TABLE DES MATIÈRES

Page

INTRODUCTION 1

CHAPITRE 1 CONTEXT CLINIQUE…….....5

1.1 Contexte clinique et prise en charge initiale du grand brûlé.....5
1.2 Évaluation de l'étendue des brûlures 7
1.2.1 Stabilisation et réanimation liquidienne..... 7
1.2.2 Méthodes cliniques 8
1.2.3 Erreurs liées à l'estimation de la surface corporelle 11

CHAPITRE 2 REVUE DE LITTÉRATURE……15
2.1 Modélisation des surfaces corporelles 15
2.1.1 Modélisation 2D/ 3D des surfaces corporelles brûlées..... 15
2.1.2 Modélisation 3D de la surface corporelle du corps humain 19
2.1.3 Un outil pertinent pour la modélisation 3D des surfaces corporelles : MakeHuman..... 22
2.2 Anthropométrie 25
2.3 Modélisation statistique en science des données 27
2.3.1 Deux cultures : les statistiques et l'apprentissage machine 27
2.3.2 Modèles et hypothèses de la régression linéaire multiple en statistique... 27
2.3.2.1 Régression linéaire multiple et ses hypothèses..... 29
2.3.2.2 Modèle linéaire général..... 33
2.3.2.3 Modèle linéaire généralisé 34
2.3.2.4 Modèle non linéaire 36
2.3.3 Problématique de l'étude et importance de la sélection de variables 36
2.3.3.1 Multicolinéarité et son effet sur les modèles linéaires généralisés..... 38
2.3.3.2 Surapprentissage en apprentissage machine 39
2.3.4 Prétraitements de données..... 40
2.3.4.1 Corrélation et allure de la distribution des variables 40
2.3.4.2 Centrage et normalisation 41
2.3.5 Les principales approches de sélection de variables en régression..... 41
2.3.6 Comparaison entre modèles linéaires et modèles d'apprentissage machine 45
2.3.7 Sélection de sous-ensembles de modèles linéaires 47
2.3.7.1 Sélection de sous-ensembles..... 47
2.3.7.2 Méthodes de réduction..... 48
2.3.8 L'apprentissage machine : application à la régression 50
2.3.8.1 Sélection de modèles et généralisation 51
2.3.8.2 Algorithmes paramétriques et non paramétriques 55

2.3.9 Algorithmes de sélection de variables non linéaires 57
2.3.9.1 Arbre décisionnel de régression – CART 58
2.3.9.2 Forêt d'arbres décisionnels 60
2.3.9.3 Boosting d'arbre de régression 62
2.3.9.4 Régression à vecteur de support 64
2.3.9.5 Réseau de Neurones 67
2.3.9.6 Moyenne de Réseau de Neurones 70
2.3.9.7 Régression multivariée par splines adaptatives 71
2.4 Création d'un échantillon stratifié 73
2.5 Métriques de validation 74
2.5.1 Métriques de validation des modèles prédictifs 74
2.5.2 Métriques de validation 3D 76
2.5.3 Validation de méthodes 79

CHAPITRE 3 HYPOTHÈSES ET OBJECTIFS DE RECHERCHE 83

CHAPITRE 4 PUBLISHED ARTICLE: TOWARD A QUANTIFIED ASSESSMENT OF TOTAL BODY SURFACE AREA FROM ANTHROPOMETRIC MEASUREMENTS FOR PATIENT WITH BURN INJURIES 87
4.1 Mise en contexte 87
4.2 Abstract 88
4.3 Introduction 88
4.4 Materials and methods 89
4.4.1 Generation of realistic three-dimensional models 89
4.4.2 Anthropometric measuring process 90
4.4.3 Multiple regression 91
4.5 Results 92
4.6 Discussion 93
4.7 Conclusion 94

CHAPITRE 5 PUBLISHED ARTICLE: THE IMPORTANCE OF A 3D-BASED APPROACHE WITH PERSONALIZED MODELS FOR ACCURATELY ASSESSING TBSA 95
5.1 Mise en contexte 95
5.2 Abstract 96
5.3 Introduction 97
5.4 Methods 100
5.4.1 Burn representation and calculation of the BSA, burned BSA, and TBSA 100
5.4.2 Importance of personalizing 3D models 101
5.4.3 Preliminary validation of the accuracy of the 3D models built with MakeHuman Software 107

5.4.4 Preliminary validation of the proposed approach 108
5.5 Results 109
5.5.1 Burn representation and calculation of the TBSA 109
5.5.2 Importance of personalized 3D models 110
5.5.3 Preliminary validation of the accuracy of the 3D models built with MakeHuman software 111
5.5.4 Preliminary validation of the proposed approach 111
5.6 Discussions and conclusion 113

CHAPITRE 6 PREDICTIVE MODELLING FOR PERSONALIZED THREE-DIMENSIONAL BURN INJURIES ASSESSMENTS 117
6.1 Mise en contexte 117
6.2 Abstract 118
6.3 Introduction 119
6.4 Methods 121
6.4.1 Reliability of anthropometric measurements in clinical settings 121
6.4.2 Generation of a stratified random sample 123
6.4.3 Selection of best feature subsets 126
6.4.4 Body Surface Area accuracy of 3D models 129
6.4.5 Morphology accuracy of 3D models 130
6.5 Results 132
6.5.1 Reliability of anthropometric measurements in clinical settings 132
6.5.2 Generation of a stratified random sample 134
6.5.3 Selection of best feature subsets 134
6.5.4 Body Surface Area accuracy of 3D models 138
6.5.5 Morphology accuracy of 3D models 139
6.6 Discussion 142
6.7 Conclusion 146

CHAPITRE 7 DISCUSSION GÉNÉRALE 147
7.1 Synthèse des travaux de recherche 147
7.2 Limites et recommandations 152
7.3 Retombés et impacts 153

CONCLUSION 157

ANNEXE I PROTOCOLE DE PRISE DE MESURES ANTHROPOMÉTRIQUES 161

ANNEXE II PROTOCOLE CLINIQUE ET UTILISATION DE L'APPLICATION SUR TABLETTE…….....167

ANNEXE III AUTRES CONTEXTES CLINIQUES NÉCESSITANT UNE ESTIMATION PRÉCISE DE LA BSA…….……..…178

ANNEXE IV SCANNER PORTABLE : PROBLÉMATIQUE POUR LES GRANDS BRÛLÉS……..……181

APPENDICES……...188

BIBLIOGRAPHIE…….190

LISTE DES TABLEAUX

Page

Tableau 1.1 Tableau synthèse des différentes méthodes cliniques....11

Tableau 2.1 Présentation des méthodes numériques du schéma de Lund & Browder .16

Tableau 2.2 Synthèse des différentes méthodes de modélisation 3D des brûlures....17

Tableau 2.3 Synthèse des systèmes d'acquisition Kinect pour la modélisation 3D des brûlures 18

Tableau 2.4 Approche de création et de reconstruction....19

Tableau 2.5 Approche d'interpolation et de modèle....21

Tableau 2.6 Présentation des différentes distributions et fonction de liaison pour des données continues positives....35

Tableau 2.7 Règle des pouces de Mukaka, (2012), Table 1, p3 41

Tableau 2.8 Avantages (+) et inconvénients (-) des différentes méthodes de sélection de variables 45

Tableau 2.9 Avantages (+) et inconvénients (-) des GLMs et apprentissage machine..46

Tableau 2.10 Critère d'acceptabilité d'une mesure répétable et reproductible 80

Tableau 4.1 Limited set of measures for each regression model....92

Tableau 4.2 Results of each regression model....93

Tableau 5.1 Summary of the paired models 105

Tableau 5.2 Mean and standard deviation of the difference of BBSA, TBSA and Parkland fluids between each paired models....111

Tableau 5.3 Summary of the BSA and error percentage of the scans and MH models....111

Tableau 5.4 Accepted standard for the surface areas of the patches and estimated fluids for each mannequin....112

Tableau 5.5 Summary of the Percent Total Body Surface Area (TBSA) assessment with the Lund & Browder chart....112

Tableau 5.6 Summary of the Percent Total Body Surface Area (TBSA) assessment with the three-dimensional (3D) rendering interface....113

Tableau 6.1 Characteristics of study participants....122

Tableau 6.2 Reliability and correlation coefficient of anthropometric measurements....133

Tableau 6.3 Results of best features subsets selection ─ data modelling....135

Tableau 6.4 Results of best features subsets selection ─ Algorithmic modelling....136

Tableau 6.5 Comparison of predictive modelling performance for algorithmic modelling including 3 measurements....136

Tableau 6.6 Comparison of predictive modelling performance for algorithmic modelling including 4 measurements....137

Tableau 6.7 Surface difference with gold standard Scan....139

Tableau 6.8 Two sided surface distances between 3D scans and 3D models built from anthropometric measurements (avNN_4)....140

Tableau 7.1 Contexte clinique d'utilisation de la BSA et leurs problématiques associées....179

LISTE DES FIGURES

Page

Figure 1.1 Résumé du parcours de soins du patient brûlé 5

Figure 1.2 Exemple réel du schéma de Lund & Browder tel qu'utilisé à l'hôpital St-Justine de Montréal. Schéma pour adulte à gauche et enfant à droite 9

Figure 2.1 Problématique des logiciels de modélisation 3D actuels 16

Figure 2.2 Aperçu des modèles de MH Modèles créés à partir du logiciel MH http://www.MH.org/ 23

Figure 2.3 Vue d'ensemble des différents modèles de régression en statistique. Le symbole « $^{+}$ » indique l'extension du modèle par rapport au sous-ensemble. Le symbole « $\in$ » indique l'inclusion du modèle dans le sur ensemble 28

Figure 2.4 a) Structure attendue des résidus. b) Structure non désirée des résidus (biais) 31

Figure 2.5 Schéma de fonctionnement des méthodes *Filter* 43

Figure 2.6 Schéma de fonctionnement des méthodes *Wrapper* 44

Figure 2.7 Schéma de fonctionnement des méthodes *Embedded* 45

Figure 2.8 Compromis biais-variance pour un modèle optimal en apprentissage machine 54

Figure 2.9 Schéma de fonctionnement d'un arbre décisionnel 59

Figure 2.10 Schéma de fonctionnement de l'algorithme Forêt d'arbres de régression. Les histogrammes représentent l'importance de chaque variable 61

Figure 2.11 Schéma de fonctionnement du Boosting d'arbres de régression 63

Figure 2.12 Schéma visuel des différences entre les méthodes Clustering et SVM pour la classification et régression 64

Figure 2.13 Structure du SVR à gauche. Fonction de perte associée à droite. Seuls les points en dehors de l'intervalle ε contribuent au coût final. 66

clicours.com

Figure 2.14 Présentation schématique du réseau de Neurones (ou perceptron multi-couche) 68

Figure 2.15 Réseau de Neurones à propagation avant (*feed-forward network*). Adapté de (Chrislb, 2005) 70

Figure 4.1 Examples of some 3D models randomly generated for this study....90

Figure 4.2 Anthropometric measurements used in this study 91

Figure 5.1 Example of the mesh model. STi is the surface area of 1 triangle and v1, v2, v3 its the adjacent vertices. 100

Figure 5.2 Examples of some three-dimensional models created with MakeHuman™ software....102

Figure 5.3 Difference between existing methods for 3D assessment of the TBSA and the proposed method. a.1 and b.1, Example of the base model. a.2, expansion algorithm of the model proposed by BurnCase3D (Giretzlehner et al., 2004) and BAI (Prieto et al., 2011). b.2, Local expansion algorithm of MakeHuman software (Bastioni et al., 2008) suggested in the proposed method. The represented models were built with the MakeHuman software and serve only as an example....104

Figure 5.4 Example of paired models. The two models have the same BSA but a different morphology 105

Figure 5.5 Burn patterns used in this study. In red the Burned Body Surface Area (BBSA) 106

Figure 5.6 On the left, the 4 commercial mannequins used with patches represented simulated burns. On the right their respective 3D scans....108

Figure 5.7 Example of the 3D rendering interface....110

Figure 6.1 Construction method of the stratified random sample....124

Figure 6.2 Obesity prevalence according to age and gender. Distribution age from 1 to 14 years (men=4116, women=3909) and from 80 to 90 years (men=793, women=1263) were not represented for graph since 3D models only had one BMI class: normal weight 125

Figure 6.3 Importance of each anthropometric measurement for each algorithm. A score close to 100 (in red) indicates an important measure. A score close to 0 (in green) indicates a non-important measurement....137

Figure 6.4 Color maps of subjects with maximum error distance (A, A', B, B') and lowest error distance (C, C', D, D'). Distance errors in 3D scans are represented on the left (A, B, C, D), while errors in their respective 3D models (built from 4 anthropometric measurements as determined by avNN_4) are represented on the right (A', B', C', D'). The blue color represents the maximum distance and the red color the minimum distance 141

Figure 7.1 Première étape de l'utilisation de l'application, remplir le formulaire du patient....170

Figure 7.2 Choix de morphologie. Exemple d'un enfant en haut et d'une femme en bas 172

Figure 7.3 Entrée des mesures à droite. Exemple avec l'ensemble taille, circonférence poitrine et hanches....173

Figure 7.4 Peinturage des brûlures sur le modèle 3D. La TBSA est calculée en conséquence. Il est possible de régler la taille du curseur pour plus de précision (figure du haut)....174

Figure 7.5 Menu proposé par l'application tablette 176

Figure 7.6 Outils proposés par l'application tablette 177

Figure 7.7 Étude préliminaire avec le Go!Scan, vue des trois scans à recaler avec leurs armatures....181

Figure 7.8 Étude préliminaire avec le Go!Scan, armature à gauche et modèle générique à droite....182

Figure 7.9 Étude préliminaire avec le Go!Scan, résultats premier essai....183

Figure 7.10 Étude préliminaire avec le Go!Scan, problématique du faible recouvrement de surface et manque de données important 184

Figure 7.11 Étude préliminaire avec le Go!Scan, vue des trois scans avec modèle générique....184

Figure 7.12 Étude préliminaire avec le Go!Scan, résultats deuxième essai à gauche. Résultats mannequins au centre et à droite....185

Figure 7.13 Étude préliminaire avec le Go!Scan, résultats d'un scan d'un sujet sain debout sans contraintes de pose 186

Figure 7.14 Exemple d'une table de chirurgie d'hôpital, Montréal 187

ClicOurs.com

Figure 7.15 Carte heuristique des différents algorithmes en apprentissage machine. Tiré de Brownlee, (s.d.). "Machine Learning Mastery" avec l'autorisation de l'auteur. En ligne <https://machinelearningmastery.com>. Consulté le 16 Novembre 2017. ✓ : algorithmes utilisés. ✗ : non applicable au projet de recherche 188

LISTE DES ABRÉVIATIONS, SIGLES ET ACRONYMES

2D	Deux dimensions
3D	Tridimensionnel
%TBSA	Pourcentage de superficie corporelle brûlée (Percent Total Body Surface Area)
AIC	Critère d'information d'Akaike (Akaike Information Criterion)
ANN	Réseau de Neurone Artificiel (Artificial Neural Network)
ASD	Distance de surface symétrique moyenne (Average Surface Distance)
AVD	Différence de volume absolue (Average Volume difference)
avNNet	Moyenne de réseau de neurones (average Neural Network)
BIC	Critère d'information Bayesien (Bayesian Information Criterion)
BSA	Surface corporelle (Body Surface Area)
BSD	Berkeley Software Distribution License
CART	Arbre décisionnel de Classification et de Régression (Classification And Regression Tree)
CHUM	Centre Hospitalier Universitaire de Montréal
CRCHUM	Centre de Recherche du Centre Hospitalier Universitaire de Montréal
ES0X	Exigence Scientifique
ÉTS	École de Technologie Supérieure
Gage R&R	Gage de Répétabilité & Reproductibilité
IMC	Indice de Masse Corporelle
LIO	Laboratoire de recherche en imagerie et orthopédie
LGPL	Licence publique générale limitée GNU

MARS	Régression par splines adaptatives multivariées (Multivariate Adaptative Regression Spline)
MAE	Erreur moyenne absolue (Mean Absolute Error)
MAPE	Pourcentage d'erreur absolue moyen (Mean Absolute Percentage Error)
MH	MakeHuman
MLP	Perceptron multicouches (Multilayer perceptron)
MSE	Erreur quadratique moyenne (Mean Square Error)
MSD	Distance de surface symétrique maximum (Mean Square Distance)
ndc	Nombre de catégories distinctes
NN	Réseau de Neurones (Neural Network)
RMSE	Racine de l'erreur quadratique moyenne (Root Mean Square Error)
SVM	Machine à support de vecteurs (Support Vector Machine)
SVR	Machine à support de regression (Support Vector Regression)
TOF	Capteur Temps de vol (Time of Flight)
VIF	Facteur d'Inflation de la Variance (Variance Inflation Factor)
VIH	Virus de l'Immunodéficience Humaine

LISTE DES SYMBOLES ET UNITÉS DE MESURE

UNITÉS DE BASES

Aire

Cm^2	Centimètre carré
M^2	Mètre carré

Distance

Cm	Centimètre

Masse

Kg	Kilogramme

Volume

mL	Millilitre

INTRODUCTION

Les brûlures sont un cas exceptionnel d'agression du corps humain qui entraîne de nombreuses altérations physiologiques et métaboliques. Leur potentiel dévastateur est aussi bien physique que psychologique et n'épargne aucune population (Durrant, Simpson, & Williams, 2008; Peck, 2011). Selon Peck, (2011) des brûlures suffisamment sévères pour nécessiter des soins médicaux touchent près de 11 millions de personnes chaque année dans le monde. L'incidence des brûlures est plus élevée que la tuberculose et les infections du VIH combinées. Les brûlures comptent pour plus de 300 000 morts chaque année dans le monde. Au Québec, on dénombre près de 300 patients dans les centres spécialisés pour grands brûlés adultes chaque année. À l'Hôpital Enfant-Jésus de Québec en 2015, c'est plus de 129 patients qui sont admis à l'unité des grands brûlés avec une hausse de 48% ces cinq dernières années (Centre d'expertise pour les victimes de brûlures graves de l'Est du Québec, 2015).

Dans la prise en charge initiale d'un grand brûlé, l'étendue des brûlures de deuxième et troisième degré est évaluée et constitue un facteur critique et déterminant pour évaluer le pronostic vital et fonctionnel du patient brûlé. La détermination de la superficie des surfaces brûlées en pourcentage est également critique pour la prise en charge médicale de ces patients et, notamment, pour le calcul des solutés de réanimation liquidienne dont ils seront dépendants dans les premiers jours post-traumatiques pour assurer leur survie. Les brûlures impliquent une perte de fluide dramatique qui doit être remplacée pour maintenir l'homéostasie du corps et assurer la survie du patient. La quantité de fluide dont le patient a besoin est fonction du pourcentage ou ratio de la superficie corporelle brûlée (nommée %TBSA). Le %TBSA est fonction de la superficie corporelle brûlée (nommée BBSA) sur la superficie corporelle totale (nommée BSA). Pour estimer ce pourcentage, les cliniciens utilisent plusieurs méthodes qui sont des schémas en deux dimensions avec des proportions corporelles fixes. Ces méthodes sous ou surestiment les calculs du fait de la transposition de brûlures représenter en 3 dimensions sur des schémas en deux dimensions, des variations morphologiques des patients et de la variabilité entre les observateurs (Wachtel, Berry, Wachtel, & Frank, 2000; Neuwalder, Sampson, Breuing, & Orgill, 2002; Giretzlehner et al., 2013; D. Parvizi et al., 2014). C'est

encore plus le cas pour des situations cliniques fréquentes où les patients présentent une obésité tronculaire (Kamolz, Parvizi, Giretzlehner, & Lumenta, 2014) ou chez les enfants puisque leurs proportions corporelles varient en fonction de l'âge (Livingston & Lee, 2000; Fairbrother, 2015). Des erreurs dans le calcul de la superficie brûlée peuvent entraîner des complications médicales et chirurgicales significatives (Livingston & Lee, 2000; Daryousch Parvizi et al., 2014) et dans certains cas, la mort du patient.

Il apparaît ainsi que l'estimation de la taille des brûlures effectuée par les méthodes cliniques sera toujours approximative. De sorte que l'on peut se demander comment améliorer la précision de l'estimation du %TBSA? Quel outil clinique développer afin de répondre aux besoins et au contexte clinique des grands brûlés dans l'évaluation initiale du patient ? Cette évaluation initiale demeure critique puisqu'elle permet non seulement d'estimer les besoins volémiques, mais également de déterminer les critères de transfert dans une unité et de statuer sur le pronostic vital du patient. L'amélioration de la précision du %TBSA permettra aux cliniciens d'améliorer cette évaluation initiale critique.

Intitulé « Modélisation et quantification 3D de la taille des superficies corporelles brûlées », cette thèse tend ainsi à démontrer l'importance d'améliorer la précision du %TBSA par la personnalisation 3D de la BSA et de la morphologie des patients. À travers la conception et validation de méthodes numériques permettant l'évaluation quantifiée de la taille des superficies corporelles brûlées à partir de modélisation 3D du corps entier obtenue à partir d'un nombre limité de mesures anthropométriques; c'est l'évaluation initiale des grands brûlés qui sera améliorée. Ce projet de recherche permettra de développer un outil calculant le pourcentage de superficie brûlée de façon précise et reproductible. Cela va contrer d'une part aux erreurs de calcul des cliniciens, spécialement lorsque les patients présentent des morphologies particulières et d'autre part, à la variabilité entre les observateurs. Les cliniciens auront un outil pour les aider dans la planification des traitements des grands brûlés et pour diminuer les morbidités associées. Les impacts visés sont la diminution des complications médicales à la suite d'une évaluation sous-optimale, raccourcir le séjour hospitalier des patients et optimiser leurs résultats cliniques.

Le projet de recherche a été réalisé en collaboration avec l'équipe du laboratoire de recherche en imagerie et orthopédie (LIO) de l'École de Technologie Supérieur (ÉTS) et du Centre de recherche du centre hospitalier universitaire de Montréal (CR-CHUM), le Centre des Brûlés du CHUM et le Centre hospitalier universitaire Sainte-Justine. C'est au sein d'un environnement clinique, en situation d'urgence, que s'inscrit ce projet de recherche. Lors de l'évaluation initiale, les patients sont soit inconscients, car intubés ou bien ambulants.

Le CHAPITRE 1 introduit le contexte clinique du projet. L'importance de l'évaluation du %TBSA pour la prise en charge initiale du grand brûlé ainsi que les problématiques associées aux méthodes cliniques y sont présentées. Ces notions fondamentales vont permettre de bien comprendre les besoins cliniques qui ont amené la problématique de recherche ainsi que la revue de littérature au CHAPITRE 2. Y est d'abord exposée une revue des différentes approches pour la modélisation des surfaces corporelles ainsi que leurs avantages et leurs limites dans le cadre de ce projet. Y est également présentée une revue indispensable sur l'anthropométrie et son lien avec la BSA. Ensuite vient la modélisation statistique en science des données où est comparée la modélisation de données (statistiques) et la modélisation algorithmique (apprentissage machine). Suivent quelques notions nécessaires sur la création d'un échantillon stratifié. Enfin, plusieurs métriques de validation sont formulées. Le CHAPITRE 3 énumère les hypothèses et objectifs spécifiques du projet. Chacun des objectifs spécifiques du projet a fait l'objet d'une publication dans une revue, d'une soumission, ou d'une présentation à un congrès et constitue les CHAPITRE 4, CHAPITRE 5 et CHAPITRE 6. Enfin, le CHAPITRE 7 présente une synthèse générale des travaux réalisés lors de ce projet doctoral. Des recommandations sont ensuite formulées pour la suite des travaux.

Clicours.com

CHAPITRE 1

CONTEXTE CLINIQUE

1.1 Contexte clinique et prise en charge initiale du grand brûlé

La prise en charge initiale d'un patient gravement brûlé est critique, spécifiquement dans les premiers 24 à 48 heures où le patient doit être évalué et stabilisé en fonction de cette évaluation. Il n'est pas rare que le patient soit pris en charge par des équipes de soins successives, notamment lorsque la brûlure survient dans une région éloignée des grands centres de trauma tertiaires et des unités de brûlés spécialisées. La Figure 1.1 ci-dessous présente le parcours de soins du patient brûlé dont la prise en charge passe par plusieurs étapes depuis l'accident jusqu'à la réhabilitation.

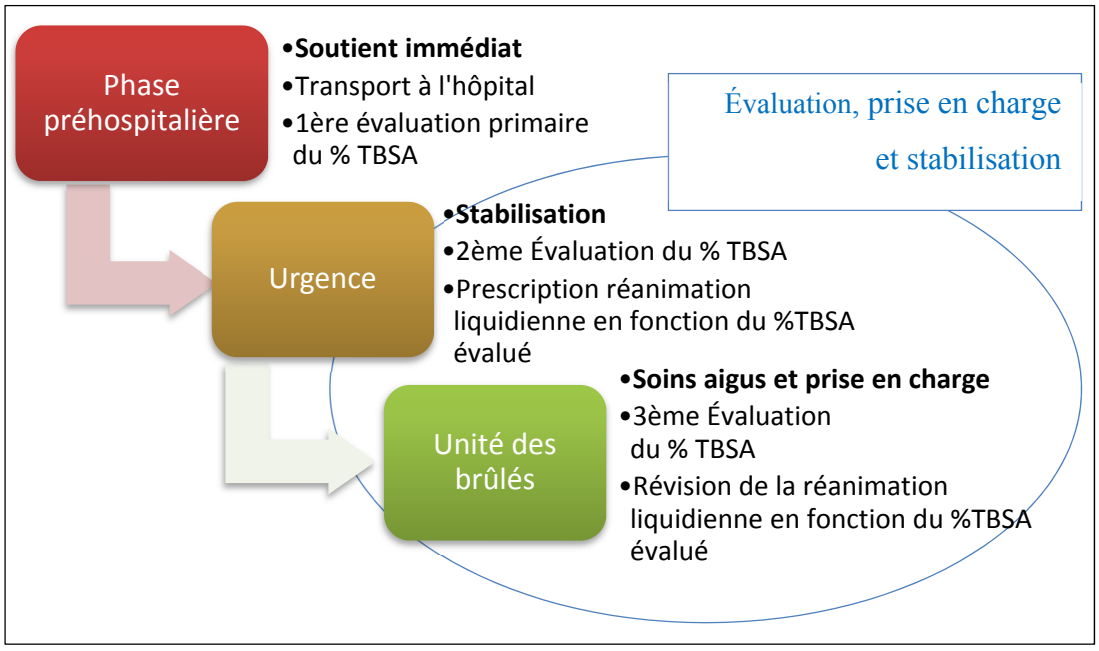


Figure 1.1 Résumé du parcours de soins du patient brûlé

Rappelons que chaque hôpital suit ses propres lignes directrices concernant la prise en charge d'un patient brûlé. Par conséquent, seul un résumé des lignes générales du parcours de soins du patient brûlé est présenté ci-après sur la base des références suivantes : Hettiaratchy et

Clicours.com

Dziewulski, (2004) et Durrant et al., (2008) du Royaume-Uni, ainsi que le Centre d'expertise pour les victimes de brûlures graves de l'Est du Québec, (2015). Il se peut donc que des variantes existent.

En premier lieu vient la prise en charge du patient par les premiers répondants et paramédics. Une première évaluation primaire du pourcentage des superficies brûlées (%TBSA) est effectuée afin de trier les patients. Le patient brûlé est orienté vers un centre de traumatologie pour toute brûlure supérieure ou égale à 10 %.

En deuxième lieu vient la stabilisation du patient par l'urgentologue et l'équipe de réanimation traumatologique en salle de réanimation de l'urgence. Cette équipe interdisciplinaire s'assure que chaque patient est évalué et traité de façon rigoureuse et systématiquement selon les protocoles de traumatologie reconnus (ABLS ─ Advanced Burn Life Support et ATLS ─ Advanced Trauma Life Support). Parmi les priorités vitales retrouvées chez la majorité des grands brûlés, le diagnostic d'obstruction des voies respiratoires par une lésion d'inhalation ou par la présence d'œdème doit être fait promptement et les voies aériennes sécurisées par le moyen d'une intubation endotrachéale précoce accompagnée de ventilation mécanique. L'évaluation de la circulation, de la pression sanguine et de la fonction cardiaque est primordiale dès les premières heures après une brûlure étendue. La survenue d'une brûlure significative s'accompagne d'une perte de liquides massive survenant dans le compartiment intra vasculaire, ce qui a pour effet de provoquer une hypotension importante, allant jusqu'au choc et au décès si le patient n'est pas réanimé avec les fluides nécessaires. L'obtention rapide d'un accès intraveineux adéquat est capitale, de préférence sur les régions non atteintes, ainsi que l'évaluation du %TBSA qui déterminera, avec le poids sec du sujet, la quantité de fluides à administrer dans les premières 24 heures. Les cliniciens administrent un premier apport en fluide (solution de Lactate- Ringer[1]) pour les brûlures de 2ième et/ou de 3ième degré de plus de 15% TBSA pour les enfants ou supérieurs à 20% pour les autres groupes d'âge. Cela permet

[1] Soluté composé de chlorure de sodium, lactate de sodium, de potassium et de calcium dans de l'eau.

le maintien du système circulatoire, cardiaque, rénal et respiratoire du patient. Celui-ci est ensuite transféré de manière appropriée à une unité des brûlés.

En troisième lieu viennent les soins aigus et la réadaptation précoce par l'équipe spécialisée interdisciplinaire du centre de brûlés, incluant le chirurgien plasticien, le chirurgien traumatologue et l'intensiviste. L'unité des brûlés assure la prise en charge du patient et son suivi auprès de sa famille. Évaluation médicale, réanimation et stabilisation sont poursuivies. Une fois l'installation de voies d'accès veineux sécurisées et les fonctions vitales stabilisées, les plaies sont nettoyées et débridées. Une troisième évaluation du %TBSA est effectuée afin de doser les solutés de réanimation nécessaire à la survie du patient. La profondeur des brûlures est évaluée afin de déterminer si des chirurgies de greffes cutanées seront nécessaires. Enfin, les infirmières procèdent aux pansements selon les recommandations du chirurgien plasticien. Sont également prodigués par cette unité, la réhabilitation physique, émotionnelle et psychologique du patient, ainsi que la reconstruction des cicatrices et le suivi du patient à long terme.

1.2 Évaluation de l'étendue des brûlures

1.2.1 Stabilisation et réanimation liquidienne

Les grands brûlés deviennent rapidement hypothermiques et ont une perte de fluide importante qui doit être remplacée afin de soutenir les organes vitaux du patient, mais également pour maintenir une perfusion tissulaire adéquate à la brûlure et ainsi prévenir une évolution de la gravité de la brûlure (Durrant et al., 2008); c'est le but de la réanimation liquidienne. La plus grande perte de fluide du brûlé a lieu dans les 24 heures après l'agression. Pour déterminer l'apport en fluide dont le patient a besoin, les cliniciens se basent fréquemment sur une formule commune qui est celle de Durrant et al., (2008), appelée formule de Parkland (1.1). Facile à utiliser, elle est la formule de réanimation liquidienne la plus répandue chez les grands brûlés. Des modifications de cette formule ont été décrites, notamment en pédiatrie (Brook modifiée) pondérant la quantité de fluide par %TBSA et poids sec. Toutes les formules utilisent le %TBSA comme point de départ pour évaluer les pertes liquidiennes à remplacer.

Cette formule donne une estimation de l'apport en fluide nécessaire pour les 24 premières heures. Cependant, cette formule n'est qu'un apport initial, puisque la quantité de fluide à administrer au patient doit constamment être ajustée en fonction des paramètres physiologiques surveillés et de la profondeur de la brûlure (Hettiaratchy & Papini, 2004; Durrant et al., 2008; Q. E. Chan, Barzi, Cheney, Harvey, & Holland, 2012).

$$Apport = 4\,ml\ \times TBSA\ (\%) \times Poids\ (kg)$$
$$TBSA = \frac{BBSA}{BSA} \tag{1.1}$$

1.2.2 Méthodes cliniques

L'évaluation du %TBSA demeure critique puisqu'elle permet non seulement d'estimer les besoins volémiques, mais également de déterminer les critères de transfert dans une unité de brûlés et de statuer sur le pronostic vital du patient (Livingston & Lee, 2000; Q. E. Chan et al., 2012), en plus d'être un critère de récupération (Livingston & Lee, 2000; Wachtel et al., 2000). Il existe trois méthodes communes pour l'évaluation de la TBSA. Ces méthodes « papier et crayon » requièrent une estimation visuelle (voir Figure 1.2).

La règle des neuf de Wallace utilise un schéma en 2D représentant un corps humain avec des proportions corporelles fixes de 9% (TBSA). La règle des 9 de Wallace ne peut s'appliquer à un enfant de moins de 15 ans car ses proportions corporelles sont différentes de l'adultes.

La surface palmaire qui consiste à considérer la surface palmaire complète de la main comme étant 1% de la BSA du patient. Elle présente cependant de nombreuses ambiguïtés quant à son utilisation (Jose, Roy, Vidyadharan, & Erdmann, 2004). En effet selon certains auteurs, c'est la paume de la main avec les doigts qui représentent les 1% (Hettiaratchy & Papini, 2004; Durrant et al., 2008), alors qu'au contraire, pour d'autres auteurs c'est seulement la paume de la main (Malic, Karoo, Austin, & Phipps, 2007). De plus selon certains auteurs, la surface

palmaire ne représenterait pas 1%, mais plutôt 0,78% (Gehan & George, 1970), chez le patient obèse entre autres.

Le schéma de Lund & Browder (Figure 1.2) est un schéma en 2D représentant un corps humain avec des proportions corporelles fixes plus adaptées aux adultes et aux enfants qui ont des proportions corporelles variables en fonction de leur âge. Des trois méthodes, c'est la plus précise et pourtant la moins utilisée (Daryousch Parvizi et al., 2014).

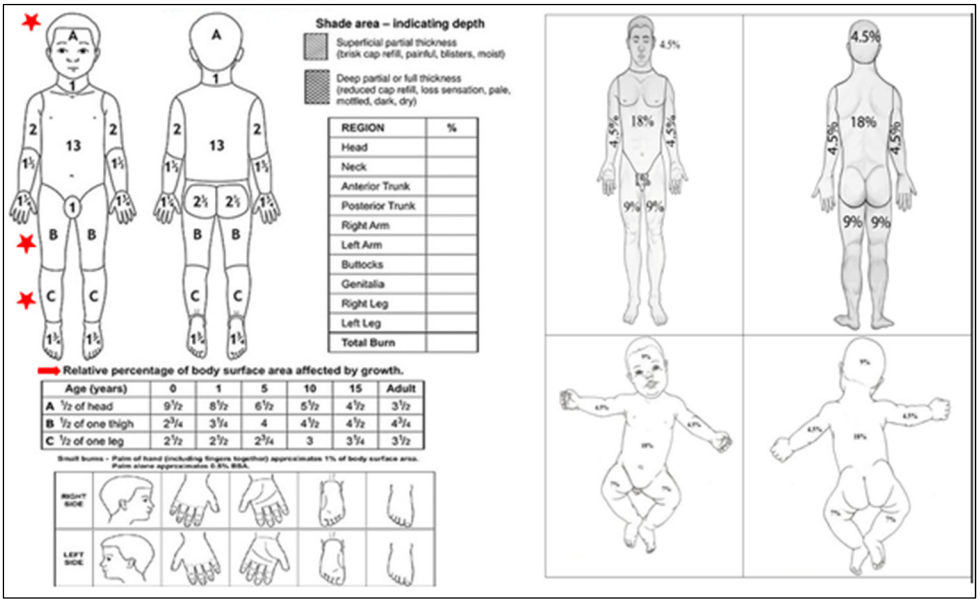


Figure 1.2 Exemple réel du schéma de Lund & Browder tel qu'utilisé à l'hôpital St-Justine de Montréal. Le schéma provient du site web (http://www.urgencehsj.ca/protocoles/brulukres/). Avec autorisation.

Le Tableau 1.1 ci-après présente un résumé des méthodes cliniques et leurs limitations sur la base des études suivantes : (Rossiter, Chapman, & Haywood, 1996; Livingston & Lee, 2000; Wachtel et al., 2000; Tikuisis, Meunier, & Jubenville, 2001; Yu, Lo, & Chiou, 2003; Hettiaratchy & Papini, 2004; Jose et al., 2004; Verbraecken, Van de Heyning, De Backer, & Van Gaal, 2006; Malic et al., 2007; Durrant et al., 2008; Neaman et al., 2011; Giretzlehner et al., 2013; Rae et al., 2013; R. Y. Williams & Wohlgemuth, 2013; D. Parvizi et al., 2014; Harish et al., 2015).

L'évaluation du %TBSA par ces méthodes a tendance à être surestimée (Hettiaratchy & Papini, 2004; Malic et al., 2007; D. Parvizi et al., 2014; Daryousch Parvizi et al., 2014) et est sujette à des variations entre les observateurs (Durrant et al., 2008). C'est encore plus le cas pour des personnes obèses (Rae et al., 2013; R. Y. Williams & Wohlgemuth, 2013), ou chez les femmes aux fortes poitrines du fait que le gras s'accumule de façon disproportionnelle (Livingston & Lee, 2000; Hidvegi, Nduka, Myers, & Dziewulski, 2004), ou encore les femmes enceintes (Kennedy, Baird, & Troiano, 2008).

Pourtant l'obésité est une des situations cliniques fréquentes (Neaman et al., 2011), puisque c'est un problème commun à tous les pays développés où le taux d'obésité augmente constamment, allant jusqu'à être supérieur à 50% dans certaines populations (Livingston & Lee, 2000). On retrouve également ces mêmes problématiques chez les enfants qui sont des cas avec la plus haute mortalité due principalement à une mauvaise évaluation de TBSA (Fairbrother, 2015) et là encore, c'est pire pour des enfants obèses (Livingston & Lee, 2000).

Or il a été démontré par certains auteurs que des erreurs dans le calcul de la superficie brûlée peuvent entraîner des complications médicales et chirurgicales significatives (Klein et al., 2007; Daryousch Parvizi et al., 2014). En effet, une surestimation dans les calculs entraîne un apport en fluide excessif et a pour conséquence de « noyer » le patient tandis qu'une sous-estimation place le patient en choc hypovolémique avec risque de collapsus. Ces deux conditions critiques peuvent entraîner la mort du patient (Wachtel et al., 2000; Klein et al., 2007). Klein et al., (2007) ont étudié la relation entre l'apport en fluide administré et la survie de 72 patients d'âge moyen de 46 ans (18 ─ 86 ans) ayant un %TBSA moyen de 44.5% (20% ─ 90%). Ils ont observé qu'une sous ou surestimation de 1.5 litres dans l'apport en fluide (due à des erreurs dans le calcul du %TBSA) entraîne, en moyenne, la mort du patient dans 25% des cas, la défaillance de plusieurs organes dans 21% des cas, un syndrome de détresse respiratoire aiguë dans 35% des cas, l'infection du sang dans 11% des cas, et une pneumonie dans 54.9% des cas.

Le plus significatif est le fait que ces méthodes exigent la transposition de brûlures représentées en 3D sur des schémas en 2D (Wachtel et al., 2000; Prieto, Acha, Gómez-Cía, Fondón, & Serrano, 2011).

Tableau 1.1 Tableau synthèse des différentes méthodes cliniques

Méthodes	**Règle des 9**	**Surface palmaire**	**Schéma de Lund & Browder**
Type de Brûlure	Moyennes et grandes brûlures	Brûlures < 15% et > 85%	Tous types de brûlures
Erreurs %TBSA	79% chez les enfants. Augmente la taille des brûlures de 3% Sous-estime et surestime les brûlures. Femmes aux fortes poitrines sous-estimées	Utilisation ambiguë	Femmes aux fortes poitrines sous-estimées
	Variabilité entre les observateurs de plus de 62% et Surestimation de 50% à 161%		
Adultes?	OUI	OUI	OUI
Enfants ?	NON. Proportions des enfants différentes de celle du schéma.	OUI	OUI
Obèses ?	NON	NON	NON
Préférence clinique	38%	37%	18%

1.2.3 Erreurs liées à l'estimation de la surface corporelle

Comme il a été vu précédemment, la BSA d'un patient est un paramètre déterminant dans l'évaluation du %TBSA. Il est donc nécessaire de s'attarder dans un premier temps sur l'estimation de la BSA et sur les méthodes utilisées pour l'estimer. À l'origine D. Du Bois et E. F. Du Bois, (1916) ont été les premiers à proposer une formule anthropométrique en fonction du poids et de la taille. À partir d'une analyse de régression effectuée sur 9 sujets, ils ont pu estimer une approximation de la BSA :

$$BSA = C \times P^A \times T^B \tag{1.2}$$

Avec P le poids en kg, T la taille en cm et A, B, C des constantes prédéfinies.

Par la suite, de nombreux auteurs ont tenté d'améliorer cette formule en améliorant les constantes prédéfinies A, B et C de l'équation. Ces mêmes auteurs affirment que le poids et la taille étaient de meilleures variables explicatives pour la BSA. Cependant, ces formules sont souvent critiquées du fait que leurs précisions sont limitées par l'estimation de coefficients de régression qui dépendent strictement de la population étudiée et ne permettent pas une représentation des morphologies variées. Par exemple, la formule de Yu et al., (2003) est basée sur l'étude de 3951 adultes chinois de morphologie normale. L'appliquer à une population nigérienne aura pour conséquence de surestimer la BSA de 6 à 22% (Nwoye, 1989). Tikuisis et al., (2001) ont grandement amélioré l'estimation de la BSA en se basant sur un échantillon de 641 Nord-Américains et Européens (jeunes adultes et adultes) de différentes morphologies. Mais là encore, appliquer leur formule à une population d'enfants conduira à des erreurs d'estimations évidentes.

Néanmoins, bien que ces formules présentent des résultats relativement satisfaisants pour l'estimation de la BSA dans certains cas, ce n'est pas forcément le cas pour l'estimation de la superficie corporelle locale (morphologie). De plus, dans la plupart des cas, il y a une surestimation de la BSA avec une tendance à augmenter la taille du corps (Bailey & Briars, 1996), mais également une sous-estimation des personnes obèses (Livingston & Lee, 2000; Verbraecken et al., 2006; Lee, Choi, & Kim, 2008). En effet, la forme du corps humain présente des caractéristiques typiques et de nombreux traits délicats. Mais encore, elles ne mesurent que les caractéristiques brutes du corps humain et ne prennent pas en compte l'âge, le sexe et les différentes morphologies humaines. Par conséquent, la superficie corporelle locale peut différer énormément de l'approximation de la BSA et ainsi impacter le %TBSA. C'est le cas des méthodes cliniques communément utilisées de nos jours qui sont issues de données modifiées de la formule de Du bois (D. Du Bois & E. Du Bois, 1916; Knaysi, Crikelair, & Cosman, 1968). Lund & Browder ont révisé cette formule pour élaborer leur schéma (Charles C Lund & Browder, 1944). De même que Gehan et George, (1970) pour estimer la surface palmaire. Ces formules présentent donc des erreurs de mesure inévitables et ne sont pas adaptées pour l'évaluation du %TBSA chez des grands brûlés.

Ainsi, à la vue du contexte clinique, il y a une nécessité d'estimer de manière précise le %TBSA (Kamolz, Parvizi, et al., 2014). La problématique du projet de recherche peut ainsi être posée : comment améliorer de manière répétable la précision de l'estimation du %TBSA? Sachant que les méthodes cliniques ne prennent pas en compte la BSA et la morphologie spécifique du patient, comment prendre en compte ces paramètres importants ? Comment mieux personnaliser la prise en charge initiale des patients ? Compte tenu du besoin technologique, existe-t-il des méthodes numériques pour l'évaluation des surfaces brûlées ? Existe-t-il des outils cliniques ? La communauté scientifique a-t-elle déjà pensé à une fusion de méthodes et/ ou technologies ? Le CHAPITRE 2 ci-après tentera de répondre à ces questions.

Clicours.com

CHAPITRE 2

REVUE DE LITTÉRATURE

2.1 Modélisation des surfaces corporelles

2.1.1 Modélisation 2D/ 3D des surfaces corporelles brûlées

Les méthodes cliniques exigent une représentation des brûlures en 3D sur des schémas en 2D (Wachtel et al., 2000). C'est pourquoi cette évaluation n'est pas objective et présente des variabilités intra-inter observateur non négligeables. De nombreux auteurs se sont donc tournés vers le développement de méthodes numériques pour estimer le %TBSA. Les premiers programmes d'évaluation de la taille des brûlures (Tableau 2.1) étaient principalement une représentation numérique du schéma de Lund & Browder avec l'ajout de fonctions supplémentaires. Cependant, ces méthodes restent toujours une représentation en 2D des brûlures et souffrent donc des mêmes problèmes mentionnés précédemment.

Afin de pallier cette problématique, certains auteurs se sont tournés vers le développement de programme de modélisation 3D des brûlures (Tableau 2.2). Ces logiciels sélectionnent un modèle déformable (non paramétrique) provenant d'une base de données plus ou moins exhaustive suivant les données inscrites par l'utilisateur. La surface globale du modèle 3D est ensuite adaptée de manière à correspondre à l'approximation de la BSA du patient (Figure 2.1).

« BurnCase 3D » et « BAI » sont les seuls de nos jours à offrir une solution adéquate pour l'évaluation du %TBSA. Cependant, les modèles 3D sont issus d'une approximation de la BSA du patient basée sur les formules anthropométriques vues précédemment (1.2). Le fait que des modèles prédéfinis soient seulement adaptés au poids et à la taille du patient implique un décalage inconnu entre la BSA du patient réelle et la BSA du modèle 3D qui n'a pas été évalué par « BurnCase 3D » et « BAI » (Johannes Dirnberger, Michael, Robert, Herbert, & Peter, 2012). Seule une mesure du décalage de la géométrie 3D peut être évaluée en comparant les modèles obtenus à des scans 3D. Ces logiciels n'ont pas pris en compte la corrélation existante

CLICOURS.COM

entre les différentes parties du corps tel qu'on peut le représenter schématiquement sur la Figure 2.1.

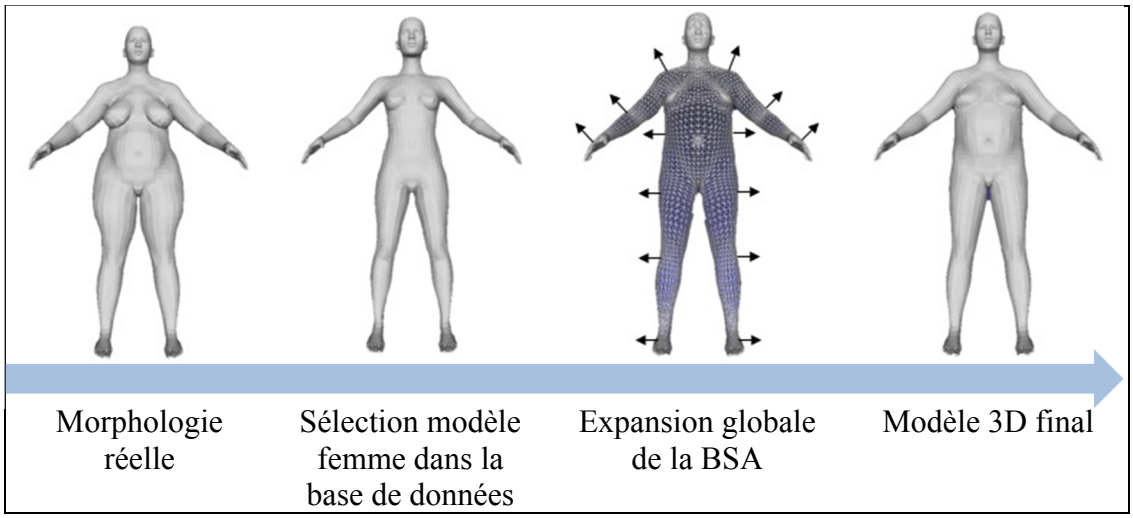


Figure 2.1 Problématique des logiciels de modélisation 3D actuels

Tableau 2.1 Présentation des méthodes numériques du schéma de Lund & Browder

Auteurs	**SAGE II, Neuwalder et al., (2002).**	**CFU/HD, Richard, Jones, et Parshley, (2015).**	**BurnCalculator, Berry, Goodwin, Misra, et Dunn, (2006).**
Avantages	Calcul le TBSA automatiquement. Informations sur le type de profondeur de brûlure et recommandation d'infusion en fluide. Meilleure évaluation du TBSA que les méthodes cliniques « papier/crayon ».		
	Réduis la surestimation et améliore la répétabilité.	Schéma plus précis. Décomposé en 401 régions corporelles.	-
	Fonction d'amputation et sites donneurs.		
Inconvénients	Erreur d'estimations. Transposition de brûlures représentées en 3D sur un schéma en 2D.		
Validation clinique?	Oui. Testé par 5 infirmières, 3 médecins, une diététicienne et 1 physiothérapeute.	Non.	Oui. Basé sur l'étude de 60 dossiers médicaux
Précision?	Inconnue.	Inconnue.	25% des cas surestiment la taille des brûlures et 15% des cas sous-estiment la taille des brûlures.

Or les caractéristiques typiques du corps ainsi que les différences de morphologies constituent une différence notable dans l'évaluation de la superficie brûlée. Les modèles 3D sont toujours une approximation de la constitution réelle du patient.

De nombreux efforts ont été réalisés pour capturer la réelle constitution d'un sujet avec une Kinect (Popescu & Lungu, 2014; Sheng et al., 2014) qui est un capteur à lumière blanche structurée. Malgré des modèles 3D personnalisés, de nombreuses erreurs et artefacts dus à la faible résolution de la Kinect ont été observés avec la nécessité d'effectuer des post-traitements (Tableau 2.3).

Tableau 2.2 Synthèse des différentes méthodes de modélisation 3D des brûlures

<table>
<tr><th>Auteurs</th><th>EPRI 3D Burn Vision. Neuwalder et al., (2002).</th><th>BurnCase3D. Giretzlehner, Dirnberger, Luckender, Haller, et Rodemund, (2004).</th><th>BAI. Prieto et al., (2011).</th></tr>
<tr><td>Avantages</td><td>Évalue régions brûlées, greffes et sites donneurs potentiels sur un modèle 3D. Semble améliorer la répétabilité et réduire la surestimation du TBSA</td><td>Modèle 3D. Interface intuitive. Calcul du TBSA. Calcul des fluides requis. Documentation des antécédents. Détails et blessures du patient. Intégration de photos dans une base de données.</td><td rowspan="2">Patient normal, en surpoids ou obèse parmi une base de données de 8 modèles. Sélection de la profondeur suivant un jeu de couleur. Calcul du TBSA.</td></tr>
<tr><td rowspan="3">Inconvénients</td><td rowspan="2">Ne génère qu'un seul modèle représenté en maillage et non en surface lisse.</td><td>Patients obèses provoquent des artefacts.</td></tr>
<tr><td colspan="2">Modélisation du modèle basé sur les formules anthropométriques.</td></tr>
<tr><td colspan="3">Sous et surestimations du %TBSA. Modèle non adapté à la morphologie des patients. Non personnalisable à leurs mensurations. Ne prends pas en compte les corrélations entre les différentes parties du corps humain et des différentes morphologies</td></tr>
<tr><td>Validation clinique?</td><td>Testé auprès de 5 infirmières, 3 médecins, 1 diététicienne et 1 physiothérapeute.</td><td>Oui. Sur des brûlures de 20% et plus.</td><td>Oui. 80 patients âgés de 9 mois à 82 ans.</td></tr>
<tr><td>Précision?</td><td>Très faible. Modèle 3D peu précis. Présente des erreurs d'estimations égales aux méthodes cliniques.</td><td>Inconnu. Erreurs de surestimations et de variabilité plutôt faibles.</td><td>Inconnue. Différence significative entre « BAI » et les méthodes cliniques.</td></tr>
</table>

De plus, ces méthodes présentent un temps d'acquisition élevé (>15 min), une impossibilité de déplacer le sujet durant la procédure d'acquisition et une absence d'interface pour l'évaluation

de la TBSA. Ces méthodes sont donc peu adaptées aux besoins cliniques nécessitant une évaluation rapide et précise du %TBSA.

Tableau 2.3 Synthèse des systèmes d'acquisition Kinect pour la modélisation 3D des brûlures

Référence	**BurnCalc. Sheng et al., (2014).**	**Popescu et Lungu, (2014)**
Méthode	Système : Kinect Reconstruction : algorithme « marching cubes » Acquisition : distance de 50cm	
Avantages	Scan 3D d'un sujet. Modèle personnalisable au sujet.	
Inconvénients	Temps total : 15 min à 23min. Hypothèse : patient immobile donc impossible pour des patients ambulants ou en soins critiques. Modèles : peuvent être incomplets ou impossible à scanner. Erronés pour une distance de 0,5m à 1m. L'erreur augmente proportionnellement avec la distance. Kinect : déplacement lent sans saccade. Faible résolution. Sensible à la réflexion de lumière. Post-traitements manuels indispensables. Pas d'interface graphique pour l'évaluation du TBSA.	
Validation clinique?	Non. Validation sur 17 sujets simulant des brûlures par un chirurgien. Comparaison du TBSA obtenue par la règle des neuf et la surface palmaire versus BurnCalc versus la formule anthropométrique de Yu, Lo et al. (2003).	Non. Testé pour trois tailles de brûlures simulées sur des objets planaires de superficie connue.
Précision?	Meilleure évaluation que les méthodes cliniques. Résultats presque proches de la formule de Yu, Lo et al. (2003). Erreurs d'estimation de la TBSA de 10%.	Pour un nuage de points d'un objet échantillon non normalisé on obtient une erreur moyenne de 48,7% (cas de l'étude de gauche, *BurnCalc*). En utilisant un filtre de moyenne, on a une erreur moyenne de 7,32 %

2.1.2 Modélisation 3D de la surface corporelle du corps humain

En dehors du contexte clinique, il existe plusieurs approches pour la modélisation 3D du corps humain. Ces approches sont résumées dans le Tableau 2.4 et Tableau 2.5 ci-dessous.

Tableau 2.4 Approche de création et de reconstruction

Approches	Création	Reconstruction
Méthode	Numérisation corps entier : scanner/Kinect/TOF. Maillage à surface ouverte (non consistant).	Reconstruction du scan à surface ouverte par correspondance avec un modèle 3D fermé afin d'obtenir un scan à surface fermée (consistant).
Avantages	✓ Plus précis et plus rapide (ms)	✓ combler les données manquantes.
Inconvénients	✗ Très cher. ✗ Données manquantes ✗ Artefacts mouvement, respiration ✗ Post-traitements manuels. ✗ Temps coûteux inacceptable. ✗ Dimension du scanner incompatible avec une salle de chirurgie. ✗ Scanne debout impossible avec des patients ambulants.	✗ À part les données manquantes, même problématiques que l'approche de création. ✗ Correspondance exacte seulement si le scan et le modèle ont la même pose. Cette contrainte est impossible à réaliser dans le contexte des grands brûlés (voir ANNEXE IX). ✗ Impossible de modifier les modèles une fois reconstruits suivant les intentions de l'utilisateur.

