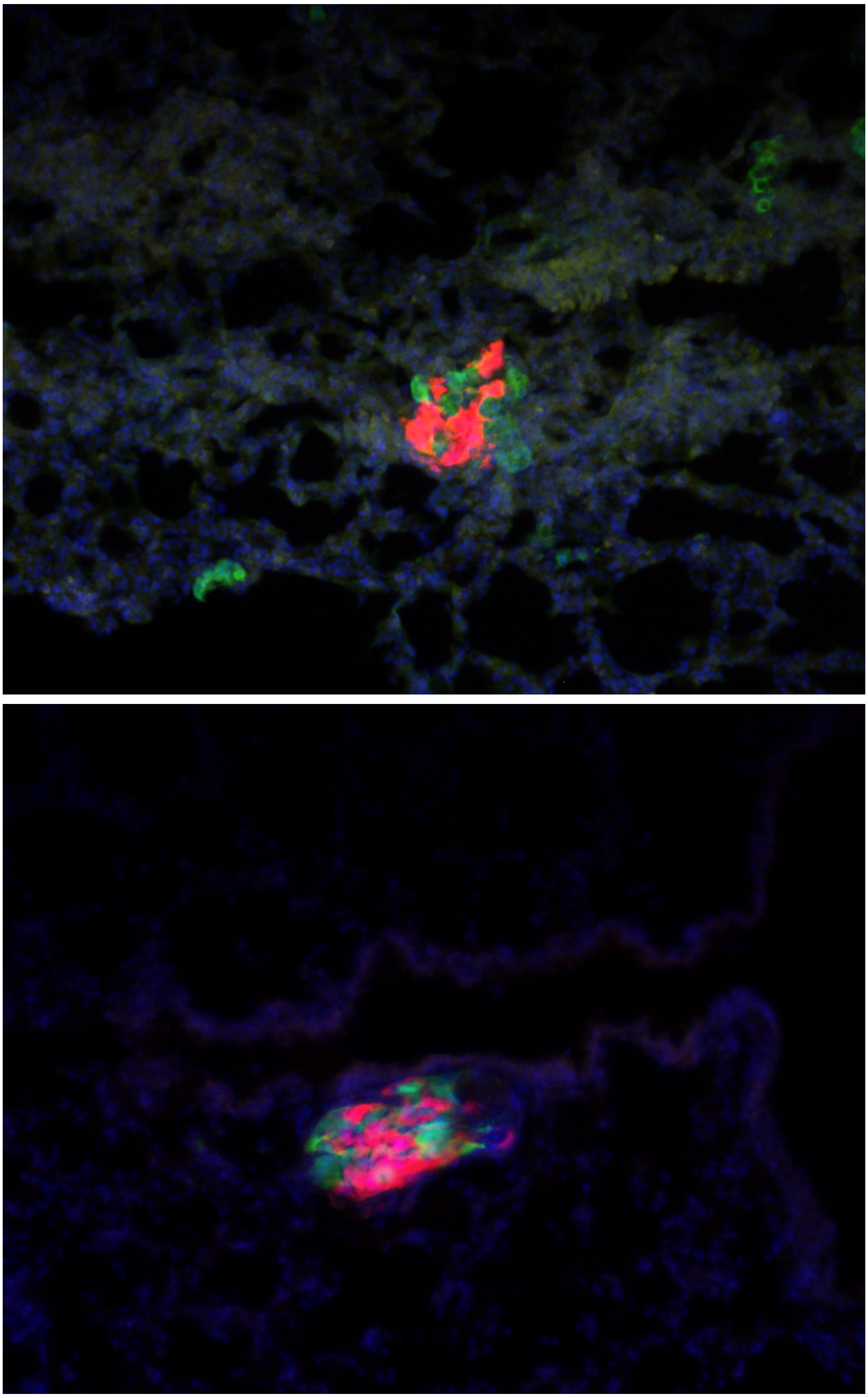


Figure 6.7: Top: lung slice in the case of renal subcapsular injection of green-tagged RENCA cells followed by intravenous injection of green-tagged RENCA cells into the tail vein. Bottom: lung slice where red-tagged and green-tagged RENCA cells are injected with equal proportions into the renal subcapsular space.

6.4 Discussion and perspectives

6.4.1 Discussion

Using a combined approach between experimental data and mathematical models, we demonstrated that the standard theory of metastasis formation and growth, where metastases grow independently from the rest of the system, was biologically unlikely. To explain our findings, we proposed several hypotheses, including the possibility of metastatic foci merging by passive motion and metastatic cells attraction. First experiments confirmed that such phenomena are likely to occur during metastasis development. We also proposed an extension of Iwata's model (the discrete version) including space limiting of metastatic growth and possible merging of metastatic foci by passive motion due to proliferation. This model, despite slight improvements, was not able to describe the visible metastatic size distributions, which could mean that other phenomena like metastatic cells attraction and aggregation are in part responsible of such metastatic patterns.

Based on the rationale that lung capillaries have a diameter of the order of one tumor cell (20μm) and that metastatic cells have lost expressions of cell-cell adhesion proteins such as cadherins [5], we assumed in our simulations, that metastases originated from one cell. This might be arguable and metastasis could start from tumor cell clumps [82, 19]. To resolve this further and assess the robustness of our results, we performed the entire data analysis (fit of the total metastatic burden and resulting prediction of the metastatic size distribution) for values of the initial number of cells of 1, 10, 100 and 500 (Fig A.3 of the appendices). Initial numbers of 10, 100 and 500 cells could be in agreement with the data at day 19. However, with $V_0 = 10$ cells, the predicted number of macro-metastases at day 26 was $3-$fold higher than in the data. For $V_0 = 100$ cells and $V_0 = 500$ cells, the predicted macro-burden was $2-$fold smaller than the observed one. Moreover the largest metastasis at day 26 was still predicted much smaller in the model than in the data (3.11 mm^3 for $V_0 = 10$ cells, 3.58 mm^3 for $V_0 = 100$ cells, 3.8 mm^3 for $V_0 = 500$ cells, against 13.6 mm^3 in the data). Furthermore, in animal experiments the vast majority of detaching tumor cell clumps has been shown to comprise less than 10 cells [82] with a range of $2-50$ cancer cells [19], which makes the theories $V_0 = 100$ cells and $V_0 = 500$ cells unlikely. This suggests that, if the metastases started from a substantial amount of cells, the grouping of these cells probably occurred at the distant site, after extravasation from the blood circulation. Similarly, we did not consider any cell loss at the moment of initial sub-capsular injection. We could make theoretical assumptions of cell loss (of 10%, 20%, etc...), which would simply consist in multiplying by the relevant factor. For instance, considering a 90% loss (i.e. that only 10% of the cells remain viable) would be equivalent to multiplying by 10. As demonstrated in supplementary Fig A.3 of the "Apendices" section, it is necessary to assume an initial size of at least $100V_0$ to recover plausible values for the number of metastases at time days. Combining the two (cell loss of 10% and initial metastatic size of 10 cells) thus gives a hypothesis that we are not able to infirm given the data we dispose.

To explain the inconsistencies of the standard theory with observations, we proposed a first hypothesis, which consisted in merging of metastases. This phenomenon has been experimentally validated by MRI data 6.5. A mathematical model consisting in a simple spatial extension of the standard theory has been proposed, allowing metastases that are growing in close vicinity to merge, giving rise to one only metastases whose the volume is the sum of the merging metastases. Including this spatial aspect to the standard model made the metastatic distribution go to the "good way", that is less but larger metastases. However, these improvements are too slight to recover the observed distributions. Indeed, we remarked that substantial changes between standard theory and merging theory are visible from day 24. It means that passive motion by proliferation alone is not efficient enough to make the metastatic foci aggregate fast enough. An other phenomenon is likely to occur, like cells attraction. In the future, we also could test the merging theory for initial metastatic volumes larger than one cell.

In the model including merging, when two metastases merge, it does not impact their respective growths and the global dynamics is the same. It alters the metastatic size distribution only. However, when two metastases enter in contact, mechanical interactions may occur, altering their respective growths. The next chapter of this thesis is focused on these quantitative aspects.

The inability of the merging theory to explain the metastatic size distribution may indicate that besides merging by passive motion due to proliferation, other mechanisms such as chemokine-mediated cells attraction occur [10, 141]. Circulating tumor cells may be attracted by some established niches and explain the abnormally fast volume expansions that we observed. Indeed, such chemokine-mediated attractions are presumed to play an important role for the pre-metastatic and metastatic niches establishment, in mediating myeloid and tumor cells attraction [10, 141, 142]. Moreover, chemo-attractants may play a role in tissue tropism of metastatic cells [44]. Chemotactic gradients can attract metastatic cells that express the chemokine receptor to specific locations. In the future, additional phenomena such as aggregation and recruitment of cells during the metastatic process from the circulation should be integrated in the standard mathematical model. Another phenomenon that could possibly explain the observed volumes would be the presence of circulating tumor cell clusters that would give rise to metastases [19]. Indeed, Aceto et al. recently showed in a breast cancer animal model that metastases do not originate from single cells only but also from tumor cells clusters that have a higher metastatic potential than single cells. However, they did not show evidence of this phenomenon for kidney cancer and in their experiments, clusters were formed by at most 50 cells. As indicated above, this order of magnitude of the initial cell numbers that colonizes the lung is not able to describe the dynamics of metastasis formation in our model and experimental data.

In order to assess if such attraction of metastatic cells was possible we proposed experimental protocols implying injections of RENCA cells with multiple color-tagging. Similar experimental protocols have been already performed in [19, 11]. The first experimental results suggested that such recruitment of metastatic cells by established metastatic foci is possible, as we observed red-tagged core of RENCA cells surrounded by green-tagged RENCA cells. For now, these results need to be confirmed with more experiments. Transcriptomic analyses could be then made to identify factors that would be responsible of metastatic cells recruitment/attraction. Whether this phenomenon could explain the observed metastatic size distribution is a quantitative question that requires further work. This question could be possibly treated with a modeling approach. It could have clinical implications as factors mediating cells attraction could be targeted.

Taken together, our results suggest that observed metastatic volumes in the lungs (and possibly in other organs) do not result from proliferation only but also from interactions between metastases and interactions between metastatic foci and circulating tumor cells. Our methodology and results illustrate, furthermore, how a combined approach using multimodal biological data on one hand, and multimodal modeling analysis on the other, provides powerful insights into tumor biology and, in particular, into the metastatic process.

6.4.2 Biological perspectives

The phenomenon of tumor cells attraction by metastatic foci and niches has to be confirmed with further experiments. The two protocols that have been tested (injection of a mixed-color RENCA tumor in the renal sucapsule and injection of red-tagged RENCA cells in the renal subcapsule followed by intravenous injection of green-tagged RENCA cells) have to be performed on more animals. In the first experiment, it would be interesting to compare the sizes of monocolor (so monoclonal) foci with multicolor foci in order to know if phenomena of cells aggregation can be responsible for the spectacular increases in size. In the second experiment, it would be interesting to quantify the proportions of intravenously injected RENCA cells that respectively form independent foci or join existing ones. These quantitative questions could help to know if the phenomenon of cells attraction is a major one during the metastatic development. Another experiment, more clinically realistic, could consist in injecting red-tagged RENCA cells intravenously first, and then (few days latter) to inject green-tagged RENCA cells into the renal subcapsule. It would indicate if "real" circulating tumor cells coming from a primary tumor are attracted by existing metastatic foci. *In vitro* experiments could also be performed with cocultures of metastatic cells extracted from the lungs with RENCA cells extracted from the primary tumor in order to observe if metastatic cells attract tumor cells. Finally, if these experiment confirm the phenomenon, a transcriptomic analysis should be performed in order to identify the factors responsible for this attraction, opening possible clinical perspectives.

6.4.3 Future modeling perspectives

This study highlighted limitations of the standard theory of metastasis development to describe *in vivo* metastatic size distributions. Further modeling efforts could be devoted to incorporate components of cells attraction and aggregation, including tumor self-seeding, in Iwata's model. Such a model could be useful if predictions on metastatic dynamics have to be made on the basis of imaging data giving information on the visible metastatic size distribution.

Another modeling clue would be to develop a "spatial version" of the discrete derivation of Iwata's model. Indeed, if a simple growth law as the Gompertz one is used to describe the growth of each metastasis, a spatial derivation could be possible in using the spatial version of the Gompertz model based on the equation (5.21) of the chapter 5. the tissue density of each metastasis would follow a conservation law as (5.21), and each metastasis would birth from one cell appearing randomly in a 3D organ. Metastases could then merge but deformations of the tissue due to the pressure would be taken in account. In order to assess if chemokine-mediated attraction of cells could explain observed metastatic size distributions, a chemotaxis component could be incorporated into the conservation laws, without altering the total mass dynamics.

Chapter 7

Modeling metastasis merging and tumor-tumor mechanical interactions

Published in Plos Computational Biology [1]

Inconsistencies of the standard theory of metastatic development highlighted in the previous chapter led us to consider the hypothesis that metastatic tumors could merge and give rise to larger metastases. In our MRI data, the phenomenon of metastases merging has indeed been observed (Fig 6.5). This led us to investigate the effect of spatial interactions between merging metastatic tumors on the dynamics of the global metastatic burden. In order to study possible mechanical interactions between merging tumors, we derived a mathematical model of spatial tumor growth, confronted it with experimental data of single metastatic tumor growth, and used it to provide insights on the dynamics of multiple tumors growing in close vicinity. The spatial model we derived was then used to explore the quantitative implications of tumors merging on global tumor growth and to estimate the numbers of required merging metastatic foci in order to obtain the observed metastatic volumes mentioned in the previous chapter, which are unlikely to result from proliferation only. More specifically, we wanted to address the following questions: do spatial interactions have an impact on the dynamics of the total metastatic burden? To what extent could these interactions correct the theoretical predictions of the unlikely fast growth rates observed in section 6.1 of the chapter 6? Answers to these questions have implications on future theoretical models of metastatic development. The possibility of merging for two neighboring metastases introduces a spatial aspect of metastatic colonization and, therefore, required a spatial modeling approach. We derived such a model which had to full-fill the following requirements: 1) it should be based on biological knowledge of macroscopic tumor growth (retrieved from the literature), 2) it should remain as parsimonious as possible (minimal number of parameters) and 3) it should be able to fit our spatial growth data.

7.1 Modeling metastasis merging

7.1.1 Without spatial aspect: metastatic burden fractionation

We first made the simplest assumption: merging of tumors results in addition of their volumes. We made a second assumption: each metastasis follows an exponential growth law: $V(t) = V_0 e^{\lambda t}$. In this case, due to the linearity of the model as a function of V_0 , a tumor starting from size $2V_0$ has the same size as two independent tumors starting from size V_0. (Fig 7.1).
However, *in vivo* tumor growth exhibits increasing doubling times when increasing in size [58, 108, 60]. If we consider the volume dynamics and model this fact using the Gompertz law, then, due to the nonlinearity of the model, merging two tumor lesions into one (in summing their volumes) leads to a slower growth and ultimately smaller volume compared to two tumors that are growing separately (Fig 7.1). The same result can be observed for more than two tumors. It means that the more fractionated the total burden is, the more fast is the growth. Larry Norton makes the same remark to emphasize his hypothesis that thanks to self-seeding, a tumor is fractionated in multiple seeds and then grows faster [50].

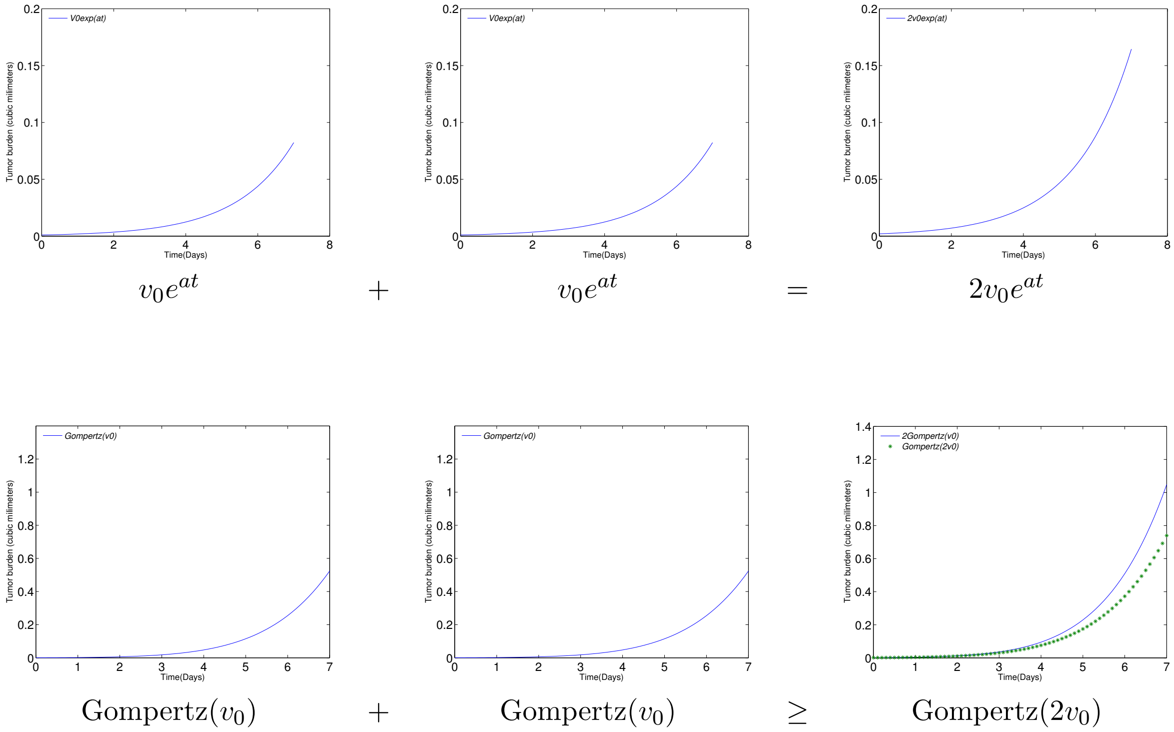


Figure 7.1: Top: two tumors growing under an exponential law lead to the same final total volume than one tumor with the same initial total volume; Bottom: two tumors growing under a Gompertz law lead to a larger final volume than one tumor with the same initial total volume

Yet, under the Gompertz law, spatial interactions are not taken into account, while they are probably of great importance. This motivated us to study the global impact of spatial interactions and to propose a spatial model of tumor growth.

7.1.2 Advection-type modeling of proliferation-mediated movement and pressure-mediated inhibition of tumor growth

The two-dimensional model we used for the spatial growth has already been presented in the chapter 5. It describes the tumor expansion thanks to a conservation equation on the tumor density P. The medium is considered as a porous medium and the tumor expansion is described by a velocity field representing the passive motion due to proliferation. Let us recall the equations of the model:

$$\frac{\partial P(t,x)}{\partial t} + \nabla \cdot (\mathrm{v}(t,x)P(t,x)) = \gamma(t,x,P,S,\Pi)P(t,x) \tag{7.1}$$

$$\frac{\partial S(t,x)}{\partial t} + \nabla \cdot (\mathrm{v}(t,x)S(t,x)) = 0 \tag{7.2}$$

$$-k\Delta\Pi = \gamma P \tag{7.3}$$

$$\mathrm{v} = -k\nabla\Pi \tag{7.4}$$

$$\tag{7.5}$$

Again, we consider a porosity constant porosity k and assumed the domain Ω large enough to consider the pressure on the boundaries as equal to the homeostatic pressure of the body, leading to a Dirichlet condition on the pressure:

$$\Pi = \Pi_{eq} \text{ on } \partial\Omega,$$

where Π_{eq} is the homeostatic pressure of the body, which corresponds, as mentioned by Perthame et al. [69], to the conditions where pressure-mediated proliferation inhibition is minimal. The size of the computational domain Ω was once again fixed to the order of magnitude of mouse lungs ($\simeq 1cm^3$).
In this model, taking γ constant leads to exponential growth of the tumor burden. However, *in vivo* growth can depend on environmental conditions, leading to increased doubling time when the conditions are not optimal. Similar models using more cellular species were used by Ribba, Colin and Schnell in [149] to predict efficacy of radiotherapy, and by Colin, Saut and colleagues in [122] to describe avascular tumor growth. In the latter work, lack of nutrients and hypoxia were considered as essential limiting factors for growth and hence included in the model, thus introducing a supplementary variable (nutrients concentration or the vasculature density). However, in our study, because we wanted to keep the model as parsimonious as possible, we focused on a more phenomenological way to describe the fact that the tumor doubling time increases with the tumor size. The natural environmental variable being the pressure, a simple way to formalize this was to model the proliferation rate as a decreasing function of the pressure. This yielded to a simple model that captured the essential features of tumor expansion and was able to describe *in vivo* tumor growth.
As in the chapter 5, and following the work of Montel et al. [125] and Stylianopoulos et al. [15, 16], we considered that the growth rate of the tumor tissue decreases with the pressure exerted on the tissue. Therefore, we modelled the growth rate with a decreasing exponential law [125]:

$$\gamma(\Pi) = \gamma_0 \exp\left(-\frac{\Pi}{\Pi_c}\right), \tag{7.6}$$

where Π represents the pressure field, γ_0 the maximal proliferation rate, and Πc a characteristic pressure. Under the assumption of a constant porosity, the value of k has no impact. Indeed, as long as the product $k\Pi_c$ remains constant, the solution remains unchanged. That is why we fixed the value of k to 1. Moreover, the boundary condition is taken homogeneous: $\Pi_{eq} = 0$, which means that the homeostatic pressure of the body is the optimal pressure of proliferation.
Under such a model, high pressure provokes decreased proliferation, but not apoptosis. Montel et al. suggested that mechanical stresses have a poor effect on apoptosis, but also indicated that this could depend on the cell line and the experimental protocol [125].

7.1.3 Merging or pushing?

What kind of interactions take place when two tumors enter in contact? The corresponding literature is not so abundant. We tried therefore to base our study on reasonable physical considerations. What is the behavior of two tumors in contact? Does the pressure at the interface become high, resulting in tumors pushing each other? Or is it quite the opposite, the tumors attracting each other when they are close? It seems to be reasonable to think that it depends on the nature of the tumors (the organ tissue, primary or metastatic tumor, dense or diffuse).

Pushing

Let us study the behavior of the model based on the equations (7.1)-(7.4) when the initial condition is two tumors that are in close vicinity, as in the Fig 7.2. Let us recall the equations:

$$\frac{\partial P(t,x)}{\partial t} + \nabla \cdot (\mathrm{v}(t,x)P(t,x)) = \gamma(t,x,P,S,\Pi)P(t,x)$$
$$\frac{\partial S(t,x)}{\partial t} + \nabla \cdot (\mathrm{v}(t,x)S(t,x)) = 0,$$

The goal is to study the possibility for both tumors to merge. In other words, going from the configuration of two disconnected domains Ω_1 and Ω_2 that represent the two tumors, to one connected domain Ω that represents the tumor resulting from merging of the both initial tumors, as illustrated by Fig 7.2.

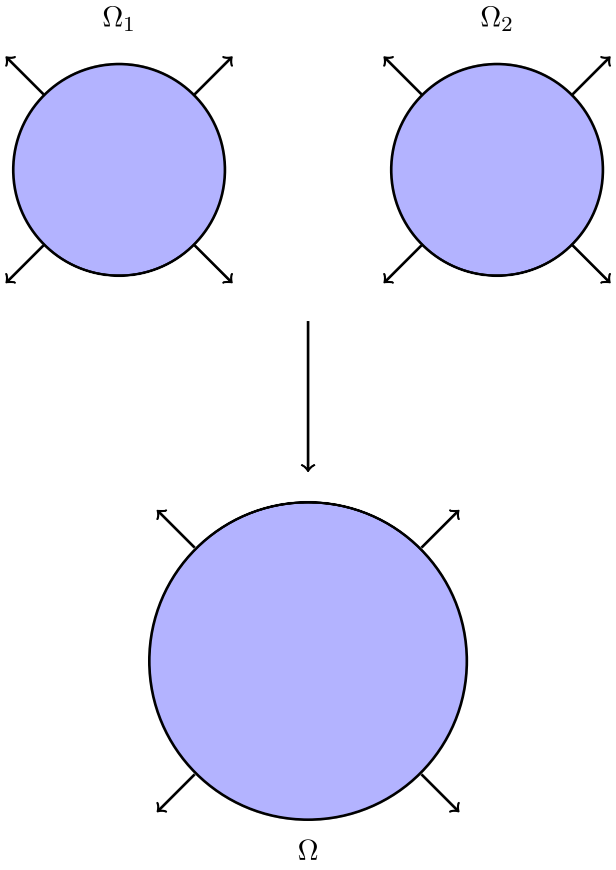


Figure 7.2: Scheme representing two tumors that are merging by passive motion due to proliferation

In order to simplify the study, we make the assumption of a tumor density equal to 1 inside the tumor and equal to 0 outside the tumor. Let us recall the characteristics equation:

$$\frac{\partial X(t,y)}{\partial t} = \mathrm{v}(t,X(t,y))$$
$$X(0,y) = y$$

Along these trajectories, the tumor density is governed by the following dynamics:

$$\frac{\partial \tilde{P}(t,y)}{\partial t} = \tilde{\gamma}(t,y)\tilde{P}(t,y)(1-\tilde{P}(t,y))$$

According to the Cauchy-Lipschitz theorem, if the velocity v is locally Lipschitz with respect to x, two characteristics with two different initial positions y_1 and y_2 can not cross each other.

If the both tumors merge, it implies that at a particular instant t, the two interfaces enter in contact in at least one point, as illustrated in Fig 7.3.

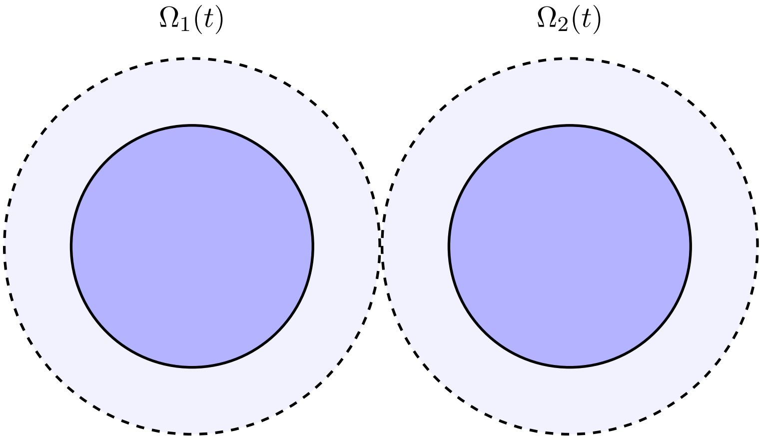


Figure 7.3: Contact in one point between two growing tumors

At this point, two characteristics $X(t, y_1)$ and $X(t, y_2)$ cross each other, with $y_1 \in \Omega_1$ and $y_2 \in \Omega_2$. Hence if the velocity field is regular enough, tumors can not merge under this spatial model. This is illustrated by a simulation performed with equations (7.1)-(7.6) and a bifocal initial condition (Fig 7.4).

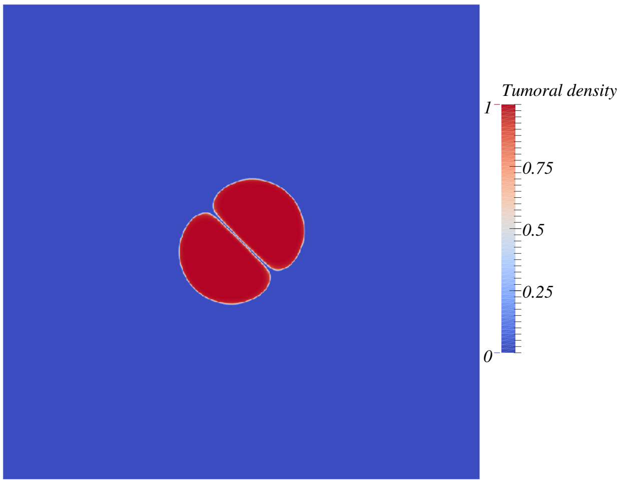


Figure 7.4: Illustration of tumors pushing each other. Simulations were obtained using Eqs (7.1)-(7.4) with the following parameter values: $\gamma_0 = 0.78$ day^{-1}; $\Pi_c = 0.0026$ Pa; time of simulation: $T = 7$ days

Here the two tumors push each other by the pressure exerted on the tissue between them, but do not merge.

Merging

Adding an invasive component in the model (7.1) allows tumors merging. We made a supplementary modeling assumption: in addition to the macroscopic passive motion due to proliferation, cells have their own motility. The active motion of each cell can be modeled by a random (brownian) motion [126]. So then, at the macroscopic scale, this active motion results in a diffusion component where the diffusion constant represents a macroscopic parameter reflecting the motility of the cells. The model now writes:

$$\frac{\partial P(t,x)}{\partial t} + \nabla \cdot (\mathrm{v}(t,x)P(t,x)) - \nabla \cdot (D\nabla P) = \gamma(t,x,P,S,\Pi)P(t,x) \tag{7.7}$$

$$\frac{\partial S(t,x)}{\partial t} + \nabla \cdot (\mathrm{v}(t,x)S(t,x)) - \nabla \cdot (D\nabla S) = 0 \tag{7.8}$$

$$P_{|\partial\Omega} = 0 \tag{7.9}$$

$$S_{|\partial\Omega} = 1 \tag{7.10}$$

This time, boundary conditions on P and S are necessary. The assumption of saturated flow $P+S=1$ is made again, and the relation on the velocity and the pressure are the same as the model without diffusion:

$$\nabla \cdot \mathrm{v} = \gamma P$$
$$\mathrm{v} = -k\nabla\Pi$$

This model contains now two phenotypic features of tumor cells: proliferation and invasion. The term $\nabla \cdot (\mathrm{v}(t,x)P(t,x))$ reflects the passive motion due to proliferation, whereas $-\nabla \cdot (D\nabla P)$ reflects the active motion due to cells motility. The model contains therefore a supplementary parameter, which is D, the cell diffusion constant. Such an invasive component allows tumor to merge, as illustrated in Fig 7.5.

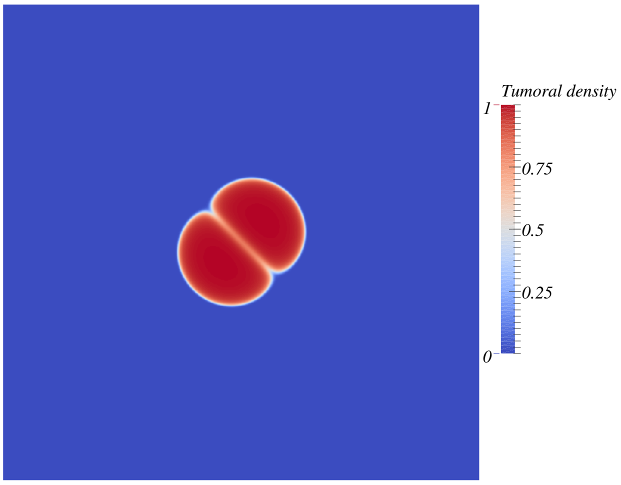


Figure 7.5: Illustration of tumors pushing each other. Simulations were obtained using Eqs (7.1)-(7.4) with the following parameter values: $\gamma_0 = 0.78$ day^{-1}; $\Pi_c = 0.0026$ Pa; time of simulation: $T = 7$ days

Integrating the equation (7.7) on the domain gives the equation on the tumor burden $\int_\Omega P\mathrm{d}x$:

$$\int_\Omega \frac{\partial P}{\partial t}\mathrm{d}x + \int_\Omega \nabla \cdot (\mathrm{v}P)\mathrm{d}x - \int_\Omega \nabla \cdot (D\nabla P)\mathrm{d}x = \int_\Omega \gamma P\mathrm{d}x,$$
$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega P\mathrm{d}x + \int_{\partial\Omega} P\mathrm{v}\cdot n_\sigma \mathrm{d}\sigma - \int_{\partial\Omega} D\nabla P \cdot n_\sigma \mathrm{d}\sigma = \int_\Omega \gamma P\mathrm{d}x,$$

where n_σ is the outgoing normal on $\partial\Omega$. Thanks to the Dirichlet condition, we know that the tumor does not reach the boundaries of the domain, which implies that the term $\int_{\partial\Omega} P\mathrm{v}\cdot n_\sigma \partial\sigma$ is equal

to zero. However, the term $\int_{\partial\Omega} D\nabla P \cdot n_\sigma \mathrm{d}\sigma$ is not necessary equal to zero and the equation on the burden $\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega P\mathrm{d}x - \int_{\partial\Omega} D\nabla P \cdot n_\sigma \mathrm{d}\sigma = \int_\Omega \gamma P \mathrm{d}x$ is different than in the case without diffusion.
If D is large enough, the term $\int_{\partial\Omega} D\nabla P \cdot n_\sigma \mathrm{d}\sigma$ cannot be negligible, representing the mass loss by diffusion across the boundaries. In this case, the burden dynamics is impacted by this outflow. In contrast, if the size of the domain is large enough over D, the term $\int_{\partial\Omega} D\nabla P \cdot n_\sigma \mathrm{d}\sigma$ is negligible and the burden dynamics is driven by the proliferative distribution only. The situation we want to model is the second one. The domain size being fixed to the order of magnitude of mouse lungs ($\simeq 1$ cm^3), we fixed the diffusion constant at a value retrieved from the literature. A range of values for this parameter based on measurements for embryonic cells is available in [126] : $2 \times 10^{-12} - 6.5 \times 10^{-11}$ cm$^2 \cdot$ s^{-1}. The parameter D has been taken in this range: $D = 10^{-11}$ cm$^2 \cdot$ s^{-1}. Considering the proliferation rate γ as a constant and fixing its value at the *in vitro* growth rate for the RENCA cell line ($\gamma = 0.67$ day^{-1}), we compared the outputs of the two models (with and without diffusion) with different initial sizes of tumor (1 mm^2 and 10 mm^2) to check if the dynamics were similar. The simulation are presented in Fig 7.6.

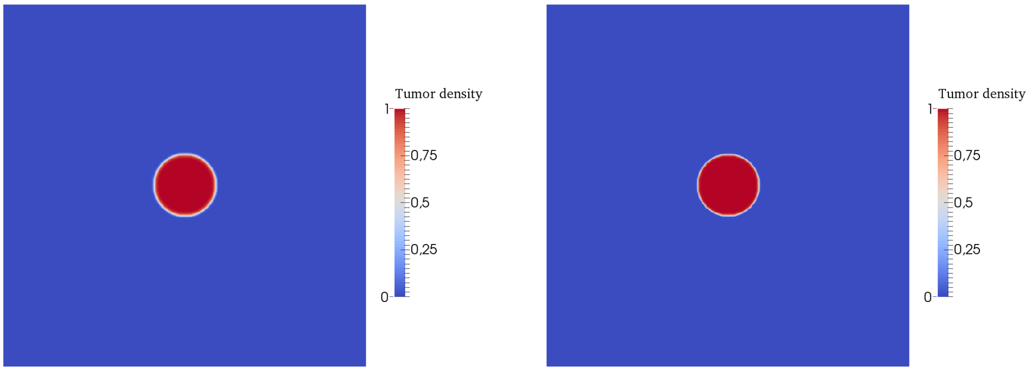


Figure 7.6: Comparison of between the model with and without cellular diffusion (invasion). Final tumor density distribution for an initial tumor surface of 10 mm^2; left: with diffusion; right: without diffusion. The proliferation rate γ is constant equal to 0.67 day^{-1}. The simulation time is $T = 1$ day.

