

Omar GALARRAGA CASTILLO
Né le 28 avril 1989 (28 ans)
Vénézuélien
Marié sans enfant
Permis B, voiture personnelle

Centre de Réadaptation de Coubert
D96 – Route du Liverdy
77170 Coubert, France
01.64.42.19.23 / 06.47.72.14.87
omar.galarraga@ugecam-idf.cnamts.fr

Expérience Professionnelle

- Depuis 2016 : **Ingénieur du Laboratoire d'Analyse du Mouvement.** Centre de Réadaptation de Coubert. UGECAM Île-de-France. Coubert, France.
- 2016 : **Ingénieur de recherche.** Unité d'analyse du mouvement. Fondation Ellen Poidatz. St Fargeau Ponthierry. France. (Remplacement de 3 mois).
- 2013 - 2016 : **Doctorant contractuel.** Laboratoire IBISC, **Université d'Evry Val d'Essonne**, Evry, France et Unité d'Analyse du Mouvement, **Fondation Ellen Poidatz**, St. Fargeau-Ponthierry, France.
Thèse : *Simulation de l'effet de la chirurgie sur la marche par apprentissage statistique chez des enfants atteints de paralysie cérébrale.*
Mots clés : *machine learning, prédiction, analyse du mouvement et de la marche, biomécanique.*
- 2012 : **Stagiaire Ingénieur de Recherche. Morpho (SAFRAN).** Issy – Les Moulineaux, France.
Projet : *Filtrage d'empreintes digitales par clustering sur des bases de données de grande taille.*

Autres Projets :

- 2012 : *Perception de mouvement du corps humain en utilisant le suivi 3D de la pose par Kinect.* TSP.
- 2011 : *Implémentation d'une carte topologique en langage C.* UVSQ et TSP, France.
Reconnaissance de chiffres manuscrits en utilisant de réseaux de neurones, SVM et HMM. TSP.
- 2010 : *Développement d'un électrocardiographe.* USB. Caracas, Venezuela.

Formation

- 2017 : **Docteur en Sciences de l'Information et de la Communication. Université Paris-Saclay** (Univ. d'Evry Val d'Essonne). France. Spécialité : Traitement du signal et des images.
- 2012 : **M2 TRIED/MSc. Automatic Data Processing.** Mention "Bien". Double diplôme. Institut Télécom SudParis et Université de Versailles Saint-Quentin-En-Yvelines. France.
Traitement de signal et d'image, biométrie, reconnaissance de formes et machine learning.
- 2012 : **Ingénieur en Électronique.** Université Simón Bolívar. Caracas, Venezuela.
Programme de double diplôme avec l'Institut Télécom SudParis.
Circuits analogiques et numériques, contrôle automatique, systèmes de télécommunications.

Langues

Espagnol : langue maternelle. **Français** : courant. **Anglais** : courant (C1. 103/120 TOEFL IBT).

Compétences Informatiques

C/C++, Matlab, Unix/Linux, Codewarrior, LabView, Assembly, PSpice, MySQL, PHP, XML, HTML, JavaScript, LaTeX, R, Python, Vicon Nexus, Codamotion Odin

Centres d'Intérêt

Tennis (Club Le Mée Sport), tennis de table (1 année en club USB), football, baseball, guitare électrique (groupe "Les Albumès"), flûte à bec, voyages, garde-forestier (groupe USB).

Galarraga C. O.A., Vigneron V., Dorizzi B., Khouri N., Desailly E., 2017. **Predicting postoperative gait in cerebral palsy.** *Gait & Posture*, 52, pp. 45-51.

Galarraga C. O.A., Vigneron V., Dorizzi B., Khouri N., Desailly E., 2016. **Predicting postoperative knee flexion and pelvic tilt at initial contact of cerebral palsy children.** *Movement & Sport Sciences*, 93, pp. 87-92.

Conférences avec actes publiés

Galarraga C. O.A., Vigneron V., Khouri N., Dorizzi B., Desailly E., 2017. **Predictive simulation of surgery effect on cerebral palsy gait.** *Computer Methods in Biomechanics and Biomedical Engineering*, 20 (S1), pp. 85-86.

Desailly E., Galarraga C. O.A., Khouri N., 2017. **Improving multilevel surgery planning and predicting post-operative outcome in cerebral palsy.** *Computer Methods in Biomechanics and Biomedical Engineering*, 20 (S1), pp. 59-60.

Desailly E., Galarraga C. O.A., Khouri N., 2017. **Apprendre de l'expérience pour mieux planifier et prédire l'effet de la chirurgie sur la marche dans la paralysie cérébrale.** *Revue de Chirurgie Orthopédique et Traumatologique*, 103 (S7), pp. 34-35.

Galarraga C. O.A., Vigneron V., Khouri N., Dorizzi B., Desailly E., 2016. **Predicting kinematic outcome of multi-level surgery in cerebral palsy.** *Gait & Posture*, 49 (S2), p. 1. (Best paper session. ESMAC 2016).

Galarraga C. O.A., Vigneron V., Dorizzi B., Khouri N., Desailly E., 2015. **Predicting surgery effect on knee kinematics in cerebral palsy.** *Gait & Posture*, 42 (S2-S3). (Best paper session. ESMAC 2015).

Galarraga C. O.A., Vigneron V., Dorizzi B., Khouri N., Desailly E., 2015. **Predicting postoperative knee flexion during gait of cerebral palsy children.** *Computer Methods in Biomechanics and Biomedical Engineering*, 18 (S1), 1940-1.

Galarraga C. O., Vigneron V., Dorizzi B., Khouri N., Desailly E., 2015. **Estimation of Postoperative Knee Flexion at Initial Contact of Cerebral Palsy Children using Neural Networks.** In *Proceedings of the International Conference on Pattern Recognition Applications and Methods*, pp. 338-342.

Autres conférences

Galarraga C. O.A., Llorca C., Le Saux A., Hareb F., 2018. **Nouvelle méthode de calcul des scores de déviation de la marche dans l'hémiplégie.** 17^{ème} congrès de la SOFAMEA.

Galarraga C. O.A., Vigneron V., Khouri N., Dorizzi B., Desailly E., 2018. **Fusion de modèles pour la prédiction de la marche postopératoire dans la paralysie cérébrale.** 17^{ème} congrès de la SOFAMEA.

Le Saux A., Galarraga C. O.A., Hareb F., 2018. **Variabilité intersession et inter-opérateur de la cinématique de la marche.** 17^{ème} congrès de la SOFAMEA.

Galarraga C. O.A., Vigneron V., Dorizzi B., Khouri N., Desailly E., 2016. **Prédiction de la marche postopératoire chez des enfants atteints de paralysie cérébrale.** 15^{ème} congrès de la SOFAMEA.

Galarraga C. O.A., Vigneron V., Dorizzi B., Desailly E., 2014. **Approximation des angles articulaires de la marche chez des enfants paralysés cérébraux.** 13^{ème} journées de la SOFAMEA.

Simulation par apprentissage statistique de l'effet de la chirurgie sur la marche des enfants atteints de paralysie cérébrale

Omar Antonio Galarraga Castillo

Objectifs

L'objectif de la thèse est de simuler l'effet de la chirurgie orthopédique sur la marche chez des enfants atteints de paralysie cérébrale en utilisant des techniques d'apprentissage statistique automatique (*statistical machine learning*). Ces simulations permettront de donner une prédiction du résultat au niveau cinématique du traitement lorsqu'il y a une chirurgie proposée.

Travaux accomplis

Une base de données avec les données d'analyses quantifiées de la marche (AQM) [1] et d'examens cliniques [2] préopératoires et postopératoires, ainsi que les données chirurgicales, a été constituée. Cette base de données est en augmentation constante, à chaque fois qu'un nouveau patient arrive à la Fondation Ellen Poidatz pour des analyses postopératoires. À la date de la dernière mise à jour (septembre 2015), la base de données compte avec les données de 134 patients opérés avec examens préopératoires et postopératoires.

Pour simuler l'effet de la chirurgie sur la marche, des modèles de régression entre la cinématique postopératoire et les données préopératoires et chirurgicales ont été conçus afin d'établir une relation mathématique entre la marche postopératoire en fonction de l'état préopératoire du patient et la chirurgie réalisée. Les angles cinématique considérés sont : antéversion, inclinaison et rotation du bassin ; flexion, adduction et rotation des hanches ; flexion des genoux, flexion dorsale de cheville et progression des pieds. Pour décrire l'état préopératoire des patients, différentes méthodes d'extraction de caractéristiques et de réduction de dimension ont été considérées : ajustement des courbes par des B-Splines [3] et des séries de Fourier [4], sélection de variables avec classement de Gram-Schmidt avec variable de sonde [5], [6] et analyse en composantes principales [7].

Les données manquantes de certaines variables d'examen clinique ont été imputées (remplacées) en utilisant l'algorithme d'imputation itérative sous modèles [8]. Ensuite ces données imputées ont été intégrées à la description de l'état préopératoire du patient.

Plusieurs méthodes de régression ont été testées à la fois des régressions linéaires avec intervalles de confiance et des régressions non-linéaires [9] en utilisant des réseaux de neurones artificiels [10]. Des modèles de régression ont été appris pour chaque catégorie de geste chirurgical considérée : osseuse de hanche, parties molles de hanche, rectus femoris, allongement des ischio-jambiers, abaissement de rotule, ostéotomie fémorale distale, osseuse jambe, parties molles du pied et osseuse du pied. Les modèles par chirurgie ont été ensuite fusionnés en utilisant des techniques d'apprentissage par ensembles [11].

Les modèles de régression ont été testés en utilisant une méthode *jackknife* de « leave-one-out » [9], [12] afin d'avoir des prédictions en test pour tous les patients dans la base de données. Les erreurs de prédiction par angle cinématique considéré ont été comparées à l'erreur de trois prédicteurs naïfs :

- NoChange-P : signal préopératoire, prédisant aucun changement cinématique).
- Mean-P : moyenne des signaux postopératoires.
- Median-P : médiane des signaux postopératoires.

Finalement, les signaux cinématiques prédits ont été utilisés pour animer un avatar personnalisable et ainsi montrer la marche résultante probable.

Résultats

Les erreurs (RMSE) de prédictions moyennes en test du système varient entre 3,69° (inclinaison du bassin) et 10,00° (progression du pied) selon l'angle cinématique considéré [table 1]. Le système proposé obtient des erreurs plus petites que celles des prédicteurs naïfs pour tous les angles cinématiques, avec des différences statistiques significatives ($p < 0.05$) [table 1].

Pour chaque membre inférieur, le système montre les courbes prédites, ainsi que les courbes préopératoires et la contribution estimées de chaque geste chirurgical [figure 1]. Si les courbes

postopératoires sont connues, le système montre aussi ces signaux et calcule l'erreur de prédiction. Les courbes prédites des deux membres inférieurs d'un même patient sont ensuite utilisées pour animer l'avatar personnalisé du patient [figure 2].

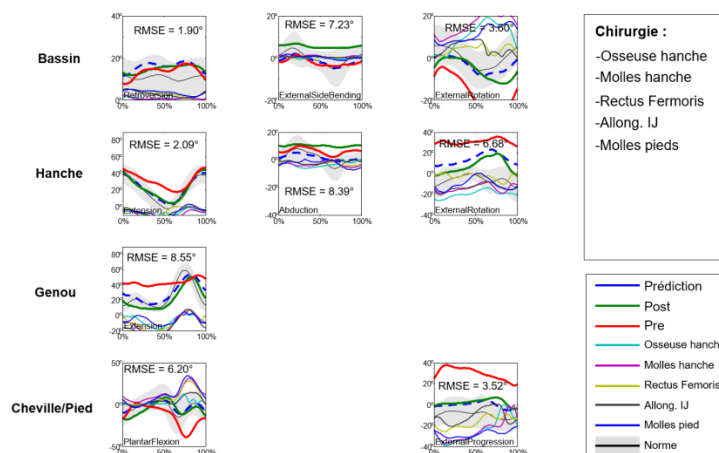


Figure 1. Sortie du système pour un membre inférieur choisi de façon aléatoire.

