



THÈSE

pour obtenir le grade de Docteur de

l'Université de Lille

Spécialité : INFORMATIQUE

Danilo DURSONIAH

Modélisation Computationnelle de l'Absorption Intestinale du Glucose pour la Prédiction du Diabète de Type 2

Thèse soutenue le 23 octobre 2024

Composition du jury :

<i>M.</i>	JEAN-PAUL COMET	Professeur, lab. I3S, Université Côte d'Azur	(Rapporteur)
<i>Mme</i>	ANNE SIEGEL	Directrice de recherche, IRISA, CNRS & Université de Rennes	(Rapportrice)
<i>M</i>	STEFAN HAAR	Directeur de recherche, Inria, Université Paris Saclay	(Examineur)
<i>M.</i>	FRANÇOIS PATTOU	Professeur, INSERM U1190, Université de Lille	(Co-Directeur de Thèse)
<i>M.</i>	CÉDRIC LHOSSAINE	Professeur, CRISAL, Université de Lille	(Co-Directeur de Thèse)
<i>M.</i>	MAXIME FOLSCHETTE	Maître de Conférences, CRISAL, Centrale Lille	(Co-encadrant de Thèse)

Équipe BioComputing du Laboratoire CRISAL

Ecole Doctorale MADIS-631, Mathématiques-Sciences du numérique et de leurs interactions
Centre de Recherche en Informatique, Signal et Automatique de Lille

Résumé

Les recherches sur le diabète de type 2 (DT2) se sont principalement concentrées sur la fonction des cellules bêta du pancréas et la sensibilité à l'insuline. De nombreux indices ont été proposés pour les mesurer, calculés à partir de modèles plus ou moins complexes utilisant des données statiques de glycémie à jeun ou des données dynamiques de test glycémique oral. La chirurgie bariatrique a révélé un troisième facteur pouvant causer le DT2 : l'absorption intestinale de glucose (IGA). Contrairement aux deux premiers facteurs causaux, aucun modèle n'a encore été proposé pour mesurer l'effet de l'IGA sur le DT2. Mesurer expérimentalement l'IGA nécessite l'accès à la veine porte, ce qui est impossible chez l'homme. Une technique utilisant le glucose marqué a été proposée en alternative, mais elle reste très difficile à réaliser et demande une expertise particulière, rendant son usage en clinique peu pratique. Notons aussi que les modèles actuels pour prédire la réponse postprandiale de glucose nécessitent ce gold standard. Les rares modèles existants sont partiellement mécanistiques et assez complexes. Cette thèse vise à résoudre ces problèmes. Premièrement, nous reproduisons le modèle postprandial de Dalla Man et les simulations de leur article de 2007. Ce modèle, décrit uniquement par des équations différentielles ordinaires (EDO), a été partiellement retranscrit en un système de réactions chimiques pour mettre en lumière les mécanismes physiologiques. Cette implémentation a permis de vérifier la reproductibilité, malgré l'absence des données originales, puis de tester le modèle avec nos données cliniques OBEDIAB, révélant ainsi ses limites en termes d'estimations et d'identifiabilité. Ainsi, pour éviter d'utiliser le gold standard du multi-traceur, nous avons employé le D-xylose, un sucre analogue au glucose, comme biomarqueur pour observer directement l'absorption intestinale de glucose (IGA), grâce à nos données précliniques obtenues sur des minipigs. Nous avons développé le premier modèle de D-xylose à notre connaissance. Ce modèle a été choisi en estimant des paramètres à partir de nos données, puis validé par une analyse d'identifiabilité pratique et une analyse de sensibilité globale. Ces analyses ont également permis d'étudier les contributions de la vidange gastrique et de l'absorption intestinale sur la dynamique du D-xylose. Enfin nous nous interrogerons sur les liens entre la modélisation de la glycémie et celle de la réponse de D-xylose postprandial tout en envisageant les applications cliniques et les limites du modèle de D-xylose.

Mots-clés : Biologie des systèmes, modélisation, réseaux de réactions chimiques, équations différentielles ordinaires, estimation de paramètres, analyse d'identifiabilité, diabète de type 2, D-xylose.

Remerciements

Je tiens à exprimer ma plus profonde gratitude à toutes les personnes qui ont contribué, de près ou de loin, à la réalisation de cette thèse. Mes remerciements les plus sincères vont d'abord à mon directeur de thèse, CÉDRIC LHOSSAINE, ainsi qu'à mon encadrant, MAXIME FOLSCHETTE, pour leur encadrement, leur expertise, leurs précieux conseils et leur soutien indéfectible tout au long de ce travail. Je tiens également à remercier mon co-directeur de thèse, FRANÇOIS PATTOU, ainsi qu'à son équipe de recherche, pour leur expertise en recherche translationnelle et dont la collaboration a été la clé de voûte de cette thèse et de différents projets corrélés. Je remercie également les membres du jury pour l'honneur qu'ils m'ont fait en acceptant d'évaluer mes travaux. Un grand merci à mes collègues et amis pour leur soutien moral et intellectuel.

Bien entendu, le fruit de ces années de recherche n'aurait pas vu le jour sans mon plus grand soutien, ma famille : leur patience, leur encouragement et leur amour constant m'ont été indispensables pour ne pas abandonner et mener à bien ce projet ambitieux.

Enfin, je souhaite exprimer ma reconnaissance envers le laboratoire CRISAL, la région Hauts-de-France et l'École Doctorale MADIS, qui ont permis la réalisation de mes travaux, que ce soit par leur dispositif de financement, leurs propositions de formations, ainsi qu'à leur réactivité pour la réalisation des tâches administratives.

Bien plus qu'une formation de l'enseignement supérieur, la thèse a été avant-tout, pour moi, une aventure intellectuelle et humaine, rythmée par de modestes accomplissements et de doutes, mais toujours au service de la science et, plus égoïstement, à mon propre épanouissement intellectuel. Ainsi, je tiens à réitérer mes remerciements les plus sincères envers les personnes et les structures susmentionnées.

Table des matières

1	Introduction	6
1.1	Diabète de type 2 : caractéristiques et diagnostics	6
1.2	Chirurgie bariatrique et mise en évidence du rôle intestinale dans le diabète de type 2	9
1.3	Suivi et prédiction de l'absorption intestinale de glucose après un repas	11
1.4	Approches de modélisation	12
1.5	Problématique	14
1.6	Contributions	15
2	Méthodologie et État de l'art	17
2.1	Méthodologie	17
2.1.1	Formalismes	17
2.1.2	Estimation des paramètres	22
2.1.3	Analyses pour la validation de modèles	26
2.2	État de l'art des modèles de glucose-insuline	28
2.2.1	Modèle minimal de Bergman	28
2.2.2	Modèle de vidange gastrique d'Elashoff	29
2.2.3	Modèle de vidange gastrique de Dalla Man	30
2.2.4	Modèle post-prandial homéostatique de Dalla Man	32
2.2.5	Modèle intestinal de Salinari	37
2.2.6	Conclusion	40
3	Étude du modèle postprandial de la dynamique glucose-insuline de Dalla Man : Reproductibilité et limites	41
3.1	Contexte et Motivations	41
3.2	Données cliniques	42
3.3	Modèle de Dalla Man <i>et al.</i>	46
3.4	Du modèle DT2 au modèle normal	48
3.5	Estimation des paramètres des diabétiques obèses et RYGB	52
3.6	Discussion	54
3.6.1	Du modèle de DT2 au modèle normal	54
3.6.2	Estimation des paramètres des diabétiques obèses et de RYGB	54

3.7	Nécessité d'un nouveau modèle	55
4	Modélisation de l'absorption intestinale de glucose à partir des données de D-xylose	56
4.1	Motivation	56
4.2	Jeux de données des minipigs	59
4.3	Volume de distribution	60
4.4	Modèles de D-xylose	60
4.4.1	Modèle oral « simple »	61
4.4.2	Modèle oral « multi-compartiments »	62
4.4.3	Modèle oral « à rétention »	62
4.4.4	Modèles intraveineux et intrajejunal	65
4.5	Exploitation des modèles	65
4.5.1	Estimation des paramètres	65
4.5.2	Analyse de l'identifiabilité pratique	67
4.5.3	Prédiction du taux d'absorption du glucose à partir du jeu de données de résection intestinale	68
4.6	Rôles relatifs de la vidange gastrique et de l'absorption intestinale	71
4.6.1	Analyse de sensibilité globale	71
4.6.2	Modèle de Vidange Gastrique « Complexe »	73
4.7	Intérêt d'un modèle de Xylose	74
5	Conclusion et Perspectives	76
5.1	Conclusion	76
5.2	Perspectives	78

Chapitre 1

Introduction

Le diabète de type 2 est considéré comme l'épidémie de ce siècle. Il consiste en un déséquilibre du taux de glucose dans le sang, provoquant des lésions dans de nombreux organes. Il est un facteur important de plusieurs autres maladies dont les maladies cardio-vasculaires.

Dans cette introduction, nous présenterons dans un premier temps les caractéristiques cliniques du diabète de type 2, ainsi que ses causes. Les principaux diagnostics et indices permettant d'évaluer plus précisément l'état diabétique d'un patient seront brièvement décrits. Ces indices mettent en lumière les deux principales causes, ou effets, du diabète traditionnellement étudiés : la fonction bêta du pancréas et la résistance à l'insuline.

Nous insisterons sur le rôle sous-estimé de l'activité intestinale dans la survenue de la maladie, révélé par la chirurgie bariatrique. Il s'agit d'un troisième paramètre du diabète sur lequel cette thèse se focalisera.

L'ensemble des hypothèses explicatives et confondantes de la maladie nous amènera à mentionner les méthodes de suivi du taux de sucre dans le sang après un repas. Les biomarqueurs et les méthodes de suivi de la dynamique du glucose nous pousseront à nous interroger sur la compréhension et la prédiction de ce que nous observons après un repas, ce qui aboutira au besoin de la modélisation informatique par un formalisme mathématique. Nous aborderons brièvement quelques modèles de référence qui seront présentés plus en détail dans le chapitre 2 : *Méthodologie et État de l'Art*. Nous évoquerons le besoin de proposer un nouveau modèle pour étudier le rôle intestinal dans la dynamique du taux de glucose dans le sang après un repas (ou *profil glycémique postprandial*) au regard des limites des modèles les plus utilisés.

1.1 Diabète de type 2 : caractéristiques et diagnostics

Le diabète de type 2 est une maladie métabolique chronique caractérisée par une *glycémie* (taux de glucose dans le sang) anormalement élevée, due à une résistance à l'insuline et à une production insuffisante de cette dernière (Ahmad et al., 2022). L'*insuline* est une hormone produite par le pancréas qui aide à réguler la glycémie en permettant aux cellules d'absorber et d'utiliser le glucose comme source d'énergie. Chez les personnes atteintes de diabète de type 2

– que nous noterons occasionnellement DT2 – les cellules du corps deviennent résistantes aux effets de l’insuline et le pancréas est dans l’incapacité de produire suffisamment d’insuline pour compenser cette résistance. En conséquence, le glucose s’accumule dans la circulation sanguine au lieu d’être utilisé par les cellules, ce qui entraîne une glycémie anormalement élevée. Plusieurs facteurs contribuent au développement du diabète de type 2, notamment la prédisposition génétique (Prasad and Groop, 2015), ainsi que les facteurs épigénétiques et environnementaux (Dhawan and Natarajan, 2019) (alimentation, obésité, âge, etc.). La maladie est souvent diagnostiquée à l’âge adulte, mais elle peut survenir à tout âge. Les symptômes courants du diabète de type 2 comprennent une soif accrue, des mictions fréquentes, une perte de poids inexplicée, de la fatigue et une vision floue. S’il n’est pas traité, le diabète de type 2 peut entraîner de graves complications telles qu’une maladie cardiaque, un accident vasculaire cérébral, des lésions rénales, nerveuses et des problèmes de vision. Le diagnostic du diabète de type 2 repose sur une combinaison de tests cliniques allant de la mesure de la glycémie à l’évaluation des facteurs de risque. En routine clinique, le diagnostic repose majoritairement sur les trois tests cités ci-après (Cox and Edelman, 2009) :

- Test de glycémie à jeun (TGO) : le test de glycémie à jeun est l’un des tests de dépistage les plus courants pour le diabète de type 2. Il mesure la glycémie après une période de jeûne d’au moins huit heures. Un taux de glucose sanguin égal ou supérieur à 126 milligrammes par décilitre (mg/dL), soit 7 millimoles par litre (mmol/L), à deux reprises est généralement considéré comme un test positif.
- Test de l’hémoglobine glyquée (HbA1c) : il mesure le pourcentage d’hémoglobine qui est glycosylée, c’est-à-dire liée au glucose, dans le sang. Il est représentatif de la glycémie moyenne sur les deux à trois mois qui précèdent le test. Un taux d’HbA1c de 6,5 % ou plus est généralement utilisé comme critère pour confirmer le diagnostic pathologique.
- Test de tolérance au glucose oral (TTGO) : il implique de boire une solution sucrée contenant une quantité spécifique de glucose, puis de mesurer la glycémie à intervalles réguliers pendant plusieurs heures. Un taux de glucose égal ou supérieur à 200 mg/dL (11.1 mmol/L) deux heures après la consommation de la solution est considéré comme un test positif.
- Évaluation des symptômes et des facteurs de risque : les médecins peuvent également prendre en compte les symptômes du patient (comme la soif excessive, la polyurie, la fatigue, etc.) ainsi que les facteurs de risque connus tels que l’obésité, l’hypertension artérielle, les antécédents familiaux de diabète, etc.

Il est important de noter que le diagnostic du diabète de type 2 ne se fait pas uniquement sur la base d’un seul test, mais plutôt sur une combinaison de tests et d’évaluations cliniques. Si un diagnostic de diabète de type 2 est posé, le patient est généralement orienté vers un professionnel de la santé pour un suivi et un plan de traitement approprié. Aussi, une surveillance plus fine de la concentration de glucose dans le sang provenant de diverses sources peut s’avérer nécessaire pour mesurer les mécanismes à l’oeuvre dans sa régulation. Dans la majeure partie des

cas, il s'agit de mesurer la capacité d'absorption du glucose par les tissus en réponse à l'insuline et la capacité de production d'insuline par le pancréas. Dans cette perspective, de nombreuses approches ont été proposées : du simple prélèvement sanguin à jeun au test de tolérance au glucose par voie intraveineuse en passant par le test par voie orale. Expérimentalement et cliniquement, comme nous allons le voir brièvement sur quelques exemples, les mesures à jeun présentent l'alternative la plus simple et la plus économique, alors que les tests s'appuyant sur l'injection intraveineuse sont les plus informatifs et les plus précis.

Les indices HOMA (Homeostasis Model Assessment) (Matthews et al., 1985) sont des outils utilisés en médecine pour évaluer la résistance à l'insuline et la fonction des cellules bêta du pancréas, c'est-à-dire la capacité de cet organe à produire de l'insuline efficacement. Ce sont des indices parmi ceux les plus utilisés en clinique. Il sont en effet calculés à partir des concentrations de glucose et d'insuline à jeun qui sont très faciles à mesurer et sont très économiques. Les deux indices HOMA sont :

- HOMA-IR (Homeostasis Model Assessment of Insulin Resistance) : cet indice évalue la résistance à l'insuline. Habituellement, une valeur élevée de HOMA-IR est associée à une résistance à l'insuline, ce qui signifie que les cellules du corps ont du mal à répondre efficacement à l'insuline.
- HOMA-B (Homeostasis Model Assessment of Beta-cell function) : cet indice évalue la fonction des cellules bêta du pancréas, qui produisent de l'insuline. Une valeur faible de HOMA-B peut indiquer une altération de la fonction des cellules bêta du pancréas, ce qui peut contribuer au développement du diabète de type 2.

Les valeurs de HOMA-IR et de HOMA-B peuvent être utilisées pour évaluer la santé métabolique d'un individu, pour diagnostiquer la résistance à l'insuline, pour évaluer les effets de certains traitements ou interventions sur l'insulinorésistance, ou encore pour prédire le risque de développer un diabète de type 2. Cependant, ces indices sont calculés à partir de modèles linéaires très simplifiés et fournissent des estimations de la résistance à l'insuline et de la fonction bêta très approximatives. Les indices HOMA2 (Wallace et al., 2004) ont été introduits pour palier à ces défauts mais ils demeurent liés au glucose et à l'insuline à jeun, mesurent essentiellement la résistance hépatique à l'insuline et ne donnent pas d'indication sur la dynamique de l'homéostasie du système glucose-insuline. L'indice de Matsuda-DeFronzo (Matsuda and DeFronzo, 1999) corrige à cette lacune : c'est un indice de sensibilité à l'insuline calculé non seulement à partir des concentrations d'insuline et de glucose à jeun mais également de tests de tolérance au glucose oral (données dynamiques). Cet indice, également très utilisé en clinique, reflète la sensibilité à l'insuline d'origine hépatique mais aussi d'origine tissulaire.

Au delà des méthodes fondées sur les données de glycémie et d'insulinémie à jeun ou à la suite d'un test oral, signalons l'existence d'une méthode, particulièrement importante pour cette thèse, fondée sur un test de tolérance au glucose par voie intraveineuse. Il s'agit du calcul de la sensibilité à l'insuline et de la sensibilité au glucose (c'est-à-dire le taux d'absorption de glucose dû à sa propre concentration dans le plasma) à partir d'un modèle mathématique de régulation de l'homéostasie du système glucose/insuline communément appelé *modèle minimal*. Ce modèle,

proposé par Bergman et Cobelli (Bergman et al., 1979; Bergman, 2021) a été développé pour représenter les interactions entre l'insuline, le glucose et les tissus périphériques à l'aide de seulement 2 équations différentielles ordinaires et 4 paramètres. Contrairement au calcul des indices précédents, cette approche nécessite d'estimer les paramètres du modèle à partir des données de glycémie mais donne une estimation plus précise de la sensibilité à l'insuline. Nous détaillerons plus précisément modèle minimal dans le chapitre 2.

En plus de ces caractérisations de la maladie basées sur l'action du pancréas et des tissus, des observations cliniques ont montré que les chirurgies bariatriques subies par des patients obèses morbides, souffrant de diabète de type 2 – entre autres comorbidités – conduisent systématiquement à une guérison de la maladie métabolique, au-delà de la perte de poids attendue. Il semble donc qu'un troisième paramètre, en lien avec le trac gastro-intestinal, soit souvent sous-estimé dans les diagnostics du diabète de type 2.

1.2 Chirurgie bariatrique et mise en évidence du rôle intestinale dans le diabète de type 2

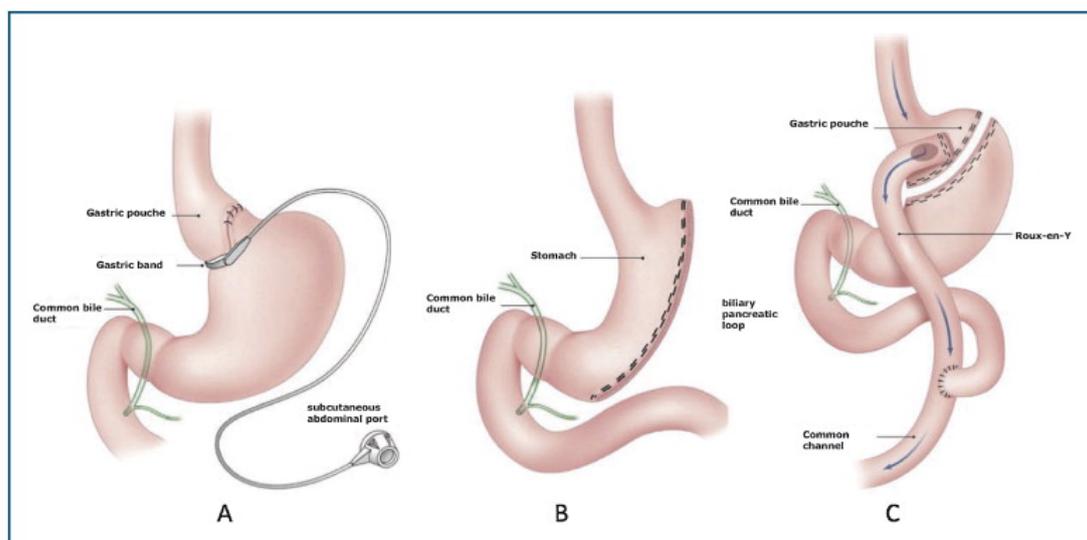


FIGURE 1.1 – A : Anneau gastrique ajustable (*adjustable gastric banding*). B : Gastrectomie en manchon, aussi dite longitudinale (*sleeve gastrectomy*). C : Court-circuit gastrique (*gastric bypass*) de Roux-en-Y. Source : (Caiazzo and Pattou, 2013)

La chirurgie bariatrique, également connue sous le nom de chirurgie métabolique, est un ensemble d'interventions chirurgicales visant à aider les personnes souffrant d'obésité sévère à perdre du poids. Ces procédures sont généralement envisagées lorsque d'autres méthodes de perte de poids, telles que le régime alimentaire et l'exercice physique, n'ont pas réussi et lorsque l'obésité présente de graves risques pour la santé.

La chirurgie bariatrique peut entraîner une perte de poids significative soutenue et entraîne souvent une amélioration ou une rémission des problèmes de santé liés à l'obésité (Caiazzo and

Pattou, 2013). Les 3 principales techniques chirurgicales sont le *Roux-En-Y*, le *gastrectomie en manchon* et l'*anneau gastrique ajustable* (voir figure 1.1). Durant une chirurgie de pontage gastrique ou *court-circuit gastrique de Roux-en-Y*, une petite poche est créée au sommet de l'estomac et est reliée directement à l'intestin grêle, contournant une partie de l'estomac et la partie supérieure de l'intestin grêle. Cela limite la quantité de nourriture que l'estomac peut contenir et *réduit l'absorption des calories et des nutriments*.

La gastrectomie en manchon (ou *gastric sleeve surgery*) implique l'ablation d'une grande partie de l'estomac, laissant un manchon ou un tube en forme de banane. La taille réduite de l'estomac limite la quantité de nourriture pouvant être consommée et affecte également les hormones intestinales qui influencent la faim et le métabolisme.

Enfin, durant la pose de l'anneau gastrique ajustable (ou *gastric band surgery*), une bande est placée autour de la partie supérieure de l'estomac, créant une petite poche et un passage étroit vers le reste de l'estomac. La bande peut être ajustée pour contrôler la taille du passage, régulant ainsi la quantité de nourriture pouvant être consommée.

La chirurgie bariatrique n'est pas une solution universelle et le choix de la procédure dépend de facteurs tels que l'état de santé du patient ou de ses antécédents médicaux. Comme pour toute intervention chirurgicale, il existe des risques et des complications potentiels, et un examen attentif et une consultation des professionnels de la santé sont cruciaux. Comme nous le montrerons de façon plus détaillée dans le chapitre 2, parmi les trois types de chirurgie présentés, on peut remarquer que le court-circuit gastrique de Roux-en-Y est le seul impliquant une diminution drastique de la glycémie postprandiale, c'est-à-dire après un repas. Cette diminution intervient bien avant la perte de poids, c'est-à-dire avant un rétablissement potentiel de la résistance à l'insuline. Cette chirurgie permet un découplage partiel du flux de suc digestif du transit du *bolus alimentaire* (masse d'aliments ingérés), ce qui conduit structurellement à une *diminution de l'absorption des aliments par l'intestin grêle*. En effet, l'intestin grêle, et plus précisément les *entérocytes* (cellules épithéliales de l'intestin impliquées dans l'absorption du glucose), utilisent un transporteur sodium-dépendant, communément appelé *SGLT1* ou *SLC5A1*, pour transférer le glucose de la lumière de l'intestin vers la veine porte (Baud et al., 2016a). La principale source de sodium facilitant le transfert du glucose est le suc pancréatique. Ce niveau de compréhension du transport du glucose de l'intestin grêle vers le sang pourrait expliquer la récupération de la concentration sanguine de glucose après une chirurgie de Roux-en-Y.

Ainsi, par corrélation, nous pouvons émettre l'hypothèse que la dynamique d'absorption intestinale du glucose, plus généralement appelée *absorption intestinale du glucose (AIG)*, serait un facteur causal majeur du DT2. Cependant, cette affirmation n'est pas communément acceptée, car de nombreux facteurs de confusion subsistent et sont mieux étudiés, tels qu'une diminution de l'activité métabolique pancréatique et hépatique ou une activité de *vidange gastrique* (expulsion du contenu de l'estomac vers l'intestin) anormale (Das et al., 2014; Goyal et al., 2019; Kumar, 2018). Les travaux réalisés dans cette thèse ont pour objectif de contribuer à clarifier le rôle de l'AIG dans le DT2 par le biais de la modélisation computationnelle. Le

mécanisme d'absorption étant directement relié à la dynamique du glucose dans le sang après un repas, c'est sur cet aspect que se concentreront nos investigations. Dans la suite de cette introduction, nous nous intéressons aux méthodes de mesure expérimentale de l'AIG ainsi qu'à la modélisation de la dynamique *postprandiale* (faisant suite à un repas) de glucose.

1.3 Suivi et prédiction de l'absorption intestinale de glucose après un repas

De nombreux indices, du type de ceux présentés précédemment, ont été proposés (Muniyappa et al., 2015; Patarrão et al., 2014) mais, lorsqu'ils sont basés sur les mesures à jeun et des tests par voie intraveineuse, ils ne rendent évidemment pas compte des caractéristiques du trac gastro-intestinal. Par ailleurs, il n'offrent pas la possibilité à l'observateur de discriminer avec certitude le glucose provenant du repas ingéré (dit glucose *exogène*) du glucose issu de l'activité hépatique (dit glucose *endogène*) survenant également après un repas.

Le *gold standard* pour atteindre cet objectif implique d'utiliser des techniques de doubles ou triples traceurs, consistant à introduire du glucose isotopique dans le sujet (Gonzalez and King, 2022), permettant ainsi de distinguer le glucose exogène (donc marqué dans le cas de cette expérience) du glucose endogène (hépatique, adipeux etc). La figure 1.2 illustre cette technique.

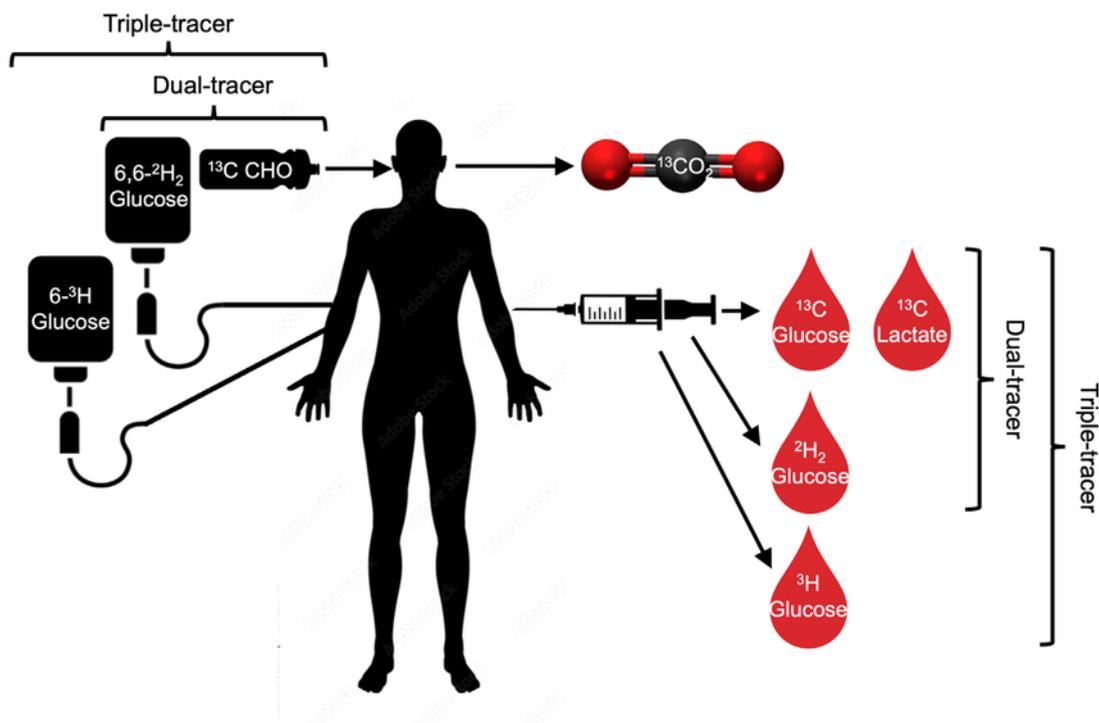


FIGURE 1.2 – Exemple d'administration de glucoses marqués isotopiquement. Source : (Gonzalez and King, 2022)

La méthode de suivi de glucoses isotopiques est utilisée dans de nombreuses applications de recherche, notamment pour étudier le métabolisme du glucose chez les personnes atteintes

de diabète, d'obésité, ou d'autres troubles métaboliques, ainsi que pour évaluer l'efficacité des traitements pharmacologiques ou des interventions nutritionnelles sur le métabolisme du glucose. Elle permet d'obtenir des informations détaillées sur les voies métaboliques impliquées dans le métabolisme du glucose et peut aider à mieux comprendre les mécanismes sous-jacents des maladies métaboliques. En ce qui concerne la problématique de cette thèse, ladite méthode permet à l'expérimentateur d'observer directement l'AIG. Cependant, cette technique est difficile à mettre en place, expérimentalement et encore plus cliniquement, car très intrusive et coûteuse, même pour une taille de cohorte raisonnable. De plus, les compétences requises pour mettre en œuvre une telle méthode sont rares, c'est pourquoi très peu d'équipes de recherche clinique l'ont réellement mise en œuvre.