The dynamics of the mass is presented Fig 7.7.

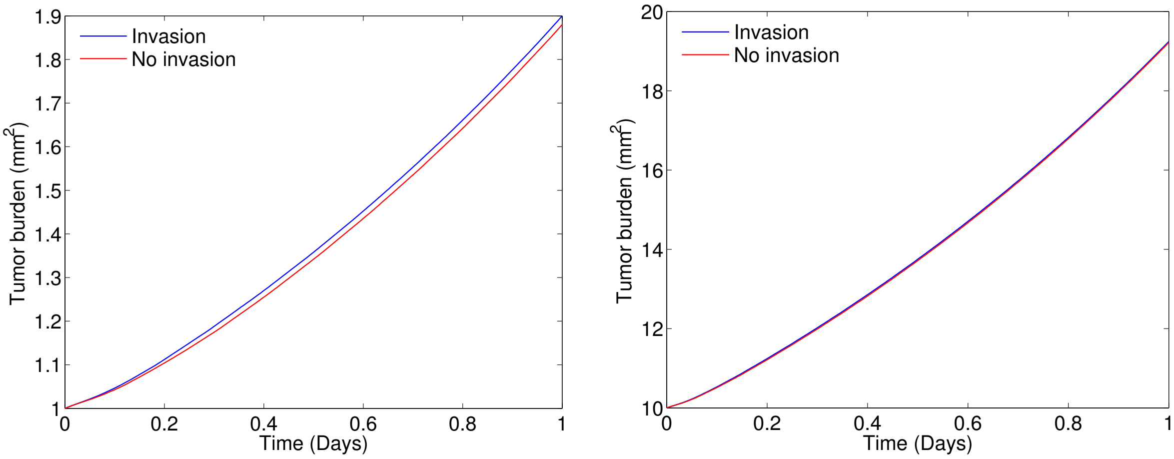


Figure 7.7: Comparison of between the model with and without cellular diffusion (invasion). Burden dynamics for two initial tumor surface; left: 1 mm^2; right: 10 mm^2; The proliferation rate γ is constant equal to 0.67 day^{-1}. The simulation time is $T = 1$ day.

The dynamics of both models are quite similar and exhibit an exponential dynamics, as expected (the growth rate being constant). Indeed, differences in the final burden were 1% for the small burden (1 mm^2) and 0.1% for the large burden (10 mm^2). We observe a larger difference for the smaller tumor, which makes sense as cellular diffusion is a small scale phenomenon having little impact on a growth at the tissue scale. The outflow due to diffusive component of the model is therefore negligible. Neglecting the motility of the cells seems therefore a reasonable assumption at the tissue scale.
In the case where the growth rate γ depends on time and space, the burden dynamics writes: $\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega P\mathrm{d}x + \int_{\partial\Omega} P\mathrm{v}\cdot n_\sigma \mathrm{d}\sigma - \int_{\partial\Omega} D\nabla P\cdot n_\sigma \mathrm{d}\sigma = \int_\Omega \gamma P \mathrm{d}x$. Even if the outflow $\int_{\partial\Omega} P\mathrm{v}\cdot n_\sigma \mathrm{d}\sigma$ is negligible, the proliferative distribution γP can be different in the both cases with or without cellular diffusion. Therefore, the burden dynamics are theoretically different too. This time, we performed simulations with the pressure-mediated growth model based the growth law (7.6). The parameters were extracted from the fits on MRI data presented in Fig 5.16 of the section 5.5 of the chapter 5. The two models with and without diffusion are compared again with the diffusion constant value extracted from the literature: $D = 10^{-11}$ cm^{2}s^{-1}. As illustrated by Fig 7.8, the differences are very small, with respectively 0.6%, 2.7%, 1.1% and 2.9% for the four parameter sets.

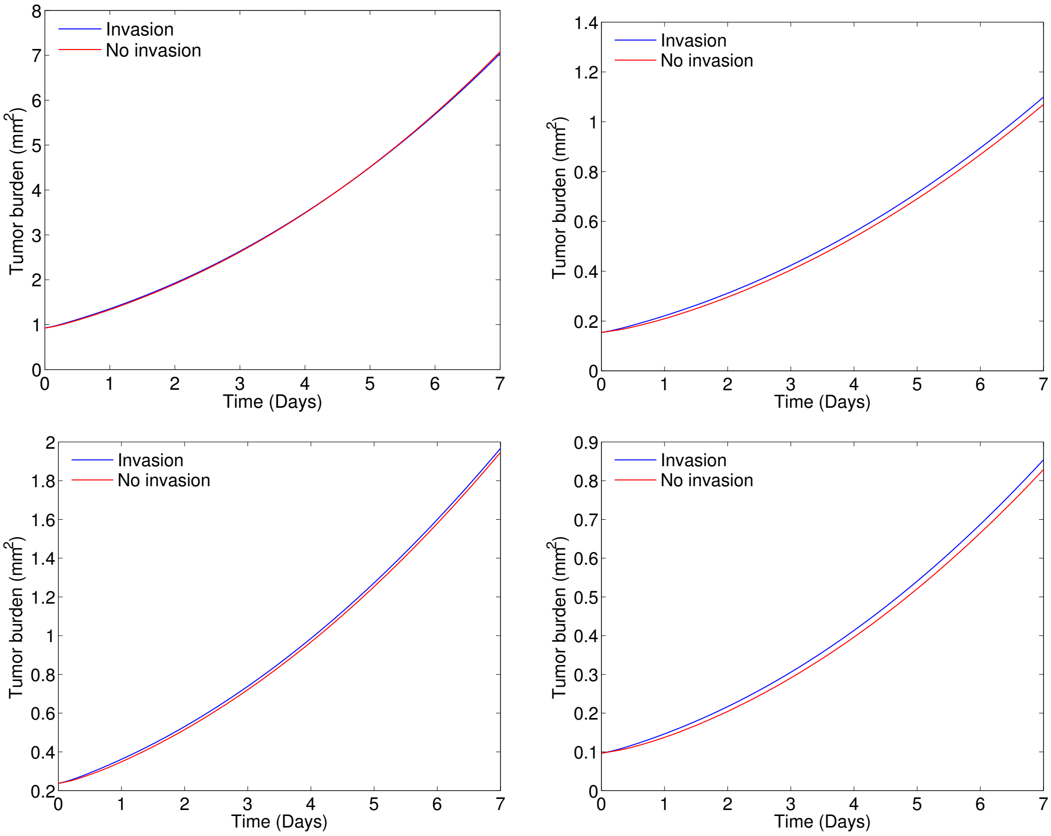


Figure 7.8: Comparison between the model with and without cellular diffusion (invasion) using the pressure mediated proliferation law (7.6). The parameters were retrieved from the fits on MRI data presented in Fig 5.16.

Although cells are known to exhibit a random (brownian) motion when left alone [126], cellular diffusion did not alter our results, so we neglected this diffusion in the study for computational cost reasons. Therefore, in our simulations, merging of tumors do not result from a physical diffusion but rather from the numerical diffusion, which is not a problem because we know that results would be quite similar in both cases.

	$\gamma_0(day^{-1})$	$\Pi_c(Pa)$
Mouse 1	0.78	0.0026
Mouse 2	1.01	0.00079
Mouse 3	0.67	0.00067
Mouse 4	0.8	0.00052

Table 7.1: Values of the parameters resulting from the fit of each metastatic dynamics. The value of γ_0 corresponds to the minimal doubling time in optimal conditions of pressure.The values of Π_c are not reliable to real pressure values because they depend on the porosity, which is not a parameter.

7.2 Quantitative study of spatial interactions between merging metastases

7.2.1 Model calibration

To perform the study, we first wanted to fix the parameters of the model to realistic values. We calibrated the model according to the growth of four metastases observed by MRI. These four tumors were selected because they were detectable on a sufficient number of time points (four for three of them and three for one of them) and were manually segmented. Fig 5.17 of the chapter 5 shows the MR images, the numerical simulation starting from the initial shape with the parameters fitted to the volume for one of the four metastases, and the dynamics of the simulated volume changes of the four metastases. A movie of the simulation on the MR image can be found at `http://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1004626#sec026` and spatial distributions of pressure (Π) and proliferation rate (γ) are presented in Fig A.5 of appendices. The model was able to describe the increase of the tumor volume for the four metastatic lesions with excellent goodness-of-fit. Table 7.1 presents the values of the two parameters for the four fitted growth curves. The fits were performed on the volume only, considering the metastases as spherical, which is a reasonable assumption because only slight differences on the mass are observed between the spherical and non-spherical cases (Fig A.6 of appendices and Fig 5.19 of the chapter 5).
These fitted parameters provided a range to perform the tumor-tumor contact interactions study: $(\gamma_0, \Pi_c) \in [0.67; 1.01] \times [0.00052; 0.0026]$.

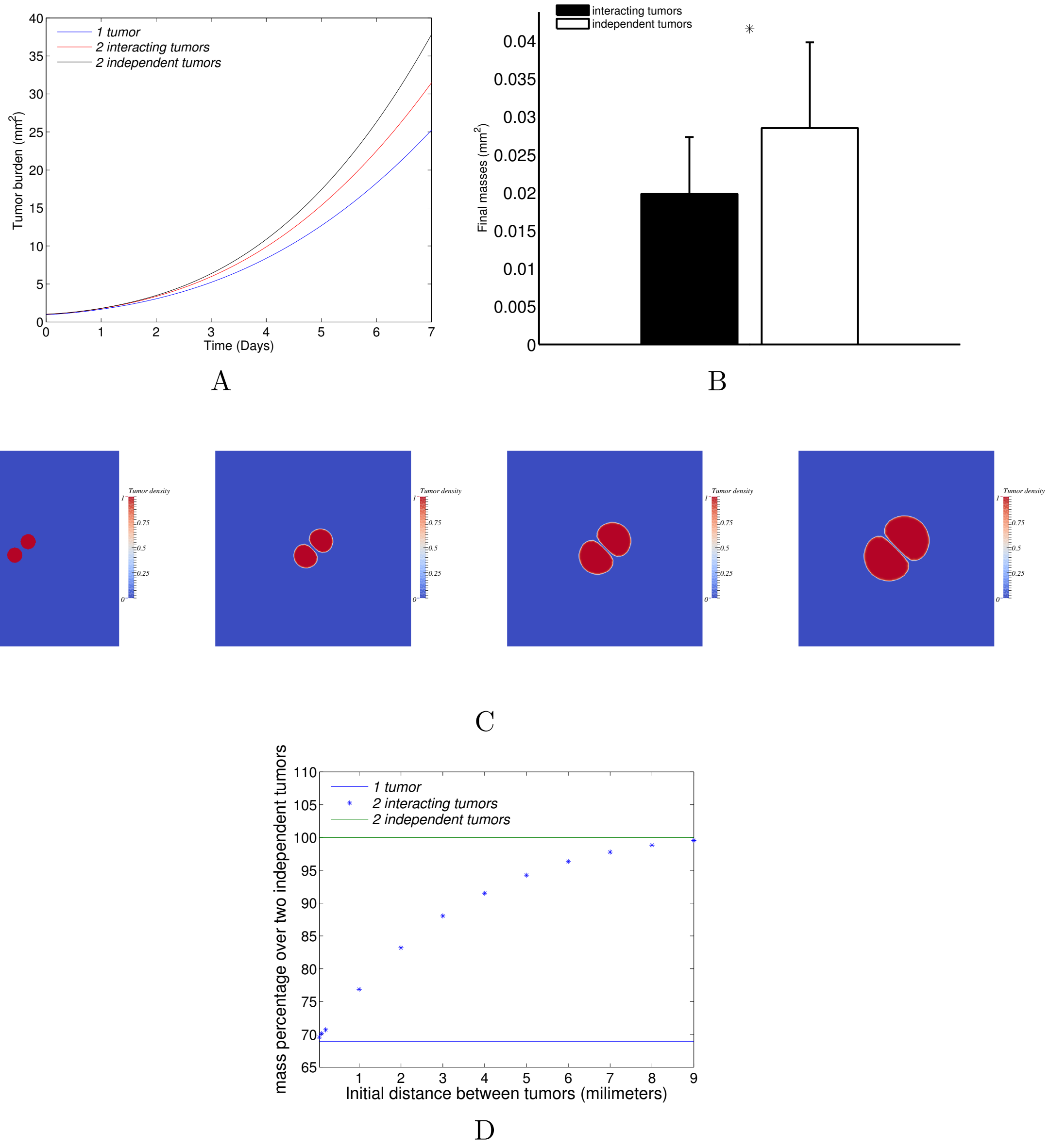
7.2.2 Quantitative impact of pressure-mediated growth interactions

We next aimed at studying the quantitative impact of the pressure that two neighbouring metastases exert on each other when they grow, and whether the merging hypothesis could explain the fast metastatic growth we observed in experiments on metastatic renal cell carcinoma (see chapter 6).

One tumor versus two tumors

Under classical non-spatial tumor growth laws (such as the Gompertz, power law, etc), two independently growing tumors lead to a larger final burden than one single tumor having double initial size. However, these models do not provide any mechanistic explanation. Our spatial model reproduced this fact while additionally providing a possible mechanistic explanation (Fig 7.9A): in a larger tumor, there are more cells that proliferate, resulting in higher mechanical constraints than in a smaller one. The pressure is therefore higher in the larger tumor, resulting in faster saturation of tumor growth over time.
Moreover, a simulation of two interacting tumors was performed by choosing a bi-focal initial condition to equations (7.1)-(7.4) and using the parameters of Meta1 ($\gamma_0 = 0.78$, $\Pi_c = 0.0026$, see a movie in the supplementary at `http://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1004626#sec026`). When comparing the total growth (sum of the surfaces of the two tumors)

to the growth of only tumor seeded with the same initial surface as the two tumors together, we observed a slower growth in the single tumor (Fig 7.9).



A

B

C

D

Figure 7.9: tumor-tumor spatial interactions. (A) Three different configurations with a same initial burden: only one tumor, two close tumors, two far tumors. The dynamics in the three configurations are compared with the parameter set inferred from the fit on one metastatic growth $(0.78, 0.0026)$ $\text{day}^{-1} \times \text{Pa}$. (B) The final burden are compared in two configurations: two close tumors and two independent tumors. The mean burdens over a set of 64 parameters (resulting from an 8×8 uniform discretization of the relevant parameter space given by the individual tumor fits, $(0.67, 1.01) \times (5.2 \cdot 10^{-4}, 2.6 \cdot 10^{-3})$ are plotted with the standard deviations (difference of $31\% \pm 1.5\%$ between the two distributions). (C) From left to right: time course of two interacting tumors growing and pushing each other. The parameters were fixed from one of the fitted MRI metastases: $\gamma_0 = 0.78$ day^{-1}; $\Pi_c = 0.0026$ Pa; time of simulation: $T = 7$ days; initial distance between the two metastases: 0.2 mm; initial surface for each metastasis: 0.46 mm^2. (D) The curve represents the time evolution of the final burden with respect to the initial distance between the two interacting tumors. The initial total burden and the parameters were taken to be the same as one of the four fitted metastases (same as C).

Two interacting tumors versus two independent tumors

To quantify the impact of the mechanical interactions, we then compared the two following situations: (1) two metastases that grow independently (the final burden consists in summing-up the two burden) and (2) two metastases that are close to each other (the exerted pressure of one metastasis impacts the growth of the other). The two configurations were studied and compared with 64 sets of parameters chosen in the parameter space established by the calibration. The results highlight the possibility for two metastases to mutually impair their growth by mechanical interactions. Indeed, by proliferating, the neighboring tumors exert pressure on each other, which leads to a decrease in proliferation in comparison to distant growing tumor foci. More precisely, under the assumption of increasing doubling time with respect to the pressure, the calibrated model revealed substantial differences in tumor burden, as shown in Fig 7.9B. Among all the parameter sets, when two tumor foci interact, at the final time ($T = 7$ days, which corresponds to the time scale of the four metastatic growths) the loss of mass was $31 \pm 1.5\%$ (mean ± standard deviation) in comparison to distant growing tumor lesions. As an example, Fig 7.9C presents a simulation of two interacting tumors at four time points.

Impact of the merging time

We observed in our data neighboring metastases growing close to each other until merging (Fig 6.5). To simulate merging of two metastases, we did not introduce any merging effect in the model. It occurred naturally in the model when the tumor density field consists of two tumor foci growing in close vicinity. Mechanical interactions occurred at the time of merging, resulting in a slow-down of tumor growth. In terms of global dynamics, two different merging times generated two different dynamics. This merging time is equivalent to the initial distance between the two metastatic foci. We therefore studied the impact of the initial distance. Under our modeling assumptions, the interactions between two metastases decrease with the initial distance between them (Fig 7.9D). This means that the later the metastases merge, the larger the final burden is. When the initial distance between the tumors went to 0, the burden corresponded to the burden of only one tumor. When the distance tended to infinity, the burden was equivalent to the burden of two independent tumors.

Combined effects of mass fractionation and interactions between tumors

We studied the effect of the fractionation of a same burden for independent metastases and for interacting metastases with a distance of 0.2 mm between metastases. Fig 7.10 depicts the evolution of the burden as a function of the number of tumors in the case of independently growing tumors or tumors that mechanically interact. As shown, the difference between both situations increases with the number of metastases. For instance, for 18 metastases growing close to each other, the loss of mass from the independent case to the interactions case is 76.3% (to be compared to the 31% for two tumors).

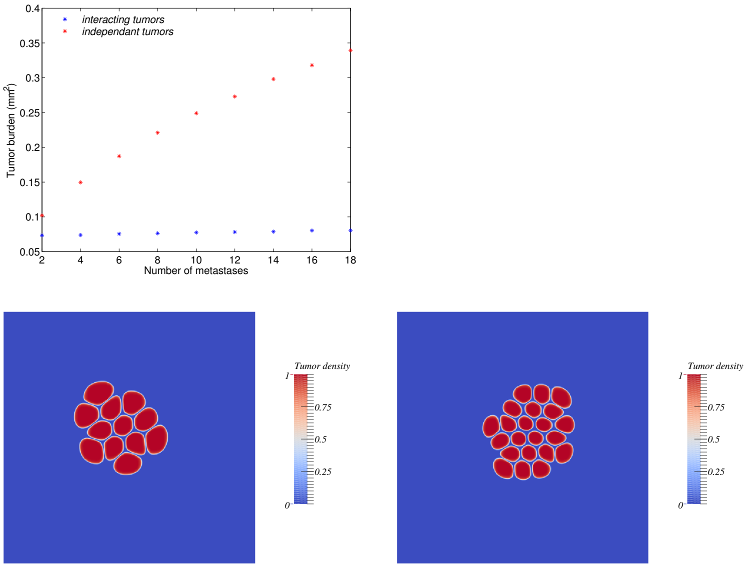


Figure 7.10: Evolution of the final burden with respect to the number of interacting metastases. Results of the simulation (Day 7) with different numbers of metastases: 4, 12 and 22. The parameters values are chosen among the sets of fitted parameters on individual metastatic growths. Simulations were obtained using Eqs (7.1)-(7.4) using the following parameter values: $\gamma_0 = 0.78$ day^{-1}; $\Pi_c = 0.0026$ Pa; time of simulation: $T = 7$ days

7.2.3 The merging hypothesis

We investigated whether the merging of metastatic foci could have generated the formation of macro-metastases visible by MRI at day 19 (see Fig 6.1) in the required time frame, with biologically realistic growth rates. We investigated the two situations: without spatial interactions (i.e., assuming the volume resulting from the merging as equal to the sum of the metastatic foci volumes), and with spatial interactions. To do so, we performed four simulations with the four fitted parameter sets, starting from one cell, to estimate the number of merging metastatic foci required to obtain the respective observed volumes (of 0.022, 0.046, 0.085 and 0.67 mm^3) seven days after initiation (day nineteen). Indeed, we chose day twelve and not day fourteen (which was the time at which the first metastatic cells were observed by direct examination of lung tissues) as the starting day because the GFP signal started to rise at day twelve (Fig 6.2). The required numbers are presented in Table 7.2, with and without spatial interactions between the tumors. For the spatial interactions case, simulations were performed as follows. Each focus was assumed to start from one cell and the foci were randomly distributed within the computational domain. To avoid metastatic foci too close from the domain boundaries and to allow the foci to merge together, the initial distance between the foci was constrained to 0.03 mm. Fig 7.11 depicts the simulations results (see a movie at `http://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1004626#sec026`). The number of required metastases reported in Table 7.2 has been estimated by dichotomy (the final burden increases with the number). Because the initial distance between the foci was small, the mechanical interactions were maximal here. The estimated number is therefore probably overestimated. Consequently, the two estimated numbers (with and without spatial interactions) give an approximate range for the exact required number.

meta index	required number of metastases	
	Without spatial interactions	With spatial interactions
meta 1	1337	2127
meta 2	20	65
meta 3	301	375
meta 4	40	70

Table 7.2: There it is the number of required merging foci to obtain the metastatic sizes measured on the MR images for each followed metastasis. Two cases are considered: with and without spatial interactions.

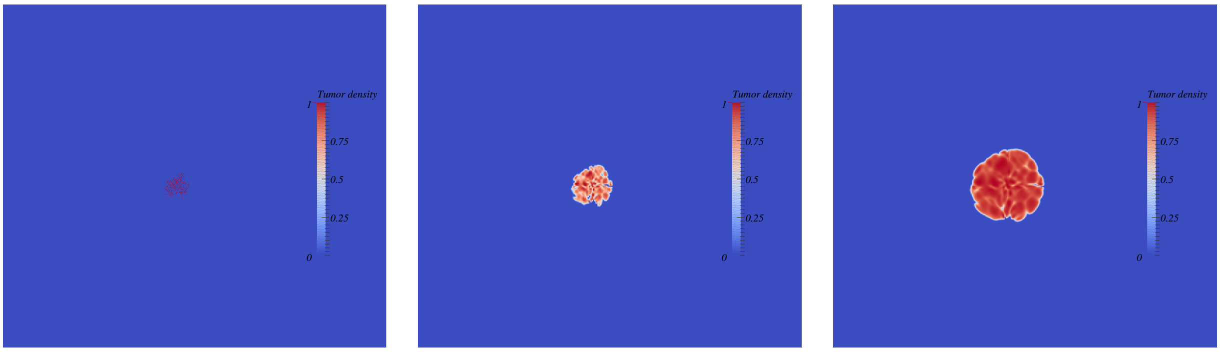


Figure 7.11: Simulation of multiple metastatic foci merging (with spatial interactions). From left to right: time course of merging metastatic germs. Each germ starts from one cell. The germs are randomly located at a distance of 0.03 mm from each other. Simulations were obtained using Eqs (7.1)-(7.4) and the following parameter values: $\gamma_0 = 0.78$ day^{-1}; $\Pi_c = 0.0026$ Pa; time of simulation: $T = 7$ days; number of germs = 200 in 2D. The corresponding number of cells in 3D is computed under a spherical symmetry assumption and is 2127.

As can be seen in Table 7.2, since spatial interactions reduce the growth velocity, the number of metastases was higher when interactions were taken into account. Because of potential variability (error measurements during the segmentation, differences between the MRI signal and the real lesion, especially for the small metastases, modeling assumptions), the estimated numbers of required metastatic foci may give only a rough estimate. For two of the metastases (termed Meta 2 and Meta 4 in Table 7.2), the estimated numbers appear to be reasonable. On the other hand, for the two other ones, the required number ranged respectively between301 and 375 and between 1300 and 2100, which are probably too large to be biologically realistic.

Besides spatial interactions, another possible phenomenon involved could be the attraction and aggregation of circulating tumor cells. This hypothesis has been discussed in chapter 6.

7.3 Discussion and perspectives

7.3.1 Discussion

In this study, we modeled tumor merging and mechanical interactions between tumors growing in close vicinity. It first allowed us to quantify the impact of such interactions on the global growth. Indeed, when they are in contact, tumors exert pressure on each other in growing, hence impairing their mutual growth. Second, we used this model to assess the required number of merging metastases to form foci as large as those observed by MRI in 6.

To investigate whether tumor merging would have quantitatively non-negligible impact on the kinetics of the total metastatic burden (thus requiring more intricate modeling for the model describing the size distribution at the scale of the organ), we introduced a parsimonious spatial model of tumor growth.

The spatial model for tumor growth that we introduced is based on a pressure-induced decrease of the growth rate. Contact inhibition between cells is a mechanism for maintaining tissue homeostasis [4]. The ability of cancer cells to ignore these inhibition signals is a hallmark of cancer. In a recent study, Stylianopoulos et al. showed that the uncontrolled proliferation of tumor cells results in mechanical stresses in the surrounding microenvironment of transplanted and human tumors [15]. Furthermore, they also showed that such an exerted pressure impairs in vivo proliferation via two mechanisms: reduced cancer cell proliferation in direct response to increased pressure, as well as a pressure-induced collapse of blood vessels within the tumor, leading to nutrient deficiency for tumor cells [16]. Based on these considerations, it seems relevant to consider that tumor expansion depends on the pressure. In our spatial growth model, the tissues motion is mediated by pressure gradients. It means that cells within a tumor tissue proliferate and that the exerted pressure pushes the neighbouring tissues. This pressure is not solely due to mechanical constraints (solid stresses, interstitial fluid pressure,. . .) exerted by the neighbouring cells on each other, but represents a more phenomenological pressure, that reflects the basic assumption of our modeling strategy for the tumor tissue being constituted by a fluid mixture in a porous medium. The effect of the pressure on proliferation has also been studied using numerical simulations. In [125], Montel et al. discussed the fact that cells proliferate faster on the surface than in the bulk of a tumor spheroid. A classical reason is that nutrients do not penetrate deeply in the spheroid. However, Montel et al. suggested a mechanical effect due to the necessity for a cell to deform its environment in order to proliferate. In a *in silico* study on two-dimensional monolayers and three-dimensional spheroids, based on experimentally determined biophysical parameters, Drasdo and Höhme suggested that pressure conditions have a higher impact on doubling time than lack of nutrients [126]. Moreover, in [125], Montel et al. performed experiments where tumor cells were submitted to different pressure constraints and observed a decrease in proliferation when pressure was applied. In their study, simulation results that were compared to experimental ones showed an exponential decreasing of proliferation with pressure, consistently with the modeling adopted here. However, the bulk and surface division rate were not affected equally by stresses. In our model, we used a similar pressure-mediated proliferation law translating direct effects of mechanical stresses on proliferation as well as indirect effects of proliferation on the micro environment (collapsing of blood vessels leading to lack of nutrients).

Interestingly, our pressure-mediated growth model gives a mechanical explanation for the fact that small tumors grow faster than large tumors. Indeed, most of the classical ODE models (Gompertz, logistic, etc) of tumor growth provide a phenomenological way to describe that the doubling time is increasing when the tumor volume is increasing, without any explanation on the phenomenon underlying it. In the spatial model we introduced, as the accumulated pressure results from the contribution of all proliferating cells, the pressure exerted in a small tumor is lower than in a large tumor, resulting in a lower inhibition of the growth in the small tumor and then in a shorter doubling time of the small tumor.

After calibrating this new spatial model to the growth of single metastases, we found (in simulations) that spatial interactions between tumors result in a significant reduction of tumor growth ($31 \pm 1.5\%$ of mass reduction) in comparison to distant growing tumor lesions. Our results indicate that spatial interactions may have a large impact on growth and should be considered in future efforts for the development of a general quantitative theory of metastatic colonization. Our proposed hypotheses should be further experimentally reinforced, by, for example, *in vivo* investigations by observing two (or more) growing tumours in close vicinity that would enter mechanical interactions and then assess with a Ki-67 staining if the proliferation is impaired in the contact area, would further reinforce our contentions.

Together, our results have implications for theories of the metastatic process and suggest that global dynamics of metastasis development is dependent on spatial interactions between metastatic lesions. However, it is unlikely that they alone control metastasis expansion. Indeed, when trying to assess whether this concept alone explains the fast growth of various metastases from the beginning of organ colonization (from the first cell at days $12-14$ to $(0.022-0.67)$ mm^3 at day 19, according to the delay observed on the data in Fig 6.1), unrealistic numbers were found for two of the tumors. The inability of the merging theory based on our spatial modeling approach to explain all of the observed volumes reinforces the results of the previous chapter suggesting that merging alone could not explain the observed metastatic size distributions. In the present study, we tried this time to obtain single metastatic volumes observed at MRI with a spatial model describing the spatial expansion of merging metastases and mechanical interactions occuring during this phenomenon. Thus, other mechanisms are probably also involved such as recruitment of additional cells from the blood stream, an hypothesis discussed in the previous chapter, and microenvironmental cues such as nutrient depletion or responses to environmental stress.

7.3.2 Future modeling perspectives

A first perspective is to extend the pressure model in 3D. Indeed, we can imagine that the evacuation of the pressure is easier in 3D and that the behavior of the modele could be different than the 2D one.

Moreover, a drawback of this model is the impact of the size of the computational domain. Indeed, the smaller the domain is, the faster the pressure decreases from the center to the boundaries of the domain. Anyway, if the domain represents the organ the tumor is located in, square boundaries are not realistic. An idea could be to consider the real organ shape (the lung for example) for the boundaries and moving boundaries (following the velocity field at the boundaries), keeping Dirichlet conditions on the pressure. In this case, the pressure at the boundaries would be still the homeostatic pressure of the organ, but deformations of the organ caused by the tumor growth would be taken into account. Of course, such a model would be still unrealistic because organs are probably elastic and tumor growth probably does not cause direct deformation, but rather compress the organ. it comes from our saturation hypothesis, which is probably a little simplistic. Moreover, the lungs are principally composed of alveoli full of air. When a tumor is growing inside, these alveoli are compressed, inducing no particular external deformation of the organ. Furthermore, in the lungs as in other organs, the vision of a tumor only pushing the surrounding tissue is probably very simplistic also, as invasive tumor cells and cells of the microenvironment at the tumor interface secrete proteases to degrade extracellular matrix like basal membranes. Therefore tumor expansion is associated with tissue deformation but also tissue degradation.

Finally, another modeling clue, overlapping perspectives highlighted in the previous chapter, could be to include spatial interactions between metastases in a model describing systemic metastatic development as the model used in the previous chapter.

Chapter 8

Quantitative modeling of the premetastatic niche

In preparation

In this chapter, we focus on the metastatic microenvironment, and more particularly a phenomenon called premetastatic niche. It consists in a precolonization of the distant organ by bone marrow derived cells (BMDCs) before the arrival of metastatic cells in the site. Experimental evidences showed that these BMDCs seem to prepare the organ to home metastatic cells, making this *a priori* inhospitable organ more permissive to support metastatic cells attraction, survival and proliferation. We propose here a mechanistic model of the premetastatic niche formation. This model has been coupled with Iwata's model of metastatic growth (6.2) in order to study the impact of the microenvironment on the metastatic output. The complete model has been fitted on the data of Kaplan et al. [10] and showed a good ability to reproduce the observed dynamics. The future perspectives in terms of clinically relevant quantitative questions are discussed further.

8.1 Elements of the metastatic microenvironment

8.1.1 Cells of the tumor and metastatic microenvironment biology

A tumor is not a mass only constituted of overproliferative mutant cells. It contains multiple kinds of cells, which are in part not malignant [44, 29].

Cancer stem cells (CSCs)

Cancer stem cells are defined as cells that are able to generate a new tumor [4]. This definition can be complemented by expression of markers of stem cells of the tissue of origin. Indeed, CSCs often share transcriptional profiles similar to classical tissues stem cells [150]. It should be noted that the definition of CSCs differs from definition of normal stem cells in that it is stronger than just self-renewal capability (which, combined to pluripotency, forms the definition of a normal stem cell) [4, 151]. Note also that self-renewal capability (= infinite replicative potential) is a cancer hallmark supposedly shared by all the cancer cells [4]. CSCs were first discovered in the late 1990's early 2000 in hematopoietic malignancies. Later on for solid tumors (for breast for instance in 2003).

Their origin during cancer progression is still unclear. They could come from normal tissue stem cells or from progenitor cells that are already more differentiated and would de-differentiate [4]. It could also happen that there is heterogeneity in "stemness" within the tumor.

Recent research has shown connections between epithelial to mesenchymal transition (EMT) and stem cells. Cells that underwent EMT were found to exhibit classical stem cells features, such as the self-renewal capability that could thus help then in seeding a new tumor at a distant site [4]. EMT being triggered by heterotypic signals (from inflammatory cells for instance), it raises the hypotheses that these stromal cells play a role in the presence of CSCs [4].

Indeed, it has been shown in a recent study that cancer cells are necessary for metastasis formation in the MMTV-PyMT mouse breast cancer model and that colonization of the distant organ by CSCs is mediated by POSTN (an extracellular matrix component) expression in the distant organ [150]. This POSTN expression is mediated by fibroblasts stimulated by tumor factors like TGF-β.

It is still open to determine whether CSCs are really a (rare) subpopulation of cancer cells. In particular, evidence of this fact is made difficult by the possibility of bidirectional transitions between CSCs and non-CSCs [4]. Malanchi et al. observed that CSCs represent $3 \pm 2.1\%$ of all tumor cells in the primary tumor and the metastases [150]. However, this proportion increases until 20% during the metastatic colonization phase.

CSCs have two important clinical implications: 1) they were shown to be more resistant to chemotherapy and radiation, thus possibly yielding an explanation of the inevitable recurrence of the disease after such treatments and 2) they could help to understand long dormancy periods followed by regeneration of a tumor [4].

CSCs could have a role in the development of the tumor-associated stroma since they were shown to be able to differentiate into supportive cells such as fibroblasts or endothelial cells, thus generating a stroma based only on cancerous material (without recruitment of normal host cells) [4].

CSCs have same genome as other cancer cells, but different phenotypic behavior and degree of differentiation (plasticity) [4].

The cellular helpers

During tumor growth, the supportive tissue or stroma is necessary for the tumor development [44, 29]. Different kinds of cells constitute this stroma. We present here the cellular helpers of tumor and metastatic development. For each cell type, we present first its physiological function, that is in normal conditions, and then its pathological function in cancer.

Endothelial cells

Endothelial cells constitute the epithelial wall of the blood vessels, as illustrated in Fig 8.1. During the "angiogenic switch", quiescent endothelial cells are activated and enter into the vasculogenesis program. It is interesting to notice that tumor-associated endothelial cells express different genes in comparison with normal endothelial cells [4]. It could be valuable to exploit these differences to develop novel therapy strategies to target the tumor vasculature.
The role of the lymphatic endothelial cells that constitute lymphatic vessels is poorly understood. Notably, intratumoral lymphatic vessels are often collapsed, but an important lymphangiogenesis takes place at the periphery of the tumor stroma [4]. These vessels are invested by cancer cells when the tumor becomes invasive, resulting in roads for cancer cells to metastasize in lymph nodes.

