

**Physical and statistical shape modelling
in craniomaxillofacial surgery: a
personalised approach for outcome
prediction**

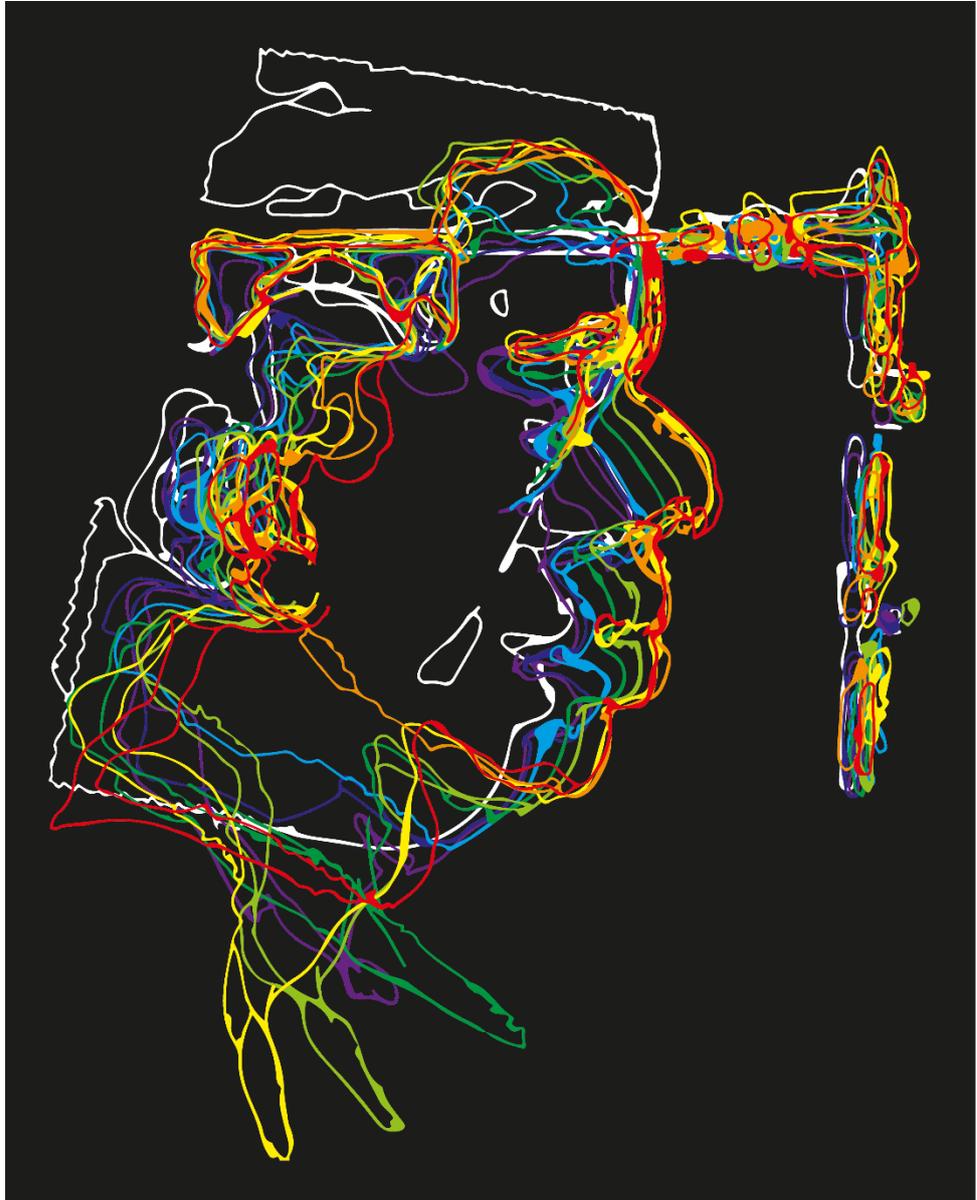
Paul G.M. Knoops

Great Ormond Street Institute of Child Health
University College London

This dissertation is submitted for the degree of

Doctor of Philosophy

London, March 2019



Changing faces/Facing changes

DECLARATION

I, Paul Gijsbertus Mari Knoops, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Date

Signed

ABSTRACT

Orthognathic surgery involves repositioning of the jaw bones to restore face function and shape for patients who require an operation as a result of a syndrome, due to growth disturbances in childhood or after trauma. As part of the preoperative assessment, three-dimensional medical imaging and computer-assisted surgical planning help to improve outcomes, and save time and cost. Computer-assisted surgical planning involves visualisation and manipulation of the patient anatomy and can be used to aid objective diagnosis, patient communication, outcome evaluation, and surgical simulation. Despite the benefits, the adoption of three-dimensional tools has remained limited beyond specialised hospitals and traditional two-dimensional cephalometric analysis is still the gold standard.

This thesis presents a multidisciplinary approach to innovative surgical simulation involving clinical patient data, medical image analysis, engineering principles, and state-of-the-art machine learning and computer vision algorithms. Two novel three-dimensional computational models were developed to overcome the limitations of current computer-assisted surgical planning tools. First, a physical modelling approach – based on a probabilistic finite element model – provided patient-specific simulations and, through training and validation, population-specific parameters. The probabilistic model was equally accurate compared to two commercial programs whilst giving additional information regarding uncertainties relating to the material properties and the mismatch in bone position between planning and surgery. Second, a statistical modelling approach was developed that presents a

paradigm shift in its modelling formulation and use. Specifically, a 3D morphable model was constructed from 5,000 non-patient and orthognathic patient faces for fully-automated diagnosis and surgical planning. Contrary to traditional physical models that are limited to a finite number of tests, the statistical model employs machine learning algorithms to provide the surgeon with a goal-driven patient-specific surgical plan. The findings in this thesis provide markers for future translational research and may accelerate the adoption of the next generation surgical planning tools to further supplement the clinical decision-making process and ultimately to improve patients' quality of life.

IMPACT STATEMENT

Craniomaxillofacial surgery aims to restore face function and shape for patients who require surgery as a result of a syndrome, due to growth disturbances in childhood or after trauma. As part of the preoperative assessment, a surgeon can test various operations in a virtual environment – before actually doing the operation in the theatre – to determine which patient-specific approach will give the best functional and aesthetic result. Commercial programs can be used for this process, known as computer-assisted surgical planning, and the benefits include cost reduction and safer and more precise surgery. Despite these advantages, the initial cost, learning curve, and time investment needed to familiarise with computer-assisted surgical planning have prevented adaptation beyond specialised hospitals, and the lack of accuracy means that patients have to be informed about the shortcoming during a consultation.

To overcome these limitations and make computer-assisted surgical planning a more accessible and accurate technology, this thesis presents a multidisciplinary approach to computer-assisted surgical simulations involving clinical patient data, medical image analysis, engineering methods, and state-of-the-art machine learning and computer vision algorithms. Two new methodologies were developed and validated; one based on readily available engineering tools and another based on a novel machine learning model. These new models can have an impact clinically, academically, and commercially. Clinically, the former model

provided insightful simulations that illustrate the potential outcomes of surgery and elucidate how the predicted result might differ from the achieved result after surgery – analogous to how a weather forecast incorporates minimum, most likely, and maximum predicted values – which can improve the patient communication and clinical decision-making processes. Moreover, as this approach relied on readily available engineering tools, after further validation in prospective studies, this methodology can be implemented in commercial software. The statistical model is a shift from current approaches as it infers data from a large collection of faces as opposed to calculate how the shape of a face changes from modifications to the underlying skeleton. The high accuracy and fast computation time suggested that the machine-learning-based method not only provides an interesting line of future research but also that such tools can have an impact in clinical decision-making and surgical planning. Both models will help to make computer-assisted surgical planning more accessible, ultimately to support and empower surgeons and to improve patients' quality of life.

ACKNOWLEDGEMENTS

This thesis would not have been possible without the help and support of many others. First and foremost, I would like to thank my supervisors Silvia Schievano, David Dunaway, and Alessandro Borghi for their mentorship, encouragement, and trust. I am proud of the work we have achieved, and it has been a great journey together. My appreciation also goes to Owase Jeelani, without your unlimited enthusiasm none of this would have been possible. Furthermore, I would like to thank the members of the craniofacial group who contributed to a very pleasurable and fruitful working climate; Will, Will, Naiara, Freida, Thanos, Lara, Selim, Greg, Juling, Justine, and Richard – many thanks! Additionally, as this work would not have been possible without the help of our collaborators, I would like to express my appreciation to Federica Ruggiero and the Bologna group, a special thank you to Bonnie Padwa for the wonderful time you facilitated at Boston Children’s Hospital, and my gratitude to Derek Steinbacher and the team at Yale who unequivocally supported our collaborative effort.

Last but not least, the support of my family and friends has been phenomenal and fundamental to the achievement this thesis presents; thank you Mario, Marianne, Erik, Lucy, and all others!

TABLE OF CONTENTS

| | |
|--|--------------|
| DECLARATION | V |
| ABSTRACT | VII |
| IMPACT STATEMENT | IX |
| ACKNOWLEDGEMENTS | XI |
| TABLE OF CONTENTS | XIII |
| LIST OF FIGURES | XVII |
| LIST OF TABLES | XXI |
| LIST OF ABBREVIATIONS | XXIII |
| CHAPTER 1 INTRODUCTION | 1 |
| 1.1 INTRODUCTION | 2 |
| 1.2 AIMS AND OBJECTIVES | 6 |
| 1.3 OUTLINE OF THE THESIS | 9 |
| CHAPTER 2 BACKGROUND | 11 |
| 2.1 COMPUTER-ASSISTED SURGICAL PLANNING | 12 |
| 2.2 EMPIRICAL MODELS | 15 |
| 2.3 PHYSICAL MODELS | 16 |
| 2.3.1 Mass-spring-models and mass-tensor models | 17 |
| 2.3.2 Finite Element Models | 19 |
| 2.4 STATISTICAL MODELS | 22 |
| 2.5 SUMMARY | 24 |
| CHAPTER 3 3D CRANIO-MAXILLO-FACIAL IMAGING | 27 |
| 3.1 INTRODUCTION | 28 |
| 3.2 METHODOLOGY | 30 |

| | | |
|---|---|-----------|
| 3.2.1 | Data acquisition and processing | 30 |
| 3.2.2 | Data analysis and statistics | 34 |
| 3.3 | RESULTS | 36 |
| 3.3.1 | Scanner overview | 36 |
| 3.3.2 | Accuracy | 37 |
| 3.3.3 | Repeatability | 38 |
| 3.3.4 | Post-processing and landmark digitisation error | 40 |
| 3.4 | DISCUSSION | 41 |
| 3.5 | SUMMARY | 44 |
| CHAPTER 4 PROBABILISTIC FINITE ELEMENT MODELLING | | 47 |
| 4.1 | INTRODUCTION | 48 |
| 4.2 | METHODOLOGY | 49 |
| 4.2.1 | Patient population | 49 |
| 4.2.2 | Probabilistic finite element modelling | 50 |
| 4.3 | RESULTS | 56 |
| 4.4 | DISCUSSION | 60 |
| 4.5 | SUMMARY | 64 |
| CHAPTER 5 CLINICAL COMPARISON OF PHYSICAL MODELS | | 67 |
| 5.1 | INTRODUCTION | 68 |
| 5.2 | METHODOLOGY | 69 |
| 5.2.1 | Patients | 69 |
| 5.2.2 | Surgical simulation and soft tissue prediction | 71 |
| 5.2.3 | Data analysis | 73 |
| 5.3 | RESULTS | 74 |
| 5.4 | DISCUSSION | 79 |
| 5.5 | SUMMARY | 82 |
| CHAPTER 6 3D MORPHABLE MODELS | | 85 |
| 6.1 | INTRODUCTION | 86 |

| | | |
|--|---|------------|
| 6.2 | METHODOLOGY | 88 |
| 6.2.1 | Data sources | 88 |
| 6.2.2 | 3D morphable model construction | 91 |
| 6.2.3 | Error quantification | 94 |
| 6.3 | RESULTS | 95 |
| 6.3.1 | Model validation | 95 |
| 6.3.2 | Qualitative and quantitative shape analysis | 95 |
| 6.3.3 | Manifold visualisation | 98 |
| 6.3.4 | Classification for diagnosis | 100 |
| 6.3.5 | Regression for surgical simulation | 101 |
| 6.4 | DISCUSSION | 105 |
| 6.5 | SUMMARY | 109 |
| CHAPTER 7 COMBINED STATISTICAL AND PHYSICAL MODELLING | | 111 |
| 7.1 | INTRODUCTION | 112 |
| 7.2 | METHODOLOGY | 113 |
| 7.2.1 | Patient characteristics | 113 |
| 7.2.2 | Combined modelling framework | 113 |
| 7.2.3 | Data analysis | 115 |
| 7.3 | RESULTS | 116 |
| 7.3.1 | Postoperative maxilla changes | 116 |
| 7.3.2 | Statistical model simulation | 117 |
| 7.3.3 | Inverse physical model simulation | 118 |
| 7.4 | DISCUSSION | 120 |
| 7.5 | SUMMARY | 122 |
| CHAPTER 8 CONCLUSIONS | | 125 |
| 8.1 | OVERVIEW | 126 |
| 8.2 | DETAILED OUTCOMES | 126 |
| 8.3 | LIMITATIONS AND FUTURE DIRECTIONS | 131 |
| 8.3.1 | Sample size and data | 131 |

| | |
|--|------------|
| 8.3.2 Error evaluation, clinical accuracy and subjectivity | 132 |
| 8.4 CONCLUSIONS | 134 |
| REFERENCES | 135 |

LIST OF FIGURES

- Figure 1.1 Le Fort I, II, and III osteotomies. 3
- Figure 2.1 Computer-assisted surgical planning framework. Tomographic data (CT or MRI) is processed and a 3D reconstruction of patient anatomy is generated, virtual surgery is performed, and various post-processing tasks can be carried out. Note that this is a simplified representation; additional steps may ensue including merging of CT data with 3D surface data to include texture, cephalometric analysis, orthodontic planning. 14
- Figure 2.2 Tetrahedral mesh of the facial bones and soft tissues. The mesh comprises different parts: mandible (green), maxilla (yellow), cranium (red), and soft tissues (transparent blue). 17
- Figure 3.1 3D scanners, mesh overview, and mesh detail. Mesh representation of the data obtained using the Avanto MRI, 3dMDface System, M4D Scan, and Structure Sensor. Note: MRI does not capture hair and 3D surface scanners have great difficulty with hair too; therefore, a cap is often used as shown in the M4D scan mesh. 33
- Figure 3.2 3D data processing steps. Following ICP registration, 3D scans were imported into Rhinoceros: a cutting plane was defined between the left and right tragus and the chin, and a second plane was constructed orthogonal to the first; these planes were used to crop the scan; and a smooth filter was applied. 34
- Figure 3.3 Facial colourmap showing the agreement with the reference. The 3dMDface System scan (grey, top) is shown, which is the reference image for the other three scans. Green indicates highly reliable areas, whilst red (underestimation) and blue (overestimation) indicate unreliable areas. Participant 6 is shown which was representative for the cohort of participants. 37
- Figure 3.4 Percentage of data points within deviation ranges. The error bars indicate the accuracy of the Avanto MRI (dark grey), M4D Scan (grey), and Structure Sensor (light grey), relative to the 3dMDface System, within various ranges: 0–1 mm (highly reliable), 1–1.5 mm (reliable), 1.5–2 mm (moderately reliable), and >2 mm (unreliable). Mean and standard deviation were computed from all scans (n=8). 38
- Figure 4.1 Preoperative CBCT reconstruction of bone and soft tissue. The fixed boundary conditions are indicated with arrows, and four soft tissue points for design of experiments were: nose (N), upper lip (UL), right cheek (RC), and left cheek (LC). 51
- Figure 4.2 Flowchart of the methodology for the probabilistic finite element analysis with colour coding: variable correlation, design of experiments I, optimisation on material properties, and design of experiments II. 52
- Figure 4.3 Soft tissue prediction results. For patient 1 the range of the soft tissue prediction: (a) minimum with $G_{\text{soft}} = 94.4\%$, $E_{\text{soft}} = 0.1 \text{ MPa}$, $\nu_{\text{soft}} = 0.45$, $E_{\text{cart}} = 0.5 \text{ MPa}$, $x_{\text{disp}} = 3 \text{ mm}$, and (b) maximum with $G_{\text{soft}} = 30.5\%$, $E_{\text{soft}} = 1 \text{ MPa}$, $\nu_{\text{soft}} = 0.49$, $E_{\text{cart}} = 5 \text{ MPa}$, $x_{\text{disp}} = 7 \text{ mm}$. Grey

depicts no change from the postoperative CBCT (0 mm), red depicts a maximum change (6 mm). 58

Figure 4.4 Design of experiments I and II. (a) For all eight patients, the vertical bars depict the range of soft tissue prediction in each point and the black dots represent the true postoperative position from CBCT. The baseline (0 mm) is the preoperative CBCT. (b) Green (patient 2, 3, 5, 7 and 8): soft prediction range after optimisation and blue (patient 1, 4 and 6) soft tissue range from design of experiments II, based on optimised material properties. 59

Figure 4.5 Histograms of material property distribution. Material properties following optimisation, based on all training data (patient 2, 3, 5, 7 and 8). A Weibull curve (solid line) was fitted to the data, and the cumulative density (dotted line) is also shown. G: viscoelastic scale factor, E: Young's Modulus, ν : Poisson's ratio, soft: soft tissue, cart: cartilage. 60

Figure 4.6 Soft tissue prediction probability. (a) For patient 1: histograms and cumulative density plots illustrate the probability of the soft tissue displacement for four landmarks (nose tip (N), upper lip (UL), right cheek and left cheek) – the histograms for patient 4 and 6 showed similar probability curves. The arrows (red) on the x-axis depict the true postoperative soft tissue position, from CBCT, for the upper lip and nose. (b) The lateral view for patient 1, 4 and 6, with the preoperative outline (solid black line), predicted range (blue), and postoperative outline (dotted black line), indicating good agreement between the prediction and the postoperative position. 61

Figure 5.1 Pipeline for soft tissue prediction with Dolphin, ProPlan CMF and PFEM. That the image processing for the three methods is identical, consisting of ICP alignment, segmentation, osteotomy, and advancement. The differences are within the soft tissue prediction algorithms: Dolphin 3D is a landmark-based method and allows for patient-specific soft-to-hard tissue ratios to be set; ProPlan CMF does not require landmarks and is, therefore, relatively straightforward; and PFEM requires two separate programs for the full process, but allows for patient or population-specific material properties to be defined. 72

Figure 5.2 Distance colour maps of various soft tissue simulations. Colourmaps are relative to the (a) preoperative CBCT: (b) postoperative CBCT, (c) Dolphin, (d) ProPlan, and (e) PFEM. Patient-specific hard-to-soft tissue ratios were investigated in Dolphin and the corresponding frontal and lateral views are shown for: (f) minimum ratios, (g) default ratios, and (h) maximum ratios. Note that in (c) the chin support of the CBCT scanner is still present; the soft tissues and support have identical grey intensity values and Dolphin does not allow for manual tracing. 75