Dans la lignée du modèle minimal, des modèles plus détaillés ont été développés pour exploiter les données expérimentales de multi-traceurs. Ils permettent, notamment, de préciser davantage la source, les flux et les différentes interactions de l'insuline dans l'organisme pour la régulation du glucose, et contribuent à démêler le paysage des hypothèses et des variables confondantes. Parmi ces modèles, l'un des plus référencés est le modèle homéostatique du glucose et de l'insuline proposé par Dalla Man et ses collègues (Dalla Man et al., 2007) avec la particularité d'introduire un sous-modèle de vidange gastrique original et complexe (Dalla Man et al., 2006a) mais avec un sous-modèle d'absorption intestinal très simplifié. À contrario, Salinari et ses collègues (Salinari et al., 2011) ont fait le choix inverse, permettant d'étudier, par exemple, l'impact que pourrait avoir une répartition non uniforme des entérocytes le long de l'intestin sur la dynamique d'absorption. Ces modèles seront présentés plus en détail dans ce manuscrit au chapitre 2. Ils présentent généralement l'inconvénient d'être complexes par leur nombre de variables et de paramètres. Leur calibration, tout particulièrement en l'absence de données issues d'expérience de multi-traceurs, ne permet pas de connaître avec certitude les valeurs des paramètres physiologiques et plus particulièrement de ceux du trac gastro-intestinal.

1.4 Approches de modélisation

Pour suivre l'évolution de nos molécules d'intérêt dans l'organisme au cours du temps, en dehors des expériences de mesures directes, plusieurs approches de modélisation sont possibles. Les plus courantes sont les approches mathématiques reposant sur les systèmes d'*équations différentielles ordinaires* (EDO en français, ODE en anglais). Ces systèmes d'équations décrivent le système dynamique de façon *déterministe*, c'est-à-dire qu'il n'y a pas de prise en compte d'éventuelles variabilités de la dynamique dues à des événements aléatoires intrinsèques ou extrinsèques. Au regard du type de données dont nous disposons, et de l'échelle macroscopique de modélisation qui nous intéresse, l'approximation déterministe est justifiée. À contrario, si, par exemple, nous modélisons des systèmes de régulation génétique à l'échelle cellulaire, alors le passage à des méthodes *stochastiques* serait requis.

L'approche de modélisation par EDO est souvent qualifiée de *boite noire* dans le sens où elle ne vise « qu'à » reproduire des données expérimentales et à les prédire. Dans ce cas, il

s'agit essentiellement de caractériser un système par ses entrées-sorties indépendamment de la façon dont celles-ci sont réellement produites. La modélisation boîte noire est particulièrement utilisée pour des systèmes biologiques particulièrement complexes et dont le fonctionnement détaillé n'est pas connu.

En biologie des systèmes, une autre approche, dite *mécanistique*, est souvent privilégiée sur l'approche boîte noire. Dans ce type de modèles, il s'agit en effet d'essayer de représenter, plus ou moins fidèlement suivant le niveau d'abstraction désiré, les caractéristiques physiologiques d'un système biologique et les processus qu'il met en oeuvre en termes d'interactions et de flux. Il nécessite bien sûr des connaissances *a priori* sur le système étudié, ou de faire des hypothèses sur son fonctionnement.

Le modèle minimal de Bergman (Bergman et al., 1979) est un modèle d'EDO mécanistiques qui a été réduit pour ne contenir que 4 paramètres. Le modèle homéostatique de Dalla Man (Dalla Man et al., 2007) est un modèle d'EDO à la fois boîte noire et physiologique. En effet, certaines parties représentent directement des flux de glucoses et d'insuline entre organes alors que, par exemple, la dynamique de sécrétion de l'insuline est modélisée par des équations complexes sans interprétation physiologique.

Lorsque des aspects spatiaux entrent en jeu, il convient d'utiliser une autre forme d'équations différentielles : les *équations aux dérivées partielles* (EDP en français, ou PDE en anglais). Ce type d'équations ajoute un degré de complexité supplémentaire car leur résolution et leur analyse sont plus complexes que pour les EDO. Le modèle de trac gastro-intestinal de Salinari (Salinari et al., 2011) est un exemple de modèle EDP mécanistique.

Une fois le modèle d'EDO ou d'EDP réalisé, les valeurs de ses paramètres doivent être déterminées. Pour les paramètres ayant une interprétation physiologique, le plus simple est de se référer à la littérature. Cependant, cette méthode mène souvent à intégrer, dans un même modèle, des valeurs de paramètres déterminées pour des organismes vivant différents et obtenus dans des conditions expérimentales très hétérogènes ce qui prête à discussion. Alternativement, des expériences peuvent être réalisées pour observer les variables d'intérêt et collecter des données quantitatives. Le modèle est ensuite calibré afin de simuler le plus fidèlement possible les données expérimentales. La calibration est un exercice qui peut s'avérer très difficile et nécessitant des données qu'il n'est pas toujours possible de collecter en pratique, en particulier dans le domaine de la santé. Nous reviendrons plus en détail sur ces questions dans le chapitre 2.

Dans cette thèse, nous utilisons un langage de modélisation alternatif aux EDO : les *réseaux de réactions chimiques* (CRN). Il s'agit d'un formalisme permettant de représenter un modèle à l'aide de réactions « chimiques ». Ces réactions transforment des espèces chimiques abstraites, appelées *réactifs*, en d'autres espèces chimiques abstraites, appelées *produits*. Elles sont associées à une *cinétique* qui détermine la vitesse à laquelle la réaction a lieu. De tels modèles sont, par nature, mécanistiques, car ils décrivent des interactions (physiologiques) entre des entités biologiques ou des flux entre compartiments biologiques. Ils peuvent être qualifiés de *méta-modèles* dans le sens où on peut leur associer (on parle alors d'*interprétation*) des dynamiques de différents types :

- déterministe, en dérivant un modèle d'EDO à partir des réactions,
- stochastique, en dérivant par exemple des chaînes de Markov à temps continu (CTMC),
- non-déterministe, en ne tenant compte que de la présence ou l'absence des espèces chimiques et en ignorant des cinétiques potentiellement inconnues.

1.5 Problématique

Les travaux sur le diabète de types 2 (DT2) se sont jusqu'à présent très majoritairement focalisés sur le rôle de la fonction bêta du pancréas et de la sensibilité à l'insuline. De nombreux indices, de précision et de pertinence variables, ont été proposés pour les mesurer. Ces indices sont calculés à partir de modèles plus ou moins complexes utilisant des données statiques de glycémie à jeun, ou des données dynamiques de test glycémique oral.

La chirurgie bariatrique a mis en évidence l'existence d'un troisième paramètre pouvant potentiellement être une cause de DT2 : l'absorption intestinale de glucose. Contrairement à la fonction bêta du pancréas et à la sensibilité à l'insuline, aucun indice n'a encore été proposé pour mesurer l'effet de ce paramètre sur le DT2. Mesurer expérimentalement l'absorption intestinale de glucose nécessite l'accès à la veine porte, ce qui est en pratique impossible sur les humains. Une technique expérimentale de multi-traceur utilisant le glucose marqué a été proposée en alternative mais celle-ci demeure très difficile à réaliser et demande une expertise qui ne permet pas son utilisation en routine clinique. Notons aussi que les approches de modélisation jusqu'à présent proposées pour prédire la réponse postprandiale de glucose nécessitent ce *gold standard*. Les rares modèles existants ne sont que partiellement mécanistiques et sont relativement complexes.

Un autre souci qui se pose lors de la construction d'un tel modèle est le caractère confondant de la vidange gastrique et de l'absorption intestinale. En effet, que l'on ait accès à la quantité de glucose exogène (par exemple par le biais de méthodes multi-traceurs) ou seulement à la glycémie totale (comprenant à la fois le glucose endogène et exogène), l'apparition du glucose exogène dépend bien de ces deux mécanismes. En effet, si l'apparition de glucose exogène dans le sang est une conséquence directe de l'absorption intestinale (*via* les entérocytes), cette absorption ne peut s'effectuer qu'après la vidange gastrique qui déverse progressivement le bolus alimentaire de l'estomac vers l'intestin. Il est donc pertinent de se demander si ces mécanismes contribuent de façon égale ou si l'un des deux prend le dessus. Comme nous le verrons dans la suite, les modèles existants se sont alternativement concentrés sur l'un ou l'autre de ces deux mécanismes.

Étant donnés ces constats, cette thèse tente de contribuer à l'*identification du rôle de l'absorption intestinale de glucose dans la survenue du diabète de type 2*. L'objectif, à long terme, est de proposer un indice permettant de diagnostiquer un défaut d'absorption intestinal susceptible d'entraîner un DT2. Notre démarche s'appuie donc sur la modélisation et sur des données qui puissent facilement être acquises en routine clinique. Dans l'immédiat, nous tentons de répondre plus précisément à ces questions :

- Un modèle informatique glucose-insuline de référence peut-il simuler l'ensemble de données cliniques que nous avons en pratique à notre disposition ? En effet, à notre connaissance, aucun modèle glucose-insuline n'a jamais été paramétré à partir de données de patients ayant subi une chirurgie bariatrique.
- Peut-on trouver une alternative à l'observation par multi-traceur qui soit simple d'utilisation tout en fournissant suffisamment d'information sur l'absorption intestinale ?
- Comment modéliser la réponse postprandiale de glucose dans une approche mécanistique qui rende compte des caractéristiques de l'intestin ?
- Un tel modèle permet-il de discriminer l'action de la vidange gastrique de celle de l'absorption intestinale sur la dynamique postprandiale de la glycémie ? En effet, il faut pouvoir distinguer ces variables confondantes pour mesurer l'AIG de façon non ambiguë.

1.6 Contributions

Pour répondre aux questions de la problématique, dans un premier temps, nous reproduisons le modèle post-prandial de Dalla Man et les simulations de l'article de référence (Dalla Man et al., 2007). Ce modèle étant décrit exclusivement à l'aide d'EDO, nous l'avons partiellement retranscrit en un système de réactions chimiques afin de mettre en perspective les aspects relevant de mécanismes physiologiques. Cette tâche ardue ne peut être automatisée car le système d'EDO est trop complexe et certaines parties « boîte noire » ont dû être directement intégrées dans les cinétiques de certaines réactions.

Des simulations des EDO qui dérivent de notre CRN ont été réalisées pour confirmer une bonne reproductibilité des simulations de l'article original lorsque les paramètres sont instanciés avec ceux donnés dans l'article (à savoir, paramètres moyens estimés à partir d'individus sains et paramètres moyens estimés à partir d'une cohorte de patients diabétiques non obèses).

Les paramètres de ce modèle ont ensuite été recalibrés à partir de nos propres données cliniques pour évaluer sa pertinence à simuler la chirurgie bariatrique et donc à étudier le rôle de l'AIG dans la réponse postprandiale. Ces travaux ont fait l'objet d'une publication en conférence internationale en 2021, à l'*International Conference on Bioinformatics, Biocomputational Systems and Biotechnologies* (Dursoniah et al., 2021).

En alternative au gold standard du multi-traceur, nous avons utilisé le D-xylose, un sucre analogue du glucose, comme biomarqueur afin d'observer directement l'AIG. Le D-xylose est un sucre naturel classé comme pentose, ce qui signifie qu'il possède cinq atomes de carbone dans sa structure moléculaire. C'est un monosaccharide, c'est-à-dire un sucre simple qui ne peut pas être hydrolysé davantage en sucres plus simples. Sa formule chimique est $C_5H_{10}O_5$. On le trouve dans de nombreuses plantes, notamment les fruits, les légumes et les fibres ligneuses. Contrairement aux sucres courants comme le glucose ou le fructose, le D-xylose n'est pas métabolisé efficacement dans le corps humain par les voies glycolytiques normales. En pra-

tique, cela signifie qu'il n'est presque pas consommé par les tissus et qu'il perdure dans le sang jusqu'à son élimination. Cela le rend donc utile pour certains tests médicaux liés à la fonction digestive, notamment pour tester le syndrome de malabsorption (Garcia-Naveiro and Udall Jr, 2011). Ici, nous proposons une utilisation originale de ce sucre : nous souhaitons notamment montrer dans cette thèse qu'il est possible d'établir un profil satisfaisant d'absorption intestinale du glucose d'un individu, à partir de ses données d'absorption de D-xylose, contournant ainsi le recours au gold standard par multi-traceur.

Nous avons réalisé le premier modèle de D-xylose à notre connaissance, diffusé dans une publication internationale. Ce modèle a été sélectionné par estimation de paramètres sur nos jeux de données, puis à l'aide d'une analyse d'identifiabilité pratique et d'une analyse de sensibilité globale. Ces analyses ont également permis d'étudier les contributions relatives de la vidange gastrique et de l'absorption intestinale sur le profil de la dynamique de D-xylose. Enfin nous nous interrogerons sur les liens entre la modélisation de la glycémie et celle de la réponse de *D-xylose* postprandial tout en envisageant les applications cliniques et les limites du modèle de D-xylose. Ces travaux ont été publiés en conférence internationale en 2024, à l'*International Conference on Bioinformatics Models, Methods and Algorithms* (Dursoniah et al., 2024).

Toutes les implantations de cette thèse ont été réalisées dans le langage de programmation Julia (Bezanson et al., 2017). Il s'agit d'un langage de programmation *open source* récent principalement développé au MIT. Il propose de très nombreux packages spécialisés dans les sciences des données, l'apprentissage machine, la programmation scientifique et la modélisation, en particulier de systèmes biologiques (Roesch et al., 2023). Les programmes Julia sont compilés et s'exécutent de façon très efficace ce qui est une caractéristique importante par exemple pour la calibration de modèles. C'est une des raisons pour lesquelles Julia est de plus en plus utilisé en alternative à Matlab, R et Python.

Dans la suite de cette thèse, nous dressons un état de l'art des méthodes de modélisation et d'analyse de modèles en relation avec notre sujet au chapitre 2. Nous y détaillons également les principaux modèles qui nous seront utiles pour nos contributions. Dans le chapitre 3, nous décrivons les travaux réalisés dans l'analyse du modèle de Dalla Man pour l'étude de l'absorption intestinale. Notre proposition de modèle reposant sur l'exploitation de données de D-xylose est décrite au chapitre 4. Nous concluons cette thèse au chapitre 5 et ouvrirons quelques perspectives scientifiques.

Chapitre 2

Méthodologie et État de l'art

2.1 Méthodologie

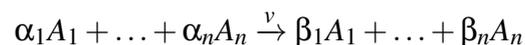
Cette section présente les formalismes et outils de modélisation utilisés dans le reste de ce manuscrit.

2.1.1 Formalismes

Pour formaliser mathématiquement un système biologique d'intérêt, plusieurs approches sont possibles. Les plus courantes sont les réseaux de réactions chimiques, les réseaux de Petri et les systèmes d'équations différentielles ordinaires.

Réseau de Réactions Chimiques

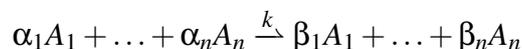
Un réseau de réactions chimiques consiste en un système de réactions interdépendantes, où chaque réaction provoque la conversion de certaines espèces chimiques (réactifs) en d'autres espèces (produits) selon une cinétique, spécifique pour chaque réaction (Feinberg, 2019). Ces réactions sont influencées par divers facteurs physico-chimiques (concentration initiale des espèces, température, présence de catalyseurs, etc). Plus formellement, un réseau de réactions chimiques peut être défini comme un ensemble $\{A_1, \dots, A_n\}$ de n espèces et un ensemble R de réactions, chacune convertissant un ensemble de réactifs en un ensemble de produits. De plus, une certaine stœchiométrie est associée à chaque réactif et chaque produit. On notera une réaction ainsi :



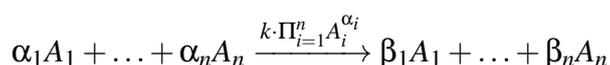
où les α_i et β_i sont les *coefficients stœchiométriques*, c'est-à-dire des nombres entiers naturels (éventuellement nuls), indiquant le nombre de molécules de chacune des espèces impliquées dans la réaction. v est une expression dépendant des concentrations de réactants, appelée *cinétique*, qui spécifie la vitesse de la réaction. La cinétique la plus courante associée à une réaction est issue de la *loi d'action de masse* qui est définie par l'expression

$$k \cdot \prod_{i=1}^n A_i^{\alpha_i}$$

où A_i représente ici la concentration de l'espèce A_i et k est un coefficient constant. Il est d'usage, lorsque la loi d'action de masse est sous-entendue, de n'indiquer que le coefficient k à la place de l'expression v dans la réaction. Pour éviter toute ambiguïté, on utilisera dans ce cas la flèche \rightarrow et on écrira simplement :



pour la réaction



y compris dans les cas où k n'est pas une constante mais une expression dépendant éventuellement de concentrations d'espèces qui ne sont pas consommées par la réaction. Une autre cinétique courante est la *cinétique de Michaelis-Menten* utilisée en particulier pour les réactions enzymatiques.

Il est également habituel de représenter un système de réactions chimiques à l'aide d'un *matrice de stœchiométrie*. Si nous avons n espèces chimiques et m réactions chimiques, alors la matrice de stœchiométrie S est une matrice $n \times m$ définie comme suit :

$$S = \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1m} \\ s_{21} & s_{22} & \cdots & s_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ s_{n1} & s_{n2} & \cdots & s_{nm} \end{bmatrix}$$

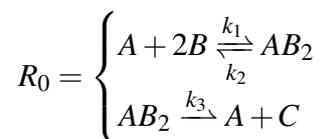
s_{ij} représente la stœchiométrie de la i -ème espèce chimique dans la j -ème réaction, c'est-à-dire la différence entre le nombre de molécules A_i en produits et le nombre de molécules A_i en réactants dans la j -ième réaction. Si la i -ème espèce chimique est davantage impliquée en tant que réactif, alors sa stœchiométrie est négative, sinon elle est positive. Notons que cette matrice n'est pas strictement équivalente à la notation précédente. En effet, pour les réactions



la stœchiométrie de A est 0 alors que la première réaction ne peut s'exercer qu'en présence de A .

Graphiquement, un réseau de réactions chimiques peut aussi être représenté par un hypergraphe dirigé, où chaque hyperarc est étiqueté par le coefficient stœchiométrique correspondant.

Exemple Définissons un réseau de 3 réactions chimiques, impliquant 3 espèces chimiques A , B et C :



Dans cet exemple, le réseau comporte 3 réactions chimiques. Dans les deux premières, réversibles l'une de l'autre, on considère une enzyme A se complexifiant à 2 substrats B . Dans la

troisième réaction, le complexe stable ainsi formé, AB_2 , permet à l'enzyme A de catalyser B_2 en produit final C , libéré de façon irréversible. On en déduit la matrice de stœchiométrie suivante :

$$S_0 = \begin{bmatrix} -1 & 1 & 1 \\ -2 & 2 & 0 \\ 1 & -1 & -1 \\ 0 & 0 & 1 \end{bmatrix}$$

Dans cette matrice, les lignes représentent respectivement les espèces A , B , AB_2 et C de haut en bas et les colonnes représentent les réactions.

Liens entre réseaux de réactions chimiques et d'autres types de réseaux Afin de capturer les dynamiques quantitatives d'un tel réseau, celui-ci peut être traduit, soit en un réseau de Petri (que l'on définira dans la sous-section suivante), soit en un système d'équations différentielles, qui expriment au moins la variation temporelle des concentrations des différentes espèces chimiques impliquées de façon continue. Dans les deux représentations mathématiques, chaque équation du système représente la vitesse de variation des concentrations des espèces en fonction de leurs concentrations et des constantes cinétiques associées à chaque réaction. Généralement, on utilise les réseaux de Petri pour modéliser les variations de concentration avec des entiers (discrets) tandis que les EDO représentent les variations avec des réels (continus).

Réseau de Petri

Les *réseaux de Petri* ont été définis par Carl Adam Petri dans les années 1960 (Petri, 1962). Ils fournissent une représentation mathématique des interactions entre différents composants ou entités au sein d'un système. On peut aussi les représenter graphiquement sous la forme répandue d'un graphe dirigé biparti, où une partie des nœuds représente les entités et l'autre partie représente les réactions consommant et produisant ces entités.

Les réseaux de Petri et les réseaux de réactions chimiques, bien qu'ils ont des formalismes distincts, voient leurs domaines d'application se chevaucher dans la modélisation de systèmes biochimiques pour la capture de propriétés telles que la concurrence et la synchronisation et des réactions chimiques (Chaouiya, 2007).

Formellement, un réseau de Petri est défini par un quadruplet d'ensembles (P, T, A, M_0) avec :

- P , un ensemble de *places* tel que $P = \{p_1, p_2, \dots, p_n\}$, chacune caractérisée par un nombre de jetons à une itération donnée ;
- T , un ensemble de *transitions* tel que $T = \{t_1, t_2, \dots, t_m\}$, représentant les événements induisant le changement d'état du système ;
- A , un ensemble d'*arcs* tel que $A \subseteq (P \times T) \cup (T \times P)$, représentant le flux de jetons entre les places et les transitions, chaque arc étant pondéré par une valeur entière positive, indiquant le nombre de jetons produits ou consommés par la transition. On peut représenter cette pondération sous la forme d'une fonction $p : A \rightarrow \mathbb{N}^+$;

— M_0 , l'état (ou *marquage*) initial du système, qui indique, pour chaque place p , le nombre initial $M_0(p)$ de jetons dans cette place.

Pour un marquage donné, on peut jouer une transition lorsque toutes les places qui l'alimentent ont un nombre de jetons supérieur ou égal à la pondération de l'arc. Si on joue cette transition depuis le marquage en question, on obtient un nouveau marquage. On parle d'évolution asynchrone lorsqu'on se limite à appliquer une seule transition à la fois, et d'évolution synchrone lorsqu'on applique plusieurs transitions en même temps.

Initialement, les réseaux de Petri manipulent des ensemble entiers de jetons dans les places. Il est possible d'étendre ce formalisme à des valeurs réelles de jetons (David and Alla, 1987). Si on ajoute à cela une évolution synchrone du modèle, où chaque transition possède une vitesse d'exécution k_1, k_2, \dots alors un tel réseau de Petri continu est équivalent à un réseau de réactions chimiques, tel que présenté dans la sous-section précédente. Un des avantages de ce rapprochement est de pouvoir utiliser la représentation graphique des réseaux de Petri pour pouvoir représenter les réseaux de réactions chimiques.

Exemple Reprenons le réseau de réactions chimiques R_0 défini dans la sous-section précédente : $A + 2B \xrightleftharpoons[k_2]{k_1} AB_2 \xrightarrow{k_3} A + C$. Une représentation graphique de ce réseau en réseau de Petri est montrée dans la figure 2.1. Les cercles représentent des places et les rectangle des transitions. Il est à noter que les coefficients stœchiométrie sont explicitement indiqués sur les arcs lorsque ceux-ci sont strictement supérieures à 1. Par défaut, lorsqu'aucune valeur de stœchimétrie n'est indiquée sur un arc, elle est égale à 1.

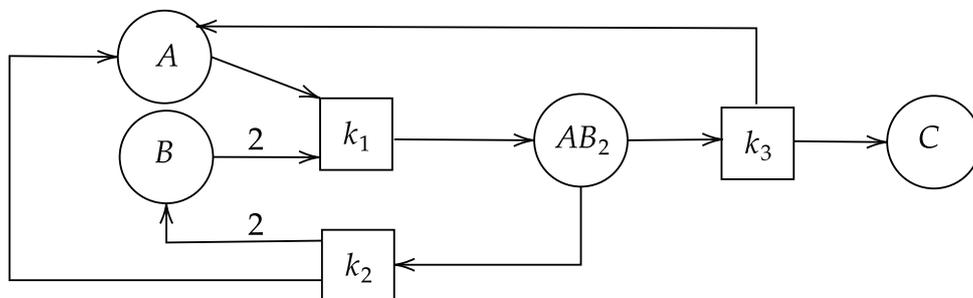


FIGURE 2.1 – Une représentation sous forme de réseau de Petri du réseau de réactions chimiques R_0 .

Système d'équations différentielles ordinaires

Un *système d'équations différentielles ordinaires* est un ensemble d'équations différentielles où les inconnues sont des fonctions dépendant d'une seule variable indépendante. Soit t la variable indépendante, et soient $y_1(t), y_2(t), \dots, y_n(t)$ les fonctions inconnues dépendant de t .

Un système d'EDO de n équations est alors représenté par :

$$\begin{cases} \frac{dy_1}{dt} = f_1(t, y_1, y_2, \dots, y_n) \\ \frac{dy_2}{dt} = f_2(t, y_1, y_2, \dots, y_n) \\ \vdots \\ \frac{dy_n}{dt} = f_n(t, y_1, y_2, \dots, y_n) \end{cases}$$

où f_1, f_2, \dots, f_n sont des fonctions. Les équations ci-dessus spécifient comment les dérivées des fonctions inconnues dépendent de la variable indépendante t et des fonctions inconnues elles-mêmes.

Afin de limiter le nombre de solutions possibles, un ensemble de conditions initiales peut aussi être fourni :

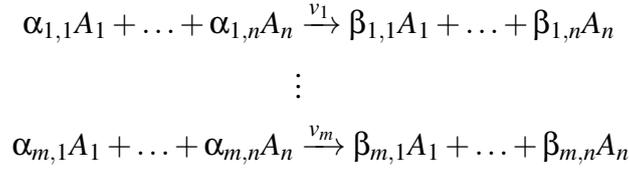
$$\begin{cases} y_1(0) = u_1 \\ y_2(0) = u_2 \\ \vdots \\ y_n(0) = u_n \end{cases}$$

où u_1, u_2, \dots, u_n sont des valeurs réelles.

La solution d'un système d'EDO est un ensemble de fonctions qui satisfont aux équations différentielles données et à toutes les conditions initiales spécifiées. Les systèmes d'EDO sont largement utilisés pour modéliser des systèmes dynamiques où le comportement du système dépend des taux de changement de certaines variables. Les exemples incluent la dynamique des réactions chimiques.

Tout comme les représentations plutôt « structurelles » que constituent les réseaux de Petri et les réseaux de réactions chimiques, les systèmes d'EDO sont des cadres de modélisation mathématique utilisés pour décrire et analyser des systèmes dynamiques. Alors que les systèmes d'EDO permettent de modéliser les taux de changement de variables continues au fil du temps, les réseaux de réactions chimiques, eux, se concentrent spécifiquement sur la modélisation des réactions au sein d'un système, dans le sens où ils se contentent de représenter quelles espèces interagissent entre elles sans être directement utilisable en tant qu'outil de suivi quantitatif des espèces impliquées. On peut toutefois observer un lien fort entre ces deux approches de modélisation, notamment dans la modélisation de la cinétique chimique et de la dynamique des réactions. En effet, toutes deux permettent de représenter une structure de réseau d'interactions biochimiques, des équations de vitesse, une cinétique d'action de masse, des états d'équilibre. De plus, toutes deux associent une matrice de stœchiométrie. Du fait de ces liens et de la possibilité d'interpréter intuitivement et directement les réactions chimiques, c'est le formalisme de systèmes d'EDO qui a été choisi dans le cadre de ces travaux de modélisation. Les réseaux de réactions chimiques seront cependant parfois utilisés pour proposer une représentation alternative structurelle des modèles étudiés, tout en étant représentés graphiquement sous la forme de réseaux de Petri.

La traduction d'un réseau de réactions chimiques en un système d'EDO est défini comme suit. Soit le réseau de réactions chimique :



Le taux de production (resp. consommation) d'une espèce A est la donné par la somme des vitesses des réactions qui produisent (resp. consomment) A pondérées par ses stœchiométries. Il en découle que l'interprétation du réseau de réactions précédent en termes d'EDO est donné par le système suivant :

$$\begin{aligned} \frac{dA_1(t)}{dt} &= \sum_{i=1}^m \beta_{i,1}v_i - \sum_{i=1}^m \alpha_{i,1}v_i \\ &\vdots \\ \frac{dA_n(t)}{dt} &= \sum_{i=1}^m \beta_{i,n}v_i - \sum_{i=1}^m \alpha_{i,n}v_i \end{aligned}$$

Exemple Reprenons le système de réactions R_0 . Comme dit précédemment, il est possible de déduire le système d'équations différentielles ordinaires directement à partir d'un réseau de réactions chimiques ou d'un réseau de Petri. Ainsi le système d'EDO de R_0 serait :

$$\begin{cases} \frac{dA(t)}{dt} &= (k_2 + k_3) \cdot AB_2(t) - k_1 \cdot A(t) \cdot B(t)^2 \\ \frac{dB(t)}{dt} &= 2(k_2 \cdot AB_2(t) - k_1 \cdot A(t) \cdot B(t)^2) \\ \frac{dAB_2(t)}{dt} &= k_1 \cdot A(t) \cdot B(t)^2 - (k_2 + k_3) \cdot AB_2(t) \\ \frac{dC(t)}{dt} &= k_3 \cdot AB_2(t) \end{cases}$$

2.1.2 Estimation des paramètres

L'estimation des paramètres pour les systèmes d'équations différentielles ordinaires (EDO) est le processus de détermination des valeurs de paramètres inconnus dans le système sur la base des données observées. Dans le contexte des systèmes d'EDO, les paramètres incluent souvent des constantes de vitesse, des conditions initiales et d'autres coefficients qui caractérisent le comportement du système.

L'inférence de paramètres implique la définition d'une fonction objectif (également appelée fonction de coût, dont des exemples sont donnés dans la sous-section suivante) qui quantifie la différence entre les prédictions du modèle (basées sur les estimations des paramètres actuels) et les données observées. L'une des fonctions objectifs les plus couramment utilisées est la somme des carrés des différences (*MSE*) entre les prédictions du modèle et les données expérimentales.