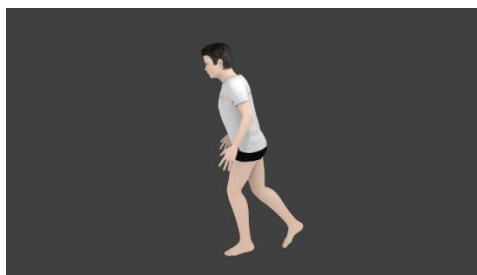


Figure 2. Capture d'écran de l'animation de la marche prédite pour un patient aléatoire.

Table 1. Erreurs de prédiction (RMSE) par angle cinématique et méthode de prédiction. Moyenne et (erreur standard) [°].

| | NoChange-P | Mean-P | Median-P | Méthode proposée |
|--------------------------|--------------|--------------|--------------|------------------|
| Antéversion bassin | 6,26 (0,27) | 6,50 (0,26) | 6,51 (0,26) | 5,55 (0,27) |
| Inclinaison bassin | 4,23 (0,15) | 3,95 (0,12) | 3,96 (0,12) | 3,69 (0,12) |
| Rotation bassin | 7,77 (0,29) | 6,91 (0,23) | 6,91 (0,23) | 6,53 (0,24) |
| Flexion hanche | 7,32 (0,31) | 7,99 (0,25) | 7,99 (0,26) | 6,76 (0,24) |
| Adduction hanche | 4,87 (0,17) | 4,23 (0,12) | 4,24 (0,13) | 4,06 (0,12) |
| Rotation hanche | 12,03 (0,50) | 10,06 (0,37) | 10,06 (0,37) | 9,50 (0,34) |
| Flexion genou | 13,63 (0,57) | 10,49 (0,34) | 10,50 (0,34) | 9,71 (0,31) |
| Flexion dorsale cheville | 13,02 (0,65) | 7,71 (0,27) | 7,71 (0,27) | 7,51 (0,27) |
| Progression du pied | 14,65 (0,66) | 10,90 (0,43) | 10,92 (0,44) | 10,00 (0,43) |

Discussion

Le système prédit la marche postopératoire des enfants atteints de paralysie cérébrale en fonction de leur marche préopératoire, leurs données cliniques et le programme chirurgical. Malgré le manque relatif d'exemples d'apprentissage, le système est capable de donner de prédictions pour un grand nombre de combinaisons chirurgicales possibles (80 combinaisons différentes dans la base de

données) et un grand nombre de patterns de marche différents. Les données cliniques sont considérées pour la prédiction, lesquelles représentent une information complémentaire à la cinématique [13]. La performance du système est meilleure que celle du meilleur prédicteur naïfs (Mean-P), donc la méthode de prédiction donne une information plus fine que les résultats chirurgicaux moyens [14].

La base de données créée est la plus grande dans son genre en France et une des plus grandes au monde.

Les angles cinématiques le plus compliqués à prédire sont la progression du pied, la flexion du genou et la rotation de hanche. En revanche, les angles les plus faciles à prédire sont l'adduction de hanche, et l'antéversion et l'inclinaison du bassin.

Intérêt

L'intérêt principal du système est la mesure anticipée du résultat fonctionnel de la chirurgie. Cela permet non seulement de prédire si ce résultat serait bon ou mauvais, mais aussi de quantifier de façon probabiliste le bénéfice ou la dégradation. Le système peut ainsi aider le chirurgien à choisir le traitement optimal pour chaque cas. Il s'agirait aussi d'une façon d'estimer la capacité d'amélioration fonctionnelle de chaque patient, une fois le traitement optimal trouvé.

Limites

La principale limite des travaux est la faible quantité d'exemples dans la base de données d'apprentissage (patients opérés disponibles dans la base de données avec analyse quantifiée de la marche avant et après la chirurgie) par rapport au nombre de paramètres que l'on veut estimer (la quantité de données disponible pour chaque patient et chirurgie, plus le nombre de variables que l'on veut prédire). Malgré avoir construit la base de données de données d'AQM pré et post-chirurgie la plus grande de France et l'une des plus grandes au monde, cela nous amène dans un contexte de données statistiquement petites (*statistical small data*) [15], dans lequel le risque de sur-apprentissage est considérable entraînant des limites de généralisation des modèles appris. Dans ce contexte, le choix de techniques de réduction de dimension est indispensable ainsi que le choix de modèles de régression suffisamment robustes. En pratique, cela se traduit par un manque de précision des prédictions sur les patients de test. Une autre limite des travaux est l'absence de contraintes mécaniques dans les modèles générés, ce qui ne permet pas d'assurer que les prédictions soient physiquement réalisables. Cependant, la notion de dynamique intrinsèque dans les données utilisées réduit de façon importante la probabilité de tomber sur des solutions physiquement impossibles.

Perspectives

Comme les modèles d'apprentissage statistique améliorent leur performance lorsqu'il y a plus de données disponibles, la suite immédiate des travaux est d'augmenter la taille de la base de données et ainsi la précision des modèles. De plus, il est envisagé d'intégrer des modèles dynamiques de synthèse de la marche pour obtenir une solution de prédiction qui soit à la fois la plus probable et mécaniquement réaliste. Lorsqu'une excellente précision de prédiction soit atteinte, il est envisagé de trouver une chirurgie optimale en considérant la combinaison chirurgicale qui donne le meilleur résultat fonctionnel, après avoir testé une grande quantité de combinaisons chirurgicales possibles.

Bénéfice pour les patients

En premier lieu, les patients et leurs familles peuvent avoir un aperçu du résultat probable du traitement. Cela aide les patients à avoir une idée plus claire du bénéfice que la chirurgie pourrait les apporter, ce qui est un point très important car la notion d'amélioration du patient est en général différente de celle du chirurgien. Cela permettrait une meilleure discussion du traitement avec le chirurgien et le médecin traitant et pourrait aussi constituer une motivation pour que le patient poursuive le traitement proposé.

En deuxième lieu, le système donne au chirurgien la possibilité de tester sa proposition chirurgicale et ainsi de la valider, la modifier (ajouter ou enlever un geste chirurgical) ou la rejeter selon l'évaluation des prédictions obtenues. Cela permettrait au chirurgien de trouver le traitement optimal pour chaque cas, surtout lorsqu'il y a plusieurs possibilités de combinaison chirurgicale.

Références

- [1] J. R. Gage et J. L. Stout, « Gait Analysis: Kinematics, Kinetics, Electromyography, Oxygen Consumption and Pedobarography », in *The Identification and Treatment of Gait Problems in Cerebral Palsy*, 2nd Edition., J. R. Gage, M. H. Schwartz, S. E. Koop, et T. F. Novacheck, Éd. London: MacKeith Press, 2009, p. 260-284.
- [2] J. P. Trost, « Clinical Assessment », in *The Identification and Treatment of Gait Problems in Cerebral Palsy*, 2nd Edition., J. R. Gage, M. H. Schwartz, S. E. Koop, et T. F. Novacheck, Éd. London: MacKeith Press, 2009, p. 181-204.
- [3] C. De Boor, *A Practical Guide to Splines*, vol. 27. Springer, 2001.
- [4] H. Motulsky et A. Christopoulos, *Fitting Models to Biological Data Using Linear and Nonlinear Regression: A Practical Guide to Curve Fitting*, 1 edition. Oxford ; New York: Oxford University Press, 2004.
- [5] I. Guyon et A. Elisseeff, « An introduction to variable and feature selection », *J. Mach. Learn. Res.*, vol. 3, p. 1157–1182, 2003.
- [6] G. Dreyfus, « L'apprentissage statistique : pourquoi, comment ? », in *Apprentissage statistique*, G. Dreyfus, Éd. Paris: Eyrolles, 2008, p. 1-71.
- [7] I. T. Jolliffe, *Principal Component Analysis*. New York: Springer-Verlag, 2002.
- [8] M. Templ, A. Kowarik, et P. Filzmoser, « EM-based stepwise regression imputation using standard and robust methods », *Res. Rep. Cs-2010-3 Dep. Stat. Probab. Theory Vienna Univ. Technol.*, 2010.
- [9] C. M. Bishop, *Pattern Recognition And Machine Learning*, 1st ed. 2006. Corr. 2nd printing 2011. New York: Springer-Verlag New York Inc., 2006.
- [10] R. O. Duda, P. E. Hart, et D. G. Stork, *Pattern Classification*, 2nd Edition. New York: Wiley-Blackwell, 2000.
- [11] R. Polikar, « Ensemble based systems in decision making », *Circuits Syst. Mag. IEEE*, vol. 6, no 3, p. 21–45, 2006.
- [12] R. G. Miller, « The jackknife-a review », *Biometrika*, vol. 61, no 1, p. 1-15, avr. 1974.
- [13] K. Desloovere, G. Molenaers, H. Feys, C. Huenaerts, B. Callewaert, et P. V. de Walle, « Do dynamic and static clinical measurements correlate with gait analysis parameters in children with cerebral palsy? », *Gait Posture*, vol. 24, no 3, p. 302-313, nov. 2006.
- [14] A. Bonnefoy-Mazure, Y. Sagawa, V. Pomero, P. Lascombes, G. De Coulon, et S. Armand, « Are clinical parameters sufficient to model gait patterns in patients with cerebral palsy using a multilinear approach? », *Comput. Methods Biomech. Biomed. Engin.*, vol. 19, no 7, p. 800-806, mai 2016.
- [15] C. M. Bishop, « Model-based machine learning », *Philos. Trans. R. Soc. Math. Phys. Eng. Sci.*, vol. 371, no 1984, déc. 2012.



Laboratoire IBISC
EA 4526
Université d'Évry, Université Paris-Saclay
*Laboratoire Informatique, BioInformatique
et Systèmes Complexes*

Vincent Vigneron

MCF-HDR Université d'évry

Tél : +33 (6) 63 568 760

Email : vincent.vigneron@ibisc.univ-evry.fr

Madame, Monsieur,

J'ai le plaisir de soutenir la candidature de Mr. Omar Galarraga au prix de thèse IFRATH 2017 pour la thèse de doctorat intitulée « Simulation of Surgery Effect on Cerebral Palsy Gait by Supervised Machine Learning », soutenue le 30 mars 2017 à l'université d'Évry et pour l'Ecole Doctorale STIC de l'université Paris-Saclay. Son doctorat, cofinancé par la région IDF (Projet PICRI 2013), en partenariat avec la Fondation Ellen Poidatz (St-Fargeau) et la Fondation Bettencourt Schueller, était centré sur la conception d'un simulateur « intelligent » simulant l'effet de la chirurgie neuro-orthopédique sur la marche afin de prédire la marche postopératoire des enfants atteints de paralysie cérébrale.

Omar Galarraga, de nationalité vénézuélienne, a obtenu son diplôme d'ingénieur à TSP en 2012 ; Il a décidé de préparer son doctorat en mathématiques appliquées à l'université d'Évry sous ma direction et celle de Bernadette Dorizzi et d'Eric Desailly.

La motivation originale de ce projet de recherche était d'améliorer la prise en charge thérapeutique des enfants en situation de handicap moteur, notamment dans le cadre d'une paralysie cérébrale (PC). Une des conséquences habituelles de cette déficience est un trouble de la marche, qui s'explique par des anomalies du contrôle et du tonus musculaire, souvent de type spastique.