Figure 5.3 Image processing and accuracy of predictions. (a) Cutting plane as defined by the left and right tragus, subnasale, and stomion superior. (b) Comparison of soft tissue simulation with postoperative CBCT for Dolphin, (c) ProPlan CMF, and (d) PFEM. The colourmap demonstrates the overestimation (blue) and underestimation (red) of the simulation compared to the postoperative CBCT. 76

Figure 5.4 Range of advancements and the associated error evaluation. (a) RMS indicating the difference between the soft tissue prediction and the postoperative soft tissues. Each dot represents a simulation made with for a range of horizontal advancements (0-7 mm). The planned maxillary position was 5 mm and postoperative maxillary position was 3.3 mm at the A-point (dotted vertical lines). Soft tissue predictions based on the postoperative maxillary position correspond to the lowest RMS for PFEM and ProPlan CMF, whilst soft tissue predictions at 7 mm correspond to the lowest RMS for Dolphin. (b) The colourmaps illustrate how the predicted soft tissue changes between the planned and postoperative position, relative to the preoperative CBCT. 78

Figure 6.1 Patient characteristics. (a) Number of patients per surgical procedures, LF1: Le Fort 1 osteotomy, BSSO: bilateral sagittal split osteotomy, GP: genioplasty. (b) Number of patients with syndromic diagnosis, HFM: hemifacial microsomia. (c) Number of patients according to the main indication for surgery: history or no history cleft lip and/or palate (CLP). For those without a history of CLP, often multiple indications are present. (d) Timing of preoperative and postoperative scans: average preoperative scan was 73 ± 86 days (red, dotted line) and the average postoperative scan 326 ± 123 days (green, dotted line). 90

Figure 6.2 Compactness, generalisation, and specificity for the three models. Characterisation of the three 3DMM compared to the large-scale facial model (LSFM), a state-of-the-art benchmark. (a) Compactness, the amount of variance retained for a certain number of principal components, is 79.6% at 10 components for the global model (red), 81.8% for the bespoke preoperative model (green) and 79.9% for LSFM (blue). (b) Generalisation demonstrates the ability to describe new (patient) faces that were not used to construct the original model, and at 100 components is 0.3 mm, 0.4 mm, and 1.4 mm for the global model, bespoke preoperative model, and LSFM, respectively. Specificity measures how well synthetic faces resemble real faces, and (c) the bespoke preoperative model shows errors of 0.40 ± 0.01 mm, and (d) the global model 0.37 ± 0.01 mm. 96

Figure 6.3 Visualisation of the mean shape and variation for a face from the general public, a preoperative face, and a postoperative face. Following generation of the (a) global LSFM, (b) bespoke preoperative and (c) bespoke postoperative shape model, mean shape (μ) and first five shape eigenvectors are displayed, with weights for the standard deviation of +3SD (top row) and -3SD (bottom row). 97

Figure 6.4 Comparison of the mean LSFM, preoperative, and postoperative face. Colourmaps illustrate deviations from the mean LSFM face. The mean preoperative face colourmap is consistent with indications for orthognathic surgery – the cohort of orthognathic patients shows upper jaw underdevelopment (red) and lower jaw overdevelopment (blue). The mean surgical correction appropriately ameliorates these jaw malformations; however, the mean postoperative face retains nose abnormality. 98

Figure 6.5 t-SNE embedding of the high-dimensional face manifold. (a) The t-SNE embedding in two dimensions was generated with randomly sampled LSFM faces, for visualisation purposes, and labelled according to LSFM (blue, n=500), preoperative patient (red, n=119), and postoperative patient (green, n=127) faces. Lateral views of (b) a patient and (c), (d) two close neighbours to illustrate shape similarity within the t-SNE embedding, particularly in the upper lip angle; (e-f) Faces that correspond to false negatives in the classification experiment. 99

Figure 6.6 Classification of preoperative patient and non-patient LSFM faces. (a) A split of 80-20% provides 95.4% classification accuracy at 96 principal components, superior to other splits. (b) Average confusion matrix, obtained from classification using preoperative patient scans (n=140) and randomly selected LSFM scans (n=280), representing the average of 1,000 iterations. With an 80-20% split, patient (n=112) and LSFM (n=224) scans were used for training and patient (n=28) and LSFM (n=56) for testing. Shape abnormality was diagnosed with 95.5% sensitivity and 95.2% specificity and with a positive and negative predictive value of 87.5% and 98.3%, respectively. 100

Figure 6.7 False negatives and false positives. (a) For 1,000 iterations, 33 unique patient scans were classified as a false positive and 4 of those very often (more than 200 out of 1,000), signifying orthognathic surgery are not driven by aesthetic indications alone, and (b) 25 unique LSFM scans were classified as false negatives, of which 3 often (more than 150 out of 1,000), suggestive of mild jaw malformation or untreated patients in the non-patient sample. 101

Figure 6.8 Regression for postoperative face shape simulation. (a) Overall error of the ground-truth postoperative face shape compared to the simulated shape using ridge regression (RR), linear regression (LR), least-angle regression (LARS), and least absolute shrinkage and selection operator regression (LASSO). (b) Mean and standard deviation for ridge regression. (c) The best simulated face, with an error of 0.6 mm between the simulated shape and the postoperative face shape, and (d) the worst simulated face, with an error of 2.4 mm. 103

Figure 6.9 Regression error evaluation. (a) Colourmaps illustrate that the patient-specific simulations are better than predicting a face closer to the mean LSFM or the mean postoperative face; a comparison of error between the postoperative 3D scan (n=127) and patient-specific ridge regression simulations (1.1 ± 0.3 mm), the average postoperative face (1.6 ± 0.5 mm), and the average LSFM face (1.8 ± 0.5 mm). (b) Colourmaps elucidate how the best patient-specific simulation compares to the ground-truth postoperative face, the mean postoperative face, and the mean LSFM face. 104

Figure 7.1 Schematic overview of the combined modelling framework. The framework comprised two steps: statistical modelling and physical modelling. In the first step, the 3DMM automatically simulated the face shape outcome. In the second step, the bone movements were obtained via inverse application of PFEM. A comparison with preoperative and postoperative CBCT and 3D scans validated the results. 114

Figure 7.2 Pipeline to determine postoperative maxilla translation and rotation. (a) The preoperative (light grey) and postoperative (dark grey) STL files were imported into Matlab. (b) An iterative closest-point (ICP) algorithm was used to register the postoperative STL to the preoperative STL using the skull base as matching region. (c) Three landmarks were placed onto each maxilla to form a triangle: the upper incisor, mesial cusp 16 and mesial cusp 26; and (d) generalised Procrustes analysis (GPA) was used to place the two triangles into correspondence from which the translation (T) and rotation (R) were determined. 116

Figure 7.3 Preoperative and postoperative CBCT. (a) Preoperative CBCT, (b) colourmap of the postoperative CBCT with distances relative to the preoperative CBCT, red indicating differences up to 10 mm. (c) Translation and (d) rotation of the maxilla (dark grey). 117

Figure 7.4 Accuracy of the 3DMM simulation. The 3DMM simulation was generated using ridge regression from the preoperative 3D scan. The colourmap illustrates the agreement between the simulation and the postoperative 3D scan with RMS = 0.9 mm. 118

Figure 7.5 Contour plots of the agreement between PFEM and 3DMM simulations. The contour plots illustrate the overall agreement in soft tissue surface between 14 PFEM simulations and the 3DMM. (a) Horizontal advancement and vertical repositioning, and (b) horizontal advancement and lateral repositioning. The best simulation result was obtained at 7.6 mm horizontal advancement, 0 mm vertical repositioning, and 0.5 mm lateral repositioning (orange dot). The contour plot is represented by a polynomial fit between the 14 simulations, and the colour indicates the magnitude of the RMS from 1.5 mm (blue) to 3.0 mm (yellow). 119

LIST OF TABLES

| | |
|--|-----|
| Table 3.1 Characteristics of 3D scanners..... | 32 |
| Table 3.2 Overview the three studies and the datasets involved. | 36 |
| Table 3.3 Accuracy of Avanto MRI, M4D Scan, and Structure Sensor. Relative to the 3dMDface System, measured as the root mean square distance (RMS) and mean and standard deviation (SD)..... | 39 |
| Table 3.4 Repeatability of M4D Scan and Structure Sensor | 39 |
| Table 3.5 Post-processing error. One dataset was analysed 5 times (i.e. steps 2 – 4 in Figure 3.2) for each of the three scanners, relative to the 3dMDface System. Root mean square distance (RMS) and standard deviation (SD) are shown. | 40 |
| Table 4.1 Details for eight patients who consecutively had orthognathic surgery. | 50 |
| Table 4.2 Input parameters and material properties E: Young’s modulus, ν : Poisson’s ratio, G: viscoelasticity (Prony series). | 53 |
| Table 4.3 Variable correlation. Correlation of input variables and output parameters. E: Young’s modulus, ν : Poisson’s ratio, G: viscoelasticity (Prony series), x: advancement, soft: soft tissue, cart: cartilage, N: nose, UL: upper lip, RC: right cheek, LC: left cheek, r: Pearson’s correlation, CI: confidence interval. | 57 |
| Table 5.1 Patient characteristics. Mean and standard deviation (SD) are shown in the bottom row except for the male-female | 70 |
| Table 5.2 Root mean square distance and percentage <2 mm of the soft tissue prediction compared to the postoperative CBCT. Mismatch is the difference in maxillary position in the sagittal plane at the A-point between the postoperative CBCT and the planning. RMS: root mean square, P: percentage..... | 77 |
| Table 6.1 Face database characteristics. n is the number of individuals for whom that measurement was available. SD: standard deviation. | 89 |
| Table 7.1 Comparison of the rotation and translation from postoperative CBCT and simulation. | 120 |

LIST OF ABBREVIATIONS

| | |
|-------|--|
| 2D | Two-dimensional |
| 3D | Three-dimensional |
| 3DMM | 3D morphable model |
| BSSO | Bilateral sagittal split osteotomies |
| CAD | Computer-aided design |
| CAM | Computer-aided manufacturing |
| CBCT | Cone beam computed tomography |
| CT | Computed tomography |
| DICOM | Digital imaging and communications in medicine |
| DOE | Design of experiments |
| FEM | Finite element model |
| GOSH | Great Ormond Street Hospital for Children |
| GPA | Generalised Procrustes analysis |
| HFM | Hemifacial microsomia |
| ICP | Iterative closest-point algorithm |
| LARS | Least-angle regression |
| LASSO | Least absolute shrinkage and selection operator regression |
| LF1 | Le Fort I osteotomy |
| LR | Linear regression |
| LSFM | Large-scale facial model |
| MRI | Magnetic resonance imaging |
| MSM | Mass-spring-model |
| MTM | Mass-tensor-model |
| OSF | Optimal space-filling |
| PCA | Principal component analysis |
| PFEM | Probabilistic finite element model |
| PROMs | Patient reported outcome measures |
| RMS | Root mean square distance |

| | |
|-------|--|
| RR | Ridge regression |
| SD | Standard deviation |
| SMAS | Superficial muscular aponeurotic system |
| SVM | Support vector machine |
| t-SNE | t-distributed stochastic neighbour embedding |
| VMTK | Vascular modelling toolkit |

Chapter 1 **INTRODUCTION**

1.1 Introduction

Orthognathic surgery, the focus of this thesis, involves repositioning of the upper, lower, or both jaw bones to correct their spatial relationship and the overall proportion of the skull and face shape. Whilst a dental malocclusion is relatively common in the general population and can often be treated by orthodontics approaches alone, a combination of orthognathic surgery and orthodontics is needed for patients with severe jaw malformation (Soh and Narayanan, 2013; Obwegeser, 2016). Discrepancies of the jaw can present due to growth disturbances in childhood, as a result of a syndrome, after trauma, or without a (known) specific cause. Growth disturbances may occur due to childhood jaw fracture or repair of cleft lip and palate, which is the most common birth defect with an incidence of 1 in 600-700 live births (Mossey and Castilla, 2001; Kuijpers-Jagtman, 2013). Syndromes that are known to cause malformation of the jaw and midfacial bones include hemifacial microsomia (HFM), Pierre Robin complex, Treacher Collins syndrome, Apert syndrome, and Crouzon syndrome (Kharbanda and Wadhawan, 2013). These syndromes can also manifest in other craniofacial abnormalities and limb malformations (Ko, 2016), for which patients may need additional surgical treatment.

The modern-day management of orthognathic surgery was revolutionised by Hugo Obwegeser (1920-2017) (Steinhäuser, 1996; Naini, 2016) who devised, amongst other techniques, the bilateral sagittal split osteotomy (BSSO) procedure, performed the first osseous genioplasty on a living patient, and reported on the first bimaxillary procedure involving simultaneous mobilisation of the maxilla and mandible (Naini, 2017). Whilst others preceded Obwegeser, his techniques have stood the test of time (Naini, 2016).



Figure 1.1 Le Fort I, II, and III osteotomies.

Whilst orthognathic surgery focuses on ameliorating dental malocclusion, craniofacial surgery aims to restore face shape and function by repositioning the neurocranium and facial bones, with much overlap between the two specialities. Rene Le Fort (1869-1951), a French orthopaedic surgeon by training, made significant contributions to orthognathic and craniofacial surgery as he performed experiments that led to a classification system of facial fractures: Le Fort I, II, and III fractures (Figure 1.1). These fracture patterns are nowadays used to indicate where an *osteotomy* (bone cut) is made in various orthognathic and craniofacial procedures.

In contemporary medical practice in the UK, oral and maxillofacial surgeons perform orthognathic surgery, whilst craniofacial surgery is a subspecialty within plastic surgery. However, complex patients who need numerous plastic and reconstructive surgeries throughout life are followed up by multidisciplinary clinical teams at specialised hospitals, including Great Ormond Street Hospital for Children (GOSH). The teams comprise oral and maxillofacial surgeons, craniofacial surgeons, plastic surgeons, neurosurgeons, as well as

orthodontists, radiologists, ophthalmologists, psychologists, and speech and language therapists. In this thesis, the term craniomaxillofacial surgery is used to refer to craniofacial surgery, maxillofacial surgery and orthognathic surgery collectively.

In addition to the personal contribution of numerous surgeons, craniomaxillofacial surgeries have benefitted considerably from advances in medical imaging (Schendel *et al.*, 2012; Naini, 2016) which have facilitated the development of surgical planning tools (Keeve *et al.*, 1998). Nowadays, different non-invasive imaging techniques including medical photography, orthopantomogram and lateral cephalic radiographs (both X-ray), computed tomography (CT), cone beam computed tomography (CBCT), and magnetic resonance imaging (MRI) can be used to obtain a representation of head anatomy (Wippold, 2007; Benson *et al.*, 2014; Eley *et al.*, 2014).

Initial computer-assisted surgical planning and simulation methods were developed based on digitised lateral cephalic radiographs. In two-dimensions, landmarks were identified and bone cuts made, and the postoperative soft tissue profile could be simulated based on empirically derived bone-to-soft tissue ratios (San Miguel Moragas, Van Cauteren and Mommaerts, 2014; San Miguel Moragas *et al.*, 2015). This two-dimensional planning approach using lateral cephalic radiographs is an established method, but the bone-to-soft tissue ratios are not easily determined and large prospective studies still are needed to delineate the influence of the type of osteotomy, magnitude of bone movement, age, gender, and ethnicity (San Miguel Moragas, Van Cauteren and Mommaerts, 2014). Recently, based on the same landmark methodology, three-dimensional (3D) planning has gained popularity. However, assessment of the bone-to-soft tissue ratios in 3D is lacking (Olate, Zaror and Mommaerts, 2017) and sparse landmark-based approaches are unable to take full advantage of high-

resolution dense imaging, thus resulting in low accuracy of predicted soft tissue models (Resnick *et al.*, 2017). In the last ten years, powerful reconstruction algorithms have been developed for more sophisticated 3D surgical planning using high-quality digital patient models (Stokbro *et al.*, 2014; Schendel, 2015), nevertheless, predicting the soft tissue changes remains challenging (Zhang *et al.*, 2016). To further improve on computer-assisted surgical tools, large databases of high-quality image data are needed in combination with advanced tools such as machine learning algorithms. In this context, the aim of this thesis is described in detail in the next paragraph and an overview of the chapter contents is provided.

1.2 Aims and objectives

In the field of orthognathic surgery, whilst many mathematically and computationally advanced methods are available in the literature, only a fraction of those are adopted in clinical practice, and the majority of hospitals relies on two-dimensional (2D) cephalometric analysis and model surgery to assess and manage patients (Anwar and Harris, 1990). Therefore, this thesis embodies an effort to translate state-of-the-art engineering, computer vision and machine learning principles into clinical tools to provide improved diagnosis, and more accurate and automated planning and simulations.

The principal aim of this thesis is to develop, apply, and validate novel 3D computational models, and, combined with large medical image datasets, enhance computer-assisted surgical planning – ultimately to improve clinical decision-making and patients’ quality of life.

Three objectives are defined to help achieve this aim.

- **Objective 1: Evaluate the range of 3D imaging tools at GOSH.**

As noted, the surgical planning framework demands high-quality 3D anatomical data as input. Therefore, the first objective is to investigate the 3D scanning systems available at GOSH to establish which image sources are the most suitable for surgical planning and facial analysis, and in parallel to establish protocols for prospective data collection.

-
- **Objective 2: Develop and validate a probabilistic physical model for orthognathic surgery.**

Once the imaging protocols are established, two methods for improved surgical simulation are developed: first a stochastic physical model, aiming to improve upon deterministic models typically used in 3D surgical simulations; and second a statistical model which represents a paradigm shift by employing a machine-learning-based algorithm, unique for surgical simulation in its mathematical formulation and use. The physical modelling approach includes a finite element model (FEM) coupled with a design of experiments (DOE) scheme, referred to as the probabilistic finite element model (PFEM). As opposed to the classical deterministic FEM formulation, which provides an identical output for each recurrent model computation, the outcome of the probabilistic model depends on the uncertainties in mechanical, geometric and loading properties (Stefanou, 2009). This is particularly relevant for face simulations when patient-specific parameters such as skin material properties are unknown and show large intra- and interpatient variation (Luboz, Promayon and Payan, 2014). Although probabilistic modelling capabilities were implemented in commercial FEM packages a decade ago (Reh *et al.*, 2006) and FEM is a popular method for surgical simulations, to the best of my knowledge probabilistic models have not yet been used for surgical simulations of the face shape. Therefore, the second objective is to develop and validate the PFEM approach with patient data and demonstrate its clinical utility.