La recherche itérative des valeurs de paramètres nécessite donc un algorithme d'optimisation qui utilise cette fonction objectif. Cet algorithme va chercher à minimiser la fonction de

coût en choisissant les meilleures valeurs de paramètres au regard des données expérimentales. Il est possible d'améliorer cette estimation par l'ajout de contraintes, qui vont réduire l'espace des solutions des valeurs de paramètres. Un exemple de contraintes utilisées est de se limiter à un intervalle communément admis comme raisonnable pour certains paramètres ou d'effectuer l'inférence sur plusieurs jeux de données pour lesquels nous admettons que le modèles partagent les mêmes paramètres pour ces observations.

Au-delà du travail d'estimation, il est important d'évaluer quantitativement l'incertitude des paramètres estimés : typiquement cette évaluation s'exprime par l'estimation d'intervalles de confiance sur les paramètres ou par une fonction de coût sous la forme d'une fonction de vraisemblance. Il est à noter que la qualité de l'inférence de paramètres est limitée par la complexité du modèle étudié (nombre de variables, nombre de paramètres etc.) et par la qualité des données expérimentales (taille, présence d'outliers etc.).

Fonctions de coût

Plusieurs fonctions de coût existent pour sélectionner le bon modèle. Plus précisément, nous avons dû choisir entre une fonction L1, une fonction L2 et une fonction « Negative Log-likelihood ».

Fonction de coût L1 La fonction de perte L1, également connue sous le nom de perte absolue ou erreur absolue moyenne (MAE), est une première mesure couramment utilisée de la différence entre les valeurs prédites et les valeurs réelles (cibles) dans les problèmes de régression. L1 est définie comme la moyenne des différences absolues entre chaque valeur prédite (\hat{y}_i) et sa valeur réelle correspondante (y_i) :

$$L_1(\theta) = \frac{1}{N} \sum_{i=1}^N |\hat{y}_i - y_i|$$

avec :

- θ : ensemble des paramètres du modèle.
- N : nombre de points de données par individu.
- \hat{y}_i : valeur prédite pour le i -ème point de données.
- y_i : valeur réelle (cible) du i -ème point de données.

L1 pénalisant les différences absolues, elle est particulièrement robuste aux valeurs aberrantes, comparés à la fonction L2, définie dans le paragraphe suivant.

Fonction de coût L2 La fonction de perte L2, également connue sous le nom de perte quadratique ou erreur quadratique moyenne (MSE), est une autre mesure très courante de la différence entre les valeurs prédites et les valeurs cibles. Elle est largement utilisée dans l'apprentissage automatique, les statistiques et l'optimisation. La fonction de perte L2 est définie comme la moyenne des carrés des différences entre chaque valeur prédite (\hat{y}_i) et sa valeur réelle correspondante (y_i) :

$$L_2(\theta) = \frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i)^2$$

La minimisation des carrés des différences pénalise davantage les erreurs importantes que les erreurs plus petites, ce qui rend le modèle plus sensible aux valeurs aberrantes.

Fonction Negative Log-Likelihood La fonction Negative Log-Likelihood (NLL), également connue sous le nom de perte de log-vraisemblance, est une fonction objectif courante utilisée dans de nombreux types de modèles, probabilistes ou déterministes, afin d'apporter une estimation du maximum de vraisemblance sur la distribution des données au regard des paramètres inférés. Ainsi, l'optimiseur qu'on choisira d'associer à cette fonction cherchera à maximiser la probabilité des données observées dans le cadre du modèle. Considérons un scénario dans lequel nous avons un ensemble de points de données observés $\{y_1, y_2, \dots, y_N\}$, et nous avons un modèle probabiliste qui prédit la distribution de probabilité $P(\theta)$ pour les données, où θ représente les paramètres du modèle. La vraisemblance des données observées compte tenu des paramètres du modèle est représentée par le produit des probabilités attribuées à chaque point de données :

$$\text{Probabilité}(\theta) = \prod_{i=1}^N P(y_i | \theta)$$

La fonction de log-vraisemblance négative est alors définie comme la négation du logarithme naturel de la vraisemblance :

$$\text{NLL}(\theta) = -\log(\text{Probabilité}(\theta))$$

Pour des raisons mathématiques, il est souvent plus courant de travailler directement avec la log-vraisemblance, car elle convertit les produits (de la vraisemblance) en sommes :

$$\text{NLL}(\theta) = -\sum_{i=1}^N \log(P(y_i | \theta))$$

Le création du modèle probabiliste dans notre fonction de coût est de trouver les valeurs des paramètres (θ) qui maximisent la vraisemblance (ou de manière équivalente, minimisent la log-vraisemblance négative). Maximiser la vraisemblance signifie rendre les points de données observés plus probables dans le cadre du modèle. Ayant des données expérimentales sur plusieurs individus, l'utilisation de cette fonction s'est avéré pratique pour estimer nos paramètres sur notre distribution de données expérimentales (que nous avons considérés comme normale), au regard de la variabilité interindividuelle pour chaque point de données temporelles, comme montrée dans la suite de ce manuscrit.

Algorithmes d'optimisation

Présentation générale L'optimisation des paramètres d'un système d'EDO implique de trouver les valeurs qui minimisent ou maximisent une fonction objectif, souvent une mesure de

la différence entre les prédictions du modèle et les données observées. Plusieurs algorithmes d'optimisation peuvent être utilisés pour l'estimation des paramètres d'un système d'EDO :

- Descente de gradient : Il s'agit d'un algorithme d'optimisation répandu et très classique qui met à jour de manière itérative les valeurs des paramètres dans la direction du gradient négatif de la fonction objectif (Haji and Abdulazeez, 2021). Ce processus peut être lent pour des problèmes complexes et de grande dimension.
- Algorithme de Levenberg-Marquardt : Cet algorithme est couramment utilisé pour les problèmes de moindres carrés non linéaires, tels que ceux découlant de l'estimation des paramètres du système d'EDO (Levenberg, 1944). Il combine des aspects de descente de gradient et des méthodes de Gauss-Newton.
- Algorithme évolutionnaire : Un algorithme évolutionnaire, ou évolutif, utilise des mécanismes inspirés par l'évolution biologique, tels que la reproduction, la mutation, la recombinaison et la sélection. Les solutions candidates au problème d'optimisation jouent le rôle d'individus dans une population, et la fonction de fitness détermine la qualité des solutions. L'évolution de la population se produit alors après l'application répétée des solutions candidates de l'itération précédente (Vikhar, 2016).
- Optimisation bayésienne : L'optimisation bayésienne est une technique d'optimisation basée sur un modèle probabiliste (Mockus, 2005). Il construit un modèle de substitution de la fonction objectif et sélectionne de manière itérative de nouveaux points à évaluer sur la base d'une fonction d'acquisition.

Lors du choix d'un algorithme d'optimisation pour l'estimation des paramètres du système EDO, les considérations incluent la nature de la fonction objectif, la dimensionnalité de l'espace des paramètres, la présence de contraintes et les ressources de calcul disponibles. Il est courant d'expérimenter plusieurs algorithmes pour trouver celui qui fonctionne le mieux pour un problème spécifique. De plus, la combinaison d'algorithmes d'optimisation avec des techniques telles que la réduction de modèle ou l'analyse de sensibilité peut améliorer l'efficacité de l'estimation des paramètres pour les systèmes d'EDO.

Optimiseur CMA-ES Dans la classe des algorithmes évolutionnaires se trouve l'approche appelée Covariance Matrix Adaptation Evolution Strategy ou CMA-ES (Hansen and Ostermeier, 1996) qui est très efficace pour l'estimation des paramètres dans les systèmes d'EDO (Hansen and Ostermeier, 2001; Hansen, 2023). En effet, les systèmes EDO impliquent souvent des relations complexes et non linéaires entre les paramètres et les données observées, et trouver l'ensemble optimal de paramètres qui correspond le mieux au modèle EDO aux données est un problème d'optimisation difficile. CMA-ES, avec sa capacité à gérer des problèmes d'optimisation de grande dimension, non linéaires et bruyants, peut être un choix approprié pour l'estimation des paramètres dans les systèmes EDO. Pour cela, cette approche génère des jeux de paramètres candidats qui sont ensuite croisés au fil des itérations, de façon à chercher des candidats qui minimisent la fonction de coût, comme pour tout algorithme évolutionnaire. Sa particularité est de reposer sur une façon particulière de mettre à jour la matrice de covariance,

élément central qui représente les variances entre chaque paire de paramètres.

De par nos différentes interactions avec d'autres modélisateurs et diverses références littéraires (Hansen and Ostermeier, 2001), notre choix s'est porté sur cet algorithme d'optimisation.

2.1.3 Analyses pour la validation de modèles

Cette section propose deux analyses à effectuer sur des modèles après identification des paramètres : l'analyse d'identifiabilité et l'analyse de sensibilité. Comme détaillé dans la suite, l'analyse d'identifiabilité permet de déterminer la pertinence du jeu de paramètres choisi en testant si d'autres jeux de paramètres pourraient aussi convenir, ce qui permet par exemple de sélectionner un modèle plus robuste parmi un ensemble de modèles. De son côté, l'analyse de sensibilité teste la sensibilité du modèle aux variations de valeurs des paramètres choisis, ce qui peut être un indice sur la qualité du modèle, mais est aussi utilisé dans la suite de ce manuscrit pour évaluer l'importance relative des variations des paramètres d'un modèle sur les valeurs des simulations.

Analyse pratique de l'identifiabilité

L'analyse d'identifiabilité est une étape importante dans la modélisation des systèmes dynamiques décrits par des équations différentielles. Dans le contexte des équations différentielles, un système peut être décrit par un ensemble de variables d'état, qui sont des fonctions qu'on suppose connues à ce stade, mais qui dépendent de la valeur de paramètres, qui sont des nombres réels. L'objectif de l'analyse d'identifiabilité est de déterminer si ces paramètres peuvent être estimés avec précision à partir des données expérimentales.

L'analyse d'identifiabilité peut être réalisée de différentes manières, notamment en utilisant des techniques mathématiques telles que l'analyse de sensibilité ou l'analyse de la variance. Certaines approches ne dépendent que du modèle et de sa structure, tandis que d'autres se reposent aussi sur les données disponibles pour calibrer le modèle. Ces méthodes permettent d'évaluer à divers degrés la sensibilité des sorties du système par rapport aux variations des paramètres, et ainsi de déterminer si ces paramètres peuvent être estimés de manière unique.

Dans ce manuscrit, nous nous intéressons à l'analyse d'identifiabilité *pratique* qui permet d'évaluer s'il existe plusieurs ensembles de valeurs de paramètres qui pourraient produire des résultats similaires ou identiques dans le cadre des observations disponibles. Dit autrement, dans un tel cas, il est difficile, voire impossible, de déterminer les valeurs réelles des paramètres à partir des données expérimentales, et le modèle est dit non identifiable. À l'inverse, si un seul ensemble de paramètres correspond de façon satisfaisante aux données expérimentales, alors le modèle est dit identifiable. De façon plus précise, nous nous intéressons ici à l'analyse d'identifiabilité *pratique locale*, basée sur la méthode du profil de vraisemblance (*profile likelihood method*) (Raue et al., 2009). Son principe consiste à explorer localement les valeurs des paramètres, c'est-à-dire à faire varier ces valeurs autour des valeurs précédemment identifiées. Pour chaque variation, une nouvelle estimation de tous les autres paramètres est effectuée afin

d'observer les variations de la fonction de coût. Si, pour chaque variation qui s'éloigne de la valeur initiale d'un paramètre, la fonction de coût augmente, alors cela signifie que le paramètre était identifiable, car sa variation ne peut pas être compensée par les autres paramètres du modèle. À l'inverse, si la fonction de coût n'augmente pas sensiblement (au-dessus d'un certain seuil) alors le paramètre n'est que partiellement identifiable (s'il n'est identifiable que d'un seul côté) ou non identifiable (des deux côtés). Un modèle est considéré comme robuste si tous ses paramètres sont localement identifiables en pratique.

Analyse de sensibilité globale

Parmi les différentes définitions de l'analyse de sensibilité présentées dans la littérature, celle proposée par Saltelli et al. (Saltelli et al., 2004) est la plus largement utilisée : il s'agit de « l'étude de la manière dont l'incertitude dans les résultats d'un modèle (numérique ou autre) peut être attribuée à différentes sources d'incertitude dans les données d'entrée du modèle ». En d'autres termes, l'analyse de sensibilité explore la relation entre les variables d'entrée et les variables de sortie du modèle.

Les techniques d'analyse de sensibilité se divisent en deux catégories : locale et globale. L'analyse de sensibilité locale est réalisée en faisant varier les paramètres du modèle autour de valeurs de référence spécifiques, dans le but d'explorer comment de petites perturbations des entrées influencent les performances du modèle. Cette approche a été largement utilisée dans la littérature, mais présente d'importantes limitations. Si le modèle n'est pas linéaire, les résultats de l'analyse de sensibilité locale peuvent être fortement biaisés, car ils sont fortement influencés par les hypothèses d'indépendance et une exploration limitée des entrées du modèle. Si les facteurs du modèle interagissent, l'analyse de sensibilité locale sous-estimera leur importance, car elle ne prend pas en compte ces effets. En général, comme l'analyse de sensibilité locale n'explore qu'en partie l'espace paramétrique d'un modèle, elle n'est pas considérée comme une approche valide pour les modèles non linéaires. Dans un modèle linéaire sans interactions entre les variables d'entrée, l'analyse de sensibilité locale peut produire des indices de sensibilité appropriés, tels que les indices de Sobol (Sobol, 1993). Cependant, si les facteurs interagissent, la considération partielle de l'espace ne peut pas correctement rendre compte des effets de chaque facteur sur la réponse du modèle, car elle n'est informative qu'à la valeur de référence où elle est appliquée.

En revanche, une analyse de sensibilité globale fait varier les facteurs incertains dans l'ensemble de l'espace faisable des réponses variables du modèle. Cette approche révèle les effets globaux de chaque paramètre sur la sortie du modèle, y compris les effets interactifs. Pour les modèles qui ne peuvent pas être prouvés linéaires, l'analyse de sensibilité globale est préférée.

Dans cette thèse, le choix s'est porté sur l'indice de Sobol, qui est une forme d'analyse de sensibilité globale. Calculé dans un cadre probabiliste, il décompose la variance de la sortie du modèle ou du système en fractions qui peuvent être attribuées aux entrées ou aux ensembles d'entrées. Une mesure directe de la sensibilité basée sur la variance S_i , appelée « indice de

sensibilité de premier ordre » ou « indice d'effet principal », est définie par :

$$S_i = \frac{V_i}{\text{Var}(Y)}$$

Cette équation représente la contribution à la variance de la sortie de l'effet principal de l'input d'un modèle, noté X_i (qui sera un paramètre dans le cadre d'un système d'EDO). Cela permet ainsi de mesurer l'effet de la variation de X_i seul, mais moyenné sur les variations des autres paramètres d'entrée. Il est standardisé par la variance totale pour fournir une contribution fractionnelle.

2.2 État de l'art des modèles de glucose-insuline

Cette section présente un état de l'art des principaux modèles d'absorption du glucose et de sa régulation par l'insuline. Ces modèles sont présentés dans l'ordre chronologique, ce qui rend compte des améliorations successives proposées par les différents auteurs. Dans la suite de ce manuscrit, certains de ces modèles seront analysés pour en montrer les limites (chapitre 3) ou serviront d'inspiration pour la création d'un nouveau modèle (chapitre 4).

2.2.1 Modèle minimal de Bergman

Le modèle minimal de Bergman, également connu sous le nom de modèle minimal de régulation du glucose, est un modèle mathématique développé par le Dr Richard N. Bergman et ses collègues dans les années 1970 pour décrire la dynamique du glucose et de l'insuline dans le corps humain, en particulier dans le contexte de métabolisme du glucose et sensibilité à l'insuline (Bergman et al., 1979; Bergman, 1989). Ce modèle a été largement utilisé dans la recherche pour évaluer la sensibilité à l'insuline et la fonction des cellules bêta, en particulier chez les personnes atteintes ou risquant de développer un diabète de type 2. Le modèle vise plus précisément à caractériser la relation entre la dynamique du glucose et de l'insuline en réponse à un test de tolérance au glucose oral, aussi appelé test de provocation au glucose. Pour rappel, ce test consiste à provoquer une hyperglycémie par ingestion d'une solution contenant une quantité définie de glucose, et à mesurer la glycémie avant et après ingestion.

Le modèle de Bergman est généralement représenté par un ensemble d'équations différentielles qui décrivent la cinétique des concentrations de glucose et d'insuline. Ces équations sont souvent simplifiées pour rendre le modèle informatiquement exploitable tout en capturant la dynamique physiologique essentielle. En particulier, le modèle comprend :

- l'efficacité du glucose, qui représente la capacité du glucose à provoquer sa propre élimination en l'absence d'insuline,
- la sensibilité à l'insuline qui représente la capacité de l'insuline à augmenter l'élimination du glucose,
- et la réponse insulinaire aiguë qui représente la sécrétion initiale d'insuline en réponse à une provocation par le glucose, reflétant la fonction des cellules bêta pancréatiques.

Le modèle est ajusté aux données expérimentales obtenues à partir d'un test de tolérance au glucose. En estimant les paramètres du modèle pour s'adapter au mieux aux concentrations de glucose et d'insuline observées au fil du temps, on peut estimer la sensibilité à l'insuline, l'efficacité du glucose et la réponse insulínique aiguë de l'individu.

Le modèle minimal de Bergman a été largement utilisé dans la recherche clinique pour évaluer la sensibilité à l'insuline et la fonction des cellules bêta dans diverses populations, notamment les personnes souffrant d'obésité, de résistance à l'insuline, d'intolérance au glucose et de diabète de type 2. Il fournit des informations sur la physiopathologie du métabolisme du glucose et peut aider à la détection précoce et à la gestion des troubles métaboliques. Dans l'ensemble, le modèle minimal de Bergman fournit un cadre précieux pour quantifier la sensibilité à l'insuline et la fonction des cellules bêta, contribuant ainsi à notre compréhension de la régulation du glucose chez des individus sains comme malades.

La figure 2.2 donne une représentation graphique du modèle minimal de Bergman. Mathématiquement, ce modèle est défini par le système d'EDO donné à la figure 2.3.

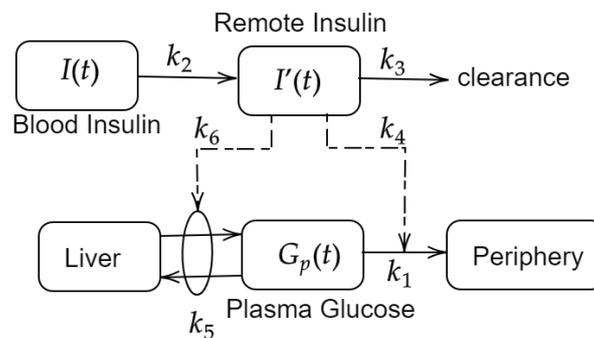


FIGURE 2.2 – Représentation graphique du modèle de Bergman. Les arcs pleins représentent les flux de glucose et d'insuline tandis que les traits en pointillés représentent les régulations.

2.2.2 Modèle de vidange gastrique d'Elashoff

Le modèle de vidange gastrique Elashoff est un modèle mathématique utilisé pour décrire le processus de vidange gastrique, c'est-à-dire le mouvement des aliments de l'estomac vers l'intestin grêle (Elashoff et al., 1982). Ce modèle a été développé par le Dr Jerry D. Elashoff dans les années 1980 et a été largement utilisé dans les études pharmacocinétiques, notamment dans le contexte de médicaments administrés par voie orale. Le modèle d'Elashoff décrit la vidange gastrique comme un processus biphasique, impliquant une phase initiale rapide suivie d'une phase plus lente. Il est souvent exprimé mathématiquement à l'aide de fonctions exponentielles pour représenter le taux de vidange au fil du temps. Le modèle inclut généralement des paramètres tels que le temps de demi-vidange pour chaque phase, qui quantifie le temps nécessaire à la moitié de la matière ingérée pour se vider de l'estomac.

$$\frac{dG(t)}{dt} = -(k_1 + k_5 + X(t)) \cdot G(t) + B_0 \quad (2.1)$$

$$\frac{dX(t)}{dt} = -k_3X(t) + k_2(k_4 + k_6) \cdot I(t) \quad (2.2)$$

$$X(t) = (k_4 + k_6) \cdot I'(t) \quad (2.3)$$

où :

- $G(t)$ est le glucose dans le sang,
- $I(t)$ est l'insuline dans le sang,
- $X(t)$ est l'insuline intervenant effectivement dans la régulation de la glycémie, obtenue à partir d'une mesure notée $I'(t)$,
- B_0 est la concentration de glucose hépatique à concentration de glucose plasmatique nulle.

FIGURE 2.3 – Modèle minimal de Bergman.

Le modèle de vidange gastrique d'Elashoff a été utilisé dans diverses applications de recherche. Il est par exemple utilisé pour étudier la cinétique d'absorption des médicaments administrés par voie orale, en particulier ceux qui peuvent être affectés par les taux de vidange gastrique. Cela permet d'optimiser les formulations de médicaments et les schémas posologiques afin de garantir une absorption efficace et des résultats thérapeutiques. Le modèle d'Elashoff est aussi utilisé dans des études portant sur les effets de différents aliments et nutriments sur les taux de vidange gastrique. En quantifiant la cinétique de vidange gastrique, les chercheurs peuvent évaluer l'impact des facteurs alimentaires sur l'absorption des nutriments et la fonction gastro-intestinale globale. Enfin, en milieu clinique, ce modèle peut être utilisé pour évaluer les anomalies de la vidange gastrique chez les patients souffrant de troubles gastro-intestinaux, tels que la gastroparésie (vidange gastrique retardée) ou le syndrome de dumping (vidange gastrique rapide). Le modèle sert à l'évaluation de la dynamique de la vidange gastrique qui peut aider à diagnostiquer et à gérer ces conditions. Dans l'ensemble, le modèle de vidange gastrique d'Elashoff fournit donc un cadre précieux pour quantifier le processus complexe de vidange gastrique et a de larges applications dans la recherche biomédicale et la pratique clinique.

Le modèle d'Elashoff est représenté graphiquement à la figure 2.4 ainsi que par l'équation de la figure 2.5.

2.2.3 Modèle de vidange gastrique de Dalla Man

Une des difficultés pour le développement et l'évaluation de la qualité des modèles de vidange gastrique a été le manque de données gold standard pour leur validation. Ainsi, bien que des données sur les concentrations plasmatiques de glucose soient disponibles, les taux d'apparition dans le plasma du glucose exogène (*Ra* pour *rate of appearance of exogenous glucose*) n'ont jamais été utilisés pour développer de tels modèles jusqu'alors. Par conséquent, Dalla Man

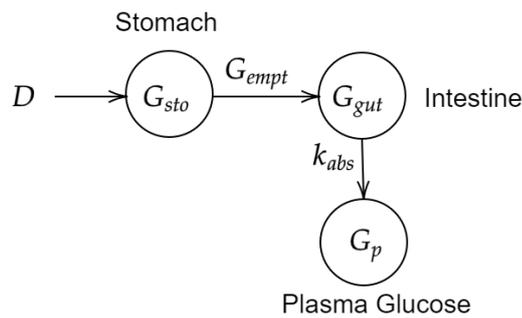


FIGURE 2.4 – Représentation graphique du modèle de vidange gastrique d'Elashoff.

$$G_{empt}(t) = D \cdot \beta \cdot k^\beta \cdot t^{\beta-1} \cdot e^{-(kt)^\beta}$$

avec :

- $G_{empt}(t)$ la vitesse de la vidange gastrique en fonction du temps,
- D la dose initiale de glucose entrant dans l'estomac,
- k un taux qui dépend directement de $t_{1/2}$, le temps de vidange de la moitié de la quantité initiale : $k = \frac{\ln(2)}{t_{1/2}}$,
- β un paramètre qui détermine la forme de la courbe : une grande valeur de β augmente le délai à partir duquel la vidange est notable.

FIGURE 2.5 – Modèle de vidange gastrique proposé par (Elashoff et al., 1982)

et ses collègues ont utilisé leurs propres données de Ra , estimées avec une technique de traceur multiple – obtenues sans modèle mécanistique – pour formuler un modèle de vidange gastrique du glucose (Dalla Man et al., 2006a). C'est un modèle qui a davantage été conçu avec une approche de conformité aux données (*fitting*) plutôt qu'une démarche mécanistique, ce qui rend l'interprétation biologique, c'est-à-dire l'explicativité des paramètres, complexe. Leur modèle a été testé par rapport à celui d'Elashoff (ainsi que deux autres) sur cet ensemble de données de traceurs multiples.

Ce modèle est non linéaire, et simule leurs données Ra de manière satisfaisante après estimation de ses paramètres. Il répond notamment à la spécificité d'un taux d'absorption intestinale présentant deux pics. Ce modèle renforce ainsi l'hypothèse de l'importance de la vidange gastrique par rapport au taux d'absorption Ra , en opposition à l'hypothèse que nous essayons de démontrer dans cette thèse. Cependant, ce modèle n'a pas été construit avec une approche explicative, ce qui pourrait montrer ses limites dans d'autres conditions expérimentales, telles que des données obtenues après chirurgie gastro-intestinale.

Le modèle décrivant la vidange gastrique a été défini comme indiqué par les équations de la figure 2.6 et une traduction de ces équations en graphe de réactions chimiques est proposée dans la figure 2.7. Dans le modèle, Q_{sto1} (équation 2.4) représente la quantité de glucose dans le bolus alimentaire non broyé tandis que Q_{sto2} (équation 2.5) représente la quantité de glucose

dans le bolus broyé pouvant faire l'objet de la vidange gastrique. Le passage du premier état au second se fait avec un taux k_{21} et la somme de ces deux quantités donne Q_{sto} (équation 2.6), soit la quantité totale de glucose dans l'estomac, qui détermine directement la vitesse de vidange gastrique appliquée à Q_{sto_2} à travers le paramètre (non constant) k_{empt} (équation 2.8). Enfin, Q_{gut} (équation 2.7) représente la quantité de glucose dans l'intestin et pouvant faire l'objet de l'absorption intestinale avec un taux k_{abs} . Les paramètres α , β , b , c , k_{min} et k_{max} influent sur la forme de la courbe du taux de vidange k_{empt} en fonction de la quantité totale Q_{sto} de glucose dans l'estomac, comme montré à la figure 2.8. Cette figure représente l'évolution du taux de la vidange gastrique en fonction de la quantité du bol alimentaire (Q_{sto}) restant dans l'estomac, à partir d'une dose initiale D . Plus précisément, D représente la dose initiale de glucose dans l'estomac, f la fraction absorbable de glucose (la partie non absorbée est excrétée ou métabolisée par les entérocytes) et BW est le poids de l'individu. Ce taux varie entre une valeur maximale k_{max} et une valeur minimale k_{min} . D'après les auteurs du modèle, le tracé de la figure 2.8 serait réaliste s'il est lu de droite à gauche, reflétant ainsi comment un estomac humain se viderait en fonction de la quantité de son contenu restant.

$$\frac{dQ_{sto_1}(t)}{dt} = -k_{21} \cdot Q_{sto_1}(t) \quad (2.4)$$

$$\frac{dQ_{sto_2}(t)}{dt} = k_{21} \cdot Q_{sto_1}(t) - k_{empt}(Q_{sto}) \cdot Q_{sto_2}(t) \quad (2.5)$$

$$Q_{sto}(t) = Q_{sto_1}(t) + Q_{sto_2}(t) \quad (2.6)$$

$$\frac{dQ_{gut}(t)}{dt} = k_{empt}(Q_{sto}) \cdot Q_{sto_2}(t) - k_{abs} \cdot Q_{gut}(t) \quad (2.7)$$

$$k_{empt}(Q_{sto}) = k_{min} + \frac{k_{max} - k_{min}}{2} \cdot \{ \tanh[\alpha(Q_{sto} - b \cdot D)] - \tanh[\beta(Q_{sto} - c \cdot D)] + 2 \} \quad (2.8)$$

$$\alpha = \frac{5}{2 \cdot D \cdot (1 - b)} \quad (2.9)$$

$$\beta = \frac{5}{2 \cdot D \cdot c} \quad (2.10)$$

$$Ra(t) = \frac{f \cdot k_{abs} \cdot Q_{gut}(t)}{BW} \quad (2.11)$$

FIGURE 2.6 – Système d'équations différentielles ordinaires du sous-système de la vidange gastrique proposé par Dalla Man.

Malgré les performances montrées dans l'article de référence du modèle (Dalla Man et al., 2006a) sur leur jeux de données, ces derniers ne sont pas disponibles, posant ainsi un problème de reproductibilité.

2.2.4 Modèle post-prandial homéostatique de Dalla Man

De nombreux modèles de simulation du système glucose-insuline pour la période postprandiale ont été développés (Mari et al., 2020). Dans ce travail, nous examinons un modèle très

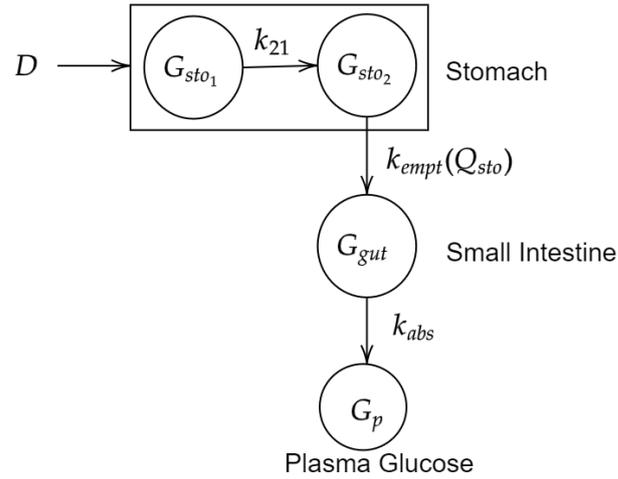


FIGURE 2.7 – Représentation graphique du modèle de vidange gastrique de Dalla Man.