La démarche d'Omar Galarraga a consisté, d'une part, à évaluer des méthodes de modélisation de la marche existantes, à les améliorer et à en développer de nouvelles basées sur l'apprentissage statistique supervisé et, d'autre part, à les mettre en œuvre dans une situation clinique réelle à la Fondation Poidatz. Quand on sait que les résultats de la chirurgie orthopédique multi-sites sont difficilement prévisibles, on comprend tout l'intérêt de ce travail.

Dès son arrivée, Omar Galarraga, a su se familiariser rapidement avec son environnement de recherche. Il a rapidement approuvé certaines procédures d'évaluation. Il s'est intéressé à la validation des simulateurs en utilisant différentes approches mathématiques.

La formation doctorale qu'il a suivie lui a assuré ainsi de solides compétences en informatique, en mathématiques, et en automatique, notamment en apprentissage statistique.

Dès son arrivée au Laboratoire, Omar a démontré un grand dynamisme, une motivation personnelle, une curiosité scientifique élevée, un sens critique et un haut niveau d'autonomie qui confirment son plein potentiel à entreprendre et à réussir dans une carrière de chercheur ou académique qui se précisera au fil du temps. Sa capacité à surmonter les différentes barrières rencontrées durant son travail de thèse démontre son sens de l'organisation; ses scrupules à respecter les échéances et un certain regard critique constituent des atouts importants ayant participé à sa réussite.

L'originalité de la démarche, le haut niveau scientifique des contributions d'Omar Galarraga ainsi que son autonomie et son aptitude à maîtriser une stratégie de recherche justifie pleinement sa candidature à ce prix.

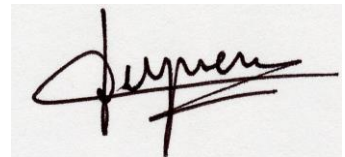
Un élément d'appréciation essentiel du bon niveau des contributions scientifiques d'Omar Galarraga est constitué par ses publications dans des revues d'audience internationale.

L'autonomie scientifique d'Omar Galarraga me semble acquise depuis plusieurs années comme on peut en juger par les collaborations qu'il a su bâtir à l'extérieur de son laboratoire et les publications qu'il rédige. Quant à l'aptitude à maîtriser une stratégie de recherche, elle apparaît de manière manifeste au travers de la progression régulière dans la difficulté des problèmes à résoudre.

En tant que directeur de thèse et directeur de l'équipe SIMOB, c'est sans réserve que j'atteste qu'Omar Galarraga fait partie des meilleurs doctorants que j'ai encadrés.

En résumé je soutiens donc avec le plus grand enthousiasme la candidature d'Omar Galarraga au prix de thèse IFRATH 2017.

Evry, le 4 février 2018

A handwritten signature in black ink, appearing to be 'J. Dupont', written on a light-colored background.



Laboratoire TIMC-IMAG

UMR 5525

Université Joseph Fourier – CNRS

*Techniques de l'Ingénierie Médicale et de la Complexité -
Informatique, Mathématiques et Applications, Grenoble*

Pierre-Yves Guméry

Professeur UJF

Tél. +33 (0)4 56 52 00 62

Fax : +33 (0)4 56 52 00 33

Email: Pierre-Yves.Gumery@univ-grenoble-alpes.fr

Grenoble le 1 mars 2017

Rapport sur le mémoire de thèse présenté par Mr Omar Antonio Galarraga Castillo pour obtenir le grade de docteur de l'université Paris-Saclay. Le titre des travaux est : 'Simulation of Surgery Effect on Cerebral Palsy Gait by Supervised Machine Learning'.

Le manuscrit s'articule autour de deux parties précédées d'une introduction. La deuxième partie est close par une conclusion et des perspectives. Le mémoire est constitué de 132 pages, annexes comprises.

Cette thèse est centrée sur la simulation par apprentissage statistique supervisé de l'effet de la chirurgie sur les signaux de marche. Le contexte clinique est celui de la paralysie cérébrale. Lorsque l'on sait que le traitement par chirurgie orthopédique multi-sites permettant de traiter les troubles engendrés par cette pathologie ont des résultats encore difficilement prévisibles, on comprend clairement l'intérêt de ce travail.

Dans la partie introductive, le chapitre n°1, Mr Galarraga Castillo décrit le consortium d'équipe participant au projet. Il précise l'objectif méthodologique et définit synthétiquement les termes clefs de la méthodologie. Il précise que les travaux précédemment menés dans le domaine sont trop limités. Ces travaux n'offrent que des estimations qualitatives ou ne portent que sur un nombre de paramètres prédits et de procédures chirurgicales insuffisants. Les objectifs spécifiques sont ensuite clairement listés. L'organisation du document est également précisée.

La première partie décrit le contexte, les données à traiter et leur conditionnement. Elle se décompose en trois chapitres.

Le chapitre n°2 est consacré à l'état de l'art. Les principes de l'analyse de la marche et les techniques chirurgicales multi-sites sont décrits. Sur le plan cinématique, le système de coordonnées de chaque segment corporel d'intérêt est explicité. Ce système est construit et réévalué à chaque instant à l'aide du suivi optique de points anatomiques. Il est précisé que ces données constituent les données principalement utilisées dans l'approche développée. Le développement de la modélisation est par ailleurs largement agrémenté de références bibliographiques. Il est précisé à ce stade, que les données cinématiques exploitées dans ce travail (angles de Cardan) sont construites à partir d'un modèle inspiré de la littérature. Le chapitre est conclu sur le constat selon lequel les méthodes de prédiction précédemment

décrites ne sont pas satisfaisantes. La valeur ajoutée de l'approche proposée est à nouveau rappelée. L'objectif et l'innovation consiste à délivrer une prédiction quantifiée de paramètres de la marche en considérant plusieurs procédures chirurgicales mais également un grand nombre de profils de marche.

Le troisième chapitre contient la description des données et de leur conditionnement. Les données sont extraites d'une base de patients ayant subi une chirurgie multi-sites. Ce chapitre est clairement rédigé. On y aborde notamment le conditionnement des données cinématiques. On pointe également le problème des données manquantes dans les examens cliniques des patients. La variabilité des approches chirurgicales est traitée et neuf classes de procédures sont proposées. Les effets des différentes procédures sont analysés sur le plan statistique. Une discussion générale est proposée de façon très opportune à la fin du chapitre. Elle permet de mettre en exergue trois questions méthodologiques centrées sur le traitement des données manquantes, la nécessité de réduire la dimension du problème et enfin le choix de la méthode de régression à retenir pour optimiser la prédiction.

Le quatrième chapitre s'intéresse plus particulièrement à l'extraction de caractéristiques et à la réduction de la dimension du problème. On s'intéresse en premier lieu à l'approximation des courbes cinématiques par des approches harmoniques (séries de Fourier tronquées) et splines périodiques. L'optimisation de l'approximation est construite par seuillage de l'erreur quadratique moyenne et de la corrélation. Ces deux approches ont le mérite de respecter la nature continue et périodique des signaux de marche. Une analyse permet de démontrer que la réduction de dimension est plus importante par approche harmonique. Une approche par sélection de variables, qu'elles soient cinématiques ou issues des examens cliniques est discutée. Cette approche se révèle peu adaptée au problème et au nombre de paramètres de sortie à considérer. Le traitement des données cliniques manquantes anticipe l'exploitation par analyse en composante principale. Une approche d'imputation de la littérature est choisie pour répondre aux spécificités de ces données. L'étape d'imputation permet de mettre en œuvre des méthodes statistiques utilisées sur des jeux de données complets. Une approche par analyse en composante principale est enfin proposée. Une discussion générale permet de conclure sur les avantages de l'analyse en composante principale. Cette dernière permet d'une part, un taux de compression plus important sur les courbes cinématiques et d'autre part et dans le même temps, une réduction des données correspondant à l'examen médical des sujets.

L'ensemble de la première partie est rédigé avec pédagogie, ceci permet de comprendre clairement les enjeux et de suivre les choix proposés.

La deuxième partie aborde la mise en concurrence de techniques de machine learning supervisées. Une introduction synthétique mais très pédagogiquement rédigée permet de présenter les principes de classification et de régression.

Le chapitre n°5 a pour but de tester et ce pour la première fois, les capacités de prédiction non-linéaires de plusieurs paramètres post-opératoires par une approche par réseaux de neurones (perceptron multicouches). Les prédictions sont limitées à deux paramètres et à l'instant du point de contact. On compare les résultats à des techniques de « pseudo-prédictions » sur une base de données de 99 patients. Le test de prédiction est mené sur l'ensemble des sujets. L'apprentissage du réseau exclue le patient sur lequel est testée la prédiction. Les résultats montrent des différences significatives entre la méthode par réseau de neurones et les trois méthodes de référence. Les prédictions se révèlent systématiquement meilleures. Il est toutefois relevé qu'un tiers des patients ont une trop mauvaise prédiction pour que la méthode puisse être employée en l'état. La solution proposée en perspective, est d'exclure les patients qui sont trop différents des sujets inclus dans la base d'apprentissage.

En règle générale il est d'usage de tester des approches linéaires avant celles qui ne le sont pas. Aussi il aurait pu paraître plus logique, mais ceci n'est qu'un point de détail et n'enlève rien à la qualité du travail, de débiter la partie II par une approche linéaire afin de renforcer la

fluidité de la lecture. Ceci d'autant plus que la méthode par réseau de neurones précédemment exposée ne vise pas à simuler l'ensemble du cycle (qui reste le but affiché).

Le chapitre n°6 est centré sur la prédiction linéaire multiple d'un cycle de marche complet. Il rejoint ainsi plus directement l'objectif assigné de ce travail. Les capacités de prédiction sont comparées à celles des techniques de « pseudo-prédiction » précédemment citées. Des analyses de performances complémentaires sont également proposées. Une explicitation de ces dernières aurait, pour le non spécialiste de l'orthopédie, complété avantageusement les références fournies dans l'article de revue qui est, seul, proposé à la lecture. Sur le plan méthodologique il est rappelé que l'analyse en composante principale permet de réduire la dimension du problème mais qu'elle entraîne une perte d'information. Le problème reste dans tous les cas complexe et on pointe ici la nécessité d'augmenter la taille de la base d'apprentissage pour optimiser les performances obtenues. Les erreurs de prédiction restent inférieures à la variabilité des mesures de paramètres. Elles restent également inférieures à celles des « pseudo-prédicteurs » de référence cités. Il n'en reste pas moins que le système est encore trop imprécis pour choisir la procédure chirurgicale la plus appropriée. La conclusion laisse en revanche place à des perspectives encourageantes.

Le chapitre n°7 expose une approche par ensemble basée sur des régressions linéaires et non-linéaires. L'objectif est d'évaluer l'effet propre de chaque procédure chirurgicale. La prise en compte de ces dernières dans la base d'apprentissage a pour effet de réduire la dimension du vecteur d'apprentissage. Une approche par ensemble est donc particulièrement adaptée à ce contexte de faible effectif. La technique proposée met en œuvre une fusion des différentes prédictions (règle arbitraire basée sur une moyenne pondérée par les significativités statistiques des différentes chirurgies sur tel ou tel paramètres). Il aurait été peut-être nécessaire ici de développer plus précisément les hypothèses sous-jacentes (indépendance des procédures chirurgicales, ...) permettant de construire cet opérateur de fusion. Une analyse en composante principale est préalablement menée. L'avantage de la méthode sur celles proposées jusque-là est qu'elle permet d'analyser l'effet d'un grand nombre de combinaisons chirurgicales y compris celles ne figurant pas nécessairement dans la base d'apprentissage. Les résultats montrent que les régressions optimales sont obtenues avec des approches linéaires ou avec des couches cachées comportant quelques unités. Ceci est cohérent avec le fait que l'apprentissage se fait sur des échantillons réduits. La méthode donne des erreurs significativement plus faibles que les prédicteurs de référence déjà utilisés dans les analyses de performances précédentes.