L'approche la plus répandue, dite de création, consiste en une acquisition directe de la surface corporelle (voir Tableau 2.4 – Création). L'acquisition est réalisé soit par un scanner 3D (fixe ou portable), tel que proposé par les études suivantes : (Daanen & Van De Water, 1998; Robinette, Daanen, & Paquet, 1999; Tneb, Seidl, Hansen, & Pruett, 2000; Yu et al., 2003; Treleaven & Wells, 2007; Weiss, Hirshberg, & Black, 2011). D'autres auteurs proposent l'utilisation de Kinect Xbox, telle que proposée par les études suivantes : (Aitpayev & Gaber, 2012; Tong, Zhou, Liu, Pan, & Yan, 2012; Cui, Chang, Nöll, & Stricker, 2013; Khongma et al., 2014; Popescu & Lungu, 2014; Sarbolandi, Lefloch, & Kolb, 2015; Yao, Dong, & Hu, 2015). Ou encore les capteurs TOF (Time of Flight), proposés par les études suivantes: (Tao, Srinivasan, Malik, Rusinkiewicz, & Ramamoorthi, 2015).

Malheureusement, des erreurs majeures sont observées : la dérive de données et la perte de données. En général, le capteur TOF tend à montrer des données aberrantes, tandis que le capteur à lumière structurée (Kinect/ scanner) souffre de perte de données. En particulier, le capteur de TOF produit un grand décalage dans les valeurs de profondeur le long des pixels de bord et les pixels transparents, ce qui correspond à une dérive des données. Dans les mêmes conditions, le capteur à lumière structurée tend à produire des trous, dans lesquels la profondeur ne peut pas être estimée. Pour les deux capteurs, les reflets spéculaires conduisent à la perte de données (Hansard, Lee, Choi, & Horaud, 2012; L. Li, 2014; Sarbolandi et al., 2015). Bien que cette approche soit la plus rapide et précise, des post-traitements sont obligatoirement nécessaires pour combler les données manquantes occasionnées par des artéfacts (surfaces d'occlusions, mouvements, recalage de différentes parties du corps, complexité du scanner, etc.), en plus de nécessiter un technicien. Toutes ces contraintes ralentissent le processus d'acquisition (voir chapitre 2.1.1 sur la Kinect également).

Des méthodes de reconstruction (voir Tableau 2.4 – Reconstruction) permettent de combler les trous (Allen, Curless, & Popović, 2003; Allen, 2005; Chu, Tsai, Wang, & Kwok, 2010). Cependant la contrainte de pose (debout jambes et bras écartés) associée à cette approche la rend inadaptée pour certains contextes cliniques où les patients sont couchés et inconscients. De plus, notre expérience a démontré que les approches par acquisition directe ou par reconstruction présentent de nombreux inconvénients ne permettant pas de répondre aux besoins et exigences cliniques des grands brûlés de manière satisfaisante, intéressante et bénéfique (voir ANNEXE IX SCANNER PORTABLE : PROBLÉMATIQUE POUR LES GRANDS BRÛLÉS). La précision de l'acquisition 3D de ces deux approches est directement liée à la précision du capteur utilisé (et souvent de l'ordre du mm).

Une autre approche populaire, dite d'interpolation (voir Tableau 2.5 – Interpolation), consiste à générer un modèle statistique à partir d'une base de données de scans d'une population (Seo & Magnenat-Thalmann, 2003; Magnenat-Thalmann, Seo, & Cordier, 2004; Seo & Magnenat-Thalmann, 2004; Allen, 2005; Baek & Lee, 2012). Bien que la modélisation est plus robuste (précision liée aux scans 3D de l'ordre du mm), le modèle construit dépend, là encore (voir

section 1.2.3), essentiellement de la population étudiée et ne permet pas une représentions des morphologies variées (Ranger, 2010). L'acquisition d'une base de données de scans de corps humains est un processus très long et fastidieux ou sinon très cher (> 10 000 $). De plus, certains auteurs ont démontré qu'elles ne représentaient pas suffisamment la population (Nadadur & Parkinson, 2010; Parkinson & Reed, 2010; Kulaga, Myers, & Gillich, 2014). C'est le cas des deux seules enquêtes anthropométriques CAESAR de Robinette et al., (1999) et ANSURE de Gordon et al., (2013). Toutes deux ayant acquis la surface corporelle des sujets avec un scanner à lumière blanche. Comme qu'observé par Ranger, (2010), ANSUR est moins représentative des civils d'aujourd'hui en raison de données provenant de militaires. CAESAR est bien plus représentative, mais pour des civils de morphologie normale. La distribution des morphologies n'est pas représentée pour des enfants, obèses, etc.

Tableau 2.5 Approche d'interpolation et de modèle

Approches	Interpolation	Modèle
Méthode	1- Entrée des mesures (âge, poids, taille) 2- L'algorithme statistique trouve la relation entre ces mesures et la forme du modèle la plus proche dans la base de données 3- Le modèle est choisi en fonction des paramètres inscrits par l'utilisateur	Production de la forme du corps par déformation d'un modèle. La somme des déformations des segments permet la déformation globale du modèle.
Avantages	✓ Base de données : meilleure ressource disponible pour modéliser et estimation des corrélations entre les mesures et la forme du corps. ✓ Permet de guider l'estimation des informations incertaines. Modélisation plus robuste.	✓ Approche la plus rapide et la plus simple. ✓ Bonne alternative pour pallier l'inconvénient de l'approche de reconstruction. ✓ Permet un contrôle de haut niveau du modèle cible tout en maintenant la qualité qui existe dans le modèle type.
Inconvénients	✗ Nécessité des prétraitements de la base de données (mêmes problématiques que l'approche de reconstruction) ✗ Prend beaucoup de temps de chargement. ✗ Ne répond pas aux exigences de temps de la prise en charge des grands brûlés. ✗ Impossible de modifier les modèles une fois reconstruits suivant les intentions de l'utilisateur.	✗ Ne prends pas en compte la corrélation entre les parties du corps.

Enfin, l'approche de déformation par segment (Seo, Cordier, Philippon, & Magnenat-Thalmann, 2001; Kasap & Magnenat-Thalmann, 2007; J. Li & Wang, 2007) aussi appelée modélisation paramétrique (voir Tableau 2.5 – Modèle) est celle proposée pour le présent projet de recherche. Cette approche a l'avantage de permettre un contrôle de haut niveau du modèle cible tout en maintenant la qualité qui existe dans le modèle de base. Cependant, un reproche fait à cette approche est le fait qu'elle ne prend pas en compte la corrélation entre les parties du corps. À noter que la précision de ces méthodes n'a pas été évaluée.

Il est a noté cependant, que ces approches s'inscrivent dans un contexte différent du présent projet de recherche, tel que l'infographie, design ergonomique, vision par ordinateur, animation, design vestimentaire, jeux vidéo ou encore environnements virtuels ludiques. L'enjeu n'est pas la précision, mais l'acquisition, ou la modélisation 3D d'un modèle de corps humain. Autrement dit, ces approches ne sont pas optimales pour les exigences du projet (voir Exigence 04 à Exigence 06 au CHAPITRE 3).

2.1.3 Un outil pertinent pour la modélisation 3D des surfaces corporelles : MakeHuman

Il existe plusieurs logiciels permettant la modélisation paramétrique du corps humain, notamment pour les jeux vidéo. Le logiciel multiplateforme open source (licence LGPL) MakeHuman™ (MH) de Bastioni, Re, et Misra, (2008) est celui le plus utilisé pour la modélisation paramétrique de corps humain en 3D dans la recherche scientifique. Basé sur l'étude de plus de 3000 morphologies, MH opère une déformation paramétrique contrôlée du modèle tout en préservant les proportions corporelles du corps et donc sa morphologie. Le modèle de base est un maillage composé de quadrilatères ayant une forme humaine androgyne. La déformation du modèle est effectuée par morphing 3D. C'est toujours le même modèle de base qui est déformé avec la même topologie. MH peut modéliser de façon réaliste toutes les formes anatomiques humaines (Fonseca, Bogaerts, Hunt, & Vanhavere, 2014), grâce à une modélisation paramétrique basée sur des paramètres simples et intuitifs tels que l'âge, le genre, la taille, plusieurs mesures anthropométriques et les proportions du corps.

Plus précisément, la déformation paramétrique ou « Keyframe morphing » permet de changer la forme d'un objet tridimensionnel donnée par la relocalisation de ses sommets. Les vecteurs de ces sommets sont définis afin de décrire un chemin de déformation linéaire (interpolation linéaire en translation et rotation). Un groupe de ces vecteurs est appelé un « morphing target », qui couvre habituellement seulement un sous-ensemble de sommets du modèle complet. MakeHuman en compte plus de 60. Cette déformation est contrôlée par un facteur de forme qui définit l'impact de ces vecteurs et, par conséquent, la dimension de cette relocalisation. Enfin, les normales aux faces sont recalculées afin de limiter la déformation locale pour un maillage plus lisse et plus uniforme. C'est la fonction « mesures » de MH qui permet, à partir de mesures anthropométriques, de déformer localement le modèle 3D en fonction des « mophing target » associé à ces mesures indépendamment des autres.

La combinaison originale et intelligente de morphing en translation et rotation, associée à un facteur de forme et un lissage du maillage permet à MH de fournir une gamme de modèles réalistes (voir Figure 2.2 ci-dessous).

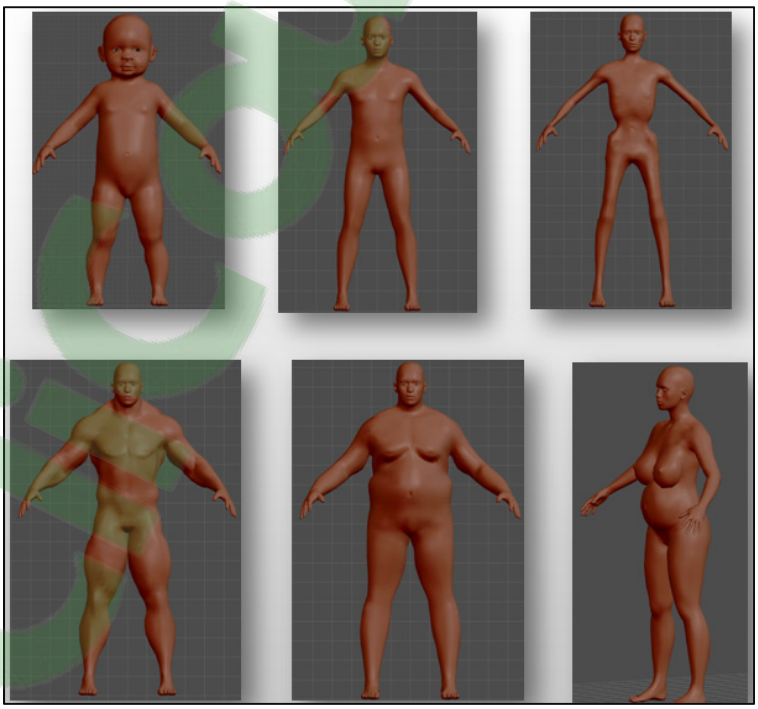


Figure 2.2 Aperçu des modèles de MH Modèles créés à partir du logiciel MH http://www.MH.org/

Plusieurs études ont utilisé MH pour la création de modèles 3D humains personnalisés: essayage de vêtements virtuels (Volz, Blum, Häberling, & Khakzar, 2007) ; médecine et robot-assistance (Buys, Deun, Laet, & Bruyninckx, 2011) ; ergonomie du sommeil (Van Deun, Verhaert, Buys, Haex, & Vander Sloten, 2011) ; simulation de l'exposition humaine aux UV solaire (Vernez et al., 2011) ; fantôme virtuel pour la radioprotection, radiologie et médecine nucléaire (Cassola, de Melo Lima, Kramer, & Khoury, 2009; Fonseca et al., 2014) ; autopsie virtuelle précise d'abus subis par des victimes, négligence et malnutrition sévère chez une enfant (Davy-Jow, Lees, & Russell, 2013); évaluation militaire de la vulnérabilité aux munitions fragmentées (Kulaga et al., 2014).

Les résultats de l'étude de Van Deun et al., (2011), sont particulièrement intéressants. Leurs besoins étaient d'obtenir des modèles suffisamment précis afin de simuler la mémoire de forme du matelas. Ils leurs étaient également souhaitable de pouvoir générer des modèles à partir d'un nombre limité de mesures anthropométriques. Au moment de l'étude, MH n'offrait pas encore la possibilité d'entrer des mesures anthropométriques. Les auteurs ont donc modifié le code source de MH de manière à inclure certaines mesures anthropométriques nécessaires. En se basant sur 60 sujets (30 hommes et femmes), ils ont comparé les paramètres estimés du modèle avec les mesures réelles des sujets. La conclusion est qu'en utilisant six paramètres d'entrée tels que le sexe, l'âge, la taille, le poids (scalaire), la largeur des hanches et la circonférence de l'abdomen, les modèles résultants avaient une précision suffisante avec une erreur quadratique moyenne de 5,0 ± 3,5%.

On retiendra également l'étude de Buys et al., (2011), qui illustre deux algorithmes basés sur des caméras pour générer automatiquement des avatars humains (modèle squelettique et maillage issu de MH). Dans un contexte médical et de robot-assistance, ces avatars sont nécessaires pour la reconnaissance et l'identification de geste. D'autre part, ils font remarquer que le principal avantage de MH est que les différents paramètres de MH sont couplés: par exemple, si l'âge estimé est modifié, les autres paramètres tels que la hauteur de l'avatar changeront automatiquement en conséquence. Ce couplage permet à l'utilisateur de réduire le nombre de paramètres nécessaires pour chaque application spécifique, afin de diminuer le

temps de calcul. L'approche de déformation par segment mentionnée dans le Tableau 2.5 à la section 2.1.2 précédente est différente de l'approche utilisé par MH. En effet, les proportions entre les différentes parties du corps sont prises en compte avec MH beaucoup plus simple et plus rapide, cet outil très prometteur a de fortes chances de répondre aux besoins et aux exigences cliniques de façon intéressante et rentable.

2.2 Anthropométrie

Selon le rapport d'un comité d'experts de l'Organisation mondiale de la Santé, l'anthropométrie est la seule et unique technique à la fois portable, universellement applicable, bon marché et non invasive, qui permette d'évaluer la corpulence, les proportions et la composition du corps humain (World Health Organization, 1995). C'est donc un outil du plus grand intérêt, mais actuellement sous-utilisé.

Comme il a été vu en à la section 1.2.3, la taille et le poids sont les variables les plus utilisées pour approximer la BSA, plus que toutes autres mesures anthropométriques. Cependant, dans un contexte de modélisation 3D, il est impossible de quantifier virtuellement le poids (qui dépend de la masse de l'eau, des muscles, du gras, des os, etc.) de manière suffisamment précise pour en représenter sa morphologie (par exemple, deux personnes peuvent avoir le même poids, mais une morphologie totalement différente). De plus, il n'est pas rare que le poids soit parfois estimé visuellement en contexte clinique (Leary, Milner, & Niblett, 2000). Leary et al., (2000) et Uesugi et al., (2002), ont observé chez les cliniciens des erreurs significatives de plus ou moins 20% dans l'estimation visuelle du poids avec pour conséquence, un impact clinique important.

Cela est d'autant plus préoccupant lorsqu'on sait que le fait d'utiliser seulement deux mesures (taille et poids) pour l'estimation de la BSA n'est pas toujours adapté aux morphologies particulières des enfants (Dubin & Zietz, 1996; C. S. Ahn, Oneill, & Maitz, 2011) et des personnes obèses (Bailey & Briars, 1996; Lee et al., 2008; Adler, Nathanson, Raghunathan, & McGee, 2012, 2013). Les formules anthropométriques habituelles (basées sur le poids et la

Clicours.COM

taille) n'estiment que la surface corporelle brute du corps humain et ne prennent pas en compte l'âge, le sexe et les proportions corporelles du corps. Par conséquent, d'autres mesures anthropométriques sont nécessaires pour quantifier la morphologie (caractéristique de formes locales du corps) et la BSA (caractéristique de formes globales du corps) afin d'obtenir une modélisation 3D précise.

Neaman et al., (2011) ont réalisé une étude intéressante sur 163 sujets hommes/ femmes (28/26 normaux, 22/34 obèses, 18/36 obèses morbides) afin de définir le pourcentage de variance expliqué de la BSA des différentes parties du corps en fonction de l'Indice de Masse Corporelle (IMC). Leurs résultats suggèrent que la partie du tronc du corps (comprenant les mesures de la longueur nuque – tour de taille, longueur tour de taille – hanches et les circonférences du tour de taille et du buste) et la partie de la cuisse (comprenant les mesures la circonférence des hanches, circonférence de la cuisse et longueur jambe supérieure) comptent pour 65% de la variabilité de la BSA. Wells, Treleaven, et Cole, (2007) ont étudié une population de 9617 adultes d'IMC différents afin de démontrer que les prédicteurs les plus élevés du poids après la taille sont la circonférence du tour de taille et de poitrine pour l'homme, ainsi que la circonférence des hanches et de poitrine pour la femme. Selon l'étude de Hidvegi et al., (2004) réalisée sur 60 volontaires (20 hommes, 20 femmes de faibles poitrines et 20 femmes de fortes poitrines), la taille de poitrine d'une femme (fortement corrélée à la circonférence de sous-poitrine) peut altérer grandement les proportions relatives de la surface du tronc antérieur au tronc postérieur. Enfin, Nadadur et Parkinson, (2010) ont observé une forte corrélation entre l'IMC et la largeur des hanches.

Ces études mettent donc en évidences l'intérêt de prendre en considération plus que deux mesures anthropométriques pour une modélisation 3D de la morphologie.

2.3 Modélisation statistique en science des données

2.3.1 Deux cultures : les statistiques et l'apprentissage machine

Un objectif spécifique majeur du projet de recherche est de déterminer un ensemble de mesures anthropométriques réduit qui explique et prédit le mieux la surface corporelle et la morphologie des patients. C'est donc un problème de sélection de caractéristiques (variables) et de prédiction. Soit la détermination de modèles de sélection de variables continues en régression. Selon Breiman, (2001b), il existe deux cultures lorsqu'on souhaite utiliser la modélisation statistique afin d'en tirer des conclusions à partir d'un ensemble de données : la culture de la modélisation de données (statistiques) et la culture de la modélisation algorithmique (apprentissage machine). L'un admet que les données sont générées suivant un modèle de données stochastiques tandis que l'autre utilise des modèles algorithmiques pour traiter un mécanisme de données qu'il considère inconnu. La validation est, là encore différente, puisqu'en statistique on considéra la qualité d'ajustement et l'examen des résidus (estimation des erreurs) alors qu'en apprentissage machine un modèle est validé par la précision de sa prédiction. Les statistiques sont fondées sur des hypothèses (est-ce que les solutions correspondent aux hypothèses?) alors que l'apprentissage machine se fonde sur des résultats (est-ce que ça marche ?).

Il est également important de différencier le problème de classification du problème de régression. Bien que les deux soient reliés à la prédiction, les méthodes et résultats associés sont différents l'un de l'autre. La classification consiste à prédire l'appartenance de classes d'un ensemble de données labélisées et de valeurs discrètes, tandis que la régression consiste à prédire des valeurs numériques continues afin de trouver la meilleure relation qui représente l'ensemble de données.

2.3.2 Modèles et hypothèses de la régression linéaire multiple en statistique

En statistique, le modèle de régression linéaire multiple permet de quantifier la relation entre plusieurs variables indépendantes (ou encore appelées variables explicatives ou prédicteurs) et

une variable dépendante (ou encore appelée variable expliquée ou prédite). Ce modèle statistique qui est dit paramétrique puisqu'il admet que la distribution des erreurs suit une loi normale présente une limitation lorsque les variables indépendantes ne sont pas linéairement indépendantes. À noter que les erreurs constituent la différence entre les valeurs observées de la variable indépendante et ses valeurs prédites par le vrai modèle statistique pour une population dans son ensemble. Cette restriction peut être surmontée par le modèle linéaire général qui est une extension de la régression linéaire multiple. Cependant, la nature de données réelles est souvent imparfaite et peut avoir une distribution des erreurs qui suit une autre distribution. Le modèle linéaire généralisé est une généralisation flexible du modèle linéaire général à d'autres distributions des erreurs et est donc considéré non paramétrique. La Figure 2.3 ci-dessous présente la vue d'ensemble des différences entre les modèles de régression et leurs généralisations.

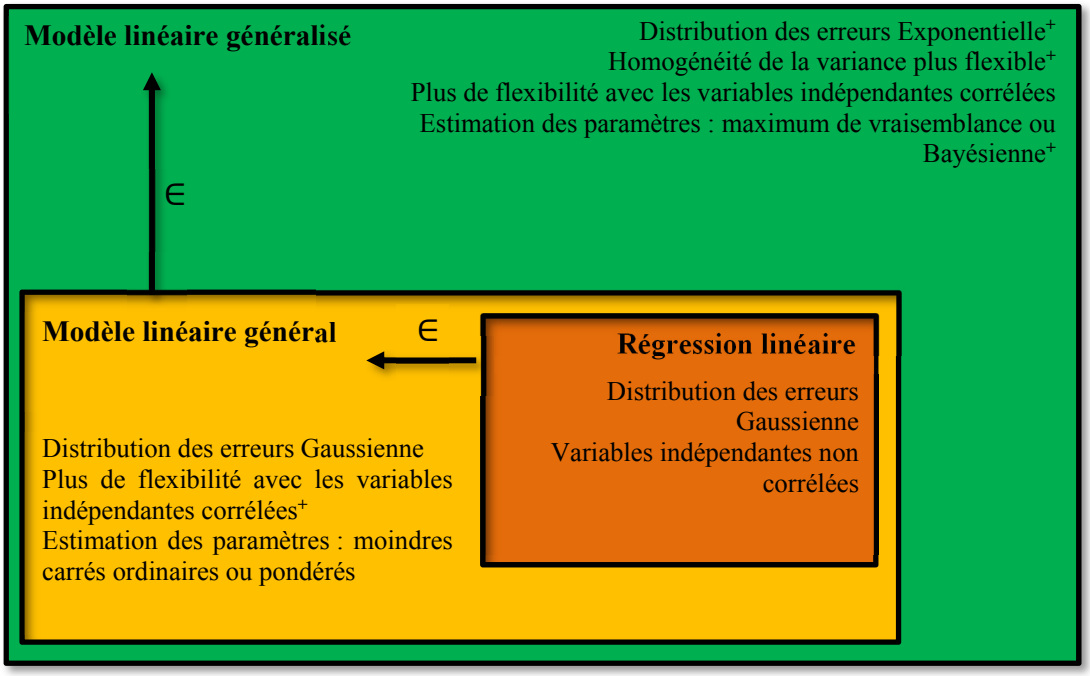


Figure 2.3 Vue d'ensemble des différents modèles de régression en statistique.
Le symbole « + » indique l'extension du modèle par rapport au sous-ensemble.
Le symbole « ∈ » indique l'inclusion du modèle dans le sur ensemble

2.3.2.1 Régression linéaire multiple et ses hypothèses

La régression linéaire multiple permet de construire un modèle afin de prédire la BSA représentée par y qui est la variable dépendante en fonction des valeurs d'une ou plusieurs mesures anthropométriques X_i qui sont appelées variables indépendantes.

Les coefficients de régression $\beta = (\beta_0, \beta_1, \dots, \beta_p)^t$ sont les inconnus du modèle qui représentent les contributions indépendantes de chaque variable (indépendante) à la prévision de y; ε est une erreur aléatoire (bruit) qui est supposée avoir une distribution normale et indépendante avec une moyenne nulle. Le modèle de régression est de la forme linéaire suivante:

$$y_i = \beta_0 + \sum_{j=1}^{p} \beta_i X_{ij} + \varepsilon_i , \qquad i = 1, \cdots, n \tag{2.1}$$

$$\varepsilon_i = y_i - \hat{y}_i \tag{2.2}$$

Les résidus ε_i de l'échantillon sont une estimation des erreurs de la population, également de la forme suivante :

$$\begin{pmatrix} y_i \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_{1,1} & \cdots & x_{1,p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1,n} & \cdots & x_{n,p} \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix} + \begin{pmatrix} \varepsilon_i \\ \vdots \\ \varepsilon_n \end{pmatrix} \tag{2.3}$$

Soit :

$$y = X\beta + \varepsilon \tag{2.4}$$

Les paramètres inconnus β sont le plus souvent estimés par des techniques d'estimation des moindres carrées ordinaires ou pondérés.

La régression linéaire, tout comme le modèle linéaire général étant des méthodes paramétriques, elles admettent une distribution normale des résidus. Certaines hypothèses

doivent donc être vérifiées afin de garantir des résultats et des conclusions non biaisés (Osborne & Waters, 2002).

Pourtant de nombreuses études vérifient les mauvaises hypothèses, ou encore utilisent des tests trop conservateurs et certaines hypothèses plutôt implicites peuvent porter à confusion (M. N. Williams, Grajales, & Kurkiewicz, 2013). Sur la base de ces deux études de Osborne et Waters, (2002) et M. N. Williams et al., (2013), une évaluation informée des hypothèses actuelles de la régression multiple est présentée ci-dessous.

Distribution normale des erreurs : Un diagramme Quantile-Quantile (ou encore *Q-Q plot*) permet d'observer si la déviation suit une distribution normale. D'un autre côté, l'hypothèse d'une distribution normale des erreurs est moins problématique lorsque la taille d'échantillon est suffisamment grande pour référer au théorème central limite. Les coefficients de régression estimés seront toujours non biaisés et consistants. Ce théorème stipule que la distribution échantillonnale des coefficients approche une distribution normale à mesure que la taille d'échantillon augmente, et ce, même si les erreurs ne sont pas normalement distribuées.

Linéarité des paramètres β : Si la relation entre les variables indépendantes et la variable dépendante n'est pas linéaire dans les paramètres, alors les coefficients calculés aboutiront à des conclusions erronées sur la force et la nature de la relation entre les variables du modèle. Visualiser le graphe des résidus ε_i par rapport aux valeurs prédites de la variable indépendante y permet d'observer une structure des résidus non désirée, indiquant des résultats biaisés.

L'équation du modèle de régression comprend une portion non seulement déterministe, mais également stochastique, donc aléatoire et imprévisible. Pour que les résidus concordent avec l'erreur stochastique, ils devraient être centrés sur zéro le long de la gamme des valeurs ajustées et présenter une structure symétrique. Une structure non aléatoire et non symétrique (Figure 2.4, b) indique que la partie déterministe du modèle non linéaire (variables indépendantes) n'explique pas toutes les possibilités qui peuvent être soit une importante variable

indépendante manquante, un coefficient de régression manquant ou encore une interaction manquante entre les coefficients existants.

Moyenne conditionnelle nulle des erreurs : Les erreurs sont assumées avoir une moyenne nulle pour chaque valeur donnée ou combinaison de valeurs des variables indépendantes. La violation de cette hypothèse peut rendre les coefficients de régression biaisés. Les raisons plausibles de cette violation incluent la modélisation d'une relation linéaire au lieu d'une relation non linéaire, et des variables indépendantes qui sont corrélées avec la variable dépendante.

Indépendances des erreurs[2]. (Ne concerne que les données de séries temporelles). Si les erreurs ne sont pas indépendantes, l'estimation des erreurs standard et tests de significativité associés seront biaisés. Néanmoins, les coefficients de régression resteront non biaisés, mais sous-optimaux.

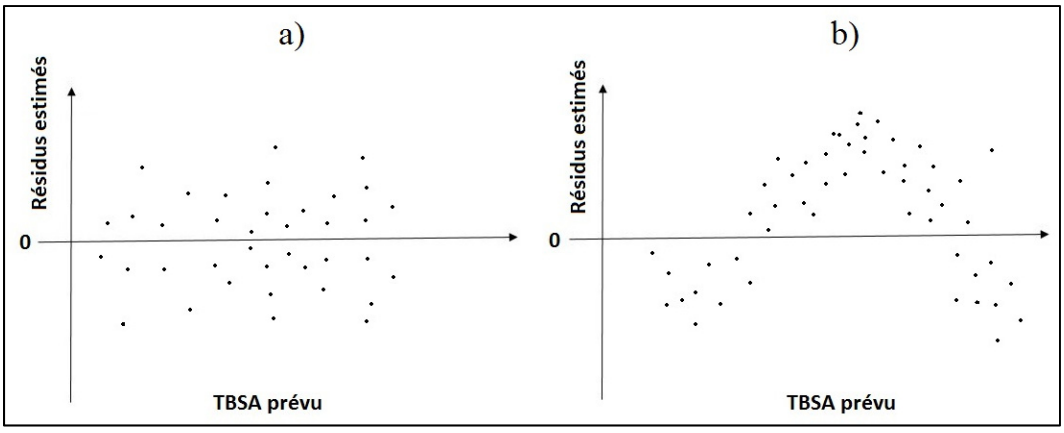


Figure 2.4 a) Structure attendue des résidus.
b) Structure non désirée des résidus (biais)

Homoscédasticité : Les erreurs sont assumées avoir une variance inconnue, mais finie qui soit constante. L'estimation des moindres carrées ordinaires nécessaires à la résolution des

[2] Le présent projet n'est pas concerné par cette hypothèse étant donné la nature transversale des données (unités statistiques de la population étudié à un instant donné).

paramètres sera non biaisée et consistante tant que les erreurs sont indépendantes, mais ne seront pas optimaux.

De plus, toutes inférences ne seront plus fiables puisque les intervalles de confiances, test *t* et de *Fisher* (pour évaluer la significativité) ne seront plus justifiés. Là encore l'homoscédasticité peut être visualisée par le graphe des résidus. Plusieurs alternatives existent pour surmonter l'homoscédasticité telles que des transformations stabilisant la variance, les méthodes *Bootstrap*, une estimation via les moindres carrés pondérés ou encore la spécification d'un modèle linéaire généralisé.

Multicolinéarité : La présence de multicolinéarité entre plusieurs variables peut être un souci dépendamment de ce qui est souhaité : si l'objectif concerne la prédiction, alors la multicolinéarité n'est pas un obstacle puisque la prédiction de la variable dépendante ne sera pas grandement affectée. Au contraire, si l'objectif est de faire des inférences sur les paramètres d'une population, la multicolinéarité sera plus problématique. Le coefficient d'inflation de la variance (VIF) est une mesure populaire de la multicolinéarité. Bien que la multicolinéarité ne soit pas une hypothèse actuelle de la régression multiple, il est important de la prendre en considération.

Les performances d'un modèle statistique dépendent de la méthode utilisée pour estimer les paramètres d'une régression. La plus populaire étant l'estimation par les moindres carrées ordinaires. Cette méthode consiste à minimiser la somme des carrés des écarts des valeurs observées y_i à la droite $\widehat{y_i} = a \cdot x_i + b$. Ce qui revient donc à estimer les valeurs de a et b qui minimisent la somme des carrés des résidus :

$$E = \sum_{i=0}^{n} {\varepsilon_i}^2 = \sum_{i=0}^{n} (y_i - (a \cdot x_i + b))^2 \tag{2.5}$$

Dans certains cas, l'hypothèse de variance constante dans les valeurs résiduelles pour les moindres carrés est contredite (hétéroscédasticité). On applique dans ce cas une pondération

adaptée, pour calculer ce que l'on appelle les estimations par les moindres carrés pondérés. Cette procédure minimise la somme des carrés des valeurs résiduelles pondérés, de manière à générer des valeurs résiduelles présentant une variance constante.

2.3.2.2 Modèle linéaire général

Le modèle linéaire général ou encore modèle de régression multivarié est une extension de la régression linéaire multiple pour une seule variable dépendante, qui peut s'écrire sous la forme suivante :

$$Y = X \cdot B + U \tag{2.6}$$

Y est un vecteur d'une série de mesures de la BSA. X est une matrice des observations des variables indépendantes qui peut être une matrice de régresseurs (qui dépend du modèle statistique choisi, ici la régression multiple), chaque colonne étant un ensemble de mesures d'une des variables indépendantes (mesures anthropométriques). B est une matrice des paramètres à estimer et U une matrice des erreurs (bruit).

L'avantage du modèle linéaire général par rapport au modèle de régression linéaire multiple repose sur son aptitude à fournir une solution pour les équations normales lorsque les variables X_i ne sont pas linéairement indépendantes (multicolinéarité). En effet, lorsque les variables X sont corrélées, la matrice inverse de $X_i{}'X_i$ (nécessaire à la résolution des β_i) n'existe pas (Rawlings, Pantula, & Dickey, 2001). Et la nature des relations existantes entre ces variables et la variable indépendante (BSA) peut en être affectée (M. N. Williams et al., 2013). Le modèle linéaire général permet de contourner ce problème. De plus, le modèle linéaire général possède cette importante propriété d'invariance et permet de mesurer le degré de multicolinéarité grâce au coefficient VIF. Une règle des pouces commune pour l'évaluation des coefficients VIF veut qu'une valeur entre 5 et 10 indique une corrélation des variables qui pourrait être problématique et une valeur supérieure à 10 indiquera la présence de multicolinéarité (Belsley, Kuh, & Welsch, 1980; Rawlings et al., 2001; Akinwande, Dikko, &

Samson, 2015). Néanmoins, le plus important étant d'identifier les groupes de variables avec un VIF élevé.

2.3.2.3 Modèle linéaire généralisé

Le modèle linéaire généralisé est une généralisation flexible du modèle linéaire général qui permet à la variable dépendante Y d'avoir une distribution conditionnelle des erreurs autre que normale (McCullagh, 1984; Lindsey, 2000) et est donc considéré non paramétrique. Le modèle linéaire général est généralisé en lui permettant d'être relié à la variable dépendante par une fonction de lien afin que l'amplitude de la variance de chaque variable indépendante soit fonction de ses valeurs prédites (plutôt que d'admettre que la variable dépendante doit varier linéairement). Si l'on reprend (2.6), le modèle linéaire généralisé est de la forme suivante :

$$E(Y) = \mu = g^{-1}(X \cdot B) \tag{2.7}$$

La variable dépendante Y est assumée suivre une distribution de la famille Exponentielle incluant un large éventail de probabilité de distribution telle que Normal, Binomial, Poisson, et Gamma, entre autres. La moyenne μ de la distribution dépend de l'indépendance des variables X avec $E(Y)$ qui est la valeur prédite de Y.

La variance est typiquement une fonction V de moyenne :

$$Var(Y) = V(\mu) = V(g^{-1}(X \cdot B)) \tag{2.8}$$

V suit une distribution de famille exponentielle ou bien est fonction de sa valeur prédite. Les paramètres inconnus B sont estimés par des techniques d'estimation de maximum de vraisemblance ou quasi maximum ou encore Bayésienne.

g qui est la fonction de liaison, assure la relation entre les variables indépendantes et la moyenne de la fonction de distribution. Il existe plusieurs fonctions de liaison couramment utilisées, leurs choix étant informés par plusieurs considérations. Étant donné la nature

continue positive des données, dans le cadre du présent projet, on se limitera aux distributions suivantes (Tableau 2.6) :

Tableau 2.6 Présentation des différentes distributions et fonction de liaison pour des données continues positives

Distribution	**Support de la distribution**	**Nom de la fonction**	**Fonction de liaison**	**Fonction de moyenne**
Normale	$Y \in R$	Identité	$X \cdot \beta = \mu$	$\mu = X \cdot \beta$
Exponentielle	$Y \in R_+$	Inverse	$X \cdot \beta = \mu^{-1}$	$\mu = (X \cdot \beta)^{-1}$
Gamma				
Gaussienne inversée		Inverse carré	$X \cdot \beta = \mu^{-2}$	$\mu = (X \cdot \beta)^{-1/2}$

Le modèle linéaire généralisé a plus de flexibilité que la régression linéaire et le modèle linéaire général au niveau des hypothèses. Seule une relation linéaire entre la variable dépendante transformée en termes de fonction de liaison et les variables indépendantes est supposée (McCullagh, 1984; Lindsey, 2000).

Cependant la multicolinéarité peut-être, là encore, un problème. Toutefois ce n'est pas la matrice inverse de $X_i{}'X_i$ qui pose problème, mais la matrice $X_i{}'W_iX_i$. W_i étant une matrice diagonale des pondérations. C'est donc la colinéarité entre les prédicteurs pondérés $W_i^{1/2}X_i$ qui affecte directement la condition de la matrice inverse $X_i{}'X_i$. Autrement dit, la multicolinéarité entre les variables indépendantes X_i pourrait ne pas affecter la matrice inverse parce que les pondérations pourraient réduire l'effet de la multicolinéarité. Par conséquent, l'évaluation de la multicolinéarité entre les variables indépendantes n'est pas équivalente à celle de la régression linéaire et le modèle linéaire général. On utilisera plutôt une généralisation du VIF qui ne mesure donc pas la multicolinéarité entre les variables indépendantes, mais analyse la matrice variance-covariance des paramètres estimés (Fox & Monette, 1992).

Clicours.COM

2.3.2.4 Modèle non linéaire

La régression non linéaire consiste à trouver l'équation non linéaire qui minimise l'erreur standard des résidus. La différence entre la régression linéaire et non linéaire est due aux fonctions mathématiques possibles du modèle statistique. N'importe quelle fonction non linéaire $f(x_i'; \theta)$ est possible et n'est pas limité à une fonction linéaire telle que l'équation (2.1). Selon Rawlings et al., (2001) le modèle devient donc le suivant :

$$y_i = f(x_i'; \theta) + \varepsilon_i \tag{2.9}$$

Les coefficients de l'équation sont estimés en utilisant l'algorithme de Levenberg-Marquardt qui permet d'obtenir une solution au problème de minimisation (Rawlings et al., 2001). Cet algorithme est parmi les plus populaires pour la résolution des problèmes des moindres carrés non linéaires.

C'est là toute la complexité de la régression non linéaire, en ce qu'un nombre infini de fonctions est possible. Le choix de la fonction appropriée dépend entièrement de l'expertise du domaine et de la connaissance *a priori*. De nombreux essais et validations seront nécessaires. En conséquence, déterminer l'effet de chaque variable indépendante sur la variable dépendante peut s'avérer moins intuitif en régression non linéaire que pour la régression linéaire.

2.3.3 Problématique de l'étude et importance de la sélection de variables

La sélection de variables (aussi appelé sélection de caractéristiques) est importante aussi bien en statistique, qu'en apprentissage machine. Non seulement cela permet d'améliorer le processus de prédiction, mais également l'estimation de la structure des corrélations en plus d'en offrir une compréhension plus facile des résultats (Guyon & Elisseeff, 2003; Karagiannopoulos, Anyfantis, Kotsiantis, & Pintelas, 2007; Saeys, Inza, & Larrañaga, 2007; De Silva & Leong, 2015). La sélection de variables permet surtout d'améliorer la

généralisation des résultats par la diminution du surapprentissage et donc de réduire la variance (G. James, Witten, Hastie, & Tibshirani, 2013).

En modélisation statistique, la multicolinéarité des données se produit lorsque le modèle inclut des variables explicatives qui ne sont pas juste corrélées à la variable expliquée, mais aussi les unes entre les autres. Contrairement aux situations habituelles où les méthodes de réduction de données sont appliquées (échantillons de données de grandes dimensions ou images médicales), l'échantillon de données de l'étude est de faible dimension (taille d'échantillon $n > p$ variables) et est caractérisé par des corrélations et interactions entre les mesures (multicolinéarité).

Puisque certaines variables peuvent être corrélées, il n'est pas nécessaire de toutes les inclure dans la modélisation statistique. Cependant, il peut y avoir des interdépendances, où deux variables ou plus apporteront plus d'informations ensemble qu'isolées. De plus, différentes variables peuvent embrouiller et masquer les différents facteurs explicatifs de la variation derrière les données. Kohavi et John, (1997) ainsi que Bell et Wang, (2000) identifient trois types de variables :

Pertinentes : variables ayant une influence sur la variable dépendante et dont les interactions ne peuvent être assumées par les autres variables. Cependant, la pertinence d'une variable n'implique pas nécessairement qu'elle fasse partie du sous-ensemble de variables optimales (Kohavi & John, 1997).

Non pertinentes : variables n'ayant aucune influence sur la variable dépendante. Cependant, une variables non pertinente par elle-même peut le devenir en étant combinée à d'autres (Kohavi & John, 1997; Guyon & Elisseeff, 2003; De Silva & Leong, 2015).

Redondantes : variables dont les interactions peuvent être assurées par d'autres variables. Ces variables n'apporteront donc aucune autre information supplémentaire (Guyon & Elisseeff, 2003).

Enfin, comme il a été vu en au chapitre 1.1, la prise en charge d'un grand brûlé s'inscrit dans un contexte clinique d'urgence. Il est donc indispensable de proposer un nombre limité de mesures anthropométriques.

2.3.3.1 Multicolinéarité et son effet sur les modèles linéaires généralisés

Lorsqu'il y a beaucoup de variables corrélées (multicolinéarité), le modèle linéaire généralisé devient instable dans le sens ou de faibles variations dans les données peuvent induire de forts changements dans le modèle (Gunst & Webster, 1975; Slinker & Glantz, 1985) :

- Les coefficients de régression estimés peuvent être élevés en valeur absolue;
- Leurs signes peuvent être contraires à l'intuition;
- Les variances des estimateurs peuvent être élevées;
- Les coefficients de régression et le coefficient de corrélation multiple sont instables par rapport aux coefficients de corrélation entre les variables explicatives.

Par exemple, admettons un ensemble de données avec comme « vrai » modèle $y = x_1 + x_2$, tandis qu'est observé $y = x_1 + x_2 + \varepsilon$. De plus, x_1 et x_2 sont linéairement corrélés de sorte que $x_1 \approx x_2$. Idéalement, le modèle sera $y = x_1 + x_2$. Mais dépendamment de la quantité de bruit, des données et des corrélations entre ces variables, le modèle pourrait également être $y = 2 \cdot x_1$ ou $y = 3 \cdot x_1 - \cdot x_2$.

Cette colinéarité statistique crée donc des difficultés importantes dans l'interprétation des résultats. Par exemple, le fait que le signe d'un coefficient de régression puisse être changé par la colinéarité peut être particulièrement gênant pour étudier l'effet d'une variable sur la variable dépendante (ou expliquée). De plus, les erreurs types des coefficients seront augmentées et des variables seront statistiquement significatives alors qu'elles ne le devraient pas (variables redondantes). Sans multicolinéarité (donc avec des erreurs types plus faibles), ces coefficients pourront être significatifs. Toutefois, la présence de multicolinéarité laisse

présager que plusieurs combinaisons de variables anthropométriques peuvent être intéressantes pour prévoir la surface corporelle.

Identifier au préalable un ensemble restreint de variables est indispensable pour améliorer les résultats obtenus et éviter de modéliser le bruit échantillonnal plutôt que la population générale (Babyak, 2004).

2.3.3.2 Surapprentissage en apprentissage machine

En apprentissage machine, ce n'est pas la multicolinéarité qui pose problème, mais le surapprentissage dû à un trop grand nombre de variables sélectionnées (Kohavi & John, 1997; Bell & Wang, 2000; Guyon & Elisseeff, 2003; Karagiannopoulos et al., 2007; Saeys et al., 2007; Yun & Yang, 2007; Khalid, Khalil, & Nasreen, 2014; De Silva & Leong, 2015).

- Plus il y a de variables, plus de bruit est introduit. Dans ce cas, l'algorithme aura tendance à surapprendre et donner une fausse impression de bonne performance. Réduire le nombre de variables améliore la performance des algorithmes.

- Réduire le nombre de variables permet de réduire la durée de chargement des algorithmes. Par conséquent, il sera donc possible d'utiliser des algorithmes plus complexes et réaliser plusieurs évaluations.

Un ensemble restreint de variables permet une meilleure compréhension des régularités inhérentes aux données. À noter que Liu et Salinas, (2015) ont observé au travers de 15 études l'intérêt grandissant de la sélection de variables en apprentissage machine chez le traitement de patients brûlés. Des performances supérieures ont été constatées pour les approches d'apprentissages machines comparées aux approches statistiques.

2.3.4 Prétraitements de données

2.3.4.1 Corrélation et allure de la distribution des variables

Avant de procéder à toute analyse statistique, il est toujours avisé de dresser le portrait complet des données en observant l'allure de leurs distributions et leurs corrélations. Quand les variables ont une distribution non normale, on préférera utiliser les coefficients de corrélation de rang de Spearman plutôt que ceux de Pearson qui requièrent une distribution bivariée normale des variables (Mukaka, 2012).

Habituellement, on préférera le test de normalité de Shapiro-Wilks. Cependant, ce test est limité à une taille d'échantillon de 5000. On utilisera donc un diagramme Quantile-Quantile ainsi que les coefficients d'aplatissement standardisé et d'asymétrie standardisée qui sont particulièrement utiles pour déterminer avec plus de précision si l'échantillon suit une loi normale. Des valeurs de ces statistiques en dehors de la plage -2 à $+2$ indiquent un écart significatif à la normalité. Ce coefficient est basé sur les rangs des données et il mesure la force de l'association entre deux variables. Ce coefficient est plus robuste aux données extrêmes et n'est pas limité au lien linéaire. Les corrélations sont calculées d'abord en remplaçant les valeurs des données de chaque variable par leurs rangs (allant de 1 à n) puis en calculant le degré de discordance entre chaque rang. Le test de corrélation ne cherche qu'à déterminer l'absence ou la présence d'association significative entre les variables.

Pour l'interprétation des coefficients de corrélation, il est proposé de suivre la règle des pouces de Mukaka, (2012) définie dans le Tableau 2.7 ci-dessous. Pour déterminer si deux variables sont significativement liées l'une à l'autre, une valeur de probabilité est calculée pour chaque coefficient de corrélation. Toute paire de variables pour laquelle la valeur de la probabilité est inférieure à 0,05 indique une association significative au niveau de signification de 5%.

Tableau 2.7 Règle des pouces de Mukaka, (2012), Table 1, p3

Coefficient de corrélation	Interprétation
.90 à 1.00 (−.90 à −1.00)	Très forte association positive (négative)
.70 à .90 (−.70 à −.90)	Forte association positive (négative)
.50 à .70 (−.50 à −.70)	Moyenne association positive (négative)
.30 à .50 (−.30 à −.50)	Faible association positive (négative)
.00 à .30 (.00 à −.30)	Association négligeable

2.3.4.2 Centrage et normalisation

En statistique, on peut également utiliser une transformation par puissance de Box-Cox qui est l'une des procédures les plus optimisées afin d'améliorer la variance. En apprentissage machine, certains algorithmes comme le réseau de neurones (voir chapitre 2.3.9.5) et la machine à vecteurs de support (voir chapitre 2.3.9.4) convergent plus rapidement avec des données redimensionnées (Nawi, Atomi, & Rehman, 2013).

Redimensionner une variable permet de modifier sa moyenne et sa variance. Le redimensionnement $min - max$, entre autres, permet d'avoir un intervalle fixe [0,1]. Ainsi, les variables deviennent donc indépendantes de leurs unités, sans impacter leurs variations d'origines. Cette pratique qui est considérée comme indispensable en apprentissage machine est très utile non seulement pour la convergence des algorithmes, mais également pour réduire la multicolinéarité et aider à l'interprétation des résultats (Dalal & Zickar, 2012).

2.3.5 Les principales approches de sélection de variables en régression

La réduction de données comprend deux approches : l'extraction de caractéristiques (aussi appelé création de caractéristiques) et la sélection de caractéristiques (aussi appelé sélection

de sous-ensembles). Une caractéristique étant un ensemble de mesures anthropométriques influe sur la surface corporelle totale dans le contexte de l'étude.

Les méthodes d'extraction de caractéristiques se basent principalement sur la projection et compression des données, telles que la régression des moindres carrés partiels (Wold, Ruhe, Wold, & Dunn, 1984; Helland, 1990; Tenenhaus, 1998) ou encore l'analyse en composantes principales (Rawlings et al., 2001). Cependant, ces méthodes ne font que remplacer l'ensemble de données par un plus petit nombre de composantes, lesquelles sont une combinaison de variables non corrélées représentant la plus grande partie possible de l'ensemble des données originales. Utiliser ces méthodes pourrait résoudre le problème de multicolinéarité, mais ne permettrait pas de répondre à l'objectif de l'étude qui est de déterminer un ensemble de mesures anthropométriques réduit qui explique et prédit le mieux la surface corporelle. De plus, on sait que ces méthodes de réductions de données sont insuffisantes lorsque la relation entre les paramètres (mesures anthropométriques) et la surface corporelle est plus complexe (Meunier, Shu, & Xi, 2009).

L'approche de sélection de caractéristiques consiste à sélectionner un sous-ensemble des caractéristiques pertinentes pour la construction d'un modèle de régression. Outre la diminution de la multicolinéarité et l'amélioration de la performance, le principal avantage de cette approche est qu'aucune information liée aux mesures anthropométriques n'est perdue (Saeys et al., 2007). Ce qui n'est pas le cas de l'extraction de caractéristiques, puisque les données sont transformées. De plus, préserver la sémantique des données facilite l'interprétation des résultats par les experts (Saeys et al., 2007). Par la suite, on appellera la sélection de caractéristiques, sélection de variables pour éviter toute confusion par rapport au contexte de l'étude.

Il y a beaucoup de méthodes de sélection de variables aussi bien en statistique qu'en apprentissage machine. Pour un ensemble de $\boldsymbol{N}$ variables, le problème de sélection de variables est de trouver le sous-ensemble le plus optimal parmi les $\mathbf{2}^{\boldsymbol{N}}$ possibles variables (De Silva & Leong, 2015). Selon Guyon et Elisseeff, (2003), Saeys et al., (2007) ainsi que De Silva et

Leong, (2015) l'ensemble de ces méthodes peuvent être regroupé en trois catégories : les méthodes *filters* et *wrappers* d'abord introduits par John, Kohavi, et Pfleger, (1994) et les méthodes intégrées dites *embedded* (ou encore *built-in*). La première (*filters*) attribue un poids à chaque variable, tandis que les deux dernières (*wrappers* et *embedded*) évaluent les combinaisons de variables différentes De Silva et Leong, (2015). Kohavi et John, (1997) et Guyon et Elisseeff, (2003) ont souligné le contraste existant entre une sélection de variables utiles (cas des méthodes *wrappers* et *embedded*) et la problématique de trouver et de classer toutes les variables potentielles (cas des méthodes *filters*). Sélectionner les variables les plus pertinentes est souvent sous optimal, en particulier si les variables sont redondantes. Inversement, un sous-ensemble de variables utiles exclut beaucoup de variables redondantes certes, mais aussi des variables pertinentes.

Les méthodes *Filter* qui sélectionnent les variables en les classant selon leurs poids attribués sont généralement utilisées en tant que pré-analyse simple et rapide des données afin d'observer/évaluer les corrélations et dépendances individuelles des variables (Guyon & Elisseeff, 2003; Karagiannopoulos et al., 2007). Par conséquent, elles ignorent la multicolinéarité. C'est le cas des méthodes statistiques univarié tel que la corrélation (Pearson/ Spearman), la Cohérence maximale et redondance minimale (*mRMR*), etc. qui attribuent un score individuel à chaque variable. La sélection de variables est indépendante de toutes analyses de régression statistique ou tout algorithme d'apprentissage machine (voir figure 2.5 ci-dessous).

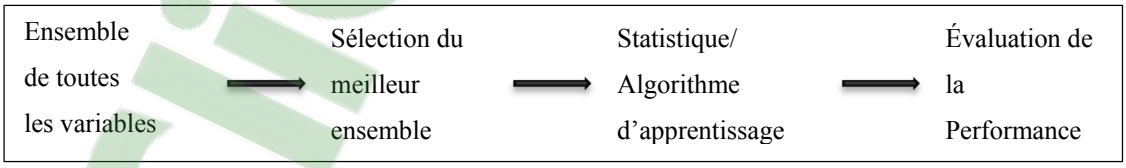


Figure 2.5 Schéma de fonctionnement des méthodes *Filter*

Les méthodes *filter* vont plutôt sélectionner les propriétés intrinsèques des variables et donc leurs pertinences. Une critique majeure de cette approche est qu'elle conduit à la sélection d'un

sous-ensemble de variables redondantes. La même performance peut être atteinte avec des variables complémentaires. Les méthodes *filters* ont donc peu de chance de déterminer les combinaisons de variables donnant les meilleures performances (Kohavi & John, 1997; Guyon & Elisseeff, 2003).

Les méthodes dites *wrappers* considèrent la sélection comme un problème de recherche où différentes combinaisons de variables utiles sont préparées, évaluées et comparées à d'autres combinaisons de variables utiles. Un modèle de régression prédictif est construit afin d'évaluer la combinaison de variables et lui attribuer un score basé sur sa précision (voir Figure 2.6 ci-dessous). Cependant, ces méthodes ont plus de risques de surapprentissage et sont des procédures très lourdes (méthodes itératives) (Kohavi & John, 1997; Karagiannopoulos et al., 2007) et sont indépendantes de l'apprentissage (De Silva & Leong, 2015). C'est le cas des analyses de régression pas-à-pas (*forward/ backward*), régression exhaustive, séquentielle ou par élimination récursive.

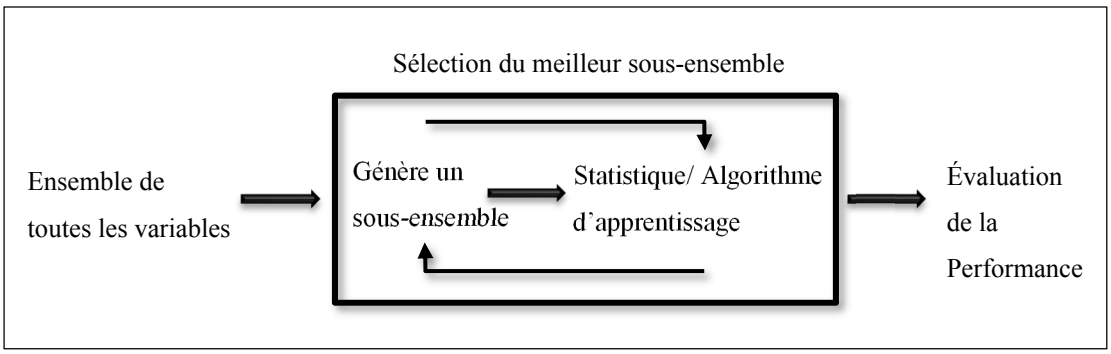


Figure 2.6 Schéma de fonctionnement des méthodes *Wrapper*

Dans bien des cas, utiliser une méthode de sélection de variables *Embedded* sera plus efficace que les méthodes *wrappers*. En effet, bien que similaire, la sélection de variables est directement optimisée par une fonction objective comprenant un terme pour la qualité d'ajustement et une pénalité pour le nombre de variables (voir Figure 2.7 ci-dessous) et est donc, dépendante de l'apprentissage (De Silva & Leong, 2015). Le Tableau 2.8 ci-après présente un résumé des inconvénients et avantages de chaque méthode.