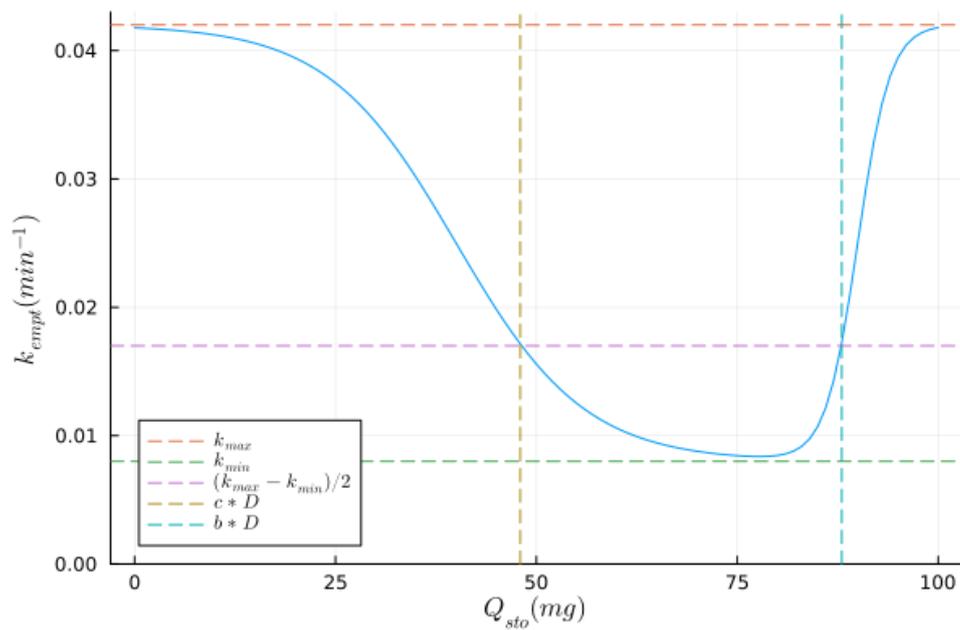


FIGURE 2.8 – Évolution de la cinétique de la vidange gastrique (k_{empt} , en min^{-1}) en fonction de la quantité totale de glucose dans l'estomac (Q_{sto} , en mg), selon le modèle proposé par Dalla Man (Dalla Man et al., 2006b).

cité, proposé dans (Dalla Man et al., 2007), pour simuler les événements physiologiques post-prandiaux de leurs propres cohortes de sujets normaux et de patients atteints de DT2. Ce modèle est composé de 12 EDO et de 36 paramètres décrivant les flux de glucose et d'insuline entre les compartiments physiologiques : tractus gastro-intestinal, plasma, foie, pancréas, muscle et tissus adipeux (Figure 2.9). En outre, 14 équations algébriques décrivent des comportements physiologiques complexes qui ne pourraient pas être abstraits sous forme d'équations différentielles ordinaires (par exemple, le retard de l'insuline pour atteindre ses cibles, les fonctions par morceaux pour éviter les valeurs négatives, etc.). Nous rappelons ci-après, en termes informels, comment les modules physiologiques interagissent. L'ensemble des flux du modèle est résumé par la figure 2.9.

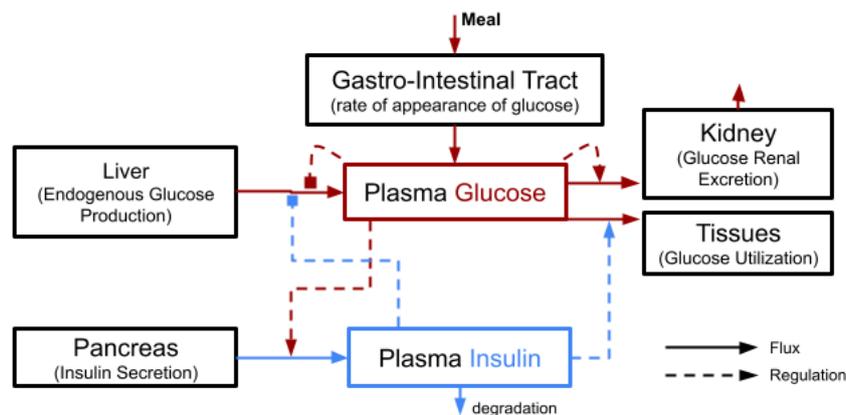


FIGURE 2.9 – Graphe d'interaction simplifié des compartiments physiologiques décrivant le modèle de (Dalla Man et al., 2007).

Le sous-réseau du *Tractus gastro-intestinal (Gastro-Intestinal Tract)* décrit le processus de digestion, de l'estomac à l'intestin, et peut être considéré comme l'entrée de l'ensemble du système, tout en faisant abstraction de la mastication. Il inclut la complexité de la vidange gastrique en fonction de la proportion des phases solides et liquides du bol alimentaire, tel que décrit dans la section 2.2.3 (Dalla Man et al., 2006a). Seule la phase liquide qui se retrouve dans l'intestin est absorbée par l'intestin et déversée dans le sang. Les incrétines (hormones stimulant la création de l'insuline) ne sont pas modélisées. Le sous-réseau de la *Cinétique du glucose (Glucose Kinetics)* est le sous-système décrivant la circulation du glucose entre le sang et les tissus. Ce sous-réseau hautement connecté englobe en particulier la glycémie plasmatique (désignée par G). Cette variable représente non seulement la quantité de glucose ingérée, mais aussi la quantité de glucose produite par l'organisme (principalement le foie). Le sous-réseau du *Foie (Liver)* décrit l'activité hépatique responsable de la production endogène de glucose, appelé glycogène, régulée par l'insuline. L'hormone glucagon, qui stimule la production endogène de glucose, n'est pas modélisée. Le *Pancréas* sécrète de l'insuline, qui est régulée en fonction de la quantité de glucose dans le sang. L'absorption du glucose par les *Tissus* (musculaires et adipeux) est régulée par l'insuline. Le *Rein (Kidney)* est responsable de l'excrétion-réabsorption du glucose. Évidemment, pour un patient en bonne santé, le modèle est paramétré de manière à ce que l'ab-

sorption compense entièrement l'excrétion, de sorte qu'il n'y ait pas de trace de glucose dans l'urine. Il est à noter que le module rénal n'inclut pas l'élimination de l'insuline. La dégradation de l'insuline est due à sa durée de vie et à l'élimination hépatique. Même si l'absorption du glucose par le cerveau, indépendante de l'insuline, est modélisée (non représentée dans la figure 2.9), la régulation du glucose par le cerveau est négligée dans (Dalla Man et al., 2007).

Il est intéressant de noter qu'un tel système est construit sans prendre en compte les régulateurs importants tels que le glucagon ou l'incrétine pour s'adapter aux données des modélisateurs. Cependant, il est admis que ces molécules jouent un rôle clé dans l'homéostasie du glucose au cours de la phase post-prandiale. Par conséquent, une telle omission dans un modèle de réseau physiologique peut conduire à un manque d'explications mécanistes pour l'homéostasie et peut être une source d'inexactitude pour les données expérimentales différentes de celles utilisées pour construire le modèle.

En utilisant la technique du clamp à rapport traceur-traceur, les auteurs de (Dalla Man et al., 2007) ont mesuré l'évolution temporelle de la concentration de glucose dans différents compartiments. Cette mesure a été effectuée après un repas mixte reçu par plusieurs sujets normaux et diabétiques. Les paramètres ont été estimés pour s'adapter à ces données expérimentales, ce qui a donné lieu à deux ensembles de valeurs de paramètres modélisant respectivement les comportements normaux et DT2. Le système d'équations différentielles ordinaires et l'ensemble des équations algébriques du modèle sont définies dans la figure 2.10. Nous verrons dans le chapitre suivant toute la difficulté de traduire le système d'équations en une représentation formelle sous forme de réseau de réactions chimiques.

Dans ce système, les observables d'intérêt sont la concentration de glucose et de l'insuline du plasma sanguin, respectivement notés G_p et I_p (équations 2.12 et 2.20) en mg/kg et pmol/kg, respectivement. La concentration de glucose dans le sang, qui s'équilibre avec celle des tissus G_t (équation 2.13), a une source endogène et exogène. La source endogène est hépatique, représentée par la fonction algébrique EGP (équation 2.15). Cette dernière est négativement régulée par G_p , I_{po} la concentration d'insuline dans la veine porte (équation 2.28) et I_d , la concentration d'insuline atteignant tardivement le foie avec un délai (équation 2.25). La source exogène de glucose est l'apparition de glucose absorbée par l'alimentation, définie par la fonction algébrique Ra (équation 2.11 de la figure 2.6), qui dépend notamment de biodisponibilité f (la proportion de glucose absorbable), de la masse du corps de l'individu BW (en kg), de la cinétique d'absorption k_{abs} et de la quantité de glucose dans l'intestin grêle Q_{gut} (en mg, cf. équation 2.7 de la figure 2.6). Cette partie du modèle est en fait directement reprise du modèle décrit à la section 2.2.3. La consommation de glucose se fait essentiellement par les tissus, formalisée par les équations U_{ii} (utilisation insulino-indépendante, cf. équation 2.17) et U_{id} (utilisation insulino-dépendante, cf. équation 2.16). La concentration I_p , qui s'équilibre avec celle dans le foie I_l (équation 2.19), est la résultante de l'activité pancréatique, formalisée par les équations différentielles I_{po} (spécifique à la circulation d'insuline dans la veine porte, cf. équation 2.28) et Y (spécifique à la sécrétion par le pancréas, cf. équation 2.30). Leur régulation, complexe, fait intervenir les équations algébriques S_{po} et S (équations 2.29 et 2.27). L'élimination rénale

Sous-modèle du tractus gastro-intestinal : Voir section 2.2.3 et figure 2.6.

Sous-modèle de la circulation interne de glucose :

$$\frac{dG_p(t)}{dt} = EGP(t) + Ra(t) - U_{ii}(t) - E(t) - k_1 \cdot G_p(t) + k_2 \cdot G_t(t) \quad (2.12)$$

$$\frac{dG_t(t)}{dt} = -U_{id}(t) + k_1 \cdot G_p(t) - k_2 \cdot G_t(t) \quad (2.13)$$

$$G(t) = \frac{G_p}{V_G} \quad (2.14)$$

$$EGP(t) = k_{p1} - k_{p2} \cdot G_p(t) - k_{p3} \cdot I_d(t) - k_{p4} \cdot I_{po}(t) \quad (2.15)$$

$$U_{id}(t) = \frac{(V_{m0} + V_{mx} \cdot X(t)) \cdot G_t(t)}{(K_{m0} + K_{mx} \cdot X(t)) + G_t(t)} \quad (2.16)$$

$$U_{ii}(t) = F_{cns} \quad (2.17)$$

$$E(t) = \begin{cases} k_{e1} \cdot [G_p(t) - k_{e2}] & \text{si } G_p(t) > k_{e2} \\ 0 & \text{si } G_p(t) \leq k_{e2} \end{cases} \quad (2.18)$$

Sous-modèle de la circulation interne de l'insuline :

$$\frac{dI_l(t)}{dt} = -(m_1 + m_3(t)) \cdot I_l(t) + m_2 I_p(t) + S(t) \quad (2.19)$$

$$\frac{dI_p(t)}{dt} = -(m_2 + m_4) \cdot I_p(t) + m_1 \cdot I_l(t) \quad (2.20)$$

$$I(t) = \frac{I_p}{V_I} \quad (2.21)$$

$$m_3(t) = \frac{HE(t) \cdot m_1}{1 - HE(t)} \quad (2.22)$$

$$HE(t) = -m_5 \cdot S(t) + m_6 \quad (2.23)$$

$$\frac{dI_1(t)}{dt} = -k_i \cdot [I_1(t) - I(t)] \quad (2.24)$$

$$\frac{dI_d(t)}{dt} = -k_i \cdot [I_d(t) - I_1(t)] \quad (2.25)$$

$$\frac{dX(t)}{dt} = -p_{2U} \cdot X(t) + p_{2U} [I(t) - I_b] \quad (2.26)$$

$$S(t) = \gamma \cdot I_{po}(t) \quad (2.27)$$

$$\frac{dI_{po}(t)}{dt} = -\gamma \cdot I_{po}(t) + S_{po}(t) \quad (2.28)$$

$$S_{po} = \begin{cases} Y(t) + K \cdot \frac{dG(t)}{dt} + S_b & , \text{ si } \frac{dG(t)}{dt} > 0 \\ Y(t) + S_b & , \text{ si } \frac{dG(t)}{dt} \leq 0 \end{cases} \quad (2.29)$$

$$\frac{dY(t)}{dt} = \begin{cases} -\alpha \cdot [Y(t) - \beta \cdot (G(t) - h)] & , \text{ si } \beta \cdot (G(t) - h) \geq -S_b \\ -\alpha \cdot Y(t) - \alpha \cdot S_b & , \text{ si } \beta \cdot (G(t) - h) < -S_b \end{cases} \quad (2.30)$$

FIGURE 2.10 – Système d'équations différentielles ordinaires du modèle postprandial du glucose et de l'insuline tel que proposé dans (Dalla Man et al., 2007).

de l'insuline est formalisée par m_3 (équation 2.22).

En plus de la multitude de paramètres cinétiques, nous pouvons observer un nombre important de variables dépendantes, ce qui soulève la question des états initiaux. Dans l'article de référence, rien ne semble indiquer comment certains sont obtenus, notamment la concentration initiale d'insuline dans le foie ou dans la concentration initiale de glucose dans les tissus. On pourrait raisonnablement douter de l'origine expérimentale des valeurs proposées par les auteurs pour les deux variables sus-mentionnées et ainsi penser que ces valeurs sont issues de travaux d'estimations préalable de concentrations à partir des mesures expérimentales des observables G_p , I_p et Ra .

2.2.5 Modèle intestinal de Salinari

Le taux d'apparition (Ra) de glucose exogène dans le plasma après l'ingestion de glucose est actuellement mesuré par des techniques de traceur multiple qui ne peuvent pas être utilisées dans les routines de tests cliniques tels que le test de tolérance au glucose oral (TTGO). C'est sur cette observation que Salinari et ses collègues ont proposé un modèle mathématique sous la forme d'une équation différentielle partielle (ou PDE) qui représente de manière simple le transport du glucose le long du tractus intestinal et son absorption depuis la lumière intestinale jusqu'au sang de la veine porte (Salinari et al., 2011). Le modèle donne l'évolution temporelle de Ra en fonction de l'évolution spatiale du bol alimentaire dans la lumière intestinale. Les performances du modèle ont été évaluées par l'estimation de paramètres calibrées sur des données de traceur multiple et d'OGTT. De par sa complexité pour décrire le transit et l'absorption intestinale, ce modèle s'inscrit en opposition aux modèles permettant de suivre Ra qui jusqu'alors mettaient l'emphase sur la complexité de la vidange gastrique pour expliquer le Ra aux détriments des événements intestinaux. Le modèle est défini dans la figure 2.11.

De par sa définition, ce sous-modèle permet de suivre l'évolution de la quantité ou de la concentration de glucose le long de l'intestin en fonction du temps. Une proposition de la représentation graphique du modèle est donnée par la figure 2.12, dans lequel D est la dose ingérée de glucose dans l'estomac. La cinétique de la vidange gastrique est donnée par le modèle d'Elashoff, que nous avons présenté en section 2.5. Le choix d'un modèle simple de vidange gastrique permet d'éviter les effets confondants entre celle-ci et la cinétique de l'absorption intestinale, observable via la variable Ra , directement dépendant de Q_{gut} , la quantité de glucose dans l'intestin et de $\gamma(z)$, la fonction de transit du bolus alimentaire le long de l'intestin, dont la position est donnée par z . La principale difficulté de ce modèle réside dans le choix de la fonction de transition $\gamma(z)$, qui change la forme du tracé de $Ra(t)$. Une des caractéristiques réelles du taux d'absorption intestinale est la présence de deux pics d'absorption, un précoce, un tardif. La cause de leur présence est source d'effet confondant. En effet, on pourrait indifféremment l'attribuer soit à une vidange gastrique complexe, soit à une absorption intestinale complexe. La première hypothèse est la spécificité du modèle de vidange gastrique de Dalla Man et al. La seconde hypothèse est la spécificité de ce modèle de Salinari, qui ne peut être attribué qu'à la dynamique complexe de l'absorption intestinale étant donné que la vidange gastrique y a été

L'équation du modèle de Salinari est la suivante :

$$\frac{\partial q(z,t)}{\partial t} + u \cdot \frac{\partial q(z,t)}{\partial z} = -\gamma(z) \cdot q(z,t), \quad t \geq 0, z \geq 0 \quad (2.31)$$

avec :

- z : la position le long de l'intestin grêle (cm),
- t : le temps (min),
- $q(z,t)$: la concentration de glucose à la position z au temps t (mg/dL),
- $\gamma(z)$: la vitesse d'absorption intestinale à la position z (mg/[dL · min]),
- u : la vitesse de transit intestinal, supposée constante (cm/min).

Les conditions aux limites sont les suivantes, modélisant une vidange gastrique progressive au cours du temps :

$$q(0,t) = \begin{cases} \frac{D}{u} \cdot G_{empt}(t), & 0 \leq t \leq \theta \\ 0, & t \geq \theta \end{cases} \quad (2.32)$$

avec :

- θ : le temps total de la vidange gastrique (min),
- $G_{empt}(t)$: le taux de vidange gastrique au temps t repris du modèle d'Elashoff présenté à la section 2.2.2 et à la figure 2.5.

On peut dériver le taux d'absorption suivant de ces équations :

$$Ra(t) = f \int_0^L \gamma(z) \cdot q(z,t) dz, \quad t \geq 0 \quad (2.33)$$

où f est la fraction absorbable de glucose.

FIGURE 2.11 – Modèle proposé d'absorption intestinale de glucose sous la forme d'une équation différentielle partielle (Salinari et al., 2011).

simplifiée via le modèle d'Elashoff (figure 2.13).

Les auteurs de l'article associé ont testé différentes fonctions de transition, notamment linéaire et exponentielle. Dans l'objectif de conserver notre approche parcimonieuse dans la définition de notre propre modèle – notamment pour la dynamique du D-Xylose post-prandial, évoqué dans le chapitre 4 – nous avons fait le choix d'une fonction de transit du bolus alimentaire linéaire. De plus, ayant fait le choix du formalisme d'équations différentielles ordinaires, nous avons discrétisé ce modèle d'équations différentielles partielles en système différentielle ordinaire.

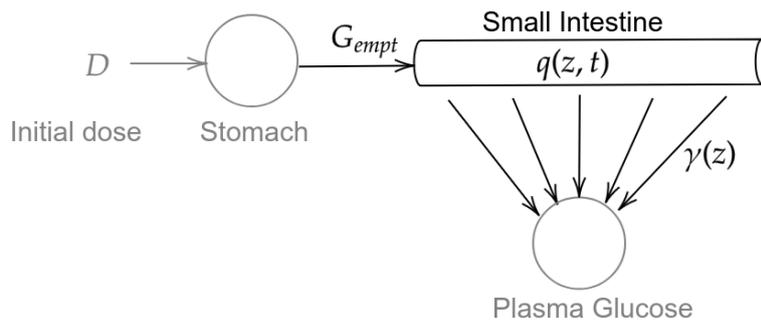


FIGURE 2.12 – Représentation graphique du modèle d’absorption intestinale de Salinari. Les parties en gris ne font pas partie du modèle mais sont représentées pour une meilleure compréhension.

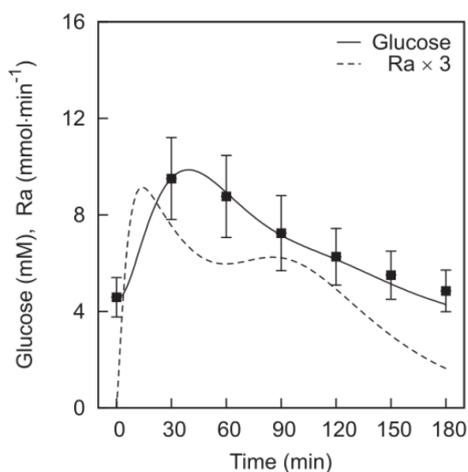


FIGURE 2.13 – Évolution du taux d’apparition de glucose, généré par le modèle de Sallinari et de la concentration mesurée de glucose dans le plasma sanguin (Salinari et al., 2011).

2.2.6 Conclusion

Parmi l'ensemble des formalismes de modélisations que nous avons évoqués – réseau de réactions chimiques, réseau de Petri, système d'équations différentielles ordinaires – nous avons fait le choix de nous orienter vers les réseaux de réactions chimiques, construits à partir de systèmes d'EDO afin de bénéficier à la fois de l'aspect mécanistique et de l'aspect quantitatif dans l'implémentation de nos modèles. Les modèles ainsi implémentés seront calibrés sur différents jeux de données, clinique ou préclinique, en utilisant un algorithme d'optimisation adapté. Comme nous le verrons dans le chapitre suivant, le premier modèle implémenté est le modèle post-prandial du glucose et de l'insuline proposé par Dalla Man, mentionné à la section 2.2.4, sur lequel les paramètres seront estimés sur le jeu de données cliniques OBEDIAB, et dont les résultats nous ont menés à réorienter notre méthodologie en y incluant les analyses d'identifiabilité et de sensibilité à partir du chapitre 4.

Chapitre 3

Étude du modèle postprandial de la dynamique glucose-insuline de Dalla Man : Reproductibilité et limites

3.1 Contexte et Motivations

Comme nous l'avons mentionné dans le chapitre précédent, de nombreux modèles homéostatiques du glucose et de l'insuline ont été développés. Parmi eux, nous nous sommes concentrés sur l'un des plus cités, proposé par *C. Dalla Man* et ses collaborateurs (Dalla Man et al., 2007), et précédemment présenté à la section 2.2.4. Cependant, ce modèle a été implémenté et paramétré sur un jeu de données cliniques classique – une cohorte de patients atteints de diabète de type 2, en surpoids – alors que nos données correspondent à une cohorte de patients atteints de diabète de type 2 en état d'obésité morbide, ayant subi une chirurgie gastro-intestinale (voir section 1.2). Il est aussi à rappeler que ce modèle accorde une grande importance à la complexité de la vidange gastrique (Dalla Man et al., 2006a), qui avait été décrit à la section 2.2.3, au détriment de l'absorption intestinale de glucose (AIG), ici modélisée par une simple loi d'action de masse. Or les données dont nous disposons et qui seront décrites dans la sous-section suivante, suggèrent que c'est la modification de l'intestin grêle et par corollaire, l'état de l'absorption intestinale de glucose, qui seraient fortement corrélés au profil glycémique des patients. Il nous a donc semblé intéressant de confronter ce modèle de référence à nos données cliniques. En résumé, dans ce chapitre, nous souhaitons étudier la contribution de l'AIG à l'homéostasie du glucose et son rôle potentiel dans le diabète. À cette fin, nous considérons le modèle homéostatique sus-mentionné, formalisé en tant que système d'équations différentielles ordinaires.

Notre objectif est double :

- tester si ce modèle peut prédire une amélioration significative de l'homéostasie du glucose en simulant le court-circuit gastrique de Roux-en-Y tel qu'il est observé expérimentalement,
- tester si ce modèle peut prédire les données chronologiques d'un ensemble de données

original de patients diabétiques.

La section suivante sera dédiée à la description des données utilisées dans ce chapitre. Ce chapitre a fait l'objet d'une publication dans la conférence BIOTECHNO en 2021 (Dursoniah et al., 2021).

3.2 Données cliniques

Le jeu de données cliniques dont nous disposons, nommé OBEDIAB, a été fourni par l'équipe de recherche dirigée par le Pr. François Pattou, la *Translational Research Laboratory for Diabetes (INSERM UMR1190)*. Ces données ont été produites par prélèvements de sang plasmatique, à différents temps, après la prise orale d'un repas test, à partir desquelles des mesures systématiques ont été effectuées. Les principales mesures d'intérêt de ces prélèvements sont la glycémie et l'insulinémie. La cohorte de cas cliniques se compose de patients en état d'obésité morbide. La cohorte de contrôle se compose d'individus sains. La population de cas cliniques (en état de morbidité) inclut :

- des individus atteints de diabète de type 2 : une partie d'entre eux ont eu recours à une pose d'un anneau gastrique (noté *GB* pour *Gastric Band* dans les résultats de ce manuscrit), aboutissant à une modification de l'estomac, l'autre partie a reçu une intervention chirurgicale appelée Roux-En-Y (notée *RY* dans la suite), aboutissant à une modification profonde de l'appareil gastro-intestinal ;
- des individus non atteints de diabète de type 2, et de la même manière que les malades : une partie ont eu recours à une pose de d'anneau gastrique tandis qu'une autre a reçu une intervention gastro-intestinale de type Roux-En-Y.

Toujours parmi tous les cas cliniques, les prélèvements ont été effectués avant et après l'intervention chirurgicale, à différents temps après ingestion orale d'un repas test (30 à 180 minutes). Pour la population contrôle, les prélèvements sont également effectués après un repas test, sans intervention chirurgicale. En complément de cette description qualitative des sous-populations qui composent nos données, il est intéressant d'observer les différences de poids moyens entre ces sous-populations, résumées dans la Table 3.1. Les données présentes dans ce tableau indiquent une amélioration nette du poids, quel que soit le type d'opération bariatrique choisi (anneau gastrique ou Roux-en-Y), tous cas cliniques confondus (diabétiques et non diabétiques). Cependant, il est à noter que les patients ayant reçu une intervention Roux-en-Y ont davantage vu leur poids diminuer que ceux ayant reçu un anneau gastrique. En effet, on observe une perte de poids de 10,6 % pour le premier type d'intervention et de 9,4 % pour le second, chez les diabétiques. De même pour les non diabétiques, on observe une perte de poids de 33,5 % et de 21,3 %, respectivement. Cette différence d'impact entre diabétiques et non-diabétiques pourrait être expliquée par la différence d'expression de gène du cotransporteur intestinal principal de glucose – SGLT1 – où l'on pourrait suspecter une bonne capacité d'adaptation du niveau d'expression de ce gène chez les diabétiques malgré la dérivation partielle de la voie de l'intestin grêle dans le cas du Roux-en-Y (Baud et al., 2016b).

Au-delà du bénéfice lié au poids, nos données indiquent également une amélioration de la glycémie postprandiale chez les patients diabétiques et non diabétiques (figures 3.1 et 3.3, respectivement) mais également de l'insulinémie postprandiale (figures 3.2 et 3.4, respectivement). L'observation de ces concentrations postprandiales indique surtout le bénéfice du Roux-en-Y sur les conséquences glycémiques et insulinémiques du diabète de type 2. Dans la figure 3.1, nous pouvons observer la *glycémie* postprandiale au cours du temps chez des patients atteints de diabète et chez une population contrôle. Il est frappant d'y constater une forte diminution de la glycémie basale, surtout après un Roux-en-Y, avec un rétablissement du pic glycémique à 30 minutes, caractéristique des personnes saines. Dans la figure 3.2, nous pouvons observer l'*insulinémie* postprandiale au cours du temps chez des patients atteints de diabète et chez une population contrôle. De la même manière que dans la figure précédemment décrite, on constate un rétablissement d'un pic insulinémique à 30 minutes, caractéristique des personnes saines, après une intervention Roux-en-Y, avec une remontée significative du niveau basal d'insuline. L'ensemble des remarques faites pour les patients obèses et diabétiques sur leur concentration postprandiale de glucose et d'insuline sont valables pour les patients obèses et non-diabétiques (voir les figures 3.3 et 3.4).

L'ensemble des observations faites sur nos données OBEDIAB semble indiquer qu'une intervention chirurgicale dans laquelle on dérive partiellement l'intestin grêle entraîne le rétablissement d'une glycémie et d'une insulinémie saines depuis un profil diabétique (OBD) ou prédiabétique (OBND). Cela nous conforte donc dans notre démarche de tester nos hypothèses liées à l'absorption intestinale de glucose tout au long de cette thèse, en commençant par démontrer le besoin de modéliser l'absorption intestinale de sucre en exposant les limites du modèle de référence – décrit à la section 2.2.4 et rappelé dans la section suivante – confronté à nos données.

	Cas clinique (59)								Cas contrôle (7)
	diabétique (45)				non-diabétique (14)				sain
	Anneau Gastrique (20)		Roux-En-Y (25)		Anneau Gastrique (5)		Roux-En-Y (9)		sans intervention
	avant	après	avant	après	avant	après	avant	après	-
Poids (kg)	128.26	116.19	148.15	132.52	134.15	105.54	149.26	99.27	68.71
Tailles (m)	1.63		1.68		1.70		1.66		1.73
BMI (kg/m ²)	47.88	43.38	52.44	47.31	46.16	36.39	53.73	35.99	22.69

TABLE 3.1 – Tableau des poids, tailles et IMC moyens des sous-populations du jeu de données cliniques interne. Les mesures les plus précoces *après* une chirurgie sont effectuées entre 3 et 5 mois.