Le chapitre n°8 propose une comparaison entre les techniques précédemment mises en œuvre et une technique impliquant une approche par approximation trigonométrique (étudiée au chapitre n°4) couplée à une régression par réseau de neurone à une couche. Les comparaisons sont rendues délicates car les estimations ne sont pas toutes faites dans des conditions identiques (estimation ou non de l'ensemble du cycle de la marche, prise en compte ou pas du code de la procédure chirurgicale). La lecture de ce chapitre est également délicate, on aurait pu choisir de limiter l'analyse à des éléments facilement comparables concernant par exemple, les méthodes prédisant l'ensemble du cycle. Ces dernières (proposées aux chapitres n°6 et n°7) affichent globalement de meilleures performances qu'un estimateur de référence de type « pseudo-prédicteur » correspondant à une approche en moyenne des paramètres post-opératoires. On souligne ici à nouveau l'intérêt de l'approche par ensemble qui pour des performances comparables permet de mettre en évidence l'apport de chaque procédure chirurgicale. Concernant l'approche par régression linéaire multiple, il est rappelé que l'intérêt réside dans le fait que des intervalles de confiance sont disponibles.

Le chapitre 9 conclut le mémoire. Il est souligné qu'au terme de ce travail la base de données est une des plus larges utilisées dans la littérature. Malgré ceci le contexte reste celui de faibles effectifs. Les erreurs sont en général plus faibles pour les méthodes proposées que pour celles de références. Ces performances sont de plus indépendantes du type de patients. Les méthodes développées permettent aux chirurgiens de choisir entre plusieurs stratégies opératoires. La précision actuellement obtenue ne permet pas cependant de garantir une bonne prédiction pour tous les patients et combinaisons de procédures. Lorsque l'estimation

est utilisable sa valeur ajoutée réside également dans le fait qu'elle propose une estimation de l'amélioration fonctionnelle du sujet.

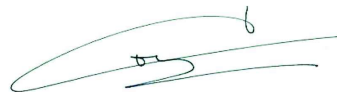
Mr Antonio Galarraga Castillo présente un mémoire de thèse de très bonne facture scientifique. Je tiens à souligner ici l'importance du travail qui s'intègre dans un environnement multidisciplinaire et qui nécessite de faire appel à des compétences de traitement statistique du signal, de classification, de robotique et de biomécanique. Les problématiques scientifiques abordées sont larges, elles comprennent notamment la modélisation en chirurgie, la statistique pour faibles échantillons, la réduction de dimension, le traitement de données manquantes, l'extraction de caractéristiques et les techniques de régression.

Les quelques remarques de forme et d'organisation du mémoire (notamment le choix d'inclure des articles dans le corps du texte) ne remettent pas en question sa qualité d'ensemble. Je tiens à souligner également la capacité de recul qui est constamment présente dans l'analyse des résultats et qui prend en compte de façon rationnelle les possibilités actuelles d'usage clinique des propositions.

Deux articles de revue (un troisième en soumission) soulignent par ailleurs la qualité du travail. Pour toutes ces raisons, Mr Omar Antonio Galarraga Castillo mérite sans réserve de présenter son travail en vue d'obtenir le diplôme de docteur de l'université Paris-Saclay.

Fait à Grenoble le 1 mars 2017.

Pierre-Yves Guméry



Opinion
about the thesis of Mr. Omar Antonio Galarraga Castillo
entitled

Simulation of Surgery Effect on Cerebral Palsy gait by Supervised Machine Learning

Cerebral Palsy (CP) indicates a movement disorder which appears in the developing infant brain. Its symptoms may turn more noticeable with the years but the underlying causes do not worsen over time. Cerebral palsy is due to a malfunction of brain areas which control movement, balance and posture. Clinical gait analysis identifies and characterizes these movement disorders which can be treated successfully by orthopedic surgery. As the latter impacts several levels of the lower limbs simultaneously, it is called *Single Event Multi-Level Surgery*. Although often highly efficient, the outcome of such a treatment is currently hard to predict. It is here where the dissertation of Omar Galarraga comes in. He developed a simulation tool to estimate the most likely outcome of such a surgical intervention. The simulations employed various machine learning techniques which were benchmarked against existing methods reported in the literature.

A short but illusive introduction motivates the study and describes the environment where the study was embedded. Obviously the project runs over a couple of years already, hence considerably experience has been accumulated by the project leaders in this field from which Omar Galarraga could profit.

Part I: Context and Data

The thesis then is divided into two parts. Part I first presents the state of the art, then moves on to present details of the data set underlying the study and finally concentrates on feature extraction and dimension reduction techniques. Concerning movement disorders there are various classification systems encompassing different views onto the disorder. Whatever, it becomes clear that the presented study excludes the most severe forms of the disorder where coordinated movements are practically impossible, and also it focuses onto the lower limbs only with the analysis. Next gait analysis needs to be considered. This is meanwhile a very active field in many disciplines, hence many data acquisition systems exist like videography, kinematics, kinetics and surface electromyography, which produce a wealth of information to which machine learning techniques for feature extraction and classification can be applied. However, because many CP - patients need walking aids, kinetic measurements were not always available and were not considered in the study reported. On the other hand, kinematics form an important source of information to characterize movement disorders. Movement descriptions rely on multi-sensor data calibrated via markers and a direct linear transform, a kinematic model which is not described in the thesis and Euler angles. Gross deviations become visible if joint angle curves of patients are compared to a normal control group. Such data are complemented by data from physical examinations and support the decision making. To do so, systematic examinations of certain variables are obligatory if sufficient statistics should be achieved. Hence, the presented study only focused on few such data from physical examinations which are performed most frequently. Given all these descriptive tools, they serve to help deciding on the ultimate action of a single event multi-level surgery (SEMLS). Thereby several goals can be followed in parallel and there is a vast number of possible combinations of surgical interventions depending on the functional and anatomical state of the patient. To improve on this combinatorial trap, gait pattern classification and surgery outcome prediction are indispensable. Classification is based on gait indexes and scores, commonly agreed upon, which measure deviations, most frequently of kinematic variables, from a normal control group. Concerning

predictive tools, most are based on mechanical or statistical models. The latter are based on linear discriminant analysis, support vector machines or random forests. However, most commonly gait variables are predicted but not surgical intervention outcomes. Considering the state of the art, a declared goal of the presented study was to quantitatively predict gait parameters considering several surgical interventions and multiple combinations of them as well as for a large number of gait patterns.

All data used in the study were collected in the same Lab. Only patients with a pre- as well as a post-operative clinical gait analysis have been considered. In addition, pre-operative a physical examination had to be performed. One set of data comprised kinematic variables. They were pre-conditioned by averaging and down-sampling to a consistent set of 765 data that could be compared to normal controls. Considering physical examinations, as not always all measurements were needed, data show many missing values and only 19 variables were analyzed in total, yielding a data set of 36 entries. Of course, a large biological variability overlays any disease-specific fluctuations in the measured values. Next, surgical interventions were grouped into 9 categories encompassing similar functional objectives and/or joint or limb. Finally, kinematic and physical examination parameters were related to surgical procedures and a Student's t-test could identify those variables which can distinguish both groups. A logic consequence of these correlations asks for a comparison of variations in kinematic variables pre- / post-operative with respect to whether a specific surgical intervention has been performed or not. A list of illusive observations could be established that way. Summarizing this data collection and preconditioning section, Omar Galarraga concludes that he had to face a large p, small n problem, indicating the necessity of a dimension reduction and feature extraction step.

The feature extraction and dimension reduction problem is exemplified on fitting functions to kinematics data. The quality of fit is measured with a mean squared error measure employing only continuous and periodic functions. Model order selection was implemented by evaluating Akaike's information criterion (AIC). Alternatively, Pearson correlation between input and output data sets was estimated and thresholded while varying the number of model parameters. Considering the approximating function to represent the measured gait cycles, a truncated Fourier series was proposed first. Referring to the model order selection problem, Omar Galarraga observes that AIC overestimates the number of parameters, corroborating an often reported observation in the context of model order selection. This is why he prefers to rely on MSE and Pearson correlation with proper thresholds for optimizing the number of parameters to achieve a minimum error approximation to the kinematic curves with the Fourier series. As a second alternative, periodic B-splines of varying order and number of segments were tested. Their approximation capability, when tested with the same kinematic gait cycles, was generally good and reasonable optimal model orders could be deduced from the threshold criteria. However, a critical comparison between both options revealed a clear preference for the Fourier series expansion. Instead of function approximation, variable selection is another alternative to approach the large p, small n problem. From the many techniques available, Omar Galarraga advocates for a combination of Gram-Schmidt orthogonalization and the probe technique which is based on a ranking of the variables. A careful analysis of this technique, however, revealed a serious inconvenience when multiple outputs are considered. As an additional drawback, both methods do not offer a possibility to include physical examination parameters. Consequently, both alternatives were rejected and principle component analysis (PCA) was considered a more promising approach. But before using this technique, the missing values problem in case of physical examination parameters needed to be solved. Only data imputation seems to be an alternative as otherwise a too strong truncation of the data sets would have resulted. As for the problems at hand, extreme values are most likely to be missing, a linear regression with k-NN initialization was chosen for data imputation. This resulted in complete data sets which, however, were too high-dimensional, thus dimension reduction was even more a must. This was done by employing a classical technique, principal component

analysis. The principal components and their related eigenvalues can be determined efficiently through a singular value decomposition of the data matrix. The normalized eigenvalues then measure the amount of information each principle direction carries with and thus provides a means to determine the minimum number of dimensions needed to preserve a certain pre-fixed amount of information. As usual, only few principal components transport a large part of the information in the data, hence a sizable reduction in dimension can be achieved with a compression ratio which outperformed all other dimension reduction techniques considered in this study.

Part II: Predicting Surgery Effect on Cerebral Palsy Gait using Supervised Machine Learning

Predicting the outcome of a surgical intervention is the declared goal of the thesis. While Part I elaborated on preprocessing techniques to condition the available data sets properly, Part II exclusively deals with regression algorithms for prediction.

First neural networks, more specifically multi-layer perceptrons (MLPs), are invoked for such a task, i.e. Finding a non-linear regression of post-operative gait from pre-operative gait and surgery. The training of the MLP was performed with a Levenberg – Marquardt algorithm with Bayesian regularization. The method has been applied to „predicting post-operative knee flexion and pelvic tilt at initial contact of cerebral palsy children“, a work that was published in the journal “Movement and Sports Sciences” in 2016.

In a second experiment, post-operative gait is predicted using multiple linear regression (MLR). As the latter involves using the Moore – Penrose pseudo-inverse, the data matrix, input to the system, has to have full rank. Hence, dimension reduction, realized via PCA, was needed here, before MLR could be applied. The whole analysis procedure was published in the journal “Gait Posture” under the title „Predicting postoperative gait in cerebral palsy“ in 2017. The careful study convincingly demonstrated that the proposed analysis methodology, though still not perfect, offers a promising tool for a general treatment outcome prediction.

A third experiment considers ensemble learning with linear and non-linear regression and applies these techniques to the prediction of post-operative gait, manifested by respective kinematics signals, in cerebral palsy. Again subspace projections via PCA had to be invoked as pre-processing tool to properly condition the pre-operative clinical gait analysis data together with data from pre-operative physical examinations. For each out of nine surgical intervention procedures either a multiple linear regression model or a non-linear regression via an MLP has been trained. The models were separately optimized for each kinematic variable and each surgical procedure considered. The single outputs were finally combined by estimating a weighted mean, where the weights specified the statistical significance of the model for each post-operative variable and thus evaluated the contribution of each surgical procedure.

Finally, all regression models have been compared concerning their performance with the conclusion that a PCA dimension reduction for conditioning the data sets and an ensemble learning obtained best results across all different surgical protocols and all considered kinematic variables.