Pericytes
Pericytes are mesenchymal cells that wrap around certain types of blood vessels (capillaries, veins, etc...), as illustrated in Fig 8.1. They are analogous to smooth muscle cells [4], which are present around other types of blood vessels (arteria, etc...). As illustrated in Fig 8.2, pericytes communicate with endothelial cells and notably secrete factors like angiopoietin-1 that are received by the Tie2 receptors expressed by endothelial cells, resulting in stabilizating antiproliferative signals [4]. To ensure vascular homeostasis, some pericytes can secrete low levels of VEGF [4]. Pericytes also have an important role in basement membrane synthetizing. This membrane allows the endothelial cells and pericytes anchorage and vessels to support the hydrostatic pressure flow [4].

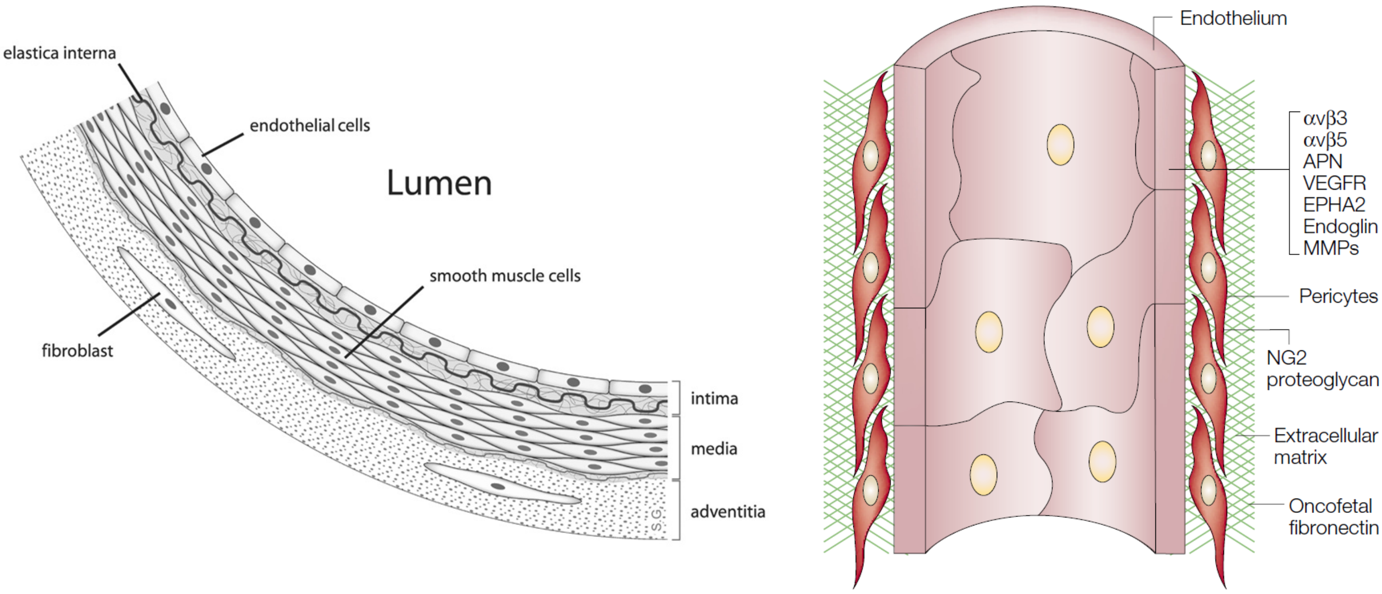


Figure 8.1: Two pictures of blood vessel at the cell scale. Cells constituting the endothelium and cells supporting the vessel are represented. Left: picture retireved from [152]. Right: picture retireved from [153]

Because their important functions in vasculature stability, pericytes may constitute an interesting therapeutic target. Targeting Platelet derived growth factor (PDGF) receptor that is expressed by tumor pericytes and bone marrow derived pericytes progenitors results in destabilized and unfunctional tumor vasculature [30]. However, such destabilized vessels would be more permeable and could favor metastatic dissemination [44, 30].

Immune cells
One can divide the immune system into two parts: 1) the innate immune system, associated with wound healing and clearing dead cells and cellular debris as well as support adaptive immune response and 2) the adaptive immune system that detects infectious agents and eliminate it.
The innate immune system includes granulocytes (neutrophils, eosinophils, basophils), macrophages,

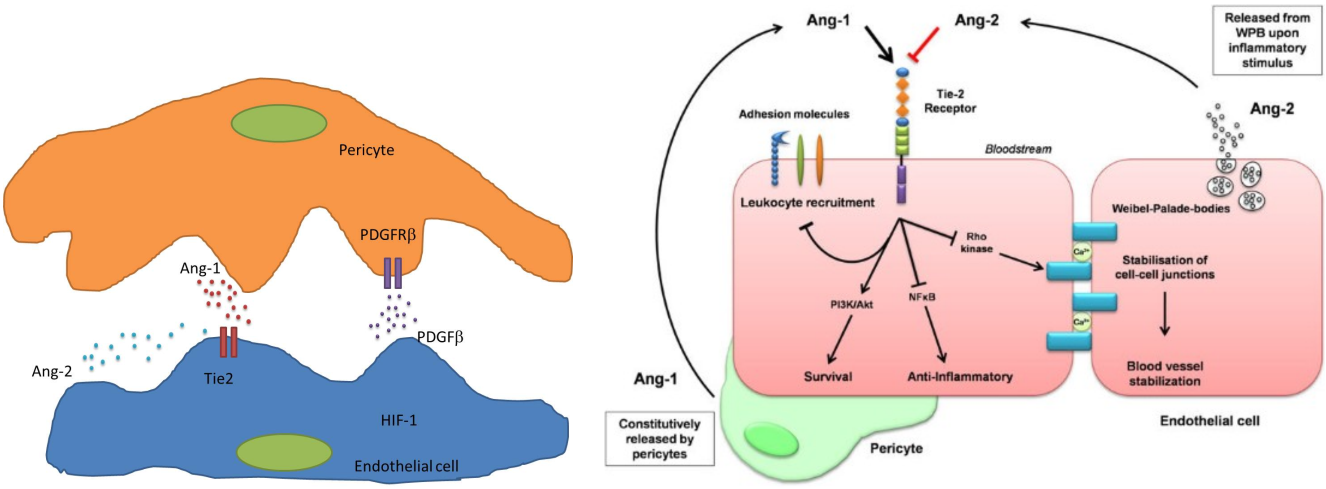


Figure 8.2: Two pictures of the crosstalk between endothial cells and pericytes. Some of the main pathways of this crosstalk (PDGF, Ang-1, Ang-2) are represented. Left: picture retrieved from [154]. Right: picture retrieved from [155]

dentritic cells and NK cells. The adaptive immune system is completely constituted by the lymphoid cell line comprising B cells, memory B cells and T cells (cytotoxic, helper, regulatory, etc...).
We can distinguish two types of inflammation: 1) immune inflammation fighting infections, which is transient and 2) pathological inflammation (fibrosis, aberrant angiogenesis, neoplasia), which is chronic [4].
There are two types of immune cells that are largely observed in all neoplasia :

- Tumor-antagonizing cells like cytotoxic T lymphocytes or NK cells
- Cells that functionally enhance hallmarks capabilities by chronic inflammation [44, 29, 4, 30]. As mentioned in [4, 156], tumors can be seen as "wounds that never heal".

The tumor-promoting immune cells are typically macrophages subtypes, mast cells, neutrophils, T (not all subtypes) and B lymphocytes and CD11b myeloid derived suppressor cells [44, 29, 4, 30]. They secrete many different factors like tumor growth factors (e.g EGF) [4, 44], proangiogenic growth factors (e.g VEGF, FGF2) [30, 4, 29], chemokines and cytokines that amplify the inflammatory state [44, 29], proangiogenic and proinvasive MMPs [29, 44, 30]. They have multiple roles in tumor progression, like cancer cells proliferation, angiogenesis stimulation, tissue invasion, metastatic dissemination and colonization [44, 29, 4, 30].
To explain the existence of two paradoxical roles of the immune system, one can notice that physiologically, the immune system has two distinct roles that are:

- Detecting and targeting infectious agents
- Wound healing and clearing dead cells and cellular debris (activated macrophages, neutrophils, and myeloid progenitors)

The second class of immune cells is thought to support neoplasia progression. This progression is controlled by the balance between presence of conventional or activated immune cells.

Fibroblasts
Fibroblasts are the most common cells of the conjonctive tissue. Fibroblasts are cells from the mesenchymal cell line, as shown in Fig 8.3 They can principally secrete and remodel the extracellular matrix (collagens, glycosaminoglycans, reticular and elastic fibers, glycoproteins) and play a role in

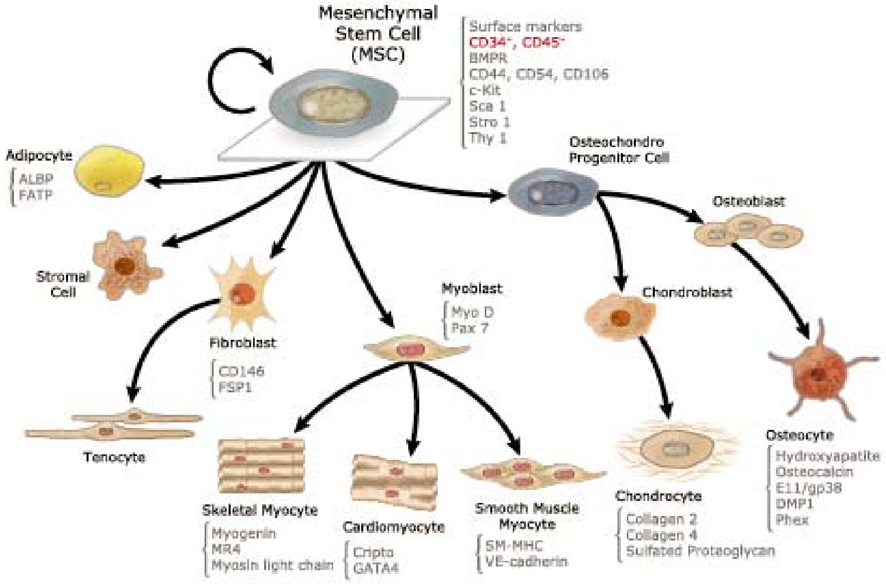
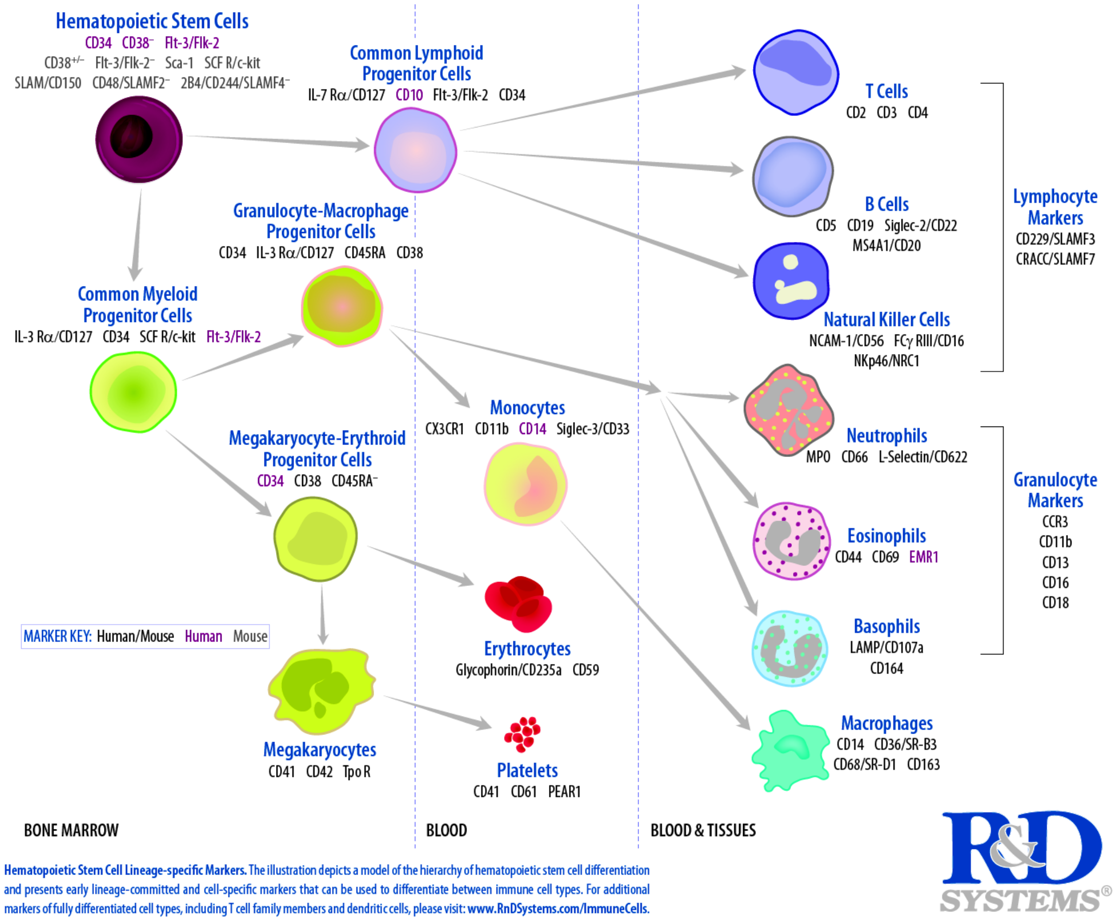


Figure 8.3: The mesenchymal cell line, picture retrieved from [157]

wound healing. They are motile and can become epithelial cells by the mesenchymal to epithelial transition. Conversely, epithelial cells can become fibroblasts by epithelial to the mesenchymal transition.

Cancer associated fibroblasts (CAFs) are the most common cell type in tumor stroma. It embraces two subtypes of cells, that are 1) tissue-derived fibroblasts and 2) myofibroblast, a differentiated state between fibroblast and smooth muscle cells (liver, pancreas, etc.).
Recruited myofibroblasts and reprogrammed variants of normal tissue-derived fibroblastic cells have been showed to enhance tumor phenotypes such as cancer cells and progenitors migration and motility by SDF-1/CXCR4 [30], cancer cells proliferation [4], angiogenesis [44], invasion (EMT by HGF, FGF, TGF-β, MMPs) [30] and metastasis (premetastatic niche, [10]).
Cancer associated fibroblasts are immortalized. They are more competent than healthy fibroblasts to support tumor growth. Furthermore, they can induce tumorigenesis in immunocompromised mice ([158]). Their matrix remodeling capacities make them responsible of demoplastic (desmoplasia is the growth of fibrous or connective tissue) stroma that characterizes many advanced carcinomas.

Stem and progenitor cells of the tumor stroma
Progenitor and stem cells are immature cells that have two principal functions: 1) differentitation into mature and functional cells, for example immune cells and 2) self-renewal to maintain a constant number of cells in the body. All the cells of the hematopoietic cell line comprising immune cells, red blood cells and platelets, originate from hematopoietic stem cells, as illustrated in Fig 8.4. A high number of stem and progenitor cells come from the bone marrow.
Most of the tumor stromal cells are recruited from the adjacent tissue but some of them are recruited from the marrow. These stem and progenitors cells exhibit functions that their more differentiated progeny don't have. Bone marrow derived progenitor cells have many cell lines types: 1) hematopoietic progenitor cells having roles in angiogenesis and premetastatic niche ([10, 30]) 2) progenitors of pericytes and of various subtypes of cancer-associated fibroblasts originated from the bone marrow 3) endothelial progenitor cells that play a role in angiogenesis. Tumor-associated stromal cells can originate from bone marrow or from differentiating stem and progenitor cells originating in the neighboring normal tissue.



Hematopoietic Stem Cell Lineage-specific Markers. The illustration depicts a model of the hierarchy of hematopoietic stem cell differentiation and presents early lineage-committed and cell-specific markers that can be used to differentiate between immune cell types. For additional markers of fully differentiated cell types, including T cell family members and dendritic cells, please visit: **www.RnDSystems.com/ImmuneCells.**

Figure 8.4: The hematopoietic cell line. Picture retrieved from [159]

8.1.2 The premetastatic and metastatic niches

As we mentioned previously, crosstalking between tumor cells and cells of the microenvironment is a critical step for tumor progression. Such cells can also be found in the metastatic microenvironment and are likely to play a critical role for metastatic progression too. The following review deals with the premetastatic and metastatic microenvironments inducing premetastatic and a metastatic niches to help survival and proliferation of metastatic cells. Moreover, the premetastatic niche, in other words the precolonization of the future metastatic site by non tumor cells to induce a viable environment for metastatic development, could explain the organ-specific nature of metastatic disease.

The organ-specific nature of metastatic dissemination

Clinically, it has been observed that metastases are more frequent in particular organs than in others and that such preferential distant sites depend on the type of cancer [47]. For example, breast cancer often metastasizes to the liver, lungs, bones and brain, whereas colorectal cancer preferentially metastasizes to the liver. In 1889, Steven Paget investigated the location of metastasis on 900 autopsies of patients with breast cancer and shed light on the non randomness of metastasis location [160]. Indeed, much more cases of liver metastases than spleen metastases were reported, although the arteries of the liver and the spleen have comparable sizes, allowing equal chances of metastatic seeds in theory. Moreover, he observed that bone metastases were found in particular bones and not in others, a tendency that could not be explained by a ‘theory of embolism alone”, according to Paget’s words. These observations led Paget to make the hypothesis that formation of metastases does not depend on the seed (tumor cell) only but rather on the compatibility between the seed and the soil (the secondary tissue). This “seed and soil” hypothesis has been challenged in 1929 by James Ewing, which proposed that metastatic dissemination depends on circulatory patterns between the primary tumor and secondary organs [161]. In breast cancer for instance, cancer cells detaching from the primary and going through the blood circulation are first trapped into the lung capillaries, whereas in colorectal cancer, they are first trapped in the liver capillaries [47]. As mentioned by Fidler in [21], Coman et al. performed experiments in 1951, showing evidences that blood flow mechanics govern the metastasis location [162]. In [47], Chambers and Groom mention a study of Leonard Weiss that observed in patients with 16 tumor types cancers that metastasis location was well explained by circulatory patterns in only 66% of the primary-secondary organs pairs. This would support that organ-specific nature of metastasis depends on both mechanical factors and seed and soil compatibility. Moreover, Chambers and Groom mentioned that seeding of tumor cells in an organ mainly depends on mechanical factors, whereas the growth of the seeds could depend on the compatibility between the seed and the soil, an that this compatibility depends on the cancer cell line [47]. In his revisited “seed and soil” hypothesis, Isaiah J. Fidler defends Paget’s hypothesis and emphasizes that such a tropism of the metastatic cells for particular organs is strongly mediated by the cells of the microenvironment like fibroblasts, endothelial cells and infiltrating leukocytes [21].

Metastasis development: an inefficient process

As mentioned in [47], metastasis development is a very inefficient process. Indeed, during all the steps of the metastatic process, tumor cells may encounter many obstacles: adhesion signals loss, blood flow pressure, immune surveillance, or lack of survival signals [5]. In [21], it is mentioned that about 0.1% of the circulating tumor cells survive within 24 hours after entry into the circulation and reach the distant organ and that only 0.01% of them manage to form a new metastasis. Experiments using *in vivo* video microscopy and intravenous injections of tumor cell lines have shown the inefficiency of the metastatic process in the lungs, the liver and the brain [163, 164, 165]. Groom and Chambers mentioned also that first steps of the metastatic process are relatively efficient and that inefficiency is mainly due to the difficulties for cancer cells to grow at the secondary sites [47]. As mentioned before, some tumor cell lines are adapted to seed in particular organ whereas other cell lines are adapted to seed in other organs. In 2005, Kaplan, Lyden et al. have shown experimental evidences of the formation of a premetastatic niche before the arrival of the metastatic cells [10]. Such a premetastatic niche can make a secondary organ more permissive for arrival, survival and proliferation of metastatic cells. As mentioned by Chambers and Groom, the success of metastasis development could be imputed both to the intrinsic metastatic abilities of the tumor cell line (seed) and how permissive the secondary organ (soil) is.

The premetastatic niche: first experimental evidences

In 2005, Kaplan et al. showed first evidences of a premetastatic niche on two mouse cell lines (Lewis lung carcinoma and B16 melanoma) [10]. The goal of this paragraph is to summarize their work.
In this study, the authors first presumed that bone marrow derived cells (BMDCs) play a role in metastasis formation. Two main types of BMDCs are considered: cells expressing VEGFR1 (vascular endothelial growth factor receptor 1), regrouping essentially hematopoietic progenitor cells (HPCs), and cells expressing VEGFR2, regrouping endothelial progenitor cells (EPCs). HPCs and EPCs were already known to contribute to tumor angiogenesis [10]. HPCs are progenitor cells, a transitional state of differentiation between stem cells and fully differentiated cells, which can differentiate into many blood cells like erythrocytes, macrophages, or granulocytes [159]. The goal of the study was first to highlight the formation a premetastatic niche by BMDCs, and then to shed light on the particular BMDCs that are implied in the process. Finally, other experiments allowed to determine what modelucar and genetic pathways are implied during the premetastatic niche formation.

First experiment: subcutaneous injection of LLC and B16 tumor cells and GFP/β-gal tracking of BMDCs

In this experiment, β-galactosidase-positive (β-gal+) or GFP+ BMDCs were transplanted in mice. It was not known *a priori* where these cells would go (staying in the blood, going back to the bone marrow, etc...). Four weeks after, a subcutaneous implantation of DsRed-tagged LLC or B16 tumor cells performed. These two cell lines are known to have different metasatic potential: LLC tumors metastasize into the lungs and sometimes the liver, whereas B16 disseminate into the lungs, the liver, the testis, the spleen and the kidney. The experimental protocol is summarized in Fig 8.5.

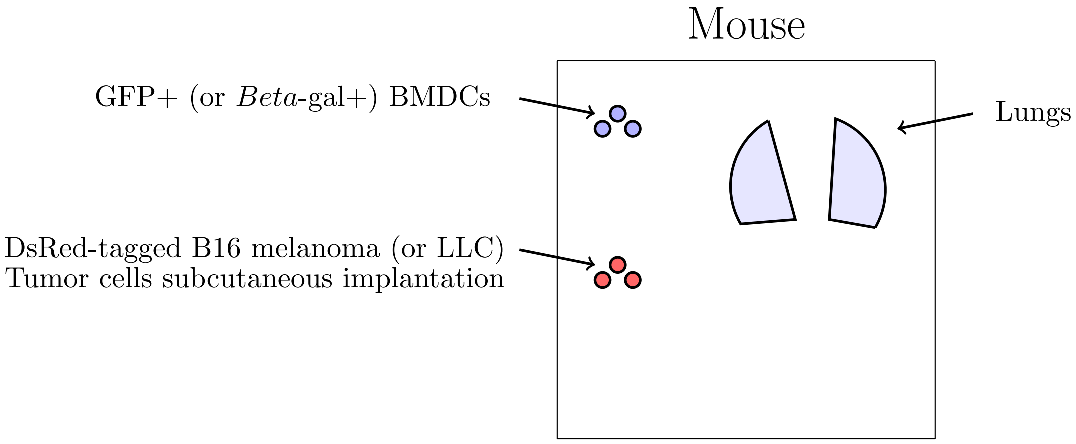


Figure 8.5: First experiment of [10]: transplantation of β-gal+ (GFP+) BMDCs followed by subcutaneous transplantation of B16 (LLC) tumor cells

At each time point, observations of lung slices were made. Before tumor implantation, the lungs do not exhibit any anomaly. Fourteen days after tumor implantation, no tumor cells were observed yet into the lungs. However, clusters of β-gal+ (GFP+) BMDCs were observed into the lungs, forming premetastatic lesions. At day 18 after tumor implantation, the first DsRed-tagged tumor cells were observed into the lungs. It was observed that the location of 95% of tumor cells coincide with the location of BMDCs clusters. The findings are summarized by the timeline in Fig 8.6

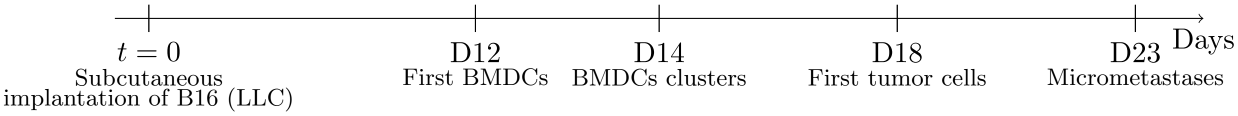


Figure 8.6: Timeline of the first experiment of [10]

Moreover, flow cytometry showed that BMDCs clusters were only formed in the future metastatic distant site depending on the tumor type: lungs and liver for LLC, lungs, liver, testis, spleen and kidney for B16.

These experiments highlighted the formation of premetastatic BMDCs clusters before the arrival of first tumor cells in the future metastatic sites.

Second experiment: injection of tumor conditioned media followed by intraveinous injection of tumor cells

The goal of this second experiment was to show that cytokines released by the primary tumor were responsible for the mobilization of BMDCs to the premetastatic site. In order to prove it, melanoma cells were maintained in culture during 18 hours and the serum called melanoma conditioned media (MCM) was extracted. Mice transplanted with β-gal+ (GFP+) BMDCs were considered with MCM four weeks after the transplantation and control mice were given serum-free media. MCM and serum free media were then injected every day and B16 (or LLC) tumor cells were injected intravenously seven days after the first MCM injection. The schemes of Fig 8.7 summarize the experimental protocol.

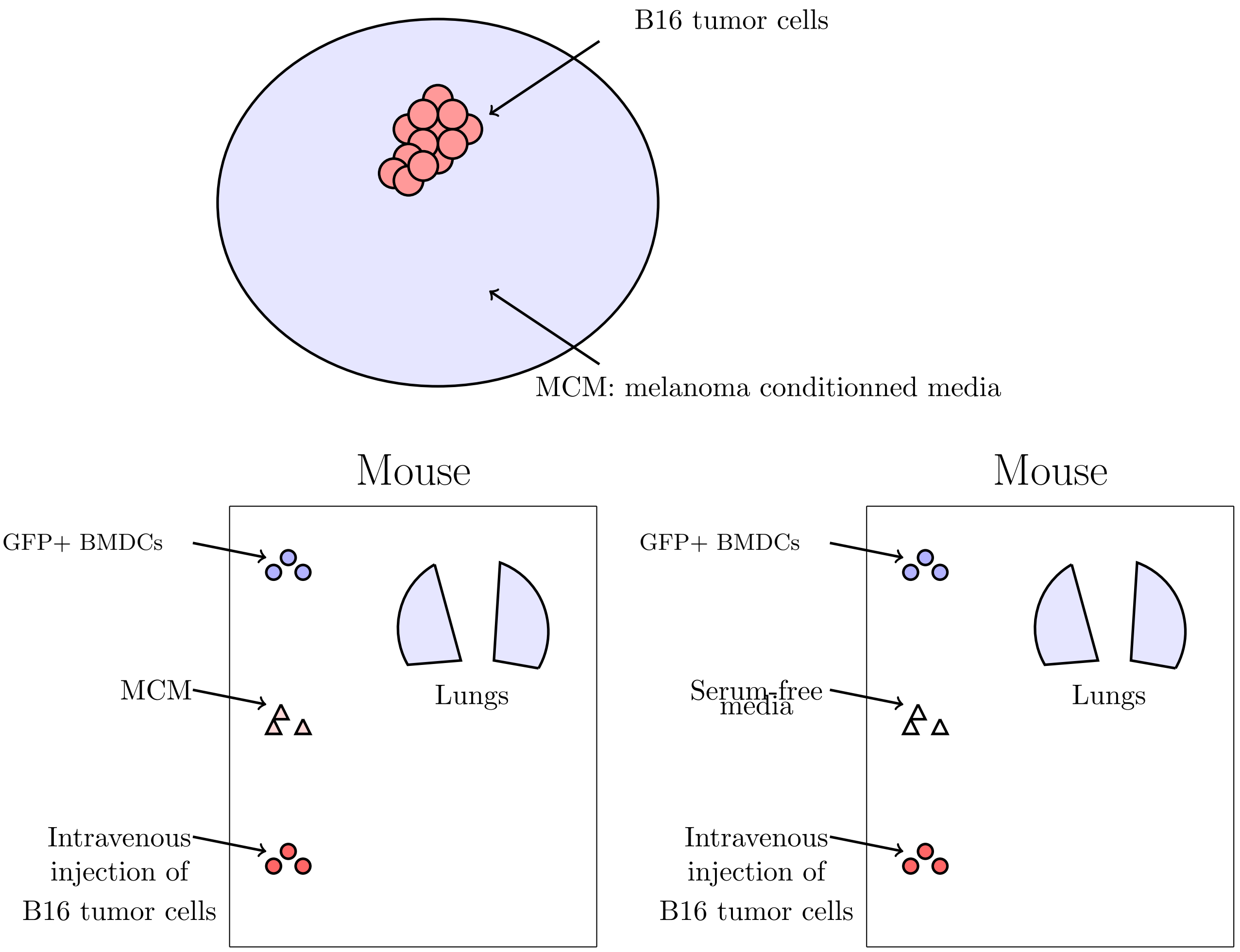


Figure 8.7: Second experiment of [10]: Melanoma conditioned media injection followed by intravenous injection of B16 (LLC) tumor cells

Few days after media injection, whereas no increase of BMDCs is observed into the lung for control mice, a significant increase of BMDCs is observed into the lungs and other potential metastatic sites. Four days after tumor cells injection, micrometastases located on the BMDCs clusters are observed in MCM mice, but not or very few tumor cells are observed in control mice. The scheme in Fig 8.8 summarizes the timeline of the protocol.

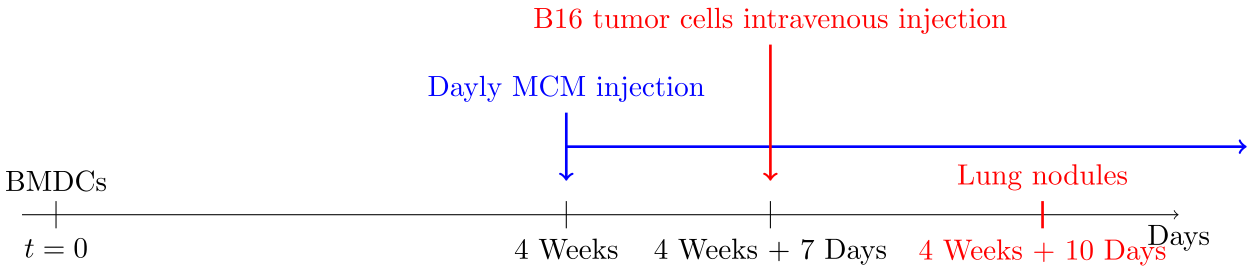


Figure 8.8: Timeline of the second experiment of [10]

This experiment showed that cytokines released by the tumor cells induced the mobilization of the BMDCs to form the premetasatic niche. Moreover, the experimental results shown these BMDCs were likely to support survival and growth of tumor in the niche.

Third step: identify the cells of the premetastatic niche
The samples of the first experiment (subcutaneous injection of tumor cells) were conserved with OCT (at a very low temperature) to be further analyzed. Clusters of differentiation (CD) allowed to identify more or less precisely the type or subset of cells, as shown in Fig 8.4. The analyses first revealed that VEGFR1+ cells appeared into the secondary sites (lungs, liver, etc...) at day 14, before the VEGFR2+ cells. Moreover, the expression of CD34 and C-kit (CD117) markers revealed the presence of myeloid progenitors (not much differentiated cells) among the VEGFR1+ cells. The marker CD11b was also present, showing that some more differentiated cells (moconytes, granulocytes) were also present. It was therefore difficult to identify precisely which cells establish the premetastatic niche but the analysis results seemed show that these cells are of the myeloid lineage with a relatively heterogeneous degree of differentiation. VEGFR2+ cells seemed to arrive shortly before the tumor cells. They are presumed to complete the HPC clusters to further support the growth of metastases, and especially to play a role during the angiogenic switch.
Fig 8.9 presents the updated timeline of the successive events of the premetastatic phase.

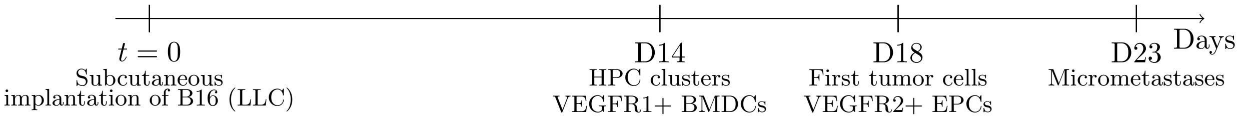


Figure 8.9: Timeline of the premetastatic phase

These experiments have shown that VEGFR1+ BMDCs form the premetastatic niche before the arrival of tumor cells for these two animal models (subcutaneous injection of LLC and B16 tumor cells). In [10], Kaplan et al also confirmed the formation of VEGFR1+ cells clusters during the premetastatic phase in a spontaneous model (c-Myc transgenic mouse) of tumor. Moreover, VEGFR1+ premetastatic lesions have also been detected on patients in organs that are often targeted by metastatic dissemination.

Targeting VEGFR1 expression or VEGFR1+ cells prevent the formation of a premetastatic niche
In an other experiment of [10], three sets of mice with purified bone marrow depleted from VEGFR1+ cells, purified bone marrow with VEGFR1+ cells only, or natural bone marrow (control) were employed.
In the case where VEGFR1+ cells were not present, no BMDC clusters and no metastases were observed. When only VEGFR1+ were present, BMDC clusters were observed, as well as micrometastases but with aberrant vasculature, whereas in the case of natural bone marrow (control), BMDC clusters and further macrometastases were observed.
Other experiments have shown that VEGFR1 antibody treatment prevents the formation of the premetasatic niche and metastases. VEGFR2 antibody treatment does not prevent the formation of a premetastatic niche and leads to micrometastases with aberrant vasculature.

Interactions between HPCs and the premetastatic microenvironment mediate the formation of the premetasatic niche
Further analyses showed an overproduction of fibronectin by the fibroblasts of the premetastatic niche before the arrival of HPCs. VEGFR1+ HPCs seem to express VLA-4 (an integrin), which is a fibronectin ligand. Binding of VLA-4 to fibronectin allows the adhesion of HPCs to the premetastatic niche. This binding leads to matrix matelloproteinase-9 (MMP-9) production by HPCs. Indeed, high expression of MMP-9 was observed in premetastatic clusters, which can degrade basement membranes, thus alterating the microenvironment by releasing Kit-ligand (or stem cell factor) and VEGF-A, supporting the arrival of new HPCs.
The Id-3 (inhibitor of differentiation 3) gene is also overexpressed in premetastatic clusters. This gene seems to be important for VEGFR1+ mobilization to the premetastatic niche.

Experiments with VLA-4, Id-3 or MMP-9 knock-out mice resulted in preventing premetastatic niche and metastasis formation.