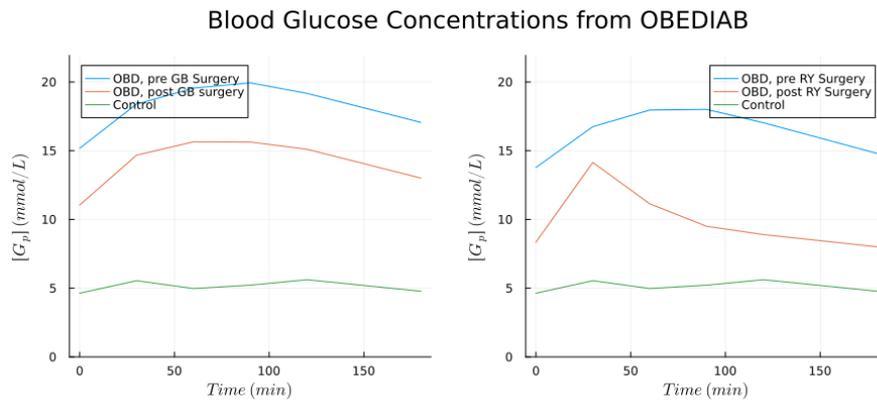


FIGURE 3.1 – Glycémie postprandiale au cours du temps chez des patients atteints de d’obésité et de diabète (OBD) et chez une population contrôlée. À gauche, population sous pose d’un anneau gastrique (GB pour *gastric band*). À droite, population sous Roux-en-Y (RY). Courbes bleues : glycémie avant intervention chirurgicale chez des patients atteints d’obésité et de diabète (OBD); courbes oranges : glycémie, après intervention chez des patients atteints d’obésité et de diabète (OBD); courbes vertes : glycémie de la population contrôlée n’ayant pas subi d’intervention.

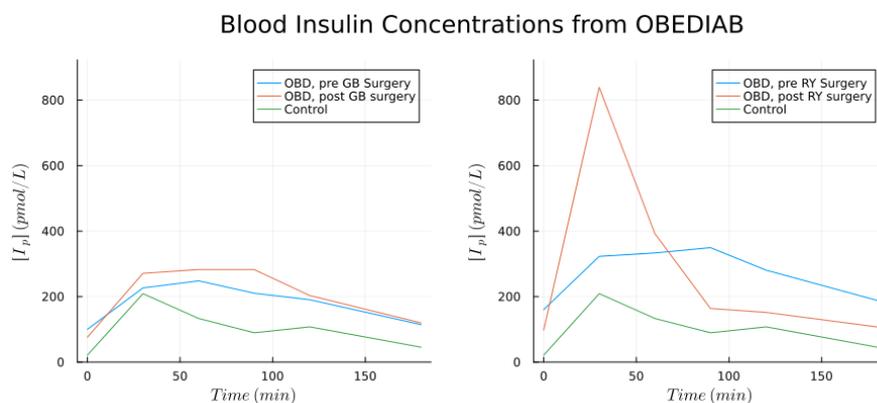


FIGURE 3.2 – Insulinémie postprandiale au cours du temps chez des patients atteints de d’obésité et de diabète (OBD) et chez une population contrôlée. À gauche, population sous pose d’un anneau gastrique (GB pour *gastric band*). À droite, population sous Roux-en-Y (RY). Courbes bleues : insulinémie avant intervention chirurgicale chez des patients atteints d’obésité et de diabète (OBD); courbes oranges : insulinémie, après intervention chez des patients atteints d’obésité et de diabète (OBD); courbes vertes : insulinémie de la population contrôlée n’ayant pas subi d’intervention.

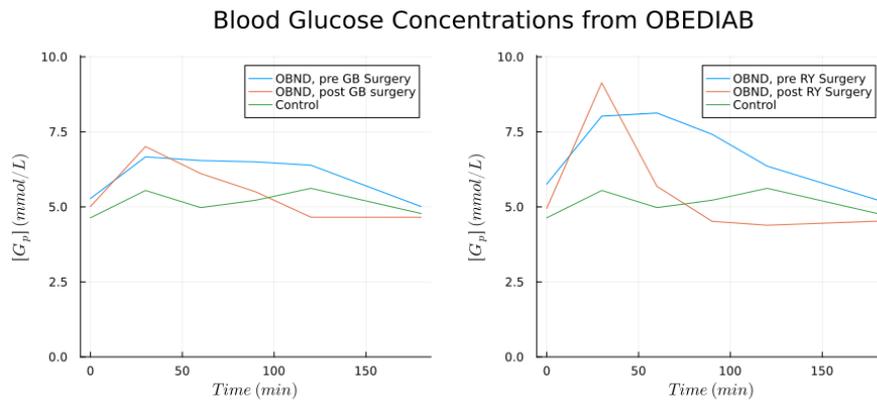


FIGURE 3.3 – Glycémie postprandiale au cours du temps chez des patients atteints de d’obésité, non diabétiques (OBND) et chez une population contrôle.

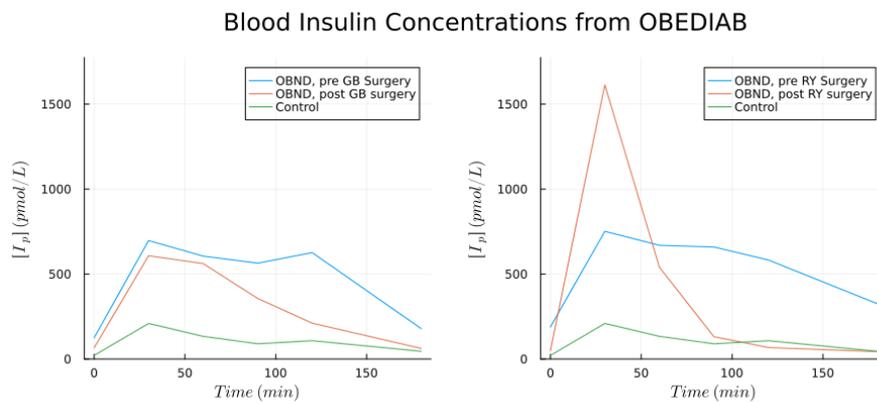


FIGURE 3.4 – Insulinémie postprandiale au cours du temps chez des patients atteints de d’obésité, non diabétiques (OBND) et chez une population contrôle.

3.3 Modèle de Dalla Man *et al.*

Le modèle proposé par Dalla Man (Dalla Man *et al.*, 2007), comme décrit dans le chapitre 2, est un modèle de référence qui permet le suivi de simulations in-silico de la concentration de glucose et de l'insuline plasmatiques au cours du temps. Sa modélisation repose sur la répartition des équations différentielles et algébriques en différents modules biologiques, pouvant librement être interprétés, soit comme des activités biologiques, soit comme des objets biologiques (systèmes, organes ou tissus) comme montré dans la figure 2.9.

L'implémentation d'un tel réseau nous a mené à nous intéresser plus en détail aux équations différentielles. Or, comme nous l'avons décrit dans la sous section dédiée à la description des formalismes – dans le chapitre 2 – un réseau de réactions chimiques, représentable graphiquement, peut être traduit en un système d'équations différentielles, mais l'inverse n'est pas forcément vrai. Pourtant cette conversion du système d'équations de notre modèle postprandial de référence en un réseau de réactions chimiques est une étape importante dans cette thèse : il s'agit de vérifier si le modèle de référence est réellement mécanistique, malgré la présence d'équations algébriques, non-différentielles. De cet exercice en a résulté le graphe d'interaction de la figure 3.6. À la différence du modèle minimal de Bergman (Bergman, 1989), présenté à la section 2.2.1, ce modèle proposé par Dalla Man présente un système de régulation complexe non seulement pour le glucose, mais également pour l'insuline, pour différents compartiments biologiques. Dans la figure 3.6, les éléments chimiques de l'insuline et du glucose, pour différents compartiments, sont représentés par des places, à l'instar des réseaux de Pétri. Les flux directs, c'est-à-dire le transfert de quantité ou de concentration de proche en proche, entre les compartiments, glycémiques ou insulinémiques, sont représentés par des flèches pleines. Ce type de flux est exclusivement formalisé par une loi d'action de masse. Pour illustrer cette traduction, prenons, par exemple, les équations suivantes, que l'on peut retrouver dans la figure 2.10 :

$$\frac{dG_p(t)}{dt} = EGP(t) + Ra(t) - U_{ii}(t) - E(t) - k_1 \cdot G_p(t) + k_2 \cdot G_t(t) \quad (3.1)$$

$$\frac{dG_t(t)}{dt} = -U_{id}(t) + k_1 \cdot G_p(t) - k_2 \cdot G_t(t) \quad (3.2)$$

$$(3.3)$$

Pour rappel, G_p et G_t sont respectivement la concentration de glucose dans le plasma sanguin et dans les tissus. Si on ne s'intéresse qu'à l'interaction entre ces deux variables, on se retrouve avec les équations suivantes, un sous-système d'équations de loi d'action de masse :

$$\frac{dG_p(t)}{dt} = -k_1 \cdot G_p(t) + k_2 \cdot G_t(t) \quad (3.4)$$

$$\frac{dG_t(t)}{dt} = k_1 \cdot G_p(t) - k_2 \cdot G_t(t) \quad (3.5)$$

Les paramètres cinétiques en commun dans les deux équations sont k_1 et k_2 . Ces deux paramètres sont alternativement des termes positifs ou négatifs en fonction de l'équation, qui se traduisent comme des flux entrants et sortants, respectivement. Par exemple, $-k_1 \cdot G_p(t)$ peut

être interprété comme un flux sortant de glucose du compartiment plasmatique avec un taux k_1 . Par symétrie, $k_1 \cdot G_p(t)$ peut être interprété comme un flux entrant de glucose dans le compartiment tissulaire depuis le compartiment plasmatique avec un taux k_1 . En procédant de la même manière pour k_2 , on peut déduire une première réaction chimique à partir des équations 3.4 et 3.4, définie de la manière suivante :



Puis, à partir de cette réaction chimique, on en déduit la représentation graphique montrée dans la figure 3.5.

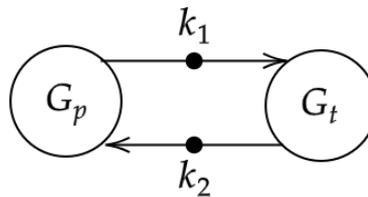


FIGURE 3.5 – Diagramme d’interactions entre le glucose sanguin G_p et le glucose des tissus G_t . Les flèches en trait plein représentent le transfert de flux de glucose entre les 2 compartiments, permettant d’établir un équilibre de concentrations de glucose entre eux.

Au-delà des flux, notre système d’équations modélise également des régulations. La complexité de représenter ces régulations sous forme de réactions chimiques réside dans le fait que quasiment aucune d’elles n’est modélisée par une fonction de Hill ou de Michaelis-Menten, généralement utilisées pour ces cas. Ainsi, pour se représenter les régulations positives et négatives, nous avons utilisé des arcs en pointillé, terminés par une flèche ou par un trait, respectivement. Or, en plus de ces régulations, certaines équations indiquent des régulations qui ne sont pas monotones, c’est-à-dire, pas absolument positives ou négatives, mais dépendent de certaines conditions de concentrations, en utilisant, par exemple, des équations par morceaux. Dans de tels cas – que l’on peut qualifier d’ambigus –, les arcs en pointillés sont terminés par un demi-cercle, sans plus de précision pour ne pas surcharger notre représentation graphique. Nous aboutissons ainsi à la figure 3.6, qui permet une lecture intuitive du modèle, au détriment d’un formalisme rigoureux qui permettrait une traduction directe en équations, sans manque d’information. Une alternative plus formelle à cela est de proposer un système de réactions chimiques défini de manière analogue à la réaction 3.6 ci-dessus. On aboutit ainsi au système d’équation de la figure 3.7. Toute cette difficulté de lisibilité des événements biologiques dans notre modèle de référence, quel que que soit le formalisme choisi, est due à la démarche de Dalla Man dans l’implémentation de son modèle, qui inclut à la fois des cinétiques dites « physiologiques » – comme évoquées dans l’article de référence – et d’autres, que l’on pourrait considérer comme « heuristiques », simplement conçues pour faire correspondre le modèle aux données, sans forcément refléter un mécanisme biologique en particulier.

Malgré tout, il est à noter que ce modèle postprandial reste intéressant car, contrairement aux autres modèles proposés jusqu’ici, propose une sous-modélisation des flux et des régulations de

l'insuline. Dans ce modèle, la concentration de l'insuline est notée I et le compartiment observé est noté en indice :

- I_p : concentration d'insuline dans le plasma sanguin,
- I_l : concentration d'insuline circulant dans le foie,
- X : concentration d'insuline dans le tissu interstitiel (tout tissu de remplissage, structurante de l'organisme, généralement intermédiaire entre la circulation sanguine et les organes),
- I_1 et I_d : concentrations d'insuline « intermédiaires », induisant un « délai » entre sa production pancréatique et l'atteinte des tissus cibles I_l ,
- I_{po} : concentration d'insuline dans la veine porte, partie précédant le foie.

Pour des raisons pratiques, nous avons entièrement réimplémenté le modèle dans le langage de programmation *Julia* (version 1.5.3) avec les packages *DifferentialEquations*(version 6.17.0) et *emphCatalyst* (version 12.3.0).

3.4 Du modèle DT2 au modèle normal

Dans un premier temps, nous n'utilisons pas les données OBEDIAB présentées précédemment car nous souhaitons uniquement tester la capacité du modèle à simuler différents types de chirurgies et de rétablissement de l'homéostasie du glucose. Dans ce qui suit, nous appelons *modèle normal*, respectivement *modèle DT2*, le modèle instancié avec les valeurs de paramètres normales, respectivement diabétiques, données dans (Dalla Man et al., 2007) (tableau 1). Nous considérons plusieurs sous-ensembles de paramètres correspondant aux modules précédents : tractus gastro-intestinal (également désigné par Ra dans (Dalla Man et al., 2007)), foie (EGP), tissus (GK+U), pancréas (IK+S) et rein (RE). Partant du modèle DT2, nous avons ré-estimé, tour à tour, chacun de ces sous-ensembles de paramètres, tout en laissant les autres paramètres fixes, afin d'ajuster la dynamique de glucose plasmatique du modèle normal. Sur la base des données de (Dalla Man et al., 2007), notre premier objectif est d'évaluer la capacité du modèle à prédire, pour chaque module seul, sa capacité à restaurer une glycémie normale à partir d'un état diabétique. Ce deuxième point pourrait également être interprété comme une stratégie thérapeutique en identifiant l'organe à cibler pour guérir du diabète de type 2. Nous estimons les paramètres deux fois : avec et sans estimation des valeurs basales de sécrétion d'insuline par le pancréas, de production de glucose par le foie et d'utilisation par les tissus. À partir des 10 modèles inférés résultants, nous avons tracé l'évolution temporelle des variables observées.

Nous ne présentons à la figure 3.8 que les variables les plus pertinentes pour notre objectif : le glucose plasmatique (G), l'insuline plasmatique (I) et le taux d'absorption intestinal (Ra) pour les deux modèles obtenus à partir des estimations des compartiments gastro-intestinal et pancréatique, avec et sans estimation basale. Afin de comparer les performances des modèles dans l'ajustement du modèle normal, nous collectons dans un graphique en barres (figure 3.9) la somme résiduelle des carrés pour chaque modèle. Ainsi, lorsque cette valeur est élevée sur le graphique, pour un compartiment donné, cela indique une incapacité de l'optimiseur à réestimer

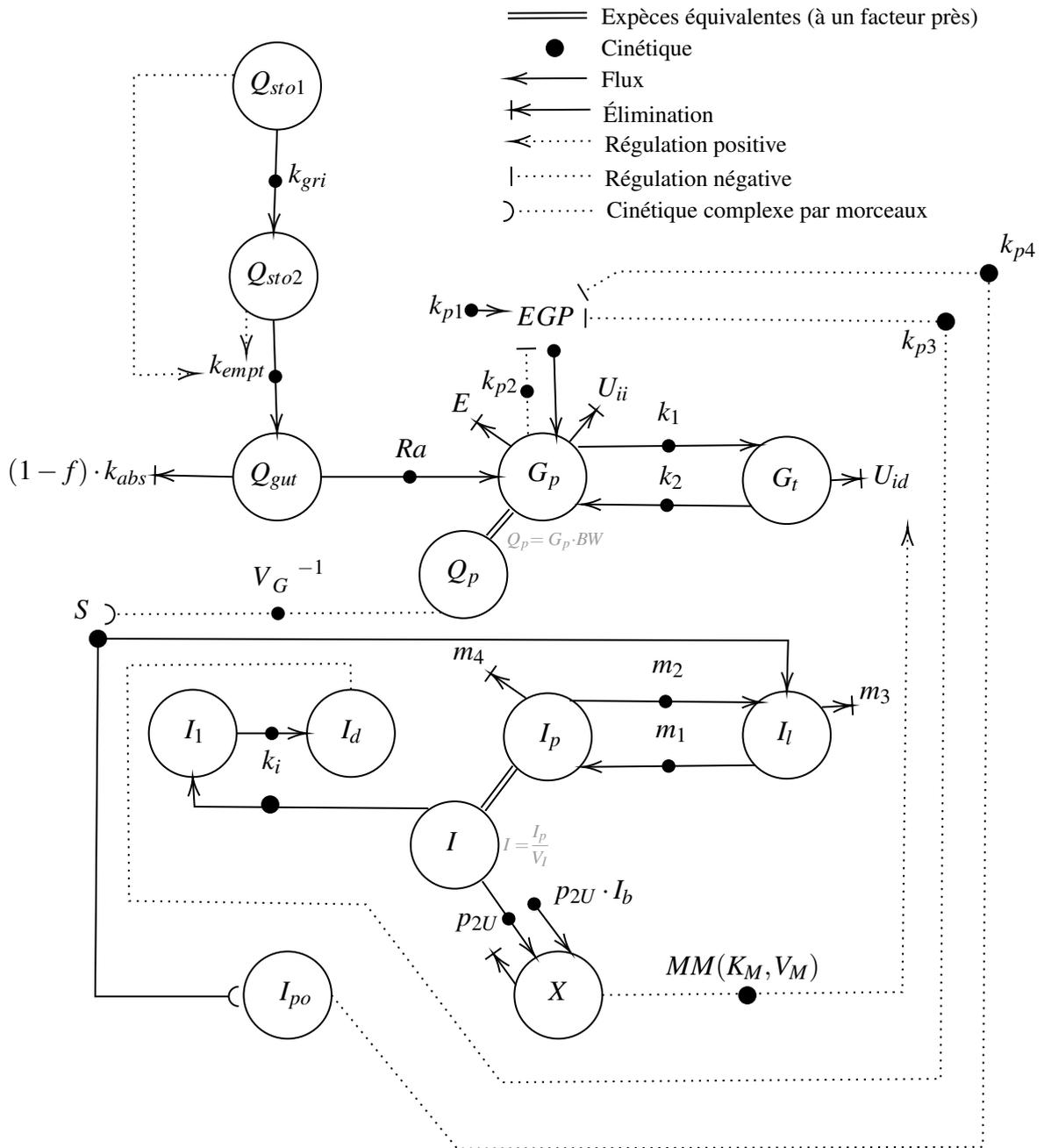


FIGURE 3.6 – Réseau de réactions d’interactions détaillé des compartiments physiologiques décrivant le modèle de *C. Dalla Man* (Dalla Man et al., 2007).

$Q_{sto1} \xrightarrow{k_{21}} Q_{sto2}$	(3.7)	Broyage du bol alimentaire par l'estomac
$Q_{sto2} \xrightarrow{k_{empt}(Q_{sto1}+Q_{sto2})} Q_{gut}$	(3.8)	Vidange gastrique
$Q_{gut} \xrightarrow{(1-f)k_{abs}} \emptyset$	(3.9)	Excrétion intestinale
$Q_{gut} \xrightarrow{fk_{abs}} Q_p$	(3.10)	Absorption intestinale
$Q_p = BW.G_p$	(3.11)	Quantité plasmatique (mg)
$\emptyset \xrightarrow{EGP} G_p$	(3.12)	Production endogène de glucose
$G_p \xrightarrow{U_{ii}} \emptyset$	(3.13)	Utilisation insulino-indépendante de glucose
$G_p \xrightarrow{E} \emptyset$	(3.14)	Excrétion rénale de glucose
$G_p \xrightleftharpoons[k_2]{k_1} G_t$	(3.15)	Circulation de glucose entre le plasma sanguin et les tissus
$G_t \xrightarrow{U_{id}} \emptyset$	(3.16)	Utilisation insulino-dépendante de glucose
$\emptyset \xrightarrow{S_{po}} I_{po}$	(3.17)	Sécrétion pancréatique d'insuline vers la veine porte
$I_{po} \xrightarrow{\gamma} I_\ell$	(3.18)	Circulation d'insuline de la veine porte vers le foie
$I_\ell \xrightleftharpoons[m_2]{m_1} I_p$	(3.19)	Circulation entre le foie et le plasma sanguin
$I_\ell \xrightarrow{m_3} \emptyset$	(3.20)	Dégradation hépatique de l'insuline
$I_p \xrightarrow{m_4} \emptyset$	(3.21)	Dégradation spontanée de l'insuline dans le plasma
$I = \frac{I_p}{V_I}$	(3.22)	Conversion de l'insuline plasmatique (pmol/dL)
$\emptyset \xrightarrow{k_i I} I_1$	(3.23)	Induction du délai de l'arrivée de l'insuline vers le foie
$I_1 \xrightarrow{k_i} I_d$	(3.24)	Délai de l'arrivée de l'insuline au foie
$I_d \xrightarrow{k_i} \emptyset$	(3.25)	Arrivée de l'insuline au foie
$X \xrightarrow{p_{2u}} \emptyset$	(3.26)	Dégradation de l'insuline dans le tissu interstitiel
$\emptyset \xrightarrow{p_{2u}(I-I_b)} X$	(3.27)	Circulation de l'insuline plasmatique vers le tissu interstitiel

FIGURE 3.7 – Ensemble des réactions chimiques permettant de reproduire le modèle de référence, à partir des équations différentielles.

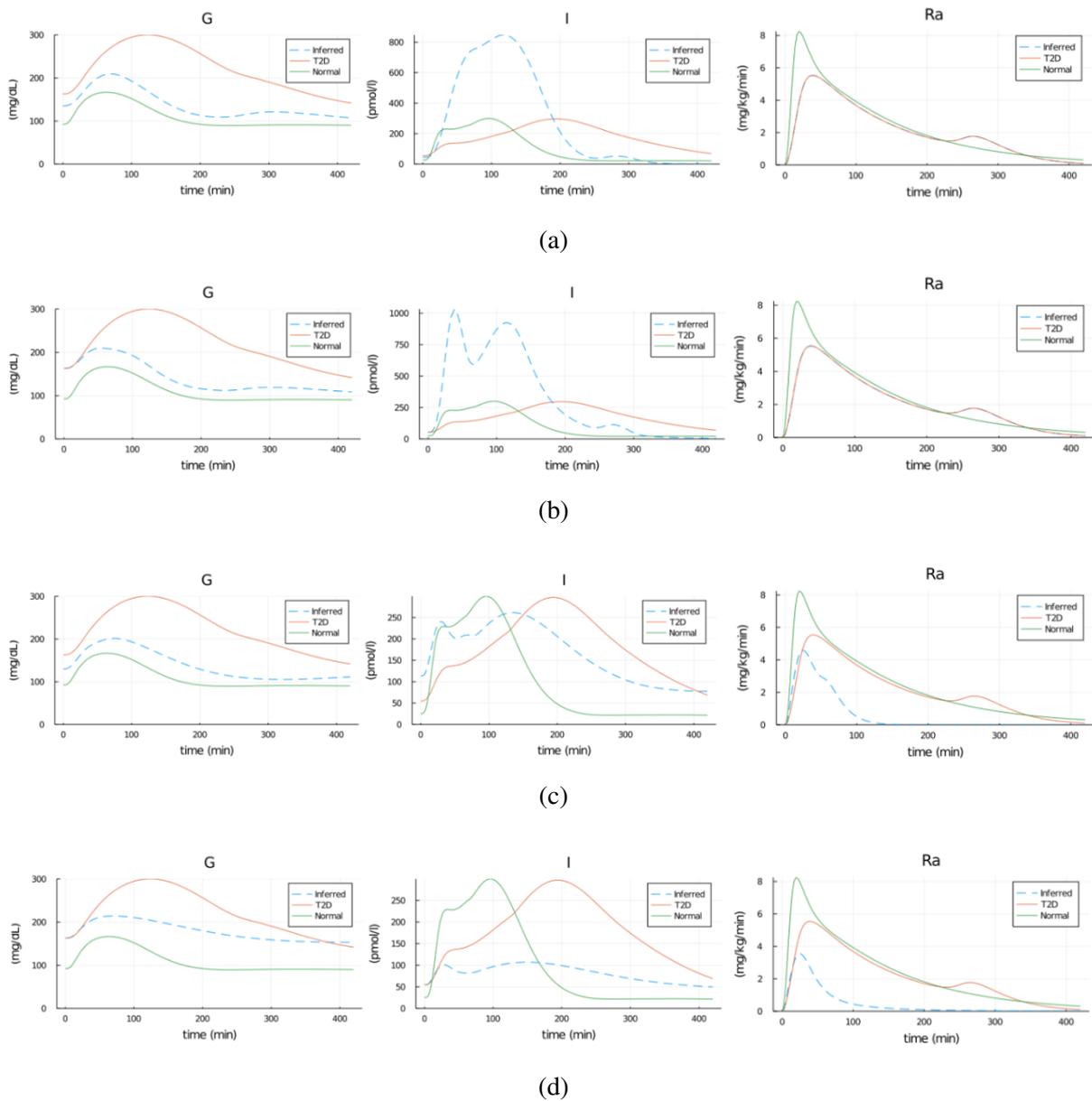


FIGURE 3.8 – Glucose (G), insuline (I), et taux d'apparition (Ra) après estimation des paramètres, avec (a et c) et sans (b et d) les concentrations basales, en cherchant à faire correspondre le modèle « DT2 » au modèle « normal ». Pour les sous-figures a et b, l'estimation n'est effectuée que sur le sous-système pancréas (IK+S). Pour les sous-figures c et d, l'estimation n'est effectuée que sur le sous-système gastro-intestinal (Ra). Courbes vertes : modèle « normal » avec les paramètres correspondant à un individu sain. Courbes rouges : modèle « DT2 » avec les paramètres correspondant à un individu diabétique. Courbes bleues en tirets : estimation des paramètres sur le modèle « DT2 » en cherchant à le faire correspondre au modèle « normal ».

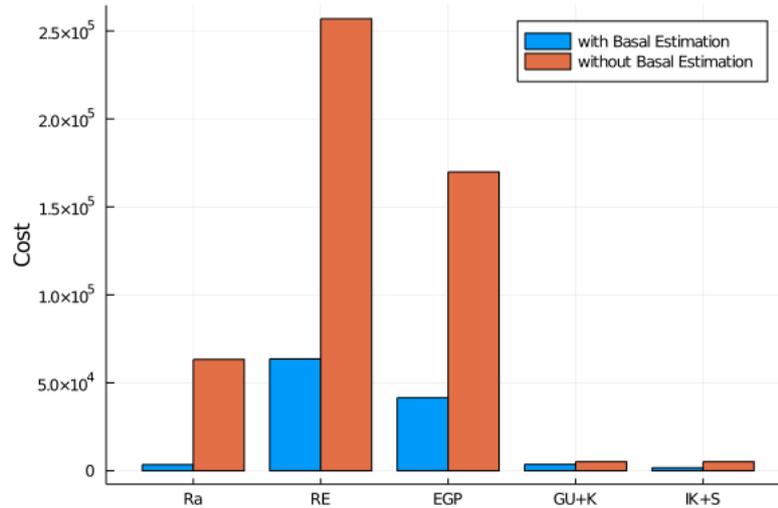


FIGURE 3.9 – Perte de l’estimation des paramètres appliquée sur les différents compartiments.

efficacement les paramètres de ce compartiment pour correspondre aux données. Cela peut être interprété biologiquement, toute proportion gardée, comme une incapacité dudit compartiment à rétablir une glycémie normale. Nous appliquons le raisonnement inverse pour les plus faibles valeurs résiduelles du graphique en barres.

3.5 Estimation des paramètres des diabétiques obèses et RYGB

Notre deuxième objectif est de tester si le modèle de (Dalla Man et al., 2007) peut prédire l’évolution des concentrations de glucose et d’insuline obtenues à partir de notre propre jeu de données de patients diabétiques ayant subi une chirurgie RYGB. Dans cette section, nous utilisons donc les données OBEDIAB présentées à la section 3.2. Pour chaque patient, nous utilisons des données avant (désignées ci-après par *visite A*) et 3 mois après (*visite B*) la chirurgie. Nous estimons d’abord tous les paramètres afin d’ajuster les données temporelles de glucose et d’insuline du jeu de données de la *visite A*. Le modèle que nous obtenons est appelé le modèle de la *visite A*. La figure 3.10 montre les concentrations plasmatiques de glucose et d’insuline prédites par ce modèle ainsi que les points de données utilisés pour l’estimation. Ensuite, comme précédemment, nous estimons chaque sous-ensemble de paramètres sur le jeu de données de la *visite B*. Ici, nous ne considérons que le cas où nous estimons également les concentrations basales, qui ont effectivement changé 3 mois après la chirurgie. Nous présentons dans la figure 3.11 l’évolution de G , I et Ra après estimation des paramètres des compartiments gastro-intestinal et pancréatique.