The thesis deals with applying machine learning techniques to medical data sets beyond the classical medical imaging or time series analysis. It discusses a wealth of low-complexity machine learning algorithms which allow to robustly analyze small data sets when one has to face a large p , small n problem. The thesis follows a general protocol with a data pre-processing stage encompassing feature

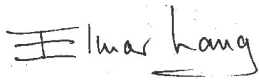
selection and dimension reduction for data conditioning followed by a data analysis stage where multi-linear and non-linear regression models, implemented through neural network models and their ensemble counterparts, were employed. Such models best avoid overfitting and guaranty robust, though not always overly precise, predictions and insights into the impact of certain variables on the former.

The thesis is very well written, technically sound and excellently illustrated. It deals with a, both scientifically as well as socially, important health problem and has far reaching consequences for the the dialogue between medical experts and their infantile patients and their parents.

I warmly recommend acceptance of this excellent thesis by the doctorate commission and evaluate it with the mark

summa cum laude

Regensburg, 22.02.2017

A handwritten signature in black ink, appearing to read 'E. W. Lang'. The signature is written in a cursive style with a horizontal line underneath the name.

(Prof. Dr. E. W. Lang)

RAPPORT DE DOCTORAT

à déposer si possible trois jours ouvrés après la soutenance et au plus tard un mois après la soutenance au service de scolarité de l'établissement de préparation de la thèse, daté et signé

Monsieur Omar GALARRAGA CASTILLO

ECOLE DOCTORALE : Sciences et Technologies de l'Information et de la Communication


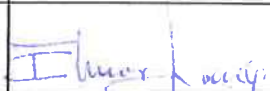
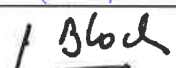



ETABLISSEMENT DE PREPARATION DU DOCTORAT: université d'Evry-Val-d'Essonne

Omar Antonio GALARRAGA CASTILLO a présenté ses travaux de thèse dans un exposé synthétique, pédagogique et d'une grande clarté, à l'image du manuscrit et de son organisation. Il a bien mis en évidence le contexte pluridisciplinaire de la thèse, l'originalité des méthodes développées et les perspectives d'application des résultats.

Les questions du jury ont été particulièrement nombreuses et ont couvert l'ensemble du travail et de ses perspectives. Le candidat y a répondu de manière argumentée. Ses réponses ont montré son recul, sa maturité, son implication dans un sujet pluridisciplinaire et complexe, en comprenant toutes les facettes. De plus, il a su instaurer une véritable discussion scientifique avec le jury.

Le jury souligne l'aspect pionnier de la recherche conduite dans la thèse.

Pour toutes ces raisons, le jury lui décerne à l'unanimité le doctorat de l'Université Paris-Saclay, préparé à l'université d'Evry

| Civilité, Nom et Prénom | Titre | Fonction dans le Jury | Signature |
|-------------------------|----------------------------|--|---|
| M. Pierre-Yves GUMÉRY | Professeur des universités | Rapporteur <input type="checkbox"/> Président |  |
| M. Elmar LANG | Professeur | Rapporteur <input type="checkbox"/> Président |  |
| Mme Isabelle BLOCH | Professeur des universités | Examineur <input checked="" type="checkbox"/> Président |  |
| M. Stéphane ARMAND | Maître de conférences | Examineur <input type="checkbox"/> Président |  |
| M. Vincent VIGNERON | Maître de conférences | Directeur de thèse |  |
| M. Eric DESAILLY | Directeur de recherche | CoDirecteur de thèse |  |



Full length article

Predicting postoperative gait in cerebral palsy



Omar A. Galarraga C.^{a,b}, Vincent Vigneron^b, Bernadette Dorizzi^c, Néjib Khouri^{a,d},
Eric Desailly^{a,*}

^a UNAM, Pôle Recherche & Innovation, Fondation Ellen Poidatz, 1 Rue Ellen Poidatz, Saint-Fargeau-Ponthierry, France

^b IBISC—EA 4526, Université d'Evry Val d'Essonne, 40 Rue du Pelvoux, Courcouronnes, France

^c SAMOVAR—UMR 5157, Télécom SudParis, Institut Mines-Télécom, 9 Rue Charles Fourier, Evry, France

^d Chirurgie Orthopédique Pédiatrique, Hôpital Universitaire Necker-Enfants Malades, 149 Rue de Sèvres, Paris, France

ARTICLE INFO

Article history:

Received 13 January 2016

Received in revised form 6 September 2016

Accepted 6 November 2016

Keywords:

Cerebral palsy

Single-event multilevel surgery

Outcome prediction

Clinical gait analysis

Machine learning

ABSTRACT

In this work, postoperative lower limb kinematics are predicted with respect to preoperative kinematics, physical examination and surgery data. Data of 115 children with cerebral palsy that have undergone single-event multilevel surgery were considered. Preoperative data dimension was reduced utilizing principal component analysis. Then, multiple linear regressions with 80% confidence intervals were performed between postoperative kinematics and bilateral preoperative kinematics, 36 physical examination variables and combinations of 9 different surgical procedures. The mean prediction errors on test vary from 4° (pelvic obliquity and hip adduction) to 10° (hip rotation and foot progression), depending on the kinematic angle. The unilateral mean sizes of the confidence intervals vary from 5° to 15°. Frontal plane angles are predicted with the lowest errors, however the same performance is achieved when considering the postoperative average signals. Sagittal plane angles are better predicted than transverse plane angles, with statistical differences with respect to the average postoperative kinematics for both plane's angles except for ankle dorsiflexion. The mean prediction errors are smaller than the variability of gait parameters in cerebral palsy. The performance of the system is independent of the preoperative state severity of the patient. Even if the system is not yet accurate enough to define a surgery plan, it shows an unbiased estimation of the most likely outcome, which can be useful for both the clinician and the patient. More patients' data are necessary for improving the precision of the model in order to predict the kinematic outcome of a large number of possible surgeries and gait patterns.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Orthopaedic surgery is usually performed in order to lessen gait abnormalities observed in patients with cerebral palsy (CP). Multiple bones and muscles are operated during a Single Event Multilevel Surgery (SEMLS) [1], which combines several procedures in the same surgery.

Clinical Gait Analysis (CGA) is used in combination with physical examination in order to propose a suitable surgery to patients with CP [1]. However, surgical decision making is not yet fully standardized. Different surgical procedures may be proposed to address the same gait deviation and different decision making algorithms may be used by medical teams to define surgical plans. Moreover, once the indication is established it is difficult for the

surgeon and furthermore for the patient to predict the effect of the surgery. Recently, several decision-making tools based on statistical machine learning have been developed for predicting surgery outcome in SEMLS. Reinbolt et al. [2] used linear discriminant analysis for predicting good or bad outcomes of rectus femoris transfer for patients with stiff knee. For predicting good or bad outcomes of hamstring lengthening, Arnold et al. [3] utilized hierarchical log-linear analysis and Sebsadji et al. [4] used support vector machines both combined with musculoskeletal models. Schwartz et al. [5] used random forests for predicting good or bad outcomes of psoas lengthening. All of the above methods give qualitative outcome predictions of improvement or non-improvement, but they do not help the surgeon nor the patient to predict how the latter will walk after surgery.

Some other methods predict some quantitative gait parameters. Hicks et al. [6] used multiple linear regression for predicting post-treatment knee flexion during stance for patients presenting crouch gait and also established good or bad outcomes based on these predictions. Sullivan et al. [7] used regression analysis and

* Corresponding author at: 1 Rue Ellen Poidatz, 77310 Saint Fargeau-Ponthierry, France.

E-mail address: eric.desailly@fondationpoidatz.com (E. Desailly).

Hersh et al. [8] used artificial neural networks to predict knee flexion during gait after rectus femoris transfer. Galarraga et al. [9] utilized artificial neural networks for predicting postoperative knee flexion and pelvic tilt at initial contact with or without hamstring lengthening. All previous works already mentioned are based on one type of abnormal gait patterns in CP (i.e. stiff knee or crouch gait) or on one principal surgical procedure, without considering the effect of other surgical procedures and their combinations. Niiler et al. [10] considered rectus femoris transfer and concurrent surgeries (hamstring lengthening, Achilles lengthening and gastrocnemius lengthening) of 68-patient series (94 lower limbs) and performed linear regressions for predicting postoperative knee range of motion during gait.

Despite these previous works, surgery planning remains difficult and global gait outcome prediction is still incomplete. Moreover, it is difficult to explain postoperative expected outcome to patients and their families, who might struggle to imagine a realistic outcome based on the predicted parameters.

The objective of this study was to use statistical machine learning techniques to develop a system able to predict postoperative kinematic curves of children with CP based on preoperative physical examination and 3-D gait analysis, and a proposed surgery plan.

2. Materials and methods

2.1. Population and data description

This retrospective study analyzed anonymous data of children with CP that have undergone SEMLS within a ten year period between 2004 and 2014. These children have had physical examination and CGA before and at least one year after surgery. Gait analysis was performed pre and postoperatively in the same laboratory. From 2004 to 2007, the acquisition was performed with a SAGA 3RT Biogesta system and, since 2008, with a Vicon system. Lower limb marker placements were identical in all the exams and kinematic data were computed from the acquisition's raw data (marker coordinates) with the same custom software based on a modified Helen Hayes [11,12] model with anatomical markers on the femoral condyles and the medial-malleolus. Fifteen kinematic angles were considered for each patient: pelvic tilt, pelvic

obliquity, pelvic rotation and hip flexion, hip adduction, hip rotation, knee flexion, ankle dorsiflexion and foot progression for both lower limbs. Surgical data were decomposed into combinations of $N_s = 9$ surgery categories: hip bony surgery, hip soft tissue surgery, rectus femoris surgery (transfer or release), hamstring lengthening, patella lowering, distal femoral osteotomy, shank bony surgery, ankle-foot soft tissue surgery and foot bony surgery. The surgical categories have been established depending on their functional objective and joint or segment that is modified. In these categories, some different surgical procedures are grouped in the same class if their functional objective and the affected joints or segments are alike [see Supplementary data for examples]. For each lower limb j , a surgery binary code $S_j = (s_{j,1} \dots s_{j,N_s})^T$ was attributed where $s_{j,i} = \begin{cases} 1 & \text{if gesture } i \text{ was conducted on patient } j \\ 0 & \text{if gesture } i \text{ was not conducted on patient } j \end{cases}$ with $i = 1, \dots, N_s$ and T is the transpose operator.

2.2. Preprocessing

The variables that have been measured during physical examination varied depending on the patient and the clinician that performed the exam. For this reason, we considered 36 parameters that were measured at a minimum rate of 80% in our database. These parameters include information about size and weight; hip, knee and ankle ranges of motion; muscle force; and spasticity (details in Supplementary data).

Fig. 1 shows all the stages of the method. Physical examination missing data were replaced using iterative robust model-based imputation (IRMI) [13]. This technique consists on initializing missing values and then iteratively perform linear regressions of each column with respect to the others. The initialization begins by searching the lower limbs with the nearest physical examination profile considering only the non-missing data with a k -Nearest Neighbor algorithm [14] for $k=5$ and ends by replacing each missing value by the median over the 5 nearest neighbors.