Once the premetastatic niche is complete (HPCs, fibroblast, fibronectin), high expression of SDF-1 (stromal derived factor 1) are detected in the premetastatic clusters. CXCR4, the SDF-1 receptor, is highly expressed on some tumor cells, allowing their attraction to the premetastatic niche.

Fig 8.10, 8.11, 8.12 and 8.13 summarize the different steps of the premetastatic niche formation.

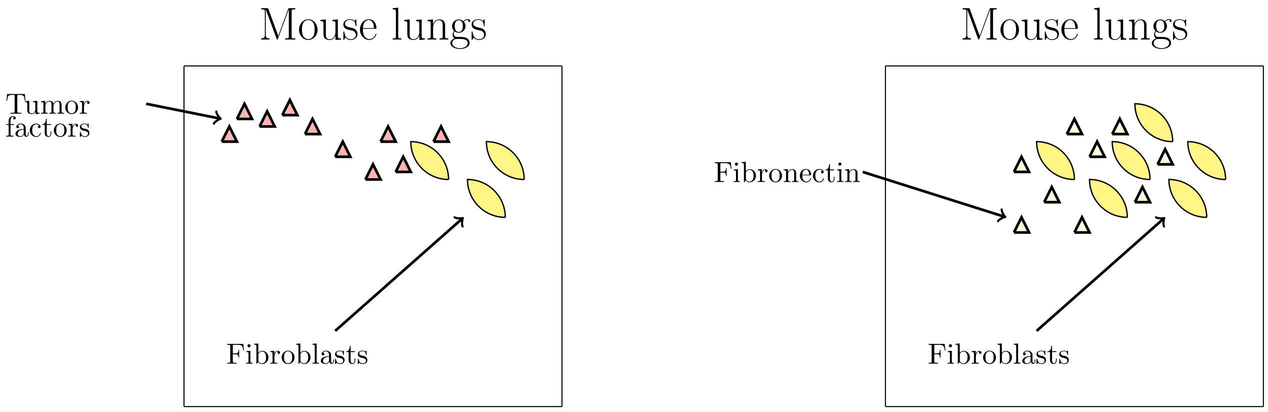


Figure 8.10: First step of the premetasatic niche formation: cytokines released by the primary tumor reach the lungs by the blood stream, stimulate the proliferation of local fibroblasts and the production of fibronectin. In parallel, tumor factors mobilize hematopoietic progenitor cells from the bone marrow to the blood.

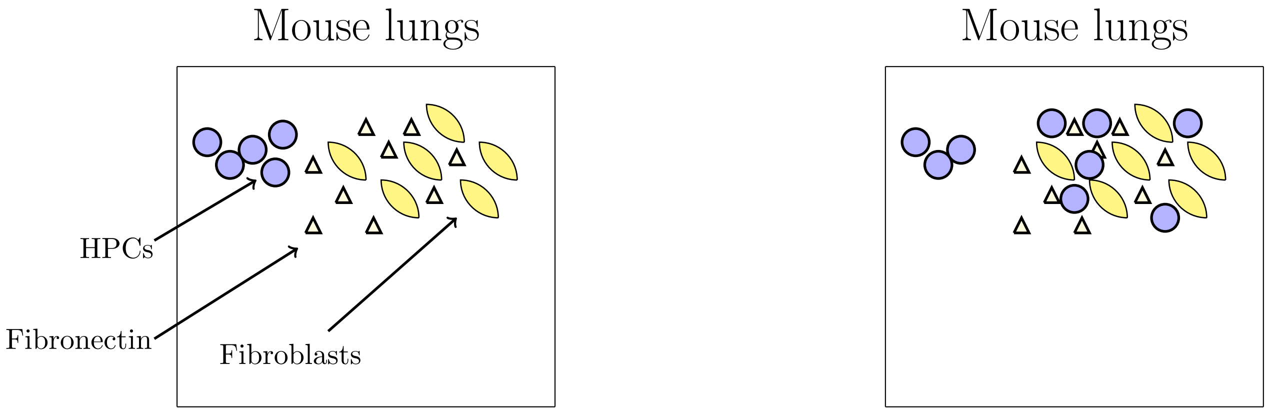


Figure 8.11: Second step of the premetasatic niche formation: HPCs adhere to the premetastatic niche thanks to fibronectin expression.

Figure 8.12: Third step of the premetasatic niche formation: Increased expression of MMP-9 in the premetastatic niche alters the microenvironment by releasing factors like c-kit and VEGF supporting the arrival of new HPCs.

Figure 8.13: Last step of the premetastatic niche formation: High expression of SDF-1 chemokine induces CXCR4 expressing tumor cells attraction to the premetastatic niche.

Primary tumor cytokines govern the metastatic potential

The two cancer cell lines used in this study, Lewis lung carcinoma (LLC) and B16 Melanoma, have different metastatic potentials. Indeed, B16 is a more aggressive cell line that metastasizes in more different organs. In order to understand these different aggressivenesses, Kaplan et al. performed a last experiment where LLC tumor is implanted after injection of MCM (melanoma conditioned media) in the mice. Fig 8.14 summarizes the protocol.

The result of this experiment is that LLC, in the presence of MCM (melanoma conditioned media), metastasized in the lungs and the liver but also in organs that are common metasatic sites for the B16 melanoma cell line (spleen, intestine, oviduct). Usually, LLC does not metastasizes in these organs. It means that metastatic potential is in part governed by the tumor factors and not only for mechanical reasons. Moreover, an analysis has been made to identify the factors secreted by the tumors and responsible of the metasatic potential. It demonstrated that vascular endothelial growth factor is overexpressed by both LLC and B16 melanoma. However, high concentrations of placental growth factor (Plgf) have been detected for B16 melanoma but not for LLC. Plgf could be in part responsible of the premetastatic niche formation in the organs where B16 metastasizes but LLC does not, in stimulating fibroblasts proliferation and fibronectin expression in these organs, whereas VEGF stimulates fibroblasts proliferation and production of fibronectin in the lungs only.

Conclusions

The article [10], shed light on the different steps of the premetastatic niche formation. Although not

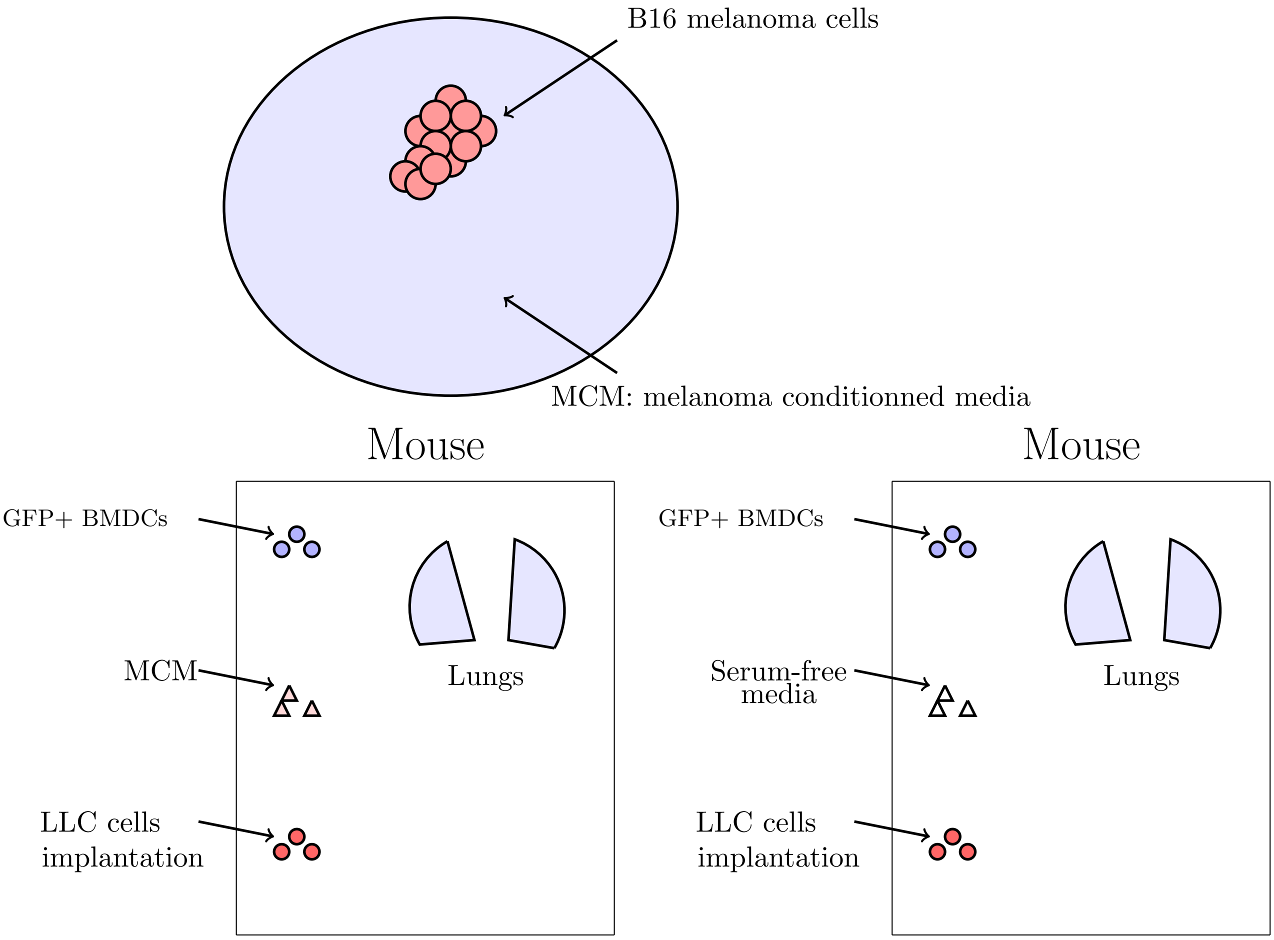


Figure 8.14: *In vitro* Culture of B16 melanoma cells to obtain the MCM and then MCM injection and serum-free media injection for the control group, followed by subcutaneous implantation of LLC cells

precisely identified, VEGFR1+ BMDCs or HPCs seem to actively participate to the formation of this niche, especially for the attraction and survival of tumor cells. These HPCs, under stimulation by tumor factors, migrate from the bone marrow to the premetastatic site. Moreover fibrolasts of the premetastatic sites, also stimulated by tumor factors, proliferate and express fibronectin, allowing HPCs adhesion and MMP-9 production by binding with VLA-4 expressed by HPCs. Then high expression of SDF-1 in the premetastatic niche mediates attraction of CXCR4 expressing tumor cells. The differentiation state of the HPCs was unclear in these experiments but they may need to maintain to an undifferentiated state, perhaps to keep a proliferation potential in the premetastatic niche. Other cells like endothelial progenitor cells could further support the growth of tumor cells in the metastatic niche.

The premetastatic and metastatic niches

Metastatic dissemination and colonization depend on multiple key factors. The invasive behavior of a tumor is in part due to genetic mutations allowing tumor cells to acquire an aggressive phenotype (loss of cadherins for cell-cell junctions, changes in cellular signals, cytoskeleton structure) [5, 30]. However, metastatic potential is also dependent on microenvironmental changes in the primary and distant sites [30]. Indeed, before the arrival of tumor cells in the distant sites, microenvironmental changes like matrix metalloproteinases (MMP) expression, inflammatory responses, stromal cells recruitment, extracellular matrix alterations (ECM) and oncogenic molecules expression are observed [30].

Further studies of the premetastatic niche have been performed after [10], focusing on other cell lines and organs. A short review of this literature body has been made here. The perspectives of quantitative studies using mathematical modeling are discussed further.

In [166], Kaplan et al. mention that HPCs proliferate and circulate into the blood stream in response to tumor chemokines before to adhere to areas of increased fibronectin in the premetastatic niche. Then high expression of SDF-1 in the premetastatic niche promote tumor cells attraction, attachment, survival and growth of tumor cells. Endothelial progenitor cells (EPCs) then appear in the metastatic niche to promote vasculogenesis and complete the metastatic lesions. Indeed, in [167], Lyden et al. suggest that the metastatic angiogenic switch depends on the arrival of the EPCs in the metastatic niche. VEGFR1 seems to play a role for activation of the HPCs but also for migration and proliferation of EPCs and VEGFR1+ tumor cells by activation of src-family kinase [166]. Moreover it has been shown *in vitro* that HPCs promote binding and proliferation of tumor cells [166]. Kaplan et al. suggested therefore that the absence of HPCs in the metastatic niche can induce metastatic dormancy and that this dormancy can be stopped in the presence of HPCs [166].

Another study highlighted the role of inflammatory chemokines S100A8 and S100A9 in the premetasatic niche formation [141] for LLC and B16 melanoma. These chemokines are expressed by premetastatic lung endothelium and lung CD11b+ myeloid cells stimulated by tumor factors (VEGF-A, TGF-β and TNF-α). Chemokines S100A8 an S100A9 then promote attraction of new CD11b+ myeloid cells and of tumor cells along chemotactic gradients [141]. Using monoclonal antibodies against S100-A8 and S100-A9 critically reduced metasatic colonization in the lungs. In [168], Rafii and Lyden emphasize particular points of [141] and make some links with the first study of the premetastaic niche [10]. First, CD11b+ cells could be the same or regroup common cells with VEGFR1+ HPCs mentioned in [10] and that form the premetastatic niche. According to Rafii and Lyden, these CD11b+ myeloid cells and VEGFR1+ HPCs attract the tumor cells and alter the premetastatic microenvironment in forming “docks” to receive the tumor cells [168]. In [141], Hiratsuka et al. showed also that S100A8 and S100A9 are overexpressed in the lungs but not in other organs like the liver and kidney, making inefficient the monoclonal antibodies against S100-A8 and S100-A9 against metastatic colonization in the liver and kidney. Rafii and Lyden suggest that organ-specific upregulation of premetastatic factors characterizes the organ-specific nature of metastatic colonization. They also mention interesting clinical perspectives as premetastatic markers like the presence of Mac1+ or VEGFR1+ cells, more easy to detect than micrometastases, in tumor-bearing patient, could suggest using adjuvant or neo-adjuvant chemotherapies, or even new therapies targeting specific pathways of the premetastatic niche formation [168].

An other paper is focused on the premetastatic niche formation in breast cancer [169]. In this study, Psaila et al. mention that breast cancer cells express CXCR4 and can be attracted by specialized breast carcinoma-associated fibroblasts, but not by normal fibroblasts, illustrating the importance of the “soil” in the metastatic process. The authors also mention evidences of premetastatic VEGFR1+ cell clusters in the lymph nodes of all patients with metastatic breast cancer and not in patients without metastasis. Interestingly, HPCs were more numerous in premetastatic lymph nodes than in metastatic lymph nodes, showing that the main role of these cells is to prime the soil [169]. Again clinical perpectives are highlighted, suggesting that prognosis of breast tumor bearing patients could be improve in checking the presence of VEGFR1+ HPCs clusters in lymph nodes. Moreover, in this paper, similarities were made between the premetastatic niche and the bone marrow niche where hematopoietic stem cells (HSCs) homeostasis is ensured by interactions with the bone marrow microenvironment. Within the premetastatic niche, HPCs exhibit markers like CD34, CD11b, c-Kit and Sca-1, characterizing their immature state, a common feature with the bone marrow niche. Moreover, HSCs dynamics in the bone marrow is transient, with proliferation phases and quiescence or dormancy phases, which can be compared to metastatic patterns of proliferation and dormancy [169]. Another important point is that VEGFR1+ has been shown to be expressed by breast cancer cells, promoting tumor growth. Targeting this pathway offers promising perspectives as it would inhibit both primary tumor growth and HPCs recruitment for the premetastatic niche formation.

Controversies: the premetastatic niche and the paradoxical role of immune cells

Other studies have been performed on mouse with breast cancer cell lines to study the premetasatic dynamics. In [170], Kowanetz et al. observed the formation of a premetastatic niche into the lungs, but this time, not by progenitor cells but by more differentiated Ly6G+ granulocytes. These cells overexpress Bv8 in the premetastatic lungs, promoting tumor cells attraction. However, in an other study with similar experimental conditions, Granot et al. showed that neutrophils, which are particular granulocytes, inhibit metastatic seeding in the lungs [171]. These paradoxical results make sense with the ambiguous role of the immune system in cancer, which comprises two types of cells, the first being anticancer, the second being cancer promoting [4].

8.2 A mechanistic model of the premetastatic niche

8.2.1 A model for what purpose?

As we previously saw, the premetastatic niche formation is a complex multistep process, which illustrates the dynamical aspect of the metastatic process. Indeed, as illustrated in Fig 8.10, 8.11, 8.12 and 8.13, we can divide the global phenomenon in three successive steps: 1)initiation of the premetastatic niche (proliferation of local fibroblasts and fibronectin expression) and mobilization of BMDCs/inflammatory cells in the blood, 2)Attraction and homing of BMDCs/inflammatory cells to the premetastatic niche and interactions with the local cells (fibroblasts, endothelial cells) to alter the microenvironment of the premetastatic niche to form a hospitable place for future metastatic cells, 3)Attraction and homing of metastatic cells by factors secreted in the premetastatic niche, metastatic cells survival and proliferation supporting by BMDCs/inflamatory cells.

Many biological and clinical questions remain open about the premetastatic niche. Is this niche already formed when a patient is diagnosed with a particular tumor, known to have a metastatic potential? With respect to the stage of the disease (premetastatic, metastatic), targeting the seed (tumor cells) is more or less efficient than targeting the soil (BMDCs/Immune cells, fibroblasts)? With respect to the stage of the disease (premetastatic, metastatic), which molecular pathway is suitable to target? The growing amount of quantitative data makes mathematical modeling a good candidate to solve these quantitative issues relied to cancer dynamical aspects.

In this section, we propose a mechanistic model describing the premetastatic niche formation, and more precisely the premetastatic cellular and molecular dynamics. This model formalizes the biological framework presented in the previous section. Some biological aspects have been simplified but we tried to synthesize and formalize the essential ones. We present the model step by step in mentioning how it is parameterized, and then we present fitting results on the data of Kaplan et al.[10]. Finally, perspectives for future work are suggested, notably the clinical issues and *scenarii* that can be tested from the calibrated model and that could suggest therapeutic strategies. In our modeling approach, we tried as much as possible to 1) build a model highly consistent with the biology of the premetastatic niche 2) build a model with parameters that have a biological meaning and that could be measured or found in the literature 3) only describe the essential biological aspects and those that are relevant to describe to answer to relevant clinical issues.

8.2.2 Model

Before to talk about the model in more details, let us draw our conceptual model of the metastatic process divided into steps comprising the premetastatic and metastatic niches formation. The scheme in Fig 8.15 summarizes this conceptual model.

Main assumptions of the model

Evidences of the premetastatic niche formation have been described for several cell lines (B16 melanoma, LLC, human breast cancer, 4T1, MDA, etc). We model here the system with the B16 melanoma cell

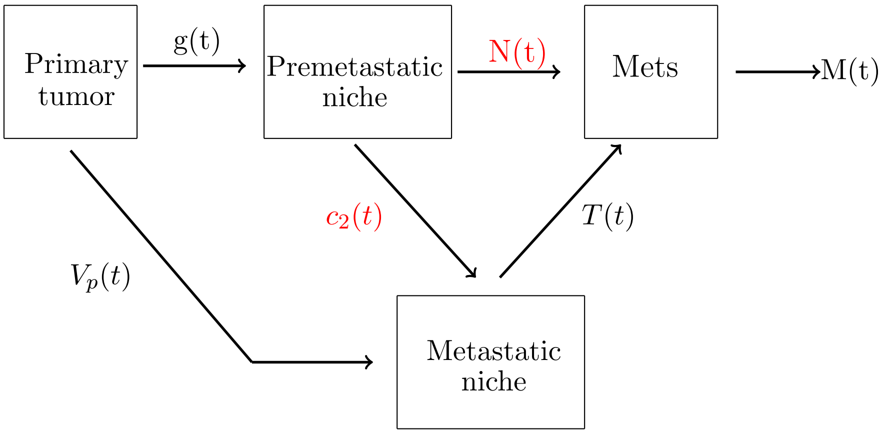


Figure 8.15: Scheme of the different parts of the model. Each part is represented by a box with an input and an output. $g(t)$ represents a concentration of growth factors emitted by the primary tumor (typically VEGF), $V_p(t)$ represents the primary tumor volume, $c_2(t)$ describes a concentration of chemokines attracting the tumor cells to the premetastatic niche, $N(t)$ is a quantitative variable called the state of the niche, which represents the capacity of the tumor cells, with respect to the soil, to survive and proliferate, $T(t)$ is the concentration of circulating tumor cells in the blood, and $M(t)$ the total metastatic burden.

line, based on the biological framework and the data of Kaplan et al. and Hiratsuka et al. [10, 141]. In the different papers dealing with the premetastatic niche formation, many different cellular and molecular species are involved during these premetastatic and metastatic phases [142] but we considered only those present in the two papers of Kaplan et al. [10] and Hiratsuka et al. [141]. In these papers, the differentiation state of the accomplice cells (VEGFR1+ HPCs in [10] and CD11b+ myeloid cells in [10, 141]) is not clear. Kaplan et al. observed undifferentiated progenitor (CD34) and more differentiated cells of the myeloid lineage (CD11b+ as Hiratsuka et al. did). In [10], Kaplan et al. emit a hypothesis (among other ones) that the function of undifferentiated progenitors in this niche is to provide a cell reserve thanks to their proliferative abilities. In our model, VEGFR1+ hematopoietic progenitor cells (HPCs) can essentially proliferate and differentiate into CD11b+ myeloid cells, which then ensure the premetastatic niche functions, like SDF-1 emission, MMP-9 production, and tumor cells survival/proliferation sustaining. For the sake of simplicity, we did not model the dynamics and the effect of VEGFR2+ endothelial progenitor cells (EPCs) on the system. Furthermore, in [10], SDF-1 is suggested to mediate chemoattraction of cancer cells to the premetastatic niche, whereas in [141], chemokines S100A8 and S100A9 are suggested to mediate both chemoattractions of CD11b+ myeloid cells and tumor cells to the premetastatic niche. To keep it simple, we considered in our model that S100 chemokines attract VEGFR1+HPCs and CD11b+ myeloid cells and that SDF-1 attract tumor cells. Moreover we also assumed that the chemotaxis velocity of VEGFR1+HPCs and CD1b+ myeloid cells for S100 chemokines is equal to the chemotaxis velocity of tumor cells for the axis SDF-1/CXCR4. A last major hypothesis of our approach was that metastatic cells need support of CD11b+ myeloid cells for survival and growth in the early stages and become autonomous once the metastasis has activated the angiogenic switch.

First step: the primary tumor growth

We first modeled the growth of the primary tumor by a Gompertz growth law:

$$\begin{cases} \dfrac{\mathrm{d}V_p(t)}{\mathrm{d}t} = g(V_p(t)), \\ V_p(0) = V_{inj}, \\ g_p(V_p) = (\alpha - \beta \ln(V_p))\, V_p. \end{cases} \tag{8.1}$$

The parameters β and α of the Gompertz have been calibrated on B16 melanoma subcutaneous growth data [172]. The result of the fit is presented in Fig 8.16 and the values of parameters α and β are presented in the parameter table 8.1.

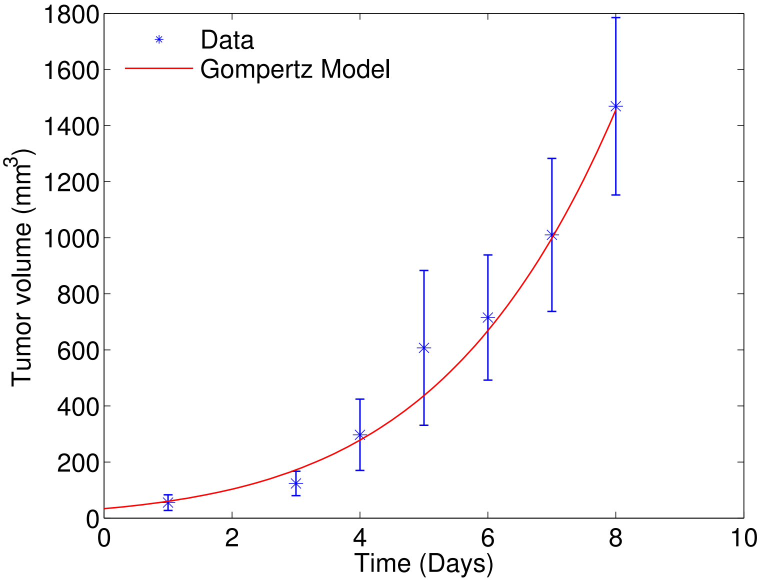


Figure 8.16: Fitting of the Gompertz model on subcutaneous B16 melanoma growth data from [172].

The premetastatic niche

We then focus on the premetastatic phase, as indicated in Fig 8.17.

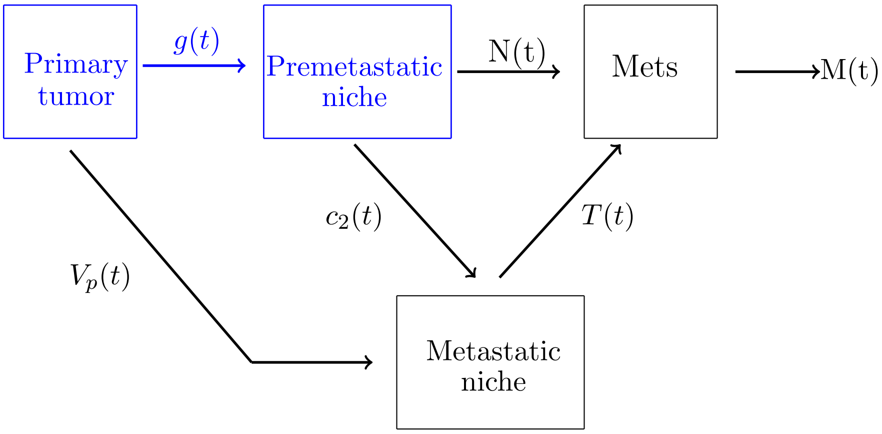


Figure 8.17: The premetastatic phase

Cellular species
Let us introduce the different cellular and molecular species:

- Fb lung fibroblasts
- H_{blood} circulating hematopoietic progenitor cells
- H_{lungs} hematopoietic progenitor cells in the lungs
- My_{blood} circulating myeloid cells

- My_{lungs} Myeloid cells in the lungs
- E extracellular matrix density
- N niche state
- T_{blood} circulating tumor cells in the blood
- B metastatic burden

The molecular species

- g_{blood} concentration of growth factors (VEGF) emitted by the tumor cells within the blood
- g_{lungs} concentration of growth factors (VEGF) emitted by the tumor cells within the lungs
- f fibronectin expression within the lungs
- c_1 Myeloid cell's chemoattractants (S100 chemokines) flux
- M matrix degrading metalloproteinases
- c_2 Tumor cell's chemoattractants (SDF1) flux

Initiation of the premetastatic niche and mobilization of BMDCs into the blood
The growth factors are produced by the primary tumor into the blood and the lungs. We assume that the production of growth factors by the B16 primary tumor is proportional to the tumor volume. Indeed, it seems reasonable to think that these growth factors are emitted from the vasculature of the primary. Formally, it would rather result in a power law for the emission of growth factors, that is an emission of the form $\Pi_g V_p^{\alpha}$ with α a coefficient comprised between 0 and 1 traducing the fractal dimension of the vasculature. We tried to determine a parameter α characterizing the B16 cell line growth. To perform it, we calibrated a power law model (see section 5.1 of the chapter 5) of growth on B16 melanoma subcutaneous growth data from [172]. The result of the fit is presented in Fig 8.18 and the estimated value of α was 0.98, resulting in a quasi-exponential growth. We therefore made the assumption that the emission of growth factors is proportional to the volume of the tumor.

A part of the growth factors produced in the lungs originates also from degradation of the extracellular matrix by MMP-9, releasing soluble factors such as c-kit ligand or VEGF that can stimulate homing survival and proliferation of HPCs. These growth factors are consumed by cells when biding to cell surface receptors. Growth factors activate mobilization and proliferation of HPCs into the blood and the lungs. Moreover, HPCs differentiate naturally into myeloid cells. The dynamics of the growth factors, HPCs, and myeloid cells concentration within the blood and the lungs are described as follows:

$$\begin{aligned}
\frac{\mathrm{d}g_{\text{blood}}}{\mathrm{d}t} &= \Pi_g V_p - \kappa_g H_{\text{blood}} g_{\text{blood}} - \delta_{g_{\text{blood}}} g,\\
\frac{\mathrm{d}g_{\text{lungs}}}{\mathrm{d}t} &= \Pi_g V_p + R \times E \times M - \kappa_g H_{\text{lungs}} g_{\text{lungs}} - \kappa_g F b g - \delta_{g_{\text{lungs}}} g,\\
\frac{\mathrm{d}H_{\text{blood}}}{\mathrm{d}t} &= \gamma_H \lambda(g, g_{\text{threshold}_1}) H_{\text{blood}} \left(1 - \frac{H_{\text{blood}}}{H_{\max}}\right) - \delta_H (H_{\text{blood}} - H_{\text{basal}}),\\
\frac{\mathrm{d}My_{\text{blood}}}{\mathrm{d}t} &= d_H H_{\text{blood}} - \delta_{My}(My_{\text{blood}} - My_{\text{basal}}).
\end{aligned}$$

The parameter Π_g represents the production of growth factor concentration per tumor cubic millimeter per unit of time. The parameter κ_g represents the consumption of growth factors by the different cellular species and δ_g represents the growth factors clearance. The parameter R represents a release of sequestrated growth factors from the extracellular matrix when the matrix is degraded

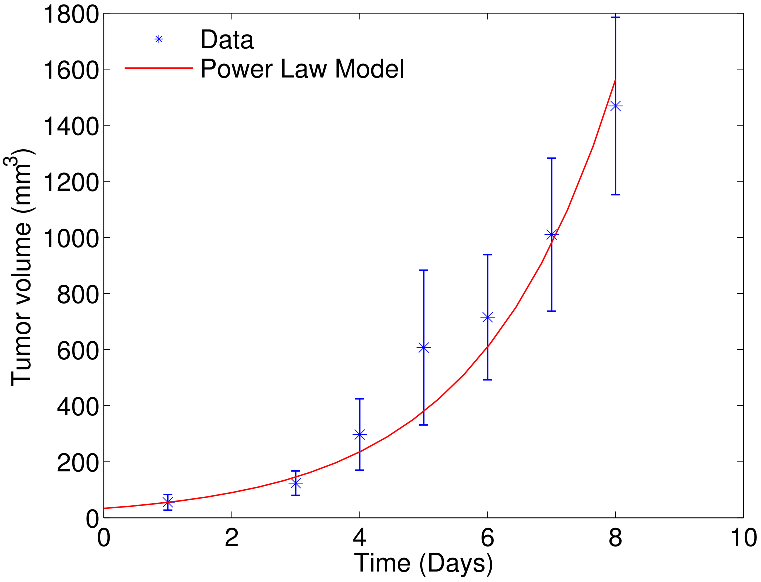


Figure 8.18: Fitting of the power law model on subcutaneous B16 melanoma growth data from [172].

[10] by MMPs. The parameter γ_H is the growth rate of the hematopoietic progenitor cells, $\lambda(g)$ is a function ensuring that proliferation is activated when g is greater than a threshold $g_{threshold}$ and not activated when lower than this threshold. Indeed, mobilization of HPCs is stimulated by tumor growth factors, but the precise mechanism is not clear in [10]. We make here the assumption that HPCs mobilization (and the other premetastatic events like fibroblasts proliferation or chemokines emission by the lung endothelium) is governed by a threshold-mediated activation. The function $\lambda(g)$ is therefore a regularization of the heaviside function:

$$\lambda(g, g_{\mathrm{threshold}}) = \frac{1 + \tanh(S(g - g_{threshold}))}{2}, \tag{8.2}$$

where S is a coefficient. Parameter d_H represents the differentiation rate of HPCs into myeloid cells. Although it is known that progenitor cells, unlike stem cells, have a limited self-renewal ability [151], we assumed here, for the sake of simplicity, an unlimited self-renewal capacity and modeled the self-renewal and differentiation of hematopoietic progenitor cells into myeloid cells as Michor et al. did for hematopoietic stem cells in [55]. The logistic term $1 - \frac{H_{\mathrm{blood}}}{H_{\max}}$ leads to a saturation of the proliferation when the cell density becomes close to the maximal cell density parameter $H_{\max}$. The value of this parameter has been fixed to 10^6 cells $\cdot$ mm^{-3}, according to the well-established conversion rule 1 mm^3 $\simeq 10^6$ cells [146]). The term $-\delta_H(H_{\mathrm{blood}} - H_{\mathrm{basal}})$ contains a death term $-\delta_H H_{\mathrm{blood}}$ and source term $\delta_H H_{\mathrm{basal}}$ ensuring homeostasis of the cell population when the density is lower than the basal value. The parameters δ_H and δ_{My} are respectively the death rates of the hematopoietic progenitor cells and the myeloid cells.

In parallel, tumor growth factors also stimulate proliferation of lung fibroblasts and fibronectin expression by fibroblasts within the lungs. Fibroblasts dynamics and fibronectin production by fibroblasts are modeled as follows:

$$\frac{\mathrm{d}Fb}{\mathrm{d}t} = \gamma_{Fb}\lambda(g, g_{\mathrm{threshold}_2})Fb - \delta_{Fb}(Fb - Fb_{basal}) - \kappa_B Fb B,$$
$$\frac{\mathrm{d}f}{\mathrm{d}t} = \Pi_f \lambda(g, g_{\mathrm{threshold}_2})Fb\left(1 - \frac{f}{f_{\max}}\right),$$

where γ_{Fb} is the proliferation rate of the fibroblasts and δ_{Fb} is the death rate of the fibroblasts. The term $-\kappa_B Fb B$ traduces that the metastatic burden (B) destroys the lung tissue in growing with a consumption rate κ. The parameter Π_f is the production rate of fibronectin by the fibroblasts.

Recruitment of bone derived cellular helpers by the premetastatic lung

For the sake of simplicity, transmigration of HPCs and myeloid cells from the blood to the premetastatic niche, including extravasation through endothelium and migration toward chemotactic gradients within the lung parenchyma, are modeled by chemotactic transport in 1D channels. The characteristic length of these channels, denoted L, has been fixed to 1 mm considering the dimensions of the lungs and the capillary beds. S100 chemokines concentration in a channel is denoted c_1. The vascular network is located in $x = 0$ and the premetastatic niche in $x = L$. The space-depending densities of circulating HPCs and myeloid cells in the channels are respectively denoted by $H_c(t, x)$ and $My_c(t, x)$.