- **Objective 3: Develop and validate a statistical model based on state-of-the-art machine learning algorithms for orthognathic surgery.**

The third objective is to progress from explorative computer-assisted approaches based on physical models to automated approaches by adopting state-of-the art computer science algorithms (Zachow, 2015). Explorative models, such as PFEM, are limited to a finite number of tests that the surgeon can explore and modify using the surgical planning platform. On the contrary, an artificial intelligence approach based on a statistical model can be fully automated to objectively and near-instantaneously predict the required surgical intervention – effectively taking out the subjective surgeon expertise. Comprehensive databases of patient and normal faces are needed to develop such a goal-driven statistical model for clinical application, and current models only incorporated several hundred faces of either normal or patient faces (Paysan *et al.*, 2009; Bolkart and Wuhrer, 2015; Staal *et al.*, 2015; Huber *et al.*, 2016; Kaya *et al.*, 2018; Maas *et al.*, 2018). Therefore, the third objective is to construct the first large-scale statistical model for clinical utility – based on a state-of-the-art machine learning framework (Booth *et al.*, 2018) – using 5,000 faces including data from the general public and from patients who had orthognathic surgery, and demonstrate its potential in diagnosis and surgical planning.

1.3 Outline of the thesis

This thesis has the following structure. **Chapter 2** presents an introduction to medical imaging and computer-assisted surgical planning, and includes an introduction to physical and statistical modelling, in line with the objectives.

Chapter 3 compares various 3D scanning systems at GOSH and evaluates their suitability for facial imaging, focussing on 3D shape analysis and the use in surgical planning. The study includes static and portable systems, and those of high and low cost.

Chapter 4 describes in detail the physical modelling pipeline central to this thesis. FEM are typically deterministic, whilst probabilistic models can be used to account for uncertainties, including the variation in material properties and the uncertainty on the bone repositioning. This approach quantifies how errors associated with the input parameters propagate into the predicted face shape, thereby aiding planning and decision-making, and potentially improving patient communication.

Chapter 5 demonstrates the retrospective application of PFEM on a cohort of patients who had orthognathic surgery, and a quantitative and qualitative analysis is performed to compare PFEM with two commercially available programs. This chapter provides insight into the strengths and limitations of each commercial program and PFEM.

Chapter 6 gives a detailed description of the statistical modelling pipeline, as an alternative approach to physical modelling. A database containing 5,000 faces is managed using an automated image processing framework and a 3D morphable model (3DMM) – a statistical shape model of face shape – is constructed. This chapter demonstrates the clinical

utility of 3DMM including fully-automated surgical planning as well as classification for diagnosis.

Chapter 7 presents a combined statistical and physical approach. To guide a surgeon on the bone movements required in an operation, the statistical model is utilised to find the optimum postoperative face shape, and the PFEM is employed inversely to obtain the bone movements. This presents an automated goal-oriented approach, as opposed to an explorative one, and the benefits and shortcomings of this method are illustrated.

Chapter 8 summarises the main outcomes and contributions to the field; conclusions are drawn and perspectives for future research are presented.

Chapter 2 **BACKGROUND**

This chapter describes the main concepts in computer-assisted surgical planning and its perioperative application, which includes patient communication, intraoperative utilisation, and outcome evaluation (Schendel, 2015). Additionally, the computational models within these planning tools are introduced and an overview of commercial software is presented.

2.1 Computer-assisted surgical planning

Orthognathic surgery, aiming at correcting a discrepancy between upper and lower jaws due to growth disturbances, syndromes, or other facial asymmetries or malformations, results in a modified dental occlusion and simultaneously a change in the facial appearance. Depending on the type and severity of the jaw discrepancy, a Le Fort I osteotomy is typically used to correct the maxilla whilst a BSSO to correct the mandible. In some cases, a bimaxillary operation is used to correct the maxilla and mandible in the same sitting. In either operation, the osteotomised segments are repositioned and prefabricated wafers are used to maintain the correct position, before the repositioned segments are fixed using metal plates and screws. To ensure that these steps lead to the best aesthetic and functional result, as an alternative to the destructive model surgery process, preoperative planning is performed using 3D imaging complied with computer-assisted surgical planning. Specifically, the computer virtual environment allows the surgeon to explore numerous therapeutic concepts, determine the necessary instrumentation, and assess the simulated outcome of the face (Chabanas, Luboz and Payan, 2003; Zachow, 2015) – for the benefit of a more precise, faster, and cost-effective surgery (Xia *et al.*, 2006). Since its inception (Vannier, Marsh and Warren, 1984), the computer-assisted surgical planning framework has largely remained unchanged – it comprises medical imaging, image processing, virtual surgery, and may involve one or multiple post-

processing tasks (Figure 2.1) (Keeve *et al.*, 1998; Chabanas *et al.*, 2002; Cevidanes *et al.*, 2010; Zachow, 2015). The framework starts with medical imaging, and CBCT is routinely used in craniomaxillofacial surgery as it provides the necessary high-quality volumetric bone data (Cevidanes *et al.*, 2010). Recently, MRI has been suggested as a non-ionising alternative, but technical difficulties remain, for example in the identification of the boundary between bone and air in the sinuses (Eley *et al.*, 2014). The Digital Imaging and Communication in Medicine (DICOM) file format is used to store tomographic data, and medical images are semi-automatically segmented to construct volumetric digital patient models of bone and soft tissue (Zachow, 2015). In the case of multimodality registration, tomographic data are merged with 3D stereophotogrammetry scans to create models of shape and texture (Cevidanes *et al.*, 2010). The digital patient model is then manipulated virtually to simulate one or multiple osteotomies, and the mobilised segment can be repositioned until appropriate dental occlusion is achieved. The response of the soft tissues to the underlying bone changes are computed using a mathematical model, which allows the simulated face shape to be obtained. Multiple iterations of the above steps may be needed to devise a plan that provides satisfactory aesthetic and function results. As part of the surgical planning process, a number of other post-processing tools can be used including intraoperative navigation; computer-aided design (CAD) and manufacturing (CAM) of intraoperative instruments including splints, bone screws, templates and implants (Aboul-Hosn Centenero and Hernández-Alfaro, 2012; Mazzoni *et al.*, 2015); and patient communication and outcome evaluation (Mollemans *et al.*, 2007) (Figure 2.1).

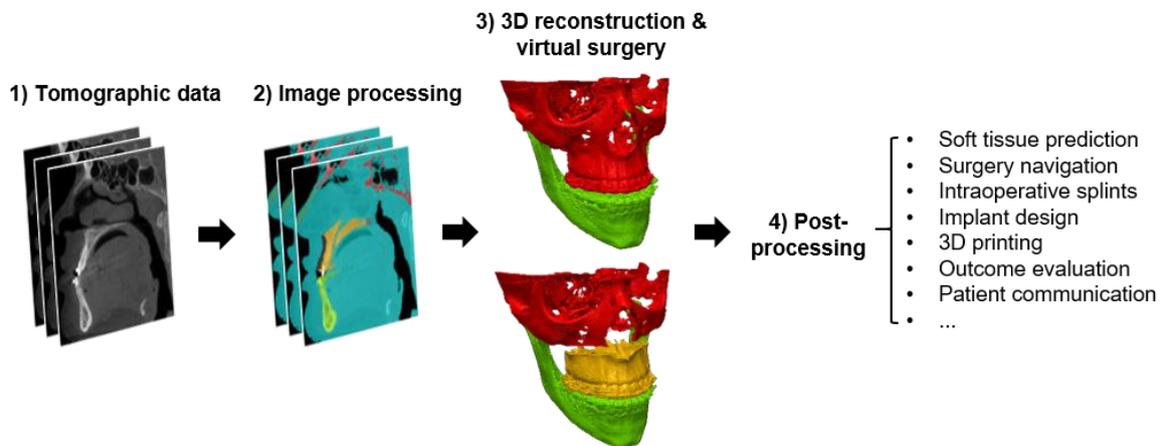


Figure 2.1 Computer-assisted surgical planning framework. Tomographic data (CT or MRI) is processed and a 3D reconstruction of patient anatomy is generated, virtual surgery is performed, and various post-processing tasks can be carried out. Note that this is a simplified representation; additional steps may ensue including merging of CT data with 3D surface data to include texture, cephalometric analysis, orthodontic planning.

In the past decade, many commercial computer-assisted surgical planning platforms have been launched onto the market (Cevitanes *et al.*, 2010), including:

- Dolphin (Dolphin Imaging & Management Solutions, Chatsworth, CA, USA)
- Maxilim (Medicim, Mechelen, Belgium)
- ProPlan CMF, Surgicase CMF (Materialise, Leuven, Belgium)
- SimPlant OMS (Dentsply-Sirona, York, PA, USA)
- 3dMDvultus (3dMD, Atlanta, GA, USA)
- IPS Case Designer (KLS Martin, Tuttingen, DE)

The above software packages employ various mathematical models to calculate the soft tissue displacements from the simulated bone movements (Mollema *et al.*, 2007). In the next

paragraphs, empirical models, physical models and statistical models are discussed and commercial packages are appraised in the applicable section.

2.2 Empirical models

As noted in chapter 1, computer-assisted surgical planning initially relied on digitised lateral cephalic radiographs (Ricketts, 1972). With advances in medical imaging, specifically 3D CT and CBCT, cephalometric analysis in 3D became feasible – a major improvement over 2D planning for clinical problems such as facial asymmetry and occlusal plane horizontality (Mollard, Lavallée and Bettega, 1998). Cephalometric-based planning in 2D and 3D involves manual placement of landmarks on the bone and on the soft tissue surfaces, repositioning of the osteotomised segment, and simulating the soft tissue movements based on empirically derived bone-to-soft tissue ratios (Keeve *et al.*, 1998). The flaws in this method limit the clinical utility (Keeve *et al.*, 1998; San Miguel Moragas, Van Caueren and Mommaerts, 2014; San Miguel Moragas *et al.*, 2015; Olate, Zaror and Mommaerts, 2017): the sparse representation with landmarks does not provide a dense and accurate representation of the anatomy and there is a lack of reliable data for the 3D bone-to-soft tissue ratios. Despite these limitations, the straightforwardness of this approach and the low computational cost makes real-time modelling feasible.

From the above-listed commercial software, Dolphin 3D (Dolphin Imaging and Management Solutions, Chatsworth, CA, USA) is the only computer-assisted surgical planning software based on an empirical model. Dolphin 3D originates from conventional lateral 2D cephalometric tracing and planning (Stokbro *et al.*, 2014) and it makes use of a landmark-based

photographic morphing algorithm that has been interpolated from 2D to 3D. Whilst Dolphin provided clinically acceptable simulations on the facial midline, inaccurate predictions of the soft tissues were observed laterally due to the sparse landmark-based algorithm (Resnick *et al.*, 2017). Chapter 5 describes a comparison between three soft tissue simulation methods including Dolphin, and the aforementioned limitations are discussed in detail.

2.3 Physical models

Physical models, like empirical models, use a mathematical description to compute the soft tissue response from the changes to the underlying facial skeleton, but rather than relying on empirically derived bone-to-soft tissue ratios, these models simulate the physical, mechanical behaviour of the face and incorporate elastic properties of facial soft tissues (Keeve *et al.*, 1998). Three main computational strategies are categorised within physical models: the mass-spring-model (MSM), the mass-tensor-model (MTM) and the finite element model (FEM) (Mollema *et al.*, 2007; Kim, Jürgens and Reyes, 2011). To construct such models and simulate soft tissues displacements, physical models first require a discretisation of the digital patient anatomical reconstruction. In the MSM, the patient features are subdivided into a connected system of points interlinked with springs and dampers – a mass-spring system (Teschner, Girod and Girod, 2000). The FEM treats the patient 3D volume as a continuum subdivided into a finite number of parts – finite elements – which behaviour is specified by a finite number of parameters (Zienkiewicz, Taylor and Zhu, 2005). Thus MSMs, MTMs and FEMs can be represented by a 3D mesh (Figure 2.2), but it is important to note the intrinsic differences.

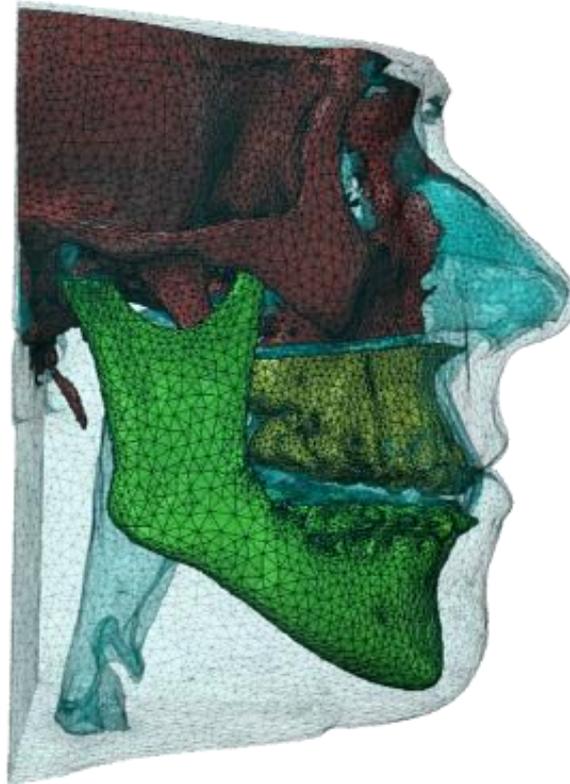


Figure 2.2 Tetrahedral mesh of the facial bones and soft tissues. The mesh comprises different parts: mandible (green), maxilla (yellow), cranium (red), and soft tissues (transparent blue).

2.3.1 Mass-spring-models and mass-tensor models

Soft tissue spring models originate from computer graphics where they were used for face animation (Terzopoulos *et al.*, 1987; Terzopoulos and Fleischer, 1988; Lee, Terzopoulos and Walters, 1995). Due to their computational simplicity and ease of implementation, MSMs became popular for application in surgical simulation (Delingette, 1998; Cotin, Delingette and Ayache, 1999; Teschner, Girod and Girod, 2000). The MSM encompasses a discrete modelling framework (Chabanas, Luboz and Payan, 2003) and in the most complex definition constitutes a layered assembly representing bone, muscles, fascia, dermis, and epidermis. The object consists of points x_i with mass m_i , interlinked by springs and dampers, which can be formulated

as a tetrahedral mesh (Mollema, 2003). The MSM constitutes of Newton's motion equation – a set of second-order ordinary differential equations (Teschner, Girod and Girod, 2000):

$$m_i \frac{d^2 \mathbf{x}_i(t)}{dt^2} + \gamma \frac{d\mathbf{x}_i(t)}{dt} = \mathbf{F}_i(t) \quad (2.1)$$

where γ is the damping factor, \mathbf{F}_i the sum of external forces and the elastic forces from neighbouring springs, and t is the time. Mechanical properties are approximated by spring constants, where different spring constants can be used to model different soft tissue layers. Furthermore, biphasic spring constants have been implemented to model non-linear stress-strain behaviour (Keeve *et al.*, 1998). An iterative minimisation towards zero forces in all points is used to obtain the dynamic solution (Mollema *et al.*, 2007). The time-dependent behaviour was relevant for animation applications but for most computer-assisted surgical planning applications the final rest position is more important than the dynamic behaviour. The final rest position can be directly computed in less than 1 second (Keeve *et al.*, 1998; Teschner, Girod and Girod, 2000). Although MSMs can be used for real-time modelling, drawbacks include the numerical stability as springs can 'flip' and inverse forces, the lack of conservation of volume, and a mismatch between the model parameters and elastic properties (Deuflhard, Weiser and Zachow, 2006; Cevitanes *et al.*, 2010).

The MTM was devised to overcome some of the above shortcomings. Real-time modelling capabilities are retained as the MTM and MSM have the same computational complexity (Cotin, Delingette and Ayache, 2000). The main improvement is in the difference of biomechanical modelling: the MSM have discrete masses and their behaviour depends on the topology of the interlinked springs whilst the MTM comprises a continuum model and its

accuracy mainly relates to the mesh resolution. Additionally, MTMs allow for direct representation of the Young's modulus and Poisson's ratio through Lamé coefficients as opposed to an approximation of spring constants in MSMs (Cotin, Delingette and Ayache, 2000). A remaining limitation is that MTMs are valid only for small deformations due to their linear formulation. Non-linear formulations were implemented, but this greatly impacted on the computation time (Picinbono, Delingette and Ayache, 2000).

Various commercial tools are based on MSM including ProPlan CMF, Maxilim, and 3dMDvultus (Mollema *et al.*, 2007; Schendel and Lane, 2009). Clinical studies with Maxilim reported clinically acceptable overall soft tissue agreement between simulation and observed postoperative face shape, although the simulations underestimated the true movements in all areas of the face except for the upper lip (Mundluru *et al.*, 2017). Therefore, when using simulations made with Maxilim for patient communication, the shortcomings in the prediction algorithm should be highlighted (Liebregts *et al.*, 2015). The comparison study in chapter 5 includes ProPlan CMF in addition to Dolphin and PFEM, and thus contains an MSM, an empirical model, and a FEM.

2.3.2 Finite Element Models

The FEM was developed for structural analysis in civil and aeronautical engineering in the 1940s (Hrennikoff, 1941; Courant, 1943) and were introduced for surgery simulation in the 1990s due to the high accuracy they provide (Koch *et al.*, 1996; Keeve *et al.*, 1998). FEMs discretise an object into elements; however, unlike MSMs, they represent mathematically defined subdivisions of a continuum problem, where biomechanical material properties can be assigned to each element (Johnson, 1987; Zienkiewicz, Taylor and Zhu, 2005; Bathe, 2014).

For face simulations, quadratic tetrahedral elements are preferred over hexahedral elements due to the complex geometry and thus the computational expense of generating good quality 3D hexahedral meshes (Zhang *et al.*, 2016).

A displacement-based FEM formulation involves solving the elastic equilibrium equations – the interested reader is referred to (Keeve *et al.*, 1998; Mollemans *et al.*, 2007), describing in detail the implementation of the FEM in computer-assisted surgical planning systems. In brief, displacements are prescribed to the surface nodes on the osteotomised bone segment, which causes internal strains in the soft tissues. In a linear-elastic scenario, Hooke's law relates strains to stress, thus the internal forces can be deducted. A system of differential equations is solved to obtain the final position where all soft tissue forces are in equilibrium, and this solution includes the displacements of all internal and surface nodes. Whilst this approach is computationally costly and real-time simulation is unattainable, computer-assisted surgical planning applications favour accuracy over computation time (Chabanas, Luboz and Payan, 2003).

From the 1990s onwards, many research groups published on FEMs in computer-assisted surgical planning (Zachow, 2015). Initial models comprised as little as 50 elements – each element covering a square inch of skin (Pieper, Rosen and Zeltzer, 1992; Pieper, Laub and Rosen, 1995). More advanced models were developed to better model large displacements (Koch *et al.*, 1996; Gladilin *et al.*, 2002), and these models were tested on a virtual patient model generated from the Visible Human Project dataset from the National Library of Medicine (Koch *et al.*, 1996; Sarti, Gori and Lamberti, 1999). Case-studies with real patient data were subsequently published, and integrated surgical planning systems were developed (Keeve *et al.*, 1998; Zachow *et al.*, 2001). Applications of FEMs other than simulation of the

postoperative face shape have included simulating facial expressions (Gladilin *et al.*, 2003) and inverse modelling where the soft tissue is manually warped to a desired shape from which the bone displacement or implant shape is deduced (Gladilin, Ivanov and Roginsky, 2004).