Kinematic data were automatically segmented into gait cycles utilizing the high pass algorithm (HPA) [15]. Then gait cycles were resampled and normalized to 51 points (2% of gait cycle) as in [16] and mean gait cycles were computed for each limb. A right and a left kinematic preoperative gait vectors were composed with the fifteen kinematic signals of both limbs normalized respectively by

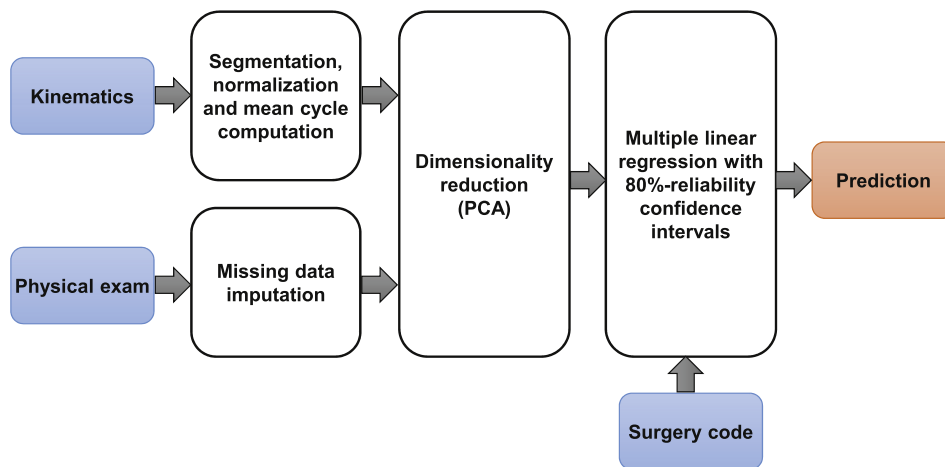


Fig. 1. Method stages from CGA and physical examination data (see Supplementary data) to prediction. Kinematic signals were segmented into gait cycles and normalized to 51 points per angle and cycle. Missing data from physical examination were imputed with the IRMI algorithm. The dimension of the concatenated vectors of preprocessed kinematics and physical examination data (see Supplementary data) was reduced using PCA. Then a multiple linear regression between postoperative kinematics and the low-dimensional preoperative vectors and surgery codes was performed. Confidence intervals with 80% reliability were computed.

the right and the left gait cycle. Postoperative gait vectors were solely described by the kinematic data of the considered limb like in [16]. Preoperative kinematic and imputed physical examination data (see Supplementary data) were gathered and projected into a lower-dimensional space using principal component analysis (PCA) [17]. Different PCA projection dimensions from 1 to 63 were tested for regression, keeping between 25% and 99% of the data inertia.

2.3. Regression model

After preprocessing, a multiple linear regression [18] was performed between the postoperative kinematic data of all lower limbs and the low-dimensional preoperative data plus the surgery code.

The multiple linear regression is represented by the equation $Y = AX + \epsilon$, where

$$Y = \begin{pmatrix} y_{1,1} & \cdots & y_{N_{LIMB},1} \\ \vdots & \ddots & \vdots \\ y_{1,N_{out}} & \cdots & y_{N_{LIMB},N_{out}} \end{pmatrix} \text{ represents the postoperative kinematic data of all lower limbs,}$$

$N_{out} = 51 \times 9 = 459$ is the total number of output points (nine

51-point kinematic angles for each limb), $X =$

$$\begin{pmatrix} x_{1,1} & \cdots & x_{N_{LIMB},1} \\ \vdots & \ddots & \vdots \\ x_{1,DIM} & \cdots & x_{N_{LIMB},DIM} \\ s_{1,1} & \cdots & s_{N_{LIMB},1} \\ \vdots & \ddots & \vdots \\ s_{1,N_s} & \cdots & s_{N_{LIMB},N_s} \\ 1 & \cdots & 1 \end{pmatrix} \text{ includes low-dimensional preoperative data } x_j \text{ and surgery codes } s_j, DIM = 1, \dots, 63 \text{ is the dimension}$$

of the PCA projection, $\epsilon = \begin{pmatrix} e_1 \\ \vdots \\ e_{N_{LIMB}} \end{pmatrix}^T$ is the training error vector

and $A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,DIM+N_s+1} \\ \vdots & \ddots & \vdots \\ a_{N_{out},1} & \cdots & a_{N_{out},DIM+N_s+1} \end{pmatrix}$ is the regression matrix to

be estimated and whose parameters were learned using the least-squares method [18] as $A = Y(X^T X)^{-1} X^T$ with $DIM \leq n$.

Parametric confidence intervals for prediction are constructed for each limb and for each point of the gait curves using the expression

$$\hat{y}_0 \pm t_{n-k-1, \alpha/2} \sqrt{1_{N_{out} \times 1} + \hat{\sigma}^2 x_0^T (X \cdot X^T)^{-1} x_0}, \quad \text{where}$$

$$\hat{\sigma}^2 = \begin{pmatrix} \hat{\sigma}_1^2 \\ \vdots \\ \hat{\sigma}_{N_{out}}^2 \end{pmatrix}, \quad \hat{\sigma}_i^2 = \frac{SSE_i}{N_{LIMB} - k - 1}, \quad SSE_i \text{ is the sum of quadratic errors}$$

for point i , $k = DIM + N_s$ is the number of input parameters, $x_0 =$

$$\begin{pmatrix} x_{0,1} \\ \vdots \\ x_{0,k} \\ 1 \end{pmatrix} \text{ is a test input vector, } \hat{y}_0 = \begin{pmatrix} \hat{y}_{0,1} \\ \vdots \\ \hat{y}_{0,N_{out}} \end{pmatrix} \text{ is the prediction of } y_0$$

and $t_{n-k-1, \alpha/2} = 1.28$ for $\alpha = 80\%$ confidence level with respect to a

unilateral Student distribution with $n - k - 1$ degrees of freedom [20].

2.4. Performance assessment

For every test limb and gait angle, the prediction performance was evaluated by the root-mean-square error:

$RMSE = \sqrt{\frac{\sum_{i=1}^{N_{points}} (\hat{y}_i - y_i)^2}{N_{points}}}$, where \hat{y}_i and y_i are respectively the prediction and the expected resampled postoperative gait signal for any i among the $N_{points} = 51$ points of a normalized kinematic angle. In addition, the overall RMSE per limb was computed with $N_{points} = N_{out} = 51 \times 9 = 459$, which corresponds to the concatenation of the nine predicted kinematic curves. The overall RMSE per patient (considering both limbs predictions at the same time) was also considered.

The method was tested using a leave-one-out jackknife procedure [14,19]: when a patient k is tested, data of both limbs of patient k were removed from the training set. Then a multiple linear regression was done with data of the other remaining patients and subsequently the data of patient k (both limbs separately) were tested on the trained model. This process was repeated for all the patients in the database.

These prediction errors were compared to errors of two pseudo-predictors: “Mean-P” and “No Change-P”. “Mean-P” is a constant pseudo-predictor equal to the mean postoperative gait cycle over all patients in the database. “No Change-P” is a pseudo-predictor that indicates the preoperative gait cycle for each patient as the postoperative prediction, which means that the surgery would have no effect on these parameters.

To measure the reliability of the parametric confidence intervals, the number of postoperative mean cycle points of test patients that were outside of the prediction band were counted for each gait angle.

The effect of the gait deviations for a single kinematic angle on the prediction performance was studied by comparing the RMSE to the preoperative Gait Variable Score (GVS) [21]. The reference database used for computing GVS is composed of 14 subjects presenting no pathology with an average age of 16 y/o (SD = 8 y/o). These reference data were acquired and processed under the same conditions of the patient's database.

Finally, the prediction performance was compared to the minimal detectable change (MDC) [22] of gait patterns in CP.

3. Results

Data of $N_{PAT} = 115$ children with CP ($N_{LIMB} = 230$ lower limbs) were included. 34 (29.6%) patients have a Gross Motor Function Classification System (GMFCS) level I, 57 (49.6%) a level II and 24 (20.9%) a level III. Male subjects represent 61% of the population and 39% are female. The mean ages of patients are 11.8 y/o (SD = 3.3) and 14.8 y/o (SD = 3.3) for preoperative and postoperative CGA respectively. Mean age at surgery is 12.6 (SD = 3.2) y/o, with postoperative CGA around 18 month after surgery. The percentages of patients that underwent each surgery category are: 33.9% hip bony surgery, 23.5% hip soft tissue surgery, 49.1% rectus femoris surgery, 50.9% hamstring lengthening, 19.6% patella lowering, 10% distal femoral osteotomy, 4.8% shank bony surgery, 50.0% ankle-foot soft tissue surgery and 23.48% foot bony surgery. There are 75 different combinations of surgical procedures and the most frequent combination has been performed over 13 lower limbs in the database (hip bony surgery with rectus femoris surgery, hamstring lengthening and ankle-foot soft tissue surgery). The average gait deviation index (GDI) [16] variation (Postoperative GDI – Preoperative GDI) is 8.5 (SD = 9.3, MAX = 37.6, MIN = –16.2).

Best results were obtained with PCA dimension of $DIM = 9$, which contains 82% of the total variance. After dimensionality reduction and reprojection into the original space, 57.6% of the original kinematic values that were outside a 2-standard-deviation (2-SD) band are preserved. Original values outside this 2-SD band represent 5% of the kinematic data. For physical examination data (see Supplementary data), 2.1% of the outlier values were preserved after reprojection. Originally 4.6% of the physical examination values were outside the 2-SD band. The kinematic angles that lost the most part of their original variance after PCA are hip adduction (bilaterally) with 61% of variance loss, knee flexion from 75% to 85% of the gait cycle with 64% of variance loss and pelvic obliquity with 51% of variance loss. The rest of the kinematic variables preserved more than 50% of their original variance, with an average loss of 18%. All the following results correspond to regressions with the optimal PCA dimension, thus to input vectors of size 18 (PCA projection of size $DIM = 9$ plus surgery code of size $N_s = 9$).

The simulator outputs the predicted kinematic curves for the test limb with their respective 80%-reliability confidence band (Fig. 2).

Mean prediction errors over all patients of the proposed method vary from 3.7° (pelvic obliquity) to 9.9° (foot progression), depending on the kinematic angle (Table 1). Best pseudo-predictors mean performances vary from 3.9° (pelvic obliquity) and 11.0° (foot progression). There are significant differences between the proposed method RMS errors and the pseudo-predictors for all the angles, except for hip adduction, hip rotation and ankle dorsiflexion. For all the kinematic angles, the mean prediction error of the proposed method is smaller or equal to the prediction error of pseudo-predictors. Overall (considering all the kinematic angles at the same time), 65.6% of the limbs are better predicted with the proposed method. Moreover, if both lower limbs of each patient are considered, the proposed method gives better predictions for 71.3% of the patients.

Unilateral mean sizes over all patients of the 80%-reliability confidence intervals vary from 5.1° (pelvic obliquity) and 14.6° (foot progression) [Table 1]. The standard deviations of the intervals' sizes are between 0.1° and 0.3° . The percentages of test points inside the parametric confidence intervals vary from 76.2% (pelvic tilt and knee flexion) to 80.6% (foot progression) [Table 1].