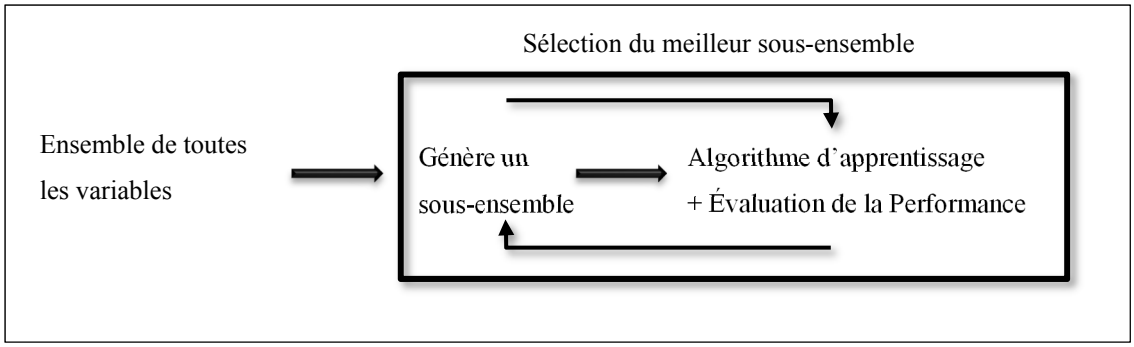


Figure 2.7 Schéma de fonctionnement des méthodes *Embedded*

Tableau 2.8 Avantages (+) et inconvénients (-) des différentes méthodes de sélection de variables

Méthodes	Critère	Recherche	Évaluation	Résultats	
« Filters »	**Pertinence** d'une variable / sous-ensembles variables	Classe par ordre d'importance	Tests statistiques	Robuste surapprentissage (+) Interactions variables ignorées (-) Échoue à sélectionner les bonnes variables les plus utiles (-)	
« Wrappers »	**Utilité** des sous-ensembles variables	Espace de toutes les variables	Validation croisée	Trouve les variables les plus utiles (+)	Tendance à surapprendre (-) Temps de calcul élevés (-)
« Embedded »		Guidé par le processus d'apprentissage			Moins coûteux en ressources (+) Moins sensible au surapprentissage (+)

Ces méthodes utilisent l'information du modèle construit afin d'inclure la structure des corrélations des variables dans le calcul de leurs importances. C'est le cas de la régression *LASSO*, *Elastic Net* ou des algorithmes d'apprentissage machine présentés au chapitre 2.3.9.

2.3.6 Comparaison entre modèles linéaires et modèles d'apprentissage machine

Un résumé des avantages et inconvénients de chacune des deux cultures sont répertoriés dans le Tableau 2.9 ci-après sur la base des études suivantes: (McCullagh, 1984; Wold et al., 1984; Lesaffre & Marx, 1993; Lindsey, 2000; Breiman, 2001b; Anderson et al., 2005; Alpaydin, 2014).

Clicours.com

Tableau 2.9 Avantages (+) et inconvénients (-) des GLMs et apprentissage machine

	Modèle linéaire général/ généralisé	**Apprentissage Machine**
Distribution conditionnelle des erreurs	Supposés être Gaussienne/ Exponentielle	Aucune hypothèse requise (+)
Relation linéaire des paramètres / Moyenne conditionnelle nulle des erreurs	Supposés être linéaire (-)	Aucune hypothèse requise (+)
Indépendances des erreurs	Supposés être non corrélés (-)	Aucune hypothèse requise (+)
Homoscédasticité des erreurs	Supposé être constante (-) / non (+)	Aucune hypothèse requise (+)
Multicolinéarité	Très sensible (-)	Peu sensible (+)
Surapprentissage	Quelques risques (+)	Risques significatifs (-)
Interprétation	Facile (+)	Difficile (-)
Ressource de l'ordinateur	Faible (+)	Peut-être lourd (-)
Précision	Bonne (+)	Bonne (+)

Le consensus des différents auteurs est que les modèles linéaires générale/ généralisé sont pratiques, faciles à implémenter et à interpréter, mais pas aussi robuste que l'apprentissage machine puisqu'ils reposent sur des hypothèses théoriques. Concernant les modèles de prédiction, par apprentissage machine, leurs principaux avantages sont leurs flexibilités (puisque programmables), le fait qu'ils ne dépendent d'aucune hypothèse statistique et leurs capacités à capturer à la fois des relations linéaires et non linéaires. Bien qu'une bonne interprétation conduit typiquement à une meilleure généralisation des résultats, c'est plutôt la combinaison de solides données (représentatif de la population, traitées, etc.) du modèle prédictif et de la compréhension du problème qui est nécessaire afin d'avoir une solution qui fonctionne mieux.

En définitive, on retiendra la suggestion de Guyon et Elisseeff, (2003) et Morozova, Levina, Uusküla, et Heimer, (2015) d'appliquer plusieurs méthodes de sélection de sous-ensembles et d'identifier les mesures communes à chacune des méthodes. Cette approche est surtout utile pour l'analyse de données, spécifiquement en situation d'incertitudes quant au choix du bon modèle prédictif, mais aussi pour améliorer la crédibilité des résultats (Morozova et al., 2015).

2.3.7 Sélection de sous-ensembles de modèles linéaires

En statistique, plusieurs classes de méthodes de sélection de variables existent. La plus connue étant la sélection de sous-ensembles de variables. Cette approche identifie un ensemble de variables indépendantes X_i pertinentes reliées à la variable dépendante y. Le modèle prédictif est ensuite ajusté par la méthode des moindres carrées ordinaires afin de réduire l'ensemble de variables.

À noter que l'estimation des moindres carrés ordinaires aura une faible variance et donc une bonne performance lorsque $n \gg p$; n étant le nombre d'observations et p le nombre de variables. Si $n \approx p$, il peut y avoir beaucoup de variabilité dans l'ajustement des moindres carrés et donc une faible prédiction. Tandis que lorsque $n < p$, il n'existe plus de solutions uniques pour les coefficients β, les approches de sélection de sous-ensembles ne peuvent donc plus être utilisées (G. James et al., 2013). Bien qu'au prix d'une augmentation négligeable du biais, des méthodes de réductions peuvent être également utilisées afin de réduire la variance et ainsi améliorer la précision du modèle prédictif (G. James et al., 2013).

2.3.7.1 Sélection de sous-ensembles

La sélection de sous-ensembles inclut deux grandes procédures, la régression pas-à-pas (*stepwise*) ascendante ou descendante et la sélection des meilleurs ensembles (*best subset selection*).

A. Régression pas-à-pas

La méthode ascendante commence avec 0 mesure dans le modèle. Ensuite X_i mesures sont ajoutées une par une en suivant un critère de sélection : améliorer le R^2 et le R^2 ajusté (voir section 2.5.1). À chaque itération, le modèle est recalculé jusqu'à ce que le critère de sélection soit satisfait. La méthode descendante commence avec l'ensemble des X_i mesures sélectionnées pour le modèle. Puis, X_j sont enlevées une par une du modèle en suivant le même

critère de sélection que précédemment. Le principal avantage de la régression pas à pas est la simplicité d'un seul modèle de régression. Cependant, bien que la régression pas à pas soit l'une des méthodes les plus souvent utilisées (Derksen & Keselman, 1992; Harrell, 2001; Flom & Cassell, 2007); Morozova et al., (2015) remettent en question son efficacité lorsque le nombre de variables corrélées est élevé.

La plupart du temps, la régression pas-à-pas n'identifie pas le bon modèle statistique, particulièrement en présence de multicolinéarité. De plus, les erreurs standards, valeurs de p et coefficients de régression sont biaisés, ce qui rend difficile l'interprétation de l'effet individuel de chacune des variables sur la variable dépendante. Plutôt d'utiliser une analyse contestable et erronée, Flom et Cassell, (2007) suggère l'analyse des meilleurs ensembles ou des méthodes de réductions.

B. Meilleurs ensembles

Cette analyse statistique considère tous les modèles de régression des moindres carrés ordinaires contenant toutes les combinaisons de p variables anthropométriques pertinentes et sélectionne les modèles en fonction des meilleures valeurs des critères MSE, R^2, R^2 ajusté et BIC (voir section 2.5.1). Toutefois, cette analyse ne choisit pas le modèle statistique final. Plusieurs meilleurs modèles statistiques sont présentés selon différents critères afin d'aider dans le choix du modèle approprié.

2.3.7.2 Méthodes de réduction

Pour surmonter les inconvénients de la régression pas à pas, la régression Ridge, tout comme le LASSO et l'Elastic-net ont été introduits. Ces méthodes de réduction des coefficients de régression permettent une meilleure généralisation des données (G. James et al., 2013).

A. Régression Ridge

Si l'on reprend les équations (2.1) et (2.2) en chapitre 2.3.2, le meilleur modèle de régression est déterminé par l'estimation des moindres carrés des résidus $min \sum_{i=1}^{n} \varepsilon_i$. L'estimation des

moindres carrés des résidus de la régression Ridge inclut toutes les variables indépendantes X_i (contrairement à la régression LASSO) et utilise la pénalisation de la norme l_2 (racine carrée de la somme des carrés) afin de contracter la valeur des coefficients β_i.

$$\min \frac{1}{2}\sum_{i=1}^{n}\left(y_i - \beta_0 - \sum_{j=1}^{p}\beta_j x_{i,j}\right)^2 + C\sum_{j=1}^{m}\left|{\beta_i}^2\right| \tag{2.10}$$

Peu sensible à la multicolinéarité, la régression Ridge à l'avantage de réduire le surapprentissage et est facile d'interprétation. Cependant, il n'y a pas de sélection parcimonieuse. La régression Ridge ne permet donc pas de sélectionner un sous-ensemble restreint de variables.

B. Régression LASSO

Pour surmonter l'inconvénient de la régression Ridge, la méthode LASSO (*Least Absolute Shrinkage and Selection Operator*), estime les moindres carrés des résidus par la pénalisation de la norme l_1 (valeurs absolues) afin de contracter la valeur des coefficients β_i. Cette réduction, aussi appelée régularisation, peut également réaliser une sélection de variables en réduisant certains coefficients $\beta_i = 0$ (contrairement à la régression Ridge). Le LASSO effectue donc à la fois une sélection de variables et une réduction afin d'améliorer les performances et l'interprétation des modèles statistiques.

$$\min \frac{1}{2}\sum_{i=1}^{n}\left(y_i - \beta_0 - \sum_{j=1}^{p}\beta_j x_{i,j}\right)^2 + C\sum_{j=1}^{m}|\beta_i| \tag{2.11}$$

Contrairement à la régression Ridge, la sélection est parcimonieuse puisque le LASSO permet de sélectionner un sous-ensemble restreint de variables dépendamment du paramètre C. Cette méthode réduit également le surapprentissage et est facile d'interprétation et meilleur que la régression Ridge (puisque moins de variables). Malheureusement cette méthode est sensible à

la multicolinéarité. S'il y a plusieurs variables pertinentes, le LASSO ne sélectionne généralement qu'une seule variable du groupe (Zhao & Yu, 2006).

C. Régression Elastic-net

Afin de surmonter la sensibilité du LASSO à la multicolinéarité et le manque de parcimonie de la régression Ridge, l'Elastic-net a été introduit par Zou et Hastie, (2005) et est une combinaison de la régression LASSO et Ridge : une pénalité Ridge est ajoutée au LASSO :

$$\min \frac{1}{2}\sum_{i=1}^{n}\left(y_i - \beta_0 - \sum_{j=1}^{p}\beta_j x_{i,j}\right)^2 + C_2\sum_{j=1}^{m}|\beta_i{}^2| + C_1\sum_{j=1}^{m}|\beta_i| \qquad (2.12)$$

Le but étant d'abord de trouver les coefficients Ridge avec une valeur de C_2 fixe, puis d'appliquer la réduction LASSO. Les coefficients sont ensuite multipliés par $(1 + C_2)$ afin d'améliorer la performance de prédiction (Zou & Hastie, 2005).

2.3.8 L'apprentissage machine : application à la régression

Alpaydin, (2014) définit l'apprentissage machine comme la programmation d'ordinateurs pour optimiser un critère de performance en utilisant des données ou une connaissance *a priori*. Plus spécifiquement, un modèle algorithmique est défini selon certains paramètres, et l'apprentissage est l'exécution du programme ordinateur pour optimiser les paramètres du modèle en utilisant les données ou une connaissance *a priori*. Étant donné la tâche principale qui est de faire des inférences à partir d'un ensemble, l'apprentissage machine utilise les théories statistiques dans la construction mathématique du modèle algorithmique. Là encore, le modèle peut être prédictif afin de prédire le futur ou bien descriptif afin d'acquérir des connaissances à partir des données ou les deux, comme c'est le cas du présent projet de recherche. En apprentissage machine, l'induction réfère aux concepts généraux d'apprentissage à partir d'exemples spécifiques. Trois cas se distinguent: l'apprentissage supervisé, non supervisé et renforcé. Dépendamment d'un échantillon de données incluant la

sortie qu'on souhaite prédire; l'apprentissage supervisé consiste à apprendre une fonction qui prédit correctement cette sortie à partir des entrées et qui peut être utilisée pour faire des prédictions sur d'autres exemples. C'est le cas de la classification et de la régression.

L'apprentissage non supervisé est différent en ce qu'il utilise un échantillon de données ou il n'y a aucune sortie à prédire. Dans ce cas, on veut apprendre un modèle qui pourrait avoir généré cet échantillon. C'est le cas des méthodes Clustering, estimation de densité ou encore réduction de bruit. Enfin, l'apprentissage renforcé est un problème de décision séquentiel, prendre des décisions au moment présent influencera celles du moment futur. Une récompense est donnée afin d'indiquer celles ayant un bon état.

La régression en apprentissage machine est un donc un problème d'apprentissage supervisé. Théoriquement, le problème est définit comme tel : étant donné un échantillon d'apprentissage de N exemples de paires d'entrées-sorties $(x_1, y_1), \cdots, (x_N, y_N)$ ou chaque y_i est généré par une fonction inconnue $y = f(x)$, trouve les hypothèses h tel que $h \approx f$. Par la suite, on désignera l'espace des hypothèses H, un ensemble de fonctions que l'algorithme d'apprentissage considère comme approximations de la vraie fonction $f()$. La méthode d'apprentissage inductive consiste donc à construire/ ajuster h de manière à ce qu'elle concorde avec f sur l'échantillon d'apprentissage (h est cohérente si elle concorde avec f pour chaque exemple).

2.3.8.1 Sélection de modèles et généralisation

L'apprentissage est un problème mal posé puisque les données à elles seules ne suffisent pas à trouver une solution unique (Alpaydin, 2014). L'ensemble des hypothèses H faites pour rendre l'apprentissage possible est appelé biais inductif de l'algorithme d'apprentissage. Par exemple, admettre une fonction linéaire ou non linéaire en régression est un biais inductif et le plus important, choisir celle qui minimise l'erreur quadratique est un autre biais inductif. L'apprentissage est donc impossible sans biais inductif.

La sélection de modèles permet de choisir parmi les *H* possibles, et donc de choisir le biais approprié. Cette étape est également importante puisque si le biais inductif contraint la recherche seulement à des hypothèses simples, alors la vraie fonction pourrait être exclue de l'ensemble des hypothèses *H*. Dans ce cas, la vraie fonction est irréalisable dans l'espace des hypothèses choisies. On parle donc de sous-apprentissage : le modèle logarithmique est non adapté, il ne peut ni modéliser l'échantillon d'apprentissage ni généraliser à de nouvelles données. Au contraire, si le biais inductif permet la recherche d'hypothèses trop complexes, alors le modèle logarithmique pourrait ajuster avec des facteurs non pertinents de l'échantillon d'apprentissage. C'est ce qu'on appelle le surapprentissage: le modèle logarithmique apprend les détails et bruits de l'échantillon d'apprentissage au détriment des impacts négatifs sur la performance du modèle à de nouvelles données.

On peut donc dire qu'il y a un compromis à faire entre la complexité des hypothèses qui correspondent mieux à l'échantillon d'apprentissage et celles, plus simples, qui se généralisent mieux. Dietterich, (2003) identifie ces problématiques qu'il définit comme étant les trois compromis (*triple trade-off*) qui sont d'importantes considérations pour tout algorithme d'apprentissage machine supervisé : la complexité de l'ensemble des hypothèses *H* (biais inductif), la taille de l'échantillon d'apprentissage et la généralisation des erreurs à de nouvelles données. En effet, à mesure que la taille de l'échantillon d'apprentissage augmente, la généralisation des erreurs diminue. De plus, comme le montre la Figure 2.8 ci-dessous, à mesure que la complexité de l'ensemble des hypothèses *H* augmente, la généralisation des erreurs diminue puis commence à augmenter.

On appelle généralisation, un modèle logarithmique qui, ayant appris avec l'échantillon d'apprentissage, prédit la bonne sortie pour de nouvelles instances. Pour une meilleure généralisation, la complexité de l'ensemble des hypothèses *H* doit correspondre à celle de la fonction sous-jacente aux données. Dans le cas contraire, on tombe dans le cas de sous-apprentissage ou surapprentissage tel que discuté précédemment (voir également Figure 2.8 ci-dessous). Le surapprentissage et sous-apprentissage sont deux causes responsables de la faible performance des algorithmes d'apprentissage machine.

La capacité de généralisation d'une hypothèse telle que $h \approx f$, soit la qualité du biais inductif, est mesurée par des données autres que celui de l'échantillon d'apprentissage. Pour cela, l'échantillon d'apprentissage est divisé en deux sous échantillons. Le premier est utilisé pour l'apprentissage (pour ajuster des hypothèses), et le deuxième appelé échantillon de validation, est utilisé pour tester la capacité de généralisation. Pour un ensemble d'hypothèses données H_i, les meilleures hypothèses $h_i \in H_i$ sont ajustées. Celle étant la plus précise sur l'échantillon de validation est la meilleure (celle ayant le meilleur biais inductif). Ce processus d'échantillonnage est appelé validation croisée, le plus populaire étant la validation croisée *k-fold.*

Dans la validation croisée *k-fold*, l'échantillon de données d'origine est aléatoirement partitionné en k sous échantillons de taille égale. Parmi les k sous échantillons, un seul est retenu pour l'échantillon de validation afin d'ajuster les hypothèses et l'erreur quadratique moyenne (*MSE*), qui est souvent utilisée comme mesure de validation. Les $k - 1$ sous échantillons restants sont utilisés comme échantillon d'apprentissage. Une validation croisée avec $k = 10$ est souvent un bon compromis entre biais et variance (Kohavi, 1995; Borra & Di Ciaccio, 2010; G. James et al., 2013). Il existe également la validation croisée *Leave-one-out* (LOOCV) qui est un cas spécial de la validation *k-fold* avec $k = n$. Étant plus couteuse en ressource informatique, on préféra la validation *k-fold* à la LOOCV. Cependant, pour estimer l'erreur d'approximation du meilleur modèle algorithmique ajusté retenu à de nouvelles données, un troisième échantillon est nécessaire, appelé échantillon de test. En effet, l'échantillon de validation ne peut être utilisé puisqu'il devient partie intégrante de l'apprentissage.

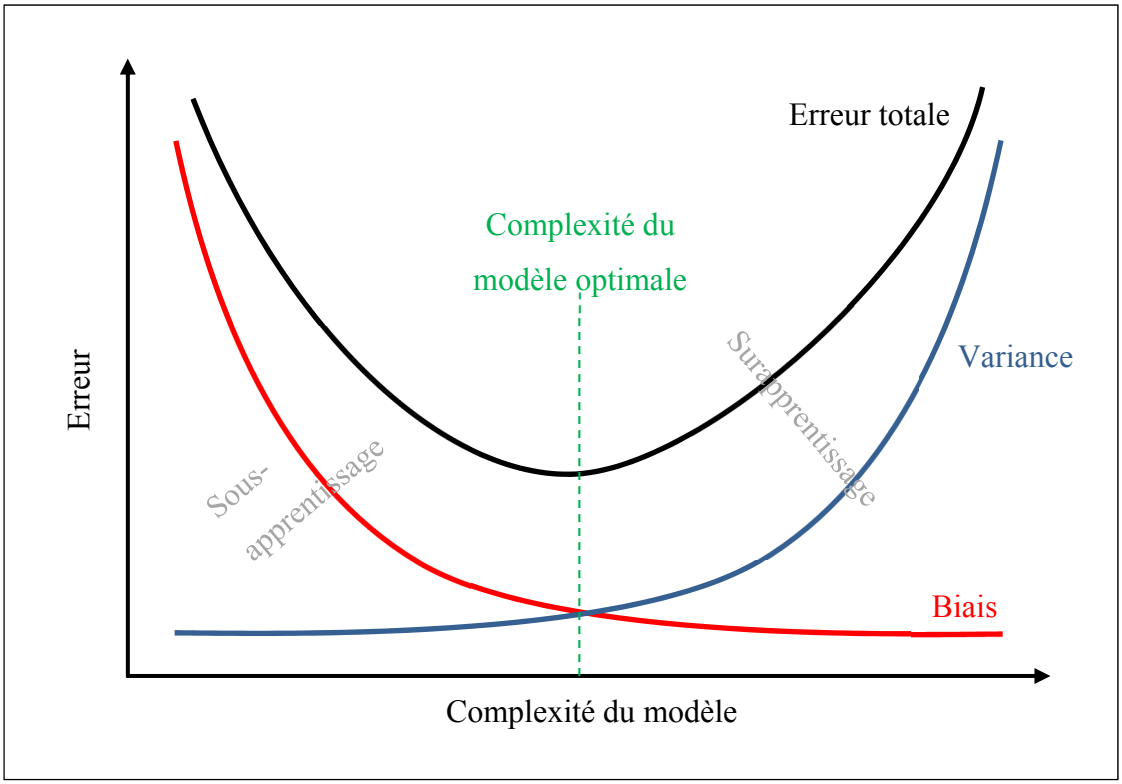


Figure 2.8 Compromis biais-variance pour un modèle optimal en apprentissage machine

En conclusion, l'apprentissage machine inductif implique de chercher à travers l'espace des hypothèses H une fonction qui fonctionnera tout aussi bien pour l'échantillon d'apprentissage que pour de nouveaux exemples en dehors de celui-ci. Trois étapes sont nécessaires à la construction d'un modèle (instance d'une hypothèse $h \in H$) :

Processus itératif

1. Construction du modèle (biais inductif) → échantillon d'apprentissage
2. Test du modèle (sélection du modèle : évaluer le biais) → échantillon de validation
3. Application du modèle à de nouveaux exemples → échantillon de test

En plus des trois compromis, il est également important de considérer la complexité de la fonction à modéliser et donc de la quantité de données nécessaire. En effet, dans le présent projet de recherche, la fonction à modéliser est complexe (surface corporelle en fonction de mesures anthropométriques), puisqu'elle implique de nombreuses interactions entre

différentes variables indépendantes. Une telle complexité ne peut-être apprise qu'à partir d'une quantité de données suffisamment large.

2.3.8.2 Algorithmes paramétriques et non paramétriques

En régression, l'algorithme cherche à approximer la fonction de modélisation sous-jacente inconnue pour une sortie Y (la BSA) selon une entrée X (les mesures anthropométriques) dépendamment d'un échantillon de données. On admet le modèle défini selon un ensemble de paramètres :

$$Y = g(X|\theta) \tag{2.13}$$

avec $g(.)$ une fonction de régression qui définit l'ensemble des hypothèses H, et une valeur particulière de θ est une instance d'une hypothèse $h \in H$. Une fonction de perte $L(\cdot)$ calcul la différence entre la sortie désirée Y et son approximation $g(X|\theta)$ pour chaque valeur des paramètres θ. L'erreur d'approximation ou perte est la somme des pertes pour chaque instance :

$$E(\theta|X) = \sum_t L(Y^t, g(X^t|\theta) \tag{2.14}$$

t étant l'index des N instances.

La fonction de perte souvent utilisée en régression est la racine quadratique moyenne. La procédure d'optimisation consiste à trouver θ' qui minimise l'erreur totale :

$$\theta' = \arg\min_\theta E(\theta|X) \tag{2.15}$$

Avec $\arg min$ qui retourne l'argument minimisant le plus la perte.

Chaque algorithme d'apprentissage machine est différent dans le choix de leurs hypothèses/ biais inductif, mesure de perte et procédure d'optimisation. Leur fonctionnement peut être divisé en quatre catégories.

clicours.com

- Approches basées sur la similarité/ analogie : trouve un problème dont on connaît la solution qui soit similaire au problème actuel. Cas des algorithmes des *k* plus proches voisins (*k-nearest neighbor*).
- Algorithmes basés sur l'information : trouve parmi les questions posées celle qui donnera le plus d'informations dans la résolution du problème. Cas des arbres décisionnels (*Decision Tree*).
- Algorithmes basés sur la probabilité : trouve parmi les solutions possibles celle qui est la plus probable compte tenu d'une preuve. Cas des naïves bayésiennes (*Naive Bayes*).
- Approches de réduction d'erreur : préparent une solution (hypothèses), puis la changent itérativement de manière à réduire l'erreur dans l'approximation des exemples donnés. Cas de la régression, réseaux de Neurones (*NN*), Machine à support de vecteurs (*SVM*).

Là encore, deux grandes catégories d'algorithmes se distinguent dans la recherche de la fonction de régression inconnue g : les algorithmes dits paramétriques et ceux dits non paramétriques (G. James et al., 2013; Alpaydin, 2014). Les algorithmes paramétriques admettent une fonction g définie et un nombre de paramètres fixent permettant d'estimer les coefficients de cette fonction. C'est le cas des réseaux de Neurones (*NN*) et de la régression par splines adaptatives multivariées (*MARS*). Bien que définir une forme de la fonction g simplifie la compréhension et l'interprétation des résultats et améliore l'apprentissage ; l'algorithme est restreint à la forme de g spécifiée et peut être limité dans la modélisation d'un problème plus complexe. Au contraire, les algorithmes non paramétriques n'admettent aucune fonction g et sont donc libres d'apprendre n'importe quelle forme de fonctions à partir de l'échantillon d'apprentissage. C'est le cas des arbres décisionnels (*Decision Tree*) et machine à support de vecteurs (*SVM*). Ces algorithmes sont donc plus performants que les algorithmes paramétriques. Cependant, leur apprentissage est plus long, plus sensible au surapprentissage et ils sont plus difficiles à interpréter.

2.3.9 Algorithmes de sélection de variables non linéaires

En apprentissage machine, le théorème « no free lunch » nous dit qu'aucun algorithme n'est plus performant en moyenne qu'une recherche à l'aveugle (Ho, 1999). Il existe de nombreux algorithmes d'apprentissage machine différents (voir carte heuristique en APPENDICES), dont la performance dépendra de la structure et de la taille des données. Bien que le choix de l'algorithme approprié reste donc souvent incertain, chacun a ses avantages et inconvénients. Ainsi, il est possible d'en sélectionner plusieurs algorithmes appropriés au contexte d'étude. Dans le contexte de sélection de variables continues en régression, certains auteurs préconisent l'utilisation de méthodes *Embedded* couplées à une sélection de variables selon des méthodes *filter* ou *wrapper* pour de meilleurs résultats.

Parmi celles-ci, on retrouve la machine à vecteurs de support, soit avec une élimination de variables récursives, plus connu sous l'appellation *SVM-RFE* (Guyon & Elisseeff, 2003; De Silva & Leong, 2015), soit avec la sélection de variables *Backward ─ SVM-BFS* ou *Forward ─ SVM-FFS* (Guyon & Elisseeff, 2003; Karagiannopoulos et al., 2007). Bien que plutôt appliquées au problème de classification, elles ont également été appliquées au problème de régression plus connu sous l'appellation régression à vecteurs de support ─ *SVR* (Smola & Schölkopf, 2004; Guyon & Elisseeff, 2006; Zhu & Wang, 2010; Fang & Tai, 2011; Were, Bui, Dick, & Singh, 2015). À noter que la machine à vecteurs de support ne fait aucune différence entre des données discrètes ou continues (K. Chan et al., 2002).

Plusieurs auteurs (Guyon & Elisseeff, 2006; Karagiannopoulos et al., 2007; De Silva & Leong, 2015; Were et al., 2015) suggèrent également les approches d'apprentissage d'ensemble tel que l'arbre décisionnel ou encore les forêts d'arbres décisionnels ou Boosting d'arbres de régression. Ces approches combinent plusieurs modèles prédictifs et ont un bon compromis entre biais et variance (*triple trade-off*, voir Figure 2.8). Ces approches ont donc une performance prédictive souvent meilleure qu'un modèle prédictif individuel (Mendes-Moreira, Soares, Jorge, & Sousa, 2012). Guyon et Elisseeff, (2006); Were et al., (2015) abordent l'utilisation du réseau de neurones pour la sélection de variables continues en régression non

linéaire. En effet, Bocco, Willington, et Arias, (2010)b ainsi que Wang, Wu et al. (2011) le démontrent plus performant que la régression multiple. Enfin, la régression multiple par Splines adaptatives (*MARS*), très connue, est un excellent algorithme pour la sélection de variables (qui est automatique) dans de nombreux domaines pour la modélisation de la relation linéaire des variables, mais aussi des fonctions sous-jacentes non linéaires et est robuste aux variables non pertinentes (Frank, 1995; Prasad, Iverson, & Liaw, 2006; Taylan, Weber, & Yerlikaya, 2008; Osei-Bryson, 2014; Kumar, Sukanya, & BIT-Campus, 2016). Ces algorithmes ont donc été retenus pour leurs pertinences pour le présent projet de recherche.

2.3.9.1 Arbre décisionnel de régression – CART

Breiman, Friedman, Stone, et Olshen, (1984) sont les créateurs de l'arbre décisionnel de Classification et de Régression (*Classification And Regression Tree - CART*) et constituera la base de l'algorithme de forêt d'arbre décisionnel (*random forest*), de boosting d'arbres de régression (*stochastic gradient boosting trees*) et de la régression par Splines adaptatives multivariées (*MARS*). L'arbre décisionnel de classification se différencie de celui de la régression non seulement par la nature de sa variable de sortie prédite, mais également par la manière dont la division s'opère dans l'arbre décisionnel.

L'apprentissage par arbre décisionnel de régression consiste à diviser (ou partitionner) en différentes régions les variables indépendantes de manière à ce que la variable dépendante (aussi appelé variable cible en apprentissage machine) puisse être mieux prédite. Tel que sur la Figure 2.9, un arbre décisionnel se définit par un nœud racine qui sera divisé, lorsqu'une variable est partitionnée, en deux nœuds « enfants » (appelé division binaire). Un nœud qui n'a pas de sous-branches est appelé nœud terminal. La valeur prédite d'un modèle de régression *CART* est la moyenne de la variable cible *Y* pour les valeurs qui tombent dans l'un des nœuds terminaux.

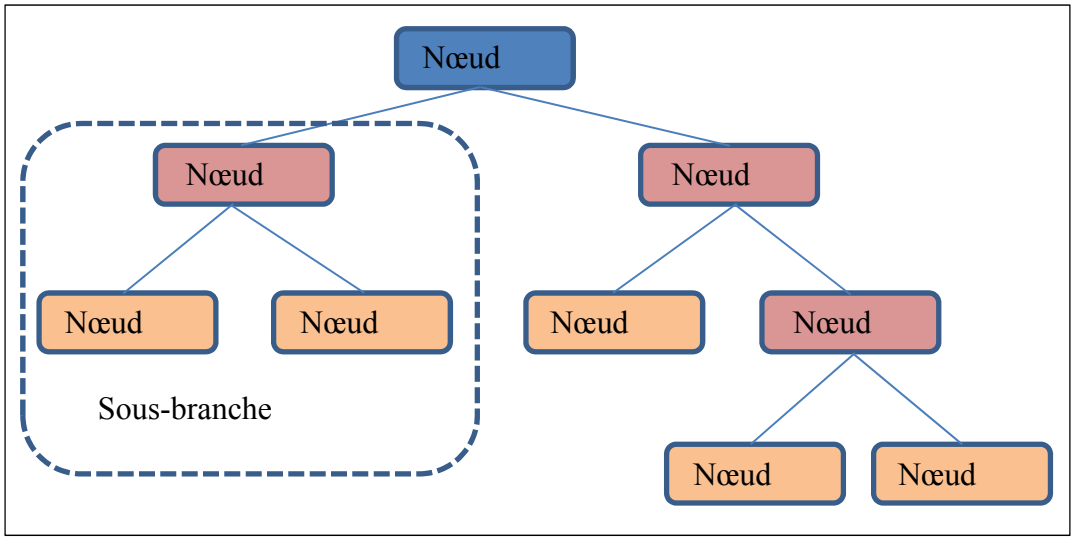


Figure 2.9 Schéma de fonctionnement d'un arbre décisionnel

L'algorithme *CART* comprend deux étapes automatiques: faire pousser un grand arbre, puis l'élaguer. D'abord, toutes les variables sont considérées à chaque division de l'arbre. Chaque nœud est réalisé en utilisant seulement une seule variable et une ou plusieurs valeurs spécifiques. Les divisions sont opérées de manière à minimiser l'erreur du modèle de régression. L'arbre pousse jusqu'à satisfaire un critère définit ou jusqu'à sa limite. Ensuite, les sous-arbres sont élagués par validation croisée afin de choisir le meilleur sous-arbre ayant la plus petite erreur. *CART* présente ainsi l'avantage de gérer automatiquement la sélection de variables puisque seules celles réduisant le plus l'erreur seront retenues. L'importance des variables est déterminée selon qu'une variable est utilisée pour la division d'un nœud (et donc, apparais dans l'arbre) ou lorsqu'elle est la division de substitution (lors d'une valeur manquante). Est également automatique, la modélisation des interactions, des effets locaux et interactions non linéaires (par des fonctions échelons), en plus d'être plus robuste aux points extrêmes en raison de la nature du processus de division (Breiman et al., 1984).

Avantages :

- ✓ Plus robuste aux interactions complexes des variables et non-linéarités. Souvent meilleur que la régression linéaire;

- ✓ Sélection automatique des variables pertinentes et robustes face aux variables redondantes (pas besoin de pré-traitements) → Très utile en exploration de données;
- ✓ Stable et robuste face aux données aberrantes ou manquantes;
- ✓ Facile d'interprétation, même par rapport au modèle de régression linéaire.

Inconvénients :

- × Tendance à surapprendre. Peut cependant être allégé par élagage approprié;
- × Performances souvent moins bonnes par rapport aux autres méthodes ci-dessous.

2.3.9.2 Forêt d'arbres décisionnels

La forêt d'arbres décisionnels, une méthode d'apprentissage d'ensemble créé par Breiman, (2001a), consiste à adapter plusieurs arbres *CART* en parallèle à des échantillons *Bootstrap* indépendants, puis d'en combiner les variables indépendantes. L'intérêt de cette méthode effectuée en parallèle est d'exploiter l'indépendance entre chacun des arbres *CART*. L'algorithme de forêt d'arbres décisionnel de Breiman, (2001b) construit d'abord un échantillon *Bootstrap* puis ajuste un grand arbre *CART* non élagué en fonction de celui-ci. À chaque division, seulement *k* variables sélectionnées aléatoirement sont considérées. Ces deux premières étapes sont répétées n fois. Enfin, les prédictions de chaque arbre *CART* sont moyennées, ce processus est appelé Boostrap Aggregation (*Bagging*). L'importance des variables est ajustée en fonction de la moyenne de chacun des arbres *CART* (voir Figure 2.10 ci-dessous).

Un échantillon *Bootstrap* est un échantillon aléatoire conduit avec replacement: à partir de l'échantillon original, n observations sont sélectionnées aléatoirement avec remise parmi les n observations initiales. Ainsi, cette technique de ré-échantillonnage permet de substituer à la distribution de probabilité inconnue dont est issu l'échantillon d'apprentissage, la distribution empirique donnant un poids $\frac{1}{n}$ à chaque réalisation (Efron, 1982).

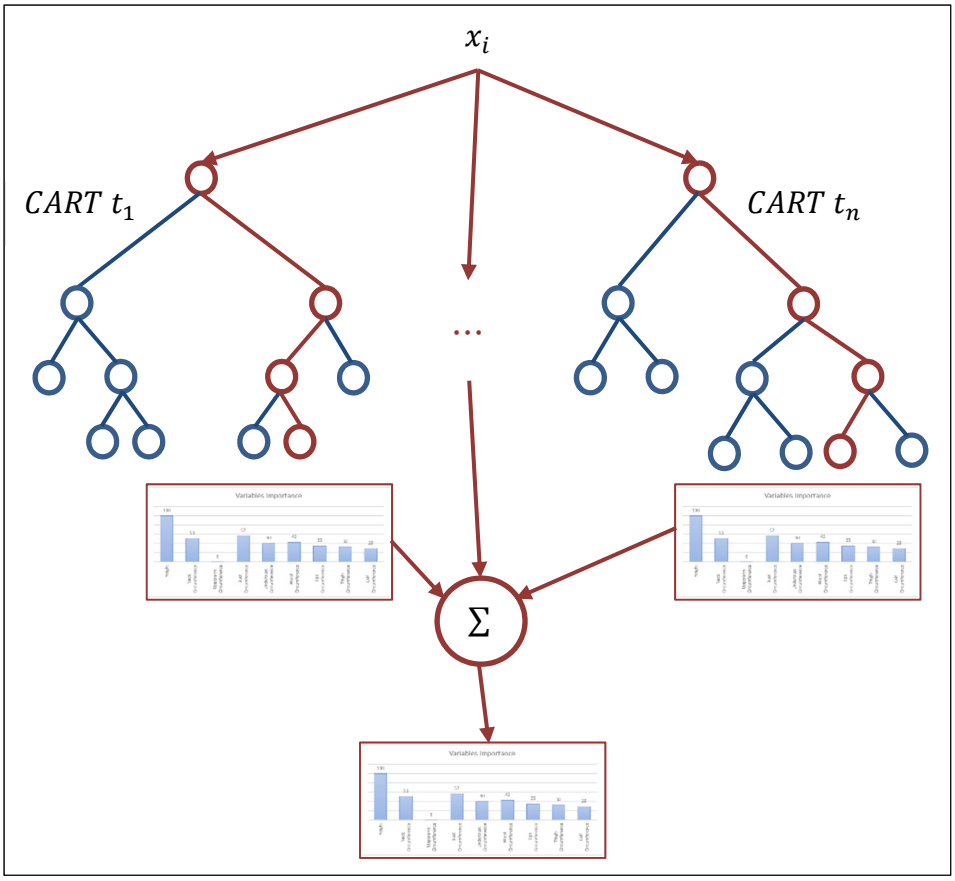


Figure 2.10 Schéma de fonctionnement de l'algorithme Forêt d'arbres de régression. Les histogrammes représentent l'importance de chaque variable

La forêt d'arbres décisionnels est généralement supérieure aux arbres *CART* en raison de sa variance plus faible (dû à la moyenne des arbres). Cet algorithme hérite des mêmes avantages que les arbres *CART*, cependant les résultats sont souvent jugés moins interprétables (Breiman, 2001a).

Avantages :

- ✓ Ajustement d'un seul hyperparamètre de l'algorithme : nombre de variables à sélectionner aléatoirement à chaque nœud;
- ✓ Stable et robuste face aux données aberrantes ou manquantes;
- ✓ L'erreur est réduite dramatiquement par la moyenne des arbres;
- ✓ Plus robuste au surapprentissage et requiert moins de paramétrages pour l'éviter;

- ✓ Plus robuste aux interactions complexes des variables et non-linéarités. Souvent meilleur que la régression linéaire et de performance égale au réseau de neurones;
- ✓ Facile d'interprétation.

Inconvénients :

- × Problème si le nombre de variables pertinentes est faible. Les arbres individuels risquent de ne pas être performants;
- × Bien que facile d'interprétation, l'algorithme offre peu de contrôle sur la sortie (perçu comme une « Boîte noire »);
- × Nécessiteux en termes de ressource mémoire.

2.3.9.3 Boosting d'arbre de régression

Le Boosting d'arbres de régression crée par Friedman, (2002), elle aussi une méthode d'apprentissage d'ensemble, mais qui ajuste itérativement des arbres *CART* afin de généraliser les résidus. L'intérêt de cette méthode effectuée séquentiellement est d'exploiter cette fois-ci la dépendance entre chacun des arbres *CART* (contrairement à l'algorithme Forêt d'arbre décisionnel). Les résidus généralisés sont calculés à partir d'une valeur initiale correspondant à la moyenne des valeurs de la variable dépendante et de l'échantillon aléatoirement construit (*Bootstrap*, voir précédemment) à partir de l'échantillon de données (Friedman, 2002; G. James et al., 2013; Alpaydin, 2014).

Un petit arbre non élagué est ajusté en fonction des résidus généralisés. Cet arbre est ensuite utilisé afin d'ajuster le modèle qui sera réduit par un taux d'apprentissage. L'idée est de modéliser les résidus avec un second modèle et de l'associer au précédent pour une meilleure prédiction. Ce second modèle compense additivement les insuffisances du premier, et ainsi de suite avec d'autres modèles (voir Figure 2.11 ci-dessous).

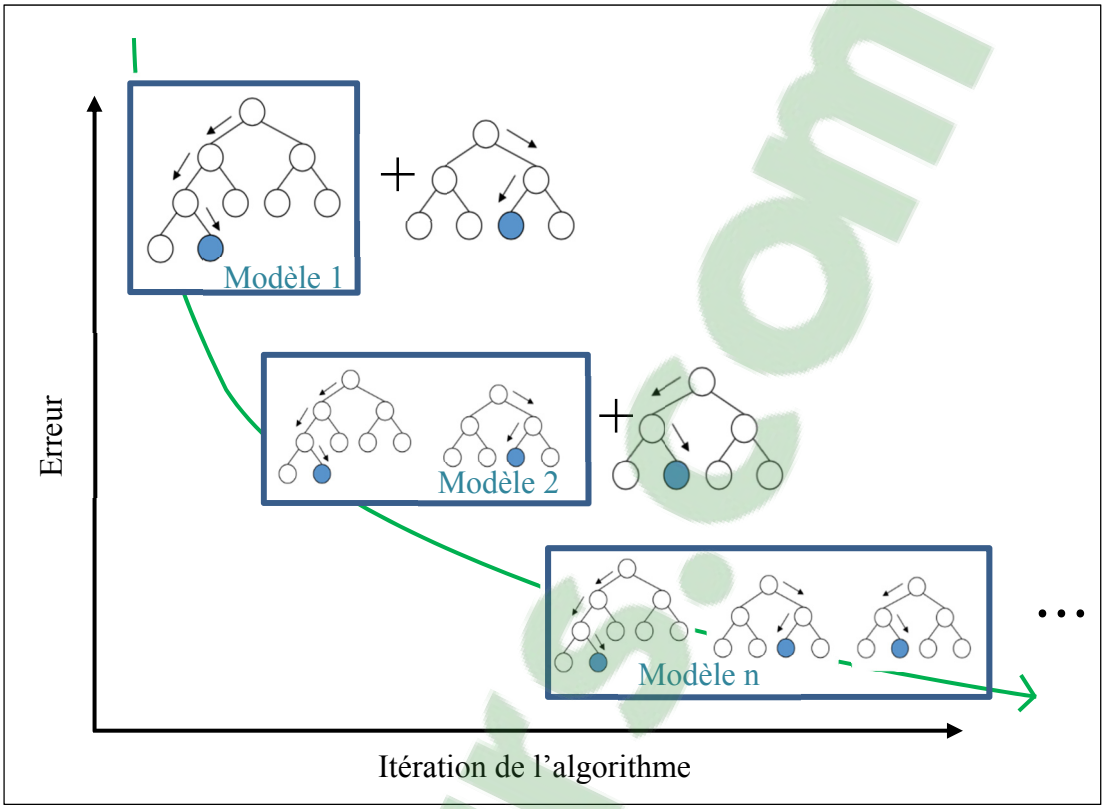


Figure 2.11 Schéma de fonctionnement du Boosting d'arbres de régression

Avantages :

- ✓ Plus robuste aux interactions complexes des variables et non-linéarités;
- ✓ Offre souvent une meilleure prédiction du fait que la variable aléatoire des résidus est augmentée par la diminution de la similarité;
- ✓ Joue sur le biais et la variance.

Inconvénients :

- × Ajustement de plusieurs hyperparamètres compliqué: nombres d'arbres, profondeur (nombres de branches) et taux d'apprentissage;
- × Nécessiteux en termes de ressource mémoire;
- × Risque de sous ou surapprentissage si mal paramétré.

2.3.9.4 Régression à vecteur de support

La régression à vecteur de support (SVR) qui cherche un modèle expliquant la sortie à partir des entrées, est à bien différencier de la machine à support de vecteur pour la classification (qui sépare les données en fonction de leurs classes) et des méthodes dites *Clustering* (qui regroupe les données en fonction de leurs caractéristiques), voir Figure 2.12 ci-dessous.

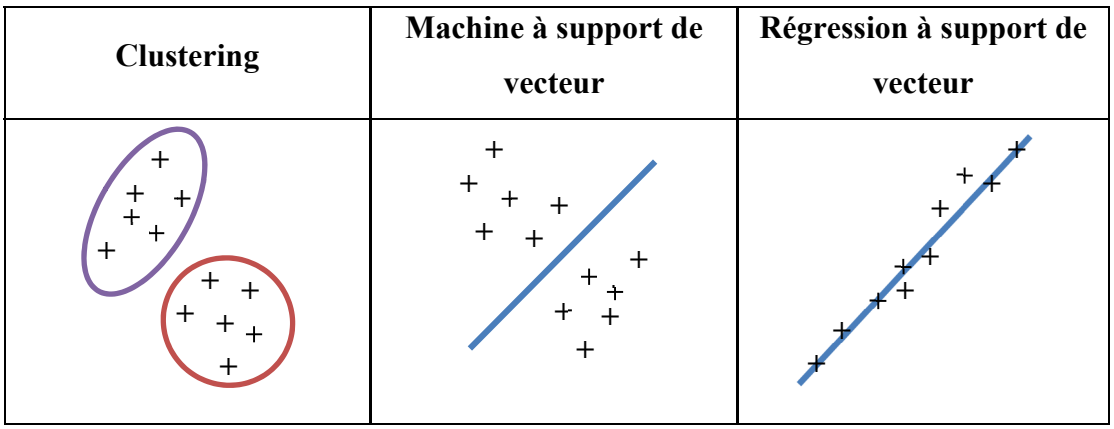


Figure 2.12 Schéma visuel des différences entre les méthodes Clustering et SVM pour la classification et régression

La SVR consiste à trouver une fonction linéaire $f(x) = w \cdot x + b$ avec une déviation d'au plus ε de la variable de sortie y. Le problème peut être écrit comme un problème d'optimisation convexe (voir Figure 2.13 ci-dessous) cherchant à minimiser la fonction $g(x)$ ci-dessous (Smola & Schölkopf, 2004; Awad & Khanna, 2015) :

$$g(x) = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}(\xi_i + \xi_i^*) \tag{2.16}$$

Sous les contraintes suivantes :

$$\begin{cases} y_i - (w \cdot x_i + b) \leq \varepsilon + \xi_i \\ (w \cdot x_i + b) - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0, \;\; i = 1, \ldots, m \end{cases} \tag{2.17}$$

Avec $\|w\|^2$ la norme des paramètres à déterminer ; ξ_i et ξ_i^* des variables de relaxation qui spécifient les erreurs d'apprentissage supérieur et inférieur relatives à une tolérance d'erreur ε. Ces variables sont importantes, car elles permettent de satisfaire les contraintes posées lorsqu'il n'existe pas de solution à la résolution de $\|w\|^2$; C est une constante positive qui détermine le degré de la perte pénalisée lorsqu'une erreur d'apprentissage survient et permet également de contrôler le sous ou surapprentissage. La perte étant la distance entre les valeurs y observées et les limites ε (voir Figure 2.13 ci-dessous). Dans le cas de la régression non linéaire, la réelle difficulté est de trouver le nombre de vecteurs de support (Smola & Schölkopf, 2004; Awad & Khanna, 2015). Les données sont représentées dans un espace de plus haute dimension où l'on cherche à minimiser la formule double $d(x)$:

$$d(x) = \frac{1}{2}\sum_{i,j=1}^{n}(\alpha_i - \alpha_i{}^*)(\alpha_j - \alpha_j{}^*)K(x_i, x_j) - \varepsilon\sum_{i=1}^{n}(\alpha_i + \alpha_i{}^*) + \sum_{i=1}^{n} y_i(\alpha_i - \alpha_i{}^*) \tag{2.18}$$

Sous les contraintes suivantes :

$$\sum_{i=1}^{n}(\alpha_i - \alpha_i{}^*) = 0, \;\; 0 \le \alpha_i, \alpha_i{}^* \le C \tag{2.19}$$

α_i et $\alpha_i{}^*$ étant des multiplicateurs non négatifs introduits pour chaque observation et $K(x_i, x_j) =< \varphi(x_i), \varphi(x_j) >$ une fonction noyau avec $\varphi(x)$ une transformation de x dans l'espace de plus haute dimension. Il n'est pas nécessaire de connaître la transformation $\varphi(x)$ puisque $K(x_i, x_j)$ est directement calculée par la fonction (2.20) permettant ainsi de trouver une fonction optimale $d(x)$ (Smola & Schölkopf, 2004; Awad & Khanna, 2015).

$K(x_i, x_j)$ peut-être soit polynomial ou à base radiale :

$$K(x,y)_{polynomiale} = (x^T y + 1)^d$$
$$K(x,y)_{base\ radiale} = e^{-\frac{1}{2\sigma^2}\|x-y\|^2} \tag{2.20}$$

Cliccours.COM

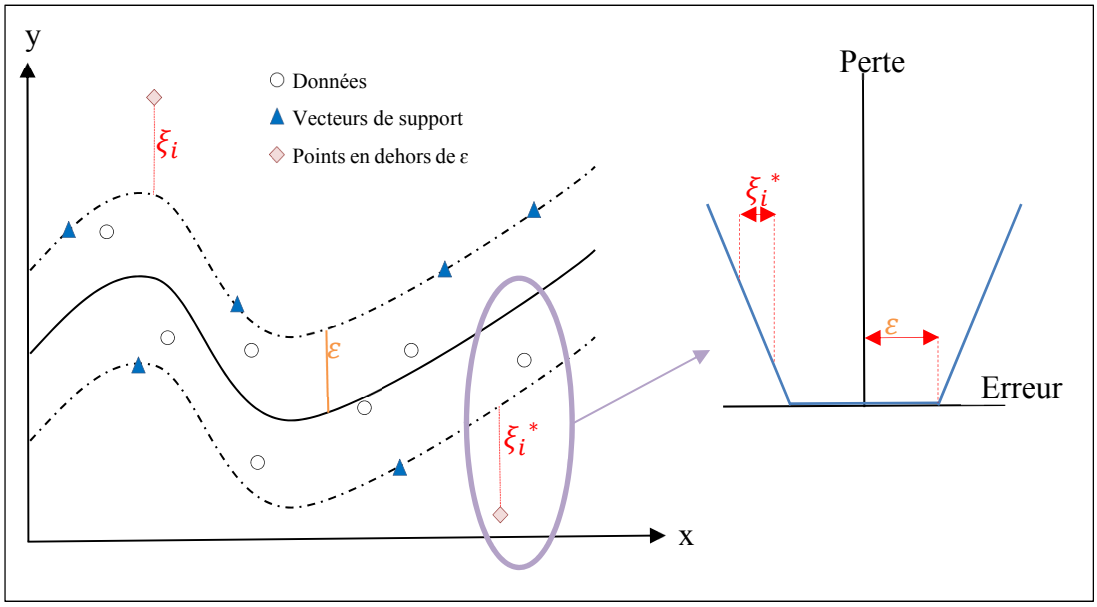


Figure 2.13 Structure du SVR à gauche. Fonction de perte associée à droite. Seuls les points en dehors de l'intervalle ε contribuent au coût final.

Afin de trouver le meilleur sous-ensemble de variables, la stratégie de recherche d'élimination de variables récursives du SVR implémente une sélection *backward* des variables en fonction de leurs ordres d'importance. À chaque itération, les variables sont d'abord ordonnées et les moins importantes sont séquentiellement éliminées. Pour la sélection du modèle finale, l'importance des variables de chaque modèle à travers toutes les itérations est moyennée pour en calculer une valeur globale. Deux informations sont essentielles pour cette stratégie de recherche :

– Sachant que le retrait d'une variable réduit toujours la marge $2 \cdot \varepsilon$, est-ce que cette réduction sera significative ? Sinon, il faut préserver la variable.
– L'erreur ε augmente à mesure que l'on retire des variables. ε est donc un paramètre important.

Avantages :

✓ Plus robuste aux interactions complexes des variables et non-linéarités;
✓ Plus robuste aux données aberrantes et plus souple : contrôlé par la constante C.

Inconvénients :

- Sensibilité aux paramètres;
- Temps de calcul élevé;
- Difficulté d'interprétation et de la pertinence des variables.

2.3.9.5 Réseau de Neurones

Le réseau de Neurones Artificiel (*ANN*) est un algorithme très populaire en science des données qui s'appuie sur l'architecture du réseau de neurones biologique. Plus spécifiquement, un réseau de neurones consiste en un groupe interconnecté de neurones appelés nœuds : les nœuds d'entrée, cachés et de sorties (voir Figure 2.14 ci-dessous). Chaque nœud est connecté aux autres et leurs associations sont caractérisées par leurs forces de connexion: inhibition (maximum étant -1.0) et excitation (maximum étant 1.0). Comme le montre la Figure 2.15 ci-dessous, globalement, chaque nœud calcule la somme de ses entrées puis passe cette valeur à travers une fonction d'activation pour produire sa sortie (ou de transfert). On oppose les réseaux de neurones récurrents dont les connexions entre les nœuds forment un graphe orienté le long d'une séquence, aux réseaux de neurones à propagation avant dans lesquels les signaux ne peuvent se diriger que dans un seul sens, d'une couche d'entrée vers une couche de sortie (tel que sur la Figure 2.15). Les Réseaux de Neurones à propagation avant comprennent le Perceptron (linéaire et non linéaire) et les réseaux à fonction de base radiale : chaque variable correspond à un nœud d'entrée et le réseau a une seule sortie qui correspondant au résultat ciblé.

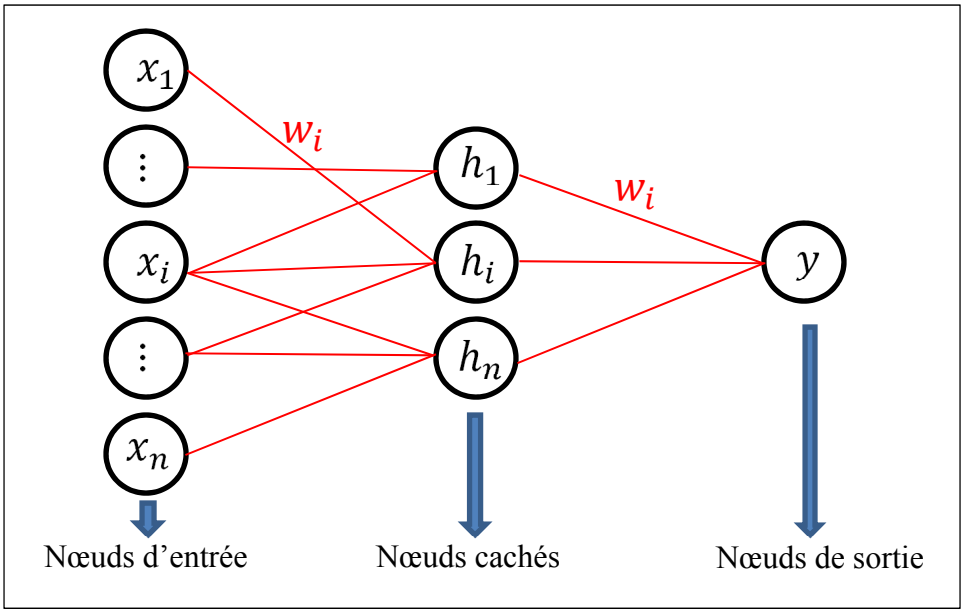


Figure 2.14 Présentation schématique du réseau de Neurones (ou perceptron multi-couche)

En effet, dépendamment de la force de connexion, inhibition et excitation ainsi que la fonction de transfert, les variables d'entrées sont passées aux neurones (nœuds) comme étant une combinaison linéaire de multiples variables. L'équation ci-dessous présente la somme pondérée d'un nœud.

$$net = b + \sum_{i=1}^{n} x_i w_i \tag{2.21}$$

$$o = \varphi(net) \tag{2.22}$$

Un poids w_i est associé à chaque valeur multipliée par chaque variable et une fonction non linéaire (2.21) est appliquée à cette combinaison linéaire de variables, ce qui lui donne la possibilité de modéliser des relations complexes non linéaires. La différence entre la valeur prédite et la vraie valeur (erreur) est propagée en aval (cas du réseau de neurones à propagation avant) de manière à la répartir en fonction des poids w_i de chaque nœud dépendamment de la quantité d'erreurs générée (algorithme de descente de gradient). Une fonction de transfert (également appelé fonction d'activation) φ, est utilisée pour quantifier les entrées pondérées (ou sortie) d'un neurone afin de produire la sortie finale. Plusieurs fonctions existent,

cependant pour la régression, une fonction d'activation linéaire est préférable (domaine restreint $[-\infty; +\infty]$).