The main limitation of FEM in the context of computer-assisted surgical planning is its computational cost (Mollema *et al.*, 2007), which has prevented its use in commercial software where MSM-based real-time simulations have been the standard. Therefore, it remains a tool that requires specialist engineering knowledge to set up and run simulations. Despite these limitations, FEM remains an active field of research due to its high accuracy and recently, finite element methodologies have been further developed in three main areas: mesh improvements, material properties optimisation, and modelling improvements. First, as CBCT does not capture distinct soft tissue layers, the anatomical representation of the mesh has been improved by constructing a face template from MR images which is then warped to a patient-specific anatomy (Chabanas, Luboz and Payan, 2003; Luboz *et al.*, 2005; Zhang *et al.*, 2016). Specifically, the face template model incorporated distinct soft tissue layers such as skin, muscle and fat, and the interaction between those layers to better mimic the true biomechanical behaviour of the face (Barbarino *et al.*, 2009). Whilst complex registration and transformation algorithms exist to warp this template to a patient-specific shape (Luboz *et al.*, 2005), the added steps lead increase complexity and the processing time, and residual error between the original shape and the warped template cancelled out the benefits of using mesh templates. For these reasons, they have not been implemented in commercial software (Chabanas, Luboz and Payan, 2003). The second area for improvement in the FEM relates to the variation in inter- and intra-patient material properties. To assign the correct material properties to a complex anatomical mesh, patient-specific parameters are required. Non-invasive measurements using an aspiration

device revealed relatively homogeneous mechanical responses in superficial regions of the face and a large variability in the deeper skin tissues (Barbarino, Jabareen and Mazza, 2011; Payan, 2017). This suggested that, for each patient, a unique set of material properties should be measured to optimise the simulation accuracy which is clinically challenging. Third, improvements to the FEM formulation have been made including non-linear models, however, no significant improvements in accuracy were found over a linear formulation (Mollema *et al.*, 2007).

In addition to FEM, other techniques for the numerical solution of differential equations include the finite difference method (FDM) and finite volume method (FVM) (Peiró and Sherwin, 2005). In computer assisted surgical simulation applications, the most popular of these is the FDM which has been implemented into the commercial software package Simplant (Sarti, Gori and Lamberti, 1999; Sarti *et al.*, 2007).

Despite these methodological advances, clinically, surgical simulation remains challenging: patients are either warned about the shortcomings in the prediction during clinical consultations (Liebregts *et al.*, 2015) or the surgeon chooses not to use 3D soft tissue simulation. This is partly because all of the above discussed models are deterministic, which means that an identical output is generated for each recurrent model computation. A probabilistic approach is presented in chapter 4 to overcome these challenges.

2.4 Statistical models

The statistical analysis of shape – the geometrical information when translation, rotation and scale are removed – is used to extract information from a set of shapes by

computing the modes of shape variation (Stegmann and Delgado, 2002; Madsen *et al.*, 2018). In the context of craniomaxillofacial surgery, statistical shape models have been proposed to streamline and automate processes in computer-assisted surgical planning (Zachow, 2015), thereby making this a more accessible technology. However, accurate statistical modelling of faces is a challenging task due to the large variation in the human population, and to build a statistical model that truthfully can represent each given face, a large collection of high-quality 3D images is required from a population diverse in age, gender, and ethnicity. Moreover, state-of-the-art computer vision and machine learning algorithms are required that automatically process these 3D images and construct a high-dimensional statistical model of face shape.

In the era of evidence-based medicine, vast quantities of patient data are collected (Kanevsky *et al.*, 2016), thus providing a great opportunity for the development of machine-learning-based methods for use in clinical decision-making and to enable automated personalised medicine approaches (Bennett and Hauser, 2013; Mirnezami and Ahmed, 2018). Although the application of shape analysis in plastic and reconstructive surgery is not new – it has been used to elucidate how syndromes affect skull growth (Staal *et al.*, 2015; Kaya *et al.*, 2018; Maas *et al.*, 2018), to quantify (Crombag *et al.*, 2014; Rodriguez-Florez *et al.*, 2017) or predict (Pluijmers *et al.*, 2012; Nikkhah *et al.*, 2013) the corrective effect of surgical techniques on skull deformities, and for outcome evaluation (Meulstee *et al.*, 2015) – its clinical usefulness has been limited due to the low number of faces in these models, absence of automated processing methods, and use of outdated mathematical models.

A popular machine learning approach, originally used to reconstruct accurate and complete 3D representations from single 2D images and for photo-realistic manipulation (Blanz and Vetter, 1999), involves 3DMM – statistical models of face shape and texture.

Current applications of 3DMM include facial recognition (Blanz, 2006), expression normalisation (Amberg, Knothe and Vetter, 2008), and face reconstruction from videos (Kittler *et al.*, 2016). Moreover, databases with a large number of normal faces have been described (Paysan *et al.*, 2009; Booth *et al.*, 2016; Huber *et al.*, 2016; Dai *et al.*, 2017) but no large-scale clinical 3DMMs exist. In chapter 6, a large-scale orthognathic 3DMM is introduced.

2.5 Summary

The main concepts in computer-assisted surgical simulation were introduced in this chapter. This included three computational methodologies; empirical, physical, and statistical models. The first originate from landmark-based cephalometric tools, they are straightforward in their use and modelling formulation, and therefore allow for real-time modelling. However, empirical models have a sparse architecture and the accuracy of the 3D simulation is limited. Physical models included MSMs, MTMs, FEMs, and FDMs. The MSM comprises a computationally efficient system of interlinked springs and masses allowing for real-time modelling. However, the spring-mass system lacks direct biomechanical correspondence and, therefore, the accuracy, although better than that of empirical models, is less than that for FEMs. The MTM was introduced to overcome these limitations but remaining difficulties prevented implementation in commercial software. The FEM involves mathematically defined subdivisions of a continuum problem. FEMs are highly accurate and provide simulations that closely match real soft tissue displacements, but they are computationally costly and do not allow for real-time modelling. Improvements to the finite element methods included the development of anatomically accurate meshes, the measurement of patient-specific material properties, and the development of novel physical models. One such example included FDM

which use a fast approximation scheme, and has been successfully implemented in a commercial software. Empirical and physical models have been implemented into commercial computer-assisted surgical planning software, but their clinical utility remains limited due to a number of factors including the deterministic nature of the computational models, the complexity of the software, and the time and cost investment needed. Statistical models, which infer meaningful information from a large collection of face shapes have been extensively described in the literature. Although they have been used in craniomaxillofacial surgery, their clinical application has been limited due to low numbers of samples and they have not been applied to surgical planning. One promising method involves 3DMMs, statistical models of face shape and texture that have been used in various computer vision applications.

In the next chapter, four different image acquisition modalities and their suitability for craniomaxillofacial imaging are investigated, focussing on 3D shape analysis and the use in surgical planning. In the chapters thereafter, a novel physical and a statistical model are introduced that aim to overcome the limitations described in this chapter.

Chapter 3 3D CRANIO-MAXILLO-FACIAL IMAGING

Part of the work described in this chapter has been published in the *Journal of Plastic, Reconstructive & Aesthetic Surgery*:

- Knoops PGM, Beaumont CAA, Borghi A, Rodriguez-Florez N, Breakey RWF, Rodgers W, Angullia F, Jeelani NUO, Schievano S, Dunaway DJ. Comparison of three-dimensional scanner systems for craniomaxillofacial imaging. *J Plast Reconstr Aesthet Surg* 2017;70(4):441-9.

Rights from Elsevier for publication automatically granted under author permissions.

As explained in chapters 1 and 2, computer-assisted surgical planning tools require high-quality facial images that accurately represent the individual face shape. At GOSH, various imaging modalities are available and include 3D surface scanning systems as well as tomographic scanners such as CBCT and MRI, and this chapter investigates and compares these imaging modalities for face scanning.

3.1 Introduction

In craniomaxillofacial surgery, 2D digitised lateral cephalic radiograms and traditional instruments, such as callipers and measurement tapes, have long been the standard for clinical assessments (Farkas, 1996; Honrado and Larrabee, 2004; Lekakis *et al.*, 2016). Parameters such as head circumference and the cephalic index are used to assess head and face proportion and deformity, to evaluate growth, or to quantify surgical outcomes (Edler *et al.*, 2010). Whilst these methods are relatively simple and cost-effective, they are time-consuming, prone to error, and 2D images or measurements lack appropriate 3D facial depth and shape (Da Silveira *et al.*, 2003; Schaaf *et al.*, 2010). Therefore, face shape analysis with 3D surface scans has recently gained popularity (Hammond, 2007; Van Loon *et al.*, 2010; Krimmel *et al.*, 2011; Van Loon *et al.*, 2015; Tenhagen *et al.*, 2016). Reported advantages of 3D surface scans include high accuracy and repeatability, quick acquisition, non-ionising and non-invasiveness, the ability to rotate and view a 3D scan from all angles, the ability to track 3D changes longitudinally, 3D video-analysis, and improved surgeon and patient satisfaction (Giovanoli *et al.*, 2003; Honrado and Larrabee, 2004; Lekakis *et al.*, 2016). The main disadvantage is the cost; a camera and computer equipment need to be purchased, a designated room is necessary for large static systems, and trained personnel is required for image acquisition and processing (Lee, 2004;

Heike *et al.*, 2009, 2010; Tzou *et al.*, 2014). Recently, novel hand-held scanning systems have entered the market that are easy to use, and cheap to purchase and operate (Occipital Inc., 2016).

Various models of such 3D surface scanners are available, each with specific advantages and disadvantages. In optical scanning, two main categories exist: active and passive scanners (Mada *et al.*, 2003). Active scanners project a pattern of visible or infrared light on a surface and infer the 3D shape from the distortion of the projected pattern using a camera (Geng, 1996; Mada *et al.*, 2003); the triangulation to infer a 3D shape occurs between the light source, the object, and the camera. On the other hand, passive scanners, including stereophotogrammetry scanners, pick up reflections from ambient light sources and compute a 3D shape from photographs taken with two or more cameras at different angles (Burke *et al.*, 1983; Lekakis *et al.*, 2016) – the triangulation occurs between the object and two camera, similar to how the human eye perceived depth (Beltran and Basañez, 2014). Stereo 3D matching algorithms are used to reconstruct a 3D image from two or more 2D photos, where corresponding points from each camera are identified using the image texture (Beltran and Basañez, 2014). In addition to optical 3D surface scanners, transmissive tomographic scanners can also be used to reconstruct 3D shapes from a stack of 2D slices, for example from CT or MRI (Papadopoulos *et al.*, 2002; Staal *et al.*, 2015). The purpose of this chapter is to describe how these different 3D scanners compare to each other in terms of accuracy and repeatability, and the analysis includes passive and active systems, static and hand-held technologies, and systems of high and low cost.

3.2 Methodology

3.2.1 Data acquisition and processing

Eight adult healthy volunteers (4 female; age 31 ± 4 years, range 24–37 years) participated in this prospective study. Four different scanners were employed for the 3D data acquisition: MR Avanto Scanner (Siemens Healthcare, Erlangen, German), 3dMDface System (3dMD Inc., Atlanta, GA, USA), M4D Scan (Rodin4D, Pessac, France), and Structure Sensor (Occipital Inc., San Francisco, CA, USA). Magnetic resonance and 3dMDface System stereophotogrammetry images were acquired by experienced operators, whilst the surface scans taken with the M4D Scan and Structure Sensor were obtained by an experienced researcher. The process for image acquisition and 3D reconstruction for each scanner was as follows (Table 3.1, Figure 3.1):

- 1.5T clinical MR Avanto scanner (Siemens Healthcare) – A standard 3D head T1-weighted Fast Low Angle Shot (FLASH) sequence with 1 mm slice thickness was used to obtain cross-sectional images in the MR scanner with the volunteer in a supine head and body position. Data were exported as DICOM files and 3D reconstructions were obtained using Mimics (Materialise) through one threshold operation (lower threshold of 80, upper threshold maximum value), followed by a volumetric reconstruction and a wrap, and then saved as stereolithography (STL) files.
- 3dMDface System (3dMD Inc.) – The static camera system uses hybrid active and passive stereophotogrammetry and structured light and comprises two modules with three cameras per module and a flash system (Heike *et al.*, 2009). A slightly tilted

backwards head position was adopted in order to capture the full face and chin area. A Macbook Pro (Apple Inc., Cupertino, CA, USA) was connected to the cameras and 3dMDpatient software was used to automatically compute 3D reconstructions as wavefront object (OBJ) files with tagged image file format (TIF) texture maps.

- M4D Scan (Rodin4D) – The hand-held scanner uses structured light emitted from white LEDs. Data was acquired with a still and neutral head position, i.e. a Frankfurt horizontal line, and the operator moved the scanner around the volunteer. 3D reconstructions were automatically generated by dedicated software Vxelements 2.0 (Creaform Inc., Quebec, Canada) and exported as STL files.
- Structure Sensor (Occipital Inc.) – The system comprises a structured light sensor, infrared LEDs and a normal camera, and it was connected to an iPad (4th generation, Apple Inc., Cupertino, CA, USA). The acquisition was performed with a still and neutral head position, and the operator moved around the volunteer. The *Scanner* application (Occipital Inc.) was used to export the automatically generated 3D reconstructions as OBJ files.

Table 3.1 Characteristics of 3D scanners.

| | Scanner | | | |
|--|------------------------------------|--|--|--|
| | Avanto MRI | 3dMDface System | M4D Scan | Structure Sensor |
| Hardware | 1 integrated full-body MRI scanner | 2 modules with 3 cameras per module; flash system; stand; computer | 1 hand-held scanner with 2 cameras, 4 white light LEDs; Computer | 1 module, i.e. iPad accessory, with 1 camera, 1 infrared sensor, 2 infrared LEDs; iPad |
| Modality | Magnetic resonance | Hybrid passive/active: stereophotogrammetry/structured light | Active: structured light (white light) | Active: structured light (infrared) |
| Accuracy [†] | 1 mm slices [‡] | 0.2 mm | 0.5 mm | 4 mm |
| Acquisition time | 360 s [‡] | 1.5 ms | ~ 30 s | ~ 20 s |
| Output files | DICOM | Textured mesh (OBJ) | Mesh (STL) | Textured mesh (OBJ) |
| Mesh density (polygons/mm ²) | 0.51 ± 0.11 | 1.10 ± 0.08 | 0.44 ± 0.01 | 0.13 ± 0.01 |
| Hand-held | No | No | Yes | Yes |
| Cost (USD)* | >250,000 | >20,000 | >15,000 | 1,000 |

[†] Manufacturers' stated accuracy varies with object distance. M4D Scan: 0.5 mm at 40 cm stand-off distance; Structure Sensor: 4 mm at 60 cm stand-off distance.

[‡] Depends on the chosen acquisition sequence and parameters.

* As of December 2015 – exact cost depends on configuration (modules, computer/iPad, software, accessories, etc.).

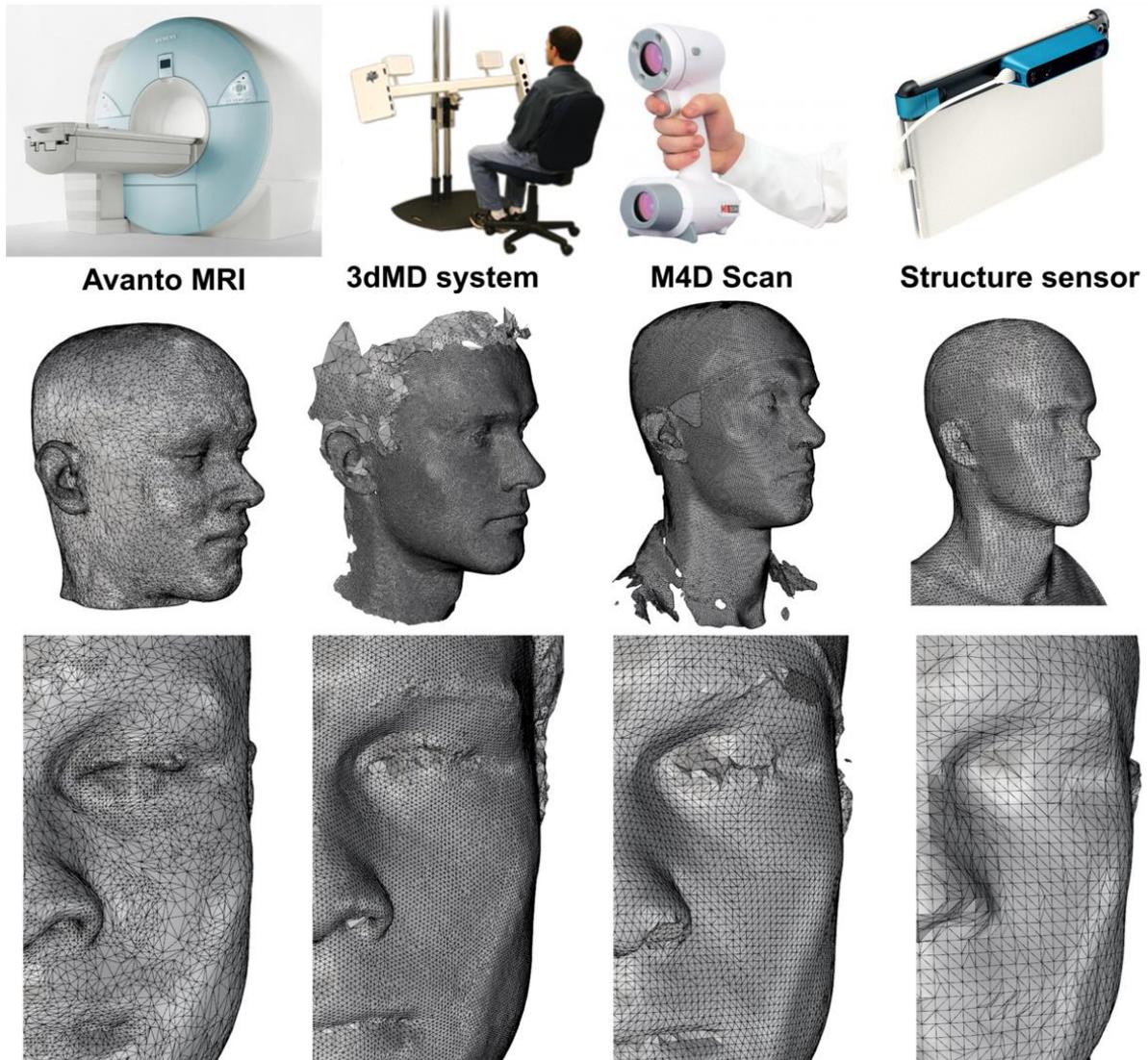


Figure 3.1 3D scanners, mesh overview, and mesh detail. Mesh representation of the data obtained using the Avanto MRI, 3dMDface System, M4D Scan, and Structure Sensor. Note: MRI does not capture hair and 3D surface scanners have great difficulty with hair too; therefore, a cap is often used as shown in the M4D scan mesh.