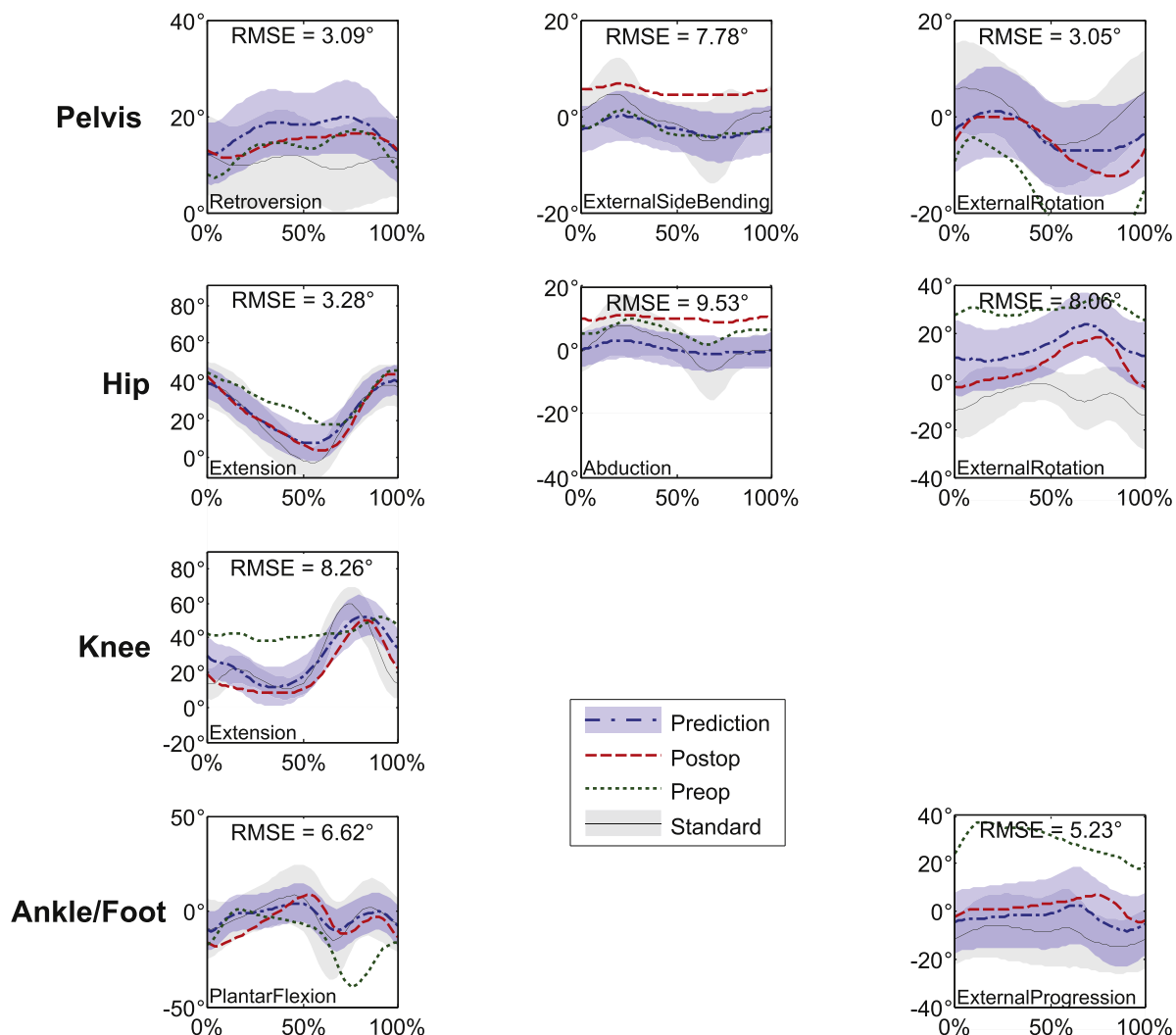


Fig. 2. System output example. Dash-dotted blue lines correspond to the estimation and blue bands represent their confidence intervals. Dashed red lines correspond to measured (real) mean postoperative gait cycle. Dotted green lines represent mean preoperative gait cycle. Solid black lines and grey bands correspond to the standards (non-pathological gait) and their two-standard-deviation bands. These standards come from the reference database described in Section 2.4. In this example the surgery for this limb consists of bony and muscle hip surgery, rectus femoris surgery (RFT), hamstring lengthening (HL) and muscle ankle/foot surgery. The contralateral limb surgery consists on rectus femoris surgery (RFT) and hamstring lengthening (HL). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1

Comparison of the performances of the proposed method and the pseudo-predictors (Mean-P, No change-P) per kinematic angle. Unilateral sizes and percentages of test points within the prediction bands of the proposed method.

| | | Pelvic Tilt | Pelvic obl. | Pelvic rot. | Hip flex. | Hip add. | Hip rot. | Knee flex. | Ankle dors. | Foot prog. |
|----------------------|---------------------------|-------------|-------------|-------------|-------------|------------|-------------|--------------|-------------|--------------|
| RMSE (°) Mean (SD) | Predictor: | | | | | | | | | |
| | No Change-P | 8.9 (5.6)* | 5.7 (3.1)* | 10.3 (5.7)* | 11.0 (5.8)* | 6.2 (3.3)* | 15.8 (9.5)* | 17.2 (10.2)* | 14.9 (9.7)* | 19.2 (12.3)* |
| | Mean-P | 7.1 (3.5)* | 3.9 (1.9)* | 7.0 (3.9)* | 7.6 (4.0)* | 4.2 (2.0) | 10.2 (6.2) | 10.6 (4.3)* | 7.5 (4.0) | 11.0 (7.3)* |
| | Proposed | 5.1 (3.2) | 3.7 (1.9) | 6.6 (3.8) | 6.8 (3.7) | 4.1 (1.8) | 9.7 (5.6) | 9.0 (3.9) | 7.5 (4.2) | 9.9 (7.3) |
| Confidence intervals | Inside points on test (%) | 76.2 | 79.3 | 79.9 | 79.0 | 76.9 | 78.3 | 76.2 | 78.0 | 80.6 |
| | Mean size (°) | 7.0 | 5.1 | 9.3 | 9.0 | 5.5 | 13.5 | 11.6 | 10.1 | 14.6 |
| | SD of size (°) | 0.1 | 0.1 | 0.2 | 0.2 | 0.1 | 0.3 | 0.2 | 0.2 | 0.3 |

* Significant difference with respect to errors of the proposed method ($p < 0.05$).

The prediction error is uniformly distributed with respect to the preoperative GVS [21], with a maximal Pearson correlation coefficient of 0.28 (hip adduction) [Fig. 3].

All the mean prediction errors are smaller than their associated MDC and at least 58.3% up to 100% of the limbs are predicted with an error smaller than the MDC depending on the predicted parameter [Table 2].

4. Discussion

For the first time postoperative lower limb kinematics are predicted with respect to preoperative kinematics, preoperative physical examination and a large number of combinations of surgical procedures in a SEMLS context for patients with CP that walk with different gait patterns. Other outcome-predictive

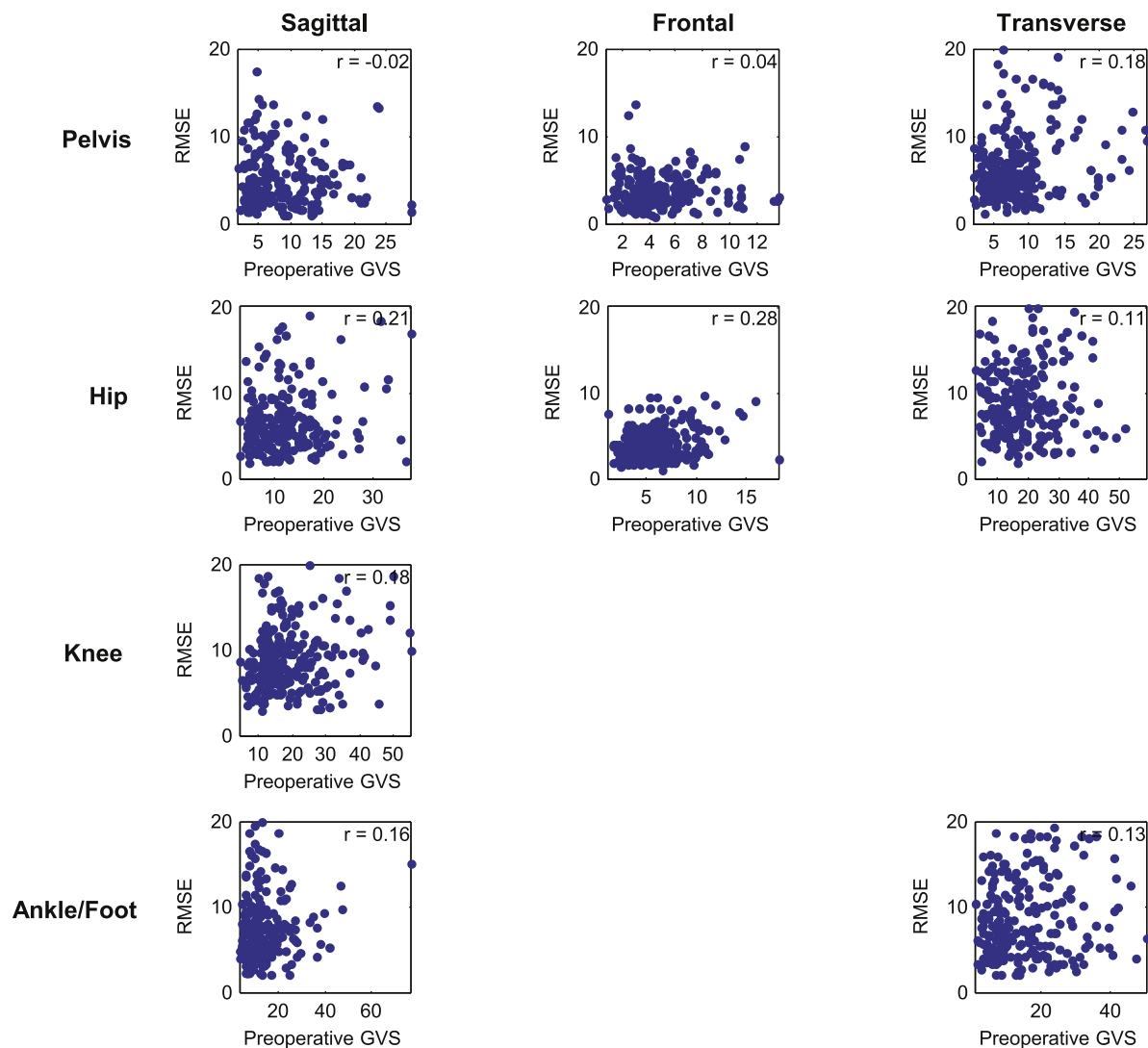


Fig. 3. Prediction Error (RMSE) with respect to preoperative GVS for each kinematic angle.

Table 2

Comparison of the proposed method performance to variability of the signals: RMSE vs. MDC.

| | | Prediction RMSE | | |
|-------------------------|--------------|-----------------|--------|-----------------------------|
| Parameter | MDC [22] (°) | Mean (°) | SD (°) | % of limbs where RMSE < MDC |
| Sagittal | | | | |
| Mean tilt | 9.3 | 4.3 | 3.5 | 89.6 |
| Pelvis ROM | 3.9 | 3.2 | 2.7 | 67.8 |
| Min hip flexion | 9.6 | 6.3 | 5.5 | 80.0 |
| Hip flexion IC | 12.0 | 5.1 | 4.1 | 93.9 |
| Hip ROM | 12.4 | 6.4 | 5.1 | 87.8 |
| Max knee flexion | 11.4 | 7.2 | 5.4 | 78.7 |
| Min knee flexion | 8.9 | 7.4 | 6.1 | 65.2 |
| Knee flexion IC | 8.1 | 7.7 | 5.7 | 58.3 |
| Knee ROM | 12.1 | 9.6 | 7.1 | 70.4 |
| Max ankle dorsi | 10.8 | 6.5 | 5.2 | 83.0 |
| Min ankle dorsi | 10.3 | 6.9 | 5.7 | 77.0 |
| Ankle dorsi IC | 10.2 | 6.6 | 5.8 | 80.0 |
| Ankle ROM | 8.5 | 5.8 | 5.2 | 75.7 |
| Frontal | | | | |
| Min up obliquity | 6.6 | 3.3 | 2.4 | 91.3 |
| Max up obliquity | 7.7 | 3.3 | 2.4 | 94.8 |
| Max hip adduction | 13.1 | 3.5 | 2.6 | 100.0 |
| Min hip adduction IC | 12.5 | 3.7 | 2.6 | 99.6 |
| Min hip adduction swing | 15.0 | 3.6 | 2.7 | 100.0 |
| Transverse | | | | |
| Max int rotation | 14.4 | 5.7 | 4.7 | 94.4 |
| Min int rotation | 10.5 | 5.5 | 4.9 | 85.2 |
| Max hip int rotation | 16.1 | 9.0 | 6.9 | 85.2 |
| Min hip int rotation | 12.6 | 9.9 | 7.6 | 70.0 |

methods are usually based on one principal surgical gesture [2,4,5,7,8,10] or give only qualitative predictions [2,4–6]. The method integrates physical examination and CGA data, which represent the usual information utilized for treatment decision making [1] and are considered complementary information because they are uncorrelated [23].