Le Perceptron multicouches (MLP), est une classe de réseaux de Neurones à propagation avant. Chaque nœud est un neurone utilisant une fonction d'activation non linéaire. L'apprentissage se fait par rétropropagation du gradient (descente de gradient), qui est une méthode pour calculer le gradient de l'erreur pour chaque neurone, de la dernière couche vers la première. À noter que le Perceptron multicouches reste toujours un modèle non linéaire, même lorsqu'une fonction d'activation linéaire est utilisée : la sortie linéaire de chaque nœud interagit d'une manière non linéaire afin d'influencer la variable prédite. Le réseau de Neurones récurrents à fonction de base radiale partage une architecture similaire au MLP. La différence réside dans les entrées du réseau de neurones caché qui est la somme des distances entre la structure d'entrée et le « centre » de la fonction de base, au lieu de la somme pondérée des entrées (tel que pour le MLP).

Avantages :

- ✓ Plus robuste aux interactions complexes des variables et non-linéarités du fait du nombre de couches et donc de paramètres possibles;
- ✓ Très flexible et robuste concernant la structure des données. Peut apprendre n'importe qu'elle fonction continue.

Inconvénients :

- × Interprétation et compréhension de l'algorithme difficile (perçu comme une « Boîte noire »);
- × Peut être très exigeant en calcul informatique si les paramètres ne sont pas bien paramétrés;
- × Requiers beaucoup de données pour atteindre une bonne performance. Est très vite surpassé par d'autres algorithmes pour de petits échantillons.

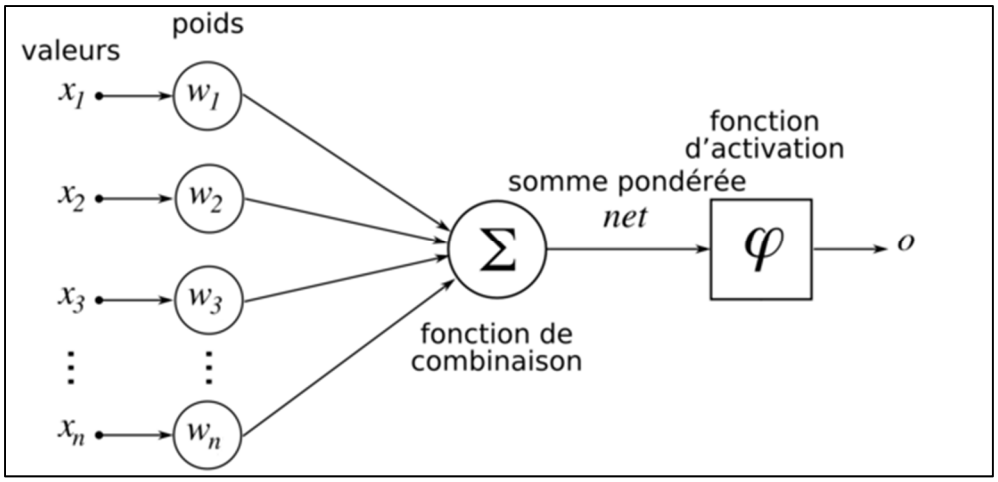


Figure 2.15 Réseau de Neurones à propagation avant (*feed-forward network*). Adapté de (Chrislb, 2005)

2.3.9.6 Moyenne de Réseau de Neurones

L'algorithme avNNet est une méthode d'apprentissage d'ensemble qui consiste à créer de multiples réseaux de neurones en les combinant ensemble afin de produire la sortie désirée. Tout comme la forêt d'arbres décisionnels, un ensemble de réseau de neurones a souvent une meilleure performance puisque les erreurs sont mieux réparties (Mendes-Moreira et al., 2012). Un ensemble de réseau de neurones est construit en deux étapes : entraîner un nombre défini de réseau de neurones puis combiner leurs prédictions par une simple moyenne ou par une moyenne pondérée.

Avantage:

✓ Idem au réseau de neurones, mais performance meilleure.

Inconvénient :

× Idem au réseau de neurones.

2.3.9.7 Régression multivariée par splines adaptatives

La régression multivariée par splines adaptatives (*Multivariate Adaptive Regression Splines - MARS*) est un algorithme de régression non paramétrique qui permet de modéliser les multiples non-linéarités des données par l'utilisation de fonctions charnières (Friedman, 1991a, 1991b; G. James et al., 2013; Osei-Bryson, 2014). Cet algorithme est différent de ceux présentés précédemment en ce qu'il permet non seulement la sélection de variables, mais également l'ajustement d'équation de régression. MARS construit un modèle logarithmique de la forme suivante :

$$\hat{f} = \sum_{i=1}^{k} c_i B_i(x) \tag{2.23}$$

Le modèle est une somme pondérée des fonctions de base $B_i(x)$ avec chaque c_i étant un coefficient constant. Chacune des fonctions de base peut prendre l'une des trois formes suivantes :

- Une constante égale à 1, appelé intercepte.
- Une fonction charnière de la forme $[\pm(x_i - c)]_+$ avec $[\cdot]_+$ la partie positive qui s'écrit également sous la forme $\max(, x - c)\, ou\, \max(0, c - x)$. L'algorithme sélectionne automatiquement les variables x_i et les valeurs pour chaque nœud c de la fonction charnière. Un nœud est un concept important en régression splines puisque c'est ce qui marque la fin d'une région de données et le début d'une autre, mais aussi connecte deux fonctions charnières et donc représente l'interaction entre elles.
- Un produit de deux ou plusieurs fonctions charnières. Ces fonctions permettant de modéliser le degré d'interactions entre deux ou plusieurs variables. Les fonctions charnières peuvent donc être des fonctions linéaires par morceaux ou non linéaires.

MARS construit un modèle en deux temps : une passe avant et une passe arrière. Cette même approche en deux temps est également utilisée par les arbres de partitionnement récursif tel que CART. L'approche MARS surapprend délibérément dans un premier temps le modèle algorithmique afin de trouver la paire de fonctions charnières qui donne la plus petite somme

des carrés résiduels (passe avant). Dans un deuxième temps, le modèle opère une deuxième itération en réduisant le modèle afin de mieux le généraliser à de nouvelles données. À chaque itération, la réduction supprime chaque terme $c_i B_i(x)$ le moins efficace afin de ne garder que les meilleurs sous-ensemble $c_i B_i(x)$. Chacun de ces sous-ensembles est ensuite comparé par validation croisée généralisée afin de sélectionner le meilleur sous-ensemble $c_i B_i(x)$ final, c'est donc la passe arrière.

Avantages :

- ✓ Plus flexible que les modèles de régression linéaire;
- ✓ Régression non linéaire : pas besoin de connaître ou de chercher la forme sous-jacente de la fonction, MARS estime une fonction non linéaire lui-même;
- ✓ Comparer à d'autres algorithmes d'apprentissage machine tel que le réseau de neurones ou encore la forêt d'arbres décisionnels, les modèles de régression MARS sont plus faciles à comprendre et à interpréter. De plus, l'algorithme propose une équation de régression;
- ✓ MARS donne de meilleurs résultats que les arbres de partitionnement récursif puisque les fonctions charnières sont plus appropriées à des données continues;
- ✓ La sélection de variables est automatique;
- ✓ Offre un bon compris entre biais et variance (fonctions restreintes). Le modèle MARS offre plus de flexibilité avec la possibilité de modéliser des non-linéarités et l'interaction des variables, mais pas trop non plus avec une fonction restreinte.

Inconvénient :

- × Régression non linéaire : la fonction non linéaire estimée est une fonction restreinte. Cela est nécessaire, car sans contraintes sur la fonction, c'est un problème inverse qui n'est pas bien posé.

2.4 Création d'un échantillon stratifié

Tel qu'il a été discuté en chapitre 0, les bases de données existantes ne sont pas statistiquement représentatives et sont très chères à acquérir. Il a également été vu en chapitre 2.3.3, l'importance d'une sélection de variables anthropométriques pertinentes pour la modélisation paramétrique du corps humain en 3D avec MH. Afin de pouvoir effectuer des analyses statistiques en science des données, il est donc proposé de créer un échantillon stratifié de la population du Canada. Selon Warfield, Zou, et Wells, (2004), une base de données de référence idéale serait une base de données précise qui refléterait les caractéristiques et la variabilité rencontrées chez une population réelle. En effet, dans le contexte de la modélisation 3D de l'enveloppe du corps humain, l'être humain présente de nombreux changements morphologiques, principalement au niveau du poids et/ou de la taille (Ranger, 2010). L'échantillon devra conserver les mêmes proportions de chacun des sous-groupes hétérogènes de la population étudiée afin d'être statistiquement représentative de celle-ci. Nadadur et Parkinson, (2010) ont discuté la pertinence d'un échantillon stratifié pour une meilleure représentation de la population. La taille de l'échantillon doit être représentative de la population et de sa variabilité (Leedy & Ormrod, 2013) et peut être calculée :

$$n = \frac{z^2}{4 * d^2};\ n_2 = \frac{n * N}{n + N} \tag{2.24}$$

z = niveau de confiance selon la loi normale centrée réduite
(pour un niveau de confiance de 95%, z = 1.96)
d = marge d'erreur tolérée (5% soit 0.05)
N : taille de la population étudiée
n : taille d'échantillon infinie
n_2 : taille d'échantillon finie pour N finit

La taille de l'échantillon est importante puisque l'analyse de nombreuses analyses statistiques en dépend et qu'un autre échantillon produira d'autres limites Sylvestre (2011).

2.5 Métriques de validation

La validation étant des plus cruciales dans ce projet, elle permet non seulement de mettre en évidence les caractéristiques intrinsèques d'une méthode, mais également l'évaluation de ses performances, limitations et contextes cliniques potentiels. De plus, la validation permet de démontrer la valeur ajoutée d'une méthode aussi bien que ses impacts. Les sections suivantes présentent les outils mathématiques et analyses statistiques nécessaires aux métriques de validation, ensuite la création d'une base de données et sa validation spécifique et enfin, la généralisation de méthodes de validation 2D au 3D.

2.5.1 Métriques de validation des modèles prédictifs

La performance de chaque modèle de prédiction est habituellement mesurée par des métriques standard de validation. Cependant, comme le discute Chai et Draxler, (2014), une seule métrique ne fournira qu'une seule projection des erreurs du modèle et par conséquent, n'accentuera qu'un certain aspect des caractéristiques d'erreur. Une combinaison de plusieurs métriques d'erreur est souvent donc requise pour une meilleure évaluation de la performance des modèles de prédictions :

- **MAE** (*Mean Absolute Error*): mesure l'amplitude moyenne des erreurs sans considération de leurs directions. Contrairement à la RMSE ci-après, la MAE est plus facile à interpréter et moins ambiguë (Willmott & Matsuura, 2005).

$$MAE = \frac{1}{n}\sum_{i=1}^{n} | y_i - \hat{y}_i | \tag{2.25}$$

- **MSE** (*Mean Square Error*): Estimation de la variance des erreurs du modèle que l'on cherche à minimiser. Cette quantité permet de déterminer si les variables du modèle sont significatives. La variance résiduelle améliore la précision des estimations des extrémités de la distribution des données (Nadadur & Parkinson, 2010)

- **RMSE** (*Root Mean Square Error*): est un score quadratique qui mesure aussi l'ampleur moyenne des erreurs sans considération de leurs directions. Cette métrique pénalise davantage les erreurs importantes et donne une meilleure vue d'ensemble de la distribution des erreurs. Pour ces raisons, on préféra donc la RMSE à la MAE (Chai & Draxler, 2014).

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{2.26}$$

- **MAPE** (*Mean Absolute Percentage Error*): mesure l'erreur relative en termes de pourcentage et est donc plus facile à interpréter.

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{|y_i - \hat{y}_i|}{y_i} \tag{2.27}$$

La RMSE, MAE et MAPE sont toutes trois des métriques qui doivent être minimisées autant que possible. Plus l'erreur sera petite, meilleur sera le modèle prédictif.

- **Coefficient R^2 et R^2 ajusté** : sont des indicateurs simples, plus ils s'approchent des 100%, plus le modèle est considéré intéressant ou parfait. En revanche, le R^2 ne permet pas de savoir si le modèle est pertinent pour expliquer les valeurs de la BSA. En effet, le R^2 a tendance à mécaniquement augmenter à mesure que l'on ajoute des mesures dans le modèle. De ce fait, il est inopérant si l'on veut comparer des modèles comportant un nombre différent de mesures. Il est conseillé dans ce cas d'utiliser le coefficient de détermination R^2 ajusté qui est corrigé des degrés de liberté. Néanmoins, ce R^2 ajusté peut être négatif, ce qui peut porter à confusion dans l'interprétation. Pour ces raisons, on essaiera de toujours prendre en considération le R^2 et le R^2 ajusté. Le R^2 ajusté est toujours inférieur au R^2. En modélisation de données (statistiques), ces critères devraient toujours être validés par la structure des résidus, car un score élevé n'indique pas forcément un bon modèle de régression ajusté.

Clicours.COM

$$R^2 = 100\left(1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}\right)\%$$

$$R^2_{ajusté} = 100\left(1 - \left(\frac{n-1}{n-p-1}\right)\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}\right)\% \quad (2.28)$$

Avec y_i les valeurs des BSA réelles, $\hat{y}_i$ est les valeurs prédites de BSA, $\bar{y}$ la moyenne des mesures de BSA et n correspond à la taille d'échantillon.

– **Critère d'information Bayesien** (*BIC*). Mesure la qualité relative du modèle statistique tout en tenant compte de la complexité du modèle. Plus la valeur du critère est petite, meilleur est le modèle statistique. Il existe deux critères d'information communément utilisés en statistique descriptive : le Critère d'information d'Akaike (*AIC*) et le Critère d'information Bayesien (*BIC*). Aucun de ces deux critères n'est universellement meilleur Lebarbier and Mary-Huard (2004). Cependant, à la différence du AIC, la pénalité du BIC dépend de la taille de l'échantillon et pas seulement du nombre de paramètres en plus adapté à la statistique descriptive Rawlings, Pantula et al. (2001). Le BIC est plus efficace dans le contexte de l'étude et c'est pour cette raison qu'il a été retenu.

$$BIC = 2 \times \ln(RMSE) + \frac{p \times \ln(n)}{n} \quad (2.29)$$

Avec p le nombre de coefficients estimés et n la taille d'échantillon.

À noter que le Coefficient R^2 ajusté et le critère d'information Bayesien sont des métriques plus utilisés en modélisation de données (statistiques) contrairement à la modélisation algorithmique (apprentissage machine).

2.5.2 Métriques de validation 3D

L'utilisation de plusieurs mesures d'erreurs permet de souligner les différents aspects de la qualité de la modélisation 3D. Il y a un consensus sur les mesures d'erreurs robustes et appropriées dans la revue de la littérature et qui sont fréquemment retrouvées (Aspert, Santa-

Cruz, & Ebrahimi, 2002; Heimann et al., 2009; Heimann, Morrison, Styner, Niethammer, & Warfield, 2010; Gotra et al., 2015; Laurent, Cresson, Vázquez, Hagemeister, & de Guise, 2016):

– Différence de volume absolue (AVD):

$$AVD(A,B) = \left|\frac{A-B}{B} \times 100\right| \tag{2.30}$$

À noter que cette mesure n'est pas symétrique. Une valeur de 0 signifie que les deux volumes sont identiques, mais ne signifie pas que A est égal à B ou qu'ils se chevauchent. En combinaison avec d'autres métriques et donnée en pourcentage, cette mesure signée permet d'évaluer les dissimilarités des volumes 3D par rapport à une référence (Heimann et al., 2009; Heimann et al., 2010).

– Distance de surface symétrique moyenne (ASD) :

Cette métrique calcule pour chaque surface A, la distance euclidienne à la plus proche surface B. La moyenne de toutes les distances de A à B et B à A donne l'ASD (Heimann et al., 2009; Heimann et al., 2010; Gotra et al., 2015) :

$$ASD = \sum_{a \epsilon A} (\min_{b \in B} \|a-b\|_2) \tag{2.31}$$

– Distance de surface symétrique moyenne quadratique (RMS) :

$$RMS(Al, Bs) = \sqrt{ASD^2} \tag{2.32}$$

Bien que les métriques *ASD* et *RMS* soient fortement corrélées, la différence est qu'avec la *RMS*, les écarts importants sont mis en évidences. Il est alors possible de savoir si un ensemble de données aberrantes dévie significativement à partir de la moyenne (Heimann et al., 2009;

Heimann et al., 2010; Laurent et al., 2016). Les métriques *ASD* et *RMS* permettent d'évaluer la fiabilité et la précision des volumes 3D.

– Distance de surface symétrique maximum (MSD) – Distance de Hausdorff :

Cette métrique permet d'évaluer l'erreur globale entre un volume 3D et une référence (Aspert et al., 2002; Heimann et al., 2009; Gotra et al., 2015; Laurent et al., 2016).

Soit un modèle 3D, qui est un maillage triangulaire définit par (B) représenté par un ensemble de points $b \in R^3$ correspondant aux vertices (sommet d'un triangle). La distance de Hausdorff est définie par :

$$h_u(A,B) = \max_{a \in A}(\min_{b \in B}\|a-b\|_2) \tag{2.33}$$

Cependant, $h_u(A,B)$ n'est pas une mesure symétrique, mais unilatérale. Par conséquent, les valeurs de distances peuvent être significativement sous-estimées, particulièrement lorsque deux maillages n'ont pas le même nombre de points. La distance de surface symétrique maximum donne une mesure plus précise des erreurs entre les deux surfaces :

$$h_s(A,B) = \max[h_u(A,B), h_u(B,A)] \tag{2.34}$$

Cette métrique calcule la distance euclidienne maximum entre deux volumes et donc la plus grosse erreur de distance possible entre les deux (Laurent et al., 2016). Cependant, les métriques présentées jusqu'ici ne renseignent que sur la qualité globale de la modélisation. Afin de déterminer la cohérence d'un modèle et déterminer son degré de précision locale, il est préférable de générer une cartographie en couleur de la qualité du modèle par rapport à sa référence. Il est proposé d'utiliser le logiciel open source MeshLab (Cignoni et al., 2008).

2.5.3 Validation de méthodes

La validation de méthodes requiert l'application de critères définis (Jannin et al., 2002), qui sont:

Degré de concordance : Une nouvelle méthode doit être évaluée par comparaison avec une technique établie plutôt qu'une vérité terrain qui ne fournit aucune mesure correcte sans équivoque. Si la nouvelle méthode concorde suffisamment avec l'ancienne, alors l'ancienne peut être remplacée. Déterminé par l'analyse de Bland-Altman, les résultats sont présentés comme le biais ± l'écart-type des différences ($d \pm k \cdot sdd$) avec 95% de l'IC et la valeur de p.

Précision ou exactitude: corresponds au degré par laquelle la mesure est vraie ou correcte. La précision locale est définie comme la différence entre les valeurs calculées et les valeurs théoriques ─ vérité terrain (Warfield et al., 2004). La précision globale étant obtenue par la combinaison des valeurs des précisions locales (Jannin et al., 2002). Déterminé par un test *t* de Student pour des données appariées, les résultats sont présentés comme la moyenne ± l'écart-type et la valeur de p.

Sensibilité ou robustesse: précision par laquelle les résultats d'un processus sont répétables. La sensibilité est un indice de fiabilité et de robustesse en la présence de facteurs perturbateurs. Ce critère est défini par l'erreur de distribution intrinsèque obtenue par la moyenne et l'écart-type (Jannin et al., 2002; Warfield et al., 2004). Ce critère comprend l'étude de :

- la variabilité intra observateurs : le même observateur étudiant la même méthode à des périodes successives, ne trouve pas le même résultat. Soit 1 observateur répétant n fois la mesure.
- la variabilité inter observateurs : la même méthode étudiée par plusieurs observateurs, n'aboutit pas au même résultat. Soit 1 même mesure effectuée par n observateurs.

La variabilité intra-inter observateur peut être observée par l'analyse de Bland-Altman telle que mentionnée précédemment. Cependant, l'outil Six Sigma de mesure de la Répétabilité &

Reproductibilité (*Gage R&R*) est plus accepté dans le domaine de l'anthropométrie en ce qu'il répond à la norme ISO13053-2, (2011). *Gage R&R* est une méthode statistique permettant de déterminer la variation introduite dans la prise de mesures anthropométriques (Dermot & Enda, 2014) et permet également de mesurer les incertitudes (avec des intervalles de tolérances) exprimées en pourcentage afin de statuer sur son acceptabilité (ISO13053-2, 2011).

Gage R&R permet ainsi d'identifier la source de l'erreur : est-ce difficilement répétable ? Dans ce cas, le clinicien n'est pas constant dans sa propre prise de mesure. Ou bien, est-ce un problème de reproductibilité où la même mesure prise par les cliniciens n'est pas constante. L'étude *Gage R&R* doit être réalisée avec un intervalle 6σ afin de capturer 99.7% de la variation et incluant 95% de l'intervalle de confiance. L'acceptabilité de chaque mesure anthropométrique sera identifiée par le pourcentage de la variation totale *R&R* avec ses valeurs de pourcentage de contribution de répétabilité et de reproductibilité de la variance totale en se basant sur le critère d'acceptabilité défini par ISO13053-2, (2011) :

Tableau 2.10 Critère d'acceptabilité d'une mesure répétable et reproductible

Gage R&R	**Nombre de catégories distinctes**	**Résultats**
< 10%	> 14	la mesure est acceptable
> 10% et < 30%	> 4 et < 13	la mesure est plutôt acceptable
> 30%	≤ 3	la mesure est non acceptable, inappropriée et devrait être améliorée

Le nombre de catégories distinctes (ncd) est également important puisqu'il permet d'identifier la capacité de la prise de mesure à faire la différence pour une mesure donnée. Il s'agit essentiellement, du nombre d'intervalles de confiance non superposés qui couvre l'étendue de la variation de la mesure et est inversement proportionnel au Gage R&R (Woodall & Borror, 2008). Les observateurs doivent répéter les mesures dans un ordre randomisé avec des délais entre les mesures pour prévenir de biais de rappel. Les résultats de la première mesure et des autres observateurs doivent être cachés.

Complexité fonctionnelle et temps de calcul : concerne les étapes de la méthode implémentée qui sont coûteuses en temps et fastidieuses pour l'opérateur. Cela traite à la fois l'interaction homme-machine et l'intégration dans un contexte clinique et est en lien avec l'acceptation par les cliniciens du système ou de la méthode. Déterminer par un test *t* de Student pour des données appariées, les résultats sont présentés comme la moyenne ± l'écart-type et la valeur de p.

CHAPITRE 3

HYPOTHÈSES ET OBJECTIFS DE RECHERCHE

À la vue de la littérature, il apparaît qu'aucune méthode actuelle, que ce soit les méthodes cliniques ou les outils de modélisation 2D/3D des brûlures, ne s'avèrent suffisamment adaptées à l'évaluation du pourcentage de TBSA. Ces méthodes ne sont pas assez précises et leurs évaluations ne restent qu'approximatives. De plus, elles ne répondent pas aux besoins de personnalisation des morphologies particulières des patients en surpoids, obèses et enfants et induisent une variabilité entre les observateurs qui est importante.

L'hypothèse fondamentale de ce projet est que la superficie corporelle locale, qui est induite par la morphologie du patient, peut différer de l'approximation de la BSA et ainsi affecter l'évaluation du %TBSA. Une deuxième hypothèse tout aussi importante est que l'utilisation d'un logiciel open source comme MH permettrait d'offrir une modélisation paramétrique à partir de mesures anthropométriques suffisantes pour une estimation précise de %TBSA.

L'objectif principal du présent projet de doctorat consiste en la conception et la validation de méthodes numériques permettant l'évaluation quantifiée de la taille des superficies corporelles brûlées à partir de modélisation 3D du corps entier obtenue à partir d'un nombre limité de mesures anthropométriques.

Les objectifs spécifiques de la thèse sont les suivants :

Objectif 1 :
Évaluer la précision de la modélisation 3D du logiciel MH à partir d'un nombre limité de mesures anthropométriques pour une BSA précise.

clicours.com

Objectif 2 :

Réaliser une première interface de rendue 3D sur ordinateur pour l'estimation visuelle des surfaces corporelles brûlées sur un modèle 3D.

Objectif 3 :

Procéder à une première validation de la précision d'une modélisation 3D à partir de 21 mesures anthropométriques sur 4 mannequins commerciaux (1 enfant, 2 femmes et 1 homme).

Objectif 4 :

Quantifier l'impact de la morphologie dans l'évaluation précise du pourcentage de TBSA.

Objectif 5 :

Valider l'évaluation de TBSA avec l'interface de rendue 3D sur ordinateur par rapport à l'évaluation de TBSA avec la méthode clinique du schéma de Lund & Browder.

Objectif 6 :

Générer un large échantillon stratifié aléatoire et représentatif de la population du Canada.

Objectif 7 :

À partir de cet échantillon, déterminer quelles sont les mesures anthropométriques les plus influentes en fonction de la variabilité des morphologies humaines et en proposer des sous-ensembles restreints.

Objectif 8 :

Valider la précision de la modélisation 3D à partir d'un nombre de mesures anthropométriques limité sur des sujets saints. Les modèles 3D doivent être suffisamment précis à la fois au niveau de la BSA et de la morphologie.

De plus, ce projet s'inscrivant dans un contexte particulier d'urgence, la démarche de conception doit constamment prendre en compte des exigences cliniques qui seront satisfaites par chacun des objectifs spécifiques :

Exigence 01. Le nombre de mesures anthropométriques doit être inférieur ou égal à six. Cette valeur a été déterminée suite à une discussion interne avec une équipe d'experts (chirurgien plasticien, chirurgiens traumatologues, urgentologues, ergothérapeutes/ physiothérapeutes et infirmières).

Exigence 02. La précision de l'estimation du TBSA avec l'approche proposée dans le présent projet de recherche doit être supérieure à la précision de l'estimation du TBSA par les méthodes conventionnelles utilisées par les cliniciens.

Exigence 03. La variabilité entre les observateurs avec l'approche proposée doit être inférieure à la variabilité observée chez les cliniciens.

Exigence 04. L'approche proposée doit se situer dans la marge d'erreur clinique qui est comprise entre 5% et 10% (%TBSA), 15% étant la marge d'erreur obtenue par les chirurgiens les plus expérimentés. Cette valeur a été déterminée suite à une discussion interne avec une équipe d'experts (chirurgien plasticien, chirurgiens traumatologues, urgentologues, ergothérapeutes/ physiothérapeutes et infirmières).

Exigence 05. L'approche proposée doit être simple d'utilisation pour une utilisation en milieu clinique et répondre à la demande de cliniciens experts.

Exigence 06. La complexité fonctionnelle et le temps de calcul de l'approche proposée doivent être acceptables pour une utilisation en milieu clinique (≤ 15 min). Cette valeur a été déterminée suite à une discussion interne avec une équipe d'experts

Clicours.COM

(chirurgien plasticien, chirurgiens traumatologues, urgentologues, ergothérapeutes/ physiothérapeutes et infirmières).

Les travaux réalisés pour répondre au 1er objectif spécifique sont décrits au CHAPITRE 4, et introduisent des résultats préliminaires afin de mettre en place le contexte du projet et sa possible pertinence clinique. Les résultats obtenus on fait l'objet d'une publication d'un article scientifique dans la revue « International Journal of Human Factors Modelling and Simulation » (2017) ainsi qu'une présentation orale au congrès « 4th International Digital Human Modeling » (Montréal, 2016). Les travaux reportés au CHAPITRE 5 permettent de répondre aux objectifs spécifiques 2 à 5 et on fait l'objet d'une publication d'un article scientifique dans la revue « Journal of Burn Care & Research » (2017) ainsi qu'une présentation poster au congrès « Canadian Special Interest Group of the 49th American Burn Association Annual Meeting » (Boston, 2017) et une présentation orale au congrès « 4th annual Canadian Burn Symposium » (Winnipeg, 2017). Les résultats obtenus ont permis de démontrer l'importance de la morphologie dans l'estimation du %TBSA et introduisent des résultats de validation justifiant la pertinence clinique du projet. Les travaux ayant permis d'atteindre les objectifs spécifiques 6 à 8 ont fait l'objet d'une soumission d'article scientifique dans la revue « IEEE – Transactions on Biomedical Engineering » (2018) et sont décrits au CHAPITRE 6. C'est toute la conception de modélisation 3D à partir de mesures anthropométriques, en passant par des méthodes de modélisation prédictive, qui y sont présentées. Enfin, le CHAPITRE 7 présente une discussion de l'ensemble de ces résultats ainsi que les limitations et les retombées du projet pour d'autres domaines d'application.

CHAPITRE 4

TOWARD A QUANTIFIED ASSESSMENT OF TOTAL BODY SURFACE AREA FROM ANTHROPOMETRIC MEASUREMENTS FOR PATIENT WITH BURN INJURIES

Adrien Desbois[1-2], Isabelle Perreault[3], Gabriel Chartrand[1-2], Thierry Cresson[2], Sylvie Gervais[1], Jacques A. de Guise[1-2]

[1] École de technologie supérieure (ÉTS), Montréal (QC), Canada
[2] Laboratoire de recherche en Imagerie et Orthopédie (LIO) – Centre de recherche du Centre Hospitalier de l'Université de Montréal (CRCHUM), Montréal (QC), Canada
[3] Division de chirurgie plastique, faculté de médecine, département de chirurgie, Université de Montréal, Montréal (QC), Canada

Article publié dans « International Journal of Human Factors Modelling and Simulation », Octobre 2017.

4.1 Mise en contexte

Une des hypothèses principales du projet suppose que l'utilisation d'un logiciel open source comme MakeHuman (MH) permettrait d'offrir une modélisation paramétrique à partir de mesures anthropométriques suffisantes pour une estimation précise de la BSA. D'autre part, il a été vu aux chapitres 2.1 et 2.2 qu'une modélisation paramétrique est l'approche la plus adaptée au contexte clinique. Ces chapitres insistent également sur le fait que plus de deux mesures anthropométriques sont nécessaires pour une modélisation 3D précise de la BSA. Des résultats préliminaires étaient donc indispensables afin de mettre en place le contexte du projet et justifier sa pertinence clinique. Ce premier article publié répond de manière non exhaustive, au 1er objectif du projet de doctorat « Évaluer la pertinence du logiciel MH pour une modélisation 3D suffisante de la BSA à partir d'un nombre limité de mesures anthropométriques ». Les résultats obtenus ont permis d'initier la réflexion exposée au chapitre 2.3 et les choix méthodologiques résultants. Les travaux relatifs à cet article ont également été

présentés au congrès « 4th International Digital Human Modeling » (Montréal, 2016) et d'un article web sur le site d'actualités scientifiques Substance (2016).

4.2 Abstract

The amount of replacement fluid a burn patient requires to survive depends on the ratio (TBSA) of burned body surface area (morphology) to the total body surface area (BSA). The 2D methods used by clinicians are imprecise. In this paper, preliminary result of a proposed approach using anthropometric measurements and MakeHuman (MH) software to evaluate TBSA is presented. To assess TBSA accurately with a personalised 3D model of the burn patient, a first critical step is to find a limited set of measures for BSA assessment. 20 anthropometric measurements were acquired virtually on 40 3D models generated with MH software. Using several multiple regression analyses, it was demonstrated that four to seven measures are sufficient to obtain an accurate BSA. These preliminary results highlight the relevance of using software such as MH, to assess BSA of patients with major burn injuries based on a limited set of measures.

4.3 Introduction

Shortly after admission to the hospital, clinicians assess the size of the burns in order to replace the amount of fluid loss and help the patient to survive. The amount of replacement fluid the patient requires to survive depends on a common clinical formula directly linked to the TBSA, the ratio of the burned body surface area to the total body surface area (BSA), (Durrant, Simpson et al. (2008). The assessment of TBSA is also important for the clinician in identifying criteria for transfer to a burn unit and ruling on the patient's outcome. Clinicians use several pen-and-paper methods to assess TBSA, such as the rule of nines, the rule of palms, and the Lund & Browder chart. Both assessments with the rule of nines and the Lund & Browder chart are based on transposing the observed burns onto a diagram of a human body with fixed body proportions. The rule of palms considers the surface of a palm as 1% of the total body surface area. Because of the 3D transposition of burns on 2D diagram, these clinical methods under- or overestimate the TBSA with a variability among observers (Wachtel et al. 2000; Neuwalder

et al. 2002; Hettiaratchy and Papini (2004); Giretzlehner et al. 2013; Parvizi et al. 2014). Children and obese patients are even more touched by these conditions.

Imprecision in burns assessment result in significant medical and surgical complications (Livingston and Lee 2000; Parvizi et al. 2014). This has led to the emergence of digital methods for a better assessment of burns size. The most interesting methods are the one representing burns onto a 3D model, (Dirnberger, Giretzlehner et al. (2004); Prieto, Acha et al. (2011). However, these methods use predefined models based on anthropometric formulas (such as Du Bois and Du Bois 1989; Boyd 1935; Gehan and George 1970, etc.) to approximate the body surface area with the patient's weight and height only. The body constitution remains an approximation and the morphology is not represented. Therefore the TBSA estimation will still be inaccurate. Moreover, the weight of the patient is often estimated visually and too complex to be modeled in 3D (depends on the mass of bones, water, muscle, etc.). More anthropometric measures will overcome this approximation by reaching some body areas in order to offer a better assessment of the total body surface area. However, the emergency context of burn patients limits the possible amount of anthropometric measures to be taken. Therefore, the aim of this preliminary study was to verify the possibility to obtain an accurate TBSA with only some anthropometric measurements (excluding the weight) in order to obtain personalized 3D model for a better TBSA assessment. This paper introduces preliminary results of a proposed approach incorporating anthropometric measurements and MakeHuman (MH) software. Section 2 is devoted to describing the proposed approach, Section 3 to the results, Section 4 to the discussion, and Section 5 to the conclusion.

4.4 Materials and methods

4.4.1 Generation of realistic three-dimensional models

MakeHumanTM (MH) software was used in this study to provide realistic and personalized 3D models. It allows the deformation of an initial 3D model through parameters such as age, gender, morphology and anthropometric measures. Moreover, the modeling algorithm takes into account complex and extreme body shapes such as children, obesity (class I to III,

including skin folds), pregnancy, etc. Random 3D models (see Figure 4.1) were generated through a plugin that randomly varies the parameters mentioned above.

Since the final tool is intended to be used in an emergency context, a first critical step for this project is to find a limited set of measurements for personalizing the model that can be performed on the patient easily. To identify this optimal set, a preliminary study was performed on 40 models (19 men, 21 women). These models were generated randomly using the plugin mentioned above, with a mean age of 29.5 ± 21.5 years, a mean height of 152 ± 21.8 cm, and a mean BSA of 1.51 ± 0.5 m².

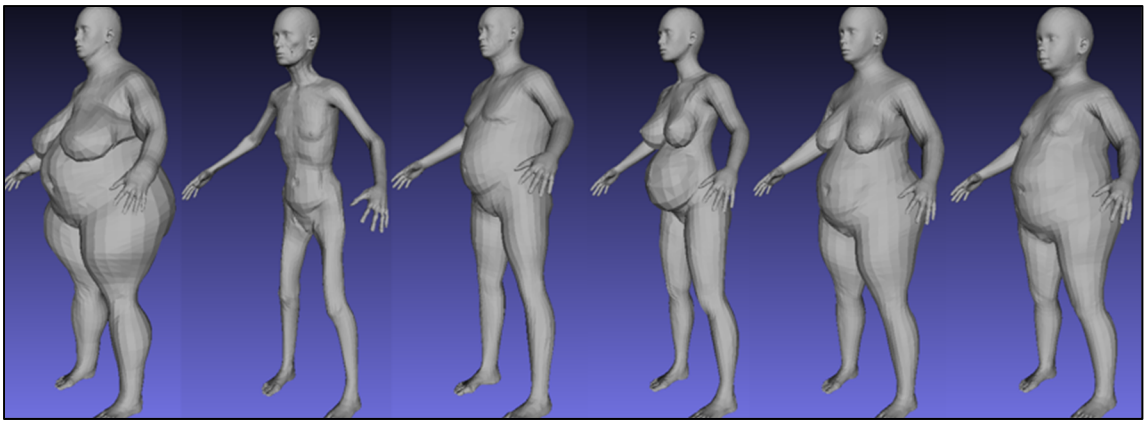


Figure 4.1 Examples of some 3D models randomly generated for this study

4.4.2 Anthropometric measuring process

Geomagic Verify Viewer software was used to perform 20 independent measurements directly on each model, 10 circumferences and 10 distances (see Figure 4.2).

The circumferences taken include the neck, bust, under-bust, waist, hips, thigh, calf, ankle, upper arm, and wrist. The distances taken measured the shoulder, chest, upper and lower legs, upper and lower arm, waist to hip, neck to waist, neck and stature.

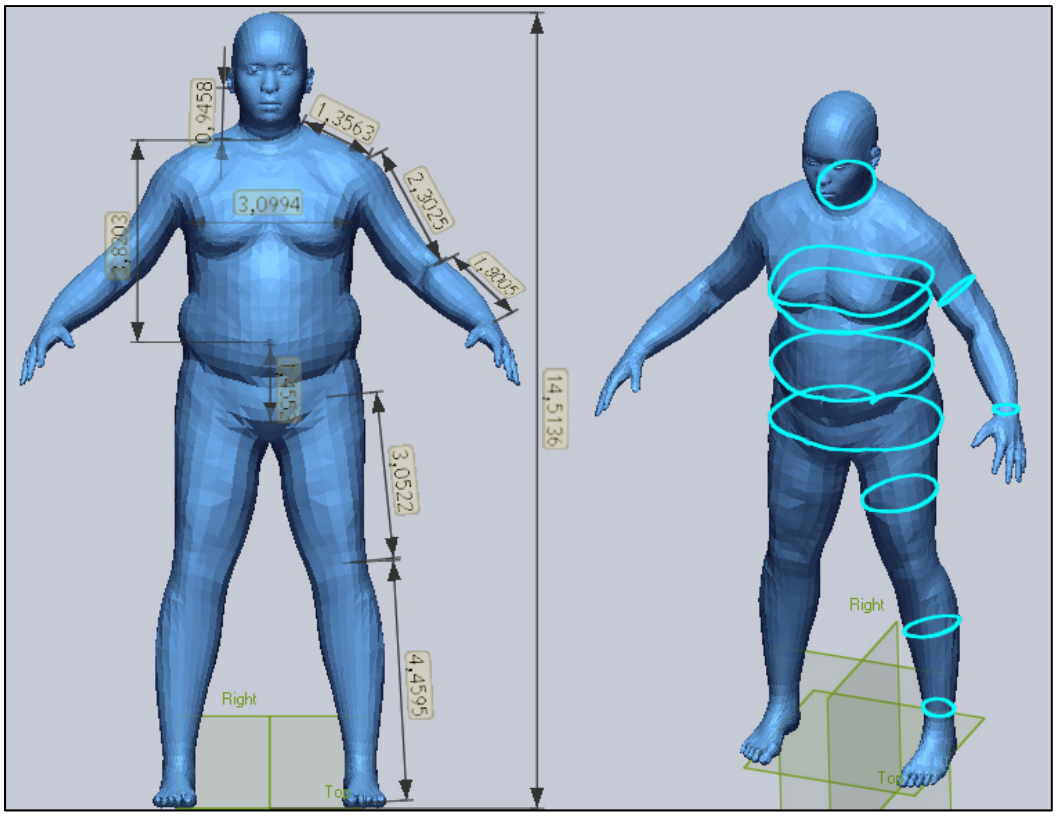


Figure 4.2 Anthropometric measurements used in this study

4.4.3 Multiple regression

By performing a multiple regression analysis on the anthropometric measures and BSA of each 3D model using the Statgraphic Centurion XVI application, we were able to build a model (4.1) to identify the anthropometric measures (X_i) that are significantly linked to the BSA (Y).

From the triangular mesh of each model, Heron's formula of Weisstein, (2015) was used to calculate each triangle's area. The BSA of a model is obtained by the sum of its triangle's area. The aim of this step is to use different stepwise regressions to reduce the 20 anthropometric measurements to a limited set. Forward and backward selection methods were used on six different sets of anthropometric measurements. The forward selection method starts with 0 variables in the model. Then, X_i variables are added one by one based on one criterion: whether they improve R^2 and adjusted R^2. For each step, the model is recalculated until the criterion is

satisfied. The backward selection method starts with X_i variables selected for the model. Then, X_j variables are removed one by one, following the same criterion.

$$Y = a_0 + \sum_{1}^{n} a_i X_i \quad i = 1, \ldots, 20 \tag{4.1}$$

4.5 Results

As this is a preliminary study, it was decided that six independent stepwise regressions would be performed. At all times, the choice of the measurements must be guided by the principle of parsimony and must be relevant in the context of patient with major burn injuries. This is why each stepwise regression was performed on different sets of anthropometric measurements considered relevant (see Tableau 4.1 and Tableau 4.2). All the regression models are statistically significant at the 95% confidence level with a R^2 and an adjusted R^2 greater than 98% (p-value < 0.05).

Tableau 4.1 Limited set of measures for each regression model

Regression model	**R^2 (%)**	**Ajusted R^2**	**p-value**
Model 1	99.32	99.18	p-value < 0,05
Model 2	99.56	99.42	
Model 3	99.50	99.41	
Model 4	99.07	98.93	
Model 5	98.86	98.73	
Model 6	99.68	99.56	

Tableau 4.2 Results of each regression model

Regression model	1	2	3	4	5	6
Stature	X	X	X	X	X	X
Upper arm circumference	X	X	X			
Lower arm length	X	X				
Bust circumference					X	
Under-bust circumference	X	X	X	X		
Waist circumference	X			X		
Waist to hips length	X	X	X	X		X
Thigh circumference		X		X	X	X
Lower leg length		X				X
Hips circumference			X		X	
Calf circumference			X			X
Neck circumference						X
Wrist circumference						X

4.6 Discussion

In this preliminary study, results show that four to seven measures are sufficient to obtain an accurate BSA. For example, only four measures (i.e. height, circumference of the chest, hips and thighs) can explain 98.9% of the variability in total surface area. The results in Table 4.2 strengthen the argument that BSA assessment requires four or more measures to be accurate. It is even more critical for obese persons, in which case both waist and bust circumference and waist to hips length should be considered for a more precise BSA (Neaman, Andres et al. (2011).

At this early stage, the choice of 6 regression models was made in order to make inferences about the limited set. However, the regression models do not take into account possible correlations existing between the anthropometric measures in MH, which could have lead to including unnecessary measures in our model. This seems likely given the fact that we were able to obtain different regression models with comparably excellent fit based on R^2. This issue could be overcome, first by assessing the presence of multicollinearity with the variance inflation factor. Second, by applying a better regression method such as subset regression or

non-linear regression. Finally, validation metrics such as the mean square error, Bayesian information criterion, standard error of the residuals should be used.

4.7 Conclusion

In the context of patient with major burn injuries, these results show that it is possible to obtain an accurate BSA with only some anthropometric measurements. Knowing there is a need to adapt 3D models due to their morphological measurements (Giretzlehner et al. 2004; Prieto et al. 2011), this preliminary study highlights the relevance of using a software such as MH, to assess TBSA of patients with major burn injuries based on a limited set of measures. The main advantage would be the increased accuracy of a personalized 3D model, which would allow for more accurate assessment of BSA. Further studies need to be done to design a better process for the determination of the limited set of measurements in order to assess the BSA accurately. As well, further studies should be performed for the 3D assessment of the TBSA on the 3D model.

CHAPITRE 5

THE IMPORTANCE OF A 3D-BASED APPROACH WITH PERSONALIZED MODELS FOR ACCURATELY ASSESSING TBSA

Adrien Desbois[1-2], Svetlana Matei[3], Florian Beguet[1-2], Isabelle Perreault[4], Sylvie Gervais[1], Jacques A. de Guise[1-2]

[1] École de technologie supérieure (ÉTS), Montréal (QC), Canada
[2] Laboratoire de recherche en Imagerie et Orthopédie (LIO) – Centre de recherche du Centre Hospitalier de l'Université de Montréal (CRCHUM), Montréal (QC), Canada
[3] Département de chirurgie, Montréal (QC), Canada
[4] Division de chirurgie plastique, faculté de médecine, département de chirurgie, Université de Montréal, Montréal (QC), Canada

Article publié dans « Journal of Burn Care & Research», Février 2018.

This is a non-final version of an article published in final form in Journal of Burn Care & Research: https://academic.oup.com/jbcr/article/39/2/235/4802244

5.1 Mise en contexte

L'hypothèse fondamentale de ce projet de doctorat est que la superficie corporelle locale, qui est induite par la morphologie du patient, peut différer énormément de l'approximation de la BSA et ainsi grandement affecter l'évaluation du TBSA. Cela justifie la nécessité de personnaliser les modèles 3D aux constitutions corporelles des patients. Ce deuxième article publié, répond aux objectifs 2 à 5 qui permettent de démontrer l'importance de la morphologie dans l'estimation de la TBSA et introduisent des résultats de validation supportant la pertinence clinique du projet. Les exigences cliniques 02 à04 ont été jugées satisfaites par ces travaux. Cet article qui a également été présenté au congrès « Canadian Special Interest Group of the 49th American Burn Association Annual Meeting » (présentation poster, Boston, 2017) et au congrès « 4th annual Canadian Burn Symposium » (présentation orale, Winnipeg, 2017), une

Clicours.COM

présentation orale au département d'urgence du CHU de Sainte-Justine (Montréal, 2017) et deux présentations orales au département de traumatologie du CHU de Sainte-Justine (Montréal, 2017). Plusieurs cliniciens ont porté un intérêt particulier au projet lors de ces congrès et ont accepté d'aider dans la recherche associée au projet.

5.2 Abstract

Objective: In this paper, the relevance of a new three-dimensional computer-based framework with personalized 3D models for accurately assessing the TBSA is demonstrated through preliminary results and validation studies.

Methods: First, a 3D rendering interface was developed for representation and calculation of TBSA. The personalized 3D models were built from anthropometric measurements using MakeHuman software. 15 paired models were randomly built with an equal body surface area but a different morphology. The difference of local body surface area, TBSA burned and Parkland fluids between each paired models were observed to highlight the impact of morphology's variation on the TBSA. Finally, a preliminary validation study was made on 4 mannequins by 14 volunteers to assess the accuracy of the 3D models built with MakeHuman software and TBSA burned assessment with the proposed method.

Results: Small variations in the morphology impacted the TBSA assessment. Mannequin's 3D models built with MakeHuman software presented an absolute error of 3 ± 2.2 % with no significant difference with their scans. The proposed approach allows for a better assessment of TBSA with a lower variability. No significant difference in the scores for expert and non-expert conditions was observed.

Conclusion: Personalized 3D model to the patient's morphology is suggested to overcome the difficulty of patients with specific morphologies such as obese and children. The proposed framework appears to be relevant for personalizing and accurately assessing TBSA and could reduce morbidity and mortality.

5.3 Introduction

Percent TBSA burned assessment is crucial for patients with major burn injuries. It helps not only in determining the quantity of fluid required but also in identifying criteria for transfer to a burn unit and evaluating the patient's outcome. Many studies have highlighted the imprecision of the two-dimensional (2D) clinical methods for TBSA assessment due to the fact they require transposing three-dimensional (3D) burns onto a 2D diagram. According to some authors, clinical methods overestimate TBSA by 50% to 161% due to morphological variations of the patients, with a variability among observers of more than 62% (Wachtel et al., 2000; Giretzlehner et al., 2013; Daryousch Parvizi et al., 2014). This is even more critical in the case of patients with central obesity (Kamolz, Parvizi, et al., 2014). However, obesity is a frequent clinical condition (Neaman et al., 2011) and is a common problem for all developed countries, where obesity rates are continuously increasing. Although children with burns represent the cases with the highest mortality, the outcomes for obese children are even worse (Livingston & Lee, 2000). Also, body proportions are not specific to the patient.

The rule of nines is not appropriate for children (Livingston & Lee, 2000). The rule of palms can be ambiguous. For some authors, the surface of a palm includes the fingers (Hettiaratchy & Papini, 2004; Durrant et al., 2008) while for others it does not (Malic et al., 2007) or represents 0.78% of the body surface area (BSA) rather than 1% (Gehan & George, 1970). The Lund & Browder chart is the most appropriate method since it is adjusted for children (Daryousch Parvizi et al., 2014). Errors in assessing TBSA can lead to significant medical and surgical complications (Livingston & Lee, 2000; D. Parvizi et al., 2014).

Given these challenges, many authors have attempted to improve TBSA assessment by developing digital methods. The first software applications were mainly digital representations of the Lund & Browder chart, with automatic computation of TBSA and additional functions (fluid requirement, types of burns, etc.) such as SAGEII Neuwalder et al., (2002) and BurnCalculator (Berry et al., 2006). The latter has been shown to overestimate the burn area in 25% of cases and underestimate it in 15% of cases. Some efforts have been made to improve

the accuracy of the Lund & Browder chart by providing a more accurate hierarchical decomposition of the diagram (Richard et al., 2015). As they still rely on a 2D diagram, these programs suffer from the same accuracy problems already mentioned above.

To address this shortcoming, some authors have proposed representing burns in 3D. One of the first tools to achieve this aim was EPRI's 3D Burn Vision (Neuwalder et al., 2002), which provides a single predefined 3D mesh. Although it affords a 3D representation of the burns, the mesh is not accurate and is difficult to adapt to the patient. The most known and reliable is BurnCase 3D (J Dirnberger, Giretzlehner, Ruhmer, Haller, & Rodemund, 2002; Johannes Dirnberger et al., 2004b; Giretzlehner et al., 2004; Giretzlehner et al., 2013; Parvizi et al., 2016), which uses a database of predefined 3D human models to supply an adapted 3D model based on the input of information such as age, gender, size, and weight. It computes TBSA and fluid requirement and has additional features (clinical records, images, etc.). However, a limitation of BurnCase 3D is the modeling algorithm which is based on empirical formulas to approximate the BSA of its adapted 3D models. These formulas can approximate BSA if weight and height are known (D. Du Bois & E. Du Bois, 1916; Gehan & George, 1970; Haycock, Schwartz, & Wisotsky, 1978). The TBSA assessment mainly depends on the BSA approximation which presents estimation error (Tikuisis et al., 2001). This is the case of the clinical methods which come from data compiled using these formulas (C. C. Lund, 1944; Gehan & George, 1970). A later solution, BAI (Prieto et al., 2011), improved the same process of selecting a model from a database of predefined 3D human models by including some obese 3D models.

Thus far, BurnCase 3D (Giretzlehner et al., 2004; Giretzlehner et al., 2013; Parvizi et al., 2016), and BAI (Prieto et al., 2011), have been the only tools to offer a suitable solution for assessing TBSA. However, the fact that predefined models are only adapted to the patient's weight and height means there is an unknown offset between the patient's real BSA and the BSA of the 3D model, which has not been evaluated (Johannes Dirnberger et al., 2012). The model selected is still an approximation of the patient's actual body constitution which sometime diverges from it's real BSA (Johannes Dirnberger et al., 2004b). Nevertheless, the studies of

Giretzlehner et al., (2013) and Parvizi et al., (2016) have highlighted the potential of computer-aided systems such as BurnCase 3D to improve TBSA assessment and inter-rater reliability for burns below 20% TBSA. A great improvement was the semiautomatic superimposition onto the 3D model of wound photography to help the tracing of burns. As well, the database offers a knowledge base of medical advice and diagnosis support. Several attempts have been made to capture the real shape of a person with a Kinect (Popescu & Lungu, 2014; Sheng et al., 2014), and thus provide personalized 3D models. Although the resulting 3D scan is personalized, it is not useable due to the Kinect's insufficient optical resolution, which necessitates post-treatment of the scan. The lengthy acquisition time (>15 min), the impossibility of moving the subject, and lack of interface make this method unsuitable at this time for initial management of burn patients in a clinical context.

An improvement to these tools would be to take into account the correlation between different human morphologies. As suggested by Johannes Dirnberger et al., (2004b), another approach whose model deformation algorithm relies on measurements entered by the surgeon should be considered.

In this paper we present a new 3D computer-based framework for evaluating TBSA, incorporating anthropometric measurements, MakeHuman (MH) software, and a 3D rendering interface. This approach aims to satisfy clinical requirements by maximizing TBSA assessment accuracy and personalizing 3D models of burn patients. The goals of this study were: first, to demonstrate the importance of personalizing 3D models to the patient's body constitution for a better assessment of TBSA. Second, to perform a preliminary validation of the 3D models built with the MakeHuman software. Third, to carry out a preliminary validation of the proposed approach.

5.4 Methods

5.4.1 Burn representation and calculation of the BSA, burned BSA, and TBSA

To assess accurately the body surface area (BSA), burned body surface area (BBSA) and TBSA, a simple 3D rendering interface was developed in C++ using Qt and VTK libraries. The 3D model is composed of triangles (see Figure 5.1), each one composed of adjacent vertices v1, v2 and v3. The surface area S_{T_i}, of a triangle T_i of the mesh, is calculated by Heron's formula (Weisstein, 2015) based on the coordinates of the adjacent vertices (Equation 5.2).

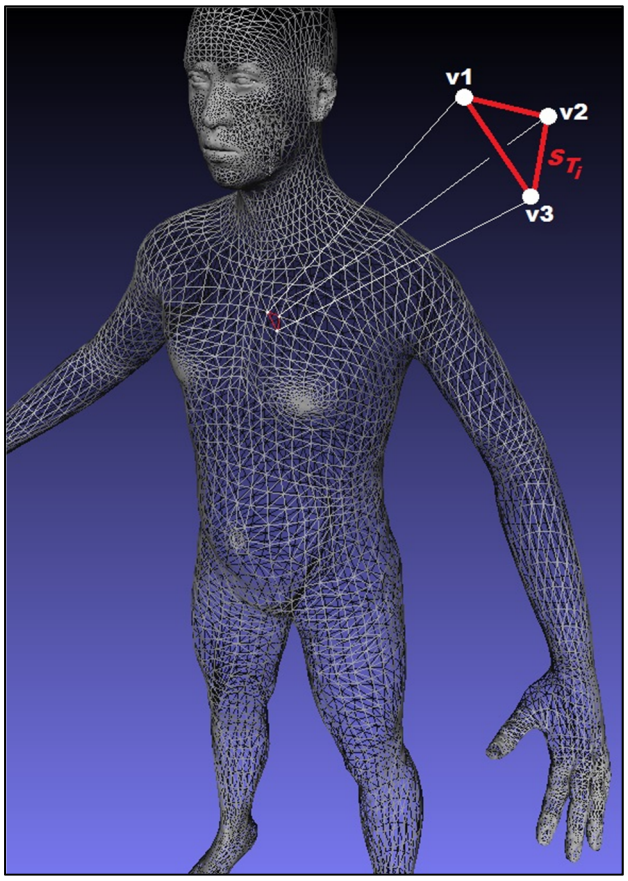


Figure 5.1 Example of the mesh model. STi is the surface area of 1 triangle and v1, v2, v3 its the adjacent vertices.