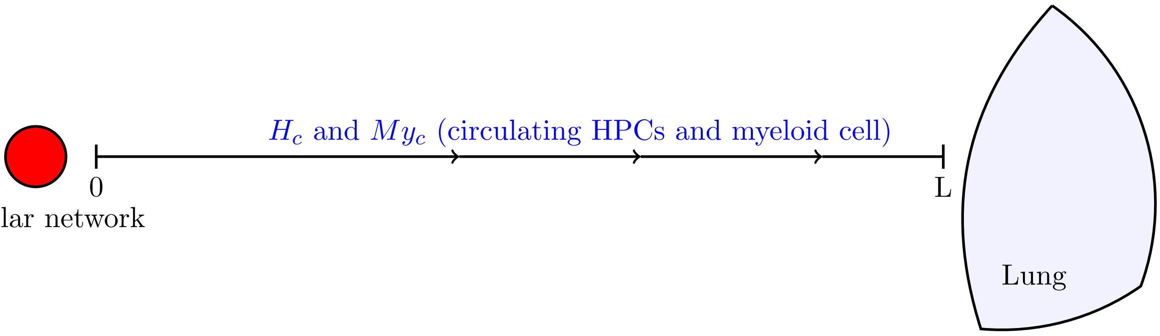


Figure 8.19: Chemoattraction of HPCs/myeloid cells from the capillary beds to the premetastatic niche

Endothelial cells are known to play an important role in cells extravasation [47] but we focus here on an other function of the endothelium. Inflammatory chemokines S100A8 and S100A9 are produced by pulmonary endothelial cells during the premetastatic phase of B16 melanoma [141] and attract HPCs/Myeloid cells toward chemotactic gradients. Once established into the premetastatic lung, myeloid cells also produce S100A8 and S100A9 chemokines [141]. For the sake of simplicity, we do not consider both S100A8 and S100A9 chemokines but chemokine S100A8 only because it is more potent chemoattractant than S100A9 in [141]. The concentration of S100A8 chemokine between the premetastatic niche and the endothelium is denoted by c_1 and its dynamics is governed by the 1D diffusion equation:

$$\begin{aligned}
&\frac{\partial c_1(t,x)}{\partial t} - D_1 \frac{\partial^2 c_1(t,x)}{\partial x^2} = 0 \qquad \forall t, x \in [0,T] \times [0,L]\,,\\
&\frac{\partial c_1(t,L)}{\partial x} = \Pi_{c_1} \lambda(g(t), g_{\text{threshold}_3})(Endo + My_{lungs}(t)),\\
&c_1(t,0) = 0,\\
&c_1(0,x) = 0,
\end{aligned}$$

where D_1 represents the diffusivity of S100 chemokines. The parameter Π_{c_1} is the S100A8 gradient produced per activated endothelial cell. The parameter *Endo* represents the density of endothelial cells producing S100A8 within the premetastatic niche. The function λ is the same as presented in (8.2) because cells of the lung endothelium produce S100A8 chemokine in response to the tumor factors. The boundary condition $c_1(t, 0) = 0$ means that we assume the concentration of S100A8 chemokine within the blood equal to 0. We make here the assumption that the diffusion of the chemokine is much faster than its production by the cells. This assumption results in a stationary state of the diffusion equation:

$$\begin{aligned}
&D_1 \frac{\partial^2 c_1(t,x)}{\partial x^2} = 0 \qquad \forall t, x \in [0,T] \times [0,L] \\
&\frac{\partial c_1(t,L)}{\partial x} = \Pi_{c_1} \lambda(g(t), g_{\text{threshold}_3})(Endo + My_{lungs}(t)) \\
&c_1(t,0) = 0 \\
&c_1(0,x) = 0
\end{aligned}$$

The solution writes $c_1(t,x) = a(t)x + b(t)$. The boundary condition in $x = L$ leads to $a(t) = \Pi_1 \lambda(g(t))(Endo + My_{lungs}(t))$. The boundary condition in $x = 0$ leads to $b = 0$.
Chemokine gradients attract HPCs and CD11b+ myeloid cells that join the niche by chemotaxis. Chemotaxis is an active motion of the cells, unrelated to the blood flow transport. We make the asumption that HPCs and CD11b+ myeloid cells have the same chemotaxis velocity χ. Denoting by $H_c(t,x)$ and $My_c(t,x)$ the respective densities of circulating HPCs and myeloid cells in a 1D channel from the blood stream to the premetastatic niche, this chemotaxis transport on $H_c(t,x)$ is written as follows:

$$\begin{aligned}
&\frac{\partial H_c(t,x)}{\partial t} + \chi \frac{\partial}{\partial x}(H_c(t,x) \frac{\partial}{\partial x} c_1(t,x)) = -\delta_H H_c(t,x) \qquad \forall t, x \in [0, t_{max}] \times [0,L] \\
&H_c(t,0) = H_{\text{blood}}(t)
\end{aligned} \tag{8.3}$$

As we can see, as the velocity $\chi \frac{\partial}{\partial x} c_1$ is always positive, one needs only a boundary condition in $x = 0$, which is $H_c(t,0) = H_{\text{blood}}(t)$, so that the cell density at the entrance of the channel is equal to the cell density in the blood. During the transmigration, we assume that cells have a life span equal to their life span in the blood and the lung. Let us recall that δ_H and δ_{My} are the death rates of the HPCs and myeloid cells. The equation on the circulating myeloid cells density is similar:

$$\begin{aligned}
&\frac{\partial My_c(t,x)}{\partial t} + \chi \frac{\partial}{\partial x}(My_c(t,x) \frac{\partial}{\partial x} c_1(t,x)) = -\delta_{My} My_c(t,x) \qquad \forall t, x \in [0, t_{max}] \times [0,L] \\
&My_c(t,0) = My_{\text{blood}}(t)
\end{aligned} \tag{8.4}$$

The gradient of c_1 is known and does not depend on the space so we can exactly solve (8.3) and (8.4) thanks to the characteristics. The characteristic curve $X(t,\tau,X_\tau)$ associated to velocity field $\chi \frac{\partial}{\partial x} c_1$, is the solution of:

$$\begin{aligned}
&\frac{\partial X(t,\tau,X_\tau)}{\partial t} = \chi \nabla c_1(t) \\
&X(\tau,\tau,X_\tau) = X_\tau
\end{aligned} \tag{8.5}$$

Let us define the following change of variable

$$\begin{aligned}
\tilde{H_c}(t,\tau,X_\tau) &= H_c(t, X(t,\tau,X_\tau)) \\
\tilde{My_c}(t,\tau,X_\tau) &= My_c(t, X(t,\tau,X_\tau))
\end{aligned}$$

It allows to write that $\frac{\partial \tilde{H_c}(t,\tau,X_\tau)}{\partial t} = -\delta_H \tilde{H_c}(t,\tau,X_\tau)$ and $\frac{\partial \tilde{My_c}(t,\tau,X_\tau)}{\partial t} = -\delta_{My} \tilde{My_c}(t,\tau,X_\tau)$.
Integrating the characteristic equation (8.5) between τ and t leads to:

$$X(t,\tau,X_\tau) - X(\tau,\tau,X_\tau) = \int_\tau^t \chi \frac{\partial}{\partial x} c_1(t)$$

Now, we are interested by the values in $x = L$ because we want to compute the number of cells entering into the premetastatic niche. Therefore, for all t, we can deduce the value $H_c(t,L)$ in writting $L = X(t,\tau,X_\tau)$ and $X(\tau,\tau,X_\tau) = 0$. We have then to determine the time τ such that:

$$\int_\tau^t \chi \frac{\partial}{\partial x} c_1(t) = L,$$

which means that the cells have traveled a distance L during a time $t-\tau$. For all time t, we finally have:

$$H_c(t,L)=\tilde{H}_c(t,\tau,0)=\exp(-\delta_H(t-\tau))\tilde{H}_c(\tau,\tau,0)=\exp(-\delta_H(t-\tau))H_{\text{blood}}(\tau)$$

The decreasing cell density along the characteristic is due to the cell death term $-\delta_H H_c$. The number of new cells entering into the lungs per unit of time at the time t is given by:

$$Nb_{cells}(t)=H_c(t,L)\chi\frac{\partial}{\partial x}c_1(t)S_{\text{extravasation}},$$

where $Nb_{cells}(t)$ is the number of new entering cells per unit of time and $S_{\text{extravasation}}$ is the total vascular surface crossed by the cells. We define this surface as follows:

$$S_{\text{extravasation}}=S_{\text{capillaries}}Perm,$$

Therefore, increase of HPCs density in the lungs by arriving of new cells is obtained by $Nb_{cells}(t)$ divided by the lungs total volume ($V_{lungs}=700\text{ mm}^3$). $S_{\text{capillaries}}$ is the total surface of the capillary wall and $Perm$ a permeability parameter comprised between 0 and 1. More precisely, the surface that circulating cells can cross is a fraction of the capillary wall surface. The dynamics of HPCs and myeloid cells in the lungs is then given by the following equations:

$$\frac{\mathrm{d}H_{\text{lungs}}(t)}{\mathrm{d}t}=\frac{H_c(t,L)\chi\frac{\partial}{\partial x}c_1(t)S_{\text{extravasation}}}{V_{\text{lungs}}}+\gamma_H\lambda(g,g_{\text{threshold}_1})H_{\text{lungs}}\left(1-\frac{H_{\text{lungs}}}{H_{\text{max}}}\right)-\delta_H(H_{\text{lungs}}-H_{\text{basal}})-\kappa_B B(t)H_{\text{lungs}}$$

$$\frac{\mathrm{d}My_{\text{lungs}}(t)}{\mathrm{d}t}=\frac{My_c(t,L)\chi\frac{\partial}{\partial x}c_1(t)S_{\text{extravasation}}}{V_{\text{lungs}}}+d_H H_{\text{blood}}-\delta_{My}(My_{\text{lungs}}-My_{\text{basal}})-\kappa_B B(t)My_{\text{lungs}}$$

The state of the niche $N(t)$ is then defined as follows:

$$N(t)=\nu My_{lungs}$$

We further explain in more detail what does represent this sate of the niche. As in the blood, the HPCs proliferate in the lungs with a growth rate γ_H and die with a death rate δ_H, and the myeloid cells are formed by differentiation of the HPCs and die with a death rate δ_{My}. As the fibroblasts, HPCs and myeloid cells are consumed by the metastatic burden with a consumption rate κ_B.
Myeloid progenitors expressing VLA4, a fibronectin ligand, produce MMP-9 that degrade the extracellular matrix. Sequestrated factors like VEGF are then released from the matrix and basal membranes are degraded by MMPs, enhancing invasion of circulating HPC/myeloid cells and circulating tumor cells. The dynamics of MMP-9 (the expression is called M in the equations) is governed by emission from myeloid cells and by an elimination of MMP-9 with a clearance noted δ_M. Indeed, experimental evidences of MMPs natural decay have been shown [173, 174]. We did not find measured values of this parameter in the litterature, therefore we calibrated it on the data. In our model, dynamics of MMP-9 and ECM are governed by the following equations:

$$\frac{\mathrm{d}M}{\mathrm{d}t}=\Pi_M(f-f_{\text{min}})My_{\text{lungs}}-\delta_M M,$$
$$\frac{\mathrm{d}E}{\mathrm{d}t}=-\kappa_E EM.$$

Π_M is the production of MMP-9 per cell per day and δ_M is the clearance of MMP-9 in the lungs. The term $\kappa_E EM$ models the consumption of ECM by MMP-9. The parameter κ_E is the consumption of ECM per unit of MMP9 expression per unit of ECM expression per day. f is the fibronectin expression in the lungs, and f_{min} is the minimal or basal expression of fibronectin in the lungs. We then define the permability $Perm$ as follows:

$$Perm=\frac{E_{\text{max}}-E}{E_{\text{max}}},$$

where E_{max} is a parameter of ECM maximal density. Consistently with the biological theory and our modeling assumptions, the more the extracellular matrix and the basal membranes of the lungs are degraded, the more cells (cellular helpers or tumor cells) can invade the lungs.

Attraction of tumor cells to the premetastatic niche

We now focus on the formation of a metastatic niche, as illustrated in Fig 8.20. Once the premetastatic niche has been formed, factors emitted from this niche like SDF-1 (CXCL12) allow homing and attraction of tumor cells to the premetastatic niche. The tumor cells transmigrate from the blood to the premetastatic niche that ensures their engraftment to the lung tissue to finally form a metastatic niche. The space-depending density of circulating tumor cells in the 1D channels between capillary beds and premetastatic niche is noted $T_c(t, x)$.

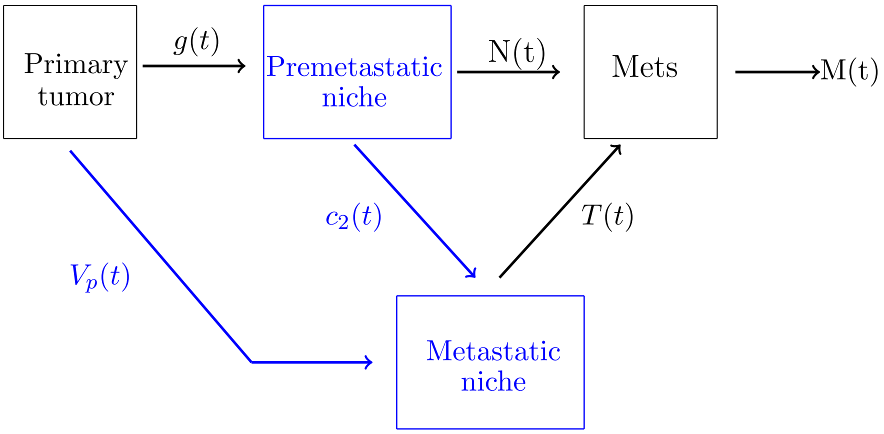


Figure 8.20: The metastatic niche formation

We consider a continuous emission of circulating tumor cells (CTCs) in the blood from the primary. Similarly to the emission of growth factors, we assume the emission rate of CTCs proportional to the volume of the tumor. This size-dependence assumption might be debatable because it is known that tumor cells invasion, extravasation and colonization abilities are in part governed by acquisition of an aggressive phenotype by tumor cells thanks to particular mutations and interactions with the microenvironment [5, 30]. However, in our case, we consider the B16 cell line, which consists in cells that have been designed to be aggressive and metastatic, so we assume that the cells have already the potential to metastasize. We therefore write the dynamics of CTCs in the blood as follows:

$$\frac{\mathrm{d}T_{\text{blood}}}{\mathrm{d}t} = \mu V_p - \delta_T T_{\text{blood}},$$

where T_{blood} is the tumor cell density in the blood, μ is the number of tumor cells emitted in the blood per tumor volume unit per day, and δ_T is the death rate of tumor cells. It has been shown in [10, 141] that tumor cells are attracted by factors emitted in the premetastatic niche. SDF-1 chemokine is implied in homing and attraction of B16 CXCR4+ tumor cells in [10], whereas S100A8 and S100A9 are suggested to play a similar role in [141]. For the sake of simplicity, we consider here only one chemoattractant specie for the transmigration and attraction of tumor cells.

The diffusion of SDF-1 from the premetastatic niche to the blood and the chemoattraction of CXCR4+ tumor cells are treated similarly to the diffusion of S100A8 and chemoattraction of HPCs/myeloid cells, as shown in Fig 8.21.

In [10], high expression of SDF-1 is observed in the premetastatic lungs before the arrival of CTCs, but the mechanisms by which myeloid cells produce SDF-1 is not clear. We therefore assume that once myeloid cells adhere to the premetastatic niche, they continuously produce SDF-1, keeping implicit interactions between myeloid cells and lung microenvironment mediating SDF-1 production. Under the same modeling asumptions as for S100A8 chemokine, SDF-1 diffusion is modeled by the following

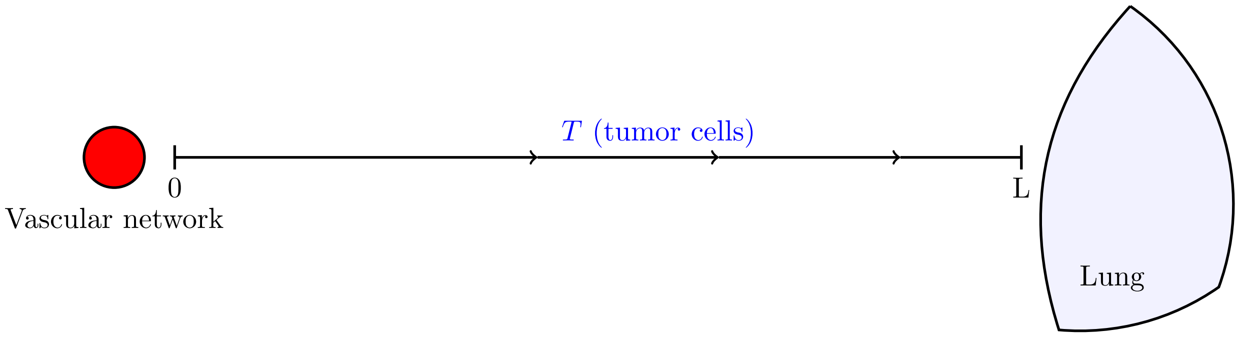


Figure 8.21: Chemoattraction of circulating tumor cells from the capillary beds to the premetastatic niche

stationary 1D-diffusion equation:

$$\begin{aligned}
&D_2\frac{\partial^2 c_2(t,x)}{\partial x^2} = 0 \qquad \forall t,x \in [0,t_{max}]\times[0,L]\\
&\frac{\partial c_2(t,L)}{\partial x} = \Pi_{c_2} M y_{lungs}(t)\\
&c_2(t,0) = 0\\
&c_2(0,x) = 0.
\end{aligned}$$

The parameter Π_{c_2} represents the chemokine flux generated per myeloid cell in the lungs. The chemottraction of circulating tumor cells is written:

$$\begin{aligned}
&\frac{\partial T_c(t,x)}{\partial t} + \chi\frac{\partial}{\partial x}(T_c(t,x)\frac{\partial}{\partial x}c_2) = -\delta_T T_c(t,x) \qquad \forall t,x \in [0,t_{max}]\times[0,L]\\
&T_c(t,0) = T_{\text{blood}}(t)
\end{aligned}$$

From $T_c(t,L)$, we can deduce the number of tumor cells entering in the metastatic niche per unit of time $\beta(t)$:

$$\beta(t) = T_c(t,L)\chi\frac{\partial}{\partial x}c_2(t)S_{\text{extravasation}}$$

Metastatic growth: coupling of the premetastatic niche model with a model of metastatic growth

The output of the previous model is the number of metastatic entering in the metastatic niche per unit of time. Now we have to model the growth of the metastases formed by these metastatic cells (see Fig 8.22.

Metastatic dissemination and colonization modeling
The model we used to describe the growth of metastatic lesions is the same as in the chapter 6. This model, introduced by Iwata et al. [17] and validated on *in vivo* mouse data for the metastatic burden dynamics [25, 26], describes the size distribution of the metastatic lesions at the distant site by means of a size-dependent density $\rho(t,v)$ of metastatic colonies of size v at time t, i.e $\rho(v,t)\mathrm{d}v$ is the number of metastatic colonies with a size comprised between v and $v+\mathrm{d}v$. Secondary emission of metastases are still neglected. This time, metastatic colonies grow into the lungs according to a growth law that depends both on the volume of the considered metastasis and of the state of the niche $N(t)$.
Each metastasis is assumed to start from one metastatic cell, as in [17, 25, 26]. In our study we also assumed that metastatic cell that manage to join the premetastatic/metastatic niche form a new

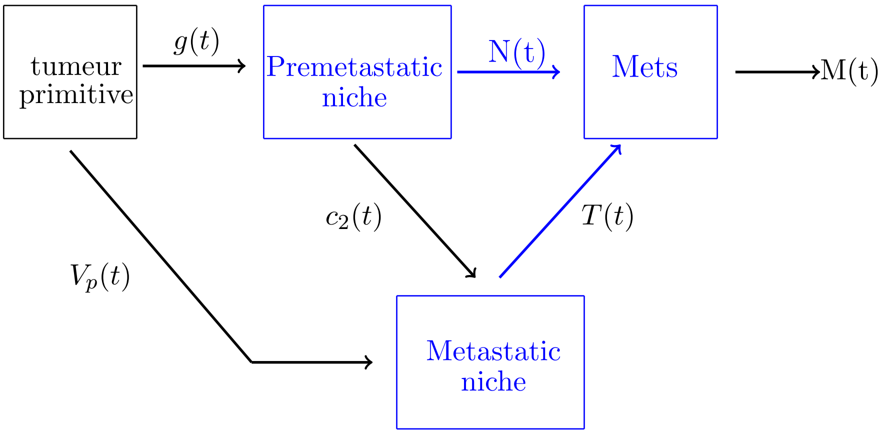


Figure 8.22: Metastatic colonization and growth

metastasis. Overall, the model writes as a transport equation on ρ, endowed with suitable boundary and initial conditions [17]:

$$\begin{cases} \partial_t\rho(t,v) + \partial_v(\rho(t,v)g(v,N(t))) = 0 & t \in]0,+\infty[, \ v \in]V_0,+\infty[\\ g(V_0,N(t))\rho(t,V_0) = \beta(t) & t \in]0,+\infty[\\ \rho(0,v) = 0 & v \in]V_0,+\infty[\end{cases} \tag{8.6}$$

where $\beta(t)$ is defined as follows:

$$\beta(t) = T_c(t,L)\chi\frac{\partial}{\partial x}c_2(t)S_{\text{extravasation}}$$

From the solution of this problem, one of the macroscopic quantities that we fitted on the data of Kaplan et al. [10] for our purpose was the total metastatic burden, defined by:

$$B(t) = \int_{V_0}^{+\infty} v\rho(t,v)dv \tag{8.7}$$

We now have to define the growth $g(v,N(t))$ of the metastases.

The early and late stages of metastatic growth

In [10], Kaplan et al. showed that HPCs/myeloid cells promote attraction and growth of LLC/B16 tumor cells *in vitro*. Whereas the mechanism of attraction seems relatively well understood (chemoattraction by SDF-1/CXCR4 axis), the role of HPCs/myeloid cells in the survival and proliferation of tumor cells is not so clear. It could be mediated by growth factors emission from HPCs/myeloid cells but up to our knowledge, no evidences of such a phenomenon has been made. It has just been shown that premetastatic clusters of HPCs/myeloid cells are required for survival of tumor cells and establishment of metastatic foci. However, it has also been shown that HPCs/myeloid cells only are not enough to promote macrometastatic lesions. VEGFR2+ endothelial progenitor cells (EPCs) are required to trigger the angiogenic switch leading to growth until formation of visible metastases [167, 10]. We did not model the dynamics of EPCs but we included two stages of growth for the metastases, one stage for small metastatic volumes (until 3 mm^3), where presence of HPCs/myeloid cells is required for survival and proliferation of metastatic cells and a second stage, where the tumor is autonomous in the sense that it mediates by itself its microenvironment, stimulating its carrying capacity by angiogenesis. We chose a volume of 3 mm^3 for the angiogenic switch volume, according to the studies of Folkman [35]. For the early stage, we modeled the proliferation of metastatic cells with a simple exponential growth but with a growth rate equal to state of the niche:

$$\frac{\mathrm{d}v(t)}{\mathrm{d}t} = N(t)v(t),$$

where $N(t) = \nu M y_{lungs}(t)$. The growth rate is hence proportional to the number of cellular helpers. For the autonomous macroscopic growth, we chose the Gompertz growth law, which has been proven able to describe *in vivo* tumor growth in numerous animal experimental systems. So we have:

$$\begin{cases} g(v, N(t)) = N(t)v & \text{if } v \leq 3 \\ g(v, N(t)) = (N(T) - \beta_m \log\left(\frac{v}{3}\right))v & \text{else,} \end{cases} \tag{8.8}$$

where T is the time of the angiogenic switch, that is the moment when the metastasis reaches a size of 3 mm^3. β_m represents the parameter of growth rate decreasing during the autonomous (angiogenic) growth. This growth law traduces the fact that a micrometastasis needs cellular helpers of the metastatic niche to sustain its growth and become able to stimulate its carrying capacity by its own during the macrometastatic phase, and that the starting growth rate of the vascular growth, that is $N(T)$ depends on the resources provided by the premetastatic niche.

Whereas the parameter ν was calibrated on data of the metastatic burden (see Fig 8.23), the parameter β_m of the Gompertz was fixed at the same value as the calibrated one for the primary tumor, characterizing that metastases and primary tumor are tumors from the same cell line. Under our modeling assumptions, the initial growth rate of the Gompertz law represents the growth potential provided by the microenvironment (nutrients availability, vasculature, cellular helpers, immune system, etc), whereas the parameter β_m of growth rate decrease characterizes the tumor cell line.

8.2.3 Results

Some parameters of the model have been fixed to values retrieved from the literature. The other ones have been calibrated on the data of Kaplan et al. [10]. These data consist of:

- VEGF kinetics in the blood
- Fibronectin expression in the lungs
- HPCs/myeloid cells dynamics in the lungs
- Metastatic burden dynamics in the lungs

Because automating the data assimilation has not been yet performed, the fit was obtained in calibrating the parameters manually, which means that a better fit would be possible. All the parameter values are presented in the table 8.1. The results of the fit are presented in Fig 8.23.

As we can observe, the simulated dynamics of the four observable variables are in accordance with the observations. Except for a few points, the global dynamics are consistent between simulations and data. More importantly, the dynamics of HPCs/myeloid cells and the dynamics of the metastatic burden are well described by the model. First HPCs/myeloid cells are detected at day 12 in the lungs, increasing until days 24-46 to finally decrease until day 30. For days 18, 26 and 30, the simulated curve is not in the error bar. However, the calibration of the parameters being manual, a better fit would be possible. For the metastatic burden, we can see that the simulated curve is in the error bars, with first tumor cells observed in the lungs at day 18 [10], and the first significant amount of tumor cells detected at day 23. The delay of arrival of these tumor cells is due to the prerequired premetastatic niche. Concerning the fibronectin, we can see that the curve crosses the error bars, with a peak at day 5. In the data, we can observe that after the peak, the fibronectin expression drops, which is not the case on the simulations. Indeed, in the model, no mechanisms of consumption or elimination of the fibronectin were incorporated, only a saturation of the production. Finally, the VEGF dynamics is globally respected by the model, but due to the consumption by cellular species and to the clearance in the blood, VEGF concentration strongly decreases at the very beginning, which is not the case in the data. A source term ensuring maintaining of VEGF concentration at the homeostatic level should perhaps be incorporated in the model.

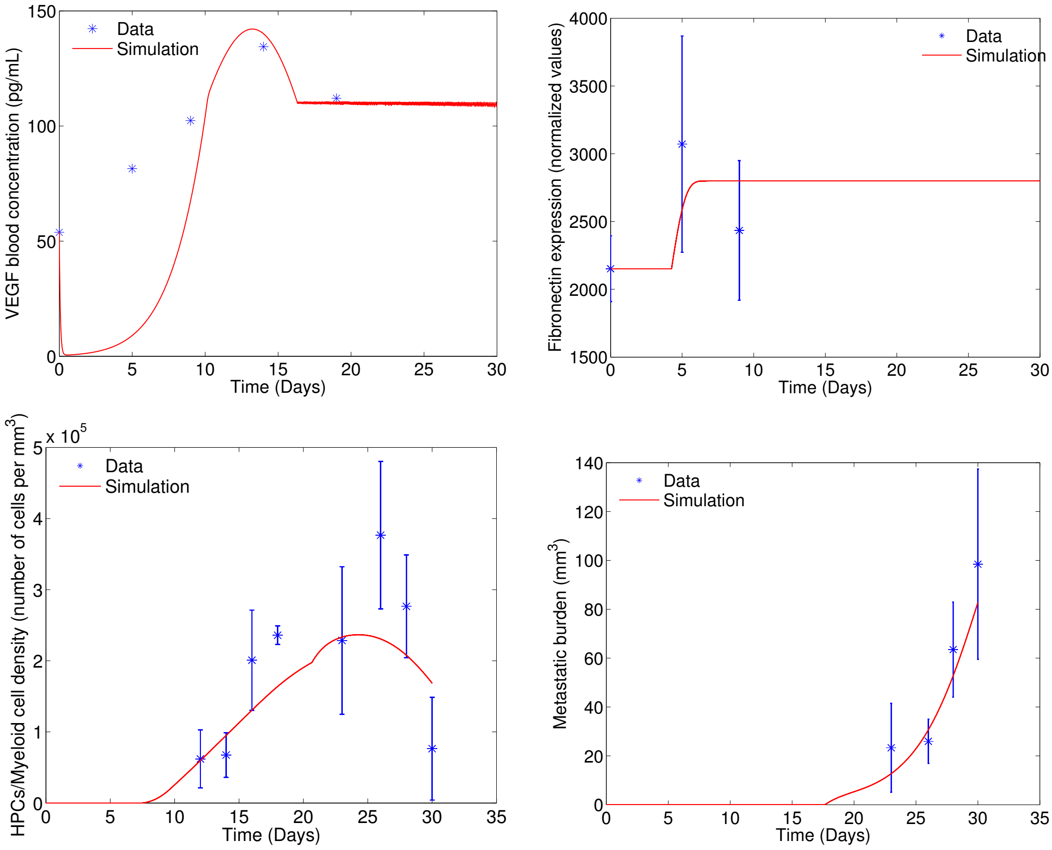


Figure 8.23: Fitting results of the model on the data on VEGF concentration in the blood (top left), fibronectin expression in the lungs (top right), BMDCs dynamics in the lungs (bottom left), metastatic burden dynamics in the lungs (bottom right).

8.2.4 The next step: testing clinical *scenarii*

Now that the model is calibrated, we can use it to assess quantitative predictions on specific situations that have not been tested or can not be tested experimentally. Many clinically relevant questions can be addressed. We propose here some of these that will be tested thanks to the model.

For a given metastatic cancer, are the premetastatic/metastatic niches already formed when the tumor is detected on a patient?
When a tumor is detected in a patient, a critical question is to know if the tumor has also disseminated cells in distant organs and if eventual metastases will blow up after tumor resection or treatment. If metastasis development is dependent on the formation of a permissive niche, as suggested by the work of Kaplan, Lyden, Hiratsuka, Rafii et al. [10, 141, 168], it could be determinant, if no metastasis is visible yet, to know if the soil was permissive to metastatic seeding or not at the time of resection. If it is the case, it would mean that premetastatic niche formation is an early event in cancer development and although representing a relevant therapeutic clue, remains hard to target because the diagnosed patient would already have developed a premetastatic niche and probably a metastatic niche also. In this case, targeting directly the cancer cells, for example with a neo-adjuvant chemotherapy, could be more relevant than targeting the soil. Quite the opposite, if the niche is not ready yet when the patient is diagnosed, targeting the soil could be relevant to prevent metastatic development. Simulations could be performed to test if the premetastatic niche is an early event or not. From the calibrated model, we could run a simulation with a tumor starting from one or few cells so that we simulate a carcinogenesis on a patient. Then we could observe if at detectable tumor sizes, the soil is yet ready or not to support metastatic development.

Param	**Meaning**	**Unit**	**Value**	**Retrieved from**
Π_g	Production of VEGF by the primary tumor	$\mathrm{pg} \cdot \mathrm{mL}^{-1} \cdot \mathrm{mm}^{-3} \cdot \mathrm{day}^{-1}$	2.6	calibrated
κ_g	Consumption of VEGF by HPCs and fibroblasts	$\mathrm{cell}^{-1} \cdot \mathrm{mm}^3 \cdot \mathrm{day}^{-1}$	0.016	calibrated
$\delta_{g_{\mathrm{blood}}}$	Clearance of VEGF in the blood	day^{-1}	3.35	[175]
$\delta_{g_{\mathrm{lungs}}}$	Clearance of VEGF in the lungs	day^{-1}	3.35·e-3	calibrated
R	Release of VEGF by the degraded matrix	$\mathrm{pg} \cdot \mathrm{mL}^{-1} \cdot \mathrm{NE}^{-1} \cdot \mathrm{day}^{-1}$	100	calibrated
γ_H	Proliferation rate of HPCs	day^{-1}	0.35	[55]
$g_{threshold_1}$	Threshold for HPCs activation	$\mathrm{pg} \cdot \mathrm{ml}^{-1}$	110	calibrated
S	Coefficient of the hyperbolic tangent	N.A	100	calibrated
δ_H	Death rate of HPCs	day^{-1}	0.02	[55]
d_H	Differentiation rate of HPCs into myeloid cells	$\mathrm{cell} \cdot \mathrm{cell}^{-1} \cdot \mathrm{day}^{-1}$	5	[55]
δ_{My}	Death rate of the myeloid cells	day^{-1}	0.05	[55]
γ_{Fb}	Proliferation rate of the fibroblasts	day^{-1}	0.92	[176]
$g_{threshold_2}$	Threshold for fibroblasts activation	$\mathrm{pg} \cdot \mathrm{ml}^{-1}$	65	calibrated
δ_{Fb}	Death rate of the fibroblasts	day^{-1}	0.05	[177]
Π_f	Production of fibronectin by fibroblasts	$\mathrm{NE} \cdot \mathrm{cell}^{-1}\ \mathrm{mm}^3 \cdot \mathrm{day}^{-1}$	60	calibrated
$Endo$	endothelial cell density	$\mathrm{cell}\ \mathrm{mm}^{-3}$	1·e6	[178, 179, 180]
Π_{c_1}	S100A9 flux generated by lung cells	$\mathrm{pg} \cdot \mathrm{mL}^{-1} \cdot \mathrm{mm}^{-1} \cdot \mathrm{cell}^{-1} \cdot \mathrm{mm}^3$	3.55·e-6	calibrated
$g_{threshold_3}$	Threshold for endothelial and myeloid cells activation	$\mathrm{pg} \cdot \mathrm{mL}^{-1}$	56	calibrated
L	characteristic length of cell transmigration	mm	1	fixed
χ	Chemotaxis velocity	$\mathrm{mm} \cdot \mathrm{day}^{-1} \cdot \mathrm{pg} \cdot \mathrm{mL}^{-1} \cdot \mathrm{mm}^{-1}$	0.0012	[181]
$S_{\mathrm{capillaries}}$	Surface of the lung capillary walls	mm^2	4.9·e5	[179, 180]
V_{lungs}	Lungs volume	mm^3	700	measured
Π_M	Production of MMP-9 by myeloid cells	$\mathrm{NE} \cdot \mathrm{NE}^{-1} \cdot \mathrm{cell}^{-1}$	1·e-4	calibrated
δ_M	Clearance of MMP-9	day^{-1}	14	calibrated
κ_E	Consumption of ECM by MMPs	$\mathrm{NE}^{-1} \cdot \mathrm{day}^{-1}$	1	calibrated
$E_{\max}$	Maximal ECM density	NA	1·e4	calibrated
μ	Emission of CTCs by the primary tumor	$\mathrm{cell} \cdot \mathrm{mm}^{-3} \cdot \mathrm{day}^{-1}$	2·e3	calibrated
Π_{c_2}	SDF-1 flux generated by myeloid cells	$\mathrm{pg} \cdot \mathrm{mL}^{-1} \cdot \mathrm{mm}^{-1} \cdot \mathrm{cell}^{-1} \cdot \mathrm{mm}^3$	1.8·e-5	calibrated
δ_T	Death rate of B16 cells	day^{-1}	0.055	[165]
ν	Proliferation rate provided by myeloid cells	$\mathrm{day}^{-1} \cdot \mathrm{cell}^{-1} \cdot \mathrm{mm}^3$	1.5·e-6	calibrated
β	First parameter of the Gompertz	day^{-1}	0.06	calibrated on [172]
α	Second parameter of the Gompertz	day^{-1}	0.80·e5	calibrated on [172]
κ_B	consumption of stromal cells by the Metastases	$\mathrm{day}^{-1} \cdot \mathrm{mm}^{-3}$	0.0018	calibrated

Table 8.1: Parameter values used for the simulation. NE = Normalized expression

Depending on the patient stage (premetastatic, metastatic), is it more efficient to target the seed or the soil to prevent metastatic development and maximize the patient survival?
Such an issue could be addressed thanks to the model in simulating proliferation inhibition of either tumor cells or HPCs or fibroblast for example, at different stages of the disease.