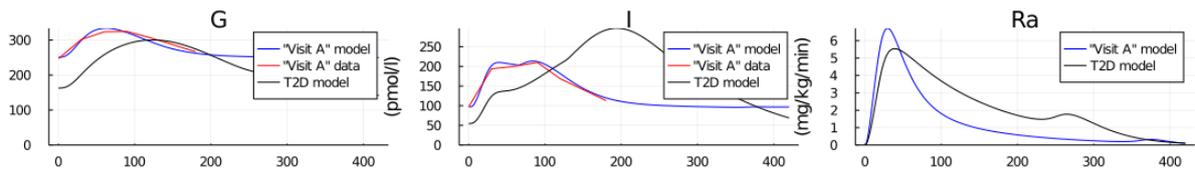
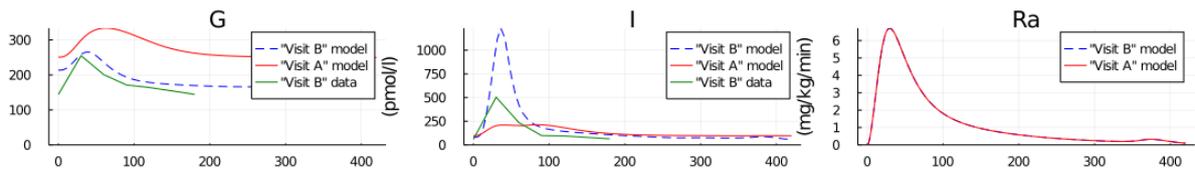
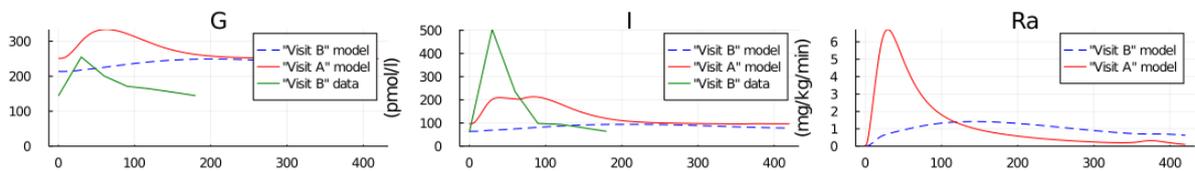


FIGURE 3.10 – Estimation de tous les paramètres, y compris les concentrations basales, pour correspondre aux données de la « Visite A » (en rouge). Courbes noires : modèle initial de Dalla Man avec les valeurs de paramètres $DT2$ fournies dans l'article. Courbes bleues : même modèle après estimation sur les données « Visite A ».



(a)



(b)

FIGURE 3.11 – Glucose (G), insuline (I), et taux d'apparition (Ra) après estimation des paramètres, comprenant les concentrations basales, en cherchant à faire correspondre le modèle « Visite A » aux données « Visite B ». Pour la sous-figure a, l'estimation n'est effectuée que sur le sous-système pancréas (IK+S) tandis que pour la sous-figure b, l'estimation n'est effectuée que sur le sous-système gastro-intestinal (Ra). Courbes vertes : données OBEDIAB « Visite B » correspondant à 3 mois après chirurgie. Courbes rouges : modèle « Visite A », voir figure 3.10. Courbes bleues en tirets : estimation des paramètres sur le modèle « Visite A » en cherchant à le faire correspondre aux données « Visite B ».

3.6 Discussion

Estimer les valeurs basales de la sécrétion d'insuline, de la production endogène de glucose et de l'utilisation du glucose insulino-dépendante peut être interprété comme une prédiction de l'effet « à long terme » des changements de paramètres *dans le meilleur des cas* (puisque le modèle n'incorpore aucun mécanisme de récupération à long terme). Ainsi, ne pas estimer ces valeurs basales peut être interprété comme une prédiction « à court terme » (ou pire cas).

3.6.1 Du modèle de DT2 au modèle normal

Nos ré-estimations des paramètres basées sur les données de (Dalla Man et al., 2007) prédisent (figure 3.9) que les compartiments les plus performants pour restaurer une glycémie normale sont le pancréas et les tissus, puis le tractus intestinal. Comme prévu, estimer les concentrations basales (c'est-à-dire, l'effet à long terme) améliore les performances notamment pour le tractus intestinal (voir également la courbe G dans les figures 3.8b et 3.8c), ce qui est cohérent avec l'observation expérimentale. Cependant, la performance du pancréas devrait être modulée. En effet, les figures 3.8b et 3.8a montrent que, pour améliorer la glycémie, une concentration plasmatique d'insuline très élevée est nécessaire si seuls les paramètres du pancréas sont modifiés. Cela semble physiologiquement irréaliste, ce qui signifie que la bonne performance de ce compartiment est surestimée. De même, l'estimation des paramètres du tractus gastro-intestinal à court terme (figure 3.8d) indique une diminution irréaliste de la concentration plasmatique d'insuline. Enfin, l'estimation « à long terme » des paramètres du tractus gastro-intestinal (figure 3.8c) permet une bonne amélioration de la glycémie avec une concentration réaliste d'insuline et une diminution de l'absorption intestinale, ce qui est cohérent avec l'observation.

3.6.2 Estimation des paramètres des diabétiques obèses et de RYGB

À partir de la Figure 3.10, il est principalement intéressant de remarquer que les données de DT2 de (Dalla Man et al., 2007) sont significativement différentes des nôtres. En effet, la glycémie dans *la visite A* (courbe rouge, titrée « G ») est plus élevée que dans le *modèle DT2* (courbe noire, titrée « G ») tandis que l'insulinémie dans *la visite A* (courbe rouge, titrée « I ») est plus faible que dans le *modèle DT2* (courbe noire, titrée « I »). Malgré cette différence, l'estimation des paramètres donne un résultat satisfaisant lorsque *tous* les paramètres sont réestimés. La qualité du résultat est sensible à la méthode d'estimation des paramètres.

Dans la figure 3.11, les paramètres sont soumis à la réestimation pour le pancréas (IK+S) et le tractus gastro-intestinal (Ra). Un tel processus peut être interprété comme des simulations d'interventions thérapeutiques (chirurgicales ou médicamenteuses) ciblant respectivement le pancréas et le tractus gastro-intestinal. Aucune tentative d'estimation ne semble satisfaisante. D'une part, libérer les paramètres IK+S semble être satisfaisant pour correspondre à la glycémie mais surestime clairement l'insulinémie (cf. figure 3.11a). D'autre part, en libérant les paramètres Ra, le taux d'apparition diminue comme observé dans les données expérimentales

(basées sur nos données de D-xylose, un marqueur du glucose alimentaire). Néanmoins, l'estimation des paramètres échoue complètement à reproduire la glycémie et l'insulinémie de la *visite B* (cf. figure 3.11b).

3.7 Nécessité d'un nouveau modèle

L'estimation des paramètres réalisée par les auteurs de (Dalla Man et al., 2007) est basée sur des données d'entraînement générées par la technique de clampage traceur-tracee (Cobelli et al., 1992), qui, malgré son efficacité, reste une méthode peu commune et complexe pour surveiller les solutés exogènes. En pratique, la glycémie plasmatique et l'insulinémie sont généralement les seules données cliniques accessibles. Cependant, comme le suggère notre évaluation du modèle, cela peut poser des problèmes d'identifiabilité des paramètres. Pour surmonter ce problème, nous prévoyons d'utiliser les données supplémentaires disponibles sur le D-xylose, un marqueur pouvant être utilisé pour ajuster le taux d'apparition (Ra) (Bjarnason and Sherwood, 2014). Nous prévoyons également, en exploitant le profil de vraisemblance et l'analyse de sensibilité, d'étudier la réduction du modèle afin d'éliminer les sources potentielles de non-identifiabilité. D'autres ensembles de données originaux sont actuellement utilisés pour l'estimation des paramètres, générés à partir d'expériences sur les mini-porcs. De tels modèles biologiques permettent plus d'expériences et de reproductibilité, et diminuent les variabilités individuelles, améliorant ainsi la fiabilité de l'estimation des paramètres. Dans cette direction, une autre possibilité est d'utiliser des ensembles de données disponibles publiquement.

Il convient de noter que l'échec de l'estimation des paramètres peut être dû à des problèmes structurels inhérents au modèle de (Dalla Man et al., 2007), dont le sous-modèle de l'AIG est largement simplifié. Par exemple, il ignore le taux d'absorption du glucose non uniforme spatialement le long de l'intestin et la sécrétion d'incrétines. Nous prévoyons d'étendre le modèle de (Dalla Man et al., 2007) avec ces aspects tout en simplifiant les autres pour surmonter les problèmes d'identifiabilité qui pourraient surgir des paramètres supplémentaires liés au tractus gastro-intestinal.

Chapitre 4

Modélisation de l'absorption intestinale de glucose à partir des données de D-xylose

À partir des limites observées au chapitre précédent sur notre modèle de référence, nous avons décidé de réorienter notre modélisation vers une approche parcimonieuse qui tient compte de la complexité de l'absorption intestinale, tout en essayant de se passer du double traceur. Pour cela, nous avons décidé de modéliser la dynamique du D-xylose, au lieu du glucose, car celui-ci pourrait simuler une absorption d'un repas complexe par simple ingestion, et sa dynamique est le reflet exclusif du comportement gastro-intestinal, puisque ce sucre n'est pas utilisé par les tissus et est directement excrété de l'organisme, une fois passée la barrière intestinale. De plus, nous avons décidé de passer d'un jeu de données cliniques (OBEDIAB) à un jeu de données pré-cliniques comprenant des cochons miniatures (qui peuvent être indifféremment évoqués sous le nom de « minipigs » dans la suite) issu de notre collaboration avec l'équipe *Translational Research Laboratory for Diabetes (INSERM UMR1190)*. Ce changement nous a permis d'obtenir des données sous des conditions expérimentales inédites, infaisable sur des humains, pour tester nos hypothèses le plus idéalement possible.

4.1 Motivation

Comme nous l'avons précédemment énoncé dans nos hypothèses, un des marqueurs de possibles risques de DT2 chez les patients est un changement dans la réponse glycémique postprandiale (Bergman et al., 2018), c'est-à-dire une modification de la dynamique dans le temps de la concentration de glucose dans le sang après un repas. Il a été démontré que l'un des principaux contributeurs de cette réponse postprandiale est l'absorption intestinale du glucose (AIG) (Tricò et al., 2019; Baud et al., 2016a). Par conséquent, le suivi de l'AIG permettrait une meilleure prévention du DT2 chez les patients à risque, améliorerait le traitement des patients affectés et, plus généralement, permettrait de mieux comprendre les mécanismes physiologiques à l'œuvre dans cette maladie. À cet égard, modéliser la dynamique glycémique postprandiale dans le sang est crucial pour prédire comment un changement de l'AIG peut affecter la concentration de glu-

cose dans le sang et concevoir de nouveaux marqueurs du diabète. Cela nécessite, en particulier, de modéliser le *taux d'apparition de glucose exogène* (noté Ra_G dans ce chapitre, pour éviter la confusion avec la même grandeur appliquée au D-xylose), c'est-à-dire le taux de glucose provenant du repas. Cependant, calibrer un tel modèle implique la mesure expérimentale de ce taux, ce qui est un défi difficile (voir chapitre 1 et notamment section 1.3, qui revient en détail sur la méthode dite du « double traceur »). En effet, la mesure directe de Ra_G nécessite un accès à la veine porte, ce qui est généralement difficilement réalisable voire impossible chez les humains. De plus, il ne peut être déduit d'autres variables facilement observables comme la concentration de glucose dans le sang, car d'autres mécanismes se produisent simultanément, à chaque instant : l'excrétion de glucose (par les reins), la production de glucose (par le foie) et la métabolisation (par les tissus), régulées par l'insuline. L'approche actuelle pour mesurer expérimentalement Ra_G consiste à réaliser des tests de glucose oral avec plusieurs traceurs isotopiques qui consistent à ingérer du glucose marqué (Toffolo et al., 2006). Ils permettent de distinguer les différents flux de glucose dans le sang et de mesurer la fraction provenant du repas. Néanmoins, ces tests sont invasifs et complexes à mettre en place dans une routine clinique ou de laboratoire, et en tant que tels, ne permettent pas de recueillir des données sur de grands cohortes de patients.

Nous proposons ici une approche alternative pour mesurer Ra_G qui utilise le D-xylose comme marqueur de l'AIG.

Pour rappel, le D-xylose est un sucre absorbé par l'intestin grêle et éliminé par les reins, de la même manière que le glucose, mais sans métabolisation significative par d'autres organes, y compris le foie, contrairement au glucose. Par conséquent, comme il n'est pas stocké, libéré ou régulé à partir d'une source endogène, nous pouvons supposer que la surveillance de la concentration de D-xylose dans le sang reflète principalement son activité gastro-intestinale. De plus, il est de loin plus simple à utiliser que les méthodes de traceurs isotopiques. L'utilisation du D-xylose comme marqueur quantitatif de l'AIG dans les paramètres cliniques et expérimentaux a déjà été proposée (Fujita et al., 1998; Baud et al., 2016a; Goutchtat et al., 2022). Cependant, une telle mesure directe n'est pas parfaitement précise car elle ignore l'effet de l'élimination du D-xylose qui a lieu en plus de l'absorption intestinale, et ne peut pas distinguer les rôles respectifs de la vidange gastrique et de l'absorption intestinale dans le taux d'apparition du D-xylose.

Jusqu'à présent, aucun modèle mécanistique de la dynamique du D-xylose n'a été proposé. En ce qui concerne la dynamique du glucose, la plupart des modèles de l'apparition du glucose dans le sang reposent sur une modélisation complexe de la vidange gastrique, comme celle d'Elashoff et ses collaborateurs (Elashoff et al., 1982), décrite dans la section 2.2.2. Plus tard, Dalla Man et ses collaborateurs ont exposé les limitations de l'approche précédente d'Elashoff, et ont proposé un modèle gastro-intestinal complet, décrivant non seulement la vidange gastrique, mais aussi l'absorption intestinale du glucose en condition postprandiale (Dalla Man et al., 2006a), décrit dans la section 2.2.4. Dans ce modèle, l'absorption intestinale est réduite à un seul flux avec un taux constant, tandis que la vidange gastrique implique une équation

complexe avec 5 paramètres. Ce modèle est capable de simuler leurs propres données de glucose exogène, obtenues avec la méthode isotopique triple traceur, considérée comme l'approche expérimentale de référence pour mesurer Ra_G . Salinari et ses collaborateurs ont proposé un modèle spatial d'absorption et de transit intestinal, défini au moyen d'un système d'équations aux dérivées partielles, dépendant du temps et de la position le long de l'intestin (Salinari et al., 2011), décrit à la section 2.2.5. Le taux de transit a été déterminé par leurs données spécifiques, dépendant principalement de la longueur de l'intestin. Plus important encore, dans ce modèle spatial, nous pouvons considérer un taux d'absorption intestinal non uniforme le long de l'intestin. Cette hypothèse est en effet considérée comme réaliste et les auteurs montrent que différentes distributions spatiales d'absorption peuvent entraîner des dynamiques d'apparition du glucose différentes.

Dans ce chapitre, nous proposons une nouvelle évaluation de l'AIG basée sur un modèle. Plus précisément, nous étudions un modèle physiologique « multicompartimenté » (nous définissons ce point crucial dans la sous-section 4.4.2) de la dynamique du D-xylose, qui est composé :

- d'une absorption intestinale répartie sur un nombre arbitraire de variables intestinales, représentant les sections successives de l'intestin grêle (d'où la notion de « multicompartiment »);
- d'un transit intestinal constant entre les variables précédemment évoquées;
- d'une vidange gastrique exponentielle et d'élimination du D-xylose;

Ce modèle peut être considéré comme une version simplifiée et discrétisée du modèle d'absorption intestinale de Salinari *et al.*. Bien que simple, nous montrons que notre modèle peut simuler nos données, qui sont des séries temporelles de D-xylose obtenues dans différentes conditions expérimentales (administration orale et jéjunale de D-Xylose) avec une bonne précision, tout en le confrontant à deux autres modèles candidats, que nous évoquerons dans les sections suivantes. Surtout, ce modèle peut prédire Ra_G , validé avec des données de traceurs. Nous montrons également que les taux de vidange gastrique et d'absorption, en particulier, sont identifiables. De plus, pour déchiffrer la contribution relative de la vidange gastrique et de l'absorption intestinale à la dynamique du D-xylose, nous montrons que le modèle alternatif de Dalla Man *et al.* (Dalla Man et al., 2006a), mettant l'accent sur la vidange gastrique, ne peut pas voir ses paramètres s'ajuster aussi bien pour correspondre à nos données expérimentales. Enfin, nous avons réalisé une analyse de sensibilité pour déterminer lequel, du taux de vidange gastrique et de l'absorption intestinale, a l'impact le plus significatif sur la quantité totale de D-xylose absorbée après 180 minutes. Nous montrons que cette quantité est plus sensible à l'absorption intestinale et que le D-xylose peut donc potentiellement faire office de marqueur de l'AIG facile à utiliser dans le cadre clinique. Toutes les données et expériences sont disponibles sur : <https://zenodo.org/records/10136595>. L'ensemble des résultats présentés dans ce chapitre a fait l'objet d'une publication dans le cadre de la conférence *BIOINFORMATICS 2024* (Dursoniah et al., 2024).

4.2 Jeux de données des minipigs

Pour répondre à notre problématique, différentes expériences ont été menées pour surveiller le glucose, et surtout le D-xylose, dans le sang après une prise d'un bolus via une administration intestinale ou une administration orale. Cela implique deux sous-populations de porcs miniatures produisant chacune plusieurs ensembles de données. Les individus de la première sous-population ont subi une administration orale et une administration jéjunale, simultanément.

Le *jeu de données du bolus oral* permet de surveiller le D-xylose sanguin à l'état normal après une administration orale du repas. Dans le *jeu de données bolus intestinal* (ou *jéjunale*), l'estomac est contourné et le repas est directement administré dans l'intestin grêle. Il s'agit d'un repas dit « mixte » dans le jargon clinique : standardisé, varié (ne comprenant pas uniquement du sucre), mixé, et comprenant 30 g de D-xylose.

Les individus de la deuxième sous-population ont subi une expérience chirurgicale (résection intestinale) pour évaluer la réponse du glucose dans le sang après un changement dans les processus d'absorption. Cette sous-population est intéressante pour comparer le taux d'apparition du glucose exogène (Ra_G) avec le taux d'apparition du D-xylose (Ra_X) dans des conditions normales et expérimentales afin de démontrer la pertinence du D-xylose pour étudier le comportement de l'AIG. En effet, cette sous-population a bénéficié d'une technique de référence pour surveiller leur AIG, connue sous le nom de double traceur. Pour rappel, cette technique implique l'utilisation de deux isotopes du glucose, donc marqués différemment, pour distinguer le glucose provenant d'une source exogène et le glucose provenant d'une source endogène (généralement produit par le foie). Cette sous-population a ainsi produit quatre ensembles de données : avant et après la chirurgie, mesurant à la fois les concentrations de D-xylose et de glucose dans le sang. Le *jeu de données bolus oral avant la résection intestinale* permet de surveiller le glucose sanguin à l'état normal (avant la chirurgie) après une administration orale du repas. Dans le *jeu de données bolus oral après la résection intestinale*, environ 80% de la partie moyenne de l'intestin grêle a été retirée. Après un temps de récupération, une administration orale du repas est effectuée. Tous ces ensembles de données sont utilisés pour calibrer les modèles.

Parmi notre jeu de données complet, nous avons également le *jeu de données bolus intraveineux*. Il s'agit de la sous-population qui s'est vu administré du D-xylose par voie intraveineuse. La mesure du D-xylose plasmatique dans cette condition d'administration nous a permis d'évaluer le volume de distribution du D-xylose, ainsi que son taux d'élimination de l'organisme (très majoritairement sous forme d'excretion rénale).

- En résumé, nous disposons d'un jeu de données sous différentes conditions expérimentales,
- administration intraveineuse
 - administration jéjunale
 - administration orale après résection intestinale

Ces conditions ont pour cas contrôle systématique, une ingestion orale de D-xylose avant toute

modification de l'organisme. Dans la section suivante, nous montrerons comment nous avons exploité le jeu de données intraveineux pour obtenir un moyen simple d'estimer le volume de distribution du D-xylose.

4.3 Volume de distribution

La première étape dans notre démarche de modélisation consiste à estimer le volume de distribution du D-xylose V_{D_X} , qui représente le volume total de liquide que le D-xylose peut occuper une fois absorbé par l'intestin : il sert de référence pour calculer les concentrations de D-xylose dans le corps au lieu des quantités. Il est généralement normalisé par le poids corporel, donc la dimension de V_{D_X} est dL/kg. Après une expérience d'injection intraveineuse, il est défini par $V_{D_X} = \frac{D_X}{BW \cdot X_p}$ où D_X est la dose administrée de D-xylose (mg), X_p est la concentration de D-xylose dans le sang (mg/dL) lorsque D_X est entièrement administré instantanément, et BW est le poids corporel (kg).

De ce travail, nous avons trouvé une corrélation linéaire significative entre le poids corporel et le volume de distribution, comme le montre la figure 4.1. Cette observation nous a permis de déduire, à partir de leur poids corporel BW , le volume de distribution, noté $V_{D_X}(BW)$, des cochons miniatures qui n'ont pas subi d'expérience d'injection intraveineuse.

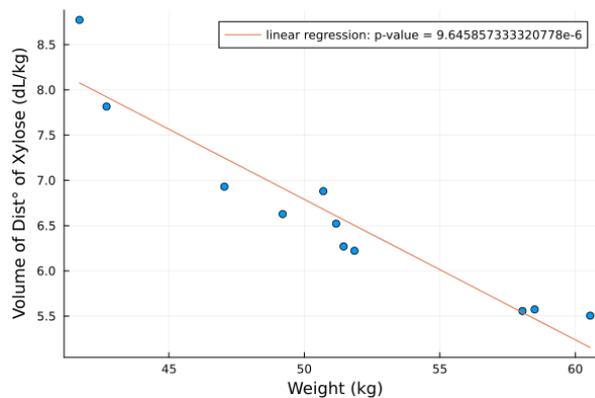
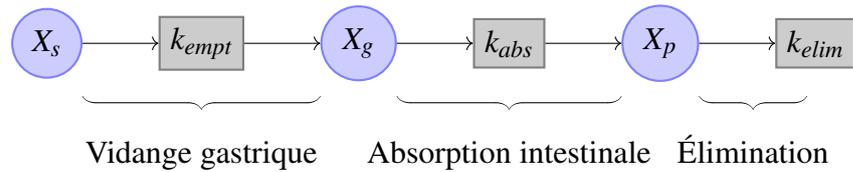


FIGURE 4.1 – Tracé de la corrélation linéaire significative entre le poids de chaque individu ayant subi une injection intraveineuse et leur volume de distribution respectif. Coefficient de Pearson : $-0,94725$.

4.4 Modèles de D-xylose

Les modèles présentés dans cette section sont tous formalisés comme des systèmes d'équations différentielles ordinaires (*EDO*). Tous les modèles ont été définis de manière à minimiser le nombre d'équations, de variables et de paramètres. Nous nous concentrerons davantage sur le modèle sélectionné dans la section suivante.

4.4.1 Modèle oral « simple »



(a) Réseau de réactions

$$\begin{aligned}
 \frac{dX_s(t)}{dt} &= -k_{empt} \cdot X_s(t) & X_s(0) &= \frac{D_X}{BW \cdot V_{DX}} \\
 \frac{dX_g(t)}{dt} &= k_{empt} \cdot X_s(t) - k_{abs} \cdot X_g(t) & X_g(0) &= 0 \\
 \frac{dX_p(t)}{dt} &= Ra_X(t) - k_{elim} \cdot X_p(t) & X_p(0) &= 0 \\
 Ra_X(t) &= k_{abs} \cdot X_g(t)
 \end{aligned}$$

(c) Conditions initiales

(b) Système d'EDO

FIGURE 4.2 – Modèle $SIMP_o$

Le premier modèle, que nous appelons $SIMP_o$, est le plus simple et est donné dans la figure 4.2a sous forme de réseau de réactions (représenté ici en utilisant une notation graphique de type réseau de Petri). Il est composé de 3 espèces chimiques correspondant au D-xylose dans 3 compartiments différents, à savoir dans l'estomac (X_s), dans l'intestin (X_g) et dans le plasma (X_p). Il implique 3 flux de D-xylose :

- la vidange gastrique, de l'estomac vers l'intestin, avec un taux k_{empt} ,
- l'absorption intestinale, de l'intestin vers le plasma, avec un taux k_{abs} ,
- et l'élimination du plasma avec un taux k_{elim} .

Le système d'EDO dérivé de ce réseau suivant la cinétique d'action de masse est donné dans la figure 4.2b avec les conditions initiales données dans la figure 4.2c. $X_s(t)$, $X_g(t)$, et $X_p(t)$ sont les concentrations (en mg/dL) de D-xylose dans les compartiments correspondants à l'instant t (min). Les taux k_{empt} , k_{abs} et k_{elim} sont exprimés en min^{-1} . La vidange gastrique aurait pu être formalisée avec des modèles courants tels que proposés par Elashoff (Elashoff et al., 1982) et par Dalla Man (Dalla Man et al., 2006b). Cependant, les deux modèles auraient introduit des paramètres supplémentaires, ce qui aurait été contraire à notre approche de modélisation minimaliste, sans nécessairement améliorer les résultats de l'estimation des paramètres. Dans cet article, le terme « élimination » est un terme générique pour désigner à la fois la clairance rénale (excrétion sous forme d'urine) du D-xylose et sa métabolisation, les deux entraînant une diminution de la concentration sanguine de D-xylose après un certain temps. Cependant, il est admis que la clairance rénale a un impact plus important sur l'élimination du D-xylose que sa métabolisation par les tissus et dans l'intestin, qui peuvent être considérées comme négligeables.

Par conséquent, un seul taux d'élimination du D-xylose (du compartiment plasma) est considéré dans nos modèles.

4.4.2 Modèle oral « multi-compartiments »

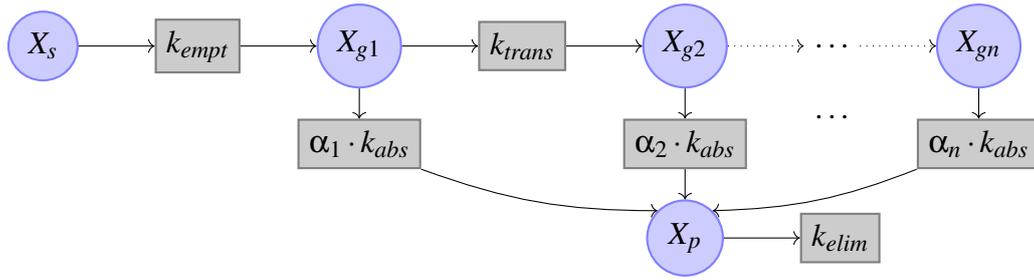
Cette section présente le modèle principal de ce travail ainsi que les résultats fournis par ce modèle. Ce modèle est celui qui sera utilisé dans le reste de ce travail et est présenté à la figure 4.3. Il est composé des variables X_s (D-xylose dans l'estomac), X_p (D-xylose dans le plasma), et X_{g1}, \dots, X_{gn} (D-xylose dans différentes portions successives du tractus intestinal, où $n = 10$ pour toutes les analyses numériques suivantes).

Le taux k_{empt} (min^{-1}) modélise la vidange gastrique de l'estomac vers l'intestin. Ce taux est volontairement simplifié, contrairement à d'autres modélisations telles que (Dalla Man et al., 2006a) et comme cela sera discuté dans la section 4.6.2. Le transit intestinal est modélisé comme un flux de D-xylose de chaque compartiment X_{g_i} vers le suivant, $X_{g_{i+1}}$. Nous supposons que ce flux est uniforme avec un taux k_{trans} (min^{-1}) défini par $k_{trans} = \frac{1}{\tau}$ et $\tau = \frac{L}{u \cdot n}$ où τ est le temps nécessaire pour le transit à travers un compartiment (min), L est la longueur de l'intestin grêle (estimée à 1100 cm, la longueur moyenne obtenue à partir de la chirurgie réalisée pour le *jeu de données bolus oral après résection*), et u la vitesse du transit intestinal (empiriquement fixée à 6 cm/min, une estimation pour le modèle intestinal de (Salinari et al., 2011)).

L'absorption intestinale globale, de l'intestin vers le plasma, est modélisée avec un taux k_{abs} (min^{-1}). Cependant, la distribution de ce taux d'absorption le long de l'intestin est supposée non uniforme. Pour cela, pour chaque variable X_{g_i} , le taux d'absorption est modulé par un paramètre strictement positif α_i . La somme de tous les paramètres $\alpha_1, \dots, \alpha_n$ est égale à 1, de sorte que le taux d'absorption globale (la somme des taux de chaque compartiment de l'intestin) est ainsi k_{abs} . Notez que si la distribution de ces paramètres est uniforme (c'est-à-dire, $\alpha_i = 1/n$ pour tous les i), alors ce modèle est équivalent au modèle simple de la section 4.4.1. Nous n'imposons aucune distribution particulière, et les paramètres α_i sont estimés par la suite.

4.4.3 Modèle oral « à rétention »

Une alternative au système précédent, au lieu d'augmenter le nombre de compartiments intestinaux, consiste à retarder l'absorption intestinale avec un compartiment de rétention X_d (mg/dL) qui s'équilibre avec la concentration de l'intestin grêle X_g avec des taux k_{d1} (min^{-1}) et k_{d2} (min^{-1}) comme donné à la figure 4.4a. Ce retard peut être interprété comme résultant de la compétition entre le glucose et le D-xylose pour l'absorption. En effet, nos expériences impliquaient toujours les deux composés, et il est connu que le glucose a une meilleure affinité avec les transporteurs que le D-xylose, dont l'absorption est ainsi retardée. Cependant, ce retard peut également modéliser le temps de digestion des glucides complexes contenus dans un repas réel et qui doivent être décomposés en glucose avant d'être absorbés, faisant du D-xylose un bon candidat pour la modélisation AIG des repas réels.