Using the computer mouse, the user can paint on the 3D model and therefore select the triangles which represent the BBSA. The BBSA is obtained using Equation 5.1 for the painted triangles. The Non-BBSA is obtained using Equation 5.1 for the non-painted triangles.

$$S = \sum_{1}^{n} S_{T_i} \tag{5.1}$$

$$\mathrm{S}_{T_{\mathrm{i}}} = \sqrt{\mathrm{P}_{T_{\mathrm{i}}}(\mathrm{P}_{\mathrm{T_i}} - \mathrm{A}_{T_{\mathrm{i-v1v2}}})(\mathrm{P}_{\mathrm{T_i}} - \mathrm{B}_{T_{\mathrm{i-v1v3}}})(\mathrm{P}_{\mathrm{T_i}} - \mathrm{C}_{T_{\mathrm{i-v2v3}}})} \tag{5.2}$$

The BSA is obtained by the sum of the BBSA and non-BBSA. TBSA is obtained by dividing BBSA by the BSA.

$\mathrm{T_{i-v1}} = (\mathrm{xT_{i-v1}}, \mathrm{yT_{i-v1}}, \mathrm{zT_{i-v1}})$,
$\mathrm{T_{i-v2}} = (\mathrm{xT_{i-v2}}, \mathrm{yT_{i-v2}}, \mathrm{zT_{i-v2}})$,
$T_{\mathrm{i-v3}} = (\mathrm{xT_{i-v3}}, \mathrm{yT_{i-v3}}, \mathrm{zT_{i-v3}})$

represent the three coordinate sets of the adjacent vertices of the triangle $\mathrm{T_i}$.

$\mathrm{A}_{\mathrm{T_{i-v1v2}}}$, $\mathrm{B}_{\mathrm{T_{i-v1v3}}}$ and $\mathrm{C}_{\mathrm{T_{i-v2v3}}}$ represent the lengths of each side of the triangle $\mathrm{T_i}$.

$\mathrm{P_{T_i}} = \frac{(\mathrm{A}_{T_{\mathrm{i-v1v2}}} + \mathrm{B}_{T_{\mathrm{i-v1v3}}} + \mathrm{C}_{T_{\mathrm{i-v2v3}}})}{2}$ represents the semi perimeter of triangle $\mathrm{T_i}$.

5.4.2 Importance of personalizing 3D models

To assess TBSA accurately with a 3D personalized model of a burn patient, the cross-platform open-source software MakeHumanTM (MH) of Bastioni et al., (2008) was used. Based on a technical and artistic study of morphological features of the human body, this software includes a system for simulating disproportion of fat and muscle, allowing realistic modeling of a whole range of human anatomical shapes (see Figure 5.2).

The initial androgynous 3D model can be deformed into a realistic human model based on simple and intuitive parameters such as age, sex, height, and anthropometric measures. One of the additional features of MakeHuman software allows to randomly vary these parameters, thereby generating random 3D models.

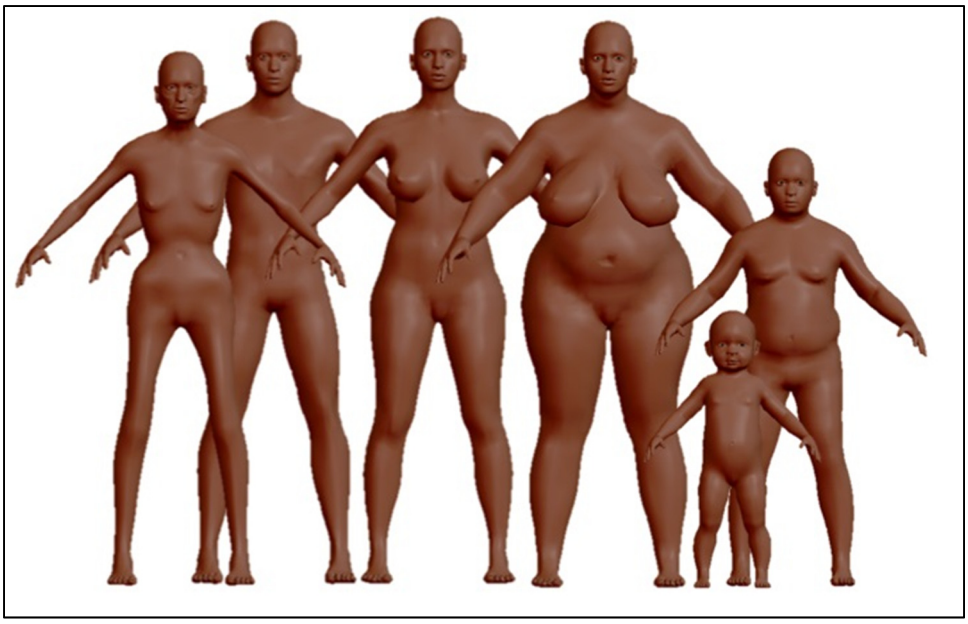


Figure 5.2 Examples of some three-dimensional models created with MakeHuman™ software

This study assumed that the typical features of the human morphology (e.g. android and gynecoid obesity, large breasts, fat disproportionality, etc.) can improve the assessment of TBSA. Indeed, as it can be seen in the Figure 5.3, the modeling algorithm of BurnCase 3D (Johannes Dirnberger et al., 2004a) and BAI (Prieto et al., 2011) are based on anthropometric formulas such as Equation 5.3 below.

$$BSA = C \times W^{A} \times H^{B} \quad (5.3)$$

A, B and C are constants defined by authors; W = weight; H = Height

The predefined 3D base model is first uniformly stretched along the longitudinal axis to fit the body height. Then, the model is adapted by expanding the surface along the normal vectors of

the mesh polygons until the approximated BSA is reached. Because the assessment of TBSA mainly depends on the BSA approximated with the Equation 5.3, the problem is that two human bodies with different morphology can produce two similar 3D projections (same BSA). To overcome this issue, the proposed approach suggests using MH software which locally expands its base model surface for each body parts depending on anthropometric measurements (see Figure 5.3). The TBSA assessment depends therefore on the BSA but also the morphology which allows for more accuracy. To verify this hypothesis, 15 paired models were first randomly generated. Paired models include two models with the same BSA but a different morphology (Figure 5.4), which allows isolating the morphology's influence from the BSA. The random modeling algorithm of paired models follows these conditions: a first model is randomly generated. Its BSA is calculated with the Equation 5.1 mentioned above. Then, a second model is randomly generated. Its calculated BSA must be equal to the BSA of the first model with an error margin of ± 0, 01 m². As well, its morphology must be different from the morphology of the first model. In total, 30 models (15 paired models), 19 men and 11 women models were randomly generated (see Tableau 5.1).

Then, to quantify the impact of the body morphology on the TBSA and Parkland fluids, six consistent burn patterns were drawn manually using the 3D rendering interface mention above. The six burn patterns were drawn by a plastic surgeon and covered TBSA from 22% to 60% (see Figure 5.5). The six burn patterns (Figure 5.5) were applied automatically and without control on the whole set of the 30 3D models.

clicours.com

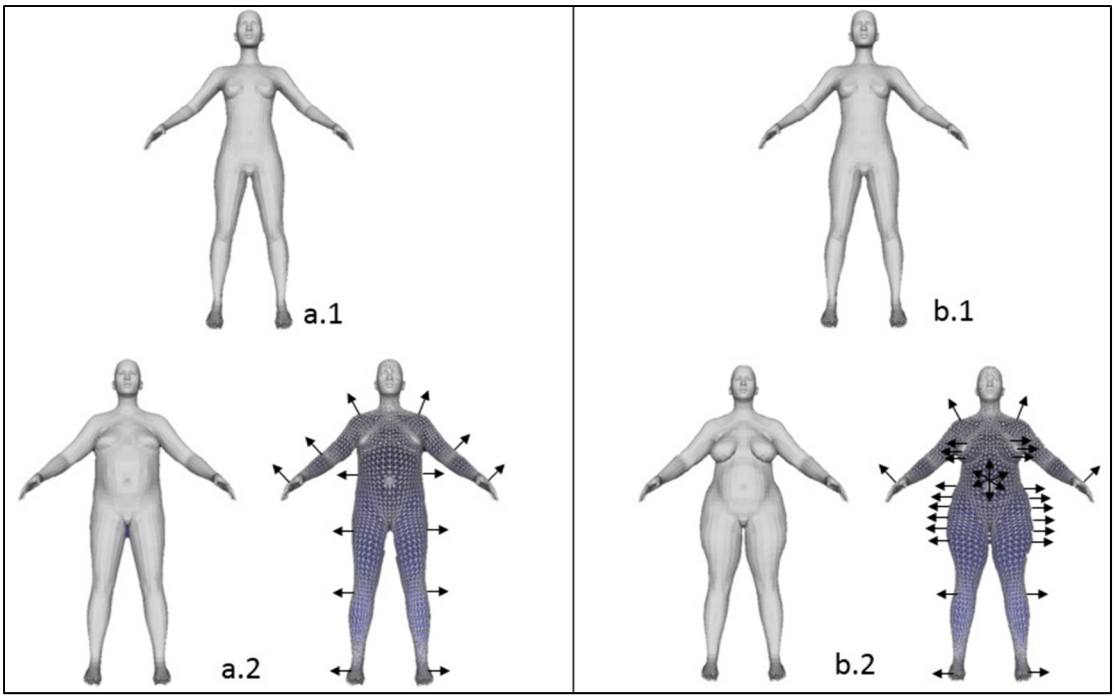


Figure 5.3 Difference between existing methods for 3D assessment of the TBSA and the proposed method. a.1 and b.1, Example of the base model. a.2, expansion algorithm of the model proposed by BurnCase3D (Giretzlehner et al., 2004) and BAI (Prieto et al., 2011). b.2, Local expansion algorithm of MakeHuman software (Bastioni et al., 2008) suggested in the proposed method. The represented models were built with the MakeHuman software and serve only as an example.

Finally, for 2 paired models, the difference between both of BBSA and TBSA were estimated. Additionally, the differences in Parkland fluid requirements (Equation 5.5) were estimated. This was estimated for the whole 15 paired models.

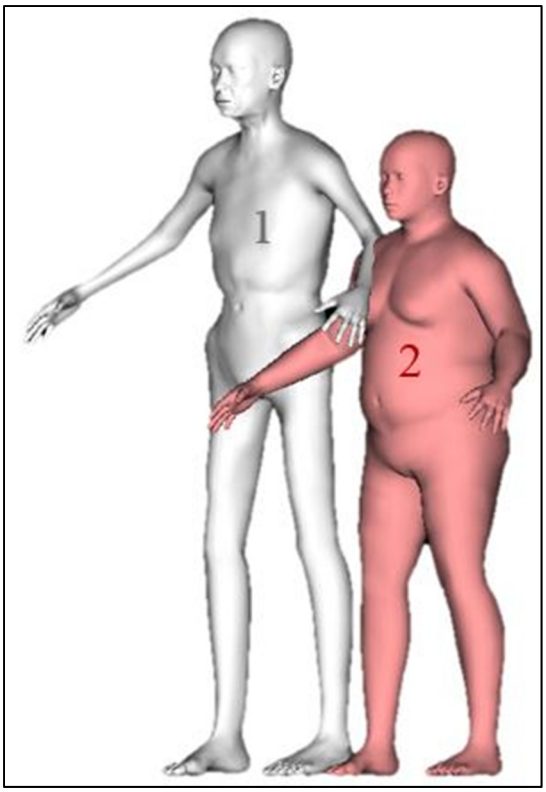


Figure 5.4 Example of paired models. The two models have the same BSA but a different morphology

Tableau 5.1 Summary of the paired models

Characteristic	**Value**
Age (years)	47 ± 29
Height (cm)	166, 73 ± 18,14
Weight* (kg)	62 ± 19
TBSA (m^2)	1,7 ± 0,3

Data presented as means ± SD.
*Weight is the approximated weight with Equation 5.3 and 5.4

Clicours.COM

$$* W = \left(\frac{BSA}{C \times H^{B}}\right)^{1/A} \quad (5.4)$$

$$Fluids\ requirements = 4\,ml\ \times TBSA\ (\%) \times Weigh\ (kg) \quad (5.5)$$

Because of the impossibility to quantify virtually the weight (which depends on the mass of the water, bones, muscle, fat, and so on) of the paired of models, the weight was approximate using Equation 5.3 and 5.4 based on the study of Tikuisis et al., (2001). This provides an overview of the impact of the variation of morphology on the fluids requirements.

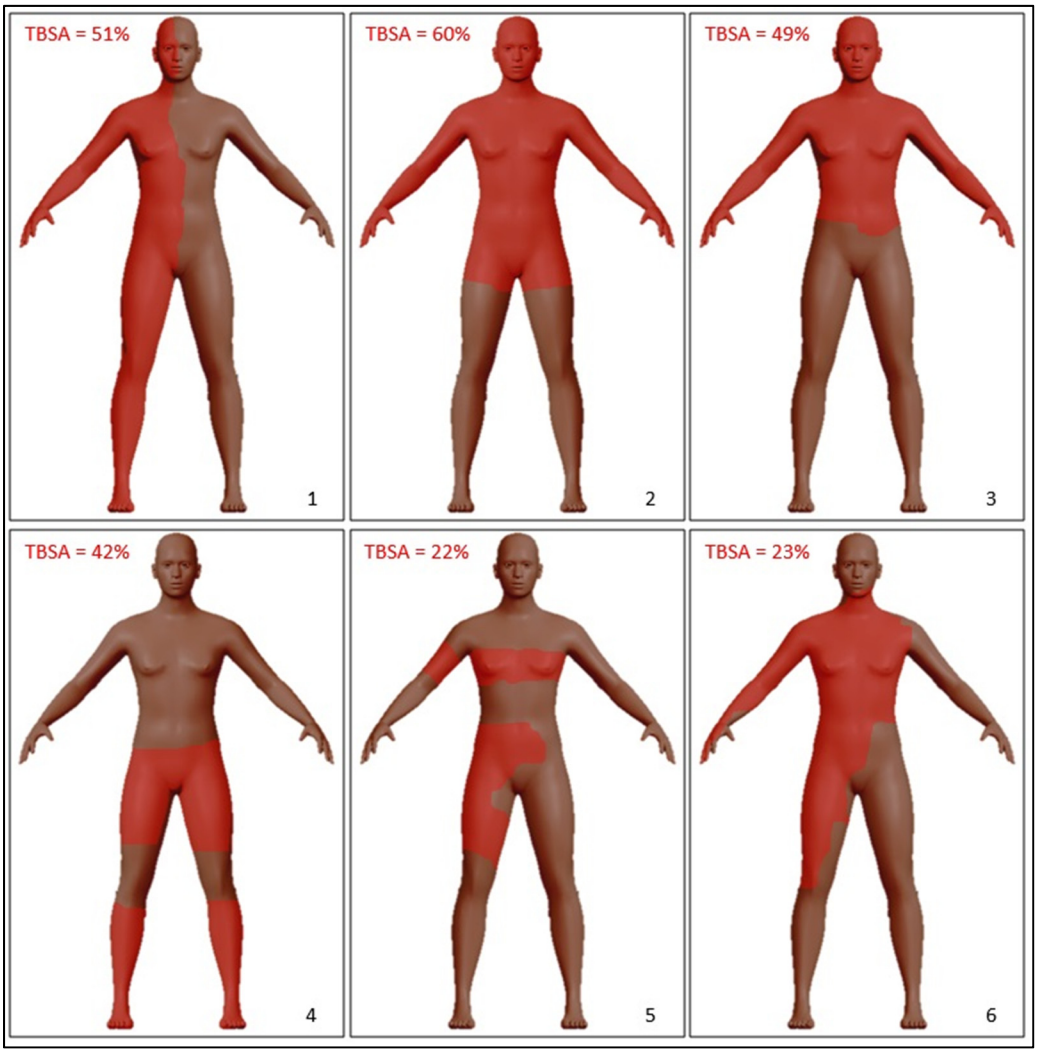


Figure 5.5 Burn patterns used in this study.
In red the Burned Body Surface Area (BBSA)

The mean and standard deviation of the differences of BBSA, TBSA and Parkland fluids were automatically calculated with the 3D rendering interface regardless intra-inter variability and transposition error.

5.4.3 Preliminary validation of the accuracy of the 3D models built with MakeHuman Software

A preliminary validation study was made to assess the accuracy of the models built from anthropometric measures. First, 19 anthropometric measurements were manually taken on 4 commercial mannequins (Figure 5.6) following a protocol. The protocol was based on (ISO, 2008) and measurer's handbook of Hotzman et al., (2011) and Gordon et al., (2013).

The 19 measures taken have been logged in MH software to generate their respective 3D models. The accepted standard (which represents the true value) has been established by scanning the commercial mannequins (Figure 5.6) with the white light portable 3D scanner Go!ScanTM of CREAFORM, Lévis, Québec, Canada. This scanner allows a resolution of 100 µm. The BSA of the five 3D scans and the five 3D models was obtained by the 3D rendering interface mentioned above.

The error percentage (Equation 5.6) was estimated for each mannequin.

$$\%\ Error = \left| \frac{BSA\ model - BSA\ scan}{BSA\ scan} \times 100 \right| \tag{5.6}$$

Using Statgraphics Centurion XVI software, a paired-samples t-test was performed to compare the mean BSA of the 3D scans to the mean BSA of the 3D models built with M.H software.

5.4.4 Preliminary validation of the proposed approach

Another preliminary validation study was made to assess the accuracy of the TBSA assessment made with the 3D rendering interface against the clinical method, the Lund & Browder chart. Each method was compared with the gold standard of the TBSA and Parkland fluids of each mannequin.

First, patches with a known surface area (m2) were put on the same commercial mannequin to simulate burns (Figure 5.6 left). Using the known BSA (m^2) of the mannequin's 3D scans (Figure 5.6 right), the gold standard of the TBSA (%) was estimated for each mannequin. Gold standards of Parkland fluids (mL) were also estimated using Equation 5.5 with the known approximated weight of the mannequin using Equation 5.4.

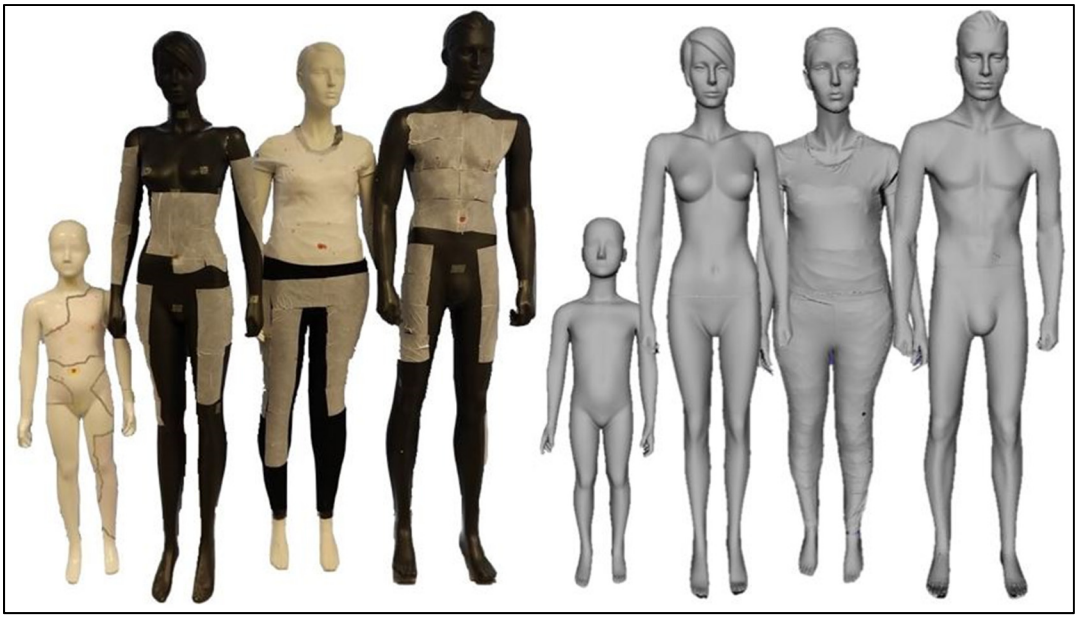


Figure 5.6 On the left, the 4 commercial mannequins used with patches represented simulated burns. On the right their respective 3D scans

Then, using the same 3D models of the mannequins previously built with MH software, 14 volunteers (8 experts and 6 non-experts) estimated the TBSA first using Lund & Browder chart and then drawing on the 3D MH models with the 3D rendering interface. A modified Lund &

Browder chart was used for the child mannequin. A training to assess the TBSA on a Lund & Browder chart was given to the non-experts. The 3D rendering interface automatically calculated the BSA, BBSA and TBSA in real time. This information was hidden to the volunteers.

The error percentage for the TBSA assessment was calculated using Equation 5.7 to measure the accuracy of each method.

$$\%\ error\ TBSA = \left|\frac{True\ TBSA - TBSA\ estimated}{True\ TBSA} \times 100\right| \tag{5.7}$$

$$\%\ Gain\ TBSA = \left|\frac{TBSA\ method\ - TBSA\ clinic}{True\ clinic} \times 100\right| \tag{5.8}$$

Paired-samples t-test has been made with Statgraphics Centurion XVI software to assess if there was a significant difference for TBSA assessment between both methods. The gain was calculated using Equation 5.8 to measure the error reduction in the TBSA assessment when using the proposed approach over the clinical method.

Finally, paired-samples t-test has been made to compare TBSA assessment made with the 3D rendering interface in experts and non-experts conditions.

5.5 Results

5.5.1 Burn representation and calculation of the TBSA

The 3D interface (Figure 5.7) allows the user to upload a 3D model to standard 3D formats, draw the burns interactively using the mouse and obtain the resulting BSA, non-BBSA and BBSA in square meter (m²) and the TBSA in percentage (%). The BBSA is represented in red. The Non-BBSA in blue. The BSA is obtained by the sum of the BBSA and non-BBSA. TBSA is obtained by dividing BBSA by the BSA.

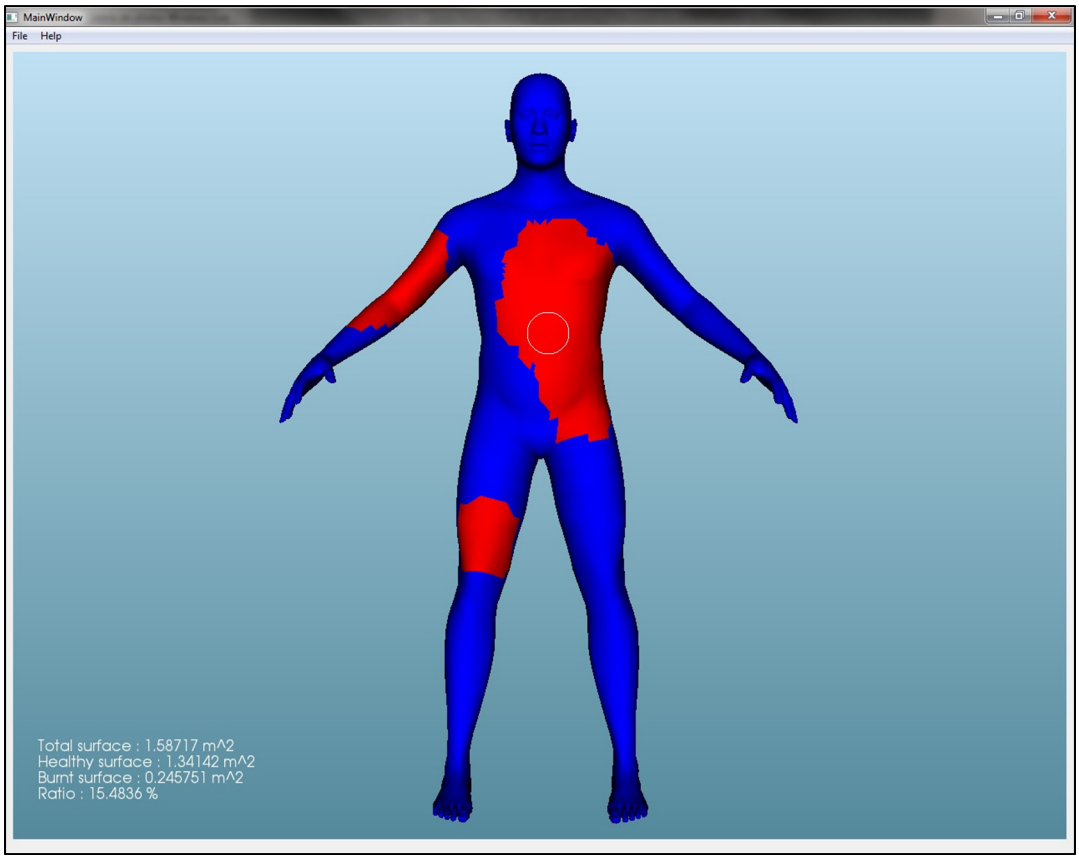


Figure 5.7 Example of the 3D rendering interface

The user has access to various intuitive tools, such as paintbrush to paint or to un-paint the burns; translation, rotation, and scaling of the model along the three axes; and representation of the model as a surface or a mesh for better visibility. The latest release is stable on Windows and Macintosh.

5.5.2 Importance of personalized 3D models

For all the burn patterns, the mean and standard deviation of the differences between each paired models were rather small, which was not the case of the TBSA. Only the burn pattern 1 showed small variation. For all burn patterns, noticeable differences were observed. For each paired models with the same BSA but a different morphology, small local variation in the morphology (characterized by the BBSA) influenced the TBSA and by consequence, the estimated Parkland fluids. The results are listed in the Tableau 5.2.

Tableau 5.2 Mean and standard deviation of the difference of BBSA, TBSA and Parkland fluids between each paired models

Burn patterns	BBSA (m^2)	TBSA (%)	Fluids (mL)
1	0,00 ± 0,01	0,27 ± 0,2	1271 ± 1190
2	0,04 ± 0,03	1,98 ± 1,40	1746 ± 1485
3	0,05 ± 0,04	2,73 ± 2,15	1537 ± 1162
4	0,05 ± 0,04	2,99 ± 2,14	1257 ± 1009
5	0.05 ± 0.03	2.58 ± 1.76	834 ± 740
6	0.02 ± 0.01	1.20 ± 0.72	618 ± 659

Data presented as mean ± SD.

5.5.3 Preliminary validation of the accuracy of the 3D models built with MakeHuman software

Results (see Tableau 5.3) have shown an average error percentage of 3 ± 2.2 % for the BSA of the MakeHuman models (M= 1.53, SD = 0.45) in comparison with the BSA of the scans (M = 1.51, SD = 0.48). A paired-samples t-test was conducted to compare the BSA in scans and MakeHuman models. There was not a significant difference between scans and MakeHuman models (p=0,4).

Tableau 5.3 Summary of the BSA and error percentage of the scans and MH models

Mannequins	Size (cm)	Weight (kg) *	BSA Scanner (m^2)	BSA MH (m^2)	Error (%)
Child	117	19	0.82	0.87	6
Woman	175.8	50	1.59	1.60	1
Overweight Woman	172	61	1.72	1.75	2
Man	187.3	76	1.99	1.92	3

*The Weight is the approximated weight with Eq 3 and 4.

5.5.4 Preliminary validation of the proposed approach

The gold standards (true values) of the TBSA (%) and estimated fluids for each mannequin are listed below in Tableau 5.4.

Tableau 5.4 Accepted standard for the surface areas of the patches and estimated fluids for each mannequin

Mannequins	BBSA (m2)	BSA (m2)	TBSA (%)	Fluids (mL)
Child	0.2	0.8	23	1773
Woman	0.4	1.6	23	4534
Overweight woman	0.6	1.7	33	7962
man	0.5	1.9	26	7925

BBSA: Burned Body Surface Area. BSA: Body Surface Area.
TBSA: Percent Total Body Surface Area assessment.

On average, TBSA and Parkland fluids estimation (see Tableau 5.5) with the Lund & Browder chart presented a greater variability among observers with a higher error percentage. Whereas TBSA and Parkland fluids estimation by drawing on the 3D models with the 3D rendering interface showed a lower variability and a lower absolute error (see Tableau 5.6).

Tableau 5.5 Summary of the Percent Total Body Surface Area (TBSA) assessment with the Lund & Browder chart

Mannequin	Experts	Non-experts	Estimated Fluids (mL)	Assessment Error (%)
Child	31 ± 3	29 ± 5	2272 ± 297	30 ± 12
Woman	34 ± 11	31 ± 4	6509 ± 1674	44 ± 37
Overweight woman	40 ± 6	41 ± 6	9865 ± 1352	24 ± 17
man	37 ± 7	34 ± 4	10853 ± 1835	37 ± 23

Data presented as mean ± SD or percentage (%)

Paired-samples t-test showed a significant difference between TBSA assessment with Lund & Browder chart and with the 3D rendering interface for each mannequin ($p < 0,05$). On average, the proposed approach allows a reduction error in TBSA assessment of 86% (see Tableau 5.6).

Tableau 5.6 Summary of the Percent Total Body Surface Area (TBSA) assessment with the three-dimensional (3D) rendering interface

Mannequin	Experts	Non-Experts	Estimated Fluids (mL)	Assessment Error (%)	Gain (%)
Child	23 ± 2	24 ± 2	1786 ± 142	6 ± 5	80 ± 58
Woman	22 ± 1	22 ± 1	4443 ± 256	4 ± 4	91 ± 89
Overweight woman	32 ± 0	32 ± 3	7829 ± 462	4 ± 4	83 ± 76
man	25 ± 1	25 ± 1	7637 ± 354	4 ± 4	89 ± 83

Data presented as mean ± SD or percentage (%).

Paired-samples t-test was conducted to compare the assessment of the TBSA with the 3D rendering interface in experts and non-experts for each mannequin. There was not a significant difference in the scores for expert and nonexpert conditions for $P > 0,05$.

5.6 Discussions and conclusion

The 3D rendering interface presented in this study showed that it is easily possible in a simple and user-friendly manner to accurately assess the TBSA (%) onto a 3D model.

Results obtained in section 5.5.2, point out the importance of taking the difference in human morphologies into consideration and so, looking for a 3D personalization of the models in the accurate assessment of TBSA. As shown in Tableau 5.2, for the same BSA, small variations in the local body surface area (due to different morphology) can induce an important clinical variation in the TBSA and Parkland fluids assessment. This is the actual improvement that misses from the current suitable 3D tools for patients with major burn injuries such as BurnCase 3D (J Dirnberger et al., 2002; Johannes Dirnberger et al., 2004b; Johannes Dirnberger et al., 2012; Parvizi et al., 2016) and BAI (Prieto et al., 2011). In most case, the BSA is overestimated with a tendency to increase the size of the body (Bailey & Briars, 1996). An underestimation of the BSA of the obese person (Lee et al., 2008), is also observed up to 20% (Tucker & Alexander, 1960). This can underestimate or overestimate morphology and therefore leads to an inadequate fluids estimation. BAI's predefined obese 3D models (Prieto et al., 2011) are insufficient to cover the variability of human morphology, specifically when

the fat is disproportionally accumulated among obese persons or women with larger sized breasts (Livingston & Lee, 2000; Hidvegi et al., 2004). As a matter of fact, the anterior trunk proportions of a woman's body can be altered by her breasts' size (Hidvegi et al., 2004). The same conditions are observed with the disproportion of the head of the children relative to their body. Moreover Wells et al., (2007), have demonstrated the significant association between body mass index and chest and waist in men and bust and hips in women. These issues had led to the aim of the proposed approach that is to personalized 3D models based on the anthropometric measurements of the patients. By doing this, it is possible to reach the typical features of the human body shape. This can make a difference not only in accurate morphology modeling but also the BSA and therefore the TBSA. It is, even more, striking with the study of Klein et al., (2007),which has highlighted the medical and surgical complication that occur when fluid requirements assessment are under or overestimated by 1,5L. The risks are the development of multi-organ failure, bloodstream infections, pneumonia, respiratory distress syndrome and death.

Results in section 5.5.3 have shown that MH 3D models are sufficiently accurate. The greatest error percentage obtained is observed for the child mannequin (Tableau 5.3). This can be explained by the fact that body shapes of the commercial mannequins do not correspond to real human body shape. Indeed, MH software takes correlation between body part and disproportion of morphology (e.g. the disproportion of the child's head as mentioned above) into consideration. The values of MH anthropometric measures are directly correlated among themselves allowing realistic 3D models. The difference of body dimensions does not allow to exactly build the respective 3D models of the mannequins. This limitation will be overcome later by validating the MH models directly to real subject scans. Nevertheless, these preliminary results allow seeing the suitability of MH software for this study.

Results in section 5.5.4 show a net decreased of the error percentage and variability in the estimation of the TBSA for the proposed approach against the Lund & Browder chart regardless of the level of expertise (experts or non-experts). This confirms the relevance and interest of a new 3D computer-based framework for evaluating TBSA, incorporating anthropometric measurements, MakeHuman (MH) software, and a 3D rendering interface. As

stated by Kamolz, Wurzer, Kamolz, Wurzer, Giretzlehner, Parvizi, et Lumenta, (2014), there is a need for objective burn size estimation such as 3D computer-aided system. A computer-aided system will help to not only compensate TBSA estimation errors but also provide an objective tool for reducing variability, (Giretzlehner et al., 2013), Parvizi et al., (2016). An approach that focuses on personalizing 3D model is highly recommended, especially when we know that many clinical scores for burn patients (Abbreviated Burn Severity Index, R-Baux, etc.) use also TBSA (%) for evaluating the patient's chances of survival or death. A slight change in the results using those scores can influence the level of care and the nature of the treatment provided, (Ryan et al., 1998; Tahir, Memon, Kumar, & Ali, 2009; Kraft et al., 2012; Finnerty et al., 2016). An assessment of TBSA (%) personalized would allow a better division of severe burned patients that really need to be taken care of.

This study presents several limitations. First, the constraint of using a triangle representation of a 3D model made the burn surface areas being roughly represented. This can hinder the TBSA assessment.

Second, both validation of the 3D models of MH and the whole proposed method were made using 19 anthropometrics measurements. In a clinical practice, using only 4 to 5 anthropometrics measurements is feasible. Moreover, only the assessment of the TBSA was studied. The whole process of taking anthropometric measures on the patient, entering them into a software for the modeling of the 3D model and finally assessing the TBSA with the rendering interface was not evaluated.

In order to overcome these limitations, further work is intended to be done. First, to identify a limited set of the most influent, feasible in clinical practice, repeatable and reproducible anthropometric measurements on the human morphology. Second, the integration of a new 3D rendering interface. The painting of the burn surface will be done on 3D textures instead of a triangle representation, which mean the painting will exactly follow the user's hand without constraint. Third, to validate the clinical study of the whole process of the method.

Clicours.COM

CHAPITRE 6

PREDICTIVE MODELLING FOR PERSONALIZED THREE-DIMENSIONAL BURN INJURIES ASSESSMENTS

Adrien Desbois, MS[1-2] , Florian Beguet, MS[1-2], Yannick Leclerc, MD[3], Angel Eduardo González Hernández, MD[4], Sylvie Gervais, PhD[1], Isabelle Perreault, MD[5], Jacques A. de Guise, PhD[1-2]

[1] École de technologie supérieure (ÉTS), Montréal (QC), Canada
[2] Laboratoire de recherche en Imagerie et Orthopédie (LIO) – Centre de recherche du Centre Hospitalier de l'Université de Montréal (CRCHUM), Montréal (QC), Canada
[3] Département de médecine générale, Université de Montréal, Montréal (QC), Canada
[4]Département de médecine générale, Escuela Superior de Medicina, Instituto Politécnico Nacional México D.F., México
[5] Division de chirurgie plastique, faculté de médecine, département de chirurgie, Université de Montréal, Montréal (QC), Canada

Article soumis pour publication, Novembre 2018.

" This work has been submitted to the IEEE for possible publication. Copyright may be transferred without notice. "

6.1 Mise en contexte

Les résultats obtenus au CHAPITRE 4 introduisent la pertinence du logiciel MakeHuman et l'utilisation d'un moins 3 mesures anthropométriques afin d'explique à plus de 98% la variabilité de la BSA. Cependant, bien que le pourcentage de variabilité expliqué ne soit pas biaisé, la multicolinéarité présente dans les modèles de régression obtenus crée des difficultés importantes dans l'interprétation des effets entre variables et leurs pertinences (voir section 2.3.3.1). Cet article présente de nouvelles analyses de prédictions afin de surmonter cette contrainte et d'offrir une meilleure généralisation des résultats. De plus, les résultats obtenus au CHAPITRE 5 ont été réalisés avec un ensemble de 19 mesures anthropométriques (cas

idéal). Cependant, en contexte clinique d'urgence, un nombre limité de mesures est préférable (Exigence 01). Cet article adresse donc cette limite et répond aux objectifs 6 à 8, à savoir : déterminer les meilleurs sous-ensembles de mesures anthropométriques à partir d'un large échantillon stratifié aléatoire et valider la précision de la modélisation 3D quant à la BSA et morphologie.

6.2 Abstract

For patients with major burn injuries, an accurate burn size estimation is essential to plan appropriate treatment and minimize medical and surgical complications. However, current clinical methods for burn size estimation lack accuracy and reliability. To overcome these limitations, this paper proposes a 3D-based approach – with personalized 3D models from a limited set of anthropometric measurements – to accurately assess the percent total body surface area affected by burns. First, a reliability and feasibility study of the anthropometric measuring process was performed to identify clinically relevant measurements. Secondly, a large representative stratified random sample was generated to output several anthropometric features required for predictive modelling. Compared to regression analysis, machine-learning algorithms better assessed the importance and the subsets of anthropometric measurements for predicting the body surface area according to specific patient morphological features. Then, the accuracy of both the morphology and Body Surface Area of 3D models built from a limited set of measurements was evaluated using error metrics and maximum distances 3D color maps. Results highlighted the height and circumferences of the bust, neck, hips and waist as the best predictors for body surface area. 3D models built from three to four anthropometric measurements showed good accuracy and were geometrically close to gold standard 3D scans. Outcomes of this study aim to decrease medical and surgical complications by decreasing errors in percent total body surface area assessments, and therefore improving patient outcomes by personalizing care.

6.3 Introduction

During the initial management of a patient with major burns injuries, an accurate burn size estimation is essential to ascertain vital and functional prognoses (Durrant et al., 2008). Burn injuries involve dramatic fluid loss and must be replaced to maintain homoeostasis and ensure the patient's chances of survival. The amount of fluid resuscitation is function of the percentage of total body surface area burned, usually referred as percent TBSA by the clinicians, which is the ratio of the burned body surface area to the body surface area (BSA).

Clinicians use several pen-and-paper methods to assess PERCENT TBSA, such as the rule of nines, the Lund & Browder chart and the rule of palms. In the literature, these clinical methods have been associated with significant inaccuracies in burn size assessment and low inter-rater reliability (Jose et al., 2004; D. Parvizi et al., 2014). Identified factors causing the lack of accuracy include transposing 3D burns onto a 2D diagram, and using a diagram with fixed body proportions to assess patients with different morphologies.

Some authors have attempted to improve percent TBSA assessment through 3D representation of burns (Haller, Dirnberger, Giretzlehner, Rodemund, & Kamolz, 2009; Prieto et al., 2011). However, limitations to these assessment tools are: the limited set of 3D predefined models, and the modeling algorithm. The global shape of the predefined 3D model is expanded in order to match the BSA of the patient. Yet, the BSA is approximated with a formula using only weight and height (D. Du Bois & E. Du Bois, 1916; Tikuisis et al., 2001). This leads to an approximation of the patient's real BSA and, in consequence, an approximation of the percent TBSA. Likewise, data compiled using these formulas to approximate BSA were also used to build the clinical methods (C. C. Lund, 1944; Gehan & George, 1970).

This major drawback is due to the local shape of the 3D model, which was ignored. Desbois et al., (2017) have shown the importance of morphology modelling (local shape), which can greatly affect the ratio of the TBSA. This is striking given that percent TBSA overestimation can lead to excessive fluid contribution, resulting in the patient drowning, whereas

underestimation can trigger hypovolemic shock, causing a cardiovascular collapse (Wachtel et al., 2000). Both of these emergency conditions can result in the patient's death. Percent TBSA assessment error also leads to clinical and surgical complications. Furthermore, percent TBSA variability increases with overweight or obese patients (Rae et al., 2013; R. Y. Williams & Wohlgemuth, 2013), children (Livingston & Lee, 2000; Fairbrother, 2015) women with larger sized breast (Livingston & Lee, 2000; Hidvegi et al., 2004), or even pregnant women (Kennedy et al., 2008). Yet, obesity is a frequent clinical condition both in adults (P. T. James, 2004) and children (Han, Lawlor, & Kimm, 2010).

Although several approaches exist for 3D modelling of the human body, none of them sufficiently satisfy clinical requirements. The widespread approach estimates the human body shape by direct acquisition with structured light sensors (scanner or Kinect) or Time-of-Flight sensors. Both suffer from outliers and data loss, which leads to unavoidable post-processing. However, burn patients are most of the time in a lying position and unconscious, which would make direct acquisition somewhat impossible, even with an approach allowing to cover data loss (Allen et al., 2003). Moreover, many artefacts occur in this context: acquisition time, weak 3D surface overlapping, pose change, movement and breathing artefacts. Another approach consists in creating a statistical model from a database of 3D scans (Seo & Magnenat-Thalmann, 2003; Chu et al., 2010) as in current databases such as CAESAR (normal morphology civilians) or ANSURE (used in the military). While modelling is more robust, the resulting 3D model depends essentially on the database and does not always allow for various morphologies (Nadadur & Parkinson, 2010). Moreover, the morphological distribution is not representative for children, obese people or particular morphologies. Finally, parameterized modelling (Kasap & Magnenat-Thalmann, 2007; J. Li & Wang, 2007), is interesting since this approach allows a high level of control over the target model while maintaining the quality of the base model. Still, correlations between body parts are not always taken into account.

To overcome these constraints for 3D modelling and the lack of accuracy in burn size estimation, Desbois et al., (2017) have shown the relevance of using a 3D computer-based framework with personalized 3D models from anthropometric measurements to accurately

assess percent TBSA. 3D modelling is performed with the MakeHuman (MH) software (Bastioni et al., 2008), which allows matching typical features of the human body shape and different correlations between body parts. Indeed, based on more than 3,000 morphologies, MH performs accurate parametric modelling while maintaining body proportions and thus, the morphology. Despite observing a more accurate and reliable percent TBSA assessment against the clinical method, this study presented a limitation: 3D modelling was performed using 19 anthropometric measurements taken on healthy subjects. In a clinical setting, a limited but relevant and accurate set of measurements is preferable. Sets of anthropometric measurements must allow an accurate representation of the BSA and morphology. Moreover, current methods fail to evaluate the intra- and inter-rater reliability of the proposed set of anthropometric measurements.

This article addresses these limitations. First, the reliability and feasibility of the anthropometric measuring process in a clinical setting were evaluated. Then, a method for generating a large stratified random sample was implemented in order for predictive modelling to determine the best feature subset of anthropometric measurements for accurate 3D modelling. Data modelling and algorithmic modelling were tested and compared to ensure accurate, reliable and consistent results on which relies the %TBSA assessment and therefore the level of care received by the burn-injured patient. Finally, the accuracy of both the morphology and BSA of 3D models built from a limited set of measurements was evaluated and comparison with existing BSA formulas used in clinical settings was performed.

6.4 Methods

6.4.1 Reliability of anthropometric measurements in clinical settings

MH software allows 19 anthropometric measurements for personalized 3D modelling. In order to assess the inter-rater reliability and feasibility of the anthropometric measuring process, six clinicians manually applied the 19 anthropometric measurements proposed by MH to eight participants (see Tableau 6.1-A) using a clinical protocol. The protocol was approved by ethics committees, and also by participants and clinicians who all offered their consent. The

measuring process was repeated twice, and each time was conducted independently to avoid recall bias. The measurement protocol was based on ISO 7250–1 (ISO, 2008) and the measurer's handbook (Hotzman et al., 2011; Gordon et al., 2013) and was revised by clinicians to ensure proper understanding of the measuring process. The height measurement was intentionally withdrawn since it is already a reliable measure widely used in clinic settings.

The reliability of each anthropometric measurement was performed using the Six Sigma tool of Gauge Repeatability and Reproducibility (Gauge R&R) of the Statgraphics Centurion XVI software (version 16.1.05). The Gauge R&R study was performed using 6σ intervals in order to capture 99.7% of variations with 95% confidence intervals (95%CI). The acceptability of each anthropometric measure was identified by the percent total variation (%RR), repeatability which is the variation due to measure ($\%_{\mathrm{VDM}}$) and reproducibility which is the variation due to clinicians ($\%_{\mathrm{VDC}}$). The output of %RR gives a standardized acceptance criteria threshold for acceptability of the anthropometric measure as established by ISO13053-2 (ISO13053-2, 2011): the measure is considered acceptable for a %RR value lower than 30% (%RR < 30%) and a number of distinct categories (ndc) greater than four (ndc > 4). The ndc value represents how many 97% confidence intervals for the true value being measured can fit within the range of expected subject-to-subject variation (Woodall & Borror, 2008). The number of positive correlations (PC) with means and standard deviation of coefficient correlation (r) for each measure was assessed with Spearman's rank coefficient correlation using Statgraphics Centurion VXI software.

Tableau 6.1 Characteristics of study participants

#	Gender	Years	Height (cm)	Weight (kg)	BSA (m^2)	BMI Category
A	4 women 4 men	39 ± 14	166 ± 9	77 ± 25	1.8 ± 0.2	4 Normal weight 3 Overweight 1 Obese class III
B	4 women 12 men	31 ± 15	163.9 ± 14.5	71.7 ± 23.5	1.7 ± 0.3	1 Underweight 7 Normal weight 4 Overweight 3 Obese class I 1 Obese class III
C	4 women	32 ± 18	158 ± 17	70.1 ± 28.3	1.7 ± 0.4	4 Normal weight

	5 men					3 Overweight 1 Obese class I 1 Obese class III

Data presented as mean ± standard deviation. # represents a subsection of the table (e.g. Tableau 6.1 -A).

6.4.2 Generation of a stratified random sample

A major constraint for this study was the lack of a representative database with a good morphological distribution which would be sufficiently large for predictive analytics. Nadadur et Parkinson, (2010) suggest a stratified sample to better represent a population.

First, the following characteristics from Canada's population data were gathered (Shields & Connor Gorber, 2008; Enquête canadienne sur les mesures de la santé, 2015) : Age distribution by gender in percentage from 1 to 90 years; obesity prevalence according to age and BMI classification. The proportion and percentage were adapted for a sample of 50,000 subjects to support the complexity and diversity of human morphologies.

Secondly, from the MH software (MakeHuman [free and open source 3D computer graphics software], 2014) (version 1.0.2), a plugin (software component) was modified in order to generate a random sample of 50,000 3D models according to the above-calculated data. As illustrated in Figure 6.1 and Figure 6.2, the known proportions were integrated to the plugin as fixed input. The remaining body parameters for modelling were randomized. The output data was an excel file containing, for each generated 3D model, the gender, age, BMI, height, 21 anthropometric measurements and BSA. The BSA data of the 3D models was obtained by summing the areas of the 3D mesh's triangles as defined by Desbois et al., (2017).

Cliquez.com

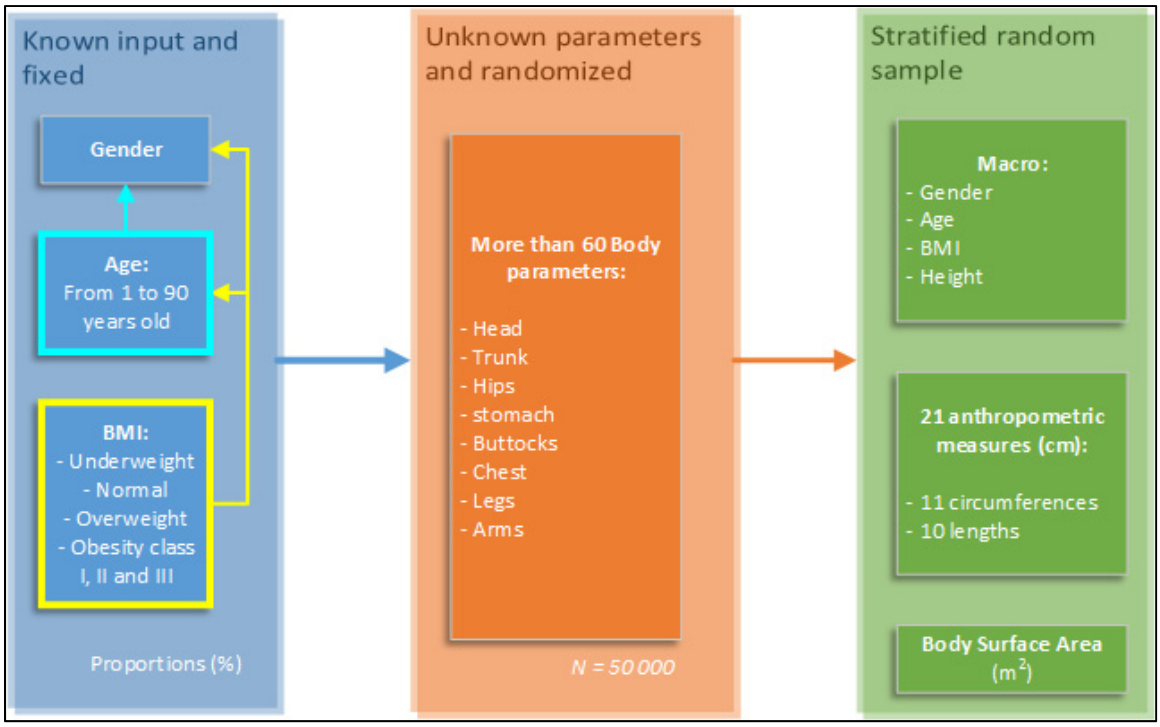


Figure 6.1 Construction method of the stratified random sample

To ensure a correct stratified random sample, a student t-test was performed using Statgraphics Centurion VXI software to compare the mean BSA of the randomized 3D model with the BSA of the data gathered from Canada's population for each age group. Since no BSA data was recorded for the Canadian population, it was approximated using the formula of Tikuisis et al., (2001) based on height and weight.

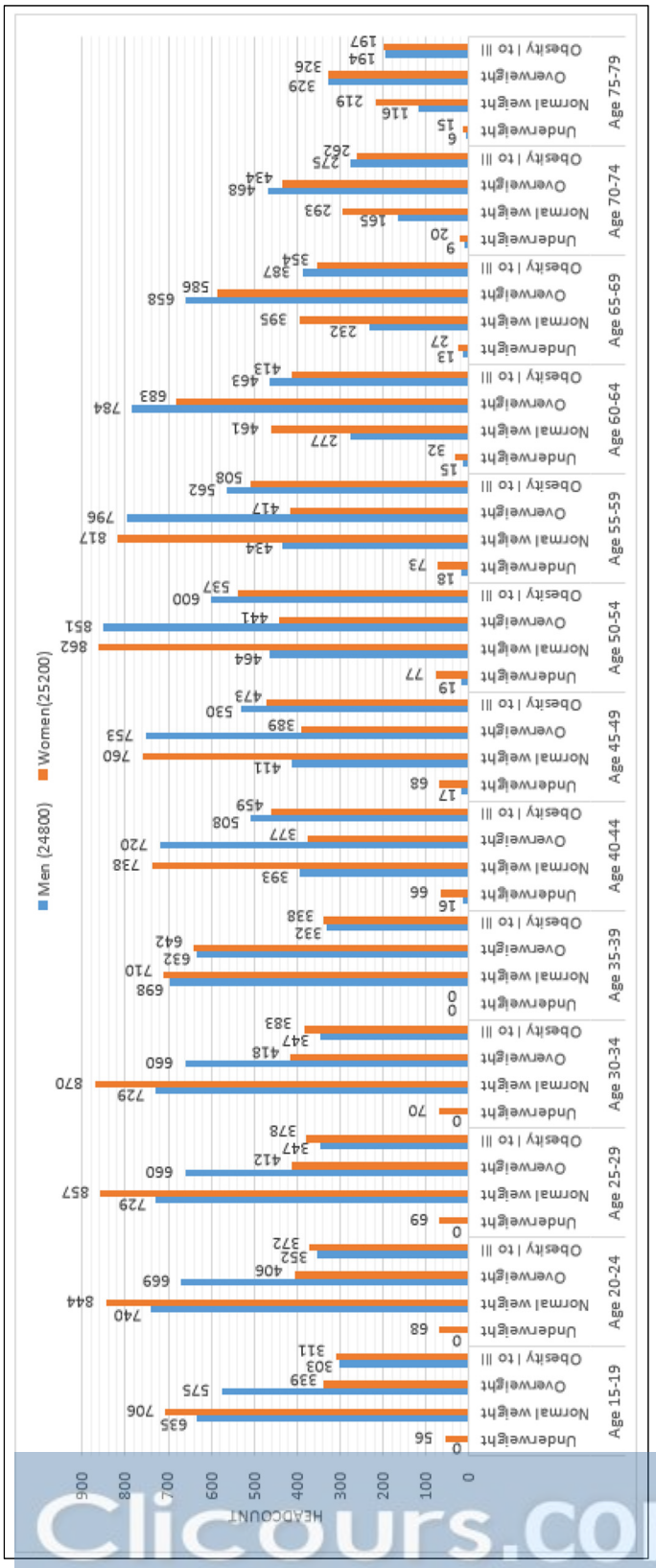


Figure 6.2 Obesity prevalence according to age and gender. Distribution age from 1 to 14 years (men=4116, women=3909) and from 80 to 90 years (men=793, women=1263) were not represented for graph since 3D models only had one BMI class: normal weight

Clicours.com

6.4.3 Selection of best feature subsets

In clinical practice with burned patients, a limited set of measurements is preferable since a patient's odds of survival critically depend on time. A team of experts (composed of plastic surgeons, trauma and general surgeons, emergency physicians and intensivists) suggested limiting the set of measurements to three or four anthropometrics measurements.

Hence, the aim of feature selection in this study is to find the most limited set of anthropometric measurements (features) satisfying clinical requirements and predicting at best the BSA (response variable) and morphology of patients. Feature selection not only improves the prediction, but also the estimation of correlation structure (De Silva & Leong, 2015). Indeed, while some measurements may be relevant and others redundant, once isolated, irrelevant measurements may become relevant when combined with other measurements (Guyon & Elisseeff, 2003; De Silva & Leong, 2015). In addition, a sufficient number of anthropometric measurements is required to characterize the morphology. Thus, features must be carefully selected according to the complexity and accuracy of predictive models.

At first, data modeling – which is more interpretable and accepted among clinician – was considered. Regression analysis, such as generalized linear models with best subsets selection (exhaustive search) or ridge regression, was used for feature selection. For the record, stepwise regression was excluded since it does not identify the correct set of anthropometric variables in instances of multicollinearity (Flom & Cassell, 2007; Morozova et al., 2015). Various predictive models had a high variance inflation factor, which highlighted the presence of multicollinearity. However, residuals studies showed that best predictive models still had heteroscedasticity (non-constant variance) and nonlinearity despite data transformation or nonlinear weighted regression resulting in questionable conclusions. This led to the selection of algorithmic modeling instead of data modeling, in which case no assumptions are needed, except for data drawn from unknown multivariate distributions. Algorithmic modeling is not sensible to multicollinearity and can capture linear as well as nonlinear relationships (Breiman, 2001b) but is less interpretable, and some algorithms could be seen as a "black-box".