Which molecular targets are suited to prevent metastatic development?
It should be interesting to simulate different pathways short-circuits among VEGF, MMPs, S100A8, SDF-1, VLA-4 at different times/stages to suggest revelant pathways to target.

All these issues will be addressed in future work.

Model improvements: chemokines diffusion

In the previous model, we modeled the dynamic over space and time of the concentration of chemokines S100A8 and SDF-1 by a 1-D diffusion equation with an influx which corresponds with the secretion by the cells. This influx is modeled by Neuman boundary conditions. We made the assumption that the diffusion of S100A8 and SDF-1 chemokines is much faster than their production by the cells and modeled this diffusion by stationary diffusion equations. However, It could be tested if this assumption is reasonable. Indeed, let us recall the 1D diffusion equation that was used in the model.

$$\begin{aligned}
&\frac{\partial c(t,x)}{\partial t} - D\frac{\partial^2 c(t,x)}{\partial x^2} = 0 \quad \forall t, x \in \left]0,T\right] \times \left]0,L\right],\\
&\frac{\partial c_1(t,L)}{\partial x} = f(t),\\
&c(t,0) = 0,\\
&c(0,x) = 0,
\end{aligned}$$

where c is the concentration of chemokines, D is the diffusion constant, L the size of the domain. The exact solution of this system can be written:

$$C(t,x) = -D\sqrt{\frac{2}{\pi}}\int_0^t \frac{f(s)}{\sqrt{2D(t-s)}}\exp\left(-\frac{x^2}{4D(t-s)}\right)\mathrm{d}s$$

The proof of this result is given in Materials and methods in Appendices.

For the time scale we are interested in and a relevant diffusion constant (see [182] for measured diffusion constant of SDF-1), this solution has to be compared to the stationary one. This will be done in a future work. If the stationary assumption is not reasonable, the non stationary solution should be included in the global model.

8.3 Conclusions and perspectives

In this part, we studied the biological literature of the premetastatic niche and from our understanding, we built a mathematical model describing the dynamics of the formation of this niche. We then coupled this model with a model of metastatic growth and after parameter calibration, assessed good agreement of the simulations with experimental data from [10]. The next step will be to use these coupled models to assess quantitative predictions on the impact of soil targeting on the metastatic outcome. Several quantitative questions have to be addressed with the model about the natural history of the premetastatic niche on a patient or the potentially relevant cellular and molecular targets.

Other modeling clues more generally concern the metastatic microenvironment. Indeed, in the literature, the role of the immune system remains unclear and seems even ambivalent. According to Dvorak et al., cancer is a wound that never heals [156] and immune cells can sustain an inflammatory state that promote cancer progression [29, 44]. Moreover, works on the premetastatic niche showed that bone marrow derived cells like hematopoietic progenitor and immune cells of the myeloid lineage like granulocytes promote metastasis in preparing the soil in the distant organ [10, 141]. However, in other studies, immune cells like granulocytes (neutrophils) have been shown to inhibit metastatic development [171]. The biological mechanisms by which immune cells either promote or inhibit metastatic seeding and growth are not yet elucidated. Why in some cases (cell line, host), a premetastatic niche is formed, promoting metastatic development, and in other ones immune reactions impair metastatic development? The premetastatic niche hypothesis suggests that the role of HPCs/myeloid cells is to attract and promote growth of tumor cells in the premetastatic niche. However, it has also been hypothesized that CD11b+ myeloid cells or myeloid derived suppressor cells could inhibit immune activity of CD8+ cytotoxic T cells and natural killers in forming an immune sanctuary [142]. In this case the role of the premetastatic niche would be more immunosuppressive rather than metastatic-promoting.

In our project, experiments have been performed to highlight the role of granulocytes in metastatic development of murine renal cell carcinoma (RENCA). The experiments consisted in a control experiment where RENCA cells were orthotopically injected in the renal subcapsule and another one where a depletion of granulocytes was performed on mice before the injection of RENCA cells. It appears that neither primary tumor nor metastatic burden changed significantly between control and treated mice, as presented in Fig 8.24.

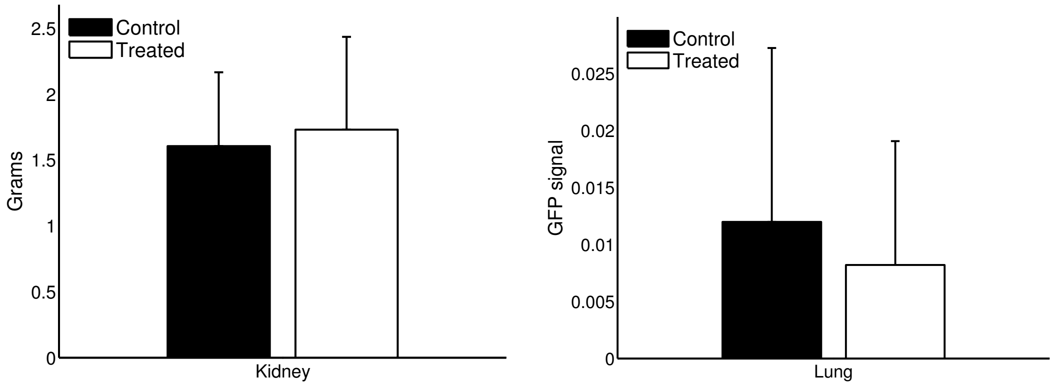


Figure 8.24:

Figure 8.25: Experimental results of an animal model of RENCA metastatic development with depletion of granulocytes. Left: kidney tumor weight at the final day (Mean±std). Right: tumor cells GFP expression in the lungs at the final day (Mean±std). Control: $N = 19$. Treated: $N = 20$

However, in the depletion experiment, a positive correlation between tumor size and metastatic burden was observed, whereas no correlation was observed in the presence of granulocytes, as shown in Fig 8.26. These results could suggest an effect of granulocytes on the metastatic burden. Indeed, in absence of granulocytes, the positive correlation between tumor and metastatic burden seemed to indicate that the larger the tumor is, the more metastases are disseminated. In contrast, in the presence of granulocytes, large tumors could induce high immune reactions, resulting in metastasis inhibition. The two pro-metastatic effect of the tumor-mediated dissemination and anti-metastatic effect of granulocytes could balance and result in an absence of correlation. It is of course a hypothesis and further experiments have to be made to confirm this tendency.

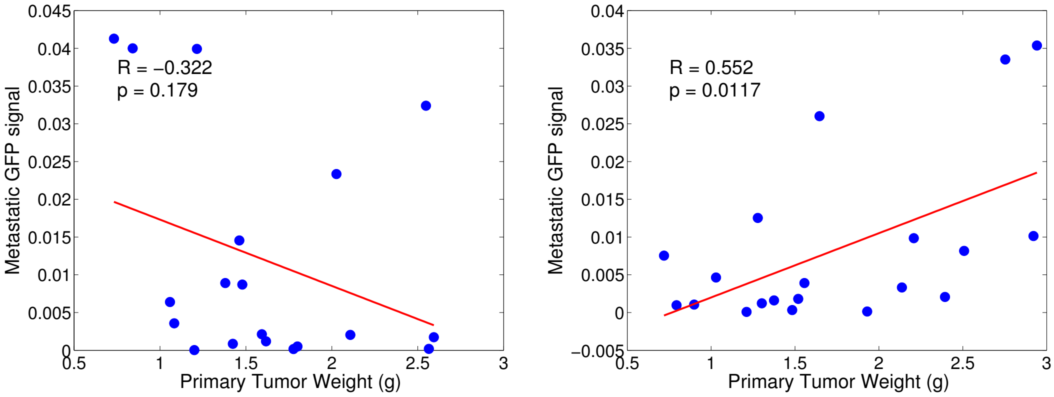


Figure 8.26: Experimental results of an animal model of RENCA metastatic development with depletion of granulocytes; Correlation between primary tumor and metastatic burden. Left: control. Right: granulocytes depletion.

In this context, it is difficult to define a general framework of the role of immune cells during premetastatic and metastatic development. In their paper showing the antimetastatic role of neutrophils, Granot et al. provide very rich data on the immune primary and metastatic dynamics [171]. Indeed, data on the metastatic size distribution are presented in this work, opening quantitative modeling perspectives. Using Iwata's model, a possible quantitative question addressable by the model could to assess whether granulocytes inhibit metastatic seeding or metastatic growth or both. As we previously saw, a mechanistic model for this kind of phenomenon is very hard to calibrate. That is why we should use a phenomenological way to model the dynamics and the effects of the immune system. An other modeling clue could be to build a model replacing immune system in a general framework where both antimetastatic and prometastatic roles of the immune system would be possible, mediated by one or several parameters. Data from different experiments and papers could be used to assess quantitative insights. The contribution of mathematical modeling could be useful, as it allows to infer insights from the dynamics that statistical analysis alone can not.

Chapter 9

Conclusions et perspectives

Le but de cette thèse était d'étudier divers aspects quantitatifs du processus métastatique à l'aide de modèles mathématiques. Pour cela, nous avons fait appel à différents types de modèles répondant à différentes problématiques : modèles EDP de populations structurées pour décrire la distribution en taille métastatique, modèles EDP spatiaux pour décrire les interactions spatiales entre tumeurs, systèmes d'EDO pour décrire les dynamiques cellulaires du microenvironment métastatique. Notre démarche a été de construire des modèles permettant d'inférer de l'information biologique qui n'était pas triviale dans les données.

9.1 Analyse mathématique et numérique d'un modèle spatial de croissance tumorale

9.1.1 Conclusions

Un modèle spatial de croissance tumorale a été introduit et étudié au cours de cette thèse. Il décrit l'expansion spatiale d'une tumeur à l'aide d'équations de conservations, ainsi que l'effet inhibiteur de la pression sur la prolifération des cellules. La dépendance non linéaire de la prolifération en la pression induit une équation elliptique non linéaire sur la pression. Une partie des travaux de cette thèse a été l'étude mathématique et numérique de cette équation. Nous avons montré l'existence d'une unique solution pour cette équation, ainsi qu'une certaine régularité de cette solution pour un domaine Lipschitz. Un algorithme de résolution a été proposé, étudié théoriquement et implémenté. Une convergence géométrique $H^{3/2}$ de cet algorithme a notamment été prouvée.
Par ailleurs, des tests numériques pour le système complet, à savoir l'équation elliptique précédente couplée avec le système d'équations de conservation sur la densité tumorale, ont été effectués. Au cours de cette étude numérique, on s'est d'abord placé dans un cas simple (taux de proliferation constant et géométrie sphérique) permettant d'exhiber une solution analytique et de tester la précision de différents schémas. Ainsi, le ratio précision sur temps de calcul a été comparé pour différents solveurs de transport. Après avoir choisi le solveur le plus performant, nous avons effectué des tests de convergence dans le cas où le taux de prolifération dépend de la pression via l'équation elliptique non linéaire citée plus haut.
Après s'être assuré de la convergence de nos schémas, nous avons présenté une méthode pour borner les paramètres du modèle basée sur une analyse du modèle et des considérations biologiques sur les paramètres. Une fois l'espace des paramètres borné, nous avons procédé a une assimilation de données de croissance métastatique pulmonaire chez la souris issues de l'imagerie. Le modèle semble capable de décrire les dynamiques volumiques des différentes métastases que nous avions segmentées. Une courte étude de l'évolution de la forme des métastases sous le modèle a montré que certaines formes étaient correctement prédites par le modèle mais que dans certains cas, les anisotropies de croissance observées n'étaient pas prises en compte par le modèle, menant à de grosses différences entre les formes prédites et observées. De plus, une comparaison de l'évolution de la masse entre une forme

circulaire et la forme segmentée a révélé que la forme semble avoir peu d'influence sur la dynamique de la masse, en tout cas sur les quelques cas étudiés.

9.1.2 Perspectives

Plusieurs perspectives peuvent être évoquées sur cette partie. Pour ce qui est de l'analyse mathématique du modèle, l'existence de solutions au problème couplé entre l'équation elliptique linéaire sur la pression et les équations de conservations pourrait être étudiée dans des travaux futurs en s'inspirant des travaux de Michel et al. [140]. Pour ce qui est du modèle, les capacités descriptives du modèle doivent être testées sur davantage de données. Une étude des capacités prédictives du modèles pourrait également être menée. Enfin l'effet de la forme de la lésion sur la dynamique de la masse doit être appronfondie sur davantage de cas.

9.2 Modélisation du processus métastatique et biologie quantitative

9.2.1 Conclusions

Lors de notre collaboration avec l'équipe de biologistes, nous nous sommes intéressés au développement métastatique du cancer du rein chez la souris. Notre approche de modélisation était de tester des théories quantitatives contre les données via la modélisation mathématique.
Nous avons confronté la théorie classique du processus métastatique à des données longitudinales en temps de la masse métastatique mais également à des données d'imagerie. Si cette théorie standard, acceptée dans de nombreuses références [20, 21, 22, 23, 24, 5], était capable de décrire la dynamique de la masse métastatique globale, elle s'avérait incapable de décrire le nombre et la taille des métastases, les métastases prédites étant trop petites et trop nombreuses.
Le motif métastatique observé, à savoir des métastases moins nombreuses mais plus massives, nous a incité à proposer plusieurs hypothèses biologiques pour expliquer les données. La première suggère que des métastases voisines pourraient, par expansion spatiale due à la prolifération, se rejoindre et fusionner, donnant naissance à une seule métastase plus massive. Un tel phénomène va dans le sens des distributions métastatiques observées, comme l'a illustré le modèle que nous avons proposé, basé sur la théorie standard mais incluant l'aspect spatial puisque les métastases croissent dans un organe limité en taille. Ce modèle permet notamment la fusion de métastases par mouvement passif, sans pour autant que ces métastases interagissent et perturbent leurs croissances mutuelles. Malgré une petite amélioration et la confirmation expérimentale que le phénomène de fusions de métastases survient au cours du développement métastatique (voir figure 6.5), l'hypothèse de fusion ne semble pas en mesure de décrire les distributions en tailles métastatiques observées.
Une deuxième hypothèse a été proposée pour expliquer le petit nombre et les tailles massives des métastases observées. Lorsqu'elles parviennent dans le parenchyme pulmonaire, les cellules tumorales pourraient être attirées, potentiellement par des facteurs chimioattractants vers des zones spécifiques et s'aggréger, formant des foyers dont la croissance serait gouvernée à la fois par la prolifération et par l'arrivée de nouvelles cellules. Afin de vérifier la validitité de cette hypothèse, un protocole expérimental a été proposé, où des injections de cellules tumorales marquées de différentes couleurs sont effectuées. Les premiers résultats de ces expériences montrent que des foyers métastatiques pourraient effectivement attirer ou recruter d'autres cellules tumorales. Ces résultats, qui doivent être renforcés par des expériences futures, font echo à d'autres travaux déjà menés sur le développement métastatique. Le *self-seeding*, décrit dans [11], est un phénomène au cours duquel les cellules métastatiques se détachent des métastases et retournent sur le site primaire pour s'y développer. Ce phénomène et celui que nous décrivons semblent analogues puisqu'il s'agit de recrutement ou d'attraction de cellules vers une tumeur déjà existante. Les mécanismes biologiques gouvernant le *self-seeding* ou l'attraction de cellules métastatique que suggèrent nos premières expériences restent très largement à définir mais il serait intéressant de voir si les mêmes mécanismes englobent les deux phénomènes. Par ailleurs, une étude menée par Aceto et al. a montré que les métastases ne se formaient pas toujours à partir

de cellules isolées mais parfois à partir de groupes de cellules tumorales [19]. Ils ont notamment montré que ces “clusters” de cellules tumorales ont un potentiel métastatique bien plus important que des cellules isolées. Le phénomène que nous proposons est un peu différent puisque les cellules se regrouperaient dans le parenchyme pulmonaire, alors que dans le phénomène décrit par Aceto et al., les cellules circulent déjà regroupées dans le sang. Cependant, les implications biologiques en termes de développement métastatique pourraient être similaires.

Ce travail illustre donc tout d’abord qu’une étude de biologie théorique peut être menée en confrontant des théories quantitatives à des données, en rejetant certaines théories et en en proposant d’autres qui soient testables expérimentalement ou quantitativement à l’aide de modèles mathématiques. De plus, les données quantitatives étant de plus en plus riches, une analyse de la dynamique des phénomènes biologiques devient possible à l’aide des modèles mathématiques, permettant ainsi d’inférer de nouvelles informations sur ces phénomènes. L’originalité de cette étude vient notamment de l’utilisation croisée de données longitudinales GFP, fournissant des informations sur toute la masse métastatique y compris la masse non visible, de données IRM, et d’un modèle mathématique décrivant la distribution en taille métastatique. Les résultats et les perspectives futures nous montrent que l’analyse de la dynamique métastatique peut permettre de mieux comprendre la biologie du développement métastatique. Les premiers résultats de cette étude, qui nécessitent l’appui de nouvelles expériences, offrent plusieurs perspectives biologiques et de modélisation.

9.2.2 Perspectives biologiques

Afin de confirmer le phénomène d’attraction de cellules tumorales par les foyers métastatiques, de nouvelles expériences doivent être menées. Tout d’abord, les deux protocoles proposés et déjà testés doivent être expérimentés sur plus de souris. Nous rappelons que ces deux protocoles consistent d’une part à l’injection orthotopique de cellules RENCA marquées en rouge et en vert dans des proportions égales, et d’autre part l’injection de cellules RENCA marquées en rouge suivie de l’injection en intraveineuse (veine de la queue) de cellules RENCA marquées en vert. Dans la première expérience, il serait intéressant de vérifier quantitativement si les plus grands foyers métastatiques sont les foyers attracteurs, c’est-à-dire multicolores avec un noyau d’une couleur et des cellules de l’autre couleur autour du noyau. Dans la deuxième expérience, il serait intéressant de pouvoir quantifier la proportion de cellules injectées par la veine de la queue qui forment des foyers indépendants par rapport à celles qui rejoignent des foyers déjà existants. En effet, ces quantifications permettraient de comprendre si l’attraction cellulaire est un phénomène majeur dans le développement métastatique et si elle contribue significativement aux lésions massives observées à l’IRM. Un autre protocole pourrait être testé afin de comprendre si les cellules tumorales circulantes sont réellement attirées par les foyers métastatiques. Une injection intraveineuse (veine de la queue) de cellules marquées en rouge pourrait d’abord être réalisée, suivie quelques jours plus tard d’une injection orthotopique (dans le rein) de cellules marquées en vert. Cela permettrait de mettre en évidence une véritable situation où les cellules tumorales circulantes sont émises par la tumeur primaire puis attirées par des foyers métastatiques. Des expériences *in vitro* pourraient également être menées pour confirmer le phénomène. En effet, des cellules métastatiques pourraient être extraites des métastases pulmonaires pour être mises en culture avec des cellules RENCA parentales ou issues de la tumeur primaire pour observer si ces cellules sont attirées par les cellules métastatiques. Si ce phénomène se confirme, des analyses plus spécifiques comme celle du transcriptome des poumons métastatiques devraient être effectuées pour identifier les facteurs responsables de cette attraction, offrant alors des perspectives cliniques potentielles en ciblant ce mécanisme d’action.

9.2.3 Perspectives de modélisation

Les phénomènes de fusions de métastases et d’attraction/aggrégation induisent potentiellement des différences dans le nombre et la taille des métastases, en comparaison avec la théorie standard où la distribution en tailles est uniquement gouvernée par la prolifération et le temps d’arrivée de chaque métastase. Ainsi, dans une perspective clinique où le but serait d’estimer, pour un patient donné, la masse métastatique occulte à partir des lésions visibles, la masse occulte pourrait être potentiellement très différente si de tels phénomènes d'aggrégation survenaient. C'est pourquoi une perspective de modélisation pourrait être d'incorporer les phénomènes d'aggrégation cellulaire et de métastases dans le modèle d'Iwata afin de prédire des distributions en taille métastatique plus réalistes. Dans un tel modèle, le phénomène de *self-seeding* pourrait naturellement être pris en compte puisque toutes les tumeurs (tumeur primaire et métastases) émettraient des cellules circulantes. Ces cellules pourraient alors former un nouveau foyer ou en rejoindre un existant.

Une autre piste consisterait à étendre ce modèle de dissémination métastatique à une description spatiale. Ceci s'avèrerait possible en utilisant des modèles basés sur des équations de conservation du type de ceux utilisés au cours ce cette thèse. En effet, on a mentionné précédemment dans ce manuscrit qu'il est possible d'obtenir un modèle spatial décrivant une dynamique Gompertz ou exponentielle lorsque les équations sont intégrées sur l'espace. Un modèle de dynamique métastatique spatiale consisterait donc à "faire apparaître" les métastases dans un organe segmenté en 3D (les poumons par exemple) avec un taux de dissémination égal au paramètre de dissémination calibré sur la masse métastatique, l'expansion spatiale de ces métastases étant ensuite gouvernée par les équations de conservation que l'on a mentionnées.

9.3 Modélisation des interactions spatiales entre métastases

9.3.1 Conclusions

L'intérêt que nous avons porté au phénomène de fusion de métastases ainsi que l'observation à l'IRM de métastases fusionnant effectivement nous ont poussé à étudier les interactions spatiales survenant lorsque deux foyers métastastiques sont en train de fusionner. Pour ce faire, nous avons introduit un nouveau modèle spatial de croissance tumorale décrivant l'expansion spatiale par des équations de conservation du type de celles utilisées dans [122, 78, 123] couplées avec une loi de prolifération dépendant de la pression environnementale. Cette loi, basée sur des études expérimentales montrant l'effet inhibiteur de la pression sur la prolifération des cellules tumorales, consiste en un taux de prolifération décroissant avec la pression. Après avoir calibré ce modèle sur des croissances de lésions métastatiques visibles à l'IRM, nous avons, à l'aide du modèle calibré, étudié quantitativement l'effet d'interactions de contact entre métastases sur leurs croissances respectives. Il s'avère, sous les hypothèses du modèle, que ces interactions inhibent de façon importante la prolifération de deux métastases en train de fusionner. En effet, on a estimé une perte de masse finale d'environ 30% par rapport à une situation où les métastases croissent indépendamment. De plus, cette différence s'accentue avec le nombre de métastases considérées. Ces résultats indiquent que les interactions spatiales entre métastases fusionnant ont un effet quantitatif important et un impact potentiellement non négligeable sur la dynamique métastatique globale, pour peu que quelques métastases se rencontrent pour fusionner. Enfin, des simulations ont été effectuées pour estimer le nombre de foyers métastatiques devant fusionner pour obtenir des métastases de la taille de celles observées à l'IRM dans le poumon de souris implantées avec une tumeur RENCA. Les simulations ont révélé que si certaines métastases semblaient pouvoir avoir été générées par la fusion de quelques foyers, d'autres nécessitaient un nombre de foyers trop important pour que cela soit réaliste. Cela appuie les résultats évoqués précédemment, rejetant l'hypothèse que la fusion de métastases par mouvement passif (dû à la prolifération) ait pu générer seule les distributions en tailles métastatiques observées à l'IRM.

9.3.2 Perspectives

L'effet inhibiteur des interactions de contact entre métastases doit encore être confirmé expérimentalement, par exemple en utilisant des marqueurs de prolifération (ki-67) pour observer si la prolifération de métastases fusionnant *in vivo* est moins importante dans la zone de contact entre les métastases. Un mouvement actif (attraction de cellules) est probablement impliqué dans le processus. Par la suite, le comportement 3D du modèle spatial introduit dans cette étude pourrait être étudié, puisque l'évacuation de la pression pourrait être différente, induisant potentiellement un comportement différent du modèle. De plus, les interactions spatiales entre métastases pourraient être prises en compte dans une description globale du processus métastatique.

9.4 Modélisation de la niche prémétastatique

Conclusions

Une dernière partie a été consacrée à la modélisation de la niche prémétastatique. Dans cette étude, après un état de l'art sur la littérature biologique de la niche prémétastatique, un modèle mathématique a été proposé pour décrire ce phénomène. Ce modèle mécanistique décrit la dynamique des principaux facteurs moléculaires et des principaux partenaires cellulaires participant à la formation de cette niche. Ce modèle a été couplé avec le modèle d'Iwata décrivant la dissémination et la colonisation métastatique, permettant d'étudier l'effet de changements dans le microenvironnement sur les métastases. Alors que certains paramètres ont été fixés à partir de valeurs tirées de la littérature, les autres paramètres ont été calibré sur les données de Kaplan et al. [10]. Le modèle est globalement capable de décrire les dynamiques observées et peut maintenant être utilisé pour répondre à des questions quantitatives cliniquement intéressantes.

Perspectives

Le modèle précédemment évoqué va maintenant pouvoir être utilisé pour étudier des questions quantitatives concernant l'effet du microenvironnement sur les métastases. La première de ces questions est la suivante : est-ce que la niche prémétastatique est un évènement précoce dans le cancer? Autrement dit, chez un patient pour lequel la présence d'une tumeur cancéreuse est diagnostiquée, est-ce que la niche prémétastatique est déjà prête à accueillir des métastases? Des questions sur la stratégie thérapeutique à adopter peuvent également être étudiées à l'aide du modèle. On peut se demander s'il est plus pertinent de cibler les cellules tumorales (seed) ou les cellules supportant leur attraction, leur survie, et leur prolifération au sein des niches prémétastatique et métastatique. Enfin, la question de la meilleure voie moléculaire à cibler peut également être étudier quantitativement.
Malgré tout, ce modèle mécanistique possède un grand nombre de paramètres à calibrer pour espérer des prédictions fiables. Même si l'identifiabilité du modèle n'a pas été étudiée, une approche plus phénoménologique est peut-être souhaitée pour étudier la dynamique. Un cadre général de biologie pourrait être modélisé, prenant en compte à la fois les effets pro et anti métastatiques du système immunitaire et du microenvironnement métastatique. Une première piste serait d'étudier les données de Granot et al. [171] consistant en des distributions en taille métastatique en présence ou absence de neutrophiles.

Appendices

Appendix A

Supplementary figures

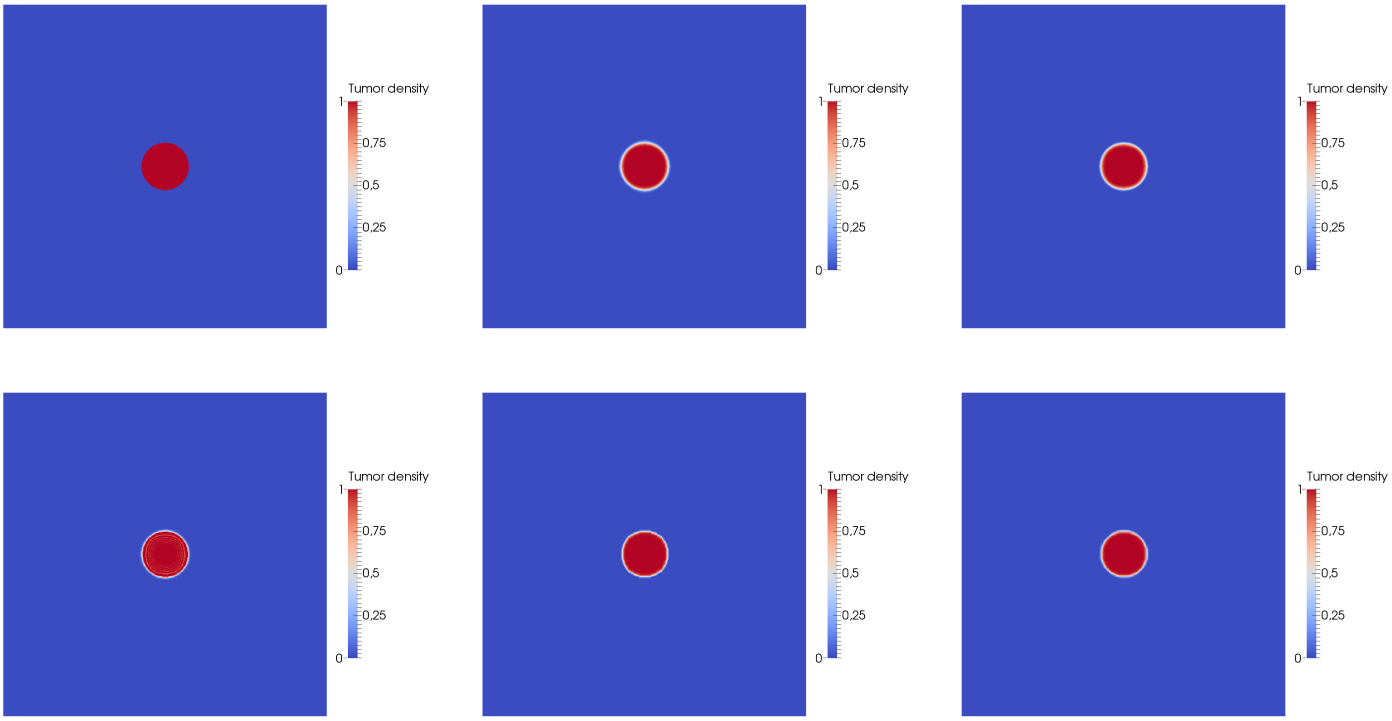


Figure A.1: Simulations from a spherical initial conditions with different transport schemes. Top: from left to right, the exact solution (on he mesh), First order Upwind scheme, second order semi Lagrangian scheme. Bottom: from left to right, scond order TVD Lax Wendroff scheme, fifth order WENO scheme, fifht order TVD WENO scheme. Simulations were obtained using Eqs (5.5)-(5.8) with a constant growth rate is constant: $\gamma_0 = 0.67$ day^{-1}, and the porosity k is equal to 1 in the included (in the square domain) circle, and equal to 1000 outside. The final time is $T = 1$ day.

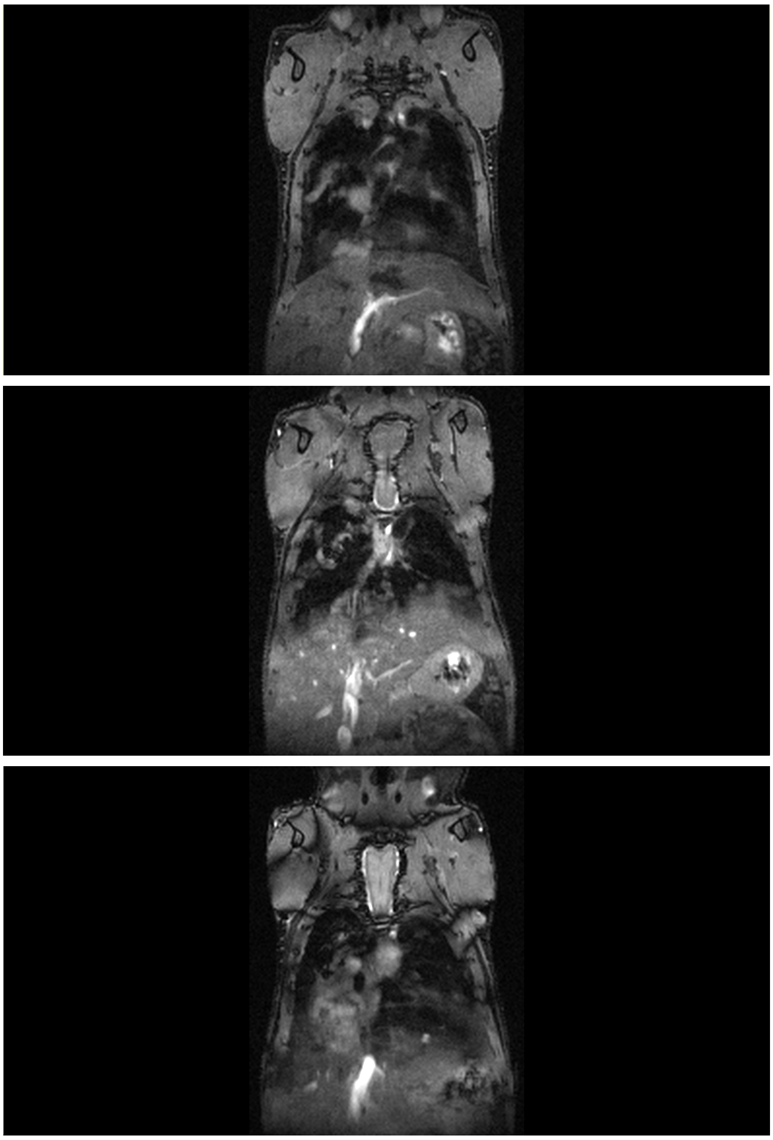
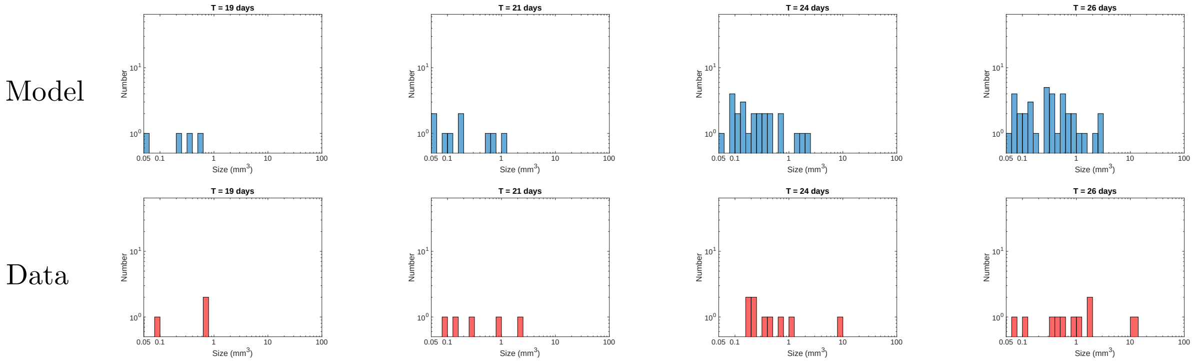


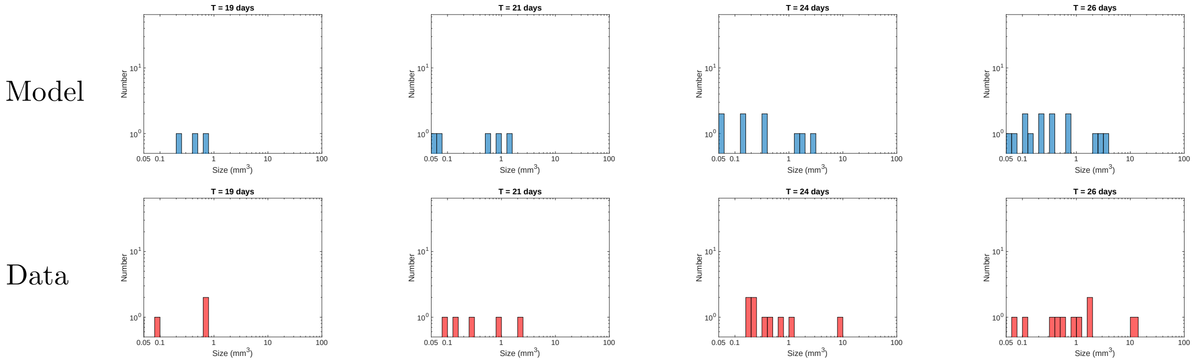
Figure A.2: Representative MR images where size distribution of metastases could not be satisfactorily assessed. Coronal slices of three mice. Top: Day 19; Middle and bottom: Day 21. The metastatic foci could not be clearly segmented because the metastatic burden was very diffuse.