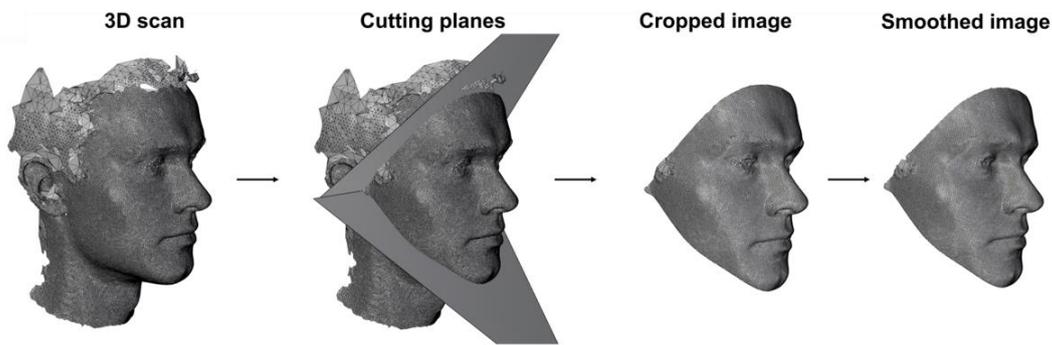


Figure 3.2 3D data processing steps. Following ICP registration, 3D scans were imported into Rhinoceros: a cutting plane was defined between the left and right tragus and the chin, and a second plane was constructed orthogonal to the first; these planes were used to crop the scan; and a smooth filter was applied.

The process for elaboration and analysis of the 3D reconstructions was the same for all imaging modalities. OBJ and STL surface scans were loaded into 3-Matic (Materialise) for each volunteer. Scans were aligned using global and N -point iterative closest-point (ICP) registration. In Rhinoceros (Robert McNeel & Associates, Seattle, WA, USA) a plane was created based on the chin and the left and right tragus, and another plane orthogonal to the first plane, on the line between the left and right tragus (Figure 3.2). The two planes were used to crop the face (Figure 3.2) and the remaining area was used to calculate mesh area, size and density (Table 3.1). Lastly, the STL files were imported into Meshmixer (Autodesk, Inc., San Rafael, CA, USA), where any voids were filled – for example in the eye region – and surfaces were minimally smoothed using a shape preserving setting to deal with artefacts and noise (Figure 3.2).

3.2.2 Data analysis and statistics

Closest-point distance vectors between scan pairs were computed using VMTK (Antiga *et al.*, 2008) (The Vascular Modelling Toolkit, Bergamo, Italy) and Matlab (The MathWorks,

Natick, MA, USA), and were visualised in ParaView (Ahrens, Geveci and Law, 2005) (Kitware, Clifton Park, NY, USA). Data analysis and statistical analysis were carried out in R (v. 3.3.0, R Foundation for Statistical Computing, Vienna, Austria).

Accuracy of the camera systems was determined by the ability of the camera to capture the facial shape in comparison to a reference shape (Table 3.2, study 1), according to the following ranges (Aung, Ngim and Lee, 1995): 0-1 mm (highly reliable), 1-1.5 mm (reliable), 1.5-2 mm (moderately reliable), and >2 mm (unreliable). The 3dMDface System was chosen as a reference shape because of its low operator dependence, low scanning time, high accuracy, and high precision (Wong *et al.*, 2008; Heike *et al.*, 2009, 2010; Lübbers *et al.*, 2010). The root mean square distance (RMS) was computed as the closest-point difference between two surfaces:

$$RMS = \sqrt{\frac{\sum_{i=1}^n (D_i)^2}{n}} \quad (3.1)$$

where D is the closest-point difference for each point-pair calculated using VMTK and n is the number of point-pairs.

Repeatability of the M4D Scan and Structure scan was determined by acquiring and analysing six scans for one participant, taken with the same scanner. First, one scan was taken as reference and compared to the remaining scans, and second, a different scan was taken as a reference and compared to the remaining scans, for a total of 10 pairwise comparisons per scanner (Table 3.2, study 2). Average RMS and standard deviation (SD) were computed for all 10 pairs. Furthermore, to quantify the post-processing error induced by the steps as laid out in the previous paragraph and in Figure 3.2, the dataset of one participant was analysed five times

successively (Table 3.2, study 3). Moreover, the accuracy of the landmark digitisation on the chin and left and right tragus was quantified to investigate systematic and random errors. 3dMD scans from all eight patients were landmarked twice and the Euclidean distances between the repeated digitisation of the same landmark was calculated.

Usability was assessed qualitatively by evaluating user-friendliness of the software and based on operators' and participants' experiences. The Mann-Whitney U test was employed for comparison of RMS, and p-values <0.05 were assumed to be statistically significant. Mean \pm SD based on all 8 datasets is given unless stated otherwise.

Table 3.2 Overview the three studies and the datasets involved.

| Study | Subject | Datasets |
|--------------|-----------------------|---|
| 1 | Accuracy | 32 (8 participants, 4 scanners, 1 scan per scanner) |
| 2 | Repeatability | 12 (1 participant, 2 scanners, 6 scans per scanner) |
| 3 | Post-processing error | 4 (1 participant, 4 scanners, 1 scan per scanner) |

3.3 Results

3.3.1 Scanner overview

An overview was given of each scanner's properties, including their imaging modality, mean mesh density in the facial area, and mean acquisition time (Table 3.1). Acquisition time was lowest for the 3dMDface System (1.5 ms; manufacturer data), followed by the Structure Sensor (20 s), M4D Scan (30 s), and Avanto MRI (300 s; dependent on acquisition sequence). In addition, mesh density was highest for the 3dMDface System, followed by the Avanto MRI, M4D Scan, and Structure Sensor.

3.3.2 Accuracy

The colourmap for participant 6, representative for the cohort, demonstrated clear differences between the scanning systems (Figure 3.3). For Avanto MRI deviations were visible in the jaw, cheek, and eyes. The M4D Scan showed minimal deviation in the eyes and around the mouth. The Structure Sensor showed moderate or unreliable agreement overall, whilst the Avanto MRI and M4D scan showed good concordance in the nose, forehead, and chin area.

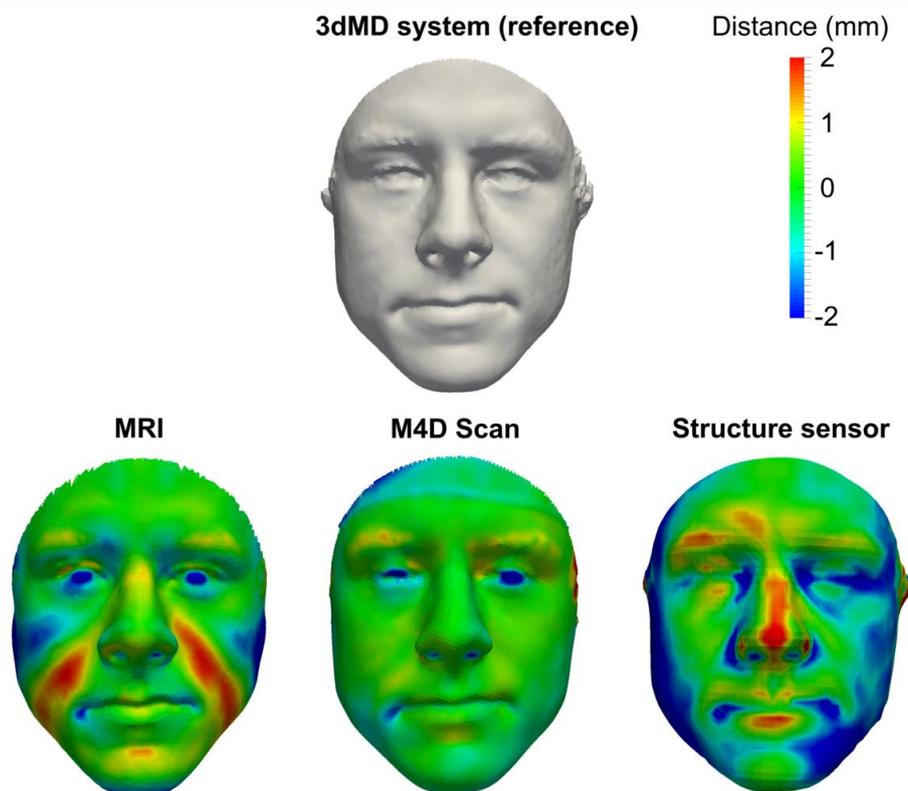


Figure 3.3 Facial colourmap showing the agreement with the reference. The 3dMDface System scan (grey, top) is shown, which is the reference image for the other three scans. Green indicates highly reliable areas, whilst red (underestimation) and blue (overestimation) indicate unreliable areas. Participant 6 is shown which was representative for the cohort of participants.

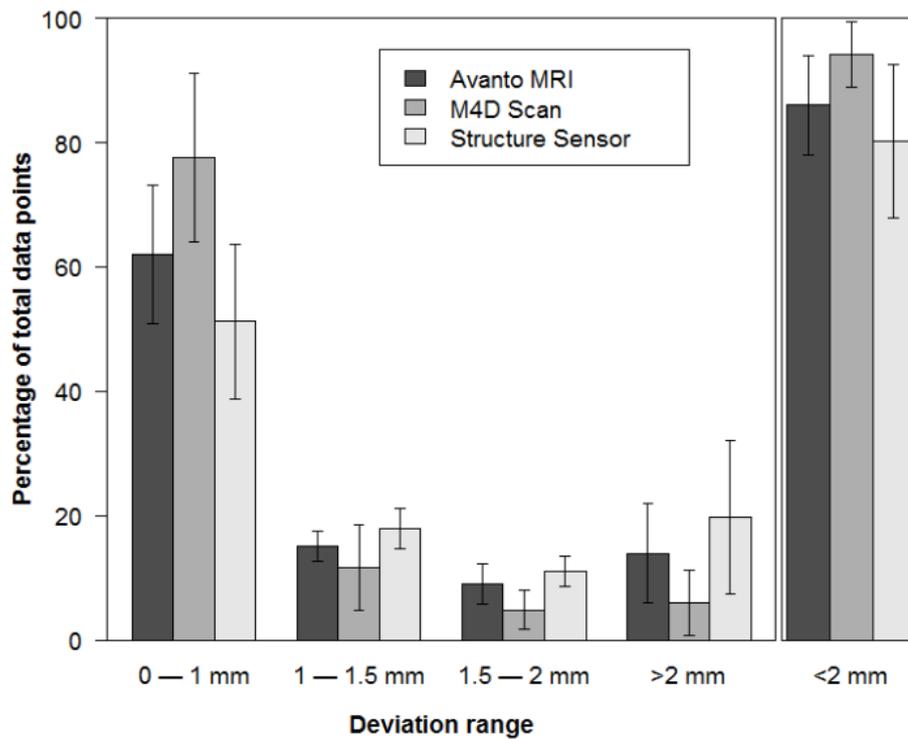


Figure 3.4 Percentage of data points within deviation ranges. The error bars indicate the accuracy of the Avanto MRI (dark grey), M4D Scan (grey), and Structure Sensor (light grey), relative to the 3dMDface System, within various ranges: 0–1 mm (highly reliable), 1–1.5 mm (reliable), 1.5–2 mm (moderately reliable), and >2 mm (unreliable). Mean and standard deviation were computed from all scans (n=8).

The M4D Scan (RMS = 0.71 mm \pm 0.28 mm) was significantly better than the Avanto MRI (RMS = 1.11 mm \pm 0.33 mm, $p = 0.008$) and Structure Sensor (RMS = 1.33 mm \pm 0.46 mm, $p = 0.008$) (Table 3.3), and there was no significant difference between the Avanto MRI and Structure Sensor ($p = 0.15$). The Avanto MRI, M4D Scan, and Structure Sensor, respectively, showed an accuracy <2 mm of 85%, 94%, and 80% (Figure 3.4).

3.3.3 Repeatability

Repeatability of the M4D Scan and Structure Sensor was determined (Table 3.4). Mean and standard deviation were 0.51 mm \pm 0.04 mm and 0.51 mm \pm 0.03 mm, respectively, and no significant difference between the two scanners was found ($p = 0.80$).

Table 3.3 Accuracy of Avanto MRI, M4D Scan, and Structure Sensor. Relative to the 3dMDface System, measured as the root mean square distance (RMS) and mean and standard deviation (SD).

| Participant | RMS (mm), relative to 3dMDface System | | |
|---------------|---------------------------------------|-----------------|------------------|
| | Avanto MRI | M4D Scan | Structure Sensor |
| 1 | 1.76 | 1.05 | 2.28 |
| 2 | 1.16 | 0.65 | 1.13 |
| 3 | 1.25 | 1.02 | 1.15 |
| 4 | 1.09 | 1.05 | 1.50 |
| 5 | 1.16 | 0.55 | 1.40 |
| 6 | 0.67 | 0.47 | 1.45 |
| 7 | 1.00 | 0.44 | 0.83 |
| 8 | 0.75 | 0.45 | 0.86 |
| Mean \pm SD | 1.11 \pm 0.33 | 0.71 \pm 0.28 | 1.33 \pm 0.46 |

Table 3.4 Repeatability of M4D Scan and Structure Sensor. Root mean square distance (RMS) mean and standard deviation (SD) are shown for 10 scan pairs from the same participant.

| Scan pair | RMS (mm) | |
|---------------|-----------------|------------------|
| | M4D Scan | Structure Sensor |
| 1 vs 6 | 0.55 | 0.49 |
| 2 vs 6 | 0.50 | 0.54 |
| 3 vs 6 | 0.51 | 0.53 |
| 4 vs 6 | 0.46 | 0.48 |
| 5 vs 6 | 0.48 | 0.48 |
| 1 vs 2 | 0.56 | 0.56 |
| 1 vs 3 | 0.49 | 0.50 |
| 1 vs 4 | 0.56 | 0.50 |
| 1 vs 5 | 0.47 | 0.50 |
| 1 vs 6 | 0.55 | 0.49 |
| Mean \pm SD | 0.51 \pm 0.04 | 0.51 \pm 0.03 |

3.3.4 Post-processing and landmark digitisation error

The post-processing error, comprising steps as outlined above (Figure 3.2) had a standard deviation of 0.04, 0.03, and 0.06 mm, for the Avanto MRI, M4D Scan, and Structure Sensor respectively (Table 3.5). To understand the influence of the errors in the landmark digitisation (Figure 3.2, step 2) on the overall post-processing error, the landmark error on the chin, left tragus and right tragus was determined (Table 3.6).

Table 3.5 Post-processing error. One dataset was analysed 5 times (i.e. steps 2 – 4 in Figure 3.2) for each of the three scanners, relative to the 3dMDface System. Root mean square distance (RMS) and standard deviation (SD) are shown.

| Repetition | RMS (mm), relative to 3dMDface | | |
|---------------|--------------------------------|-----------------|------------------|
| | Avanto MRI | M4D scan | Structure Sensor |
| 1 | 1.00 | 0.44 | 0.83 |
| 2 | 0.99 | 0.49 | 0.75 |
| 3 | 1.06 | 0.47 | 0.79 |
| 4 | 0.99 | 0.44 | 0.78 |
| 5 | 1.06 | 0.41 | 0.67 |
| Mean \pm SD | 1.02 \pm 0.04 | 0.45 \pm 0.03 | 0.76 \pm 0.06 |

Table 3.6 Landmark digitisation error. Mean and standard deviation (SD) Euclidean distance of the three landmarks used to define the area of the face for surface comparison based on measurements on all eight subjects.

| Landmark | Mean (mm) | SD (mm) | 95% confidence interval | |
|-----------------|-----------|---------|-------------------------|-------|
| | | | Lower | Upper |
| Chin (gnathion) | 1.25 | 0.17 | 1.13 | 1.37 |
| Left tragus | 1.29 | 0.18 | 1.16 | 1.41 |
| Right Tragus | 1.09 | 0.17 | 0.97 | 1.20 |
| Mean | 1.21 | 0.17 | 1.09 | 1.33 |

3.4 Discussion

In orthognathic surgery, 3D shape analysis has been extensively used to assess surgical outcomes objectively (Da Silveira *et al.*, 2003; Honrado and Larrabee, 2004; Hammond, 2007; Van Loon *et al.*, 2010, 2015; Krimmel *et al.*, 2011; Staal *et al.*, 2015; Tenhagen *et al.*, 2016). In this chapter, the accuracy and repeatability of four 3D scanning systems to capture the face shape was assessed. The 3dMDface System was chosen as gold-standard for comparison with other scanners, as previous studies have shown accuracy of this system to be within 1 mm when compared with conventional anthropomorphic measurements (Wong *et al.*, 2008). The Avanto MRI and the M4D Scan demonstrated similar levels of accuracy, although the high cost of MRI may prevent routine surface scanning, contrary to the other three more affordable surface scanners.

RMS was computed as a measure of overall accuracy relative to the 3dMDface System. The RMS was found to be lowest for the M4D Scan and significantly better than that of the Avanto MRI and Structure Sensor. Clinically, deviations larger than 2 mm are considered unreliable (Aung, Ngim and Lee, 1995). All systems showed large percentages of data points within the reliable range: 85%, 94% and 80% respectively for the Avanto MR, M4D Scan and Structure sensor. However, the usefulness of using an average measure can be limited for the assessment specific landmarks (Kouchi *et al.*, 2012) and the clinical usability of some 3D scanners may be limited due to a lack of local accuracy, even when overall RMS is satisfactory. Moreover, an accuracy measure that is relative to the maxillary movement could be used to better compare between patients with large or minor anterior displacement – an accuracy range

of 2 mm might be clinically acceptable for patients with 10 mm maxillary advancement but less so for patients with less than 5 mm advancement.

In this study, qualitative analysis using colourmaps showed good agreement for the Avanto MRI and M4D scan, but limited agreement for the Structure Sensor. Repeatability for the M4D Scan and Structure Sensor, expressed in RMS, were found to be 0.50 mm for both systems. Thus, whilst the Structure Sensor was as precise as the M4D scan, it was less accurate, which is supported by the colourmaps that revealed relatively large deviations in areas with high curvatures. Scans acquired with the Structure Sensor, due to the automatic post-processing and the limited accuracy, did not truthfully define high curvature areas such as the nose. However, the results suggested it may be used to describe less complex areas such as head shape, cheek and chin contour (Beaumont *et al.*, 2017).

Factors that influence scan quality are lighting, scanner alignment and placement, facial expression of the subject, adequate coverage of hair, the examiner, and software post-processing (Kovacs *et al.*, 2006; Heike *et al.*, 2009). A limitation of this study is the use of different head positions. A supine position in the MRI scanner, in contrast to a neutral head position, induces some deviations in the jaw and cheek regions as seen in the colourmaps. These differences are likely due to the effects of gravity on deformable soft tissues of the face and reflect the fact that the facial form is different in supine and upright positions.