Guyon et Elisseeff, (2003) and Morozova et al., (2015) suggested applying several feature selection algorithms and identifying the common feature between them. Following these recommendations, seven supervised machine learning algorithms were applied to the regression and feature selection problem using (RStudio Team, 2016).

Elastic-net (Enet) of the Caret package (Max et al., 2017) is a regularized method which linearly combines the L_1 norm of the lasso regression and L_2 norm of the ridge regression. The Random Forest (RF) ensemble learning method – from the "randomForest" package (Liaw & Wiener, 2002) – builds several fully grown decision trees (low bias but high variance) in parallel to random subsamples and merges them together to obtain more accurate and stable predictions. The trees are made uncorrelated to maximize the decrease in variance. On the other hand, Boosting tree (Boost) from the gbm package (Ridgeway, 2017), sequentially adjusts shallow decision trees (high bias but low variance) in order to generalize residuals which are reduced mainly by reducing bias (and to some extent variance, by aggregating the output from many trees). Support vector regression (svrRadial) of the Caret package (Max et al., 2017) can be formulated as a convex optimization problem by first defining a convex ε-insensitive loss function to be minimized and individualizing the hyperplane, which maximizes the margin. The cost function involves using a radial basis function kernel, which transforms the data into a higher dimensional feature space to enable nonlinear regressions. The most common type of neural network model, a feed-forward multi-layer perceptron (nnet) of the Caret package (Max et al., 2017), consists in an input and an output layer with one or more hidden layers. Each node is a neuron that uses a linear activation function, while the linear output of each node interacts in a nonlinear way. Learning is made by gradient backpropagation. Average neural network (avNNet) of the Caret package (Max et al., 2017) is also an ensemble learning method, which creates several feed-forward multilayer perceptrons and combines them in order to produce the desired output. Multivariate adaptive regression spline (MARS) of the Earth package (Milborrow, 2017) is a non-parametric regression algorithm which models multiple nonlinearities using hinge functions. A detailed description of the algorithms is beyond the scope of this article. However, for more details about each

algorithm and their parameters, please refer to the respective documentation of the R packages in reference. R script can be sent upon request.

Eighty percent of the dataset was allocated to the training set and 20% to the test set. All algorithms were performed using a 10-fold cross-validation for a good compromise between bias and variance (Borra & Di Ciaccio, 2010). Min-max normalization was applied to the data to improve convergence of the algorithms.

In order for the machine learning algorithms to yield accurate estimates, only the training data were used to perform all aspects of model-fitting, including feature selection. A validation set, which is a subset of the training set created by cross-validation, was used to evaluate model performance for each run/set of hyperparameter values and the final selection of the best predictive models. The test set was kept hidden from the algorithms and was only used after the training took place to compute some metrics that could provide hints on how algorithms behaved. For each error metric in the test set, values were predicted using the built model and compared against the real values. This ensured generalization of the best model to new data. Finally, predictive modelling performance was assessed by four error metrics: the root mean square error (RMSE), the mean absolute (MAE), the mean absolute percentage error (MAPE) and the coefficient of determination (R2). In predictive model m2, y_i is the real BSA, $\hat{y}_i$ is the predicted BSA, $\bar{y}$ is the mean BSA, and n is the sample size.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{6.1}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{6.2}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{|y_i - \hat{y}_i|}{y_i} \tag{6.3}$$

$$R^2 = 100\left(1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}\right)\% \quad (6.4)$$

RMSE (6.1) is a quadratic score measuring the average size of errors and penalizes important errors to give an overview of the error distribution. MAE (6.2) is similar to RMSE but easier to interpret and therefore comes in addition. MAPE (6.3) measures the relative error as a percentage. The closest these error metrics are to 0, the better the predictive model will be. R^2 coefficient (6.4) is the proportion of variance in the dependent variable (BSA) that is predictable from the independent variables (anthropometric measurements). Contrarily to the previous measurement, the closest it is to 100%, the better the predictive model is.

Anthropometric measurements were ranked by each machine-learning algorithm according to their importance determined by their influence on the related root mean square error. For ease of interpretation, importance score x of each algorithm within an observed min_{old} and max_{old} range of values was rescaled from range 0 (no influence) to 100 (strong influence) by the following formula:

$$x = \frac{x - min_{old}}{max_{old} - min_{old}} \times 100 \quad (6.5)$$

6.4.4 Body Surface Area accuracy of 3D models

In order to assess the accuracy of MH software-generated 3D models (MH 3D Models) based on three or four anthropometric measurements, the BSA of 16 participants (see Tableau 6.1-B) were used as a gold standard for comparison. The BSA of 10 men, four women, and two children were obtained through full-body scans, using a white light portable scanner (Go!Scan™ of CREAFORM). The protocol was approved by ethics committees and participants who offered their consent, while anthropometric measurements were gathered by clinicians as described previously in section A.

For each subject, the gold standard was compared to the BSA obtained from MH 3D models initially built from 19 anthropometric measurements, and then to those built from four and

three anthropometric measurements determined by the first two best feature subsets. The BSA of 3D models was obtained by summing the areas of the 3D mesh's triangles as defined by Authors (Desbois et al., 2017).

The BSA was also compared to existing anthropometric formulas (based on height and weight) of Gehan et George, (1970), Haycock et al., (1978), Mosteller, (1987) and Tikuisis et al., (2001). An error percentage (6.6) was estimated for each method based on the BSA of 3D models from 19, four, and three anthropometric measurements and the BSA from existing anthropometric formulas.

$$\%Error = \left| \frac{BSA_{method} - BSA_{scan}}{BSA_{scan}} \times 100 \right| \quad (6.6)$$

Using Statgraphics Centurion XVI software, a Kruskal-Wallis test and a Levene test were performed to compare medians and standard deviations of each method altogether.

6.4.5 Morphology accuracy of 3D models

As mentioned before, it is important to assess the accuracy of M.H morphology modelling (local shape) since it can greatly impact the percent TBSA (Desbois et al., 2017) . Hence, 3D scans of nine participants (Tableau 6.1-C) and their respective 3D MH models were used to assess distance between surfaces at any location. For each subject, 3D MH models were built with three and four anthropometric measurements according to algorithms with best performances (see Tableau 6.4 to Tableau 6.6) such as Boost_3 (height, and waist, neck circumferences), RF_3 (height, and bust, hips circumferences), avNN_4 (height, and bust, neck, waist circumferences), and RF_4 (height, and waist, bust, hips circumferences).

Unfortunately, a major constraint of these models was the difference of pose. MH models had legs and arms straight and apart whereas scan models had leg slightly apart with arms and hands on the head and the whole trunk twisted. First, as automatic registration can be tedious and not optimal (ICP registration has been tested unsuccessfully), two experts performed a manual registration using Blender software (2.78 version). MH models were registered on the

scan models using 3D rigging. Manual registration was performed twice in order to assess intra- and inter-rater reliability.

Secondly, surface distances between superimposed meshes were calculated using MeshLab software (1.3.3 version), and color-coded 3D models were generated. Thus, the distance between surfaces at any location was evaluated through quantitative metrics such as Average Symmetry Distance (ASD), Root Mean Square Distance (RMSD) and Hausdorff Distances (h_s), all expressed in centimeters (cm). The Hausdorff Distance unit calculations between corresponding points of the two pairs of meshes were visually complemented by generated color-coded models, which also allowed for a 3D visualization of surface distance.

For a 3D model, defined by a triangular mesh A with a set of points $a \in R^3$, being the vertex and another triangular mesh B with a set of points $b \in R^3$:

$$ASD = \sum_{a \epsilon A} (\min_{b \in B} \|a - b\|_2) \tag{6.7}$$

$$RMSD = \sqrt{ASD^2} \tag{6.8}$$

$$\begin{cases} h_s(A,B) = \max[h_u(A,B), h_u(B,A)] \\ h_u(A,B) = \max\limits_{a \in A}(\min\limits_{b \in B} \|a - b\|_2) \end{cases} \tag{6.9}$$

The ASD (6.7) is an average of all distances from a 3D model A to another 3D model B (Heimann et al., 2010; Gotra et al., 2015). Although $RMSD$ (6.8) and ASD are strongly correlated, $RMSD$ highlights significant outliers (Heimann et al., 2010; Laurent et al., 2016). Both highlight the reliability and accuracy of 3D surfaces. The h_s assesses global error between surface A and B. Using Statgraphics Centurion XVI software, a Kruskal-Wallis test and a Levene test were performed to compare medians and standard deviation of error metrics (ASD, $RMSD$, h_s) obtained by each evaluator.

6.5 Results

6.5.1 Reliability of anthropometric measurements in clinical settings

The Gauge R&R study allowed differentiating acceptable anthropometric measurements from non-acceptable ones (see Tableau 6.2). All lengths had a %RR greater than 30 (%RR > 30) with a ndc value lower than four (ndc < 4) and were therefore too variable to be clinically acceptable. The same applies to some circumferences such as the knee, wrist and ankle.

For all length measurements, the variation was mostly due to reproducibility with an average $\%_{VDC} = 63 \pm 22$ %. This means clinicians were not constant among themselves. However, each measure was more constant during the measuring process with an average $\%_{VDC} = 37 \pm 22$ %. On the contrary, measurements of the neck, under-bust, bust, waist, hips, upper arm, thigh and calf circumference had an acceptable variance with a %RR below 30 (%RR < 30) with a ndc value greater than four (ndc > 4) and were therefore clinically acceptable. For all these circumference measurements, the variation was considerably due to repeatability with an average $\%_{VDM} = 81 \pm 14$ %. This means clinicians were not constant within their own measuring process. That said, each measure was more constant among clinicians with an average $\%_{VDC} = 19 \pm 14$ %. Furthermore, these results allowed reducing the original dataset of 19 anthropometric measurements to eight clinically feasible, reproducible and repeatable measurements. Maximum measuring time was recorded at 15 min for the first time and went down to 12 min for the second time.

A significant number of positive correlations (PC) was observed among the measurements, associated with high or very high significant ($p < 0.05$) positive association (r). This highlighted the existing complex body proportion of the MH software along with possible multicollinearity within variables (anthropometric measurements). The fact that many variables can be correlated means it is not necessary to include them all in the 3D model. Moreover, some variables can confuse and hide different explicative factors of variation behind

the data (Bell & Wang, 2000; De Silva & Leong, 2015), which reinforces the purpose of the feature selection.

Tableau 6.2 Reliability and correlation coefficient of anthropometric measurements

Measurements	$\%_{VDM}$	$\%_{VDC}$	**%RR**	**ndc**	$PC(r)$
Neck circumference	84	16	20	6	15 (0.8 ± 0.08)
Shoulder length	27	73	97	0	7 (0.7 ± 0.08)
Chest length	14	86	64	1	20 (0.8 ± 0.1)
Under-bust circumference	91	9	17	8	15 (0.8 ± 0.16)
Waist circumference	70	30	11	12	15 (0.8 ± 0.19)
Upper leg length	1	99	84	0	9 (0.7 ± 0.08)
Knee circumference	73	27	33	4	15 (0.8 ± 0.15)
Upper arm length	46	54	87	0	19 (0.8 ± 0.06)
Wrist circumference	100	0	34	3	15 (0.8 ± 0.17)
Hips circumference	86	14	27	4	14 (0.8 ± 0.16)
Ankle circumference	47	53	40	3	15 (0.8 ± 0.15)
Upper arm circumference	56	44	21	6	14 (0.8 ± 0.16)
Lower arm length	46	54	67	1	9 (0.7 ± 0.08)
Bust circumference	84	16	20	6	15 (0.8 ± 0.16)
Waist to hips length	46	54	90	0	2 (0.5 ± 0.15)
Thigh circumference	77	23	28	4	14 (0.8 ± 0.16)
Calf circumference	100	0	25	5	15 (0.8 ± 0.14)
Lower leg length	73	27	82	0	17 (0.8 ± 0.07)
Nape to waist length	41	59	72	1	13 (0.7 ± 0.05)

All values are expressed within 95% confidence intervals. $\%_{VDM}$, $\%_{VDC}$ and %RR are expressed in percentages. $PC(r)$ are expressed as numbers of positive correlation (Mean ± Standard deviation of the correlation coefficient)

6.5.2 Generation of a stratified random sample

The generated sample preserves the same proportion of each heterogeneous group of the population as shown in Figure 6.2. Thus, the stratified sample is correctly generated since the sample mean is an unbiased estimate of the mean age distribution, obesity prevalence and BSA according to gender.

As well, a student t-test showed no significant difference between the mean BSA of the generated sample (BSA=1.9 ± 0.2 m^2 for men, BSA=1.6 ± 0.2 m^2 for women) and that of Canada's population (BSA= 2.0 ± 0.1 m^2 for men, BSA=1.8 ± 0.1 m^2 for women) for each different age group: $p > 0.05$, 95% CI = [-0.11; 0.27] for men and $p > 0.05$, 95% CI = [-0.05; 0.52] for women. The MH plugin utilized to generate these data can be shared upon request.

6.5.3 Selection of best feature subsets

First, results in Tableau 6.3 list the best subset of three to four anthropometric measurements according to data modeling analysis. When studying residual plots, only one regression model presented a goodness of fit and was obtained by a non-linear weighted least square regression. With a constant variance (homoscedasticity) and good linearity, this regression model (NonLinearWLS_3) had a RMSE=0.0004, MAE=0.04, MAPE=3.02%, R^2=99 and included height, and the circumferences of the chest and hips. All other regression models did not meet satisfactory requirements regarding goodness of fit and were therefore considered questionable.

Secondly, results in Tableau 6.4 list the best subset of anthropometric measurements according to algorithmic modeling limited to three anthropometric measurements (Tableau 6.5) or four anthropometric measurements (Tableau 6.6). Best predictive performance limited to three anthropometric measurements (see Tableau 6.5) was obtained for the Boosting tree (boost_3) with a RMSE=0.012, MAE=0.0001, MAPE=2.11%, R^2=99 and included height, and the circumferences of the waist and neck. The random forest (RF_3) had very close performance with a RMSE=0.013, MAE=0.0002, MAPE=2.19%, R^2=99 and included height, and the

circumferences of the bust and hips. The other algorithms had slightly lower performances but were equally adequate. When limited to four anthropometric measurements (see Tableau 6.6), RMSE and MAPE score slightly decreased resulting in a better prediction performance. Best predictive performance was obtained for the Averaging neural network (avNN_4) with a RMSE=0.011, MAE=0.0001, MAPE=1.70%, R^2=99 and included height, and the circumferences of the bust, neck and waist. The random forest (RF_4) was also very close in performance with a RMSE=0.011, MAE=0.0001, MAPE=1.82%, R^2=99 and included height, and the circumferences of the waist, bust and hips.

Tableau 6.3 Results of best features subsets selection ─ data modelling

Statistical analysis	**Goodness of fit?**		**RMSE**	**MAE**	**MAPE (%)**	**R2 (%)**	**Best feature subset**
NonLinear WLS_3	Good	Homoscedasticity and good linearity	0.0004	0.04	3.02	99	Height, Circumference of Bust and Hips
NonLinear WLS_4	Biased	Medium skewness	0.001	0.08	7.96	97	Height, Circumference of Bust, Waist and Hips
GLM_4	Biased	Medium heteroscedasticity and non-linearity (curvature)	0.088	0.07	5.92	99	Height, Circumference of Underbust, Thigh and Calf
GLM_3	Biased	Little heteroscedasticity and non-linearity (low-curvature)	0.095	0.07	6.01	99	Height, Circumference of Underbust and Hips

Best subset results are ranked by performance in a descending order (best performance at the top).

Clicours.COM

Tableau 6.4 Results of best features subsets selection ─ Algorithmic modelling

Prediction algorithm	Best feature subset
Boost	Height, Waist Circumference, Neck Circumference, Bust Circumference
RF	Height, Hips Circumference, Bust Circumference ,Waist Circumference
avNN	Height, Bust Circumference, Neck Circumference, Waist Circumference
svrRadial	Height, Bust Circumference, Neck Circumference, Waist Circumference
nnet	Height, Underbust Circumference, Neck Circumference, Bust Circumference
Enet	Height, Bust Circumference, Neck Circumference, Waist Circumference
MARS	Height, Bust Circumference, Thigh Circumference, Neck Circumference

Best feature subset for each algorithm with variable importance ranked in order. For example, the best subset for the "Boost" algorithm with three measurements includes: "Height, Waist Circumference, Neck Circumference"

Tableau 6.5 Comparison of predictive modelling performance for algorithmic modelling including 3 measurements

Prediction algorithm	RMSE	MAE	MAPE (%)	R^2 (%)
Boost_3	0.012	0.0001	2.11	99
RF_3	0.013	0.0002	2.19	99
avNN_3	0.015	0.0002	2.55	99
svrRadial_3	0.016	0.0003	3.54	99
nnet_3	0.017	0.0003	3.26	98
Enet_3	0.020	0.0005	8.02	98
MARS_3	0.040	0.0010	2.08	99

Best subset results are ranked by performance in a descending order (best performance at the top)

Tableau 6.6 Comparison of predictive modelling performance for algorithmic modelling including 4 measurements

Prediction algorithm	RMSE	MAE	MAPE (%)	R^2 (%)
avNN_4	0.011	0.0001	1.70	99
RF_4	0.011	0.0001	1.82	99
Boost_4	0.011	0.0001	1.92	99
svrRadial_4	0.011	0.0001	2.69	99
nnet_4	0.014	0.0002	2.29	99
Enet_4	0.020	0.0004	7.55	98
MARS_4	0.030	0.0008	1.56	99

Best subset results are ranked by performance in a descending order (best performance at the top)

Finally, Figure 6.3 showed the agreement between algorithms and the importance of some anthropometric measurements.

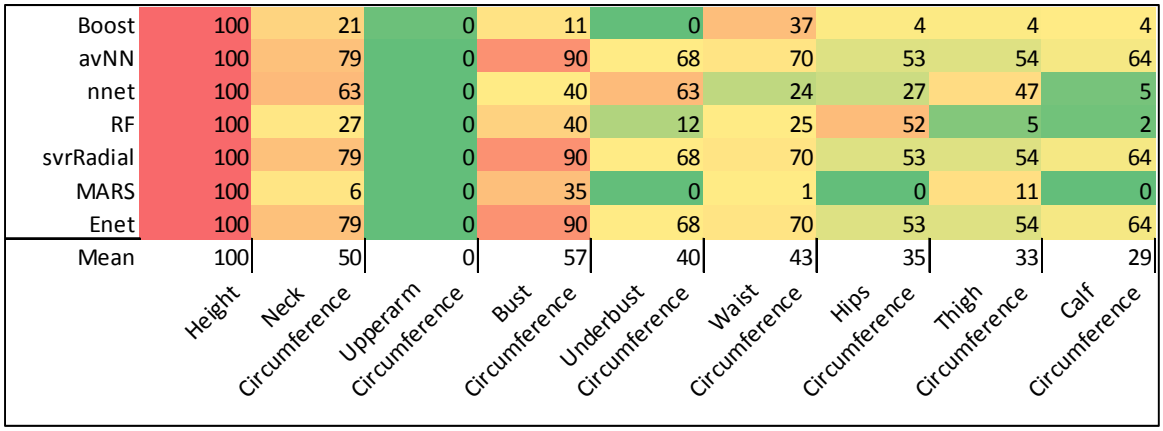


Figure 6.3 Importance of each anthropometric measurement for each algorithm. A score close to 100 (in red) indicates an important measure. A score close to 0 (in green) indicates a non-important measurement

Height was the most influent measurement on the BSA with an average score of 100 in agreement with all algorithms. Likewise, all algorithm agreed that upper arm circumference was not an important measure.

However, for other anthropometric measurements, there were differences among algorithms regarding their importance. On average, bust circumference came in second position with an average score of 57, followed by neck circumference in third position with an average score of 50. Waist circumference came in fourth position with an average score of 43.

6.5.4 Body Surface Area accuracy of 3D models

3D MH models built with 19 anthropometric measurements had the lowest error = 1 ± 1% (see Tableau 6.7). Subsets of three anthropometric measurements selected by boost and random forest algorithm (see Tableau 6.8) had an equal average error = 4 ± 3%, so did the anthropometric formulas of Gehan et George, (1970) and Haycock et al., (1978).

Subsets of four anthropometric measurements selected by avNN (see Tableau 6.4) had a decreased average error = 3 ± 2% which was equal of the anthropometric formulas of Mosteller, (1987) and Tikuisis et al., (2001). However, no improvement was observed for the random forest (see Tableau 6.7) with an average error = 4 ± 3%.

Yet, no statistically significant differences were observed when comparing medians (Kruskal-Wallis test, $p > 0.05$, 95% IC) and standard deviation (Levene test, $p > 0.05$, 95% IC) between gold standard BSA and 3D modeled BSA (regardless the number of anthropometric measurements) and anthropometric formulas (see Tableau 6.7).

Tableau 6.7 Surface difference with gold standard Scan

#	3D models built with MH from anthropometric measurements					Anthropometric formulas			
	All	**Boost_3**	**RF_3**	**avNN_4**	**RF_4**	**Gehan & George**	**Haycock**	**Mosteller**	**Tikuisis**
A	2	4	5	4	5	0	1	1	0
B	2	3	4	3	4	2	2	1	0
C	1	4	4	4	4	7	7	6	4
D	0	0	7	0	6	7	7	6	6
E	1	1	9	1	8	1	0	0	2
F	4	6	5	5	5	5	4	4	5
G	0	11	11	9	9	7	8	4	2
H	0	1	1	0	1	5	4	4	3
I	3	3	6	3	5	6	7	5	3
J	0	1	0	1	0	9	9	8	6
K	0	2	7	2	6	5	5	3	3
L	0	1	2	0	2	1	0	0	1
M	2	4	2	3	2	3	3	3	3
N	0	6	3	6	3	2	0	0	2
O	3	5	2	3	2	3	2	2	3
P	1	5	1	2	1	5	5	4	3

All values are expressed in error percentage (%) as defined by (6.6). For each column, the highest error percentage is represented in red and the lowest in green

6.5.5 Morphology accuracy of 3D models

Manual registration showed good intra- and inter-rater reliability among all 3D models built from four different subsets of anthropometric measurements (Boost_3, RF_3, avNN_4 and RF_4). From all 3D models, Hausdorff distances had an average $\%RR = 24 \pm 3\ \%$ with the total variability mostly due to repeatability $= 98 \pm 2\ \%$. Experts were consistent among themselves with good reproducibility $= 2 \pm 2\ \%$. To avoid redundancy, surface distances of only one subset (avNN_4, see Tableau 6.4) is displayed since surface distances of the other three subsets of anthropometric measurements were consistent. Indeed, there were no statistically significant differences between group medians and standard deviation of ASD, $RMSD$ and H_s among each subset as determined by the Kruskal-Wallis test ($p > 0.05$, 95%

IC) and by the Levene test ($p > 0.05$, 95%IC). Among subjects, two-sided surface distances had a very good $ASD = 1.1 \pm 0.4$ cm and $RMSD = 1.4 \pm 0.5$ cm with maximum surface distances $H_s = 7.4 \pm 1.1$ cm.

The color maps allowed to locate Hausdorff surface distances (H_s) between overlaid 3D models. The absence of any change received a Hausdorff Distance of 0.00 and was indicated in red, as shown in the MeshLab color scale (Figure 6.4). Maximum values were mapped from red to blue (presence of changes). A frequency density histogram allowed quantifying the number of total Hausdorff surface distances. Figure 6.4 shows four case study of subjects with the highest Hausdorff Distance (see Tableau 6.8).

Tableau 6.8 Two sided surface distances between 3D scans and 3D models built from anthropometric measurements (avNN_4)

#	h_s	ASD	$RMSD$
A	6.8±0.02	0.8±0.02	1.0±0.03
B	6.8±0.1	0.9±0.05	1.3±0.06
C	9.3±0.0	1.1±0.07	1.6±0.07
D	7.1±0.2	0.9±0.06	1.2±0.08
E	6.5±0.0	0.6±0.03	0.9±0.06
G	10.0±0.4	2.4±0.05	3.2±0.05
K	6.7±0.1	1.0±0.06	1.3±0.08
N	6.0±0.2	0.8±0.07	0.9±0.25

All error metric are expressed in cm. Results are presented as mean ± standard deviation.

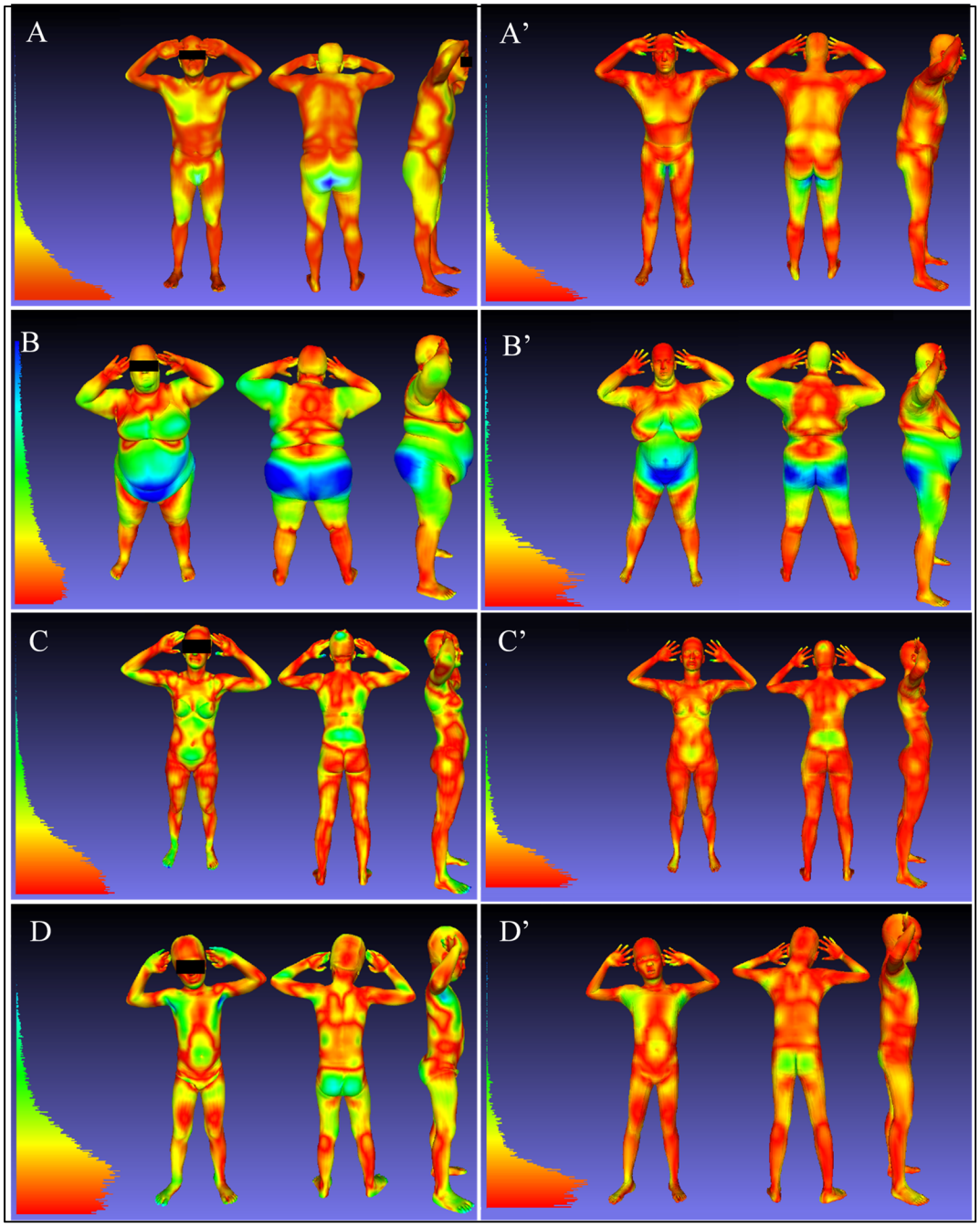


Figure 6.4 Color maps of subjects with maximum error distance (A, A', B, B') and lowest error distance (C, C', D, D'). Distance errors in 3D scans are represented on the left (A, B, C, D), while errors in their respective 3D models (built from 4 anthropometric measurements as determined by avNN_4) are represented on the right (A', B', C', D'). The blue color represents the maximum distance and the red color the minimum distance

The first case study showed an overweight man of 59 years old (see Tableau 6.1) with H_s ranging from 4 cm to 9.3 cm mostly located on the buttocks and on the genital area. Most of the frequency density was located on the trunk, arms and legs with a H_s ranging from 0 cm to 2.5 cm.

The second case study showed a woman of 59 years old with class III obesity (see Tableau 6.1). The highest H_s ranging from 7 cm to 10 cm was mostly located on the buttocks and lower part of the belly. Frequency density was more scattered with a H_s ranging from 0 cm to 6 cm and located on the upper part of the belly, chest, arms and legs. The third case study showed a 26-year-old woman (see Tableau 6.1). The highest H_s ranging from 0.27 cm to 0.39 cm was mostly located on hand and foot extremities. Most of the frequency density was located on the trunk, arms and legs with a H_s ranging from 0 cm to 0.25 cm. The fourth case study showed a 8-year-old boy (see Tableau 6.1). The highest H_s ranging from 0.24 cm to 0.37 cm was mostly located on the underarms and hand and foot extremities. Most of the frequency density was located on the trunk, arms and legs with a H_s ranging from 0 cm to 0.2 cm. For other subjects, color maps showed the highest H_s always with a minimal frequency density located on hand extremities, in between legs (genital area), on underwear or bathing cap. Since most of the main body parts showed good surface distances, results highlighted the morphology accuracy of 3D modelling.

6.6 Discussion

The study of inter- and intra-rater reliability and feasibility of the anthropometric measuring process highlighted the relevance of circumference measurements rather than length in a clinical context. Finding bony palpation landmarks needed for anthropometric measurements was difficult, and appropriate measurement sites were often hard to determine because of increased skinfolds and soft-tissue deformation, especially with overweight and obese subjects. Repeatability was less of a problem than reproducibility due to the learning curve. This was not the case with reproducibility, where measurements were not constant among clinicians. Therefore, it was not surprising that length measurements were not clinically acceptable since

they were less reproducible. On the contrary, circumference measurements were more reproducible and less repeatable, but could be improved over time with practice. Moreover, since the global body shape was already sufficiently adapted to height, adding other length measurements tended to overadapt the BSA and create disparities. Circumferences, on the other hand, were more responsible for the adaptation of local shape, therefore the morphology. The fact that many correlations were observed among anthropometric measurements means there were some interdependences where two variables or more provided more information together than individually.

As mentioned previously in the introduction, current databases lack good morphological distribution, represent a tedious process, and are very expensive (> 10 000$). With MH software, it was possible to generate a large stratified random sample according to fixed parameters. If needed, 3D models could be saved along with other parameters of interest. The generated dataset was representative of Canada's population, hence with accessible data, other datasets representative of specific populations can be generated in a free, simple and reproducible manner.

When considering predictive modelling, two approaches can be contemplated (Breiman, 2001b): data modelling (regression analysis) or algorithmic modelling (machine learning). As stated by Breiman, (2001b) the focus should be on the problem and not on the choice of the predictive model. Data presented a nonlinear distribution with relevant outliers (which characterize extreme populations such as children or obese people) and multicollinearity among variables. Data modelling was first considered for ease of interpretability and simplicity. One regression model met requirements for goodness of fit with a good score for error metrics. However, it did not satisfy the clinical requirement of offering various accurate subsets of anthropometric measurements in case a specific measurement cannot be taken on the patient. This led to choosing algorithmic modeling, which has proven to be suitable and robust, especially with data constraints (heteroscedasticity, nonlinear pattern and multicollinearity). With the same subset of predicted features, the best data model (NonLinearWLS_3) had a better RMSE score and was consistent with the best algorithmic

model (RF_3) which had better MAE, MAPE and R^2 scores. Freed from any assumptions and more robust, algorithmic modelling provided better information about the underlying data mechanism by finding the best subsets of anthropometric measurements allowing for the best BSA predictions. Most algorithms had similar errors, but some were slightly better due to their stochastic structure.

The best subset suggested height, bust, neck, hips and waist circumferences which were consistent with other studies. Indeed, Wells et al., (2007) studied a population of 9,617 adults (children excluded) with various BMI in order to highlight the most influent measurements which were, after height, the waist and chest circumferences for men and the hips and chest circumferences for women. Hidvegi et al., (2004) highlighted the importance of the chest circumference which can greatly affect relative proportion of the anterior and posterior trunk. Nadadur et Parkinson, (2010) observed a strong correlation between BMI and hips circumference. Haslam et James, (2005) identified the waist circumference as a feature of obesity. In several populations, neck circumference is also considered as a feature of overweightness and obesity (Ben-Noun, Sohar, & Laor, 2001; Yang et al., 2010; Hingorjo, Qureshi, & Mehdi, 2012), even in children (Hatipoglu, Mazicioglu, Kurtoglu, & Kendirci, 2010). The importance of each anthropometric measurement (see Figure 6.3) provided useful complementary information, especially in situations where the proposed set of measurements cannot be satisfied, in which case, other spare sets can be selected. Lastly, measured times satisfied the clinical requirement fixed by several clinicians, which needed to be under 5 min.

When compared to the gold standard (3D scans), MH 3D models showed good accuracy for BSA modelling. As percentage errors decreased, the number of anthropometric measurements increased. With only three to four anthropometric measurements, results were as good as anthropometric formulas, Gehan et George, (1970) and outperformed them when all measurements (19) were included. In fact, these anthropometric formulas are limited to the approximation of BSA, which is not the case for MH 3D models which allow personalizing BSA and morphology as well. Actually, these formulas are often criticized due to their accuracy being limited to the estimation of regression coefficients. They strictly depend on the

population being studied and do not offer a variety of morphological representation. Moreover, there is currently no literature on BSA formulas for children. Anthropometric measurements do not suffer from these constraints since measurements are directly dependent of the patient's body constitution. By offering different subsets of anthropometric measurements, it is possible to match typical features of the human body, especially when fat is disproportionally accumulated among obese people (such as subject G, see Figure 6.4-B-B') or woman with large breasts (Brown et al., 2012). It is also interesting for children morphology since MH takes into consideration the disproportion of children's heads relative to their bodies. Where burn estimation software – such as BurnCase3D (Haller et al., 2009) and BAI (Prieto et al., 2011) – show limitations for approximately modelling BSA, the proposed study makes a difference not only in terms of BSA modelling accuracy, but also morphology modelling accuracy, which leads to better percent TBSA estimations. It is even more striking in light of the study from Desbois et al., (2017) which demonstrated that even a small variation of BSA (0.1 m^2) due to a difference of morphology in men and women of normal weight and same BSA (1.6 m^2) resulted in a difference of Parkland fluids of more than 2 L. Of note, Klein et al., (2007) highlighted the medical consequences of a 1.5 L under- or overestimation of Parkland fluid: death (up to 25%), multi-organ failure (up to 23%), acute respiratory distress syndrome (up to 44%), blood infection (up to 12%) and pneumonia (up to 50%). Again, these odds are worse for overweight or obese people, large-breasted women or children.

Hausdorff distances among vertices were well below average. Therefore, 3D MH models were geometrically very similar to scans at any point. Color maps showed highest H_s in hand extremities or between legs with a minimal frequency density. Genitals were not modelled because of their extremely low percent TBSA (<1%). These results highlighted the morphology accuracy of 3D models.

Finally, MH is a very modular open source software which offers many possibilities to create a portable application. Indeed, ongoing research includes multicentric clinical validation of the proposed approach with a tablet application on burn patients and an online survey to assess acceptability, perception and utility of the application in clinical setting.

Cliccours.com

Several improvements concern feature selection, which can be performed on a subgroup of the dataset, children, men and women separately and according to their BMI and age distribution, rather than as a whole group.

6.7 Conclusion

This study presented the framework of a 3D-based approach – with personalized 3D models from a limited set of anthropometric measurements – to accurately assess the percent total body surface area affected by burns. 3D models were personalized to the subjects' body constitution through a limited set of anthropometric measurements clinically reliable and feasible. Predictive modelling allowed determining the most important subset of measurements. By improving percent TBSA assessments, clinicians will have a tool to help them plan appropriate treatment for major burn injuries on adults, children and overweight or obese patients since the proposed 3D model take into consideration BSA and morphology modelling. It is believed that such an application could improve initial management, improve burn outcomes, and decrease debilitating consequences and complications. Moreover, since any kind of improvement and features can be added to the proposed application, it could also be used for wound illustrations and monitoring, scars documentation and rehabilitation. It presents large perspectives in areas of physiology and medicine where the BSA needs to be assessed accurately.

CHAPITRE 7

DISCUSSION GÉNÉRALE

7.1 Synthèse des travaux de recherche

De la réanimation liquidienne initiale, aux traitements chirurgicaux jusqu'à l'évaluation de l'acte médical en matière de traitement des brûlures et au pronostic vital du grand brûlé, c'est toute la prise en charge des grands brûlés qui repose, à un moment ou à un autre, sur la mesure obtenue du %TBSA, une mesure approximative qui démontre une grande variabilité intra-inter évaluateurs.

Les méthodes actuelles de quantification 3D de la TBSA, cliniquement pertinentes, se basent sur des méthodes graphiques bidimensionnelles ne dépeignant qu'un seul type de morphologie (Prieto et al., 2011; Johannes Dirnberger et al., 2012). Ces modèles ne prennent donc aucunement en compte la variabilité des morphologies et illustrent le plus souvent un enfant, une femme ou un homme, tous minces et de taille moyenne. Ils ne prennent donc pas en considération, par exemple, l'impact de l'obésité ni l'impact de sa distribution (gynécoïde chez la femme le plus souvent et androïde chez l'homme), pas plus qu'ils ne tiennent en compte les tailles extrêmes, la musculature, etc. Chez l'enfant, les variations de proportions corporelles en fonction de l'âge sont d'autant plus significatives. Tout ceci complique l'estimation du %TBSA induisant des complications médicales et chirurgicales significatives. Bien que de nombreux efforts ont été réalisés pour offrir une personnalisation des modèles 3D, aucune ne satisfait les exigences cliniques. Seule une approche de déformation par segment permettrait de répondre à ces exigences. Cependant, là encore, une contrainte majeure est que les corrélations entre les différentes parties du corps ne sont pas considérées.

L'étude préliminaire présentée au CHAPITRE 4 propose l'utilisation du logiciel open source MH pour la personnalisation de modèle 3D à la BSA des patients. Les résultats préliminaires ont démontré que 4 à 7 mesures anthropométriques sont suffisantes et nécessaires pour une estimation précise de la BSA. Considérer plus que deux mesures anthropométriques proposées

dans la revue de littérature (poids et taille seulement) sont d'autant plus indispensables chez les patients en surpoids ou obèses (Livingston & Lee, 2000; Verbraecken et al., 2006; Lee et al., 2008) ou les enfants (Dubin & Zietz, 1996; Y. Ahn & Garruto, 2008) dont la BSA est d'autant plus sous ou surestimée. De plus, aucune étude n'avait encore proposé l'utilisation du logiciel open source MH comme choix technologique pertinent en contexte clinique tel que les grands brûlés. La régression pas-à-pas utilisée pour produire les résultats de cet article (*stepwise*) est très souvent utilisée pour la sélection de sous-ensembles et pour sa simplicité. Cependant, l'étude des différentes valeurs de R^2 indique une possible multicolinéarité des variables (mesures anthropométriques). Ceci ne biaise pas les valeurs de R^2, mais rend difficile l'interprétation de l'effet individuel de chacune des variables sur la variable dépendante ─ BSA (Derksen & Keselman, 1992; Harrell, 2001; Flom & Cassell, 2007; Morozova et al., 2015). De plus, les résultats préliminaires de cet article ont surtout servi de genèse pour la suite du projet de doctorat, l'échantillon d'étude n'était donc pas suffisamment représentatif. Compte tenu de ces problématiques, des analyses prédictives étaient donc nécessaires afin de pouvoir soulever ces limitations et faire des inférences sur les paramètres d'une population souhaitée.

Ces analyses ont été réalisées dans le cadre du CHAPITRE 6. Cependant, avant même d'effectuer ces analyses, il était primordial de vérifier l'hypothèse fondamentale du projet de doctorat : la superficie corporelle locale, qui est induite par la morphologie du patient, peut-elle différer énormément de l'approximation de la BSA et ainsi grandement affecter l'évaluation du %TBSA ?

L'étude du CHAPITRE 5 répond à cette question en y apportant également des résultats de validation. Les résultats de cette étude ont mis en évidence que pour des modèles 3D ayant une même BSA, de petites variations locales (due à la morphologie) peuvent induire une variation clinique importante du %TBSA et, par conséquent, des fluides de Parkland avec des surestimations allant jusqu'à être supérieures à 1.5L. Cette démonstration est une des contributions du projet de doctorat. En effet, aucun outil de modélisation 3D n'avait encore pris en considération la morphologie des patients dans l'estimation du %TBSA. De plus, la BSA est approximée par des formules anthropométriques ne dépendant que du poids et de la

taille, alors qu'il a été démontré dans l'étude précédente du CHAPITRE 4 que plusieurs mesures sont nécessaires. Personnaliser les modèles 3D de MH à partir de mesures anthropométriques, à la fois à la BSA et morphologie, est crucial pour modéliser les caractéristiques typiques du corps humain. Non seulement les patients obèses ont leurs BSA sous-estimées (Tucker & Alexander, 1960; Lee et al., 2008), mais également chez les enfants qui ont leurs têtes disproportionnées par rapport à leurs corps, ou encore la poitrine des femmes qui peut altérer les proportions du tronc antérieur (Hidvegi et al., 2004; Brown et al., 2012). De plus, on sait avec l'étude de Klein et al., (2007) que des sous ou surestimations des fluides de Parkland de 1.5L peuvent entraîner la défaillance de plusieurs organes, le syndrome de détresse respiratoire aiguë, une septicémie ou encore la mort du patient. D'autre part, 8 cliniciens et 6 non-experts ont évalués le pourcentage de TBSA de brûlures simulées sur 4 mannequins commerciaux (enfant, femme et homme de morphologie normale et une femme de morphologie en surpoids) d'abord avec la méthode clinique du schéma de Lund & Browder, puis avec une interface de rendue 3D. Les résultats ont montré une diminution de l'erreur d'estimation significative du %TBSA et de sa variabilité pour l'interface de rendue 3D ($Erreur_{TBSA} = 5 \pm 1\%$), comparativement au schéma de Lund & Browder ($Erreur_{TBSA} = 34 \pm 11\%$) et ce, peu importe le niveau d'expertise. Cela répond au besoin d'une estimation objective du %TBSA (Giretzlehner et al., 2013; Kamolz, Wurzer, et al., 2014). Ainsi, ces résultats ont permis de démontrer l'importance de la morphologie dans l'estimation de TBSA et introduisent des résultats de validation justifiant la pertinence clinique du projet. Notamment, de nombreux scores cliniques (ABSI, R-Baux, etc.) utilisent également la TBSA pour évaluer les chances de survie ou de mort du patient. De faibles variations dans ces scores peuvent grandement influencer le niveau de soin et la nature des traitements appropriés (Ryan et al., 1998; Tahir et al., 2009; Kraft et al., 2012; Finnerty et al., 2016).

Cette étude présente certaines limitations. Les modèles 3D sont représentés sous forme de maillage par l'interface de rendue, ce qui peut impacter les surfaces brûlées qui sont représentées plutôt grossièrement. Ce qui signifie que les résultats obtenus pourraient être meilleurs. Un plaquage de texture sur le modèle 3D pourrait améliorer la précision de segmentation des surfaces brûlées. D'autre part, toute cette étude a été réalisée sur des modèles

3D construits à partir de 19 mesures anthropométriques définis en ANNEXE I. Or, en contexte clinique d'urgence, un ensemble restreint de mesures est préférable (Exigence 01).

L'étude réalisée au CHAPITRE 6 surmonte ces limitations en proposant plusieurs ensembles de mesures anthropométriques pertinentes et validées, ainsi qu'une nouvelle application 3D pour la représentation des surfaces brûlées sur une texture 3D. Avant même de procéder aux analyses prédictives, une première étape importante était d'évaluer la répétabilité et reproductibilité de la prise de mesures afin d'identifier celles cliniquement pertinentes. L'étude de la mesure Répétabilité & Reproductibilité de l'outil six sigma à mise en évidence la pertinence clinique des circonférences plus reproductibles et de variance acceptable, contrairement aux longueurs plus répétables, mais trop variables entre cliniciens. Une plus faible répétabilité est moins problématique en raison de la courbe d'apprentissage. Avec l'expérience, les cliniciens deviendront meilleurs dans l'identification des points de repère anatomiques chez des patients en surpoids ou obèses. De plus, la taille est une mesure suffisante pour l'adaptation de la forme globale du corps, ajouter d'autres longueurs aura tendance à surestimer la BSA. Cependant, les circonférences changent la forme locale du corps et donc de la morphologie. Ainsi, l'intérêt initial quant aux 19 mesures anthropométriques a pu être réduit à 9 mesures anthropométriques ayant une pertinence clinique (taille et circonférences). Cette étude a également démontré l'intérêt de l'apprentissage machine en comparaison avec les méthodes d'analyses statistiques, spécifiquement lorsque les résidus présentent toujours une hétéroscédasticité et une structure non linéaire, et ce, même avec une transformation des données et des méthodes de régression linéaire et non linéaire pondérées. Ces résultats concordent avec l'étude du CHAPITRE 4 et de plusieurs études remettant en cause le manque de flexibilité des méthodes statistiques reposant sur des hypothèses théoriques à des données réelles (McCullagh, 1984; Wold et al., 1984; Lesaffre & Marx, 1993; Lindsey, 2000; Breiman, 2001b; Anderson et al., 2005; Alpaydin, 2014). Partant de ce constat, un échantillon de 50000 modèles 3D représentatif de la population du Canada à permit aux algorithmes d'apprentissage machine d'identifier plusieurs sous-ensembles de mesures anthropométriques prédisant le mieux la BSA et la morphologie. De plus, un autre intérêt de l'apprentissage machine est que l'importance de chaque mesure a permis d'être évaluée en

concordance avec les différents algorithmes appliqués. Ceci est particulièrement intéressant pour le contexte des grands brûlés où une situation clinique particulière pourrait empêcher la prise de mesures anthropométriques sur le patient. Par exemple, le meilleur ensemble de mesures suggère la taille et la circonférence du cou et du tour de taille. Malheureusement, le patient pourrait présenter des complications médicales sur sa brûlure empêchant toute prise de la circonférence du tour de taille. Il serait alors possible de choisir un autre sous ensemble de mesures tout autant pertinent, tel que la taille et la circonférence de poitrine et des hanches par exemple. Dans cette étude, les meilleures mesures jugées les plus importantes sont (mais ne sont pas limités à) la taille et la circonférence de poitrine, du cou, des hanches et du tour de taille. Ces résultats concordent avec ceux obtenus de plusieurs études réalisées dans d'autres contextes sur des adultes (Hidvegi et al., 2004; Wells et al., 2007; Nadadur & Parkinson, 2010) également en surpoids et obèse (Ben-Noun et al., 2001; Haslam & James, 2005; Yang et al., 2010; Hingorjo et al., 2012), ainsi que des enfants en surpoids et obèses (Hatipoglu et al., 2010).

Une contribution majeure de cette étude a été de montrer que là où des outils de modélisation 3D des brûlures tel que BurnCase 3D (Haller et al., 2009) et BAI (Prieto et al., 2011), sont limités à une modélisation 3D approximative de la BSA, l'approche proposée dans ce projet de doctorat, fait une différence non seulement dans la modélisation 3D de la BSA (non approximative), mais également de la morphologie. De plus, les modèles 3D prédéfinis proposés par ces outils sont issus d'une approximation de la BSA du patient basé sur des formules anthropométriques (Gehan & George, 1970; Haycock et al., 1978; Mosteller, 1987; Tikuisis et al., 2001) et sont donc limités par l'estimation de coefficients de régression. Ces formules anthropométriques sont par conséquent strictement dépendantes de la population étudiées et ne permettent pas une représentation des morphologies variées. L'anthropométrie ne souffre d'aucune de ces contraintes puisque les mesures sont directement dépendantes de la constitution corporelle du sujet. Lorsque comparés au gold standard (scan 3D de sujets sains), les modèles 3D reconstruits à partir de 3 à 4 mesures anthropométriques ont montré une très bonne précision au niveau de la BSA avec en moyenne une erreur inférieure à 4%. Cette erreur diminue à mesure que le nombre de mesures anthropométriques augmente. Pareil pour la

morphologie avec une erreur moyenne de 1.1cm. Enfin, MH s'avère être un choix technologique justifié puisqu'étant open source et donc très modulables, de nombreuses possibilités sont ouvertes pour la création de logiciels de modélisation 3D des brûlures gratuit et accessible à tous.

7.2 Limites et recommandations

Les résultats d'apprentissage machine du CHAPITRE 6, ont été effectués sur un échantillon représentatif de la population du Canada (enfants et adultes). Il est recommandé de pousser cette étude plus loin en refaisant les analyses pour des sous-groupes de cet échantillon : enfants, femmes et hommes séparément dépendamment de leurs IMC et la distribution d'âges.

Les études des CHAPITRE 5 et CHAPITRE 6 ont été validées en contexte préclinique pour une preuve de concept, soit sur des mannequins commerciaux ou bien sur des sujets sains. Il est recommandé de valider l'approche proposée dans ce projet sur une population de patients brûlés représentative. Selon le Centre d'expertise pour les victimes de brûlures graves de l'Est du Québec, (2015), à l'hôpital Enfant-Jésus de Québec, c'est plus de 127 patients qui sont admis à l'unité des grands brûlés avec une hausse de 48% ces cinq dernières années. En se basant sur cette donnée, il est possible d'estimer une taille d'échantillon théorique minimale (Barlett, Kotrlik, & Higgins, 2001) au niveau de confiance 95% et une marge d'erreur de 5%, n = 96 patients. Pour une marge d'erreur acceptable de 10% au niveau de confiance 95%, on souhaitera n = 55 patients. Il est recommandé d'évaluer l'intégration et l'acceptation de l'approche proposée dans ce projet de doctorat. Un questionnaire en ligne anonymisé et non exhaustif a déjà été envoyé à plusieurs cliniciens. Cependant, il est recommandé de renvoyer ce questionnaire suite à la validation clinique sur des patients brûlés. Ce questionnaire devra être complété avec plusieurs thématiques en lien avec:

- Avis sur l'innocuité à courts et longs termes
- Facilité de la prise de mesures anthropométriques sur des grands brûlés
- Facilité d'utilisation de l'application Tablette sur des grands brûlés

- Avis sur les impacts sur les pratiques cliniques
- Avis sur les impacts sur l'organisation des soins, incluant les autres secteurs de la santé
- Avis sur l'adoption de l'approche proposée
- Avis sur les barrières à l'implantation (peut-être institutionnelles, sociales ou éthiques à l'adoption)
- Événements indésirables reportés
- Tout autre commentaire/ recommandations

Il est recommandé de développer et de valider une application de cette preuve de concept (voir en ANNEXE II, le prototype de l'application sur tablette). D'autres outils additionnels peuvent y être également ajoutés tels que la profondeur de brûlures ou des images photographiques. On peut également ajouter les recommandations en fluide selon la formule de Parkland, ainsi que l'intégration de fonction d'amputation, les greffes et sites donneurs potentiels comme c'est le cas de certaines études (Neuwalder et al., 2002). Ou encore toutes les informations qui peuvent être ajoutées au dossier électronique du patient tel que des options pour la documentation des antécédents, détails et blessures des patients, etc. Un système de numérotation pour l'identification et le suivi. L'information de l'activité d'intubation. Des mesures de sécurité pour préserver la confidentialité du patient (Neuwalder et al., 2002). Également, d'autres fonctions prometteuses peuvent être pensées telles que dans l'étude de Richard et al., (2015), avec l'utilisation des unités fonctionnelles cutanées pour la prédiction de la cicatrice des brûlures qui sont susceptibles de devenir encore plus significatives.

7.3 Retombés et impacts

Ce projet de doctorat propose une preuve de concept d'un outil calculant le pourcentage de superficie brûlée de façon précise, rapide, simple et reproductible. Les erreurs de calcul des cliniciens seront moins importantes, en particulier lorsque les patients présentent des morphologies particulières. Cet outil permettra également de diminuer la variabilité entre les observateurs. Les cliniciens auront donc un outil pour les aider dans la planification du traitement des grands brûlés, aussi bien pour les adultes que les enfants ou encore les personnes

obèses. Cet outil permettra de diminuer la morbidité liée aux imprécisions d'évaluation de la TBSA et aux erreurs de traitements pouvant en découler. Spécifiquement, la diminution des complications médicales à la suite d'une évaluation sous-optimale, la diminution de la durée du séjour hospitalier des patients et l'optimisation des soins des patients souffrant de brûlures sévères permettant ainsi la diminution des coûts associés au traitement des grands brûlés (donnée très importante dans le contexte actuel du système de santé). De plus, certaines applications potentielles pourraient faciliter visuellement le suivi postopératoire des sites greffés et des sites donneurs, et éventuellement faciliter la représentation des zones d'atteinte cicatricielle et de contractures et, ainsi, faciliter la communication d'information en réadaptation physique.

L'approche proposée dans le cadre de cette thèse pourra également être appliquée à l'estimation du pourcentage de surface corporelle ayant un traumatisme, ou divers troubles de surfaces tels que les maladies de la peau, lésions, maladies épidermolytiques (TENS ou le Steven-Johnson), etc., et qui utiliseraient la même évaluation en 3D.

D'autre part, l'utilisation de modèles personnalisés à la morphologie des patients présente un intérêt pour la fabrication de moulages, masques et vêtements de compression sur mesures pour la réadaptation des brûlés (réduction de l'œdème et prévention de cicatrice hypertrophique/ chéloïde). La durée des séances de prises de mesures directes diverses sur les patients sera beaucoup moins longue (puisque limité à 3-4 mesures anthropométriques). De plus les moules nécessaires à la confection des masques et qui peuvent être invasifs pour les patients seront remplacés par les modèles 3D.

Enfin, l'estimation précise en 3D de la BSA présente également de larges perspectives en physiologie et médecine (voir ANNEXE III) pour le calcul des doses de chimiothérapie (Madarnas, Sawka, Franssen, & Bjarnason, 2001; Gurney, 2002; T Vu, 2002; Sparreboom et al., 2007; Field et al., 2008; Gao, Klumpen, & Gurney, 2008; Hunter et al., 2009); l'indice de l'efficacité de la pompe cardiaque (Connors, McCaffree, & Gray, 1983; Eisenberg, Jaffe, & Schuster, 1984; Celoria et al., 1990; Steingrub, Celoria, Vickers-Lahti, Teres, & Bria, 1991;

Mimoz et al., 1994; Tibby, Hatherill, Marsh, & Murdoch, 1997; Lehot & Arvieux, 2011), ou encore en clairance rénale (Hoste et al., 2005; Poggio et al., 2005; Conil et al., 2007; Herrera-Gutierrez et al., 2007; Roberts, Kruger, Paterson, & Lipman, 2008; Baptista et al., 2011; Martin et al., 2011).