$V_0 = 10$ cells



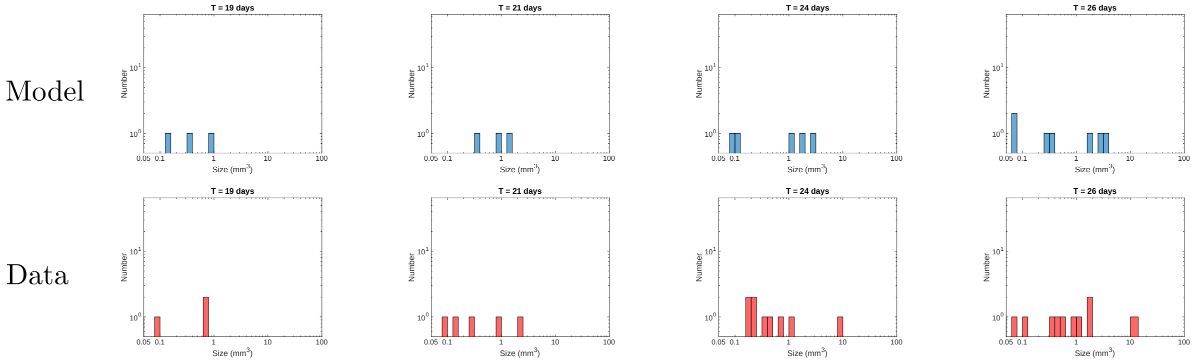
T	19		21		24		26	
	Data	Model	Data	Model	Data	Model	Data	Model
Number	4	4	5	9	9	24	11	36
Macro-burden (mm^3)	1.58	1.13	3.55	3	11.5	10.5	21.2	20.8
Size of largest met (mm^3)	0.78	0.526	2.19	1.05	8.17	2.19	13.6	3.11

$V_0 = 100$ cells



T	19		21		24		26	
	Data	Model	Data	Model	Data	Model	Data	Model
Number	4	3	5	5	9	9	11	14
Macro-burden (mm^3)	1.58	1.37	3.55	2.88	11.5	7.01	21.2	11.7
Size of largest met (mm^3)	0.78	0.758	2.19	1.38	8.17	2.64	13.6	3.58

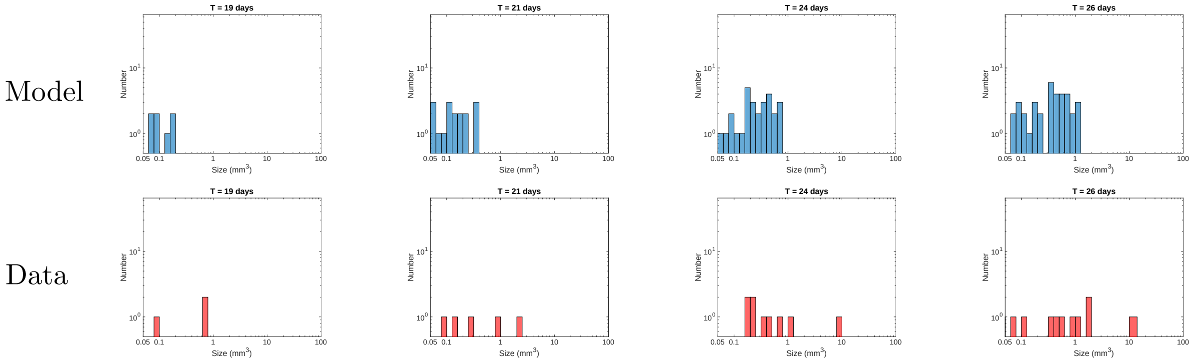
$V_0 = 500$ cells



T	19		21		24		26	
	Data	Model	Data	Model	Data	Model	Data	Model
Number	4	3	5	5	9	9	11	14
Macro-burden (mm^3)	1.58	1.37	3.55	2.88	11.5	7.01	21.2	11.7
Size of largest met (mm^3)	0.78	0.758	2.19	1.38	8.17	2.64	13.6	3.58

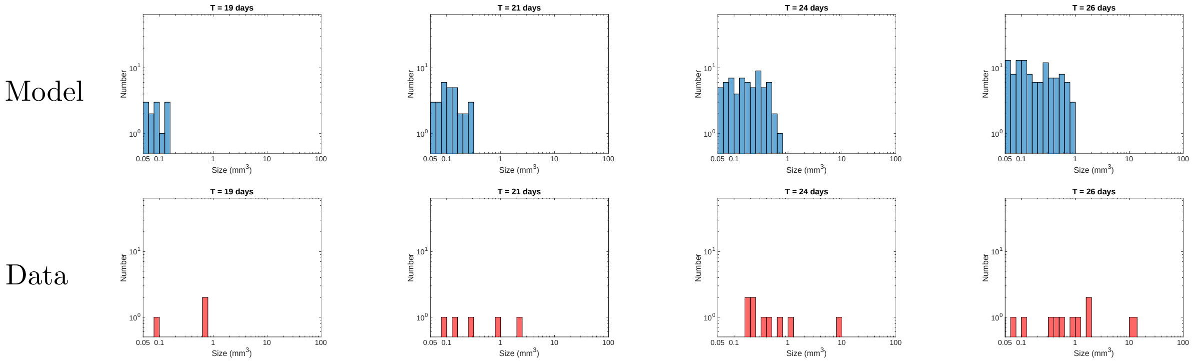
Figure A.3: Time course of the macro-metastases size distribution for different initial metastatic sizes Top to down: simulation of the mathematical formalism of the standard theory (i.e. dissemination and independent growth of the resulting tumor foci), using the parameter values inferred from the total metastatic burden data (total GFP signal in the lungs) using four different initial numbers of initiating metastatic cells. The results are compared to observations of macro-metastases numbers and sizes in one mouse on MRI data.

$\gamma = 0.1$



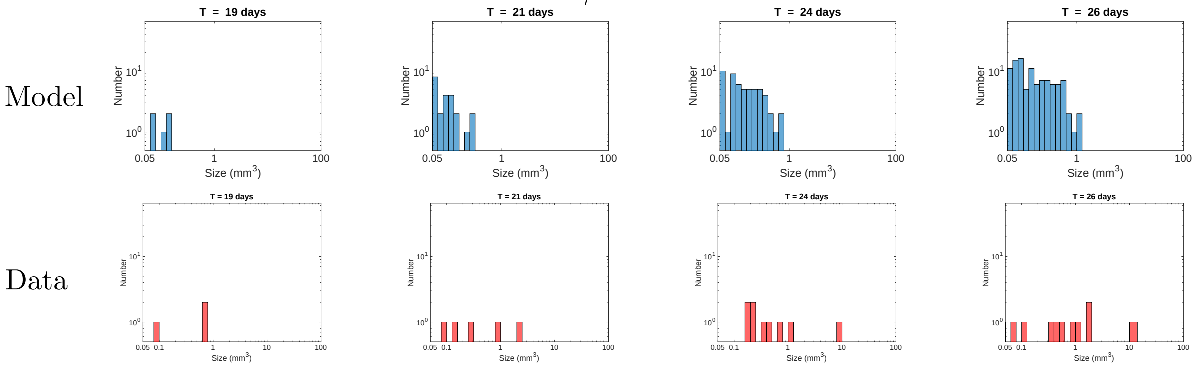
T	19		21		24		26	
	Data	Model	Data	Model	Data	Model	Data	Model
Number	4	4	5	9	9	24	11	36
Macro-burden (mm^3)	1.58	1.13	3.55	3	11.5	10.5	21.2	20.8
Size of largest met (mm^3)	0.78	0.526	2.19	1.05	8.17	2.19	13.6	3.11

$\gamma = \frac{2}{3}$



T	19		21		24		26	
	Data	Model	Data	Model	Data	Model	Data	Model
Number	4	3	5	5	9	9	11	14
Macro-burden (mm^3)	1.58	1.37	3.55	2.88	11.5	7.01	21.2	11.7
Size of largest met (mm^3)	0.78	0.758	2.19	1.38	8.17	2.64	13.6	3.58

$\gamma = 1$



T	19		21		24		26	
	Data	Model	Data	Model	Data	Model	Data	Model
Number	4	3	5	5	9	9	11	14
Macro-burden (mm^3)	1.58	1.37	3.55	2.88	11.5	7.01	21.2	11.7
Size of largest met (mm^3)	0.78	0.758	2.19	1.38	8.17	2.64	13.6	3.58

Figure A.4: **Time course of the macro-metastases size distribution for different values of** γ. The fit analysis of the GFP data was re-performed for values of γ ranging from 0.1 to 1, generating each time new distributions of the parameters α, β and μ, and simulations equivalent to Fig 6.4 were re-performed for the median values of parameters (inter-animal variability not shown here). Results only for $\gamma = 0.1$, $\frac{2}{3}$ and 1 are shown here. Qualitatively similar results are observed concerning the size distribution metrics (in particular, number of metastases and size of the largest lesion).

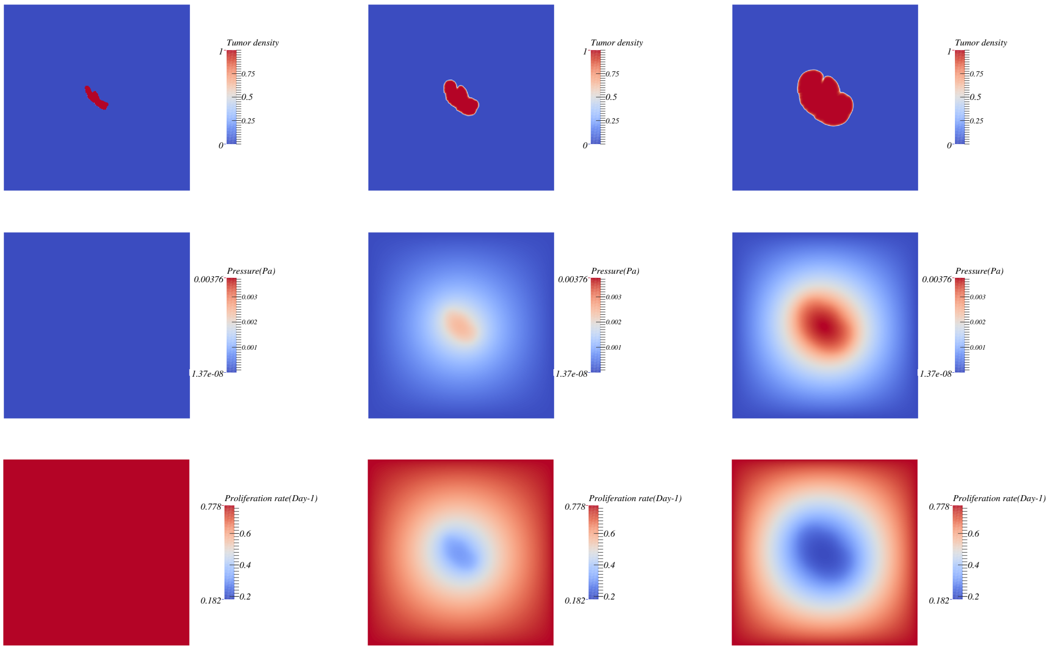


Figure A.5: The simulated growth by the model using the fitted parameters and starting from the real shape of the observed metastasis at day 19 on the coronal MRI slice. Time course of the tumor density (up), pressure (middle), and proliferation rate fields. From left to right: day 0, day 3 and day 7. Simulations were obtained using Eqs (7.1)-(7.4) and the following parameter values: $\gamma_0 = 0.78$ day^{-1}; $\Pi_c = 0.0026$ Pa; time of simulation: $T = 7$ days.

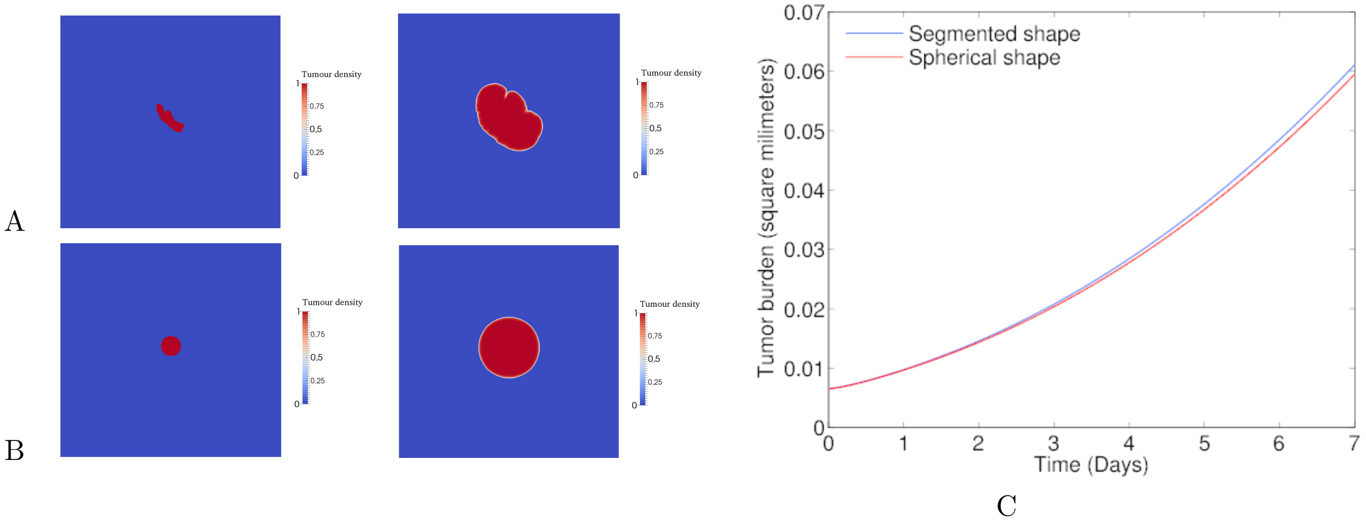


Figure A.6: Spherical and non-spherical shapes. (A) Simulation from the segmented shape. Simulations were obtained using Eqs (7.1)-(7.4) and the following parameter values: $\gamma_0 = 0.78$ day^{-1}; $\Pi_c = 0.0026$ Pa; time of simulation: $T = 7$ days. (B) Simulation with the same parameters and same initial burden from a spherical shape. (C) Volume dynamics of the two simulations. The final relative difference is 2.5%.

Appendix B

Materials and methods

Ethics statement

Ethical approval for all animal studies was obtained from the Institutional Animal Care and Use Committee of the INSERM Institute in accordance with the National Advisory Committee for Laboratory Animal Research Guidelines licensed by the French Authority. Animal facility: Animalerie mutualisée de Bordeaux 1, authorisation number: B33-522, Date: February 8^{th},2012. Investigator: Andreas Bikfalvi (authorisation number: R-45GRETA-F1-10)

Cell line and mouse experiments

These experiments were designed by Andreas Bikfalvi, Lindsay S. Cooley and Wilfried Souleyreau and were performed by Lindsay S. Cooley, Raphäel Pineau and Wilfried Souleyreau.

RENCA-GFP cells

The mouse renal adenocarcinoma cell line RENCA was cultured in RPMI media (Gibco) supplemented with 10% foetal calf serum, 1mM sodium pyruvate, 2mM glutamine, 100U/ml de penicillin and 100μg/ml streptomycin, at 37oC/5% CO2. RENCA-GFP cells were produced via infection of RENCA cells with a GFP lentivirus, a gift of Dr. C. Grosset (u889 Bordeaux).

Orthotopic implantation of RENCA-GFP cells

RENCA-GFP were cultured in exponen- tial growth phase, and harvested by trypsinisation (Gibco). After washing in basal RPMI media, the cells were counted and concentration adjusted to 100000 cells per 25μl in basal media. 25μ of cell suspension was then injected underneath the renal capsule of the left kidney of female Balb/c mice aged 6 weeks.

Tissue harvest

Mice were sacrificed at the specified intervals and the left kidney (bearing the primary tumour) and lungs were dissected and snap frozen in liquid nitrogen. RNA was extracted using TRIzol Reagent (Life Technologies) as per the manufacturer's protocol.

Reverse transcription and Q-RT-PCR

RNA samples were quantified using a nanodrop ND-1000 spectrophotometer (Nanodrop Technologies). 1μg of total RNA was reverse transcribed to cDNA using High Capacity cDNA Reverse Transcription Kit (Applied Biosystems). Real-Time PCR was carried out using the Step One Plus Real-Time PCR system. Reactions were carried out in a total volume of 20μl containing 2ng of cDNA, Power SYBR Green PCR Matsermix (Applied Biosystems), and 200nM of each of the forward and reverse primers. The reaction conditions were as follows: 10 mins at 95oC followed by 40 cycles of 15 secs at 95oC and 1 min at 60oC. Data were analysed using Step One Software v2.3. The housekeeping gene PPIA was used as an endogenous control to nomalize for differences in the amount of total RNA in each sample. Expression of GFP is thus presented as an N-fold difference relative to the total RNA per sample. The sequences of the primer used was as follows: eGFP: Forward primer 5'-CGACCACTACCAGCAGAACA-3' Reverse primer: 5'GAACTCCAGCAGGACCATGT-3'.

Mouse model of lung metastasis from kidney primary tumor

In the first set of experiments, $GFP+$ RENCA cells orthotopic injection (into the kidney) was performed to observe metastatic $GFP+$ cells in the lungs. Metastases were observed in the lungs only, not in the liver.
The first $GFP+$ metastatic cells were observed at day 14, and the first macrometastases at day 18. Another sample of mice has been injected in the same conditions and a GFP analysis has been performed on it, to follow primary tumor and metastatic dynamics. Three mice have been sacrificed at each time point to mesure the gfp signal in the kidney and in the lungs. Moreover, a last sample of mice has been injected in the same conditions to be imaged with MRI scheduling. First metastases were observed at day 18 with sizes comprised between 0.05 and 12 cubic millimeters at day 19.

MR materials and methods

The imaging process was performed by Emeline Ribot.

MR materials

The experiments were carried out on a horizontal 7T magnet (Bruker Biospec 70/20, Germany), equipped with a 12 cm gradient insert capable of 660 mT/m maximum strength and 110 μs rise time. Lung imaging was performed using a quadrature emission/reception birdcage coil (inner diameter: 2.5 cm, 5 cm length).
Mice were anesthetized with 1.5% isoflurane in air during the imaging session. Mouse respiration was monitored during the entire experiment by using an air balloon placed on the abdomen (SA Instruments).

MR sequence

A 3D water-selective balanced Steady State Free Precession sequence was used [183]. This sequence induces a T2-like contrast, allowing to detect metastases with hyper-intense signals without injecting contrast agents [184]. The main parameters of the sequence were as follow: FOV=30×22×22; matrix=192×142×142; TE/TR=3.1/6.2ms; flip $angle = 20\,\text{deg}$; reception bandwidth= 178kHz; number of excitation=1; acquisition time=2min3s. In order to suppress banding artifacts inherent to this sequence, the sum-of-square method was performed [185]. Thus, the total acquisition time was 8min12s.

Numerical methods

Scientific computing library

The numerical solvers used to solve the conservation equations (5.5)-(5.9) were coded in the C++ librairy Cadmos, which is the is the library used by the INRIA team Monc to solve partial derivative equations. The library, coded by Olivier Saut, is object oriented and contains also routines to read parameters in file, deal with parameter maps, cartesian grids and convert dicom into VTK files. The numerical solvers are finite differences based and deal with cartesian grids. The library used to solve linear systems is eigen.

Numerical methods to simulate the partial derivative equations model of spatial tumor growth

The system of partial derivative equations that models the spatial metastatic growth was solved with the following numerical methods:

- a Strang splitting method for the time scheme, as described in 5.3.1.
- a fifth order Weno finite differences scheme for the spatial resolution in the chapter 7.
- a fixed point method to solve the nonlinear equation on the pressure (5.26). At each step of the fixed-point algorithm, the equation (5.33) was numerically solved thanks to a finite differences method leading to a linear system that was solved thanks to eigen library.

Numerical methods for the ordinary differential equations model of tumor growth

In order to fit the models of the section 5.1 to tumor growth data, a Monte-Carlo method followed by a gradient descent were performed, in order to search first for a global optimum and then to converge toward the local optimum.

Numerical methods for the premetastatic niche model

The ODE system describing the dynamics of cellular and molecular species was solved thanks to a forward Euler method.
In order to solve partial derivative equation (8.6) of Iwata's model, the Lagrangian approach, based on the method of characteristics, that was proposed by Benzekry has been adopted [186, 124].

Solution of the 1D stationary diffusion equation

Searching for a solution

Let us define the following problem :

$$\frac{\partial C(x,t)}{\partial t} - D\frac{\partial^2 C(x,t)}{\partial x^2} = 0 \quad \forall (x,t) \in \mathbb{R}^{+*} \times \mathbb{R}^{+*} \tag{B.1}$$

$$\frac{\partial C(0,t)}{\partial x} = f(t) \quad \forall t \in \mathbb{R}^{+*} \tag{B.2}$$

$$C(x,0) = 0 \quad \forall x \in \mathbb{R}^{+} \tag{B.3}$$

$$f(0) = 0 \tag{B.4}$$

We are searching for a regular solution in $\mathbb{R}^+$. To achieve it, we assume that such a solution C exists and we define the following function:

$$v(t,x) = C(t,x) + f(t)\exp(-x).$$

This function is solution of the following problem :

$$\begin{aligned}
&\frac{\partial v(t,x)}{\partial t} - D\frac{\partial^2 v(t,x)}{\partial x^2} = \left(\frac{\mathrm{d}f(t)}{\mathrm{d}t} - Df(t)\right)\exp(-x) \quad \forall (x,t) \in \mathbb{R}^{+*} \times \mathbb{R}^{+*}\\
&\frac{\partial v(t,0)}{\partial x} = 0 \quad \forall t \in \mathbb{R}^{+*}\\
&v(0,x) = 0 \quad \forall x \in \mathbb{R}^{+}\\
&f(0) = 0
\end{aligned}$$

We define now :

$$w(t,x) = \begin{cases} v(t,x) & \text{if} \quad x \geq 0 \\ v(t,-x) & \text{if} \quad x \leq 0 \end{cases}$$

Such a function is a solution of the following problem :

$$\begin{aligned} &\frac{\partial w(t,x)}{\partial t} - D\frac{\partial^2 w(t,x)}{\partial x^2} = \left(\frac{\mathrm{d}f(t)}{\mathrm{d}t} - Df(t)\right)\exp(-|x|) \quad \forall (x,t) \in \mathbb{R}\times\mathbb{R}^{+*} \\ &w(0,x) = 0 \quad \forall x \in \mathbb{R} \\ &f(0) = 0 \end{aligned} \tag{B.5}$$

We have now a problem which is posed on all $\mathbb{R}$. The goal is to apply the Fourier Transform at this partial differential equation.

We define the Fourier Transform on $L^2(\mathbb{R})$ as follows:

$$\hat{u}(\xi) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} u(x)\exp\left(-ix\xi\right)\mathrm{d}x$$

To define the Fourier transform of $\frac{\partial^2 w}{\partial x^2}$, we need to have $\frac{\partial^2 w}{\partial x^2}$ being in $L^2(\mathbb{R})$ over x. Thus, w must be $H^2(\mathbb{R})$ over x. Moreover, to define the Fourier transform of $\frac{\partial w}{\partial t}$, it is necessary and sufficient that there exists $g \in L^2(\mathbb{R})$ such that $\frac{\partial w}{\partial t} \leq g$, where g does not depend on t. The function $\exp(-|x|)$ is $L^2(\mathbb{R})$ and its Fourier Transform is $\sqrt{\frac{2}{\pi}}\frac{1}{1+\xi^2}$.

The problem (B.5) becomes, after Fourier Transform, an ordinary differential equation :

$$\begin{aligned} &\frac{\partial \hat{w}(t,\xi)}{\partial t} + D\xi^2\hat{w}(t,\xi) = \sqrt{\frac{2}{\pi}}\left(\frac{\frac{\mathrm{d}f(t)}{\mathrm{d}t} - Df(t)}{1+\xi^2}\right) \\ &\hat{w}(0,\xi) = 0 \end{aligned}$$

We are searching for a solution of the following form:

$$\hat{w}(t,\xi) = k(t)\exp(-D\xi^2 t)$$

By injecting this solution in the equation, we find :

$$\frac{\mathrm{d}k}{\mathrm{d}t}\exp(-D\xi^2 t) = \sqrt{\frac{2}{\pi}}\left(\frac{\frac{\mathrm{d}f(t)}{\mathrm{d}t} - Df(t)}{1+\xi^2}\right)$$

To respect the initial condition, k has to verify $k(0) = 0$. Thus we have :

$$k(t) = \sqrt{\frac{2}{\pi}}\int_0^t \frac{\frac{\mathrm{d}f(t)}{\mathrm{d}t} - Df(t)}{1+\xi^2}\exp(D\xi^2 s)\mathrm{d}s$$

Therefore, it comes the following expression for $\hat{w}$:

$$\hat{w}(t,\xi) = \sqrt{\frac{2}{\pi}}\int_0^t \frac{\frac{\mathrm{d}f(t)}{\mathrm{d}t} - Df(t)}{1+\xi^2}\exp(D\xi^2(s-t))\mathrm{d}s$$

After having integrated it by parts, we finally obtain the following expression:

$$\hat{w}(t,\xi) = \sqrt{\frac{2}{\pi}}\left(\frac{f(t)}{1+\xi^2} - D\int_0^t f(s)\exp(-D\xi^2(t-s))\mathrm{d}s\right) \tag{B.6}$$

The Fourier transform of a Gaussian function is known as:

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\exp\left(-ix\xi\right)\exp\left(-\alpha x^2\right)\mathrm{d}x = \sqrt{\frac{1}{2\alpha}}\exp\left(-\frac{\xi^2}{4\alpha}\right)$$

In our case, $\frac{1}{4\alpha} = D(t-s)$. One can thus apply the inverse Fourier transform to the expression (B.6) so as to obtain :

$$w(t,x) = f(t)\exp(-|x|) - D\sqrt{\frac{2}{\pi}}\int_0^t \frac{f(s)}{\sqrt{2D(t-s)}}\exp\left(-\frac{x^2}{4D(t-s)}\right)\mathrm{d}s \tag{B.7}$$

We can deduce the expression of v defined on $\mathbb{R}^+ \times \mathbb{R}^+$:

$$v(t,x) = f(t)\exp(-x) - D\sqrt{\frac{2}{\pi}}\int_0^t \frac{f(s)}{\sqrt{2D(t-s)}}\exp\left(-\frac{x^2}{4D(t-s)}\right)\mathrm{d}s$$

Then C is determined as follows : $C(t,x) = v(t,x) - f(t)\exp(-x)$. We finally get:

$$C(t,x) = -D\sqrt{\frac{2}{\pi}}\int_0^t \frac{f(s)}{\sqrt{2D(t-s)}}\exp\left(-\frac{x^2}{4D(t-s)}\right)\mathrm{d}s$$

Checking that the solution is satisfied together with the boundary condition

Computations show that this function verifies the diffusion equation (B.2). But does it verify the Neuman boundary condition? *A priori*, we cannot differentiate the expression with respect to x at 0 under the integral sign.

We propose an heuristic proof to show that the Neuman boundary condition is met in the case where f is a constant. Let remember the expression of the function :

$$C(t,x) = -D\sqrt{\frac{2}{\pi}}\int_0^t \frac{f(s)}{\sqrt{2D(t-s)}}\exp\left(-\frac{x^2}{4D(t-s)}\right)\mathrm{d}s$$

We pose the change of variable $\tau = t - s$ and then we have:

$$C(t,x) = -D\sqrt{\frac{2}{\pi}}\int_0^t \frac{f(\tau)}{\sqrt{2D\tau}}\exp\left(-\frac{x^2}{4D\tau}\right)\mathrm{d}\tau$$

The function $g : (\tau, x) \mapsto \frac{f(\tau)}{\sqrt{2D\tau}}\exp\left(-\frac{x^2}{4D\tau}\right)$ defined on $\mathbb{R}^{+*} \times \mathbb{R}^+$ is uniformly dominated by the function $t \mapsto \frac{f(\tau)}{\sqrt{2D\tau}}$ which is summable on $]0;t[$. The derivative exists and is defined as follows:

$$\frac{\partial C(t,x)}{\partial x} = -D\sqrt{\frac{2}{\pi}}\int_0^t \frac{\partial g(\tau,x)}{\partial x}\mathrm{d}\tau,$$

which gives, for all $(t,x) \in \mathbb{R}^+ \times \mathbb{R}^+$:

$$\frac{\partial C(t,x)}{\partial x} = D\sqrt{\frac{2}{\pi}}\int_0^t \frac{2xf(\tau)}{4D\tau\sqrt{2D\tau}}\exp\left(-\frac{x^2}{4D\tau}\right)\mathrm{d}\tau$$

Let us define the change of variable $y = \frac{x^2}{\tau}$. The previous integral becomes :

$$\begin{aligned}\frac{\partial C(t,x)}{\partial x} &= \frac{f}{2D\sqrt{\pi}}\int_0^t \frac{x^2}{\tau^2}\frac{\sqrt{\tau}}{x}\exp\left(-\frac{x^2}{4D\tau}\right)\mathrm{d}\tau \\ &= \frac{f}{2D\sqrt{\pi}}\int_{\frac{x^2}{t}}^{+\infty}\frac{\exp\left(-\frac{y}{4D}\right)}{\sqrt{y}}\mathrm{d}\tau\end{aligned}$$

Another change of variable $s = \sqrt{y}$ gives :

$$\frac{\partial C(t,x)}{\partial x} = \frac{f}{D\sqrt{\pi}}\int_{\frac{x}{t}}^{+\infty}\exp\left(-\frac{y^2}{4D}\right)\mathrm{d}s$$

The limit of this function when tends to 0 is a Gauss integral, which has the known value:

$$\begin{aligned}
\frac{\partial C(t,0)}{\partial x} &= \frac{f}{D\sqrt{\pi}} \int_0^{+\infty} \exp\left(-\frac{y^2}{4D}\right) \mathrm{d}s \\
&= \frac{f}{D\sqrt{\pi}} \frac{\sqrt{4D\pi}}{2} \\
&= f
\end{aligned}$$

Bibliography

[1] Etienne Baratchart, Sébastien Benzekry, Andreas Bikfalvi, Thierry Colin, Lindsay S. Cooley, Raphäel Pineau, Emeline J Ribot, Olivier Saut, and Wilfried Souleyreau. Computational modelling of metastasis development in renal cell carcinoma. *PLoS Comput Biol*, 11(11):e1004626, 11 2015.

[2] American Cancer Society. Cancer facts and figures 2015, September 2015. http://www.cancer.org/acs/groups/content/@editorial/documents/document/acspc-044552.pdf.

[3] World Health Organization. Cancer, key facts, September 2015. http://www.who.int/mediacentre/factsheets/fs297/en/.

[4] Douglas Hanahan and Robert A Weinberg. Hallmarks of cancer: the next generation. *Cell*, 144(5):646–674, March 2011.

[5] Gaorav P Gupta and Joan Massagué. Cancer metastasis: building a framework. *Cell*, 127(4):679–695, November 2006.

[6] Fondation ARC pour la recherche sur le cancer. Le cancer en chiffres, October 2015. http://www.fondation-arc.org/Face-au-cancer/le-cancer-en-chiffres.html#.

[7] Fondation ARC pour la recherche sur le cancer. La survie des cancers progresse en france, October 2015. http://www.fondation-arc.org/Actualites/la-survie-des-cancers-progresse-en-france.html.

[8] Don X Nguyen, Paula D Bos, and Joan Massagué. Metastasis: from dissemination to organ-specific colonization. *Nat Rev Cancer*, 9(4):274–284, April 2009.

[9] Scott Valastyan and Robert A Weinberg. Tumor metastasis: molecular insights and evolving paradigms. *Cell*, 147(2):275–92, October 2011.

[10] Rosandra N Kaplan, Rebecca D Riba, Stergios Zacharoulis, Anna H Bramley, Loïc Vincent, Carla Costa, Daniel D MacDonald, David K Jin, Koji Shido, Scott A Kerns, Zhenping Zhu, Daniel Hicklin, Yan Wu, Jeffrey L Port, Nasser Altorki, Elisa R Port, Davide Ruggero, Sergey V Shmelkov, Kristian K Jensen, Shahin Rafii, and David Lyden. VEGFR1-positive haematopoietic bone marrow progenitors initiate the pre-metastatic niche. *Nature*, 438(7069):820–7, December 2005.

[11] Mi-Young Kim, Thordur Oskarsson, Swarnali Acharyya, Don X Nguyen, Xiang H F Zhang, Larry Norton, and Joan Massagué. Tumor self-seeding by circulating cancer cells. *Cell*, 139(7):1315–26, December 2009.

[12] D Ambrosi, L Preziosi, and Mathematical Models. On the closure of mass balance models for tumor growth. *Math. Models Method A Appl. Sci.*, 12(5):737–754, 2002.

[13] Thierry Colin, François Cornelis, Julien Jouganous, Jean Palussière, and Olivier Saut. Patient-specific simulation of tumor growth, response to the treatment, and relapse of a lung metastasis: a clinical case. *J Comput Surg*, 2(1):841–17, February 2015.

[14] Fabien Montel, Morgan Delarue, Jens Elgeti, Danijela Vignjevic, Giovanni Cappello, and Jacques Prost. Isotropic stress reduces cell proliferation in tumor spheroids. *New J Phys*, 14(5):055008, May 2012.

[15] Triantafyllos Stylianopoulos, John D Martin, Vikash P Chauhan, Saloni R Jain, Benjamin Diop-Frimpong, Nabeel Bardeesy, Barbara L Smith, Cristina R Ferrone, Francis J Hornicek, Yves Boucher, et al. Causes, consequences, and remedies for growth-induced solid stress in murine and human tumors. *Proc Natl Acad Sci*, 109(38):15101–15108, 2012.

[16] Triantafyllos Stylianopoulos, John D Martin, Matija Snuderl, Fotios Mpekris, Saloni R Jain, and Rakesh K Jain. Coevolution of solid stress and interstitial fluid pressure in tumors during progression: implications for vascular collapse. *Cancer Res*, 73(13):3833–3841, July 2013.

[17] K Iwata, K Kawasaki, and N Shigesada. A Dynamical Model for the Growth and Size Distribution of Multiple Metastatic Tumors. *J Theor Biol*, 203(2):177–186, March 2000.