In addition to the factors mentioned above, clinical considerations including patient compliance and scanner mobility are of importance, in particular in the context of paediatric imaging (Heike *et al.*, 2010). An advantage of the static 3dMDface System is its low acquisition time of 1.5 ms, thereby minimising motion artefacts and reducing the need for patient

compliance (Wong *et al.*, 2008). However, hand-held systems bear the advantage that they can be used in wards, operating theatres, and in outpatient clinics, contrary to 3dMDface System and other static systems.

Moreover, the post-processing error and landmark error are important considerations. Landmark errors below 0.5 mm have been previously reported (Almukhtar *et al.*, 2017), which is lower than the values found in this chapter. This is explained by the fact that landmarks were placed on mesh nodes which have an inter-node distance of approximately 1 mm in unprocessed 3dMD scans. This landmark digitisation could be improved by resampling the mesh to have a greater number of nodes, by allowing placement on vertices in between nodes, or by labelling landmarks on the subject face before image acquisition (Aynechi *et al.*, 2011).

Despite the Structure Sensor presenting the lowest accuracy, it demonstrated major advantages over the other scanners. The user-friendliness of the *Structure* app and the portability of the iPad provided a pleasant user experience. Moreover, it comes with an open source software development kit, which allows for custom-made software and as a result, multiple applications are available. Occipital's own software was used in this study, but purposely built apps for craniomaxillofacial assessment may improve accuracy. Another advantage is the use of infrared LED is because it does not disturb the subject in the way white light does. Based on these findings it is foreseeable that tablet-based hand-held systems play a more prominent role in future craniomaxillofacial 3D scanning.

In addition to the scanners included in this study, there are numerous other models on the market. The surface scanners used in this study represent those available at GOSH and of various cost, portability and quality. Amongst others, companies that produce 3D scanners

include 3dMD, Axisthree (Belfast, Ireland), Canfield Scientific (Fairfield, NJ, USA), Crisalix 3D (Bern, Switzerland), and Di3D (Glasgow, UK). A review of high-end static scanning systems can be found in literature (Tzou *et al.*, 2014). Furthermore, a recent study with 41 volunteers on the accuracy of Artec EVA (Artec Group, Luxembourg) and FaceScan3D (3D-Shape, Erlangen, Germany), found mean errors of a phantom between 0.23-0.24 mm and 0.52-0.63 mm for the handheld and static system respectively (Patete *et al.*, 2013).

3.5 Summary

Many different 3D scanning systems are on the market for a wide range of applications. In the context of craniomaxillofacial imaging, four scanning systems at GOSH were investigated including static and hand-held systems, and systems of high and low cost. This prospective study on eight volunteers demonstrated that the M4D Scan showed significantly best RMS, better than the Avanto MRI and Structure Sensor, although all systems showed results that were in a clinically acceptable range. For Avanto MRI, deviations occurred from a different head position (supine versus neutral), suboptimal slice thickness, and the inability to capture facial hair. The Structure Sensor lacks hardware and software to accurately characterise areas with complex shape and high curvature but is good at describing general facial form. Nevertheless, it still shows fair agreement with systems more than tenfold its cost and portability, and the direct visualisation on an iPad showed great promise for clinical use. Moreover, it may be used longitudinally for measures such as head circumference and cephalic index (Beaumont *et al.*, 2017).

These results suggest that images from different high-quality image sources may be combined into a single dataset, although it is essential to standardise the acquisition procedure including head position, lighting, facial expression, and hair coverage. Additionally, appropriate balancing the technical requirements and clinical needs will determine which scanner suits a specific application. In the next chapter a novel physical modelling framework is discussed, which uses patient clinical imaging data as input.

Chapter 4 **PROBABILISTIC FINITE ELEMENT MODELLING FOR SURGICAL SIMULATION**

Part of the work described in this chapter has been published in *PLoS ONE*:

- [Knoops PGM](#), Borghi A, Ruggiero F, Badiali G, Bianchi A, Marchetti C, Rodriguez-Florez N, Breakey RWF, Jeelani NUO, Dunaway DJ, Schievano S. A novel soft tissue prediction methodology for orthognathic surgery based on probabilistic finite element modelling *PLoS ONE* 2017;70(4):441-449.

Published under CC BY licence allows anyone to copy, distribute, or reuse the content as long as the author and original article are cited.

Computer-assisted surgical planning, after medical imaging acquisition, requires the following steps: to construct a virtual patient model, to perform virtual surgery, and to compute the soft tissue response from changes to the underlying skeleton. As discussed in chapter 2, physical models have been frequently used in computer-assisted surgical planning; however, assumptions in material properties and simplifications in the physical model may lead to inaccuracies in the simulated surgical results. This chapter describes a novel physical modelling approach to overcome the limitation of current models.

4.1 Introduction

In surgical planning, the simulation of the postoperative facial appearance can be computed as the response of the soft tissues from changes to the underlying skeleton (Koch *et al.*, 1996; Xia *et al.*, 2001; Swennen, Mollemans and Schutyser, 2009). In order to calculate these changes, various mathematical models can be employed. As described in chapter 2, popular approaches include empirical models and physical models. Despite advances in the modelling formulation and computational power, the accuracy and reliability of such software remain controversial and no consensus has been reached in the craniomaxillofacial community, likely due to the deterministic nature of these computational predictions (Xia *et al.*, 2007; Kretschmer *et al.*, 2009; Marchetti *et al.*, 2011; Aboul-Hosn Centenero and Hernández-Alfaro, 2012; Terzic, Combescure and Scolozzi, 2014; Janakiraman *et al.*, 2015; Peterman *et al.*, 2016; Zhang *et al.*, 2016; Resnick *et al.*, 2017). Moreover, assumptions and simplifications of the material properties may negatively impact the accuracy of these soft tissue predictions, as well as a simplified representation of the mesh that does not include distinct muscles and other soft tissue constituents. Another important consideration is the mismatch between preoperative

planning and actual surgical outcomes in terms of location of osteotomies and amount of repositioning. In Le Fort I surgery, a mismatch of 1.0 mm has been reported and in bimaxillary procedures of 1.2 mm (Badiali *et al.*, 2015; Baan *et al.*, 2016).

An alternative to single deterministic solutions is a probabilistic approach which can be used to obtain a range of possible solutions (Pataky, Koseki and Cox, 2016). PFEM have been already adopted in many biomedical engineering applications (Stefanou, 2009; Mangado *et al.*, 2016); for example, a probabilistic model was developed to investigate the effects of uncertainty and variability of material properties on stress and strain in a primate skull model, which showed that high non-homogeneity, anisotropy, and material property randomness give large variability in strains and low variability in stresses (Berthaume *et al.*, 2012). Soft tissues and facial appearance, however, were not assessed. Therefore, in this chapter, a PFEM approach for soft tissue prediction in orthognathic surgery is described that considers the uncertainties arising from material properties and bone repositioning mismatch, to address the difficulties associated with deterministic soft tissue prediction models.

4.2 Methodology

4.2.1 Patient population

Eight consecutive patients (3 female, mean age 24 ± 7 years, Table 4.1) who had orthognathic surgery by means of Le Fort I osteotomy and BSSO at St Orsola-Malpighi University Hospital in Bologna between October 2012 and July 2013 were retrospectively included in this study. All patients had preoperative and postoperative CBCT scans, and

Table 4.1 Details for eight patients who consecutively had orthognathic surgery.

| Patient | Age | Gender | Time (days) | | | Planned maxillary advancement (mm) |
|-----------|--------|--------|-------------------|--------------------|-----------------------------------|------------------------------------|
| | | | Preoperative CBCT | Postoperative CBCT | Preoperative – postoperative CBCT | |
| 1 | 17 | M | 56 | 101 | 157 | 5.0 |
| 2 | 19 | F | 12 | 42 | 54 | 4.5 |
| 3 | 20 | M | 24 | 80 | 104 | 3.5 |
| 4 | 27 | M | 26 | 43 | 69 | 4.5 |
| 5 | 18 | M | 18 | 94 | 112 | 5.5 |
| 6 | 35 | F | 26 | 59 | 85 | 4.5 |
| 7 | 17 | M | 13 | 64 | 77 | 5.5 |
| 8 | 32 | F | 580 | 49 | 629 | 4.0 |
| Mean ± SD | 24 ± 7 | | 94 ± 197 | 67 ± 23 | 161 ± 192 | 4.6 ± 0.7 |

bespoke cutting guides and surgical navigation (eNlite navigation system, Stryker, Freiburg, Germany) was employed to deliver the surgery (Mazzoni *et al.*, 2010, 2015). All patients provided written consent and the study was approved by the independent ethical committee of the Sant’Orsola Malpighi University Hospital, Bologna (349/2017/O/OssN).

4.2.2 Probabilistic finite element modelling

Preoperative and postoperative CBCT scans were imported and segmented in Simpleware ScanIP (Synopsis, Mountain View, USA) to generate digital patient models that included the maxilla, mandible, skull base, nasal cartilage, and soft tissue (Figure 4.1) Tetrahedral meshes were created using Simpleware ScanIP for the preoperative reconstructed 3D anatomy and imported into Ansys (v17.2, Ansys Inc, Canonsburg, USA). Linear elastic FEM simulations were set up to replicate the Le Fort I osteotomy and repositioning of the

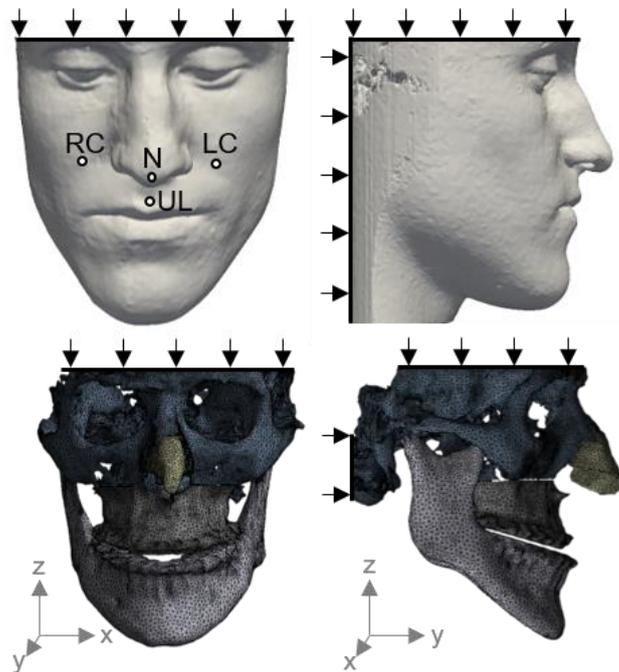


Figure 4.1 Preoperative CBCT reconstruction of bone and soft tissue. The fixed boundary conditions are indicated with arrows, and four soft tissue points for design of experiments were: nose (N), upper lip (UL), right cheek (RC), and left cheek (LC).

mobilised segment as determined from the postoperative CBCT, and the BSSO was not modelled. First, nodes were fixed in various planes (Figure 4.1): soft tissue posteriorly (y_{\min}), superiorly (z_{\max}), and inferiorly (z_{\min}); skull bone posteriorly (y_{\min}) and superiorly (z_{\max}). The mandible was fixed on the y-axis (anteriorly-posteriorly) and z-axis (superiorly-inferiorly) but set free on the x-axis (laterally-medially). Advancement of the Le Fort I segment was achieved by anterior displacement of all maxilla surface nodes. Each part of the subsequent probabilistic analysis was implemented in Ansys (Figure 4.2) (Reh *et al.*, 2006). The next paragraphs describe these steps in detail, but in brief: a correlation between input and output variables was performed, PFEM was carried out with a DOE which comprised a series of FEMs in which the value of each variable is changed upon each computation, the material properties were optimised, and a second DOE was carried out using these optimised input values.

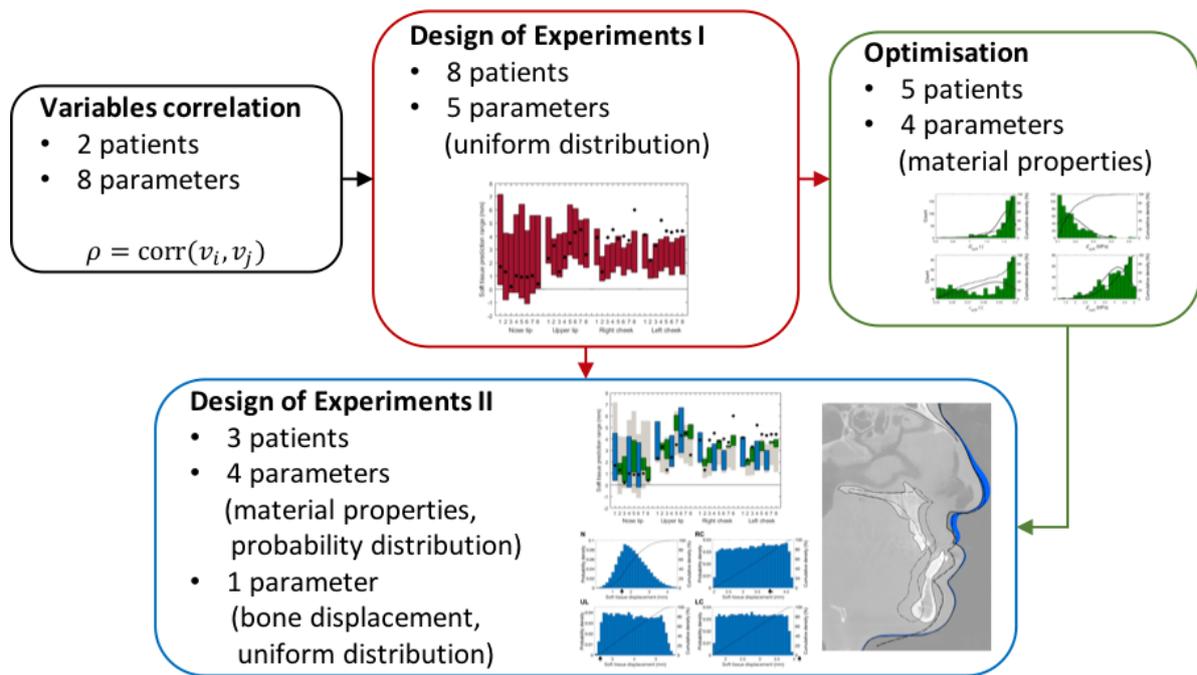


Figure 4.2 Flowchart of the methodology for the probabilistic finite element analysis with colour coding: variable correlation, design of experiments I, optimisation on material properties, and design of experiments II.

4.2.2.1 Variable correlation

A correlation analysis was carried out to assess which input variables have a significant effect on output variables in the parametric model. This correlation was performed in patient 1 and validated in patient 2. Significant variables were then implemented with a uniform range and non-significant variables were implemented using the mean value of the range. Input variables considered were: E_{BONE} and ν_{BONE} ; E_{CART} and ν_{CART} ; E_{SOFT} , ν_{SOFT} and G_{SOFT} ; and x_{BONE} , where E describes Young's Modulus, ν the Poisson's ratio, G is viscoelasticity, x is the horizontal bone repositioning, $cart$ is the abbreviation for cartilage and $soft$ for soft tissues.

All material properties were based on literature data (Table 4.2), where the minimum and maximum values reflect the range of values found in various reports. Bone material

Table 4.2 Input parameters and material properties E: Young's modulus, ν : Poisson's ratio, G: viscoelasticity (Prony series).

| | Bone | | Cartilage | | Soft tissue | | |
|---------|-----------|-----------|-----------|-----------|-------------|-----------|---------|
| | E (MPa) | ν (-) | E (MPa) | ν (-) | E (MPa) | ν (-) | G (%) |
| Minimum | 5000 | 0.2 | 0.5 | 0.26 | 0.1 | 0.45 | 31 |
| Maximum | 15000 | 0.4 | 5 | 0.38 | 1 | 0.499 | 94 |

properties were based on adult (Rupin *et al.*, 2009) and infant (Wang *et al.*, 2014) cadaver studies as well as computational models that previously described such parameters (Willinger, Kang and Diaw, 1999; Horgan and Gilchrist, 2003; Yan and Pangestu, 2011). Viscoelastic soft tissue properties were based on in vivo facial skin measurements (Barel, Lambrecht and Clarys, 2004; Jachowicz, McMullen and Prettypaul, 2007; Luboz, Promayon and Payan, 2014), in vivo human viscoelastic muscle, skin and fat properties (Then, Vogl and Silber, 2012) and a FEM with facial skin, muscle and fatty tissues (Chabanas and Payan, 2000); and cartilage properties were based on human nasal septum samples (Grellmann *et al.*, 2006; Richmon *et al.*, 2006) as well as computational studies (Lee *et al.*, 2010).

Volumetric relaxation was not implemented as the soft tissues were assumed to be incompressible due to their high water content (Gilchrist *et al.*, 2014). Normalised shear relaxation was implemented using Prony series with values from literature (Wu *et al.*, 2003, 2006):

$$G(t) = 1 - \sum_1^n g_i (1 - e^{-t/\tau_i}) \quad (4.1)$$

where $G(t)$ describes normalised shear relaxation over time, g_i the relaxation and τ_i the characteristic relaxation constant for n components, and t time. Shear relaxation (g_i) was implemented with two terms (Wu *et al.*, 2006), representing a total of 62.5% relaxation, see (4.2). Subsequently, a scaling factor (G_{soft}) ranging between 0.5 and 1.5 was applied to represent a minimum of 31% and a maximum of 94% relaxation. The characteristic relaxation components (τ_i) were assumed constant in the DOE experiment.

$$g_{1,2} = \begin{bmatrix} 0.325 \\ 0.30 \end{bmatrix}, \tau_{1,2} = \begin{bmatrix} 1 \\ 6 \end{bmatrix} \quad (4.2)$$

The ability of the surgeons to reproduce the preoperative planning was also considered and based on the literature, a mismatch in the range of ± 2 mm was implemented (Kretschmer *et al.*, 2009; Mazzoni *et al.*, 2015). The outputs for correlation were the simulated soft tissues, measured in four points selected as representative of the various areas of the face likely to be affected by orthognathic Le Fort I surgery (Figure 4.1): the nose tip (N, cephalometric equivalent: pronasale), upper lip (UL, labrale superius), and right (RC) and left cheek (LC). RC and LC are defined laterally to the nose tip and inferiorly to the centre of the eye, defined as the middle between both canthi. Spearman correlation with significance at 95% was set as inclusion criteria for the variables in the subsequent DOE approach.

4.2.2.2 Design of experiments I

DOE was used to investigate the relationship between those variables that satisfied the correlation threshold. The following steps were carried out for all eight patients. An optimal space-filling algorithm was used for input parameter generation, which is an extended Latin

Hypercube Sampling method which ensures uniform parameter distribution amongst the set range while minimising the number of simulations (Reh *et al.*, 2006).

For each patient, 27 simulations were required to compute the DOE matrix. Response surfaces based on genetic aggregation, depicting the input-output relationships, were generated from the DOE as well as sensitivity curves. The predicted ranges were compared to the postoperative CBCT scans in the four points described above: N, UL, RC and LC.