Clicours.COM

CONCLUSION

Ce projet de doctorat a abordé la problématique du manque de précision de l'évaluation du %TBSA chez les grands brûlés. Les méthodes cliniques présentent des erreurs de sous et surestimations ainsi qu'une grande variabilité pouvant entraîner des complications médicales et chirurgicales significatives. De ce constat a été formulé l'objectif principal du projet qui consiste en la conception et validation de méthodes numériques permettant l'évaluation quantifiée de la taille des superficies corporelles brûlées à partir de modélisation 3D du corps entier obtenue à partir d'un nombre limité de mesures anthropométriques.

La première contribution de ce projet de doctorat a été de démontrer quantitativement que la superficie corporelle locale, qui est induite par la morphologie du patient, peut différer énormément de l'approximation de la BSA et ainsi grandement affecté l'évaluation du TBSA. Aucune étude n'avait encore pris en considération la morphologie dans l'estimation de la TBSA.

La deuxième contribution concerne l'intérêt de générer un large échantillon stratifié aléatoire et représentatif de la population du Canada. De nombreuses analyses requièrent un échantillon représentatif, pourtant les bases de données actuelles ne représentent pas suffisamment les populations d'intérêts puisque la distribution des morphologies n'y ait pas représentée pour des enfants, obèses, etc. L'utilisation du logiciel MH permet de surmonter ces limitations et offre une excellente alternative à cette contrainte.

La troisième contribution de ce projet de doctorat porte sur l'importance des mesures anthropométriques dans la modélisation 3D de la BSA et de la morphologie. Plusieurs sous-ensembles de mesures pertinentes cliniquement ont été proposés. D'autre part, la validation de ces résultats a apporté de nouvelles informations quant à la faisabilité d'une telle approche et l'évaluation du décalage entre la BSA réelle d'un sujet et la BSA du modèle 3D. Bien qu'aucune étude ne se soit penchée sur ce point particulier, seul une mesure de ce décalage pouvait être évaluée en comparant les modèles obtenus à des scans 3D.

Les résultats finaux présentent une preuve de concept validée permettant l'évaluation quantifiée de la taille des superficies corporelles brûlées et remplissant les objectifs visés en termes de précision et de temps d'interaction. Ce projet de recherche espère démontrer l'importance d'une telle approche pour l'améliorer de la prise en charge initiale des grands brûlés qui saura faire la différence entre la mort, des séquelles débilitantes à long terme et une bonne récupération.

La première contribution de cette thèse a été de démontrer la pertinence clinique du logiciel MH et la nécessité de prendre en considération plus de deux mesures anthropométriques pour une modélisation 3D précise de la BSA. Ces résultats ont fait l'objet d'une publication d'un article scientifique dans la revue « International Journal of Human Factors Modelling and Simulation » (2017) ainsi qu'une présentation orale au congrès « 4th International Digital Human Modeling » (Montréal, 2016) et d'un article web sur le site d'actualités scientifiques Substance (2016).

La deuxième contribution de cette thèse a été de démontrer l'importance de la morphologie dans l'estimation du %TBSA et introduisent des résultats de validation justifiant la pertinence clinique du projet. Ces travaux ont fait l'objet d'une publication d'un article scientifique dans la revue « Journal of Burn Care & Research » (2017), une présentation poster au congrès « Canadian Special Interest Group of the 49th American Burn Association Annual Meeting » (Boston, 2017), une présentation orale au congrès « 4th annual Canadian Burn Symposium » (Winnipeg, 2017), une présentation orale au département d'urgence du CHU de Sainte-Justine (Montréal, 2017) et deux présentations orales au département de traumatologie du CHU de Sainte-Justine (Montréal, 2017).

La troisième contribution de cette thèse présente la comparaison des analyses prédictives (statistiques et apprentissage machine) pour des données non linéaires présentant une sévère multicolinéarité. Ces résultats ont permis d'identifier les meilleurs sous-ensembles de mesures anthropométriques permettant d'obtenir une modélisation 3D précise de la BSA et de la morphologie. La méthodologie et les résultats des analyses statistiques ont fait l'objet d'une

présentation orale pour un cours de maîtrise en statistique (Montréal, 2016). Le code source des algorithmes d'apprentissage machine pour les analyses prédictives a été publié sur le site web RPubs. La précision de la modélisation 3D à partir d'un ensemble de mesures limités a été également validé par des métriques d'erreurs en 3D. Tous ces résultats ont fait l'objet d'une soumission d'article scientifique dans la revue « IEEE – Transactions on Biomedical Engineering » (2018), une présentation orale sur écran numérique à la conférence Innove-Action du CRCHUM (Montréal, 2018) et une présentation orale au Centre de coordination de la télésanté du CHUM (Montréal, 2018).

Enfin, la quatrième contribution de cette thèse porte sur la rédaction de plusieurs documents qui seront utiles pour la suite du projet : protocole de prise de mesures anthropométriques (voir ANNEXE I), protocole clinique (voir ANNEXE II) et questionnaire d'acceptation clinique.

ANNEXE I

PROTOCOLE DE PRISE DE MESURES ANTHROPOMÉTRIQUES

Le protocole de prise de mesures s'appuie sur les mesures et points de repère définis des protocoles suivant : Measurer's Handbook: US Army and Marine Corps Anthropometric Surveys, 2010-2011; Hotzman et al., (2011) et Anthropometric Survey of US Marine Corps Personnel: Methods and Summary Statistics, Gordon et al., (2013) et respect la norme 7250-1, ISO, (2008) et du Bureau de normalisation du Québec (BNQ), (1995). À noter que les mesures et points de repère peuvent différer des cliniciens. Une version de ce protocole révisé par les cliniciens a été effectuée dans le cadre d'un projet de maîtrise par les co-auteurs des études du CHAPITRE 4, CHAPITRE 5 et CHAPITRE 6, le Dre Isabelle Perreault et le Dr Yannick Leclerc.

Circonférence du cou

Prendre la circonférence du cou au niveau de la pomme d'Adam avec un ruban métrique. Le plan de mesure est perpendiculaire à l'axe du cou. Exercer seulement une tension du ruban suffisante pour maintenir le contact entre le ruban et la peau. Pomme d'Adam: Palper la surface latérale lisse du cartilage de la thyroïde puis localiser le point bas du cartilage de la thyroïde dans le plan sagittal médian. Point de repère : proéminence thyroïde.

Distance épaule

Prendre la distance entre le point de repère trapèze droit et le point de repère acromion droit avec un ruban métrique. Point de repère trapèze droit: En se déplaçant de l'épaule au cou, palper la masse du muscle. Point de repère : point d'encolure. Point de repère acromion droit: Point de la surface de la peau obtenu en palpant le bord le plus latéral de l'acromion de la scapula. Point de repère : point acromial.

Distance poitrine

Prendre la distance de poitrine horizontale au niveau du point de repère antérieur de la poitrine avec un caliper. Point de repère antérieur de la poitrine: Apercevoir le point droit de la poitrine le plus protubérant. Point de repère : thélion.

Circonférence sous poitrine

Prendre la circonférence horizontale de la poitrine directement sous le buste au niveau du point de repère infrasternal avec un ruban métrique. Exercer seulement une tension du ruban suffisante pour maintenir le contact entre le ruban et la peau. Point de repère infrasternal: point sternal le plus bas du rebord du sternum.

Circonférence tour de taille

Prendre la circonférence horizontale du tour de taille, passant par le point de repère omphalion avec un ruban métrique. Exercer seulement une tension du ruban suffisante pour maintenir le contact entre le ruban et la peau. Point de repère omphalion: repérer le centre du nombril par inspection. Pour des patients lourds, le nombril peut pointer vers le bas.

Taille de jambes supérieure

Prendre la distance verticale entre le point de repère trochantérion et celui de l'épicondyle fémoral latéral avec un ruban métrique tendu. Point de repère trochantérion: point le plus latéral du grand trochanter. Point de repère épicondyle fémoral latéral: point le plus latéral proéminent du condyle distal du fémur.

Circonférence du genou

Prendre la circonférence du genou avec un ruban en saisissant les proéminences osseuses de la partie inférieure du fémur situé à gauche et à droite du genou. Fléchir un peu le genou du patient pour aider à localiser ces structures. Point de repère épicondyle fémoral latéral: point le plus latéral proéminent du condyle distal du fémur.

Longueur bras supérieur

Prendre la distance entre le point de repère acromion droit (a) et le point de repère radial avec un caliper. Point de repère acromion droit (a): Point de la surface de la peau obtenu en palpant le bord le plus latéral de l'acromion de la scapula. Point de repère radial: palper la zone supérieure du bord extérieur du radius (le creux sur le côté latéral du coude). Si vous avez des difficultés à localiser ce point, saisir le poignet du patient et faire pivoter le bras avant et en arrière tout en palpant l'os dans la zone du point de repère.

Circonférence poignet

Prendre la circonférence du poignet, perpendiculaire le long de l'axe de l'avant-bras avec un ruban métrique passant sur le point de repère du stylion. Exercer seulement une tension du ruban suffisante pour maintenir le contact entre le ruban et la peau. Point de repère du stylion: Point situé à l'extrémité distale du radius.

Circonférence hanches

Prendre la circonférence horizontale du tronc au niveau des points de repère des fesses, postérieur, latéral droit et gauche avec un ruban métrique. Exercer seulement une tension du ruban suffisante pour maintenir le contact entre le ruban et la peau. Points de repère des fesses, postérieur, latéral droit et gauche: chercher le point de saillie maximum de la fesse, point le plus latéral du grand trochanter.

Circonférence chevilles

Prendre la circonférence horizontale minimum de la cheville avec un ruban métrique. Exercer seulement une tension du ruban suffisante pour maintenir le contact entre le ruban et la peau. Points de repère : point de malléole péronière et malléole tibiale.

Circonférence bras supérieur

Prendre la circonférence du bras supérieur autour du muscle non fléchi du biceps brachial au point de repère du biceps avec un ruban métrique maintenu perpendiculaire le long de l'axe du bras supérieur. Exercer seulement une tension du ruban suffisante pour maintenir le contact

entre le ruban et la peau. Point de repère du biceps: par inspection, localiser sur le bras relâché le milieu entre le point de repère acromion droit et le point de repère radial.

Longueur bras inférieur

Prendre la distance entre le point de repère radial et stylion avec un caliper tenu parallèlement le long de l'axe de l'avant-bras. Si besoin, cette mesure peut être effectuée avec un ruban métrique tendu. Point de repère radial: point extérieur le plus proéminent de l'extrémité proximale du radius. Point de repère stylion: point situé à l'extrémité distale du radius.

Circonférence buste

Prendre la circonférence horizontale maximum de la poitrine au niveau du point de repère du thélion avec un ruban métrique. Point de repère thélion: point situé au centre du mamelon.

Longueur tour de taille-hanches

Prendre la distance verticale du point de repère iliocristal et au niveau du point de repère trochantérion avec un ruban métrique tendu. Point de repère trochantérion: point le plus latéral du grand trochanter. Point de repère iliocristal : utilisez les deux mains pour localiser les points des crêtes iliaques antérieures et postérieures et prendre la moitié de la distance entre eux. À ce milieu, utilisez la pointe du doigt pour localiser le point le plus élevé palpable.

Circonférence cuisse

Prendre la circonférence de la cuisse au point de repère du sillon fessier avec un ruban métrique. Exercer seulement une tension du ruban suffisante pour maintenir le contact entre le ruban et la peau. Point de repère du sillon fessier: Le point le plus bas du sillon le plus bas ou pli à la jonction de la fesse et de la cuisse.

Circonférence mollet

Prendre la circonférence horizontale maximum du mollet avec un ruban métrique au niveau de la zone de circonférence maximale observée. Exercer seulement une tension du ruban suffisante pour maintenir le contact entre le ruban et la peau.

Taille jambes inférieure

Prendre la distance verticale entre le point de repère pternion et le point de repère épicondyle fémoral latéral avec un ruban métrique tendu. Point de repère du pternion : point le plus proéminent du talon du pied. Point de repère épicondyle fémoral latéral: point le plus latéral proéminent du condyle distal du fémur.

Longueur nuque - tour de taille

Prendre la distance de surface entre le point de repère cervical et de l'omphalion postérieur avec un ruban métrique. Le ruban suit les contours du corps. Placer le point zéro du ruban sur la cervicale. Exercer seulement une pression suffisante pour empêcher le mou du ruban. Point de repère cervical: le point palpable le plus proéminent de la 7ème cervicale de la colonne vertébrale. Point de repère omphalion postérieur: repérer le centre du nombril par inspection. Pour des patients lourds, le nombril peut pointer vers le bas. Puis transférer sur la colonne vertébrale du patient.

Clicours.COM

ANNEXE II

PROTOCOLE CLINIQUE ET UTILISATION PROTOTYPE DE L'APPLICATION SUR TABLETTE

Cette version non exhaustive du protocole clinique pourra servir pour la validation clinique sur des patients grands brûlés (voir chapitre 7.2). Il est a noté également que l'application présentée n'est qu'un prototype et ne sert qu'à titre de preuve de concept.

Plusieurs patients majeurs et mineurs seront invités à prendre part à cette validation clinique selon des critères d'admissions. Le recrutement sera effectué en personne par les infirmières du groupe de recherche du département de l'urgence. En collaboration avec ce groupe de recherche, le patient et/ou son parent seront approchés et l'étude sera présentée par un professionnel de la santé (infirmière de recherche) qui a une relation indépendante de la prise en charge du patient, une connaissance adéquate du protocole, du sujet étudié ainsi que des bonnes pratiques cliniques applicables à la recherche. Les infirmières de recherche vérifient aussi l'admissibilité du patient en fonction d'une liste de critères préétablie.

Critères d'inclusions:
Selon l'American Burn Association & American College of Surgeons, (2007) être un patient admis à l'urgence pour une brûlure couvrant une TBSA = 10 % et plus. Un patient avec une TBSA = 10 % ou plus sont habituellement admis à l'hôpital. Le projet de recherche ne retardera donc pas la prise en charge, ainsi que la fluidité du débit à l'urgence. Critère choisi selon l'étude de Giretzlehner et al., (2013): surestimation de plus de 62% chez l'enfant de 2 ans pour un %TBSA < 20%.

Critères d'exclusions :
Âge du patient inférieur à 12 mois. En effet, l'algorithme de modélisation 3D ne permet de modéliser que des enfants âgés de 12 mois minimum.

Tous les cliniciens impliqués dans cette validation clinique qui ne seront pas déjà familiers avec le protocole recevront une formation sur la prise de mesures anthropométriques. Une formation sur l'utilisation de l'application sur tablette s'ensuivra pour tous. Une évaluation pratique sera effectuée afin de s'assurer de l'efficacité de la formation.

Une des considérations particulières est que cette étude sera observationnelle. Le protocole sera effectué simultanément à la prise en charge du patient et ne perturbera en aucun cas celle-ci. De plus, les risques associés sont minimes. La prise en charge du patient (ex. calcul du Parkland) ne sera pas basée sur les résultats obtenus par l'application proposée sur Tablette. Les traitements seront basés sur les schémas d'évaluation classiques en 2D qui est la méthode qui a cours aux deux centres universitaires.

Voici une proposition du déroulement des acquisitions approuvées par le Dr Perreault, Chirurgienne plasticienne au CHU Sainte-Justine (Montréal):

Ci-dessous, le protocole de validation sur patients mineurs qui présente la procédure habituelle de soins et l'utilisation simultanée de l'application sur tablette proposée dans le présent projet de doctorat :

1. Admission du patient à l'urgence. L'infirmière de recherche de l'urgence qui s'occupe du recrutement des patients s'assure que le patient répond aux critères d'inclusion le réfère comme candidat potentiel. Des assistants de recherche qui sont des étudiants en médecine peuvent également recruter des patients. Ils remplissent le même rôle que les infirmières de recherche.

2. En collaboration avec l'urgence du CHU Sainte-Justine, cette étape sera effectuée en même temps que le patient est pris en charge à l'urgence (durée moyenne de 10 min): Le patient et/ou son parent sont approchés et l'étude est présentée par l'infirmière de recherche qui a une relation indépendante de la prise en charge du patient, une connaissance adéquate du protocole, du sujet étudié et des bonnes pratiques cliniques

applicables à la recherche. Le consentement du patient et/ou de son parent est recueilli par l'infirmière de recherche de l'urgence. Le patient est référé à l'un des cliniciens qui connaissent le protocole de recherche ainsi que sa méthodologie et les étapes à réaliser en lien avec le protocole. L'infirmière de recherche de l'urgence lance l'application sur la tablette et vérifie l'incrémentation de l'identifiant numérique - IDN (voir Figure 7.1 ci-après), correspondant au numéro du patient pour l'anonymisation). Elle marque le numéro IDN dans le dossier médical du patient et remet le dossier patient ainsi que la tablette au clinicien désigné.

3. Pour une brûlure de TBSA entre 10% et 30 %, le patient va être évalué selon la procédure clinique spécifique. Selon ce protocole, le service de traumatologie et de chirurgie plastique est avisé, ainsi que les soins intensifs selon la sévérité et le type de brûlure. La première évaluation de la TBSA est habituellement faite à l'urgence par l'urgentologue, l'équipe de traumatologie et de plastie. Les pertes volémiques sont calculées et répétées en fonction de l'évaluation du % TBSA brûlé. Le patient est ensuite admis en chirurgie traumatologie ou aux soins intensifs, selon la gravité de ses brûlures. Stabilisation du patient et traitement précoce.

4. Une fois le patient stabilisé selon le protocole habituel de prise en charge des brûlures étendues et extensives, l'évaluation ATLS (Advanced Trauma Life Support) définitive du patient brûlé a lieu. Cette évaluation comprend le calcul détaillé des surfaces corporelles brûlées en pourcentage (méthode standard d'évaluation des brûlures de Lund & Browder), afin d'évaluer, entre autres, le débit de soluté requis pour traiter les pertes volémiques et, par conséquent, stabiliser les fonctions vitales du patient. Le calcul détaillé des surfaces corporelles brûlées en pourcentage effectué avec la méthode de Lund & Browder par le médecin en charge du patient sera récolté de manière anonyme (par l'identification numérique) pour comparer ces résultats avec l'évaluation de la TBSA par l'application sur tablette.

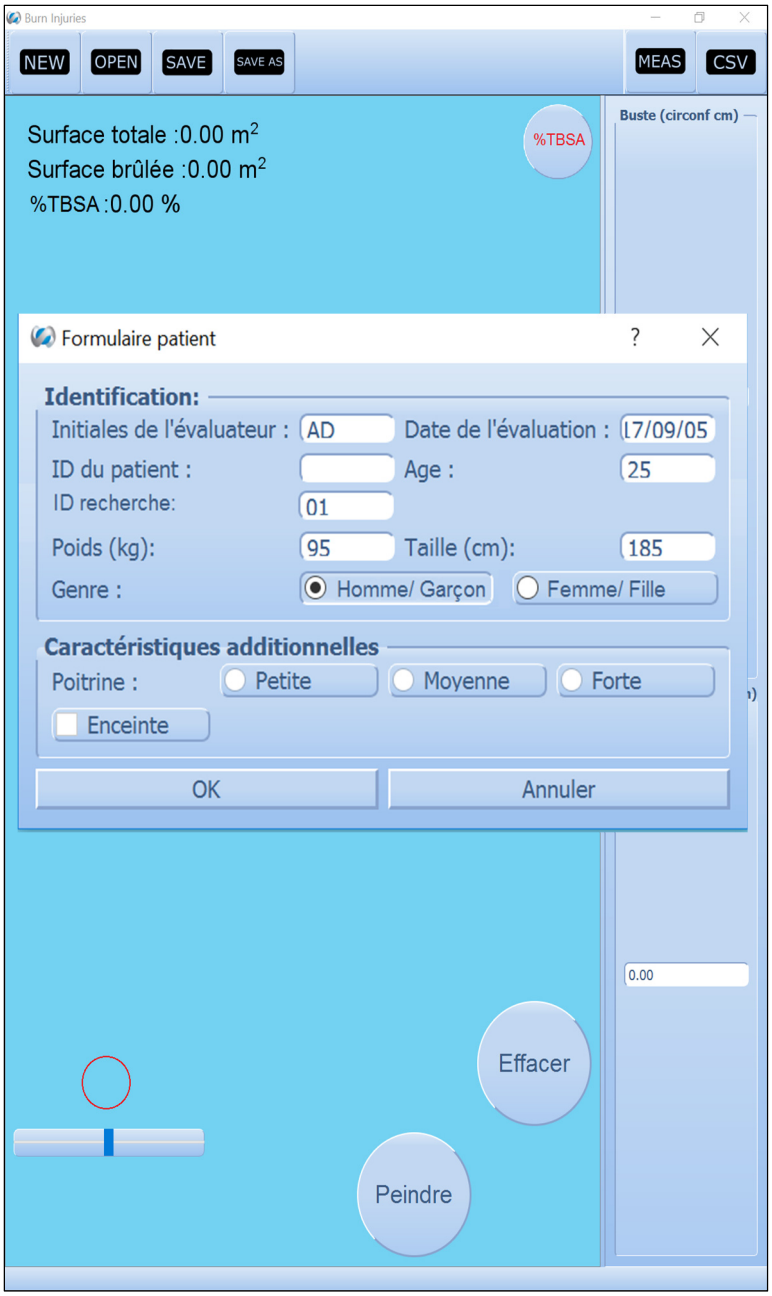


Figure 7.1 Première étape de l'utilisation de l'application, remplir le formulaire du patient

5. Le médecin en charge du patient choisi la morphologie qu'il juge appropriée au patient (voir Figure 7.2 ci-après). Il prendra ensuite 3 à 4 mesures anthropométriques (voir CHAPITRE 6) sur le patient afin de générer son avatar en trois dimensions (voir Figure 7.3). Cette tâche requiert environ 3 min et est effectuée simultanément aux soins

cliniques. Il peut être aidé des infirmières si besoin. Sachant que 4 mesures sont préférables pour de meilleurs résultats, choisir un ensemble de 3 mesures donnent des résultats tout aussi satisfaisants. Le médecin aura le choix parmi les ensembles de 3-4 mesures anthropométriques suivants :

- Taille, circonférence de poitrine, du cou et du tour de taille ou bien
- Taille, circonférence du tour de taille, de poitrine et des hanches, ou bien
- Taille, circonférence du tour de taille et du cou ou bien
- Taille, circonférence de poitrine et des hanches.

Le risque de contamination des brûlures est minime puisque l'outil qui sera utilisé pour la prise de mesure est un ruban métrique médicale utilisé habituellement en clinique pour la planification des chirurgies chez les grands brûlés. Pour une TBSA = 30% ou plus, le patient est intubé et inconscient, sous sédation et analgésique. La prise de mesures anthropométriques ne devrait pas occasionner d'inconfort ou douleur.

6. Simultanément aux soins cliniques habituels, le médecin en charge du patient évalue ensuite la taille des brûlures avec l'application sur la tablette. Il peinture les brûlures directement sur le modèle 3D soit avec un stylet, soit avec son doigt. Puis sauvegarde. S'il y a une pose de pansements standards minces, ils devront être enlevés pour l'évaluation des brûlures.

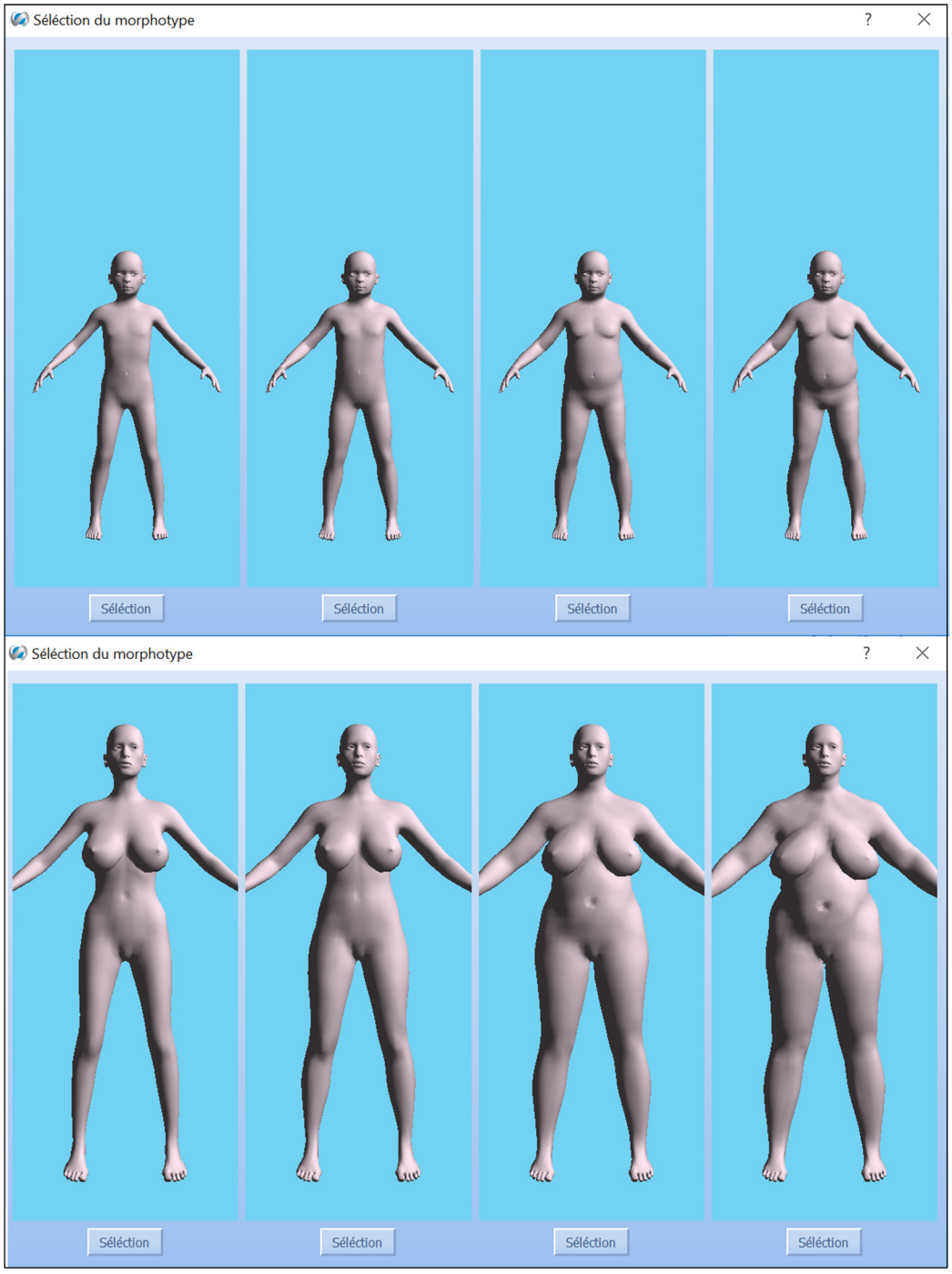


Figure 7.2 Choix de morphologie. Exemple d'un enfant en haut et d'une femme en bas

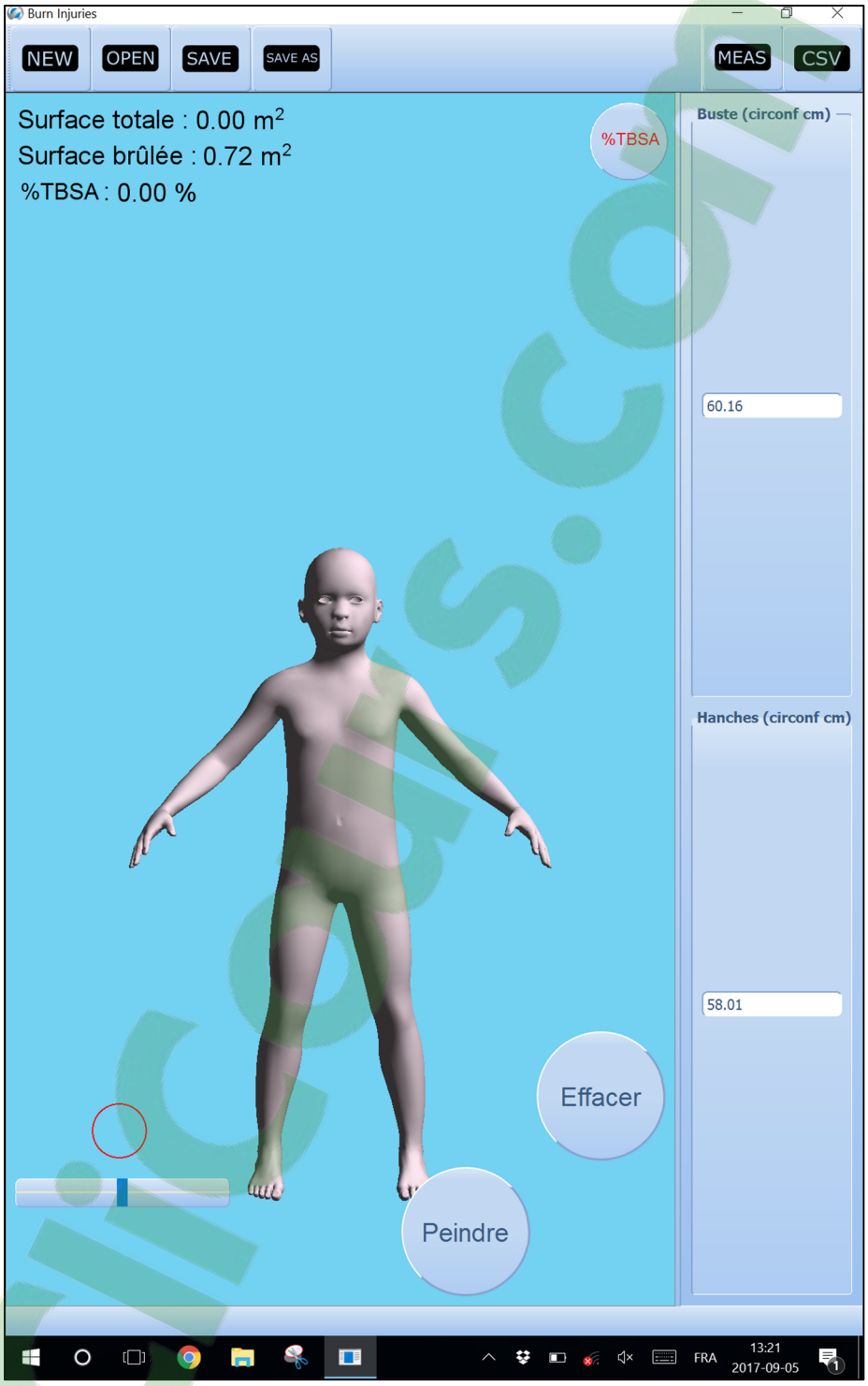


Figure 7.3 Entrée des mesures à droite. Exemple avec l'ensemble taille, circonférence poitrine et hanches

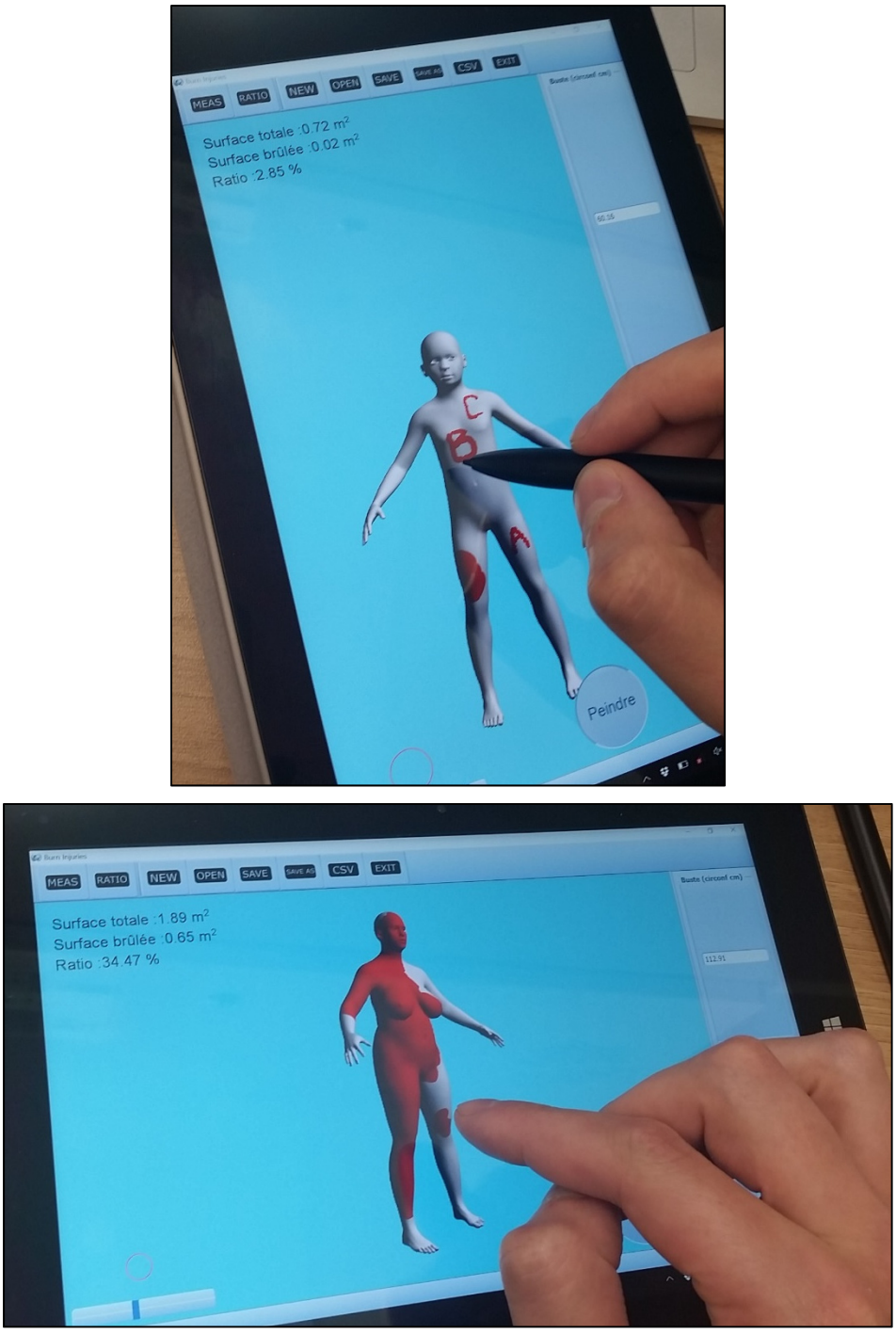


Figure 7.4 Peinturage des brûlures sur le modèle 3D. La TBSA est calculée en conséquence. Il est possible de régler la taille du curseur pour plus de précision (figure du haut)

7. Deuxième évaluation conventionnelle avec la méthode de Lund & Browder par les chirurgiens ou résidents plasticiens.

8. Si requis, deuxième évaluation avec l'application sur Tablette par les chirurgiens ou résidents plasticiens. Le modèle 3D précédemment construit du patient peut être directement chargé pour l'évaluation du patient avec son IDN (environ 3-5 min):

9. La tablette est remise à l'infirmière de recherche.

Pour des visites médicales subséquentes en salle de procédures ou en hydrothérapie, suivre les étapes 7 à 9.

Le protocole de validation sur patients majeurs est fortement similaire à celui pour des patients mineurs. Pour éviter toute redondance, seules les étapes qui pourraient différer sont présentées :

1. Une fois la réception du patient à l'unité des brûlés, l'infirmière de recherche affecte le patient au clinicien responsable du protocole et prépare la tablette (tel que définie par l'étape 2 pour les mineurs).

2. Aviser la traumatologie et les pharmaciens de l'arrivée du patient. Admission en salle de réanimation et stabilisation du patient, puis évaluation ATLS (procédure habituelle de soins).

3. Simultanément aux soins cliniques, prise de mesures anthropométrique (≈ 3 min).

4. Simultanément aux soins cliniques, évaluation du pourcentage de TBSA avec la Tablette (≈ 10 min).

Clicours.COM

5. Si requis, deuxième évaluation avec la méthode de Lund & Browder par les chirurgiens ou résidents plasticiens.

6. Si requis, deuxième évaluation avec la Tablette par les chirurgiens ou résidents plasticiens (≈ 3 min).

7. La tablette est remise à l'infirmière de recherche.

L'application sur Tablette comprend le menu et outil suivant :

Mesures : **Utilité pour la recherche seulement.** Permet l'entrée d'autres mesures anthropométriques.	
New: Créer une nouvelle évaluation.	
Open: Ouvre une ancienne évaluation auparavant effectuée. Utile pour le suivi du patient. Également pour la recherche.	
Save: Sauvegarde l'évaluation en cours.	
Save as: Sauvegarde une nouvelle évaluation sans effacer l'ancienne.	
CSV: **Utilité pour la recherche seulement.** Sauvegarde toutes les informations du patient dans un fichier Excel .csv	

Figure 7.5 Menu proposé par l'application tablette

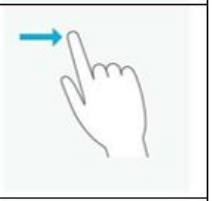
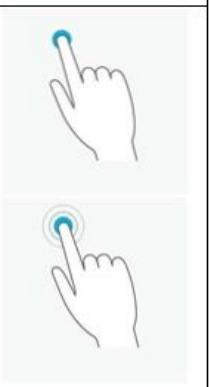
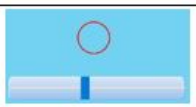
Zoom : Utiliser deux doigts.	
Rotation en 3 axes : Avec un doigt ou le stylet, glisser dans la direction que vous souhaitez tourner.	
Peindre: Cliquer sur le bouton « Peindre ». Maintenir appuyer le doigt ou stylet pour peindre. **Effacer le tracer:** Cliquer sur le bouton « Effacer ». Maintenir appuyer le doigt ou stylet pour effacer. *-> Dans ces situations, la rotation et le zoom ne sont pas disponibles.*	
Taille du curseur: Faites glisser le curseur de gauche à droite pour taille de curseur petite ou grande.	
Calculer le %TBSA: Une fois terminer de peindre les brûlures, cliquer sur le bouton « %TBSA » pour calculer le %TBSA.	

Figure 7.6 Outils proposés par l'application tablette

Notes concernant l'application sur tablette:

Cette application a été développée par Florian Beguet, assistant de recherche au LIO, en C++ sous système Windows sur une tablette CHUWI Hi10pro avec un processeur Intel Core x5-z8350 (4x 1.44 Ghz) et 4 GB RAM. L'interface a été développée en utilisant Qt avec un rendu OpenGL. Le peinturage des superficies corporelles brûlées est réalisé sur la texture 3D (UV map des modèles 3D de MH). Les scripts python de modélisation 3D de MH ont été embarqués dans l'application C++.

ANNEXE III

AUTRES CONTEXTES CLINIQUES NÉCESSITANT UNE ESTIMATION PRÉCISE DE LA BSA

Tableau 7.1 Contexte clinique d'utilisation de la BSA et leurs problématiques associées

Contexte clinique	Problématiques	Recommandations
Calcul des doses de chimiothérapie ≈ Calculé directement par les formules BSA.	Sous-estimation & Surestimation ↗ Variabilité intra-inter observateurs ↗ ⇒ ↗ risque de récidive ⇒ ↗ taux de toxicité & ↘ taux de survie ⇒ ↘ efficacité de certains traitements	considérer d'autres paramètres que la BSA seulement. Médecine personnalisée.
Indice de l'efficacité de la pompe cardiaque ≈ (exploration hémodynamique) C à l'échographie ⟶ non pratique en SSI. Débit cardiaque / BSA.	Imprécision & Sous-estimation ⇒ traitements sous-optimaux & inappropriées ⇒ complications médicales & chirurgicales significatives	
Clairance rénal ≈ (Insuffisance rénale aiguë, traitement thérapeutiques) C radio-nucléotide ⟶ non pratique en SSI. Indicateur clairance créatinine rénal / BSA.	Imprécision & Sous-estimation ⇒ sous-dosage thérapeutique ⇒ ↗ résistance aux thérapies ⇒ échecs ⇒ complications médicales & chirurgicales significatives	

Légende : SSI : Service de Soins Intensifs. ≈ approximé ; ⇒ conséquences ; ↗ augmentation ; ↘ diminution ; C complémentaire; ⟶ limites.

ANNEXE IV

SCANNER PORTABLE : PROBLÉMATIQUE POUR LES GRANDS BRÛLÉS

1. Matériels utilisés et sujets

Le numériseur portable à lumière blanche Go!Scan 3D de la société CREAFORM a été utilisé pour cette étude préliminaire. Ce scanner commercialisé a une résolution de 0,1 à 0,5 mm, n'émet aucune radiation et permet la reproduction d'un objet en trois dimensions.

Pour cette étude préliminaire, un sujet sain a participé à l'étude à titre de donnée préliminaire selon son consentement. Dans le contexte des grands brûlés, le patient devra être scanné en trois parties (dos et côtés). Ces contraintes ont été appliquées à notre sujet préliminaire afin de simuler les conditions cliniques d'un patient intubé (impossibilité de scanner debout). Ceci implique des changements de poses importants (Figure 7.7). L'objectif dans un premier temps, est donc d'arriver à obtenir un scan complet pour ensuite, procéder à la déformation d'un modèle suivant les données du scan.

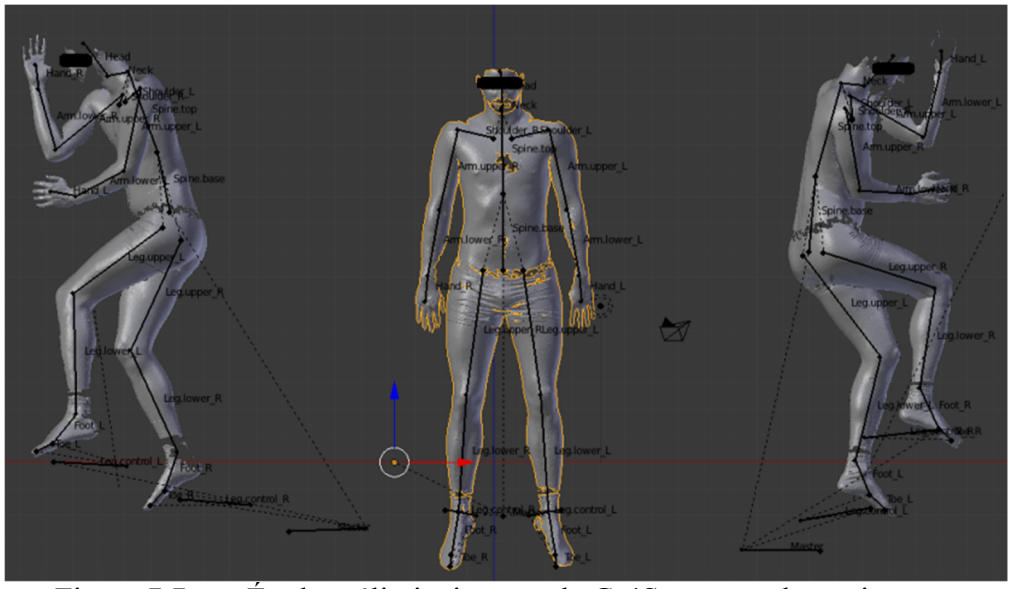


Figure 7.7 Étude préliminaire avec le Go!Scan, vue des trois scans à recaler avec leurs armatures

2. Méthodologie

Les données incomplètes sont une problématique commune à cette méthode et à la revue de la littérature. Le problème clé est de compléter ces données manquantes et d'assurer une correspondance qui soit acceptable. Les techniques les plus performantes et les plus couramment utilisées dans ce genre de situation sont l'utilisation d'un modèle générique pour compléter les données manquantes d'une part, puis l'utilisation d'armatures pour une première correspondance d'autre part. Des algorithmes de recalage non rigide permettent d'optimiser l'appariement des maillages et de diminuer l'erreur de correspondance.

Un premier pré-recalage avec le logiciel Blender (logiciel de rendu 3D) a été réalisé afin d'observer la faisabilité de cette méthode. L'approche populaire existante proposée fut l'utilisation d'un modèle générique et d'une armature (Figure 7.8) :

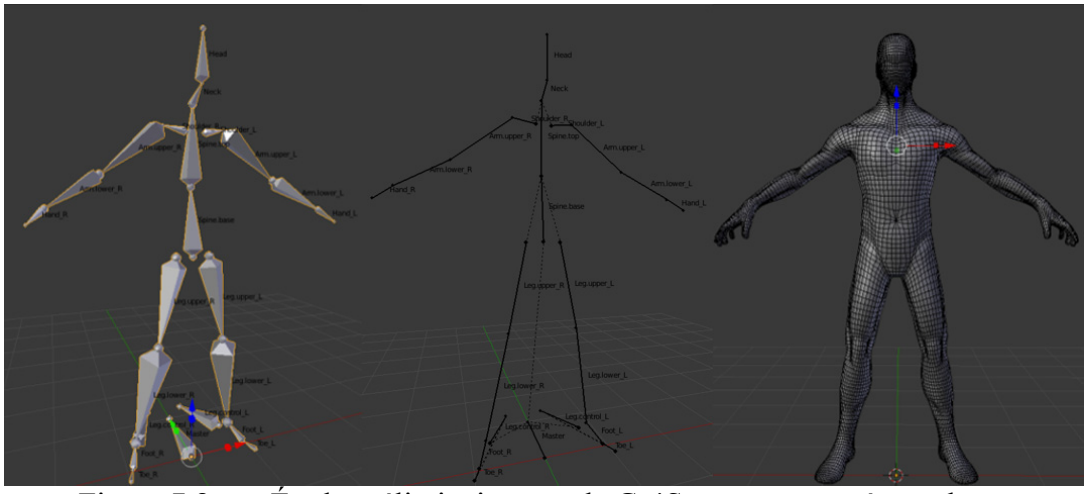


Figure 7.8 Étude préliminaire avec le Go!Scan, armature à gauche et modèle générique à droite

Une armature est constituée de plusieurs « os ». Chaque os pouvant être assigné à un groupe de vertex avec une pondération choisie. Ci-après les différentes étapes de cette étude préliminaire.

1) <u>Premier essai du pré-recalage sous Blender avec armature seul</u>.

Il fut d'abord placé une armature manuellement pour chaque scan. Chaque os fut assigné à son groupe de vertex comme sur la Figure 7.7. À l'aide d'une fonction de transformation sous Blender chaque armature fut superposée, de manière à opérer un premier « recalage » qui aurait permis de faciliter l'appariement des points de correspondances.

Comme la Figure 7.9 et Figure 7.10 ci-dessous le présente, les résultats sont décevants. De nombreux artefacts de recalage causés par la pondération de chaque os sont observés (en rouge par exemple). Ces artefacts résultent du placement manuel des armatures sans information de profondeur (données manquantes, Figure 7.11), avec pour chaque armature, une échelle et un repère différent. Ces erreurs font partie des limitations de cette méthode.

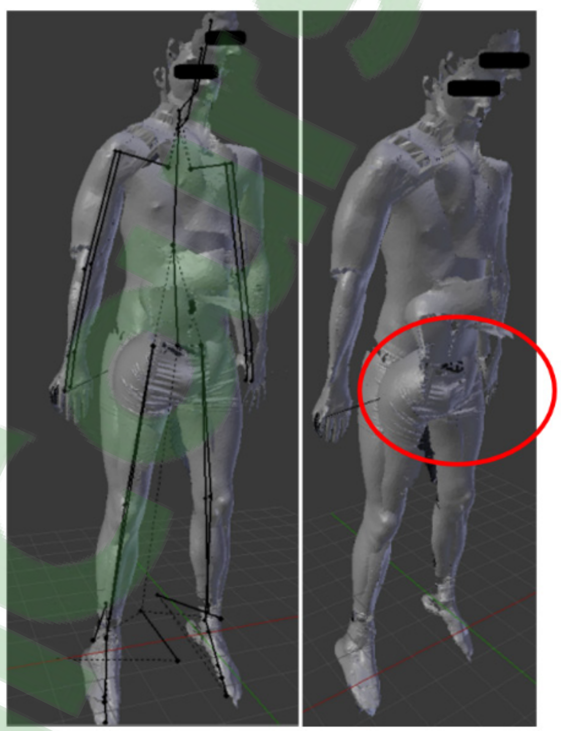


Figure 7.9 Étude préliminaire avec le Go!Scan, résultats premier essai

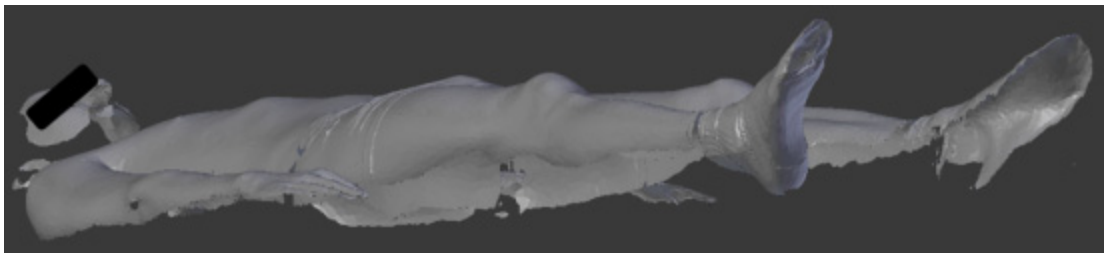


Figure 7.10 Étude préliminaire avec le Go!Scan, problématique du faible recouvrement de surface et manque de données important

2) Deuxième essai du pré-recalage sous Blender avec armature et modèle.

Ensuite, il fut tenté d'effectuer le même processus, mais en corrigeant certaines erreurs dues aux différences d'échelles et de rotations. Pour cela, un modèle générique fut utilisé pour assurer l'orientation de chaque os suivant un même axe de rotation. De plus, une seule armature commune fut placée dans chacun des scans comme sur la Figure 7.11 ci-dessous. L'utilisation d'une seule armature permettait une même mise à l'échelle et une même orientation pour chaque scan. Le modèle permettait d'aider au placement de l'armature (informations de profondeur).

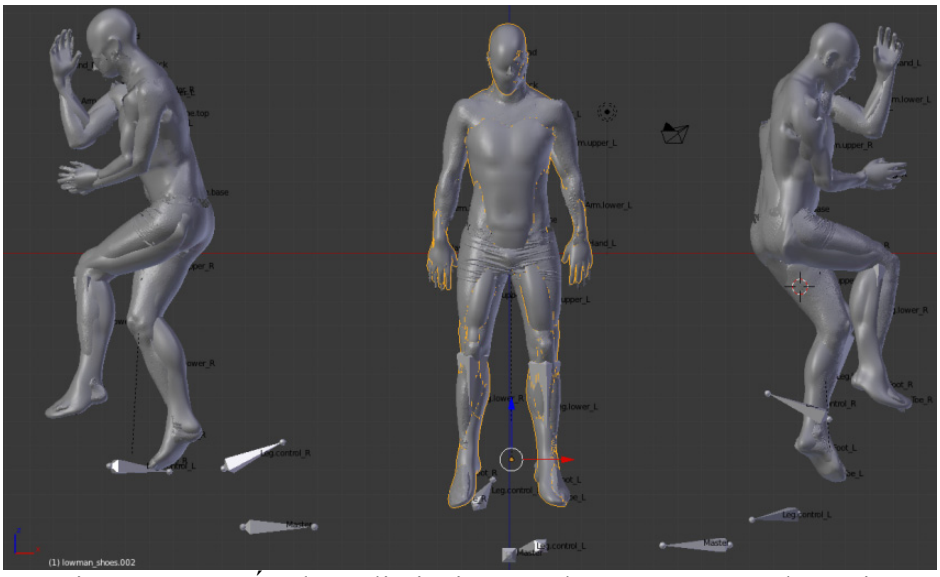


Figure 7.11 Étude préliminaire avec le Go!Scan, vue des trois scans avec modèle générique

Là encore, les résultats furent décevants (Figure 7.12). La pondération de chaque os et les rotations lors de la superposition des armatures sont des paramètres difficiles à contrôler.

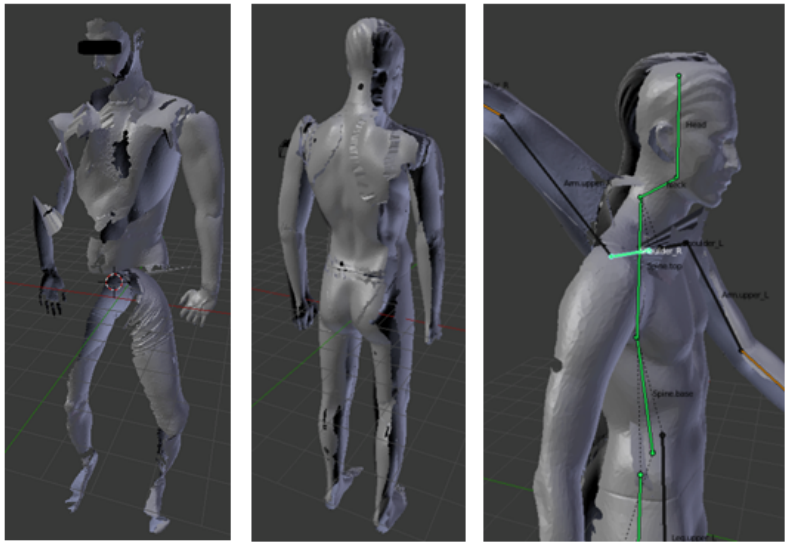


Figure 7.12 Étude préliminaire avec le Go!Scan,
résultats deuxième essai à gauche.
Résultats mannequins au centre et à droite

La caractéristique particulière de ces données par rapport à celle de la littérature est qu'elles manquent énormément d'informations pour guider la déformation. Associés aux changements de poses significatives, l'appariement des maillages devient impossible. Actuellement dans la revue de la littérature il n'existe aucune étude similaire et donc peu d'informations sur ce type d'approche. La plupart n'ont pas de changement de poses et les données ne sont pas aussi incomplètes.

Un troisième essai fut réalisé dans le but de scanner un sujet sain debout sans contrainte de pose.

Le scanner à lumière blanche étant très sensible à la réflexion de la lumière, certains tissus ou régions du corps étaient impossibles à scanner. Pour pallier cette contrainte, le sujet devait porter une combinaison en tissus. Ce tissu favorisait l'acquisition. Cependant, la non-rigidité

Clicours.COM

du corps humain, associé aux artefacts de mouvements ont conduit à l'incapacité du Go!Scan à scanner un sujet sain debout (Figure 7.13).

Figure 7.13 Étude préliminaire avec le Go!Scan, résultats d'un scan d'un sujet sain debout sans contraintes de pose

3. Discussion et conclusion

Dans un cas réel, le patient est couché sur un film plastique sur le dos dans un premier temps, puis il est tourné sur les côtés (Figure 7.14). D'une part, le fait que les patients soient allongés ou assis conduit évidemment à des changements de pose et obstructions lors du scan avec le numériseur. D'autre part, les réflexions de la lumière du plastique sur le scanner peuvent fortement gêner l'acquisition et même empêcher le scan du patient. De plus, la salle de chirurgie comprend de nombreux appareils médicaux. Le patient lui-même est doté de dispositifs médicaux nécessaires à sa survie tels que : le tube de l'appareil à oxygène placé par intubation dans la trachée du patient, les électrodes ECG (pour la représentation de l'activité électrique du cœur) placées à différentes parties du corps, ou encore des cathéters (dispositifs

médicaux destinés à être insérés dans un vaisseau sanguin et permettant le drainage ou l'infusion de liquides) qui sont placés sur le dessus des mains.