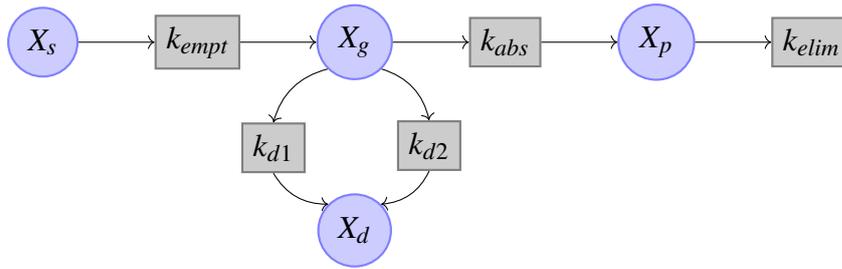
(a) Réseau de réaction

$$\begin{aligned}
 \frac{dX_s(t)}{dt} &= -k_{empt} \cdot X_s(t) \\
 \frac{dX_{g1}(t)}{dt} &= k_{empt} \cdot X_s(t) - (\alpha_1 \cdot k_{abs} + k_{trans}) \cdot X_{g1}(t) \\
 &\vdots \\
 \frac{dX_{gn}(t)}{dt} &= k_{trans} \cdot X_{g_{n-1}}(t) - \alpha_n \cdot k_{abs} \cdot X_{gn}(t) \\
 \frac{dX_p(t)}{dt} &= Ra_X(t) - k_{elim} \cdot X_p(t) \\
 Ra_X(t) &= k_{abs} \cdot \left(\sum_{i=1}^n \alpha_i \cdot X_{gi}(t) \right) \text{ and } \sum_{i=1}^n \alpha_i = 1
 \end{aligned}
 \quad
 \begin{aligned}
 X_s(0) &= \frac{D_X}{BW \cdot V_{D_X}(BW)} \\
 X_{g1}(0) &= 0 \\
 &\vdots \\
 X_{gn}(0) &= 0 \\
 X_p(0) &= 0
 \end{aligned}$$

(c) Conditions initiales

(b) Système d'EDO

FIGURE 4.3 – Modèle multi-compartiment.



(a) Réseau de réactions

$$\frac{dX_s(t)}{dt} = -k_{empt} \cdot X_s(t)$$

$$\frac{dX_g(t)}{dt} = k_{empt} \cdot X_s(t) - (k_{abs} + k_{d1}) \cdot X_g(t) + k_{d2} \cdot X_d(t)$$

$$\frac{dX_d(t)}{dt} = k_{d1} \cdot X_g(t) - k_{d2} \cdot X_d(t)$$

$$\frac{dX_p(t)}{dt} = Ra_X(t) - k_{elim} \cdot X_p(t)$$

$$Ra_X(t) = k_{abs} \cdot X_g(t)$$

$$X_s(0) = \frac{D_X}{BW \cdot V_{D_X}}$$

$$X_g(0) = 0$$

$$X_p(0) = 0$$

$$X_d(0) = 0$$

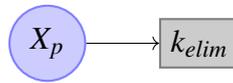
(c) Conditions initiales

(b) Système d'EDO

FIGURE 4.4 – Modèle RET_o

Le système d'EDO correspondant et les conditions initiales sont donnés dans la figure 4.4b et la figure 4.4c.

4.4.4 Modèles intraveineux et intrajéjunal



(a) Réseau de réactions

$$\frac{dX_p(t)}{dt} = -k_{elim} \cdot X_p(t) \quad X_p(0) = \frac{D_X}{BW \cdot V_{D_X}}$$

(b) Système d'EDO

(c) Conditions initiales

FIGURE 4.5 – Modèle IV

Pour l'estimation des paramètres, nous considérons également un modèle « intraveineux », appelé *IV*, qui est utilisé pour correspondre aux données *IV*. Comme décrit dans la figure 4.5, il consiste simplement en une espèce et une équation différentielle. Nous utilisons également des variantes jéjunales des modèles précédents, que nous appelons *SIMP_j*, *MC_j* et *RET_j*, et que nous utilisons pour correspondre aux données jéjunales. Celles-ci sont obtenues simplement en supprimant les variables X_s des modèles oraux et en fixant la valeur initiale de X_g (X_{g1} pour *MC_j*) à $\frac{D_X}{BW \cdot V_{D_X}}$.

4.5 Exploitation des modèles

Sur la base des différents modèles présentés dans la section précédente, nous proposons dans la suite des estimations de paramètres, permettant d'obtenir des modèles complets simulables. Parmi ces modèles, nous avons sélectionné celui qui correspond au mieux à nos données. Le modèle choisi est ensuite soumis à d'autres critères de sélections. Nous commençons par l'analyse d'identifiabilité pratique des paramètres. Puis, nous appliquons ce modèle à la question de la prédiction de la dynamique post-prandiale du glucose (et non du D-xylose) et montrons que le modèle sélectionné se révèle satisfaisant.

4.5.1 Estimation des paramètres

En utilisant nos divers jeux de données expérimentaux disponibles, nous adoptons une stratégie d'estimation des paramètres qui minimise les risques de non-identifiabilité. Dans un premier temps, disposant de données intraveineuses, il a été possible d'estimer le paramètre d'élimination, k_{elim} , à $0,00628 \text{ min}^{-1}$ (figure 4.6). Cette valeur est ensuite utilisée en tant que constante à la prochaine étape d'estimation.

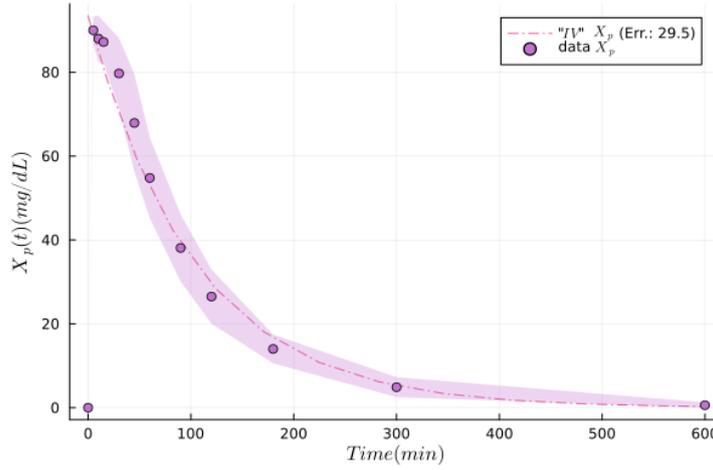


FIGURE 4.6 – Résultats d’estimation du paramètre k_{elim} du modèle intraveineux.

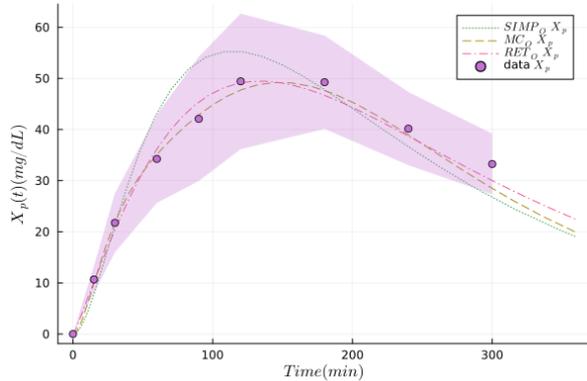
	k_{empt} (min^{-1})	k_{abs} (min^{-1})	k_{elim} (min^{-1})
<i>SIMP</i>	0.779	1.31	0.00628
<i>MC</i>	0.0379	0.222	0.00628
<i>RET</i>	0.0247	0.538	0.00628

TABLE 4.1 – Valeurs de paramètres obtenues en effectuant l’estimation simultanément sur les jeux de données *oral* et *jéjunale* avec, respectivement, un modèle et son variant jéjunale, parmi le modèle simple (*SIMP*), le modèle multi-compartimenté (*MC*) et le modèle à rétention (*RET*). NLL-Loss : 46.581,44, 035,44.264, respectivement.

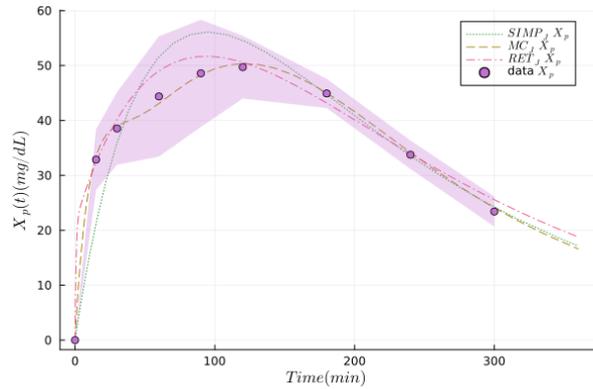
À l’étape suivante, nous estimons les paramètres en utilisant deux jeux de données simultanément : le jeu de données *bolus oral* (sur le modèle principal) et le jeu de données *bolus jéjunale* (sur sa variante jéjunale). Les paramètres estimés sont les taux k_{empt} , k_{abs} et k_{elim} , en pour le modèle multi-compartiment uniquement, les paramètres de distribution d’absorption $\alpha_1, \dots, \alpha_n$. Tous ces paramètres ont été considérés comme communs aux deux modèles, sauf pour k_{empt} qui n’existe pas dans la variante jéjunale.

Techniquement, nous estimons les paramètres sur les valeurs moyennes des données plasmatiques de D-xylose X_p (points violets dans la Figure 4.7) en tenant compte de l’écart-type (zone ombrée en violet) pour minimiser le NLL-Loss (negatif de la log-vraisemblance). Cela est réalisé en utilisant l’algorithme d’optimisation numérique CMA-ES (Covariance Matrix Adaptation Evolution Strategy) (Hansen, 2023). Toutes les étapes de mise en œuvre (prétraitement des données, mise en œuvre du modèle et analyses numériques) ont été réalisées dans le langage de programmation *Julia* (v1.8.2) avec les packages suivants : CMAEvolutionStrategy (v0.2.6), DifferentialEquations (v7.7.0), DiffEqParamEstim (v2.0.1), ModelingToolkit (v8.46.1), Catalyst (v12.3.2), LikelihoodProfiler (v0.5.0) et Plots (v1.38.5).

Les valeurs des paramètres obtenues pour k_{empt} , k_{abs} et k_{elim} sont reportées dans le Tableau 4.1. Comme on peut le voir sur la Figure 4.7, le modèle multi-compartimenté réalise un bon fitting des jeux de données oral et jéjunale. Nous nous sommes ensuite focalisé sur ce mo-



(a) Simulation du modèle obtenu par estimation des paramètres sur le jeu de données bolus oral.



(b) Simulation du modèle jéjunal obtenu par estimation des paramètres sur le jeu de données bolus jéjunal.

FIGURE 4.7 – Simulations des modèles simple (*SIMP*), multi-compartimenté (*MC*) et à rétention (*RET*) et du variant jéjunal, obtenus par estimation des paramètres sur les données de D-xylose des jeux de données *bolus oral* et *bolus jéjunal* (NLL-Loss : 46.581,44,035 et 44.264, respectivement). Les points et les lignes en pointillés représentent la moyenne des valeurs expérimentales, l’enveloppe est l’écart type, et les lignes pleines sont les simulations produites par les différents modèles.

dèle multi-compartimenté pour le soumettre à des critères de sélection additionnelles, comme l’analyse d’identifiabilité, l’analyse de sensibilité, toutes deux traités dans les sous-sections suivantes.

4.5.2 Analyse de l’identifiabilité pratique

L’estimation des paramètres permet de trouver *un* ensemble de valeurs de paramètres qui permet à un modèle de simuler les données, mais ne garantit pas qu’il n’existe pas d’autres valeurs qui pourraient *également* ou *satisfaisamment* simuler ces mêmes données. En effet, les données expérimentales sont bruitées et une partie de l’écart au bon fitting est à attribuer à l’erreur expérimentale. Intuitivement, en supposant des intervalles d’erreur acceptables pour les variables observées, s’il existe un ensemble « unique » de valeurs de paramètres qui permet aux variables observées de correspondre aux données dans ces intervalles, alors le modèle est dit *pratiquement* identifiable. L’analyse de l’identifiabilité est une étape importante pour évaluer la qualité d’un modèle. Dans ce chapitre, nous considérons l’identifiabilité pratique basée sur la méthode du profil de vraisemblance (Raue et al., 2009). Cette méthode examine l’identifiabilité pratique *localement*, c’est-à-dire, près de la valeur estimée d’un paramètre donné. Pour cela, nous avons utilisé le package Julia LikelihoodProfiler (v0.5.0). Cet outil analyse localement chaque paramètre dans un intervalle donné à explorer, ce qui donne une borne d’intervalle de confiance si le paramètre est identifiable, ou aucune si l’outil a atteint les bornes de l’intervalle d’exploration donné ou si aucun gain d’identifiabilité n’est détecté le long de cet intervalle.

	C.I. lower bound	C.I. upper bound
k_{empt}	0.0374	0.0920
k_{abs}	0.222	0.328
k_{elim}	0.00622	0.00708

TABLE 4.2 – Résultats de l’analyse de l’identifiabilité pratique. Ici, chaque valeur représente une borne d’intervalle de confiance (notée *C.I.*) qui est atteinte, signifiant une identifiabilité pratique complète pour le modèle. Les valeurs manquantes auraient indiqué des bornes non atteintes et donc une identifiabilité incomplète.

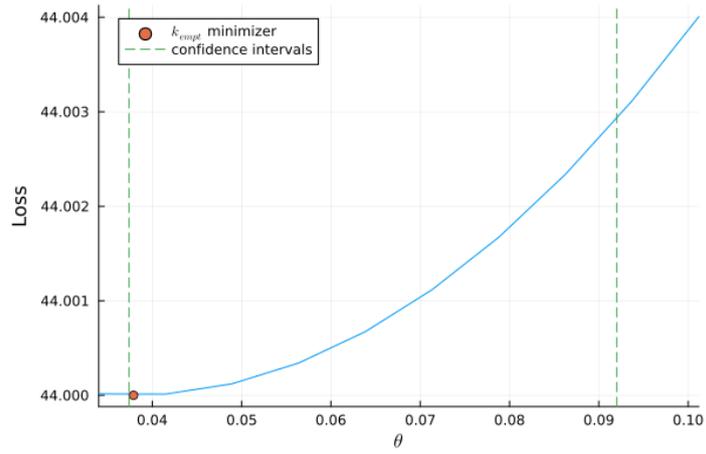
Comme il s’agit d’une étape exploratoire, nous avons donné un intervalle relativement large à explorer pour chaque paramètre d’intérêt. Nous avons fixé l’intervalle de confiance à 95%.

Les intervalles de confiance trouvés pour chaque paramètre sont regroupés dans la table 4.2. L’outil utilisé permet d’obtenir une appréciation visuelle de la recherche d’identifiabilité locale via l’identification des intervalles de confiance (Figure 4.8). Malgré la confiance que nous accordons à cet outil, les courbes de la figure 4.8, notamment les courbes 4.8a et 4.8b, semblent pointer un minimum constant à gauche du minimiseur trouvé (pour k_{empt} et k_{abs}), ce qui pourrait indiquer, en réalité, une identifiabilité partielle bien plus qu’une identifiabilité pratique confirmée pour ces deux paramètres. Une investigation approfondie sur les méthodes de recherche d’identifiabilité pratique pourrait être envisagée dans de futurs travaux.

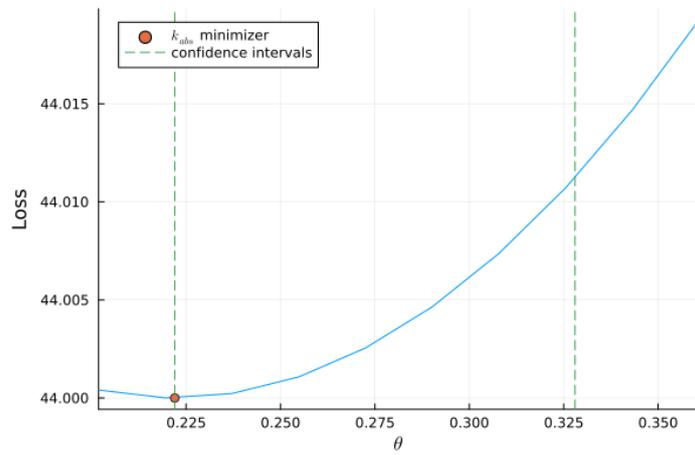
Nous ignorons en fait l’identifiabilité de la vitesse de transit intestinal et de la distribution des paramètres d’absorption qui sont sans pertinence pour le présent travail et, par conséquent, fixés comme constantes pour cette analyse. Ces intervalles indiquent une identifiabilité totale pour les trois paramètres pertinents : k_{empt} , k_{abs} et k_{elim} . Les résultats peuvent être interprétés comme une indication de la bonne pertinence des jeux de données collectés et en particulier de la bonne fiabilité de l’estimation de notre paramètre principal d’intérêt, k_{abs} . Cette analyse a été réalisée dans le même cadre que celui utilisé pour l’inférence dans la Section 4.5.1, c’est-à-dire, sur le modèle principal et sa variante jéjunale simultanément.

4.5.3 Prédiction du taux d’absorption du glucose à partir du jeu de données de résection intestinale

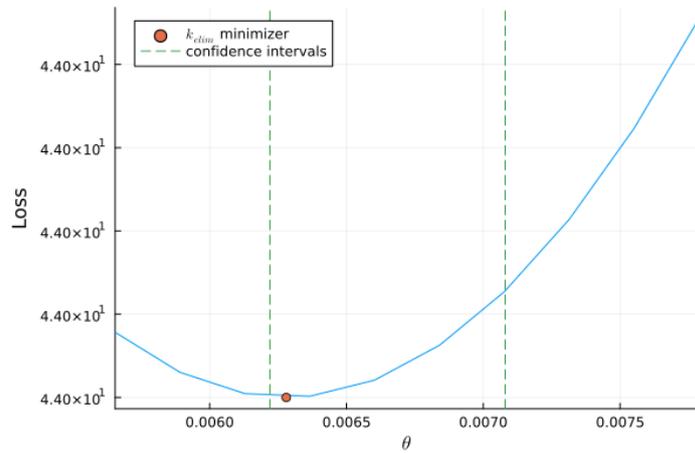
Afin de valider l’utilité de notre modèle, nous testons sa capacité à prédire le *taux d’apparition de glucose exogène* (Ra_G) aussi bien en condition normale qu’après une résection intestinale. Rappelons que ce taux correspond à la partie de la concentration de glucose par unité de temps apparaissant dans le sang et provenant du repas. Ce taux a été surveillé expérimentalement en utilisant le protocole à double traceur. Nous montrons dans la suite que le modèle est capable de s’adapter aux données obtenues après une résection intestinale, ce qui est considéré comme une simulation expérimentale d’un changement dans les mécanismes d’absorption du glucose. Dans cette étude, étant donné que le contexte et les individus diffèrent des jeux



(a) Identifiabilité de la vidange gastrique k_{empt}

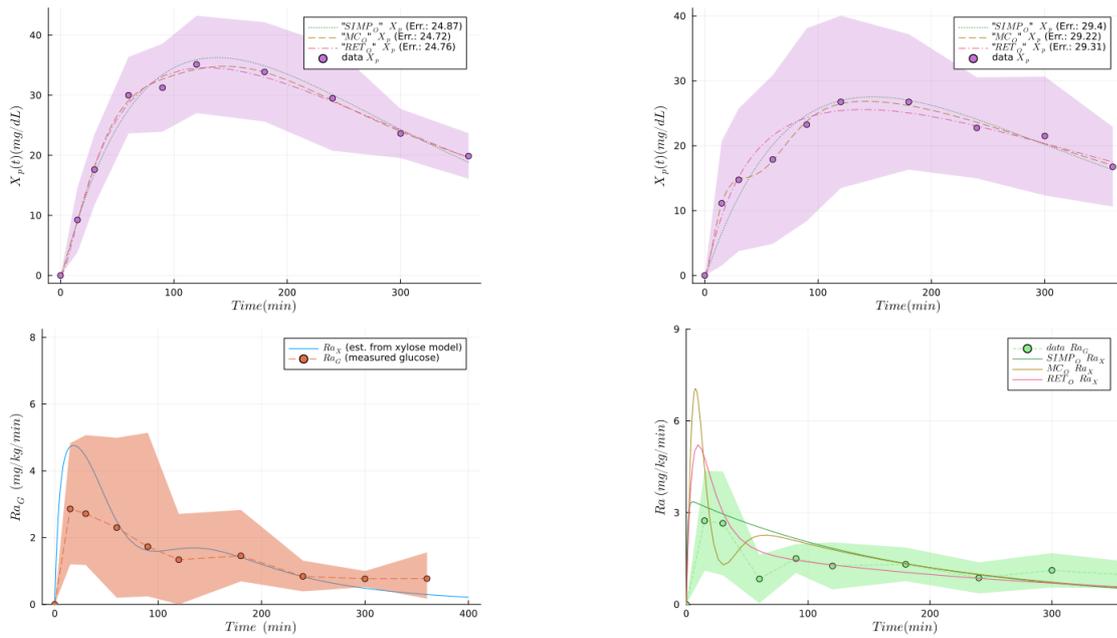


(b) Identifiabilité de l'absorption intestinale k_{abs}



(c) Identifiabilité de l'élimination k_{elim}

FIGURE 4.8 – Identifiabilité pratique de la vidange gastrique, de l'absorption intestinale et de l'élimination du glucose.



(a) Résultats du fitting du modèle sur le jeu de données bolus oral avant résection intestinale.

(b) Résultats du fitting du modèle sur le jeu de données bolus oral après résection intestinale.

FIGURE 4.9 – Comparaison du taux d'apparition exogène du glucose (Ra_G) obtenu par l'expérience à double traceur, la méthode de référence, avec le taux d'apparition généré de D-xylose (Ra_X), obtenu par estimation des paramètres (NLL-Loss : 52.459) sur la même population. Les points et les lignes en pointillés représentent les valeurs expérimentales moyennes, l'enveloppe est l'écart type, et les lignes pleines sont les simulations produites avec le modèle. Les figures du haut représentent le D-xylose plasmatique (utilisé pour estimer les paramètres). Les figures du bas représentent le taux d'absorption simulé de D-xylose (Ra_X) à partir du modèle, et le taux d'absorption observé de glucose (Ra_G) à partir des expériences à double traceur. Dans les deux conditions expérimentales (avant et après résection), nous pouvons observer un curve fitting relativement bon entre le glucose et le D-xylose, malgré l'absence d'un modèle de glucose dans ce travail.

de données utilisés ci-dessus, nous réévaluons tous les paramètres (taux et paramètres d'absorption) sauf l'élimination (considérée inchangée par l'opération) avant et après la résection intestinale. Cependant, puisque notre modèle est conçu pour le D-xylose, nous ne l'entraînons pas directement sur les données de glucose disponibles : au lieu de cela, nous l'entraînons sur les ensembles de données de D-xylose disponibles (ne comportant pas de données à double traceur, mais uniquement la concentration de D-xylose dans le sang au fil du temps) et comparons les résultats avec la dynamique du glucose à partir des ensembles de données de glucose.

Enfin, nous comparons le taux d'apparition de D-xylose (Ra_X) calculé à l'aide du modèle (avec la formule donnée dans la Figure 4.3b) et le comparons avec les données expérimentales de Ra_G (le taux d'apparition de *glucose* exogène) obtenues avec la méthode à double traceur. Ce résultat est présenté dans la Figure 4.9. Comme nous pouvons le voir, bien que les valeurs des paramètres aient été estimées sur les mesures plasmatiques de *D-xylose*, le modèle donne une prédiction relativement satisfaisante du taux d'apparition de glucose exogène (Figure 4.9, graphiques inférieurs). Cela tend à indiquer que le D-xylose pourrait être un marqueur acceptable pour l'absorption du glucose. Remarquez que la différence des taux d'apparition entre le glucose et le D-xylose après une résection pourrait refléter, d'une part, que Ra_X est le reflet exclusif de la vidange gastrique et de l'absorption intestinale, tandis que, d'autre part, Ra_G reflète ces deux mécanismes en plus de la métabolisation inévitable du glucose hépatique, malgré l'utilisation de la méthode de référence.

4.6 Rôles relatifs de la vidange gastrique et de l'absorption intestinale

Dans cette section, nous proposons de comparer les rôles relatifs que jouent la vidange gastrique et l'absorption intestinale dans l'apparition de D-xylose dans le sang, selon notre modèle. Pour cela, nous réalisons d'abord une analyse de sensibilité globale, conçue pour évaluer l'impact des paramètres du modèle sur une sortie de modèle choisie. Dans notre cas, une telle analyse évaluerait quel paramètre a le plus d'impact sur l'apparition de D-xylose, en particulier entre la vidange gastrique et l'absorption intestinale. Étant donné que nous avons considéré le D-xylose comme un biomarqueur pertinent pour l'apparition exogène de glucose, on s'attend à ce que le modèle soit plus sensible à l'absorption intestinale qu'à la vidange gastrique. En plus de l'analyse de sensibilité, nous avons estimé les paramètres sur un modèle inspiré par Dalla Man et ses collègues (Dalla Man et al., 2006a) caractérisé par une modélisation détaillée de la vidange gastrique et une modélisation intestinale simplifiée.

4.6.1 Analyse de sensibilité globale

Le taux d'apparition de glucose exogène (soit Ra_X pour le D-xylose ou Ra_G pour le glucose) dépend non seulement du taux d'absorption intestinale mais aussi du taux de vidange gastrique. Par conséquent, à la fois la vidange gastrique et les événements d'absorption intestinale contri-

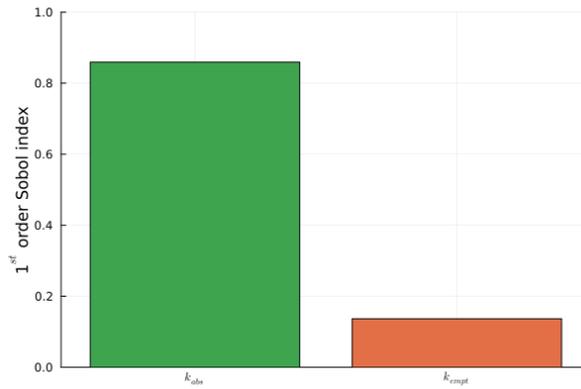


FIGURE 4.10 – Résultats de l’analyse de sensibilité globale de la moyenne de X_p au fil du temps.

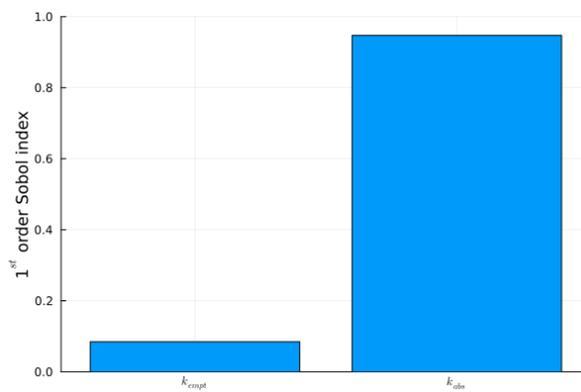
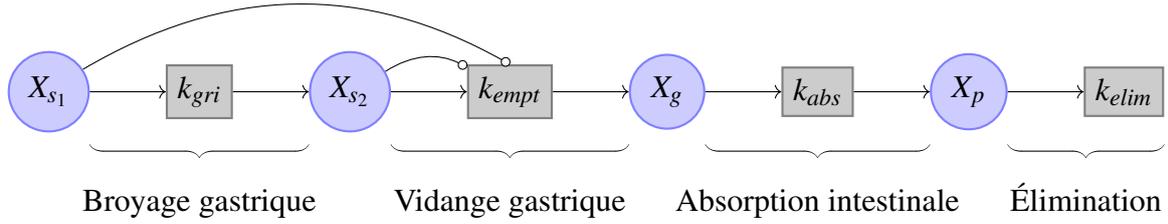


FIGURE 4.11 – Résultats de l’analyse de sensibilité globale de AUC_{RaX} à 180 minutes.

buent potentiellement à l’AIG. Comme nous cherchons un modèle capable d’évaluer l’activité intestinale pour profiler n’importe quel individu, il est important de vérifier quel facteur a le plus d’impact sur l’AIG. *L’analyse de sensibilité globale* permet de comprendre comment l’incertitude ou la variabilité des entrées d’un modèle affecte la sortie ou le résultat du modèle. Elle aide à identifier quels paramètres ont le plus d’impact significatif sur les résultats du modèle. Dans ce travail, l’analyse de sensibilité a été réalisée sur le modèle (sans la variant jejunal) en utilisant les indices de Sobol (Sobol, 1993). Pour la sortie du modèle, nous considérons l’aire sous la courbe du taux d’apparition de D-xylose à 180 minutes, notée AUC_{RaX} . Elle correspond à l’intégration de Ra_X , c’est-à-dire à la quantité totale de D-xylose qui a atteint le sang à un moment donné t , indépendamment de l’influence du taux d’élimination k_{elim} . En l’absence de méthodes de traçage (comme c’est le cas pour le D-xylose dans ce travail), le calcul de AUC_{RaX} est intéressant pour évaluer l’absorption du D-xylose car observer uniquement sa concentration dans le plasma (X_p) serait également influencé par le taux d’élimination. De plus, en vérifiant la sortie au temps maximal surveillé (180 min), nous voulions nous assurer que la vidange gastrique a beaucoup moins d’influence sur le taux d’apparition que l’absorption intestinale, garantissant ainsi que le D-xylose peut potentiellement être utilisé comme biomarqueur pour évaluer Ra_X (et éventuellement Ra_G).