4.2.2.3 Optimisation of material properties

The optimisation process sought to match the soft tissue predictions from the PFEM to the soft tissue location in the postoperative CBCT and, thereby, reduce the corresponding material property input ranges. Out of the eight patients, five were randomly selected for optimisation (patients 2, 3, 5, 7 and 8) and three for validation with a second DOE iteration (patients 1, 4 and 6). For each of the four soft tissue points, the optimisation algorithm attempted to find the exact solution as determined by CBCT if the value lies within the predicted range, otherwise, it minimised or maximised the value accordingly.

The optimisation problem is nontrivial – due to conflicting optimisation goals no unique solution exists that satisfies all optimisation objectives (Fonseca and Fleming, 1993; Miettinen, 1998, 2008). Therefore, a multi-objective optimisation method was adopted that seeks Pareto optimal solutions. An *a priori*, goal driven multi-objective genetic algorithm was selected, with equally weighted objectives as listed in the above paragraph. Within this algorithm, based on the Ansys default settings, 1,000 synthetic solutions were generated per iteration based on the response surfaces, with a maximum of 10 iterations. Convergence stability was set at 2% with a Pareto criterion of 70%. 100 synthetic simulations were extracted that best matched the above

criteria, and these simulations represented an optimised subset within the bounds of the original DOE. Finally, the corresponding optimised input variables were extracted for the next step.

4.2.2.4 Design of experiments II

From the best 100 simulations for each of the five patients, a set of four material property inputs (those properties that significantly correlated to the outputs – see Results) was extracted which provided a 500 x 4 matrix. A Weibull distribution was fitted to each material property using Matlab and these population-specific material property variable distributions were subsequently used as inputs for the second iteration of DOE on the test set, comprising patients 1, 4 and 6. The key difference with the first DOE iteration is that each material property had a probability distribution rather than a uniform distribution. These Weibull curves and their maximum values are referred to as population-specific material properties.

4.3 Results

The variable correlation analysis for patient 1 showed that five out of eight input variables were significantly correlated to at least one of the five output variables: E_{CART} , E_{SOFT} , v_{SOFT} , G_{SOFT} , and x_{BONE} (Table 4.3). These results were confirmed for patient 2. There was a strong correlation with bone advancement (maximum correlation, $r = 0.98$) followed by the material properties including the soft tissue Young's modulus ($r = 0.50$), viscoelastic relaxation ($r = 0.37$), cartilage Young's modulus ($r = -0.32$), and cartilage Poisson's ratio ($r = 0.30$).

The first DOE iteration, using values from the literature for the material properties, was used to simulate a range of outcomes (Figure 4.3), including a minimum and a maximum

predicted outcome. From this minimum and maximum predictions, for each patient the four soft tissue output points were sampled (Figure 4.4a, red bars) and compared to the actual position from postoperative CBCT (black dots). The baseline value (0 mm) represents the preoperative CBCT. These figures illustrate that the PFEM accurately simulated the nose and upper lip – seen by the black dot to be within the red bar – whilst for the left and right cheek, the displacement predictions were underestimating the postoperative position.

Table 4.3 Variable correlation. Correlation of input variables and output parameters. E: Young’s modulus, ν : Poisson’s ratio, G: viscoelasticity (Prony series), x: advancement, soft: soft tissue, cart: cartilage, N: nose, UL: upper lip, RC: right cheek, LC: left cheek, r: Pearson’s correlation, CI: confidence interval.

| | Output | | | |
|--------------|-----------------------------------|--------------------------------|--------------------------------|--------------------------------|
| | N | UL | RC | LC |
| E_{SOFT} | r = 0.50*** CI = 0.33, 0.64 | r = -0.03 | r = 0.07 | r = 0.09 |
| ν_{SOFT} | r = 0.24* CI = 0.03, 0.43 | r = 0.18 | r = 0.16 | r = 0.03 |
| G_{SOFT} | r = -0.30** CI = -0.48, -0.10 | r = 0.21* CI = 0.00, 0.40 | r = 0.08 | r = -0.05 |
| E_{BONE} | r = -0.13 | r = 0.01 | r = -0.07 | r = 0.07 |
| ν_{BONE} | r = -0.01 | r = -0.03 | r = -0.01 | r = 0.00 |
| E_{CART} | r = -0.32*** CI = -0.49, -0.12 | r = 0.08 | r = 0.09 | r = 0.06 |
| ν_{CART} | r = 0.02 | r = 0.05 | r = -0.04 | r = -0.03 |
| x_{BONE} | r = 0.60*** CI = 0.45, 0.72 | r = 0.98*** CI = 0.97, 0.98 | r = 0.98*** CI = 0.97, 0.99 | r = 0.97*** CI = 0.96, 0.98 |

*** / dark grey: $p < 0.001$, ** / medium grey: $p < 0.01$, * / light grey: $p < 0.05$, white: no correlation

The optimisation step produced 100 synthetic simulations per patient that provided subsets of the predictions for patients 2, 3, 5, 7 and 8 (Figure 4.4b, green), and Weibull curves were fitted to the corresponding material property input data (Figure 4.5, green).

The second DOE iteration was performed on the three remaining patients (patient 1, 4 and 6), with these population-specific material properties. This gave optimised and narrower predictions with a probability density (Figure 4.4b, blue), whilst retaining the accurate simulation of the nose and upper lip landmarks as compared to the CBCT. The soft tissue predictions following DOE II showed the expected predicted probability ranges (Figure 4.6a). The nose tip had a Bell-curve like probability density whilst the other variables showed a more uniform density. Lastly, the lateral view illustrates the expected variability in the face shape following surgery (Figure 4.6b) and how this compares to the preoperative and postoperative CBCT.

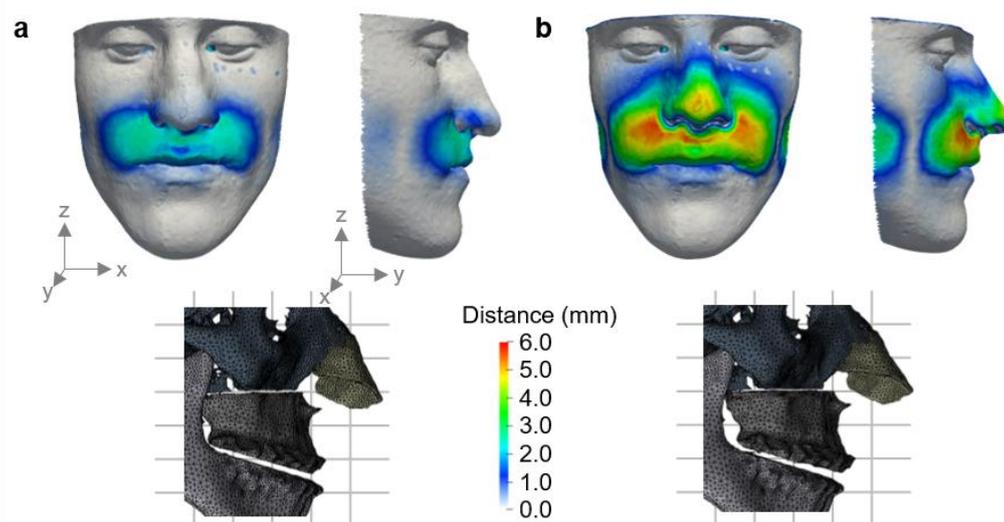


Figure 4.3 Soft tissue prediction results. For patient 1 the range of the soft tissue prediction: **(a)** minimum with $G_{\text{soft}} = 94.4\%$, $E_{\text{soft}} = 0.1$ MPa, $v_{\text{soft}} = 0.45$, $E_{\text{cart}} = 0.5$ MPa, $x_{\text{disp}} = 3$ mm, and **(b)** maximum with $G_{\text{soft}} = 30.5\%$, $E_{\text{soft}} = 1$ MPa, $v_{\text{soft}} = 0.49$, $E_{\text{cart}} = 5$ MPa, $x_{\text{disp}} = 7$ mm. Grey depicts no change from the postoperative CBCT (0 mm), red depicts a maximum change (6 mm).

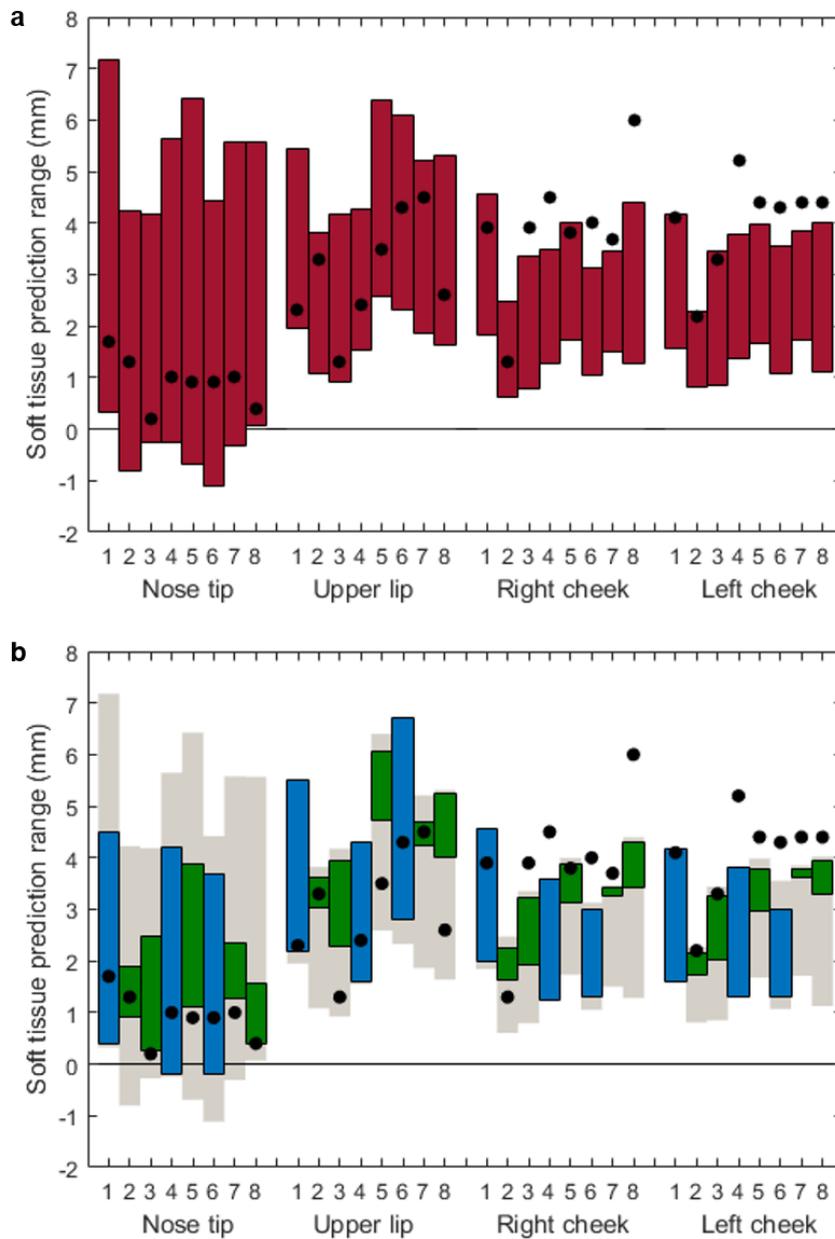


Figure 4.4 Design of experiments I and II. (a) For all eight patients, the vertical bars depict the range of soft tissue prediction in each point and the black dots represent the true postoperative position from CBCT. The baseline (0 mm) is the preoperative CBCT. (b) Green (patient 2, 3, 5, 7 and 8): soft prediction range after optimisation and blue (patient 1, 4 and 6) soft tissue range from design of experiments II, based on optimised material properties.

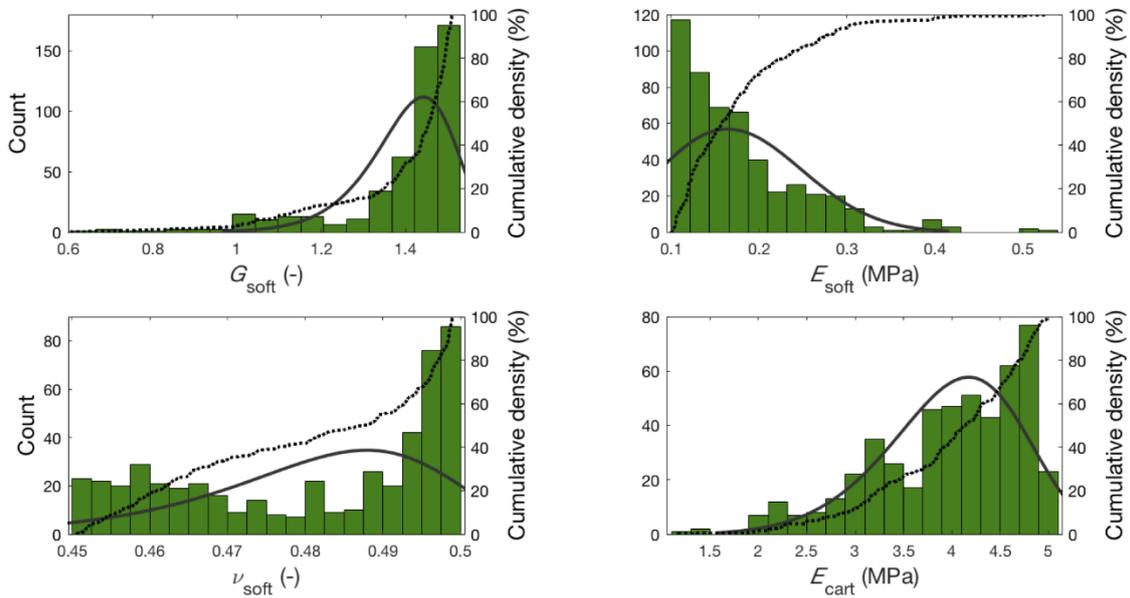


Figure 4.5 Histograms of material property distribution. Material properties following optimisation, based on all training data (patient 2, 3, 5, 7 and 8). A Weibull curve (solid line) was fitted to the data, and the cumulative density (dotted line) is also shown. G: viscoelastic scale factor, E: Young's Modulus, ν : Poisson's ratio, soft: soft tissue, cart: cartilage.

4.4 Discussion

In this chapter, a PFEM approach is described for the simulation of soft tissues in orthognathic surgery. The advantage of the PFEM is that it considers uncertainties in material properties as well as uncertainties related to the osteotomy and maxillary repositioning. The soft tissue simulation, therefore, does not consist of a single deterministic outcome, but of a range of soft tissue outcomes including a minimum and a maximum. This addresses the inability of current soft tissue prediction software to accurately capture the 3D facial outcome with one single result, and thus may be beneficial for patient communication as it demonstrates

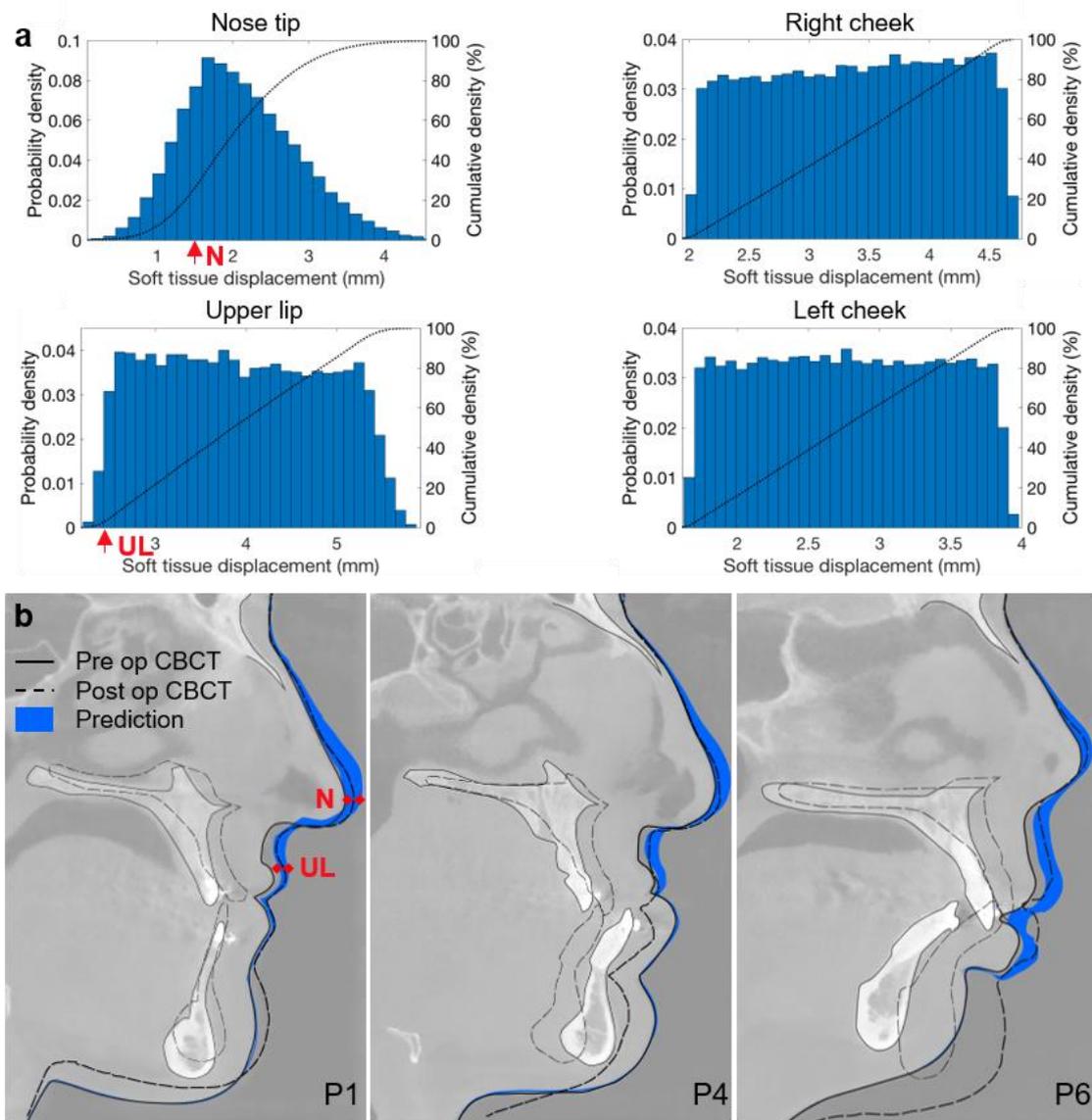


Figure 4.6 Soft tissue prediction probability. (a) For patient 1: histograms and cumulative density plots illustrate the probability of the soft tissue displacement for four landmarks (nose tip (N), upper lip (UL), right cheek and left cheek) – the histograms for patient 4 and 6 showed similar probability curves. The arrows (red) on the x-axis depict the true postoperative soft tissue position, from CBCT, for the upper lip and nose. (b) The lateral view for patient 1, 4 and 6, with the preoperative outline (solid black line), predicted range (blue), and postoperative outline (dotted black line), indicating good agreement between the prediction and the postoperative position.

the potential impact of maxillofacial surgery on their face more realistically. Clinical data from eight patients who had Le Fort I maxillary repositioning were used to train and validate the model and more patients are needed to validate the results in this chapter. Following validation on orthognathic surgery patients, this approach may also be applied to other craniomaxillofacial interventions as well as other types of aesthetic and reconstructive surgery.