Enfin, il très difficile de scanner un sujet sain, et ce, dans des conditions saines, à savoir une salle vide. Comme il l'a été constaté dans cette étude préliminaire, le scanner n'est pas capable de scanner les régions où la pilosité est abondante tels que les cheveux ou les poils. Dans certains cas le scanner était incapable de scanner certains tissus. Ce qui a amené à penser la réelle possibilité de scanner un grand brûlé avec un scanner 3D. De plus, l'acquisition des trois parties (vue de face et de côtés) nécessite plus de 30 min. Associé au temps de reconstruction des modèles, cette approche dépasserait largement la contrainte clinique de temps (<15min).

Figure 7.14 Exemple d'une table de chirurgie d'hôpital, Montréal

Les inconvénients de cette approche sont nombreux : temps d'acquisition et de traitement nécessairement long, faible recouvrement entre les vues, peu d'informations de surfaces dues aux obstructions, changement de pose entre les deux vues prises de côté qui est trop grand pour permettre un alignement, etc. En somme, ces résultats permettent de conclure que cette approche ne permettrait pas de répondre aux besoins et exigences cliniques de manières satisfaisantes, intéressante et bénéfique pour les grands brûlés. Ces travaux ont été réalisés en collaboration avec la Chaire de recherche industrielle CRSNG-Creaform sur la numérisation 3D de l'Université Laval.

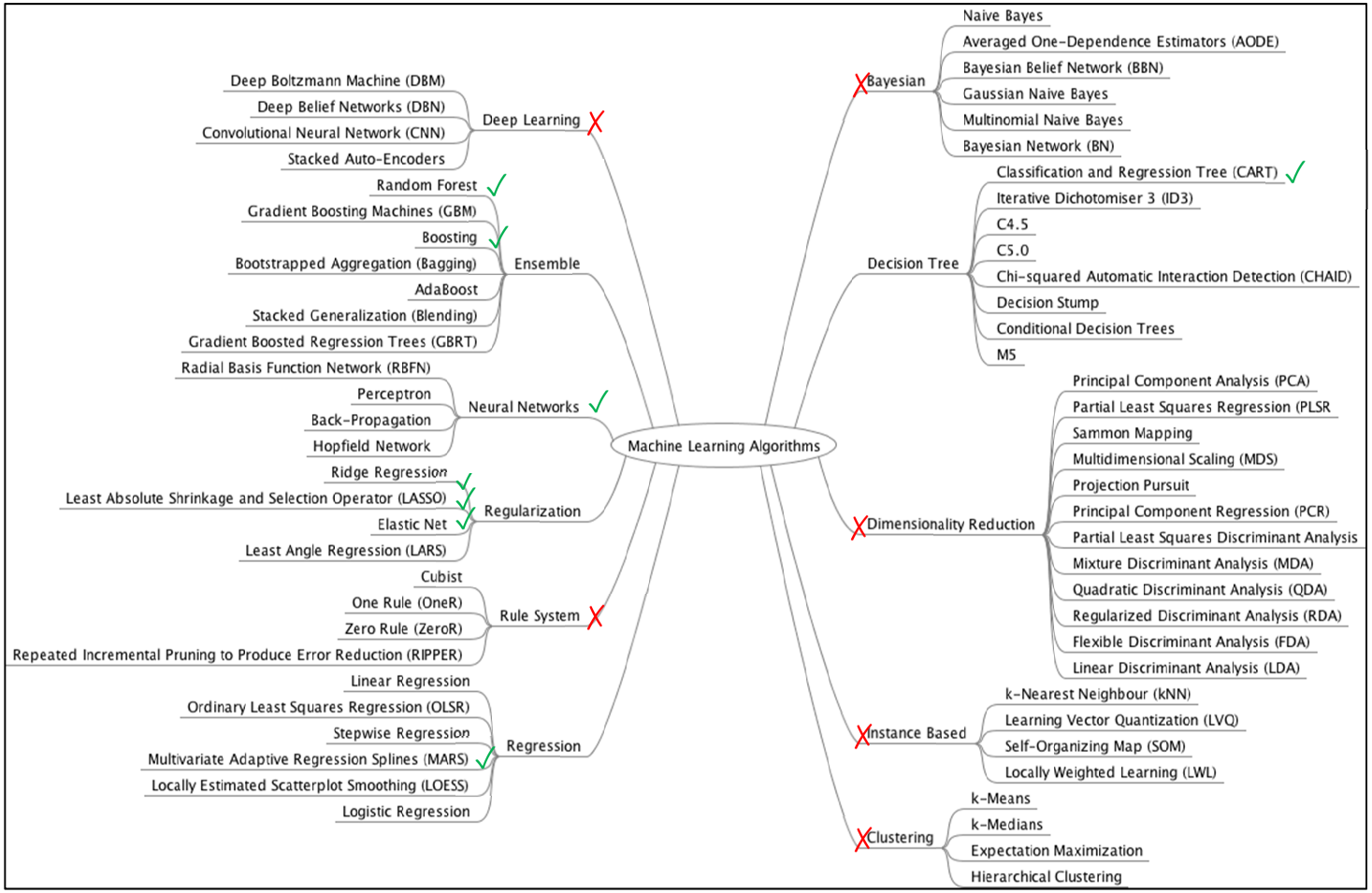


Figure 7.15 Carte heuristique des différents algorithmes en apprentissage machine. Tiré de Brownlee, (s.d.). "Machine Learning Mastery" avec l'autorisation de l'auteur. En ligne <https://machinelearningmastery.com>. Consulté le 16 Novembre 2017.
✓ : algorithmes utilisés. ✗ : non applicable au projet de recherche

BIBLIOGRAPHIE

Adler, A. C., Nathanson, B. H., Raghunathan, K., & McGee, W. T. (2012). Misleading indexed hemodynamic parameters: the clinical importance of discordant BMI and BSA at extremes of weight. *Critical Care, 16*(6), 471. doi: 10.1186/cc11876. Repéré à http://dx.doi.org/10.1186/cc11876

Adler, A. C., Nathanson, B. H., Raghunathan, K., & McGee, W. T. (2013). Effects of Body Surface Area-Indexed Calculations in the Morbidly Obese: A Mathematical Analysis. *Journal of Cardiothoracic and Vascular Anesthesia, 27*(6), 1140-1144. doi: https://doi.org/10.1053/j.jvca.2013.06.011. Repéré à http://www.sciencedirect.com/science/article/pii/S105307701300373X

Ahn, C. S., Oneill, S. P., & Maitz, P. K. (2011). Improving accuracy of burn referrals through the use of an internet-based burns chart. *European Journal of Plastic Surgery, 34*(5), 331-335. doi: 10.1007/s00238-010-0519-0. Repéré à https://doi.org/10.1007/s00238-010-0519-0

Ahn, Y., & Garruto, R. M. (2008). Estimations of body surface area in newborns. *Acta paediatrica, 97*(3), 366-370.

Aitpayev, K., & Gaber, J. (2012). Creation of 3d human avatar using kinect. *Asian Transactions on Fundamentals of Electronics, Communication & Multimedia, 1*(5), 12-24.

Akinwande, M. O., Dikko, H. G., & Samson, A. (2015). Variance inflation factor: As a condition for the inclusion of suppressor variable (s) in regression analysis. *Open Journal of Statistics, 5*(07), 754.

Allen, B. (2005). *Learning body shape models from real-world data* (Citeseer).

Allen, B., Curless, B., & Popović, Z. (2003). The space of human body shapes: reconstruction and parameterization from range scans. Dans *ACM Transactions on Graphics (TOG)* (Vol. 22, pp. 587-594). ACM.

Alpaydin, E. (2014). *Introduction to machine learning*. MIT press.

American Burn Association & American College of Surgeons. (2007). Guidelines for the operation of burn centers. *Journal of burn care & research: official publication of the American Burn Association, 28*(1), 134.

Anderson, D., Feldblum, S., Modlin, C., Schirmacher, D., Schirmacher, E., & Thandi, N. (2005). A Practitioner's Guide to Generalized Linear Models: A Foundation for Theory. *Interpretation and Application, Watson Wyatt Worldwide, London*.

Aspert, N., Santa-Cruz, D., & Ebrahimi, T. (2002). Mesh: Measuring errors between surfaces using the hausdorff distance. Dans *Multimedia and Expo, 2002. ICME'02. Proceedings. 2002 IEEE International Conference on* (Vol. 1, pp. 705-708). IEEE.

Awad, M., & Khanna, R. (2015). Support Vector Regression. Dans *Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers* (pp. 67-80). Berkeley, CA: Apress. doi: 10.1007/978-1-4302-5990-9_4. Repéré à https://doi.org/10.1007/978-1-4302-5990-9_4

Babyak, M. A. (2004). What You See May Not Be What You Get: A Brief, Nontechnical Introduction to Overfitting in Regression-Type Models. *Psychosomatic Medicine, 66*(3), 411-421. Repéré à http://journals.lww.com/psychosomaticmedicine/Fulltext/2004/05000/What_You_See_May_Not_Be_What_You_Get__A_Brief,.21.aspx

Baek, S.-Y., & Lee, K. (2012). Parametric human body shape modeling framework for human-centered product design. *Computer-Aided Design, 44*(1), 56-67.

Bailey, B., & Briars, G. (1996). Estimating the surface area of the human body. *Statistics in medicine, 15*(13), 1325-1332.

Baptista, J. P., Udy, A. A., Sousa, E., Pimentel, J., Wang, L., Roberts, J. A., & Lipman, J. (2011). A comparison of estimates of glomerular filtration in critically ill patients with augmented renal clearance. *Critical Care, 15*(3), R139. doi: 10.1186/cc10262. Repéré à http://dx.doi.org/10.1186/cc10262

Barlett, J. E., Kotrlik, J. W., & Higgins, C. C. (2001). Organizational research: Determining appropriate sample size in survey research. *Information technology, learning, and performance journal, 19*(1), 43.

Bastioni, M., Re, S., & Misra, S. (2008). Ideas and methods for modeling 3D human figures: the principal algorithms used by MakeHuman and their implementation in a new approach to parametric modeling. Dans *Proceedings of the 1st Bangalore Annual Compute Conference* (pp. 10). ACM.

Bell, D. A., & Wang, H. (2000). A formalism for relevance and its application in feature subset selection. *Machine learning, 41*(2), 175-195.

Belsley, D. A., Kuh, E., & Welsch, R. E. (1980). Regression diagnostics. J: Wiley and Sons, New York, USA.

Ben-Noun, L., Sohar, E., & Laor, A. (2001). Neck circumference as a simple screening measure for identifying overweight and obese patients. *Obesity research, 9*(8), 470-477.

Berry, M. G., Goodwin, T. I., Misra, R. R., & Dunn, K. W. (2006). Digitisation of the total burn surface area. *Burns, 32*(6), 684-688. doi: http://dx.doi.org/10.1016/j.burns.2006.01.013. Repéré à http://www.sciencedirect.com/science/article/pii/S0305417906000271

Bocco, M., Willington, E., & Arias, M. (2010). Comparison of regression and neural networks models to estimate solar radiation. *Chilean Journal of Agricultural Research, 70*(3), 428-435.

Borra, S., & Di Ciaccio, A. (2010). Measuring the prediction error. A comparison of cross-validation, bootstrap and covariance penalty methods. *Computational Statistics & Data Analysis, 54*(12), 2976-2989. doi: http://dx.doi.org/10.1016/j.csda.2010.03.004. Repéré à http://www.sciencedirect.com/science/article/pii/S0167947310001064

Breiman, L. (2001a). Random forests. *Machine learning, 45*(1), 5-32.

Breiman, L. (2001b). Statistical modeling: The two cultures (with comments and a rejoinder by the author). *Statistical science, 16*(3), 199-231.

Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). *Classification and regression trees*. CRC press.

Brown, N., White, J., Milligan, A., Risius, D., Ayres, B., Hedger, W., & Scurr, J. (2012). The relationship between breast size and anthropometric characteristics. *American Journal of Human Biology, 24*(2), 158-164.

Brownlee, J. (s.d.). Machine Learning Mastery. Repéré à En ligne <https://machinelearningmastery.com>. Consulté le 16 Novembre 2017.

Bureau de normalisation du Québec (BNQ). (1995). Anthropométrie – Définition et désignation des mesures. *2e édition. NQ, 9990-100. Québec.*, 138.

Buys, K., Deun, D., Laet, T., & Bruyninckx, H. (2011). On-line generation of customized human models based on camera measurements. Dans *International Symposium on Digital Human Modeling*.

Cassola, V., de Melo Lima, V., Kramer, R., & Khoury, H. (2009). FASH and MASH: female and male adult human phantoms based on polygon mesh surfaces: I. Development of the anatomy. *Physics in medicine and biology, 55*(1), 133.

Celoria, G., Steingrub, J. S., Vickers-Lahti, M., Teres, D., Stein, K. L., Fink, M., & Friedmann, P. (1990). Clinical assessment of hemodynamic values in two surgical intensive care units. Effects on therapy. *Arch Surg, 125*(8), 1036-1039.

Centre d'expertise pour les victimes de brûlures graves de l'Est du Québec. (2015). Programmation intégrée du Centre d'expertise pour les victimes de brûlures graves de l'Est du Québec. *Québec : CEVBGEQ*, 155.

Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)?–Arguments against avoiding RMSE in the literature. *Geoscientific model development, 7*(3), 1247-1250.

Chan, K., Lee, T.-W., Sample, P. A., Goldbaum, M. H., Weinreb, R. N., & Sejnowski, T. J. (2002). Comparison of machine learning and traditional classifiers in glaucoma diagnosis. *IEEE Transactions on Biomedical Engineering, 49*(9), 963-974.

Chan, Q. E., Barzi, F., Cheney, L., Harvey, J. G., & Holland, A. J. A. (2012). Burn size estimation in children: Still a problem. *EMA - Emergency Medicine Australasia, 24*(2), 181-186. doi: 10.1111/j.1742-6723.2011.01511.x. Repéré à http://www.scopus.com/inward/record.url?eid=2-s2.0-84859424605&partnerID=40&md5=50ffd4f3aa7e3014f9eab7cd53235793

Chrislb. (2005). Schéma d'un neurone artificiel. License Creative Commons Attribution-Share Alike 3.0 Unported. https://commons.wikimedia.org/wiki/File:ArtificialNeuronModel_english.png.

Chu, C.-H., Tsai, Y.-T., Wang, C. C., & Kwok, T.-H. (2010). Exemplar-based statistical model for semantic parametric design of human body. *Computers in Industry, 61*(6), 541-549.

Cignoni, P., Callieri, M., Corsini, M., Dellepiane, M., Ganovelli, F., & Ranzuglia, G. (2008). Meshlab: an open-source mesh processing tool. Dans *Eurographics Italian chapter conference* (Vol. 2008, pp. 129-136).

Conil, J. M., Georges, B., Fourcade, O., Seguin, T., Lavit, M., Samii, K., . . . Saivin, S. (2007). Assessment of renal function in clinical practice at the bedside of burn patients. *British Journal of Clinical Pharmacology, 63*(5), 583-594. doi: 10.1111/j.1365-2125.2006.02807.x. Repéré à http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2000748/

Connors, A. F., Jr., McCaffree, D. R., & Gray, B. A. (1983). Evaluation of right-heart catheterization in the critically ill patient without acute myocardial infarction. *N Engl J Med, 308*(5), 263-267. doi: 10.1056/nejm198302033080508

Cui, Y., Chang, W., Nöll, T., & Stricker, D. (2013). KinectAvatar: Fully Automatic Body Capture Using a Single Kinect. Dans J.-I. Park & J. Kim (Éds.), *Computer Vision -*

ACCV 2012 Workshops: ACCV 2012 International Workshops, Daejeon, Korea, November 5-6, 2012, Revised Selected Papers, Part II (pp. 133-147). Berlin, Heidelberg: Springer Berlin Heidelberg. doi: 10.1007/978-3-642-37484-5_12. Repéré à http://dx.doi.org/10.1007/978-3-642-37484-5_12

Daanen, H. A. M., & Van De Water, G. J. (1998). Whole body scanners. *Displays, 19*(3), 111-120. Repéré à https://www.scopus.com/inward/record.uri?eid=2-s2.0-0032204505&partnerID=40&md5=4dca8f23f6c35504d7fd781aebfd4ded

Dalal, D. K., & Zickar, M. J. (2012). Some common myths about centering predictor variables in moderated multiple regression and polynomial regression. *Organizational Research Methods, 15*(3), 339-362.

Davy-Jow, S. L., Lees, D. M. B., & Russell, S. (2013). Virtual forensic anthropology: Novel applications of anthropometry and technology in a child death case. *Forensic science international, 224*(1), e7-e10. doi: 10.1016/j.forsciint.2012.11.002. Repéré à http://dx.doi.org/10.1016/j.forsciint.2012.11.002

De Silva, A. M., & Leong, P. H. (2015). Feature Selection. Dans *Grammar-Based Feature Generation for Time-Series Prediction* (pp. 13-24). Springer.

Derksen, S., & Keselman, H. J. (1992). Backward, forward and stepwise automated subset selection algorithms: Frequency of obtaining authentic and noise variables. *British Journal of Mathematical and Statistical Psychology, 45*(2), 265-282.

Dermot, H., & Enda, F. F. (2014). "Assessing the Precision of Anthropometric Measurements: A Six Sigma Approach". In Advances in Applied Digital Human Modeling, sous la dir. de Duffy Vincent. Dans *5th International Conference on Applied Human Factors and Ergonomics* (Vol. 20, pp. 41-48). AHFE Conference.

Desbois, A., Matei, S., Beguet, F., Perreault, I., Gervais, S., & De Guise, J. (2017). The Importance of a 3D-Based Approach With Personalized Models for Accurately Assessing TBSA. *Journal of Burn Care & Research*.

Dietterich, T. (2003). Nature Encyclopedia of Cognitive Science. *Machine learning. Macmillan*.

Dirnberger, J., Giretzlehner, M., Luckeneder, T., Siegl, D., Haller, H. L., & Rodemund, C. (2004a). BurnCase 3D–Realistic Adaptation of 3-Dimensional Human Body Models. Dans *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2004* (pp. 363-370). Springer.

Dirnberger, J., Giretzlehner, M., Luckeneder, T., Siegl, D., Haller, H. L., & Rodemund, C. (2004b). BurnCase 3D – Realistic Adaptation of 3-Dimensional Human Body Models. Dans C. Barillot, D. R. Haynor & P. Hellier (Éds.), *Medical Image Computing and*

Computer-Assisted Intervention – MICCAI 2004: 7th International Conference, Saint-Malo, France, September 26-29, 2004. Proceedings, Part II (pp. 363-370). Berlin, Heidelberg: Springer Berlin Heidelberg. doi: 10.1007/978-3-540-30136-3_45. Repéré à http://dx.doi.org/10.1007/978-3-540-30136-3_45

Dirnberger, J., Giretzlehner, M., Ruhmer, M., Haller, H., & Rodemund, C. (2002). Modelling human burn injuries in a three-dimensional virtual environment. *Studies in health technology and informatics, 94*, 52-58.

Dirnberger, J., Michael, G., Robert, O., Herbert, H., & Peter, K. L. (2012). Comments on "A system for 3D representation of burns and calculation of burnt skin area". *Burns, 38*(7), 1091-1093.

Du Bois, D., & Du Bois, E. (1916). A formula to estimate the approximate surface area if height and weight be known. *Arch Intern Med, 17*, 863-871.

Du Bois, D., & Du Bois, E. F. (1916). Clinical calorimetry: tenth paper a formula to estimate the approximate surface area if height and weight be known. *Archives of internal medicine, 17*(6_2), 863-871.

Dubin, S., & Zietz, S. (1996). Body surface estimation: a critical evaluation. Dans *Biomedical Engineering Conference, 1996., Proceedings of the 1996 Fifteenth Southern* (pp. 397-400). IEEE.

Durrant, C. A. T., Simpson, A. R., & Williams, G. (2008). Thermal injury – The first 24h. *Current Anaesthesia & Critical Care, 19*(5), 256-263. Repéré à http://linkinghub.elsevier.com/retrieve/pii/S0953711208000926?showall=true

Efron, B. (1982). *The jackknife, the bootstrap and other resampling plans*. SIAM.

Eisenberg, P. R., Jaffe, A. S., & Schuster, D. P. (1984). Clinical evaluation compared to pulmonary artery catheterization in the hemodynamic assessment of critically ill patients. *Crit Care Med, 12*(7), 549-553.

Enquête canadienne sur les mesures de la santé. (2015). Graphique 2: Répartition des adultes de 18 à 79 ans, selon l'indice de masse corporelle (IMC), le sexe et le groupe d'âge, population à domicile, Canada, 2012 à 2013.

Fairbrother, H. (2015). Optimizing Emergency Management To Reduce Morbidity And Mortality In Pediatric Burn Patients.

Fang, J., & Tai, D. (2011). Evaluation of mutual information, genetic algorithm and SVR for feature selection in QSAR regression. *Current drug discovery technologies, 8*(2), 107-111.

Clicours.COM

Field, K. M., Kosmider, S., Jefford, M., Michael, M., Jennens, R., Green, M., & Gibbs, P. (2008). Chemotherapy Dosing Strategies in the Obese, Elderly, and Thin Patient: Results of a Nationwide Survey. *Journal of Oncology Practice, 4*(3), 108-113. doi: 10.1200/jop.0832001. Repéré à http://ascopubs.org/doi/abs/10.1200/JOP.0832001

Finnerty, C. C., Jeschke, M. G., Branski, L. K., Barret, J. P., Dziewulski, P., & Herndon, D. N. (2016). Hypertrophic scarring: the greatest unmet challenge after burn injury. *The Lancet, 388*(10052), 1427-1436.

Flom, P. L., & Cassell, D. L. (2007). Stopping stepwise: Why stepwise and similar selection methods are bad, and what you should use. Dans *NorthEast SAS Users Group Inc 20th Annual Conference: 11-14th November 2007; Baltimore, Maryland*.

Fonseca, T. F., Bogaerts, R., Hunt, J., & Vanhavere, F. (2014). A methodology to develop computational phantoms with adjustable posture for WBC calibration. *Physics in medicine and biology, 59*(22), 6811.

Fox, J., & Monette, G. (1992). Generalized collinearity diagnostics. *Journal of the American Statistical Association, 87*(417), 178-183.

Frank, I. E. (1995). Modern nonlinear regression methods. *Chemometrics and Intelligent Laboratory Systems, 27*(1), 1-19. doi: https://doi.org/10.1016/0169-7439(95)80003-R. Repéré à http://www.sciencedirect.com/science/article/pii/016974399580003R

Friedman, J. H. (1991a). Multivariate adaptive regression splines. *The annals of statistics*, 1-67.

Friedman, J. H. (1991b). *Estimating functions of mixed ordinal and categorical variables using adaptive splines*. STANFORD UNIV CA LAB FOR COMPUTATIONAL STATISTICS.

Friedman, J. H. (2002). Stochastic gradient boosting. *Computational Statistics & Data Analysis, 38*(4), 367-378.

Gao, B., Klumpen, H.-J., & Gurney, H. (2008). Dose calculation of anticancer drugs. *Expert Opinion on Drug Metabolism & Toxicology, 4*(10), 1307-1319.

Gehan, E. A., & George, S. L. (1970). Estimation of human body surface area from height and weight. *Cancer chemotherapy reports. Part 1, 54*(4), 225-235.

Giretzlehner, M., Dirnberger, J., Luckender, T., Haller, H., & Rodemund, C. (2004). BurnCase 3D: A research product for effective and time-saving documentation of burn injuries. *Annals of Burns and Fire Disasters, 17*(2), 64-72.

Giretzlehner, M., Dirnberger, J., Owen, R., Haller, H., Lumenta, D., & Kamolz, L.-P. (2013). The determination of total burn surface area: How much difference? *Burns, 39*(6), 1107-1113.

Gordon, C. C., Blackwell, C. L., Bradtmiller, B., Parham, J. L., Hotzman, J., Paquette, S. P., . . . Hodge, B. M. (2013). *2010 Anthropometric Survey of US Marine Corps Personnel: Methods and Summary Statistics*. DTIC Document.

Gotra, A., Chartrand, G., Massicotte-Tisluck, K., Morin-Roy, F., Vandenbroucke-Menu, F., de Guise, J. A., & Tang, A. (2015). Validation of a Semiautomated Liver Segmentation Method Using CT for Accurate Volumetry. *Academic radiology*.

Gunst, R., & Webster, J. (1975). Regression analysis and problems of multicollinearity. *Communications in Statistics-Theory and Methods, 4*(3), 277-292.

Gurney, H. (2002). How to calculate the dose of chemotherapy. *Br J Cancer, 86*(8), 1297-1302. Repéré à http://dx.doi.org/10.1038/sj.bjc.6600139

Guyon, I., & Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of machine learning research, 3*(Mar), 1157-1182.

Guyon, I., & Elisseeff, A. (2006). An introduction to feature extraction. *Feature extraction*, 1-25.

Haller, H., Dirnberger, J., Giretzlehner, M., Rodemund, C., & Kamolz, L. (2009). "Understanding burns": Research project BurnCase 3D—Overcome the limits of existing methods in burns documentation. *Burns, 35*(3), 311-317.

Han, J. C., Lawlor, D. A., & Kimm, S. Y. S. (2010). Childhood obesity. *The Lancet, 375*(9727), 1737-1748. doi: http://dx.doi.org/10.1016/S0140-6736(10)60171-7. Repéré à http://www.sciencedirect.com/science/article/pii/S0140673610601717

Hansard, M., Lee, S., Choi, O., & Horaud, R. P. (2012). *Time-of-flight cameras: principles, methods and applications*. Springer Science & Business Media.

Harish, V., Raymond, A. P., Issler, A. C., Lajevardi, S. S., Chang, L.-Y., Maitz, P. K. M., & Kennedy, P. (2015). Accuracy of burn size estimation in patients transferred to adult Burn Units in Sydney, Australia: An audit of 698 patients. *Burns, 41*(1), 91-99. doi: http://dx.doi.org/10.1016/j.burns.2014.05.005. Repéré à //www.sciencedirect.com/science/article/pii/S0305417914001703

Harrell, F. E. (2001). Regression modeling strategies, with applications to linear models, survival analysis and logistic regression. *GET ADDRESS: Springer*.

Haslam, D. W., & James, W. P. T. (2005). Obesity. *The Lancet, 366*(9492), 1197-1209. doi: http://dx.doi.org/10.1016/S0140-6736(05)67483-1. Repéré à http://www.sciencedirect.com/science/article/pii/S0140673605674831

Hatipoglu, N., Mazicioglu, M. M., Kurtoglu, S., & Kendirci, M. (2010). Neck circumference: an additional tool of screening overweight and obesity in childhood. *European journal of pediatrics, 169*(6), 733-739.

Haycock, G. B., Schwartz, G. J., & Wisotsky, D. H. (1978). Geometric method for measuring body surface area: a height-weight formula validated in infants, children, and adults. *The Journal of pediatrics, 93*(1), 62-66.

Heimann, T., Morrison, B. J., Styner, M. A., Niethammer, M., & Warfield, S. (2010). Segmentation of knee images: A grand challenge. Dans *Proc. MICCAI Workshop on Medical Image Analysis for the Clinic* (pp. 207-214).

Heimann, T., Van Ginneken, B., Styner, M., Arzhaeva, Y., Aurich, V., Bauer, C., . . . Bekes, G. (2009). Comparison and evaluation of methods for liver segmentation from CT datasets. *Medical Imaging, IEEE Transactions on, 28*(8), 1251-1265.

Helland, I. S. (1990). Partial least squares regression and statistical models. *Scandinavian Journal of Statistics*, 97-114.

Herrera-Gutierrez, M. E., Seller-Perez, G., Banderas-Bravo, E., Munoz-Bono, J., Lebron-Gallardo, M., & Fernandez-Ortega, J. F. (2007). Replacement of 24-h creatinine clearance by 2-h creatinine clearance in intensive care unit patients: a single-center study. *Intensive Care Med, 33*(11), 1900-1906. doi: 10.1007/s00134-007-0745-5

Hettiaratchy, S., & Dziewulski, P. (2004). Introduction. *BMJ : British Medical Journal, 328*(7452), 1366-1368. Repéré à http://www.ncbi.nlm.nih.gov/pmc/articles/PMC420296/

Hettiaratchy, S., & Papini, R. (2004). Initial management of a major burn: II—assessment and resuscitation. *BMJ : British Medical Journal, 329*(7457), 101-103. Repéré à http://www.ncbi.nlm.nih.gov/pmc/articles/PMC449823/

Hidvegi, N., Nduka, C., Myers, S., & Dziewulski, P. (2004). Estimation of breast burn size. *Plastic and reconstructive surgery, 113*(6), 1591-1597.

Hingorjo, M. R., Qureshi, M. A., & Mehdi, A. (2012). Neck circumference as a useful marker of obesity: a comparison with body mass index and waist circumference. *JPMA-Journal of the Pakistan Medical Association, 62*(1), 36.

Ho, Y.-C. (1999). The no free lunch theorem and the human-machine interface. *IEEE Control systems, 19*(3), 8-10.

Hoste, E. A., Damen, J., Vanholder, R. C., Lameire, N. H., Delanghe, J. R., Van den Hauwe, K., & Colardyn, F. A. (2005). Assessment of renal function in recently admitted critically ill patients with normal serum creatinine. *Nephrol Dial Transplant, 20*(4), 747-753. doi: 10.1093/ndt/gfh707

Hotzman, J., Gordon, C. C., Bradtmiller, B., Corner, B. D., Mucher, M., Kristensen, S., . . . Blackwell, C. L. (2011). *Measurer's Handbook: US Army and Marine Corps Anthropometric Surveys, 2010-2011*. DTIC Document.

Hunter, R. J., Navo, M. A., Thaker, P. H., Bodurka, D. C., Wolf, J. K., & Smith, J. A. (2009). Dosing chemotherapy in obese patients: Actual versus assigned body surface area (BSA). *Cancer Treatment Reviews, 35*(1), 69-78. doi: https://doi.org/10.1016/j.ctrv.2008.07.005. Repéré à http://www.sciencedirect.com/science/article/pii/S030573720800265X

ISO13053-2. (2011). Quantitative methods in process improvement -- Six Sigma -- Part 2: Tools and techniques. (1), 49.

ISO, I. (2008). 7250–1: Basic Human Body Measurements for Technological Design—Part 1: Body Measurement Definitions and Landmarks. *International Organization for Standardization, Geneva, Switzerland*.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning with Applications in R* (Vol. 112). Springer.

James, P. T. (2004). Obesity: The worldwide epidemic. *Clinics in Dermatology, 22*(4), 276-280. doi: http://dx.doi.org/10.1016/j.clindermatol.2004.01.010. Repéré à http://www.sciencedirect.com/science/article/pii/S0738081X04000112

Jannin, P., Fitzpatrick, J. M., Hawkes, D., Pennec, X., Shahidi, R., & Vannier, M. (2002). Validation of medical image processing in image-guided therapy. *IEEE Transactions on Medical Imaging, 21*(12), 1445-1449.

John, G. H., Kohavi, R., & Pfleger, K. (1994). Irrelevant features and the subset selection problem. Dans *Machine learning: proceedings of the eleventh international conference* (pp. 121-129).

Jose, R. M., Roy, D. K., Vidyadharan, R., & Erdmann, M. (2004). Burns area estimation—an error perpetuated. *Burns, 30*(5), 481-482.

Kamolz, L.-P., Parvizi, D., Giretzlehner, M., & Lumenta, D. B. (2014). Burn surface area calculation: what do we need in future. *Burns: journal of the International Society for Burn Injuries, 40*(1), 171.

Kamolz, L.-P., Wurzer, P., Giretzlehner, M., Parvizi, D., & Lumenta, D. B. (2014). Burn surface area calculation instead of burn size estimation: our opinion. *Burns, 40*(8), 1813-1814.

Karagiannopoulos, M., Anyfantis, D., Kotsiantis, S., & Pintelas, P. (2007). Feature selection for regression problems. *Proceedings of the 8th Hellenic European Research on Computer Mathematics & its Applications, Athens, Greece, 2022*.

Kasap, M., & Magnenat-Thalmann, N. (2007). Parameterized human body model for real-time applications. Dans *Cyberworlds, 2007. CW'07. International Conference on* (pp. 160-167). IEEE.

Kennedy, B. B., Baird, S. M., & Troiano, N. H. (2008). Burn injuries and pregnancy. *The Journal of perinatal & neonatal nursing, 22*(1), 21-30.

Khalid, S., Khalil, T., & Nasreen, S. (2014). A survey of feature selection and feature extraction techniques in machine learning. Dans *Science and Information Conference (SAI), 2014* (pp. 372-378). IEEE.

Khongma, A., Ruchanurucks, M., Koanantakool, T., Phatrapornnant, T., Koike, Y., & Rakprayoon, P. (2014) Kinect quality enhancement for triangular mesh reconstruction with a medical image application. du volume)*: Vol. 543. Studies in Computational Intelligence* (pp. 15-32).

Klein, M. B., Hayden, D., Elson, C., Nathens, A. B., Gamelli, R. L., Gibran, N. S., . . . Schoenfeld, D. (2007). The association between fluid administration and outcome following major burn: a multicenter study. *Annals of surgery, 245*(4), 622-628.

Knaysi, G. A., Crikelair, G. F., & Cosman, B. (1968). The rule of nines: its history and accuracy. *Plastic and reconstructive surgery, 41*(6), 560-563.

Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. Dans *Ijcai* (Vol. 14, pp. 1137-1145). Stanford, CA.

Kohavi, R., & John, G. H. (1997). Wrappers for feature subset selection. *Artificial intelligence, 97*(1-2), 273-324.

Kraft, R., Herndon, D. N., Al-Mousawi, A. M., Williams, F. N., Finnerty, C. C., & Jeschke, M. G. (2012). Burn size and survival probability in paediatric patients in modern burn care: a prospective observational cohort study. *The Lancet, 379*(9820), 1013-1021.

Kulaga, A., Myers, T., & Gillich, P. (2014). Design upgrade to the plywood-penetration mannequin. *BALLISTICS 2014: 28th International Symposium on Ballistics*, 10 pages.

Kumar, D. S., Sukanya, S., & BIT-Campus, T. (2016). FEATURE SELECTION USING MULTIVARIATE ADAPTIVE REGRESSION SPLINES. *International Journal of Research and Reviews in Applied Sciences And Engineering (IJRRASE), Vol 8.*(No.1), p.17-24.

Laurent, P., Cresson, T., Vázquez, C., Hagemeister, N., & de Guise, J. A. (2016). A multi-criteria evaluation platform for segmentation algorithms. Dans *Engineering in Medicine and Biology Society (EMBC), 2016 IEEE 38th Annual International Conference of the* (pp. 6441-6444). IEEE.

Leary, T., Milner, Q., & Niblett, D. (2000). The accuracy of the estimation of body weight and height in the intensive care unit. *European journal of anaesthesiology, 17*(11), 698-703.

Lee, J.-Y., Choi, J.-W., & Kim, H. (2008). Determination of Body Surface Area and Formulas to Estimate Body Surface Area Using the Alginate Method. *Journal of PHYSIOLOGICAL ANTHROPOLOGY, 27*(2), 71-82. doi: 10.2114/jpa2.27.71

Leedy, P. D., & Ormrod, J. E. (2013). *Practical research : planning and design*. Boston: Pearson.

Lehot, J.-J., & Arvieux, C. (2011). *Réanimation et urgences*. Springer Science & Business Media.

Lesaffre, E., & Marx, B. D. (1993). Collinearity in generalized linear regression. *Communications in Statistics-Theory and Methods, 22*(7), 1933-1952.

Li, J., & Wang, Y. (2007). Automatically constructing skeletons and parametric structures for polygonal human bodies. Dans *Proceedings of computer graphics international* (Vol. 7, pp. 151-156).

Li, L. (2014). Time-of-Flight Camera–An Introduction. *Technical White Paper, May*.

Liaw, A., & Wiener, W. (2002). Classification and Regression by randomForest. R News 2(3), 18--22. . Repéré à http://CRAN.R-project.org/doc/Rnews/

Lindsey, J. K. (2000). *Applying generalized linear models*. Springer Science & Business Media.

Liu, N. T., & Salinas, J. (2015). Machine learning in burn care and research: A systematic review of the literature. *Burns, 41*(8), 1636-1641. doi: http://dx.doi.org/10.1016/j.burns.2015.07.001. Repéré à http://www.sciencedirect.com/science/article/pii/S0305417915002004

Livingston, E. H., & Lee, S. (2000). Percentage of Burned Body Surface Area Determination in Obese and Nonobese Patients. *Journal of surgical research, 91*(2), 106-110. doi: http://dx.doi.org/10.1006/jsre.2000.5909. Repéré à http://www.sciencedirect.com/science/article/pii/S0022480400959092

Lund, C. C. (1944). The estimation of areas of burns. *Surg Gynecol Obste, 79*, 352-358. Repéré à http://ci.nii.ac.jp/naid/20001139318/en/

Lund, C. C., & Browder, N. C. (1944). The estimation of areas of burns. *Surg Gynecol Obstet, 79*(352), 8.

Madarnas, Y., Sawka, C. A., Franssen, E., & Bjarnason, G. A. (2001). Are medical oncologists biased in their treatment of the large woman with breast cancer? *Breast Cancer Research and Treatment, 66*(2), 123-133. doi: 10.1023/a:1010635328299. Repéré à http://dx.doi.org/10.1023/A:1010635328299

Magnenat-Thalmann, N., Seo, H., & Cordier, F. (2004). Automatic modeling of virtual humans and body clothing. *Journal of Computer Science and Technology, 19*(5), 575-584.

MakeHuman [free and open source 3D computer graphics software]. (2014). Retrieved from http://www.makehumancommunity.org/.

Malic, C., Karoo, R., Austin, O., & Phipps, A. (2007). Resuscitation burn card—a useful tool for burn injury assessment. *Burns, 33*(2), 195-199.

Martin, J. H., Fay, M. F., Udy, A., Roberts, J., Kirkpatrick, C., Ungerer, J., & Lipman, J. (2011). Pitfalls of using estimations of glomerular filtration rate in an intensive care population. *Internal medicine journal, 41*(7), 537-543.

Max, K., Jed, W., Steve, W., Andre, W., Chris, K., Allan, E., . . . Tyler, H. (2017). caret: Classification and Regression Training. R package version 6.0-76. https://CRAN.R-project.org/package=caret.

McCullagh, P. (1984). Generalized linear models. *European Journal of Operational Research, 16*(3), 285-292. doi: https://doi.org/10.1016/0377-2217(84)90282-0. Repéré à http://www.sciencedirect.com/science/article/pii/0377221784902820

Mendes-Moreira, J., Soares, C., Jorge, A. M., & Sousa, J. F. D. (2012). Ensemble approaches for regression: A survey. *ACM Computing Surveys (CSUR), 45*(1), 10.

Meunier, P., Shu, C., & Xi, P. (2009). Revealing the internal structure of human variability for design purposes. Dans *IEA World Congress on Ergonomics*.

Milborrow, S. (2017). earth: Multivariate Adaptive Regression Splines. R package version 4.5.0. https://CRAN.R-project.org/package=earth.

Mimoz, O., Rauss, A., Rekik, N., Brun-Buisson, C., Lemaire, F., & Brochard, L. (1994). Pulmonary artery catheterization in critically ill patients: a prospective analysis of outcome changes associated with catheter-prompted changes in therapy. *Crit Care Med, 22*(4), 573-579.

Morozova, O., Levina, O., Uusküla, A., & Heimer, R. (2015). Comparison of subset selection methods in linear regression in the context of health-related quality of life and substance abuse in Russia. *BMC Medical Research Methodology, 15*, 71. doi: 10.1186/s12874-015-0066-2. Repéré à http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4553217/

Mosteller, R. (1987). Simplified calculation of body-surface area. *N Engl J Med, 317*, 1098.

Mukaka, M. M. (2012). A guide to appropriate use of Correlation coefficient in medical research. *Malawi Medical Journal : The Journal of Medical Association of Malawi, 24*(3), 69-71. Repéré à http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3576830/

Nadadur, G., & Parkinson, M. B. (2010). Consideration of demographics and variance in regression approaches to estimating body dimensions for spatial analysis of design. *Journal of Mechanical Design, 132*(2), 021007.

Nawi, N. M., Atomi, W. H., & Rehman, M. Z. (2013). The Effect of Data Pre-processing on Optimized Training of Artificial Neural Networks. *Procedia Technology, 11*, 32-39. doi: http://dx.doi.org/10.1016/j.protcy.2013.12.159. Repéré à http://www.sciencedirect.com/science/article/pii/S2212017313003137

Neaman, K. C., Andres, L. A., McClure, A. M., Burton, M. E., Kemmeter, P. R., & Ford, R. D. (2011). A new method for estimation of involved BSAs for obese and normal-weight patients with burn injury. *Journal of Burn Care & Research, 32*(3), 421-428.

Neuwalder, J., Sampson, C., Breuing, K., & Orgill, D. (2002). A review of computer-aided body surface area determination: SAGE II and EPRI's 3D Burn Vision. *Journal of Burn Care & Research, 23*(1), 55-59.

Nwoye, L. O. (1989). Body surface area of Africans: a study based on direct measurements of Nigerian males. *Human biology*, 439-457.

Osborne, J. W., & Waters, E. (2002). Four Assumptions Of Multiple Regression That Researchers Should Always Test.

Osei-Bryson, K.-M. (2014). Overview on multivariate adaptive regression splines. Dans *Advances in Research Methods for Information Systems Research* (pp. 93-107). Springer.

Parkinson, M. B., & Reed, M. P. (2010). Creating virtual user populations by analysis of anthropometric data. *International Journal of Industrial Ergonomics, 40*(1), 106-111.

Parvizi, D., Giretzlehner, M., Dirnberger, J., Owen, R., Haller, H. L., Schintler, M. V., . . . Kamolz, L. P. (2014). The use of telemedicine in burn care: development of a mobile system for TBSA documentation and remote assessment. *Annals of Burns and Fire Disasters, 27*(2), 94-100.

Parvizi, D., Giretzlehner, M., Wurzer, P., Klein, L. D., Shoham, Y., Bohanon, F. J., . . . Kamolz, L.-P. (2016). BurnCase 3D software validation study: Burn size measurement accuracy and inter-rater reliability. *Burns, 42*(2), 329-335. doi: http://dx.doi.org/10.1016/j.burns.2016.01.008. Repéré à //www.sciencedirect.com/science/article/pii/S0305417916000115

Parvizi, D., Kamolz, L.-P., Giretzlehner, M., Haller, H. L., Trop, M., Selig, H., . . . Lumenta, D. B. (2014). The potential impact of wrong TBSA estimations on fluid resuscitation in patients suffering from burns: Things to keep in mind. *Burns, 40*(2), 241-245.

Peck, M. D. (2011). Epidemiology of burns throughout the world. Part I: Distribution and risk factors. *Burns, 37*(7), 1087-1100.

Poggio, E. D., Nef, P. C., Wang, X., Greene, T., Van Lente, F., Dennis, V. W., & Hall, P. M. (2005). Performance of the Cockcroft-Gault and modification of diet in renal disease equations in estimating GFR in ill hospitalized patients. *Am J Kidney Dis, 46*(2), 242-252. doi: 10.1053/j.ajkd.2005.04.023

Popescu, C. R., & Lungu, A. (2014). Real-Time 3D Reconstruction Using a Kinect Sensor. *Computer Science and Information Technology, 2*(2), 95-99.

Prasad, A. M., Iverson, L. R., & Liaw, A. (2006). Newer classification and regression tree techniques: bagging and random forests for ecological prediction. *Ecosystems, 9*(2), 181-199.

Prieto, M. F., Acha, B., Gómez-Cía, T., Fondón, I., & Serrano, C. (2011). A system for 3D representation of burns and calculation of burnt skin area. *Burns, 37*(7), 1233-1240. Repéré à http://www.sciencedirect.com/science/article/pii/S0305417911001677

Rae, L., Pham, T. N., Carrougher, G., Honari, S., Gibran, N. S., Arnoldo, B. D., . . . Herndon, D. N. (2013). Differences in resuscitation in morbidly obese burn patients may contribute to high mortality. *Journal of burn care & research : official publication of the American Burn Association, 34*(5), 507-514. doi: 10.1097/BCR.0b013e3182a2a771. Repéré à http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3800202/

Ranger, F. (2010). Évaluation de l'offre anthropométrique pour la conception de produits.

Rawlings, J. O., Pantula, S. G., & Dickey, D. A. (2001). *Applied regression analysis: a research tool*. Springer Science & Business Media.

Richard, R., Jones, J. A., & Parshley, P. (2015). Hierarchical Decomposition of Burn Body Diagram Based on Cutaneous Functional Units and Its Utility. *Journal of Burn Care & Research, 36*(1), 33-43.

Ridgeway, G. (2017). gbm: Generalized Boosted Regression Models. R package version 2.1.3. Repéré à https://CRAN.R-project.org/package=gbm

Roberts, J. A., Kruger, P., Paterson, D. L., & Lipman, J. (2008). Antibiotic resistance--what's dosing got to do with it? *Crit Care Med, 36*(8), 2433-2440. doi: 10.1097/CCM.0b013e318180fe62

Robinette, K. M., Daanen, H., & Paquet, E. (1999). The CAESAR project: a 3-D surface anthropometry survey. Dans *3-D Digital Imaging and Modeling, 1999. Proceedings. Second International Conference on* (pp. 380-386). IEEE.

Rossiter, N., Chapman, P., & Haywood, I. (1996). How big is a hand? *Burns, 22*(3), 230-231.

RStudio Team. (2016). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA URL http://www.rstudio.com/.

Ryan, C. M., Schoenfeld, D. A., Thorpe, W. P., Sheridan, R. L., Cassem, E. H., & Tompkins, R. G. (1998). Objective estimates of the probability of death from burn injuries. *New England Journal of Medicine, 338*(6), 362-366.

Saeys, Y., Inza, I., & Larrañaga, P. (2007). A review of feature selection techniques in bioinformatics. *bioinformatics, 23*(19), 2507-2517.

Sarbolandi, H., Lefloch, D., & Kolb, A. (2015). Kinect Range Sensing: Structured-Light versus Time-of-Flight Kinect. *Computer Vision and Image Understanding*.

Seo, H., Cordier, F., Philippon, L., & Magnenat-Thalmann, N. (2001). Interactive modelling of mpeg-4 deformable human body models. Dans *Deformable Avatars* (pp. 120-131). Springer.

Seo, H., & Magnenat-Thalmann, N. (2003). An automatic modeling of human bodies from sizing parameters. Dans *Proceedings of the 2003 symposium on Interactive 3D graphics* (pp. 19-26). ACM.

Seo, H., & Magnenat-Thalmann, N. (2004). An example-based approach to human body manipulation. *Graphical Models, 66*(1), 1-23.

Clicours.COM

Sheng, W.-b., Zeng, D., Wan, Y., Yao, L., Tang, H.-t., & Xia, Z.-f. (2014). BurnCalc assessment study of computer-aided individual three-dimensional burn area calculation. *Journal of translational medicine, 12*(1), 1-12.

Shields, M., & Connor Gorber, S. (2008). Questions de méthodologie en anthropométrie: taille et poids déclarés versus mesurés. *Vol Produit,* (11-522).

Slinker, B., & Glantz, S. (1985). Multiple regression for physiological data analysis: the problem of multicollinearity. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology, 249*(1), R1-R12.

Smola, A. J., & Schölkopf, B. (2004). A tutorial on support vector regression. *Statistics and computing, 14*(3), 199-222.

Sparreboom, A., Wolff, A. C., Mathijssen, R. H., Chatelut, E., Rowinsky, E. K., Verweij, J., & Baker, S. D. (2007). Evaluation of alternate size descriptors for dose calculation of anticancer drugs in the obese. *J Clin Oncol, 25*(30), 4707-4713. doi: 10.1200/jco.2007.11.2938

Steingrub, J. S., Celoria, G., Vickers-Lahti, M., Teres, D., & Bria, W. (1991). Therapeutic impact of pulmonary artery catheterization in a medical/surgical ICU. *Chest, 99*(6), 1451-1455.

T Vu, T. (2002). Standardization of body surface area calculations. *Journal of Oncology Pharmacy Practice, 8*(2-3), 49-54.

Tahir, S., Memon, A. R., Kumar, M., & Ali, S. A. (2009). Prediction of Mortality After Major Burn: Physiological Versus Biochemical Measures. *Wounds, 21*(7), 177-182.

Tao, M. W., Srinivasan, P. P., Malik, J., Rusinkiewicz, S., & Ramamoorthi, R. (2015). Depth from Shading, Defocus, and Correspondence Using Light-Field Angular Coherence. Dans *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 1940-1948).

Taylan, P., Weber, G.-W., & Yerlikaya, F. (2008). Continuous optimization applied in MARS for modern applications in finance, science and technology. Dans *ISI Proceedings of 20th Mini-EURO Conference Continuous Optimization and Knowledge-Based Technologies* (pp. 317-322).

Tenenhaus, M. (1998). *La régression PLS: théorie et pratique*. Editions technip.

Tibby, S. M., Hatherill, M., Marsh, M. J., & Murdoch, I. A. (1997). Clinicians' abilities to estimate cardiac index in ventilated children and infants. *Archives of Disease in Childhood, 77*(6), 516-518. doi: 10.1136/adc.77.6.516

Tikuisis, P., Meunier, P., & Jubenville, C. (2001). Human body surface area: measurement and prediction using three dimensional body scans. *European journal of applied physiology, 85*(3-4), 264-271.

Tneb, R., Seidl, A., Hansen, G., & Pruett, C. (2000). 3-d body scanning-systems, methods and applications for automatic interpretation of 3d surface anthropometrical data. Dans *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 44, pp. 844-847). SAGE Publications.

Tong, J., Zhou, J., Liu, L., Pan, Z., & Yan, H. (2012). Scanning 3D full human bodies using kinects. *IEEE Transactions on Visualization and Computer Graphics, 18*(4), 643-650. doi: 10.1109/TVCG.2012.56. Repéré à https://www.scopus.com/inward/record.uri?eid=2-s2.0-84863363342&doi=10.1109%2fTVCG.2012.56&partnerID=40&md5=d1600bd964fd9af6d0026c900671998e

Treleaven, P., & Wells, J. (2007). 3D body scanning and healthcare applications. *Computer, 40*(7), 28-34.

Tucker, G. R., & Alexander, J. K. (1960). Estimation of body surface area of extremely obese human subjects. *Journal of applied physiology, 15*(5), 781-784.

Uesugi, T., Okada, N., Sakai, K., Nishina, K., Mikawa, K., & Shiga, M. (2002). Accuracy of visual estimation of body height and weight in supine paediatric patients. *Pediatric Anesthesia, 12*(6), 489-494.

Van Deun, D., Verhaert, V., Buys, K., Haex, B., & Vander Sloten, J. (2011). Automatic generation of personalized human models based on body measurements. *status: published*.

Verbraecken, J., Van de Heyning, P., De Backer, W., & Van Gaal, L. (2006). Body surface area in normal-weight, overweight, and obese adults. A comparison study. *Metabolism, 55*(4), 515-524.

Vernez, D., Milon, A., Francioli, L., Bulliard, J. L., Vuilleumier, L., & Moccozet, L. (2011). A numeric model to simulate solar individual ultraviolet exposure. *Photochemistry and photobiology, 87*(3), 721-728.

Volz, A., Blum, R., Häberling, S., & Khakzar, K. (2007). Automatic, body measurements based generation of individual avatars using highly adjustable linear transformation. Dans *Digital Human Modeling* (pp. 453-459). Springer.

Wachtel, T. L., Berry, C. C., Wachtel, E. E., & Frank, H. A. (2000). The inter-rater reliability of estimating the size of burns from various burn area chart drawings. *Burns, 26*(2), 156-170.

Warfield, S. K., Zou, K. H., & Wells, W. M. (2004). Simultaneous truth and performance level estimation (STAPLE): an algorithm for the validation of image segmentation. *Medical Imaging, IEEE Transactions on, 23*(7), 903-921.

Weiss, A., Hirshberg, D., & Black, M. J. (2011). Home 3D body scans from noisy image and range data. Dans *2011 International Conference on Computer Vision* (pp. 1951-1958). doi: 10.1109/ICCV.2011.6126465

Weisstein, E. W. (2015). Heron's Formula. *MathWorld*. Repéré à http://mathworld.wolfram.com/HeronsFormula.html

Wells, J. C., Treleaven, P., & Cole, T. J. (2007). BMI compared with 3-dimensional body shape: the UK National Sizing Survey. *The American journal of clinical nutrition, 85*(2), 419-425.

Were, K., Bui, D. T., Dick, Ø. B., & Singh, B. R. (2015). A comparative assessment of support vector regression, artificial neural networks, and random forests for predicting and mapping soil organic carbon stocks across an Afromontane landscape. *Ecological Indicators, 52*(Supplement C), 394-403. doi: https://doi.org/10.1016/j.ecolind.2014.12.028. Repéré à http://www.sciencedirect.com/science/article/pii/S1470160X14006049

Williams, M. N., Grajales, C. A. G., & Kurkiewicz, D. (2013). Assumptions of multiple regression: correcting two misconceptions.

Williams, R. Y., & Wohlgemuth, S. D. (2013). Does the "Rule of Nines" Apply to Morbidly Obese Burn Victims? *Journal of Burn Care & Research, 34*(4), 447-452. doi: 10.1097/BCR.0b013e31827217bd. Repéré à http://journals.lww.com/burncareresearch/Fulltext/2013/07000/Does_the__Rule_of_Nines__Apply_to_Morbidly_Obese.9.aspx

Willmott, C. J., & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate research, 30*(1), 79-82.

Wold, S., Ruhe, A., Wold, H., & Dunn, I., WJ. (1984). The collinearity problem in linear regression. The partial least squares (PLS) approach to generalized inverses. *SIAM Journal on Scientific and Statistical Computing, 5*(3), 735-743.

Woodall, W. H., & Borror, C. M. (2008). Some relationships between gage R&R criteria. *Quality and Reliability Engineering International, 24*(1), 99-106.

World Health Organization. (1995). Physical status: The use of and interpretation of anthropometry, Report of a WHO Expert Committee.

Yang, G.-r., Yuan, S.-y., Fu, H.-j., Wan, G., Zhu, L.-x., Bu, X.-l., . . . Ji, Y. (2010). Neck circumference positively related with central obesity, overweight, and metabolic syndrome in Chinese subjects with type 2 diabetes: Beijing Community Diabetes Study 4. *Diabetes care, 33*(11), 2465-2467.

Yao, L., Dong, G., & Hu, G. (2015). Kinect-Based 3D Color Reconstruction. Dans *Proceedings of the 4th International Conference on Computer Engineering and Networks* (pp. 521-529). Springer.

Yu, C.-Y., Lo, Y.-H., & Chiou, W.-K. (2003). The 3D scanner for measuring body surface area: a simplified calculation in the Chinese adult. *Applied ergonomics, 34*(3), 273-278.

Yun, C., & Yang, J. (2007). Experimental comparison of feature subset selection methods. Dans *Data Mining Workshops, 2007. ICDM Workshops 2007. Seventh IEEE International Conference on* (pp. 367-372). IEEE.

Zhao, P., & Yu, B. (2006). On model selection consistency of Lasso. *Journal of machine learning research, 7*(Nov), 2541-2563.

Zhu, M., & Wang, L. (2010). Intelligent trading using support vector regression and multilayer perceptrons optimized with genetic algorithms. Dans *Neural Networks (IJCNN), The 2010 International Joint Conference on* (pp. 1-5). IEEE.

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 67*(2), 301-320.