(a) Réseau de réactions

$$\begin{aligned}
 \frac{dX_{s_1}(t)}{dt} &= -k_{gri} \cdot X_{s_1}(t) & X_{s_1}(0) &= \frac{D_X}{BW \cdot V_{D_X}} \\
 \frac{dX_{s_2}(t)}{dt} &= k_{gri} \cdot X_{s_1}(t) - k_{empt} \cdot X_{s_2}(t) & X_{s_2}(0) &= 0 \\
 \frac{dX_g(t)}{dt} &= k_{empt} \cdot X_{s_2}(t) - k_{abs} \cdot X_g(t) & X_g(0) &= 0 \\
 \frac{dX_p(t)}{dt} &= k_{abs} \cdot X_g(t) - k_{elim} \cdot X_p(t) & X_p(0) &= 0
 \end{aligned}$$

(b) Système d'EDO

(c) Conditions initiales

$$k_{empt} = k_{min} + \frac{k_{max} - k_{min}}{2} \left(2 - \tanh \left(\frac{5(X_{s_1}(t) + X_{s_2}(t) - b \cdot D_X)}{2 \cdot b \cdot D_X} \right) + \tanh \left(\frac{5(X_{s_1}(t) + X_{s_2}(t) - a \cdot D_X)}{2 \cdot (1 - a) \cdot D_X} \right) \right)$$

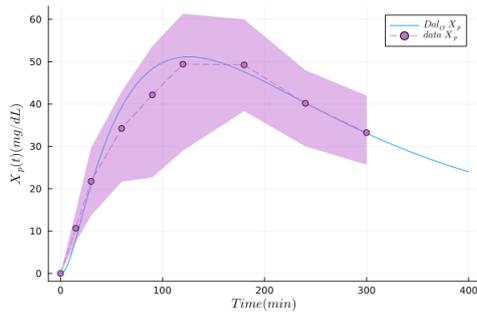
(d) Expression de k_{empt}

FIGURE 4.12 – Modèle avec vidange gastrique complexe.

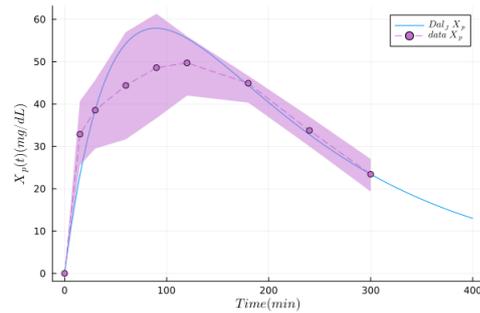
Nous utilisons le package Julia GlobalSensitivity (v2.1.4) pour effectuer cette analyse pour AUC_{RaX} . Cette analyse indique systématiquement l'importance de l'absorption intestinale, sans nier le rôle de la vidange gastrique, pour les deux paramètres. En effet, elle indique un indice de Sobol du premier ordre de 0,05 pour k_{empt} et un indice de Sobol du premier ordre de 0,95 pour k_{abs} .

4.6.2 Modèle de Vidange Gastrique « Complexe »

Pour valider davantage le degré d'implication de l'absorption intestinale sur la vidange gastrique concernant l'apparition du glucose ou du D-xylose dans le sang, nous avons comparé nos résultats avec un autre modèle comportant une partie de vidange gastrique plus complexe, inspiré des travaux de Dalla Man et de ses collègues (Dalla Man et al., 2006a). Ce modèle présente deux compartiments pour le contenu de l'estomac, le premier (X_{s_1}) représentant les aliments non broyés et le deuxième (X_{s_2}) représentant les aliments broyés (par opposition à un seul compartiment pour le modèle de la Figure 4.3) mais un seul compartiment (X_g) pour l'intestin (par opposition à plusieurs compartiments pour le modèle de la Figure 4.3). De plus, le taux de vidange gastrique k_{empt} de X_{s_2} vers X_g n'est pas une valeur constante mais dépend de la somme



(a) X_p correspondant à un bolus oral.



(b) X_p correspondant à un bolus jéjunal.

FIGURE 4.13 – Résultats des estimations de paramètres du modèle inspiré par Dalla Man et al. sur les ensembles de données de bolus *oral* et *jejunale* (NLL-Loss : 47,416). Les points et les lignes en pointillés représentent les valeurs expérimentales moyennes, l'enveloppe est l'écart type, et les lignes pleines sont les simulations produites avec le modèle de la Section 4.6.2.

des deux variables qui représentent le contenu total de l'estomac ($X_{s_1} + X_{s_2}$), du bolus initial (D_X) et d'autres paramètres constants (k_{min} , k_{max} , a et b). Intuitivement, ce taux est en forme de U et atteint sa valeur maximale (k_{max}) au début et à la fin du broyage (lorsque l'estomac est presque plein ou presque vide) et atteint sa valeur minimale (k_{min}) entre les deux.

Les valeurs de tous les paramètres constants ont été obtenues avec les mêmes données (ensemble de données de bolus *oral* et *jejunale*) et la même méthode de curve fitting que le modèle de la Section 3. Pour rappel, l'ensemble de données expérimentales présente la concentration de D-xylose au fil du temps, mesurée dans le sang, à la fois après un bolus oral et après un bolus injecté directement dans le jéjunum, et la paramétrisation est réalisée en utilisant les deux conditions expérimentales simultanément. L'idée est de vérifier si un modèle avec un estomac et une vidange gastrique plus complexes couplés avec un intestin plus simple modélisé comme un seul compartiment est capable de simuler cet ensemble de données aussi efficacement que le modèle de la Figure 4.3. Le résultat de cette expérience est donné dans la simulation de la Figure 4.13, montrant que la partie gastrique plus complexe du modèle n'est pas capable de générer une dynamique capable de correspondre aux données de façon aussi satisfaisante que le modèle de la Figure 4.3. Ainsi, combinée avec l'analyse de sensibilité sur le modèle à compartiments multiples, nous démontrons la nécessité d'utiliser le modèle de la Figure 4.3 pour refléter l'apparition de D-xylose.

4.7 Intérêt d'un modèle de Xylose

Dans ce chapitre, nous proposons un modèle à compartiments multiples de la dynamique postprandiale du D-xylose comme première étape vers un modèle prédictif de l'absorption intestinale du glucose. Ce modèle est basé sur trois paramètres majeurs représentant les taux (linéaires) de vidange gastrique, d'absorption intestinale et d'élimination, et modélise l'intestin comme une succession de compartiments, introduisant ainsi un retard qui représente le transit

intestinal. Nous avons calibré le modèle en utilisant un ensemble de données adapté provenant de plusieurs populations de miniporcs ayant subi une administration orale, intraveineuse ou jéjunale de D-xylose, ainsi qu'une résection intestinale. Nous avons étudié l'identifiabilité et la sensibilité de ses paramètres. Ce modèle présente de bonnes performances en termes de qualité d'estimation de paramètres, même avec les données d'injection jéjunale, surtout en comparaison avec un autre modèle où la partie gastrique est plus complexe mais la partie intestinale est simplifiée, et qui n'ajuste pas aussi bien les données d'injection jéjunale. Cela suggère que la modélisation à compartiments multiples choisie de l'intestin est pertinente, et souligne le rôle important de l'absorption intestinale. De plus, le modèle semble être identifiable pour tous les paramètres pertinents.

Enfin, nous avons également comparé le taux d'apparition de D-xylose prédit par le modèle avec le taux d'apparition réel de glucose exogène (Ra_G), c'est-à-dire le glucose provenant uniquement du repas et non du stockage rénal, par exemple. Ces résultats sont très intéressants car ils corroborent que le D-xylose pourrait être un marqueur précieux de l'absorption intestinale. Cela renforce le fait que notre modèle est un bon candidat pour prédire Ra_G , du moins qualitativement. Outre les investigations expérimentales, des travaux supplémentaires sont nécessaires pour améliorer, ou mieux tirer parti, de la capacité du modèle à prédire Ra_G . De plus, nous prévoyons de proposer un modèle simplifié du système de régulation glucose-insuline basé sur le modèle minimal de (Bergman et al., 1979) avec un modèle précis basé sur le D-xylose de l'AIG. Enfin, des ensembles de données sur des humains ayant subi des administrations de bolus de glucose et de D-xylose pourraient aider à traduire ce modèle chez l'homme. À long terme, nous espérons que ce modèle pourra être appliqué chez l'homme et pourra aider dans un contexte médical à diagnostiquer les patients présentant une absorption anormale du glucose intestinal. Le chapitre suivant, traitera, entre autres, d'un modèle issu des connaissances acquises tout au long de ce présent chapitre.

Chapitre 5

Conclusion et Perspectives

5.1 Conclusion

Le diabète de type 2, bien que diagnosticable et cliniquement bien caractérisé (hyperglycémie et hypoinsulinémie plasmatiques), demeure une maladie métabolique dont les causes, multiples, et les mécanismes métaboliques sous-jacents de son incidence restent à découvrir. Ainsi, nous proposons, dans cette thèse, au regard des données cliniques (OBEDIAB) et précliniques (issues d'expériences effectuées sur des porcs miniatures) dont nous disposons, d'aborder l'hypothèse de l'absorption intestinale du glucose par le biais de la modélisation mathématique pour expliquer l'importance de ce facteur dans la survenue de la maladie.

Le modèle post-prandial proposé par Dalla Man (Dalla Man et al., 2007) – un système d'ODE avec des équations algébriques – est reproductible à partir de l'article de référence, et ce, malgré l'absence de données de concentrations originales, car non-fournies par les auteurs : il est possible de reproduire les tracés de l'article à partir des jeux de paramètres donnés dans celui-ci. Pour rappel, ce modèle a été implémenté à partir des données obtenues grâce à la méthode de clampage avec un double-traceur (Cobelli et al., 1992). Bien que cette technique soit efficace, elle demeure peu courante et complexe à mettre en place pour le suivi du glucose exogène. On se retrouve ainsi avec un modèle qui, pour correspondre aux données gold-standard dont les auteurs ont pu disposer, présente à la fois des équations « physiologiques » (basées sur une démarche de modélisation des événements biologiques qui interviennent dans le modèle) et des équations « heuristiques » (c'est-à-dire abstraites, avec une démarche purement mathématique sans envisager aucune interprétation biologique associée à ces équations).

Bien qu'efficace sur le jeu de données sur lequel il se base, lorsque le modèle de Dalla Man est confronté aux données OBEDIAB, celui-ci montre ses limites par des courbes qui correspondent très mal aux données : les paramètres estimés se sont avérés très insatisfaisants, malgré diverses stratégies d'optimisation. Ce problème aurait plusieurs sources. La source majeure, corrélée à toutes les suivantes qu'on évoquera, réside dans le nombre de paramètres à estimer. En effet, quel que soit l'optimiseur utilisé, celui-ci montre toujours ses limites au delà d'un certain nombre de paramètres, surtout si ceux-ci font partie d'un modèle comprenant des équations

tions non-linéaires. De ce nombre élevé de paramètres découlerait aussi une non-identifiabilité suspectée de ces derniers. Une non-identifiabilité du modèle impliquerait que, même si une estimation de paramètres était réussie sur un jeu de données, il serait probable qu'il existe d'autres jeux de paramètres pouvant correspondre aux données de façon tout aussi satisfaisante, faussant ainsi toute interprétation biologique des valeurs de paramètres. Enfin, il est fort possible que le modèle soit intrinsèquement limité pour être réellement représentatif de l'ensemble des patients atteints de DT2, notamment ceux ayant subi une chirurgie gastro-intestinale, liée à la comorbidité de surcharge pondérale. En effet, au delà de la présence d'équations « heuristiques » (donc « non-physiologiques »), le modèle privilégie une modélisation de la vidange gastrique complexe au détriment d'une sursimplification de la modélisation de l'absorption intestinale de glucose.

À partir de nos conclusions sur le modèle de référence, nous avons décidé de résoudre les points limitant potentiels en :

1. proposant un nouveau modèle post-prandial représentatif de tous les profils gastro-intestinaux. La particularité de cet exercice de modélisation a été d'exploiter les données de D-xylose, un analogue du glucose, qui serait un bon marqueur du taux d'apparition du glucose intestinal (Ra), contournant ainsi le recours systématique à la technique double traceur. Pour ce nouveau marqueur potentiel, il a été judicieux de délaissier notre jeu de données cliniques OBEDIAB pour des données précliniques originales de cochons miniatures, laissant moins de contraintes sur les conditions expérimentales possibles.
2. contraignant la sélection de modèle à une identifiabilité pratique complète, au moins pour les paramètres critiques de vidange gastrique, d'absorption intestinale et d'élimination, grâce à une approche parcimonieuse en nombre de paramètres impliquées et en inférant les paramètres sur 2 jeux de données de D-xylose simultanément.

Ainsi, nous sommes partis de trois modèles candidats dans la première étape de sélection de modèle : un modèle « simple », un modèle « à compartiments multiples » et un modèle « à rétention ». Confrontés aux données d'administration jéjunale et orale, simultanément, pour l'estimation de leur paramètres respectifs, c'est le modèle à compartiments multiples qui s'est avéré le plus optimal pour correspondre aux données. De surcroît, ce modèle s'avère identifiable sur les 3 paramètres clés du modèle (vidange gastrique, absorption intestinale et élimination). En approfondissant les propriétés de ce modèle multicompartimenté uniquement entraîné sur des données de D-xylose, ce dernier s'avère intéressant pour refléter aussi le taux d'apparition exogène de glucose. En effet, lorsque l'on superpose les simulations du taux d'apparition de D-xylose aux données expérimentales des taux d'apparitions exogène de glucose Ra_G (exceptionnellement obtenues par double traceur, avant et après résection intestinale sur une petite cohorte de cochons miniatures), le modèle sélectionné semble refléter efficacement le profil d'absorption intestinale de glucose.

Tous ces résultats indiquent qu'en délaissant une modélisation complexe de la *vidange gastrique* (que nous avons modélisée par une simple fonction de décroissance exponentielle) au

profit d'une modélisation *intestinale* plus complexe, basée sur les travaux de Salinari (Salinari et al., 2011), tout en s'inscrivant dans une démarche parcimonieuse en nombre de paramètres et d'interprétabilité biologique dans notre modélisation, nous avons surmonté les problèmes rencontrés avec le modèle post-prandial de Dalla Man.

5.2 Perspectives

L'ensemble des travaux menés dans cette thèse pourrait nous amener à confronter notre nouveau modèle, à compartiments multiples, au jeu de données cliniques OBEDIAB ou à d'autres données cliniques de D-xylose postprandial. Cela permettrait d'établir un lien entre modèle postprandial de D-xylose et modèle postprandial de glucose-insuline. Ce lien pourrait mener à un modèle de glucose-insuline partageant les caractéristiques structurelles principales de notre modèle de D-xylose, à savoir suivre une introduction parcimonieuse en nombre de variables et de paramètres cinétiques. Cela pourrait se faire en partant du modèle minimal de Bergman (Bergman, 1989) comme base de cette modélisation. De plus, on pourrait conserver la notion de spatialisation de l'intestin, introduite dans le modèle de D-xylose, c'est-à-dire en incluant le sous-modèle discrétisé de Salinari. Une résultante de cette démarche de modélisation donnerait, par exemple, le graphe de la figure 5.1.

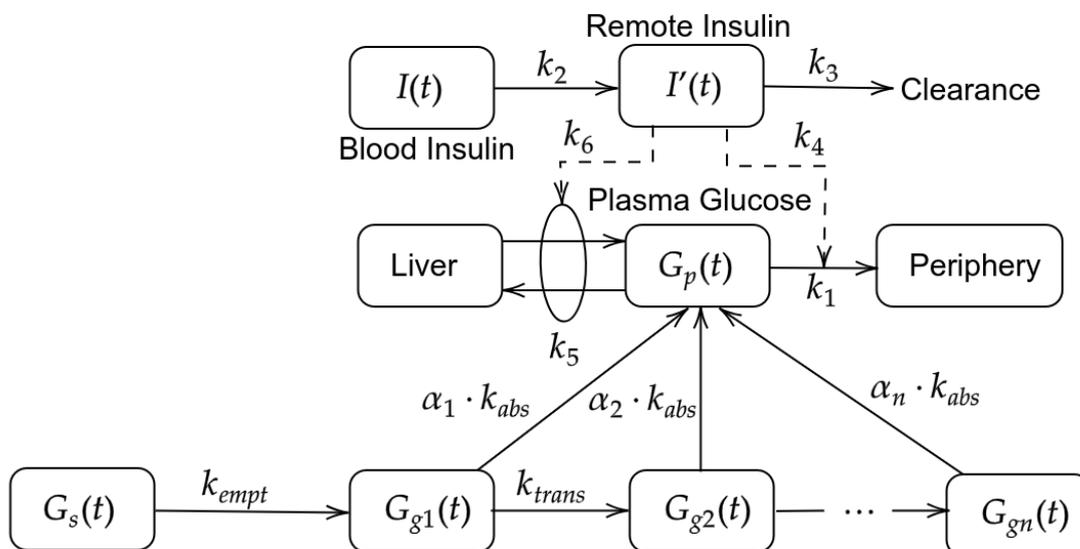


FIGURE 5.1 – Graphe d'un modèle post-prandial du glucose et de l'insuline construit à partir du modèle minimal de Bergman et le sous-modèle discrétisé de Salinari

Pour aller plus loin, il serait intéressant de trouver une méthode de mise en évidence de l'hétérogénéité des causes du DT2 à partir d'un modèle postprandial, en complément des travaux de mise en évidence statistique, déjà existants. Pour cela, plusieurs modèles pourraient être construits, en prenant en compte les spécificités déjà connues des différents types de DT2 :

insulino-résistance, métabolisme du tissu adipeux etc. Les données postprandiales de patients cliniques pourraient être utilisées pour effectuer des estimations de paramètres sur ces modèles. Ainsi, en constatant les modèles qui permettent de simuler le mieux les données d'entrée, nous pourrions effectuer une segmentation des patients en différents sous-types de DT2. Cela permettrait d'aider au diagnostic médical tout en améliorant les connaissances sur cette maladie. Malgré l'ampleur de la tâche qu'il reste à accomplir dans l'investigation du diabète de type 2 sous l'angle de la modélisation, l'ensemble de cette thèse nous mène à penser que le D-xylose serait un bon marqueur diagnostique du DT2, prenant en compte le taux d'apparition intestinal de sucre, si de futurs travaux de recherche vont dans ce sens. Même si de nombreux essais seront encore nécessaires pour que le D-xylose s'établisse comme un marqueur d'intérêt, l'existence d'un modèle simple comme le nôtre permet d'avoir une meilleure compréhension des cinétiques de vidange, d'absorption et d'élimination, ce que ne permettent pas toujours des modèles plus complexes. Il permet par exemple d'évaluer les rôles relatifs de ces trois mécanismes, là où une simple mesure de la concentration de D-xylose dans le sang ne permet évidemment pas de les distinguer. Notre modèle a ainsi été utilisé dans une publication en cours de soumission traitant de l'utilisation du D-xylose en routine clinique : le modèle a permis d'ajuster certains résultats en retirant la composante d'élimination qui perturbait les mesures, et ainsi obtenir une meilleure estimation de l'absorption.

Cette thèse permettra, nous l'espérons, de mettre en évidence l'utilité d'implémenter et d'utiliser des modèles mécanistiques dans la découverte et le raffinement de la compréhension de maladies complexes, comme ici, le diabète de type 2, en marge des analyses statistiques, davantage mises en avant dans les publications scientifiques, et ce, malgré des travaux encore nécessaires de mise en place de méthodologies claires et standardisées qui iraient dans ce sens. Cette thèse démontre par ailleurs que la modélisation mathématique n'exempte pas des efforts multidisciplinaires essentiels dans l'élucidation d'une maladie aussi complexe que le diabète de type 2. Nous espérons que ce travail ouvrira la voie à d'autres approches du même type à l'avenir.

Bibliographie

- Ahmad, E., Lim, S., Lamptey, R., Webb, D. R., and Davies, M. J. (2022). Type 2 diabetes. *The Lancet*, 400(10365) :1803–1820.
- Baud, G., Raverdy, V., Bonner, C., Daoudi, M., Caiazzo, R., and Pattou, F. (2016a). Sodium glucose transport modulation in type 2 diabetes and gastric bypass surgery. *Surgery for Obesity and Related Diseases*, 12(6) :1206–1212. Diabetes Special Issue.
- Baud, G., Raverdy, V., Bonner, C., Daoudi, M., Caiazzo, R., and Pattou, F. (2016b). Sodium glucose transport modulation in type 2 diabetes and gastric bypass surgery. *Surgery for Obesity and Related Diseases*, 12(6) :1206–1212.
- Bergman, M., Jagannathan, R., Buyschaert, M., Pareek, M., Olsen, M. H., Nilsson, P. M., Medina, J. L., Roth, J., Chetrit, A., Groop, L., et al. (2018). Lessons learned from the 1-hour post-load glucose level during ogtt : Current screening recommendations for dysglycaemia should be revised. *Diabetes/metabolism research and reviews*, 34(5) :e2992.
- Bergman, R. N. (1989). Toward physiological understanding of glucose tolerance : minimal-model approach. *Diabetes*, 38(12) :1512–1527.
- Bergman, R. N. (2021). Origins and history of the minimal model of glucose regulation. *Frontiers in endocrinology*, 11 :583016.
- Bergman, R. N., Ider, Y. Z., Bowden, C. R., and Cobelli, C. (1979). Quantitative estimation of insulin sensitivity. *American Journal of Physiology-Endocrinology and Metabolism*, 236(6) :E667.
- Bezanson, J., Edelman, A., Karpinski, S., and Shah, V. B. (2017). Julia : A fresh approach to numerical computing. *SIAM review*, 59(1) :65–98.
- Bjarnason, I. T. and Sherwood, R. A. (2014). *Clinical biochemistry of the gastrointestinal tract*, pages 214–230. Elsevier.
- Caiazzo, R. and Pattou, F. (2013). Adjustable gastric banding, sleeve gastrectomy or gastric bypass. can evidence-based medicine help us to choose? *Journal of Visceral Surgery*, 150(2) :85–95.
- Chaouiya, C. (2007). Petri net modelling of biological networks. *Briefings in bioinformatics*, 8(4) :210–219.
- Cobelli, C., Toffolo, G., and Foster, D. M. (1992). Tracer-to-tracee ratio for analysis of stable isotope tracer data : link with radioactive kinetic formalism. *American Journal of Physiology-Endocrinology And Metabolism*, 262(6) :E968–E975.
- Cox, M. E. and Edelman, D. (2009). Tests for screening and diagnosis of type 2 diabetes. *Clinical diabetes*, 27(4) :132–138.

- Dalla Man, C., Camilleri, M., and Cobelli, C. (2006a). A system model of oral glucose absorption : Validation on gold standard data. *IEEE Transactions on Biomedical Engineering*, 53(12) :2472–2478.
- Dalla Man, C., Rizza, R. A., and Cobelli, C. (2006b). Mixed meal simulation model of glucose-insulin system. In *2006 International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 307–310. IEEE.
- Dalla Man, C., Rizza, R. A., and Cobelli, C. (2007). Meal simulation model of the glucose-insulin system. *IEEE Transactions on biomedical engineering*, 54(10) :1740–1749.
- Das, S. L., Singh, P. P., Phillips, A. R., Murphy, R., Windsor, J. A., and Petrov, M. S. (2014). Newly diagnosed diabetes mellitus after acute pancreatitis : a systematic review and meta-analysis. *Gut*, 63(5) :818–831.
- David, R. and Alla, H. (1987). Continuous petri nets. In *8th European Workshop on Application and Theory of Petri nets*, pages 275–294.
- Dhawan, S. and Natarajan, R. (2019). Epigenetics and type 2 diabetes risk. *Current diabetes reports*, 19 :1–14.
- Dursoniah, D., Folschette, M., Goutchtat, R., Raverdy, V., Pattou, F., and Lhoussaine, C. (2024). Modeling intestinal glucose absorption from d-xylose data. In *15th International Conference on Bioinformatics, Models, Methods and Algorithms*.
- Dursoniah, D., Folschette, M., Lhoussaine, C., Goutchtat, R., Pattou, F., and Raverdy, V. (2021). Limits of a glucose-insulin model to investigate intestinal absorption in type 2 diabetes. In *BIOTECHNO 2021 : The Thirteenth International Conference on Bioinformatics, Biocomputational Systems and Biotechnologies*, pages 4–7. IARIA.
- Elashoff, J. D., Reedy, T. J., and Meyer, J. H. (1982). Analysis of gastric emptying data. *Gastroenterology*, 83(6) :1306–1312.
- Feinberg, M. (2019). *Foundations of chemical reaction network theory*. Springer.
- Fujita, Y., Kojima, H., Hidaka, H., Fujimiya, M., Kashiwagi, A., and Kikkawa, R. (1998). Increased intestinal glucose absorption and postprandial hyperglycaemia at the early step of glucose intolerance in otsuka long-evans tokushima fatty rats. *Diabetologia*, 41 :1459–1466.
- Garcia-Naveiro, R. and Udall Jr, J. N. (2011). Maldigestion and malabsorption. In *Pediatric Gastrointestinal and Liver Disease*, pages 337–349. Elsevier.
- Gonzalez, J. T. and King, A. J. (2022). For flux sake : isotopic tracer methods of monitoring human carbohydrate metabolism during exercise. *International journal of sport nutrition and exercise metabolism*, 33(1) :60–70.
- Goutchtat, R., Marciniak, C., Caiazzo, R., Verkindt, H., Quenon, A., Rabier, T., Lapiere, S., Raverdy, V., Hubert, T., and Pattou, F. (2022). 1361-p : D-xylose test as a biomarker for glucose intestinal absorption in humans and minipigs. *Diabetes*, 71(Supplement_1).
- Goyal, R. K., Cristofaro, V., and Sullivan, M. P. (2019). Rapid gastric emptying in diabetes mellitus : Pathophysiology and clinical importance. *Journal of Diabetes and its Complications*, 33(11) :107414.
- Haji, S. H. and Abdulazeez, A. M. (2021). Comparison of optimization techniques based on gradient descent algorithm : A review. *PalArch's Journal of Archaeology of Egypt/Egyptology*, 18(4) :2715–2743.

- Hansen, N. (2023). The CMA Evolution Strategy : A Tutorial. Préprint.
- Hansen, N. and Ostermeier, A. (1996). Adapting arbitrary normal mutation distributions in evolution strategies : the covariance matrix adaptation. In *Proceedings of IEEE International Conference on Evolutionary Computation*, pages 312–317.
- Hansen, N. and Ostermeier, A. (2001). Completely derandomized self-adaptation in evolution strategies. *Evolutionary computation*, 9(2) :159–195.
- Kumar, R. (2018). Hepatogenous diabetes : an underestimated problem of liver cirrhosis. *Indian journal of endocrinology and metabolism*, 22(4) :552–559.
- Levenberg, K. (1944). A method for the solution of certain non-linear problems in least squares. *Quarterly of applied mathematics*, 2(2) :164–168.
- Mari, A., Tura, A., Grespan, E., and Bizzotto, R. (2020). Mathematical Modeling for the Physiological and Clinical Investigation of Glucose Homeostasis and Diabetes. *Frontiers in Physiology*, 11.
- Matsuda, M. and DeFronzo, R. A. (1999). A method for quantitation of insulin sensitivity and insulin resistance in vivo. *American Journal of Physiology-Endocrinology and Metabolism*, 276(6) :E1023–E1030.
- Matthews, D. R., Hosker, J. P., Rudenski, A. S., Naylor, B., Treacher, D. F., and Turner, R. C. (1985). Homeostasis model assessment : insulin resistance and β -cell function from fasting plasma glucose and insulin concentrations in man. *diabetologia*, 28 :412–419.
- Mockus, J. (2005). The bayesian approach to global optimization. In *System Modeling and Optimization : Proceedings of the 10th IFIP Conference New York City, USA, August 31–September 4, 1981*, pages 473–481. Springer.
- Muniyappa, R., Madan, R., and Varghese, R. T. (2015). Assessing insulin sensitivity and resistance in humans. In Feingold, K. R. et al., editors, *Endotext*. MDText.com, Inc.
- Patarrão, R. S., Lutt, W. W., and Macedo, M. P. (2014). Assessment of methods and indexes of insulin sensitivity. *Revista Portuguesa de Endocrinologia, Diabetes e Metabolismo*, 9(1) :65–73.
- Petri, C. A. (1962). *Kommunikation mit Automaten*. PhD thesis, Universitat Bonn.
- Prasad, R. B. and Groop, L. (2015). Genetics of type 2 diabetes—pitfalls and possibilities. *Genes*, 6(1) :87–123.
- Raue, A., Kreutz, C., Maiwald, T., Bachmann, J., Schilling, M., Klingmüller, U., and Timmer, J. (2009). Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. *Bioinformatics*, 25(15) :1923–1929.
- Roesch, E., Greener, J. G., MacLean, A. L., Nassar, H., Rackauckas, C., Holy, T. E., and Stumpf, M. P. (2023). Julia for biologists. *Nature methods*, 20(5) :655–664.
- Salinari, S., Bertuzzi, A., and Mingrone, G. (2011). Intestinal transit of a glucose bolus and incretin kinetics : a mathematical model with application to the oral glucose tolerance test. *American Journal of Physiology-Endocrinology and Metabolism*, 300(6) :E955–E965.
- Saltelli, A., Tarantola, S., Campolongo, F., Ratto, M., et al. (2004). *Sensitivity analysis in practice : a guide to assessing scientific models*, volume 1. Wiley Online Library.
- Soboń, I. (1993). Sensitivity estimates for nonlinear mathematical models. *Math. Model. Comput. Exp.*, 1 :407.

- Toffolo, G., Basu, R., Dalla Man, C., Rizza, R., and Cobelli, C. (2006). Assessment of postprandial glucose metabolism : conventional dual-vs. triple-tracer method. *American Journal of Physiology-Endocrinology And Metabolism*, 291(4) :E800–E806.
- Tricò, D., Mengozzi, A., Frascerra, S., Scozzaro, M. T., Mari, A., and Natali, A. (2019). Intestinal glucose absorption is a key determinant of 1-hour postload plasma glucose levels in nondiabetic subjects. *The Journal of Clinical Endocrinology & Metabolism*, 104(6) :2131–2139.
- Vikhar, P. A. (2016). Evolutionary algorithms : A critical review and its future prospects. In *2016 International conference on global trends in signal processing, information computing and communication (ICGTSPICC)*, pages 261–265. IEEE.
- Wallace, T. M., Levy, J. C., and Matthews, D. R. (2004). Use and abuse of homa modeling. *Diabetes Care*, 27(6) :1487–1495.