If we consider the original problem with an input vector of 810 elements (765 kinematic elements, plus 36 physical examination variables and the 9-element surgery code), before PCA, and with 228 lower limbs in the training set (114 patients, excluding the test patient), there are infinite solutions of the least-squares optimization for the regression. PCA allows to reduce the size of the input vector in order to find a unique least-squares solution. PCA dimensionality reduction maximizes the variance and eliminates data redundancy [17], however it implies some information loss. The principal variables that were affected by this information loss were physical examination data (see Supplementary data) as well as some kinematic data such as hip adduction (bilaterally), knee flexion from 75% to 85% of the gait cycle and pelvic obliquity. Even with this dimensionality reduction stage, given the heterogeneity of gait patterns in CP and the number of different possible surgical combinations with respect to the number of patients in the database, the regression problem remains very complex. This represents a limitation for the precision of the system. If there were more patients in the training database, the dimensionality reduction stage might be less important or even omitted. In any case, with or without dimensionality reduction, prediction precision would be higher if more patients' data were available. For this reason, it would be interesting to consider data from other gait analysis laboratories to considerably increase the database size, and we would like to encourage data sharing between these laboratories, which is very limited at the present. This would also facilitate the use of other regression methods, especially nonlinear regression, which may also improve the performance.

The prediction errors are generally smaller than the variability of the CP gait parameters measured by the MDC [22]. The MDC is a

measure of intersession variability of these parameters and is an estimation of the minimal amount of change that is needed to exceed measurement error.

In addition, the prediction errors are in average smaller than the naive predictors' errors, especially Mean-P. This means that the proposed model gives richer information than just knowing the average SEMLS outcome [24], which is a tendency to slightly improve gait shown by the GDI [16] variation of the considered series. Nevertheless, the system is not yet accurate enough to find the most suitable surgery or treatment.

On the other hand, the prediction errors and the preoperative Gait Variable Scores (GVS) [21] are uncorrelated (Fig. 3). This means that the system is able to perform in the same way independently of the severity of the preoperative state of patients. However, the usage of the system should be limited to patients and surgeries similar to those composing the training database.

The system does consider contralateral preoperative state (kinematics and physical examination) but does not consider contralateral surgery, which has an effect on the postoperative kinematics. Another limitation of the system is the grouping of similar but not identical surgical procedures in one surgical category. Both the grouping and the omission of the contralateral surgery were necessary to reduce the variability of the treatment, but it also introduces a source of error that cannot be controlled by the system.

Even if further validation and especially external validation should be considered before clinical use, the adjunction of thresholds for preoperative parameters (kinematics and physical examination) could be established (i.e. within ± 2 SD) in order to filter patients having gait parameters that are far from those in the training base. Also surgery plans might be unpredicted if their combinations are too different from those in the training database.

Frontal plane angles (pelvic obliquity and hip adduction) are the angles predicted with the smallest mean errors over all the patients. However, this does not mean that this is always the case for all the patients (see Fig. 2). In addition, the global performances

for these angles are the same as predicting the postoperative mean, with no significance statistical difference for hip adduction.

Sagittal plane angles (pelvic tilt, hip flexion, knee flexion and ankle dorsiflexion) are better predicted than transverse plane angles, but have bigger prediction errors than frontal plane angles. Performances for sagittal plane angles are significantly more accurate than predicting the postoperative mean (Mean-P) except for ankle dorsiflexion.

Highest prediction errors are obtained for transverse plane angles (pelvis rotation, hip rotation and foot progression), which are also the angles with highest variability in children with CP (Table 2) [22]. In addition, confidence intervals for angles in transverse plane are the largest in general.

Computed parametric confidence intervals give prediction intervals with 80% reliability. This reliability has been confirmed for test data (76–81% of tested points per angle). The sizes of these confidence intervals are specific to the patient, surgery, angle and instant of gait cycle. These intervals show an estimation of the prediction uncertainty. For example, if a confidence interval were considered too large, its associated prediction would be considered as inaccurate. Prediction bands give a set of probable solutions that might help both clinicians and patients to better understand and discuss the most likely surgery outcome.

After evaluation of the most likely outcome, a qualitative prediction could also be presented in order to decide if the outcome would be good or bad, as most of the previous works do [2,4–6]. Decision thresholds over the gait variables variation could be established for this purpose, and each medical team could adjust those threshold according to theirs needs.

The proposed system could be used as a decision-making tool for SEMLS that shows the most likely surgery outcome in terms of gait to both clinicians and patients with a confidence level of at least 76%. The system represents an easy way for visualizing the most likely surgery outcome, which allows to facilitate clinician-patient treatment discussion. The output of the system could serve as a motivation for patients in order to proceed with a surgical treatment.

In conclusion, the proposed method is able to give a preview of the most likely surgery outcome for different surgical combinations and gait patterns, given a relatively small data set. The effect of each surgical procedure and preoperative parameters is not well estimated for frontal plane angles and ankle dorsiflexion, because these predictions tend too closely to the postoperative mean kinematics. Nevertheless, the postoperative mean is a good predictor for frontal plane angles, presenting tenuous mean prediction errors and standard deviations. On the other hand, the system is more sensitive to the input variables for both sagittal and transverse plane angles (except for ankle dorsiflexion), but is more precise for sagittal plane angles, especially for pelvic tilt and hip flexion. Finally, although there are still some developments and validations to be done, this work represents an encouraging progress towards general treatment outcome prediction, in order to help clinicians to choose optimal treatment and to help patients to better understand it.

Conflict of interest

None.

Acknowledgements

This study is part of the SIM-PC2 project promoted by the Fondation Ellen Poidatz and supported by Fondation Ellen Poidatz, Fondation Bettencourt Schueller and Region Île de France.

The authors would like to thank UNAM medical and technical team of Fondation Ellen Poidatz for useful discussions and for recording all the data used in this work.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.gaitpost.2016.11.012>.

References

- [1] J.R. Gage, M.H. Schwartz, S.E. Koop, T.F. Novacheck, *The Identification and Treatment of Gait Problems in Cerebral Palsy*, 2nd edition, MacKeith Press, London, 2009.
- [2] J.A. Reinbolt, M.D. Fox, M.H. Schwartz, S.L. Delp, Predicting outcomes of rectus femoris transfer surgery, *Gait Posture* 30 (July (1)) (2009) 100–105.
- [3] A.S. Arnold, M.Q. Liu, M.H. Schwartz, S. Öunpuu, S.L. Delp, The role of estimating muscle-tendon lengths and velocities of the hamstrings in the evaluation and treatment of crouch gait, *Gait Posture* 23 (April (3)) (2006) 273–281.
- [4] A. Sebsadji, N. Khouri, K. Djemal, D. Yepremian, F. Hareb, P. Hoppenot, E. Desailly, Description and classification of the effect of hamstrings lengthening in cerebral palsy children multi-site surgery, *Comp. Methods Biomech. Biomed. Eng.* 15 (September (Suppl. 1)) (2012) 177–179.
- [5] M.H. Schwartz, A. Rozumalski, W. Truong, T.F. Novacheck, Predicting the outcome of intramuscular psoas lengthening in children with cerebral palsy using preoperative gait data and the random forest algorithm, *Gait Posture* 37 (April (4)) (2013) 473–479.
- [6] J.L. Hicks, S.L. Delp, M.H. Schwartz, Can biomechanical variables predict improvement in crouch gait? *Gait Posture* 34 (June (2)) (2011) 197–201.
- [7] K. Sullivan, J. Richards, F. Miller, P. Castagno, N. Lennon, Predicting the outcome of surgery for children with cerebral palsy using pre-operative gait analysis, *Gait Posture* 3 (2) (1995) 92.
- [8] L.A. Hersh, J.Q. Sun, J.G. Richards, F. Miller, The prediction of post-operative gait patterns using neural networks, *Gait Posture* 5 (2) (1997) 151.
- [9] O.A. Galarraga, C.V. Vigneron, B. Dorizzi, N. Khouri, E. Desailly, Predicting postoperative knee flexion and pelvic tilt at initial contact of cerebral palsy children, *Mov. Sport Sci. - Sci. Mot.* 93 (2016) 87–92.
- [10] T.A. Nüller, J.G. Richards, F. Miller, Concurrent surgeries are a factor in predicting success of rectus transfer outcomes, *Gait Posture* 26 (June (1)) (2007) 76–81.
- [11] R.B. Davis III, S. Öunpuu, D. Tyburski, J.R. Gage, A gait analysis data collection and reduction technique, *Hum. Mov. Sci.* 10 (October (5)) (1991) 575–587.
- [12] M.P. Kadaba, H.K. Ramakrishnan, M.E. Wootten, et al., Measurement of lower extremity kinematics during level walking, *J. Orthop. Res.* 8 (3) (1990) 383–392.
- [13] M. Templ, A. Kowarik, P. Filzmoser, Iterative stepwise regression imputation using standard and robust methods, *Comput. Stat. Data Anal.* 55 (10) (2011) 2793–2806.
- [14] R.O. Duda, P.E. Hart, D.G. Stork, *Pattern Classification*, 2nd edition, Wiley-Blackwell, New York, 2000.
- [15] E. Desailly, Y. Daniel, P. Sardain, P. Lacouture, Foot contact event detection using kinematic data in cerebral palsy children and normal adults gait, *Gait Posture* 29 (January (1)) (2009) 76–80.
- [16] M.H. Schwartz, A. Rozumalski, The gait deviation index: a new comprehensive index of gait pathology, *Gait Posture* 28 (October (3)) (2008) 351–357.
- [17] I.T. Jolliffe, *Principal Component Analysis*, Springer-Verlag, New York, 2002.
- [18] C.M. Bishop, *Pattern Recognition And Machine Learning*, 2nd ed., Springer-Verlag, New York, 2006.
- [19] R.G. Miller, The jackknife—a review, *Biometrika* 61 (April (1)) (1974) 1–15.
- [20] Stuart, Kendall's *Advanced Theory of Statistics: Distribution Theory*, 6th edition, Wiley-Blackwell, Chichester, West Sussex, 1994.
- [21] R. Baker, J.L. McGinley, M.H. Schwartz, S. Beynon, A. Rozumalski, H.K. Graham, O. Tirosh, The gait profile score and movement analysis profile, *Gait Posture* 30 (October (3)) (2009) 265–269.
- [22] S. Klejman, J. Andrysek, A. Dupuis, V. Wright, Test-retest reliability of discrete gait parameters in children with cerebral palsy, *Arch. Phys. Med. Rehabil.* 91 (May (5)) (2010) 781–787.
- [23] K. Desloovere, G. Molenaers, H. Feys, C. Huenaearts, B. Callewaert, P.V. de Walle, Do dynamic and static clinical measurements correlate with gait analysis parameters in children with cerebral palsy? *Gait Posture* 24 (November (3)) (2006) 302–313.
- [24] A. Bonnefoy-Mazure, Y. Sagawa Jr., V. Pomero, P. Lascombes, G. De Coulon, S. Armand, Are clinical parameters sufficient to model gait patterns in patients with cerebral palsy using a multilinear approach? *Comput. Methods Biomech. Biomed. Eng.* 19 (7) (2016) 800–806.