[18] L A Liotta, M G Saidel, and J Kleinerman. The significance of hematogenous tumor cell clumps in the metastatic process. *Cancer Res*, 36(3):889–894, March 1976.

[19] Nicola Aceto, Aditya Bardia, David T Miyamoto, Maria C Donaldson, Ben S Wittner, Joel A Spencer, Min Yu, Adam Pely, Amanda Engstrom, Huili Zhu, Brian W Brannigan, Ravi Kapur, Shannon L Stott, Toshi Shioda, Sridhar Ramaswamy, David T Ting, Charles P Lin, Mehmet Toner, Daniel A Haber, and Shyamala Maheswaran. Circulating Tumor Cell Clusters Are Oligoclonal Precursors of Breast Cancer Metastasis. *Cell*, 158(5):1110–1122, August 2014.

[20] James E Talmadge and Isaiah J Fidler. AACR centennial series: the biology of cancer metastasis: historical perspective. *Cancer Res*, 70(14):5649–69, July 2010.

[21] Isaiah J Fidler and S Paget. The pathogenesis of cancer metastasis: the 'seed and soil' hypothesis revisited. *Nat Rev Cancer*, 3(6):453–458, 2003.

[22] J E Talmadge, S R Wolman, and I J Fidler. Evidence for the clonal origin of spontaneous metastases. *Science*, 217(4557):361–363, July 1982.

[23] I J Fidler and J E Talmadge. Evidence that intravenously derived murine pulmonary melanoma metastases can originate from the expansion of a single tumor cell. *Cancer Res*, 46(10):5167–5171, October 1986.

[24] Christoph A Klein. Parallel progression of primary tumours and metastases. *Nat Rev Cancer*, 9(4):302–312, April 2009.

[25] Niklas Hartung, Séverine Mollard, Dominique Barbolosi, Assia Benabdallah, Guillemette Chapuisat, Gérard Henry, Sarah Giacometti, Athanassios Iliadis, Joseph Ciccolini, Christian Faivre, and Florence Hubert. Mathematical modeling of tumor growth and metastatic spreading: validation in tumor-bearing mice. *Cancer Res*, 74(22):6397–6407, November 2014.

[26] Sébastien Benzekry, Amanda Tracz, Michalis Mastri, Ryan Corbelli, Dominique Barbolosi, and John M L Ebos. Modeling spontaneous metastasis following surgery: an in vivo-in silico approach. *Cancer Res*, page canres.1389.2015, October 2015.

[27] D Hanahan and R A Weinberg. The hallmarks of cancer. *Cell*, 100(1):57–70, January 2000.

[28] B Vogelstein, S Sur, and C Prives. p53: the most frequently altered gene in human cancers. *Nature Education*, 3(9):6, 2010.

[29] Karin E de Visser, Alexandra Eichten, and Lisa M Coussens. Paradoxical roles of the immune system during cancer development. *Nat Rev cancer*, 6(1):24–37, 2006.

[30] Jared Wels, Rosandra N Kaplan, Shahin Rafii, and David Lyden. Migratory neighbors and distant invaders: tumor-associated niche cells. *Genes Dev*, 22(5):559–574, 2008.

[31] Wikipedia. Necrosis, June 2015. `https://en.wikipedia.org/wiki/Necrosis`.

[32] Gallery Administrator. Dna-replication, October 2015. `http://chemistrypictures.org/index.php/miscellaneous/dna-replication`.

[33] Wikipedia. Epigenetics, October 2013. `https://en.wikipedia.org/wiki/Epigenetics`.

[34] Peter Carmeliet and R K Jain. Angiogenesis in cancer and other diseases. *Nature*, pages 249–257, 2000.

[35] J Folkman. Tumor angiogenesis: therapeutic implications. *New England J Med*, 285:1182–1186, 1971.

[36] D Hanahan and J Folkman. Patterns and emerging mechanisms of the angiogenic switch during tumorigenesis. *Cell*, 86(3):353–364, August 1996.

[37] J Folkman. Angiogenesis inhibitors generated by tumors. *Mol Med*, 1(2):120–2, January 1995.

[38] Peter Carmeliet. Angiogenesis in life, disease and medicine. *Nature*, 438(7070):932–936, 2005.

[39] Hanhua Huang, Abhijit Bhat, Gary Woodnutt, and Rodney Lappe. Targeting the angpt–tie2 pathway in malignancy. *Nat Rev Cancer*, 10(8):575–585, 2010.

[40] Yin-Shan Ng and Patricia A D'Amore. Therapeutic angiogenesis for cardiovascular disease. *Curr Control Trials Cardiovasc Med*, 2(6):278–285, 2001.

[41] Michael S. O'Reilly, Lars Holmgren, Yuen Shing, C Chen, Robert A Rosenthal, Marsha Moses, William S Lane, Yichen Cao, E. Helene Sage, and Judah Folkman. Angiostatin: a novel angiogenesis inhibitor that mediates the suppression of metastases by a Lewis lung carcinoma. *Cell*, 79(2):315–328, October 1994.

[42] M S O'Reilly, Thomas Boehm, Yuen Shing, Naomi Fukai, George Vasios, William S Lane, Evelyn Flynn, James R Birkhead, Bjorn R Olsen, and Judah Folkman. Endostatin: an endogenous inhibitor of angiogenesis and tumor growth. *Cell*, 88(2):277–285, 1997.

[43] Judah Folkman. Angiogenesis: an organizing principle for drug discovery? *Nature Rev Drug Discov*, 6(4):273–86, April 2007.

[44] J A Joyce and J W Pollard. Microenvironmental regulation of metastasis. *Nat Rev Cancer*, 9:239–252, 2009.

[45] Gallery Administrator. Glycolyse, October 2015. `http://uel.unisciel.fr/biologie/module1/module1_ch02/co/apprendre_ch2_02.html`.

[46] R A Gatenby and E T Gawlinski. A reaction-diffusion model of cancer invasion. *Cancer Res*, 56(24):5745–5753, December 1996.

[47] Ann F Chambers, Alan C Groom, and Ian C MacDonald. Dissemination and growth of cancer cells in metastatic sites. *Nat Rev Cancer*, 2(8):563–72, August 2002.

[48] Sébastien Benzekry, Alberto Gandolfi, and Philip Hahnfeldt. Global Dormancy of Metastases Due to Systemic Inhibition of Angiogenesis. *PLoS ONE*, 9(1):e84249–11, January 2014.

[49] Larry Norton. Conceptual and practical implications of breast tissue geometry: toward a more effective, less toxic therapy. *The Oncologist*, 10(6):370–381, June 2005.

[50] Larry Norton and Joan Massagué. Is cancer a disease of self-seeding? *Nat Med*, 12(8):875–878, August 2006.

[51] Andy J Minn, Gaorav P Gupta, Peter M Siegel, Paula D Bos, Weiping Shu, Dilip D Giri, Agnes Viale, Adam B Olshen, William L Gerald, and Joan Massagué. Genes that mediate breast cancer metastasis to lung. *Nature*, 436(7050):518–524, 2005.

[52] Tom E Wheldon. *Mathematical models in cancer research.* Hilger Bristol, 1988.

[53] L Norton and R Simon. The Norton-Simon hypothesis revisited. *Cancer Treat Rep*, 70(1):163–169, January 1986.

[54] Franziska Michor, Martin A Nowak, and Yoh Iwasa. Stochastic dynamics of metastasis formation. *J Theor Biol*, 240(4):521–530, June 2006.

[55] Franziska Michor, Timothy P Hughes, Yoh Iwasa, Susan Branford, Neil P Shah, Charles L Sawyers, and Martin A Nowak. Dynamics of chronic myeloid leukaemia. *Nature*, 435(7046):1267–1270, June 2005.

[56] P Hahnfeldt, D Panigrahy, J Folkman, and L Hlatky. Tumor development under angiogenic signaling: a dynamical theory of tumor growth, treatment response, and postvascular dormancy. *Cancer Res*, 59(19):4770–4775, October 1999.

[57] Hideaki Miyake, Isao Hara, Kazuo Gohji, and K Yamanaka. Relative Expression of Matrix Metalloproteinase-2 and Tissue Inhibitor of Metalloproteinase-2 in Mouse Renal Cell Carcinoma Cells Regulates Their Metastatic Potential. *Clin Cancer Res*, pages 2824–2829, 1999.

[58] L Norton. A Gompertzian model of human breast cancer growth. *Cancer Res*, 48(24):7067–7071, December 1988.

[59] P W Sullivan and S E Salmon. Kinetics of tumor growth and regression in IgG multiple myeloma. *J Clin Invest*, 51(7):1697–1708, July 1972.

[60] D Hart, E Shochat, and Z Agur. The growth law of primary breast cancer as inferred from mammography screening trials data. *Br J Cancer*, 78(3):382–7, August 1998.

[61] R P Araujo and D L S McElwain. A history of the study of solid tumour growth: the contribution of mathematical modelling. *Bull Math Biol*, 66(5):1039–1091, September 2004.

[62] W V Mayneord. On a Law of Growth of Jensen's Rat Sarcoma. *Cancer Res*, 16(4):841–846, 1932.

[63] Helen M Byrne. Dissecting cancer through mathematics: from the cell to the animal model. *Nat Rev Cancer*, 10(3):221–230, March 2010.

[64] A K Laird. Dynamics of tumor growth. *Br J Cancer*, 13:490–502, September 1964.

[65] Sébastien Benzekry, Clare Lamont, Afshin Beheshti, Amanda Tracz, John M L Ebos, Lynn Hlatky, and Philip Hahnfeldt. Classical mathematical models for description and prediction of experimental tumor growth. *PLoS Comput Biol*, 10(8):e1003800, August 2014.

[66] Kristin R Swanson, Carly Bridge, JD Murray, and Ellsworth C Alvord. Virtual and real brain tumors: using mathematical modeling to quantify glioma growth and invasion. *J Neurol Sci*, 216(1):1–10, 2003.

[67] Kristin R Swanson, EC Alvord, and JD Murray. A quantitative model for differential motility of gliomas in grey and white matter. *Cell prolif*, 33(5):317–329, 2000.

[68] Didier Bresch, Thierry Colin, Emmanuel Grenier, and Benjamin Ribba. A viscoelastic model for avascular tumor growth. *Discrete Contin Dyn S*, 2009(Special, sept. 2009):101–108, 2009.

[69] Benoît Perthame, Fernando Quirós, and Juan Luis Vázquez. The hele–shaw asymptotics for mechanical models of tumor growth. *Arch Ration Mech An*, 212(1):93–127, 2014.

[70] H P Greenspan. Models for the growth of a solid tumor by diffusion. *Studies in Applied Mathematics*, 52:317–340, 1972.

[71] J Folkman. Proceedings: Tumor angiogenesis factor. *Cancer Res*, 34(8):2109–2113, August 1974.

[72] Markus R Owen, Helen M Byrne, and Claire E Lewis. Mathematical modelling of the use of macrophages as vehicles for drug delivery to hypoxic tumour sites. *J Theor Biol*, 226(4):377–391, 2004.

[73] Yi Jiang, Jelena Pjesivac-Grbovic, Charles Cantrell, and James P Freyer. A multiscale model for avascular tumor growth. *Biophys. J.*, 89(6):3884–3894, 2005.

[74] Tiina Roose, S Jonathan Chapman, and Philip K Maini. Mathematical models of avascular tumor growth. *SIAM Rev*, 49(2):179–208, 2007.

[75] Jonathan A Sherratt and Mark AJ Chaplain. A new mathematical model for avascular tumour growth. *J Math Biol*, 43(4):291–312, 2001.

[76] MAJ Chaplain. Avascular growth, angiogenesis and vascular growth in solid tumours: The mathematical modelling of the stages of tumour development. *Math Comput Model*, 23(6):47–87, 1996.

[77] Alexander RA Anderson and MAJ Chaplain. Continuous and discrete mathematical models of tumor-induced angiogenesis. *Bull Math Biol*, 60(5):857–899, 1998.

[78] Frédérique Billy, Benjamin Ribba, Olivier Saut, Hélène Morre-Trouilhet, Thierry Colin, Didier Bresch, Jean-Pierre Boissel, Emmanuel Grenier, and Jean-Pierre Flandrois. A pharmacologically based multiscale mathematical model of angiogenesis and its use in investigating the efficacy of a new cancer treatment strategy. *J Theor Biol*, 260(4):545–562, October 2009.

[79] Alexander R A Anderson, Alissa M Weaver, Peter T Cummings, and Vito Quaranta. Tumor morphology and phenotypic evolution driven by selective pressure from the microenvironment. *Cell*, 127(5):905–915, December 2006.

[80] J G Scott, P Gerlee, D Basanta, and A G Fletcher. Mathematical modeling of the metastatic process. In A Malek, editor, *Experimental Metastasis: Modeling and Analysis*. Springer Netherlands, Dordrecht, December 2013.

[81] G M Saidel, L A Liotta, and J Kleinerman. System dynamics of metastatic process from an implanted tumor. *J Theor Biol*, 56(2):417–434, February 1976.

[82] L A Liotta, G M Saidel, and J Kleinerman. Stochastic model of metastases formation. *Biometrics*, 32(3):535–550, September 1976.

[83] Patrick Alfred Pierce Moran et al. The statistical processes of evolutionary theory. *Eugen Rev*, 1962.

[84] Robert Bartoszyński, Lutz Edler, Leonid Hanin, Annette Kopp-Schneider, Lyudmila Pavlova, Alexander Tsodikov, Alexander Zorin, and Andrej Yu Yakovlev. Modeling cancer detection: tumor size as a source of information on unobservable stages of carcinogenesis. *Math Biosci*, 171(2):113–142, 2001.

[85] Leonid Hanin, Jason Rose, and Marco Zaider. A stochastic model for the sizes of detectable metastases. *J Theor Biol*, 243(3):407–417, December 2006.

[86] Hiroshi Haeno, Mithat Gonen, Meghan B Davis, Joseph M Herman, Christine A Iacobuzio-Donahue, and Franziska Michor. Computational modeling of pancreatic cancer reveals kinetics of metastasis suggesting optimum treatment strategies. *Cell*, 148(1):362–375, 2012.

[87] R Demicheli, M W Retsky, D E Swartzendruber, and G Bonadonna. Proposal for a new model of breast cancer metastatic development. *Ann Oncol*, 8(11):1075–1080, November 1997.

[88] Romano Demicheli, Michael W Retsky, William J M Hrushesky, and Michael Baum. Tumor dormancy and surgery-driven interruption of dormancy in breast cancer: learning from failures. *Nat Clin Rev Oncol*, 4(12):699–710, December 2007.

[89] Jacob Scott, Peter Kuhn, and Alexander R A Anderson. Unifying metastasis — integrating intravasation, circulation and end-organ colonization. *Nat Rev Cancer*, 12(7):445–446, May 2012.

[90] Jacob G Scott, Alexander G Fletcher, Philip K Maini, Alexander R A Anderson, and Philip Gerlee. A filter-flow perspective of haematogenous metastasis offers a non-genetic paradigm for personalised cancer therapy. *Eur J Cancer*, 50(17):3068–3075, October 2014.

[91] Jacob G Scott, David Basanta, Alexander R A Anderson, and Philip Gerlee. A mathematical model of tumour self-seeding reveals secondary metastatic deposits as drivers of primary tumour growth. *J R Soc Interface*, 10(82):20130011–20130011, May 2013.

[92] Arturo Araujo, Leah M Cook, Conor C Lynch, and David Basanta. An integrated computational model of the bone microenvironment in bone-metastatic prostate cancer. *Cancer Res*, 74(9):2391–2401, 2014.

[93] Mathworks. kstest - one-sample kolmogorov-smirnov test, 2015. `http://fr.mathworks.com/help/stats/kstest.html`.

[94] Mathworks. chi2gof - chi -square goodness-of-fit test, 2015. `http://fr.mathworks.com/help/stats/chi2gof.html`.

[95] Mathworks. kstest2 - two-sample kolmogorov-smirnov test, 2015. `http://fr.mathworks.com/help/stats/ttest.html`.

[96] Kenneth Levenberg. A method for the solution of certain non–linear problems in least squares. *Q J Math*, 1944.

[97] Donald W Marquardt. An algorithm for least-squares estimation of nonlinear parameters. *SIAM J Numer Anal*, 11(2):431–441, 1963.

[98] G AF Seber and C J Wild. *Nonlinear regression*. Wiley-Interscience, 2003.

[99] S Benzekry. Introduction à l'estimation statistique des paramètres. February 2015.

[100] A Raue, V Becker, U Klingmüller, and J Timmer. Identifiability and observability analysis for experimental design in nonlinear dynamical models. *Chaos*, 20(4):045105, 2010.

[101] Clemens Kreutz, Andreas Raue, and Jens Timmer. Likelihood based observability analysis and confidence intervals for predictions of dynamic models. *BMC Syst Biol*, 6(1):120, 2012.

[102] Marc Lavielle. *Mixed Effects Models for the Population Approach*. Models, Tasks, Methods and Tools. CRC Press, July 2014.

[103] Estelle Kuhn and Marc Lavielle. Maximum likelihood estimation in nonlinear mixed effects models. *Comput Stat Data Anal*, 49(4):1020–1038, 2005.

[104] G G Steel and L F Lamerton. The growth rate of human tumours. *Br J Cancer*, 20(1):74–86, March 1966.

[105] Esmaeil Mehrara, Eva Forssell-Aronsson, Viktor Johanson, Lars Kolby, Ragnar Hultborn, and Peter Bernhardt. A new method to estimate parameters of the growth model for metastatic tumours. *Theor Biol Med Model*, 10:31, 2013.

[106] IJ Jacobs, MF Kohler, RW Wiseman, JR Marks, R Whitaker, BAJ Kerns, P Humphrey, A Berchuck, BAJ Ponder, and RC Bast. Clonal origin of epithelial ovarian carcinoma: analysis by loss of heterozygosity, p53 mutation, and x-chromosome inactivation. *J Natl Cancer Inst*, 84(23):1793–1798, 1992.

[107] MR Novelli, JA Williamson, IPM Tomlinson, G Elia, SV Hodgson, IC Talbot, WF Bodmer, and NA Wright. Polyclonal origin of colonic adenomas in an xo/xy patient with fap. *Science*, 272(5265):1187–1190, 1996.

[108] J A Spratt, D von Fournier, J S Spratt, and E E Weber. Decelerating growth and human breast cancer. *Cancer*, 71(6):2013–2019, March 1993.

[109] PF Verhulst. La loi d'accroissement de la population. *Nouv Mem Acad Roy Soc Belle-lettr Bruxelles*, 18:1, 1845.

[110] Van M Savage, Alexander B Herman, Geoffrey B West, and Kevin Leu. Using fractal geometry and universal growth curves as diagnostics for comparing tumor vasculature and metabolic rate with healthy tissue and for predicting responses to drug therapies. *Discrete Contin Dyn Syst Ser B*, 18(4), 2013.

[111] Alexander B Herman, Van M Savage, and Geoffrey B West. A Quantitative Theory of Solid Tumor Growth, Metabolic Rate and Vascularization. *PLoS ONE*, 6(9):e22973, September 2011.

[112] L a Dethlefsen, J M Prewitt, and M L Mendelsohn. Analysis of tumor growth curves. *J Natl Cancer Inst*, 40(2):389–405, March 1968.

[113] Albert E Casey. The Experimental Alteration of Malignancy with an Homologous Mammalian Tumor Material : I . Results with Intratesticular Inoculation. *Am J Cancer*, 21:760–775, 1934.

[114] A K Laird. Dynamics of Tumour Growth: Comparison of Growth Rates and Extrapolation of Growth Curve To One Cell. *Br J Cancer*, 19:278–91, June 1965.

[115] Larry Norton, Richard Simon, Harmar D Brereton, and Arthur E Bogden. Predicting the course of Gompertzian growth. *Nature*, 264:542–544, 1976.

[116] G F Brunton and T E Wheldon. Prediction of the complete growth pattern of human multiple myeloma from restricted initial measurements. *Cell Tissue Kinet*, 10(6):591–594, November 1977.

[117] G F Brunton and T E Wheldon. Characteristic species dependent growth patterns of mammalian neoplasms. *Cell Tissue Kinet*, 11(2):161–175, March 1978.

[118] Christine L Chaffer and Robert A Weinberg. A perspective on cancer cell metastasis. *Science*, 331(6024):1559–1564, March 2011.

[119] Patricia S Steeg et al. Tumor metastasis: mechanistic insights and clinical challenges. *Nat Med*, 12(8):895–904, 2006.

[120] Xuefeng Gao, J Tyson McDonald, Lynn Hlatky, and Heiko Enderling. Acute and fractionated irradiation differentially modulate glioma stem cell division kinetics. *Cancer Res*, 73(5):1481–90, January 2013.

[121] Elena De Angelis and Luigi Preziosi. Advection-diffusion models for solid tumour evolution in vivo and related free boundary problem. *Math Mod Meth Appl S*, 10(03):379–407, 2000.

[122] Didier Bresch, Thierry Colin, Emmanuel Grenier, Benjamin Ribba, and Olivier Saut. Computational modeling of solid tumor growth: the avascular stage. *SIAM J Sci Comput*, 32(4):2321–2344, 2010.

[123] Olivier Saut, Jean-Baptiste Lagaert, Thierry Colin, and Hassan M Fathallah-Shaykh. A multilayer grow-or-go model for gbm: effects of invasive cells and anti-angiogenesis on growth. *Bull Math Biol*, 76(9):2306–2333, 2014.

[124] Sébastien Benzekry. *Modélisation, analyse mathématique de thérapies anti-cancéreuses pour les cancers métastatiques.* PhD thesis, Aix Marseille 1, 2011.

[125] Fabien Montel, Morgan Delarue, Jens Elgeti, Danijela Vignjevic, Giovanni Cappello, and Jacques Prost. Isotropic stress reduces cell proliferation in tumor spheroids. *New J Phys*, 14(5):055008, May 2012.

[126] D. Drasdo and S. Hohme. A single-cell-based model of tumor growth in vitro: monolayers and spheroids. *Phys Biol*, 2(3):133–147, Sep 2005.

[127] Guillaume Lefèbvre, Thierry Colin, Clair Poignard, Olivier Saut, and François Cornelis. Modélisation de résistance aux traitements contre le cancer.

[128] Chohong Min and Frédéric Gibou. A second order accurate level set method on non-graded adaptive cartesian grids. *J Comput Phys*, 225(1):300–321, 2007.

[129] Dongbin Xiu and George Em Karniadakis. A semi-lagrangian high-order method for navier–stokes equations. *J Comput Phys*, 172(2):658–684, 2001.

[130] Frédéric Couderc. *Développement d'un code de calcul pour la simulation d'écoulements de fluides non miscibles. Application à la désintégration assistée d'un jet liquide par un courant gazeux.* PhD thesis, Ecole nationale superieure de l'aeronautique et de l'espace, 2007.

[131] Guang-Shan Jiang and Danping Peng. Weighted eno schemes for hamilton–jacobi equations. *SIAM J Sci Comput*, 21(6):2126–2143, 2000.

[132] Sigal Gottlieb and Chi-Wang Shu. Total variation diminishing runge-kutta schemes. *Math Comp*, 67(221):73–85, 1998.

[133] William P Ziemer. *Weakly differentiable functions: Sobolev spaces and functions of bounded variation*, volume 120. Springer Science & Business Media, 2012.

[134] Haïm Brezis. *Analyse fonctionnelle.* 1983.

[135] Elias M Stein. *Singular integrals and differentiability properties of functions*, volume 2. Princeton university press, 1970.

[136] Laurent Baratchart, Alexander Borichev, and Slah Chaabi. Pseudo-holomorphic functions at the critical exponent. *To appear in J Eur Math Soc*, 2013.

[137] Grégoire Allaire. *Analyse numérique et optimisation: Une introduction à la modélisation mathématique et à la simulation numérique.* Editions Ecole Polytechnique, 2005.

[138] David Jerison and Carlos E Kenig. The inhomogeneous dirichlet problem in lipschitz domains. *J Funct Anal*, 130(1):161–219, 1995.

[139] Julien Jouganous. *Modélisation et simulation de la croissance de métastases pulmonaires*. PhD thesis, Bordeaux, 2015.

[140] Thierry Colin, Thomas Michel, and Clair Poignard. Mathematical study and asymptotic analysis of a model for tumour drug resistance. Research Report 8784, Inria Bordeaux Sud-Ouest, October 2015.

[141] Sachie Hiratsuka, Akira Watanabe, Hiroyuki Aburatani, and Yoshiro Maru. Tumour-mediated upregulation of chemoattractants and recruitment of myeloid cells predetermines lung metastasis. *Nat Cell Biol*, 8(12):1369–1375, 2006.

[142] Bethan Psaila and David Lyden. The metastatic niche: adapting the foreign soil. *Nat Rev Cancer*, 9(4):285–293, 2009.

[143] John M L Ebos, Christina R Lee, William Cruz-Munoz, Georg A Bjarnason, James G Christensen, and Robert S Kerbel. Accelerated metastasis after short-term treatment with a potent inhibitor of tumor angiogenesis. *Cancer Cell*, 15(3):232–239, March 2009.

[144] Bartlomiej Waclaw, Ivana Bozic, Meredith E Pittman, Ralph H Hruban, Bert Vogelstein, and Martin A Nowak. A spatial model predicts that dispersal and cell turnover limit intratumour heterogeneity. *Nature*, 525(7568):261–264, 2015.

[145] B Ribba, O Saut, T Colin, D Bresch, E Grenier, and J P Boissel. A multiscale mathematical model of avascular tumor growth to investigate the therapeutic benefit of anti-invasive agents. *J Theor Biol*, 243(4):532–541, December 2006.

[146] J S Spratt, J S Meyer, and J A Spratt. Rates of growth of human neoplasms: Part II. 61(1):68–83, January 1996.

[147] The Mathworks. Matlab with statistics and optimization toolboxes. 2013.

[148] N Hartung. Efficient Resolution of Metastatic Tumour Growth Models by Reformulation into Integral Equations. *Discrete Contin Dyn Syst-B*, 2014.

[149] Benjamin Ribba, Thierry Colin, and Santiago Schnell. A multiscale mathematical model of cancer, and its use in analyzing irradiation therapies. *Theor Biol Med Model*, 3(1):7, 2006.

[150] Ilaria Malanchi, Albert Santamaria-Martínez, Evelyn Susanto, Hong Peng, Hans-Anton Lehr, Jean-Francois Delaloye, and Joerg Huelsken. Interactions between cancer stem cells and their niche govern metastatic colonization. *Nature*, 481(7379):85–89, 2012.

[151] Hematopoietic Products R&D Albertus W. Wognum, Stephen J. Szilvassy. Mini-review hematopoietic stem and progenitor cells, April 2015. `http://www.stemcell.com/~/media/Technical%20Resources/F/B/7/E/9/MR019HematopoiesisOnline_29784WEB.pdf`.

[152] Wikipedia. Endothelium, 2015. `https://en.wikipedia.org/wiki/Endothelium`.

[153] Erkki Ruoslahti. Specialization of tumour vasculature. *Nat Rev Cancer*, 2(2):83–90, 2002.

[154] Jun-ichi Kawabe Yoshinori Minami, Takaaki Sasaki and Yoshinobu Ohsaki. Accessory cells in tumor angiogenesis — tumor-associated pericytes, research directions in tumor angiogenesis, doi: 10.5772/54523, 2013. `http://www.intechopen.com/books/research-directions-in-tumor-angiogenesis/accessory-cells-in-tumor-angiogenesis-tumor-associated-pericytes`.

[155] Matijs van Meurs, P Kumpers, JJ Ligtenberg, JH Meertens, Grietje Molema, and Jan G Zijlstra. Bench-to-bedside review: Angiopoietin signalling in critical illness-a future target. *Crit Care*, 13(2):207, 2009.

[156] Harold F Dvorak. Tumors: wounds that do not heal: similarities between tumor stroma generation and wound healing. *N Engl J Med*, 315(26):1650–1659, 1986.

[157] Regeneration center of Thailand. Mesenchymal stem cells isolation & expansion, 2015. https://www.yumpu.com/en/document/view/22993688/hematopoietic-stem-cell-amp-lineage-specific-markers.

[158] Aria F Olumi, Gary D Grossfeld, Simon W Hayward, Peter R Carroll, Thea D Tlsty, and Gerald R Cunha. Carcinoma-associated fibroblasts direct tumor progression of initiated human prostatic epithelium. *Cancer Res*, 59(19):5002–5011, 1999.

[159] R&D systems. Hematopoietic stem cell and lineage-specific markers, November 2014. https://www.yumpu.com/en/document/view/22993688/hematopoietic-stem-cell-amp-lineage-specific-markers.

[160] Stephen Paget. The distribution of secondary growths in cancer of the breast. *The Lancet*, 133(3421):571–573, 1889.

[161] James Ewing. Neoplastic diseases. a treatise on tumors. *Am J Med Sci*, 176(2):278, 1928.

[162] Dale R Coman, Morton McCutcheon, et al. Studies on the mechanisms of metastasis. The distribution of tumors in various organs in relation to the distribution of arterial emboli. *Cancer Res*, 11(8):648–651, 1951.

[163] M Dianne Cameron, Eric E Schmidt, Nancy Kerkvliet, Kishore V Nadkarni, Vincent L Morris, Alan C Groom, Ann F Chambers, and Ian C MacDonald. Temporal progression of metastasis in lung: cell survival, dormancy, and location dependence of metastatic inefficiency. *Cancer Res*, 60(9):2541–2546, 2000.

[164] Keith J Luzzi, Ian C MacDonald, Eric E Schmidt, Nancy Kerkvliet, Vincent L Morris, Ann F Chambers, and Alan C Groom. Multistep nature of metastatic inefficiency: dormancy of solitary cells after successful extravasation and limited survival of early micrometastases. *Am J Pathol*, 153(3):865–873, 1998.

[165] Yvonne Kienast and Louisa Von Baumgarten. Real-time imaging reveals the single steps of brain metastasis formation. *Nat Med*, 16(1):116–123, 2009.

[166] Rosandra N Kaplan, Shahin Rafii, and David Lyden. Preparing the "soil": the premetastatic niche. *Cancer Res*, 66(23):11089–11093, 2006.

[167] Shahin Rafii and David Lyden. A few to flip the angiogenic switch. *Science*, 319(5860):163, 2008.

[168] Shahin Rafii and David Lyden. S100 chemokines mediate bookmarking of premetastatic niches. *Nat Cell Biol*, 8(12):1321–1323, 2006.

[169] Bethan Psaila, Rosandra N Kaplan, Elisa R Port, and David Lyden. Priming the soil for breast cancer. *Metastasis*, 26(2006):65–74, 2007.

[170] Marcin Kowanetz, Xiumin Wu, John Lee, Martha Tan, Thijs Hagenbeek, Xueping Qu, Lanlan Yu, Jed Ross, Nina Korsisaari, Tim Cao, et al. Granulocyte-colony stimulating factor promotes lung metastasis through mobilization of ly6g+ ly6c+ granulocytes. *Proc Natl Acad Sci*, 107(50):21248–21255, 2010.

[171] Zvi Granot, Erik Henke, Elizabeth A Comen, Tari A King, Larry Norton, and Robert Benezra. Tumor entrained neutrophils inhibit seeding in the premetastatic lung. *Cancer Cell*, 20(3):300–314, 2011.

[172] Takashi Minami, Shuying Jiang, Keri Schadler, Jun-ichi Suehiro, Tsuyoshi Osawa, Yuichi Oike, Mai Miura, Makoto Naito, Tatsuhiko Kodama, and Sandra Ryeom. The calcineurin-nfat-angiopoietin-2 signaling axis in lung endothelium is critical for the establishment of lung metastases. *Cell reports*, 4(4):709–723, 2013.

[173] Elizabeth Hahn-Dantona, Jose F Ruiz, Paul Bornstein, and Dudley K Strickland. The low density lipoprotein receptor-related protein modulates levels of matrix metalloproteinase 9 (mmp-9) by mediating its cellular catabolism. *J Biol Chem*, 276(18):15498–15503, 2001.

[174] Stanley Zucker and Wen-Tien Chen. *Cell surface proteases*, volume 54. Academic Press, 2003.

[175] Vanessa Hsei, Geralyn G DeGuzman, Allison Nixon, and Jacques Gaudreault. Complexation of vegf with bevacizumab decreases vegf clearance in rats. *Pharmaceut Res*, 19(11):1753–1756, 2002.

[176] Motoaki Ohtsubo and James M Roberts. Cyclin-dependent regulation of g˜ 1 in mammalian fibroblasts. *Science*, 259:1908–1908, 1993.

[177] Wikipedia. Fibroblast, September 2015. `https://en.wikipedia.org/wiki/Fibroblast`.

[178] Lane K Christenson and Richard L Stouffer. Isolation and culture of microvascular endothelial cells from the primate corpus luteum. *Biol Reprod*, 55(6):1397–1404, 1996.

[179] Christophe Delacourt. Développement alvéolaire normal et pathologique, September 2015. `http://b2pcr-esi.bcpp.master.univ-paris-diderot.fr/M2/R/diapo_rpw/UE1ImmunoR/courspw/DELACOURT_Lung%20development2_2012.pdf`.

[180] Sandrine Launois-Rollinat. Circulation pulmonaire, September 2015. `http://www.uvp5.univ-paris5.fr/wikinu/docvideos/Grenoble_1112/launois_rollinat_sandrine/launois_rollinat_sandrine_P07/launois_rollinat_sandrine_P07.pdf`.

[181] Emma E Pepperell and Suzanne M Watt. A novel application for a 3-dimensional timelapse assay that distinguishes chemotactic from chemokinetic responses of hematopoietic cd133+ stem/progenitor cells. *Stem Cell Res*, 11(2):707–720, 2013.

[182] Roeland MH Merks, Erica D Perryn, Abbas Shirinifard, and James A Glazier. Contact-inhibited chemotaxis in de novo and sprouting blood-vessel growth. *PLoS Comput Biol*, 4(9):e1000163, 2008.

[183] EJ Ribot, D Wecker, AJ Trotier, B Dallaudière, W Lefrançois, E Thiaudière, J-M Franconi, and S Miraux. Water selective imaging and bssfp banding artifact correction in humans and small animals at 3t and 7t, respectively. *PLoS ONE*, 10(10):e0139249, 2015.

[184] Ribot EJ, FM Martinez-Santiesteban, C Simedrea, PS Steeg, AF Chambers, BK Rutt, and PJ. Foster. In vivo single scan detection of both iron-labeled cells and breast cancer metastases in the mouse brain using balanced steady-state free precession imaging at 1.5 t. *J Magn Reson Imaging*, (34):231–238, 2011.

[185] S Miraux, P Massot, EJ Ribot, JM Franconi, and E. Thiaudiere. 3d truefisp imaging of mouse brain at 4.7t and 9.4t. *J Magn Reson Imaging*, (28):497–503, 2008.

[186] S Benzekry. Mathematical and numerical analysis of a model for anti-angiogenic therapy in metastatic cancers. *ESAIM, Math Model Numer Anal*, 46(2):207–237, 2012.