Four specific landmarks were used for comparison between the PFEM simulations and the postoperative CBCT: the predicted nose tip and upper lip position accurately described the postoperative position; the simulated cheek position accurately described the postoperative position in three patients and underestimated it in five patients. The underestimation of the cheeks may partially be attributed to swelling, since the time between surgery and postoperative CBCT was 67 ± 23 days and swelling in orthognathic surgery is reported to reduce by approximately 60% after 1 month and 83% after 3 months (Kau, Cronin and Richmond, 2007). This limitation is due to the retrospective nature of this study; the postoperative CBCT scans were acquired on an ad-hoc basis to assess skeletal outcomes prior to orthodontic movements rather than the soft tissue results, which also explains the variability in the postoperative scan date (Table 4.1).

The uncertainty in bone position had the largest influence on the predicted soft tissue displacements, followed by the soft tissue material properties. This suggests that the accuracy of any FEM soft tissue simulation is mostly defined by the ability of the surgeon to match the pre-procedural plans for bone repositioning. The postoperative bone position may differ from the planned position up to ± 2 mm due to the difficulty of accurately reproducing pre-operatively planned maxillary movements, even when CAD/CAM and surgery navigation is used (Mazzoni *et al.*, 2010, 2015; Aboul-Hosn Centenero and Hernández-Alfaro, 2012; Badiali *et al.*, 2015).

Another cause of underestimation might be the lack of titanium bone plates in the model, which have a standard thickness of 1 mm (Gilardino, Chen and Bartlett, 2009). Looking at the lower lip, there is a large variability in mandible position across CBCT scans which may influence lower lip position as the patient cohort comprised patients who had an isolated Le Fort I osteotomy and patients who had bimax osteotomies. Therefore, to ensure uniformity in the analysis, the lower lip and mandible position was omitted from the analysis. Not including the BSSO is a limitation of this proof-of-concept study, as well is that maxillary advancement was modelled solely in the anterior-posterior plane without rotation. Moreover, the optimisation process to improve the soft tissue simulation involved training data from five patients. Prospective studies with a large number of patients from a single surgical procedure – e.g. an isolated Le Fort I cohort – are required to validate the results in this chapter, including the optimised population-specific variables and test the applicability for interpolation onto other cohorts, or comparison with other statistical approaches (Suh *et al.*, 2012; Lee *et al.*, 2014; Pan *et al.*, 2016).

Another limitation relates to the complexity of the soft tissues. The PFEM has a homogeneous soft tissue layer with no subdivision between skin, superficial muscular aponeurotic system (SMAS) and fat, which is a relatively simplistic anatomy compared to inhomogeneous tissue models (Westermarck, Zachow and Eppley, 2005), and constitutive model parameters have been reported for skin, SMAS and fat (Barbarino, Jabareen and Mazza, 2009; Mazza and Barbarino, 2011). An anatomically detailed mesh with distinct soft tissue layers can improve accuracy and reduce the range between minimum and maximum predictions, for example by developing a detailed digital patient model based on MRI (Chabanas, Luboz and Payan, 2003; Luboz *et al.*, 2005; Zhang *et al.*, 2016), but technical

limitations as discussed in chapter 2 make such strategies implausible. Lastly, the surface distances were computed using the closest-point Euclidean distance, without any point-to-point correspondence, which may result in error underestimation in the coronal and sagittal plane (Badiali *et al.*, 2015).

4.5 Summary

In conclusion, this retrospective proof-of-concept study demonstrated the implementation of a PFEM for the computer-assisted surgical simulation on a cohort of eight patients who had orthognathic surgery. The shortcomings traditional of deterministic models prevent clinicians from using them for patient communication or require clinicians to warn their patients about the limitations on face predictions. The advantage of a probabilistic approach compared is that it provides insight into how inaccuracies in the modelling and uncertainties in executing surgical planning can influence the soft tissue simulation. PFEM better illustrates the impact of maxillofacial surgery on the face by providing a confidence interval, which may be beneficial for patient communication and surgical planning. Specifically, the PFEM used a DOE scheme which first quantified the relationship between various input and output variables, showing that the bone repositioning and skin material properties were most strongly correlated to the simulated soft tissue. PFEM was then used to provide simulations with a confidence interval, and, after optimisation, with a confidence interval and a non-uniform probability, whereas traditional deterministic simulations provide a single result, without any information on the (in)accuracy of these simulations.

Whilst this PFEM comes at a substantial computational cost, a similar probabilistic approach could be implemented based on the FDM, thereby improving speed. The proposed PFEM framework is not limited to the prediction of soft tissue changes following a Le Fort I osteotomy, but it may be applied to other orthognathic and craniomaxillofacial interventions, as well as other types of plastic and reconstructive surgery. Whilst the PFEM showed benefits over traditional deterministic models, fundamentally all the physical models discussed so far, including PFEM, are time-consuming as they require a user to manually define the surgeon plan. In chapter 6, a statistical model is introduced that automates much of the computer-assisted surgical planning pipeline in an attempt to streamline the technology and appeal to an audience beyond specialised centres. First, in the next chapter, PFEM and established commercial tools will be compared to show in detail the benefits and pitfalls of each method.

Chapter 5 **CLINICAL COMPARISON OF PHYSICAL MODELS**

Part of the work described in this chapter has been published in the *International Journal of Oral and Maxillofacial Surgery*:

- Knoops PGM, Borghi A, Breakey RWF, Jeelani NUO, Ong J, Bruun R, Schievano S, Dunaway DJ, Padwa BL. 3D soft tissue prediction in orthognathic surgery: a clinical comparison of Dolphin, ProPlan CMF, and probabilistic finite element modelling. *Int J Oral Maxillofac Surg*, 2018.

Rights from Elsevier for publication automatically granted under author permissions.

Following the introduction of a novel probabilistic physical model for surgical simulation, this chapter involves a retrospective clinical study comparing PFEM and two commercial deterministic programs: Dolphin 3D and ProPlan CMF. A qualitative and quantitative evaluation unravels the benefits and pitfalls of each of the three modelling approaches, and simultaneously illustrates the advantages and disadvantages of empirical and physical models.

5.1 Introduction

Computerised planning and prediction of procedural outcomes in orthognathic surgery traditionally rely on manual tracing of 2D cephalometric radiographs (Dvortsin *et al.*, 2008) and the use of established 2D hard-to-soft tissue ratios (San Miguel Moragas, Van Cauteren and Mommaerts, 2014). However, 3D computer planning has gained popularity in recent years because it can be effectively used for patient communication, surgical planning and assessment of operative outcomes (Xia *et al.*, 2011; Aboul-Hosn Centenero and Hernández-Alfaro, 2012; Mazzoni *et al.*, 2015). Various commercial programs are available for 3D planning and soft tissue simulation, with the main difference between them being the computational model they employ as noted in chapter 2. Irrespective of the computational model, accuracy of these 3D prediction tools remains ambiguous: some studies showed a prediction errors of <2 mm, which is considered clinically acceptable, whilst other studies contradicted these findings (Marchetti *et al.*, 2011; Nadjmi *et al.*, 2013; Terzic, Combescure and Scolozzi, 2014; Liebrechts *et al.*, 2015; Mundluru *et al.*, 2017; Resnick *et al.*, 2017). Inaccuracies originate from the prediction algorithm, for example the computational model and its assumptions and simplifications, and the mismatch between the bone position in preoperative planning and the operation itself.

The purpose of this chapter is to evaluate three different programs – Dolphin 3D (Dolphin Imaging & Management Solutions, Chatsworth, CA, USA), ProPlan CMF (Dentsply-Sirona, York, PA, USA) and PFEM (chapter 4) – and to compare their soft tissue predictions in a group of patients who had isolated Le Fort I maxillary osteotomy. The specific aims are to investigate the features and limitations of the three different 3D soft tissue simulation programs, and to determine how these limitations may affect clinical utility in orthognathic surgery. The hypothesis was that all three methods would provide clinically meaningful simulations, in line with previously reported results, but that each method and its underlying computational model has advantages and disadvantages over the other methods.

5.2 Methodology

5.2.1 Patients

Seven patients (5 female, mean age 18 ± 1 years, Table 5.1) who had single-jaw Le Fort I maxillary advancement with vertical repositioning and an alar base cinch suture were retrospectively included in this study. All patients were seen at Boston Children’s Hospital, Boston, MA, USA between December 2011 and January 2015 and had CBCT images acquired three months preoperatively and one-year postoperatively. The indication for an operation was maxillary sagittal hypoplasia for all patients ($n = 7$), and some were simultaneously treated for vertical hyperplasia or hypoplasia ($n = 5$), specifically: anterior maxillary vertical hypoplasia ($n = 3$), anterior vertical maxillary excess ($n = 1$), and posterior vertical maxillary excess with anterior open bite ($n = 1$). All patients had preoperative and postoperative orthodontic treatment, with fixed orthodontic appliances in place during the Le Fort I osteotomy and no

Table 5.1 Patient characteristics. Mean and standard deviation (SD) are shown in the bottom row except for the male-female distribution and diagnosis, and the mismatch displays the absolute mean and SD. MSH: maxillary sagittal hypoplasia, MSH: maxillary sagittal hypoplasia, MVH: maxillary vertical hypoplasia, MVH: maxillary vertical hypoplasia, VME: vertical maxillary excess, AOB: anterior open bite.

| Patient | Gender | Ethnicity | Age (years) | Diagnosis | Planned advancement (mm) | Postoperative advancement (mm) | Mismatch (mm) | Preoperative CBCT (days) | Postoperative CBCT (days) |
|---------------|--------|------------------|-------------|---------------|--------------------------|--------------------------------|---------------|--------------------------|---------------------------|
| 1 | M | Caucasian | 19 | MSH | 4 | 3.0 | -1.0 | 97 | 385 |
| 2 | F | Caucasian | 18 | MSH | 5 | 3.3 | -1.7 | 65 | 422 |
| 3 | M | Caucasian | 18 | MSH, AOB | 6 | 5.8 | -0.2 | 115 | 365 |
| 4 | F | African American | 16 | MSH, MVH | 6 | 5.8 | -0.2 | 100 | 608 |
| 5 | F | African American | 17 | MSH, MVH, VME | 6 | 6.5 | +0.5 | 134 | 231 |
| 6 | F | Caucasian | 18 | MSH | 7 | 7.3 | +0.3 | 117 | 369 |
| 7 | F | Caucasian | 18 | MSH, MVH | 8.5 | 9.5 | +1.0 | 92 | 580 |
| Mean \pm SD | 5F/2M | | 18 \pm 1 | | 6.1 \pm 1.4 | 5.8 \pm 2.1 | 0.7 \pm 0.6 | 103 \pm 22 | 423 \pm 131 |

appliances in place during the postoperative CBCT. This study was approved by the Institutional Review Board of the Center for Applied Clinical Investigation at Boston Children's Hospital (#00019505) and all patients provided consent.

5.2.2 Surgical simulation and soft tissue prediction

Surgical simulation including soft tissue prediction was performed retrospectively based on the preoperative CBCT, making use of Dolphin 3D (version 11.95), ProPlan CMF (version 3.0.1) and PFEM (chapter 4). The framework for surgical simulation comprised 4 steps (Figure 5.1): image acquisition, image processing, soft tissue prediction, and post-processing & visualisation. DICOM files were imported and the head structures segmented, resulting in 3D reconstructions of bone and soft tissue, with negligible differences between the three methods. Reconstructed volumes from preoperative and postoperative CBCT were aligned on the skull base using an ICP algorithm and a Le Fort I osteotomy was virtually performed according to the postoperative CBCT (Table 5.1) with average movements of: 5.8 mm sagittal advancement and 1.1 mm vertical shortening measured at the bone A-point, as well as 2.4° steepening of the occlusal plane measured as the angle formed by the intersection of a line drawn through anterior and posterior nasal spine in the midsagittal plane and the Frankfort horizontal. The soft tissue simulation from each method was exported as a STL file for analysis.

Differences between the three methods exist. First, Dolphin 3D utilises a sparse landmark-based algorithm for soft tissue prediction which enables bespoke hard-to-soft tissue ratios to account for inter-patient differences such as upper lip thickness (San Miguel Moragas, Van Cauteren and Mommaerts, 2014). Three sets of hard-to-soft tissue ratios were investigated to observe how adjusting the ratios affects the predicted soft tissues, based on literature and

further assessed by an orthodontist (Figure 5.2): a default, minimum, and maximum set (San Miguel Moragas, Van Cauteren and Mommaerts, 2014). Second, ProPlan CMF uses a FDM-based computational framework (see chapter 2) which has no manual setting for bespoke patient-specific variables. Third, PFEM with population-specific parameters was used to compute a single soft tissue simulation per patient.

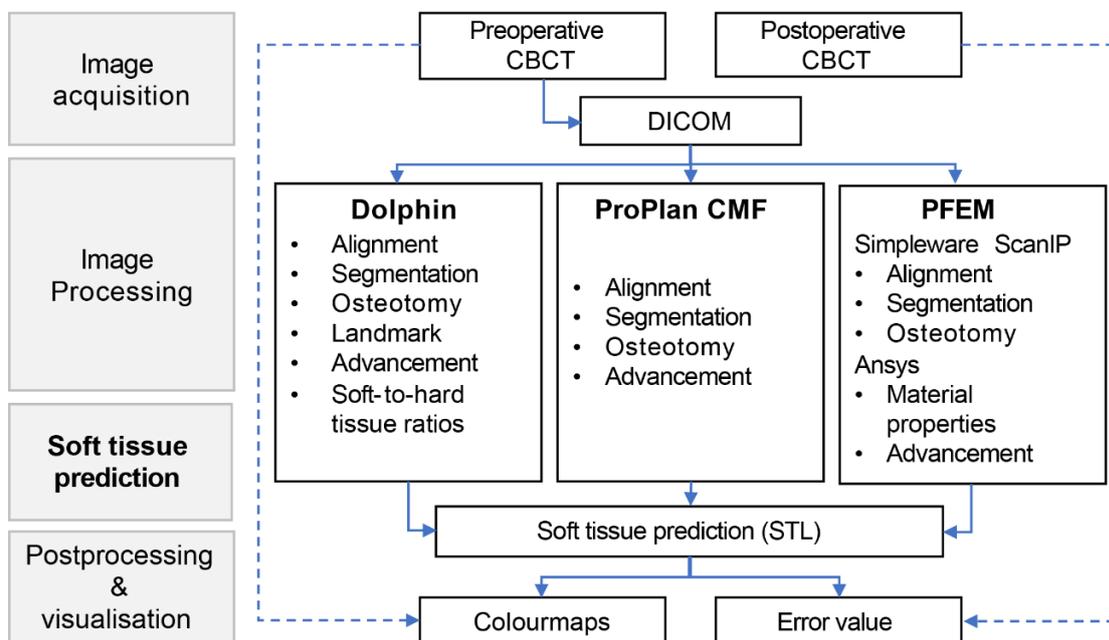


Figure 5.1 Pipeline for soft tissue prediction with Dolphin, ProPlan CMF and PFEM. That the image processing for the three methods is identical, consisting of ICP alignment, segmentation, osteotomy, and advancement. The differences are within the soft tissue prediction algorithms: Dolphin 3D is a landmark-based method and allows for patient-specific soft-to-hard tissue ratios to be set; ProPlan CMF does not require landmarks and is, therefore, relatively straightforward; and PFEM requires two separate programs for the full process, but allows for patient or population-specific material properties to be defined.

5.2.3 Data analysis

Differences between the various datasets were visualised via colourmaps: the preoperative CBCT was compared to the postoperative CBCT and the three sets of soft tissue predictions (Figure 5.2); additionally, the postoperative CBCT was compared to the three sets of soft tissue predictions (Figure 5.3). These comparisons of surfaces were computed as the closest-point distance vectors as in chapter 3 using VMTK (Antiga *et al.*, 2008) in Matlab and visualised in Paraview (Ahrens, Geveci and Law, 2005) (Kitware, Clifton Park, NY, USA). For comparison to the postoperative CBCT, only the upper lip and paranasal regions were of interest as these are the areas of the face affected by Le Fort I maxillary advancement. The full-face surfaces were cropped (Figure 5.3) in Meshmixer with a plane created between the stomion superius, left tragus, and right tragus; and another plane between the subnasale, left tragus, and right tragus. The accuracy was measured as RMS and as the agreement of two surfaces $< 2\text{mm}$, described as a percentage (P).

To investigate how the mismatch between the planning and surgery (Table 5.1) influences the predicted soft tissue, a simulation was performed using the planned bone repositioning as well as the postoperatively measured bone repositioning. The planned position was based on the preoperative clinical notes and the postoperative position was measured from CBCT. Furthermore, for the patient with the largest difference between planned and delivered maxillary position, a range of bone displacements (0-7 mm) was tested to further assess the effect on soft tissue prediction (Figure 5.4). The Friedman test was used to verify the null hypothesis that data come from a continuous distribution with equal means for more than two groups, followed by *post-hoc* Wilcoxon signed rank testing for pairs within the group and

Bonferroni correction – significance for three groups was set at $0.05/3 = 0.017$ – to minimise the likelihood of Type I error (Armstrong, 2014).

5.3 Results

The postoperative CBCT and soft tissue predictions were compared to the preoperative CBCT scan (Figure 5.2a-e). For Dolphin 3D, with default values for hard-to-soft tissue ratios, soft tissue displacements were localised in the areas of the nose and upper lip and the paranasal region showed limited movement. Changing the hard-to-soft tissue ratios had a limited effect on the 3D simulation (Figure 5.2f-h). ProPlan CMF and PFEM showed continuous displacement distributions across the nose, upper lip and paranasal region (Figure 5.2d,e). A comparison between the simulations and the postoperative CBCT, to assess how well the simulated surfaces capture the true postoperative position (Figure 5.3, patient 2), showed that Dolphin underestimated the displacement of the paranasal region across the population whilst ProPlan CMF and PFEM overestimated the displacement of the area above the chelion. Average RMS and average percentage of points <2 mm was: $RMS_{DOLPHIN} = 1.8 \pm 0.8$ mm, $RMS_{PROPLAN} = 1.2 \pm 0.4$ mm, and $RMS_{PFEM} = 1.3 \pm 0.4$ mm, and mean percentages <2 mm $P_{DOLPHIN} = 83 \pm 12\%$, $P_{PROPLAN} = 91 \pm 9\%$, and $P_{PFEM} = 88 \pm 10\%$ (Table 5.2). The Friedman test showed significant differences amongst RMS in the groups ($\chi^2 = 10.57$, $df = 2$, $p = 0.005$) and *post-hoc* Wilcoxon signed-rank tests with Bonferroni correction showed significantly lower RMS for ProPlan CMF compared to Dolphin 3D ($p = 0.016$) and for PFEM compared to Dolphin ($p = 0.016$), and no significant differences between ProPlan and PFEM ($p = 0.219$).

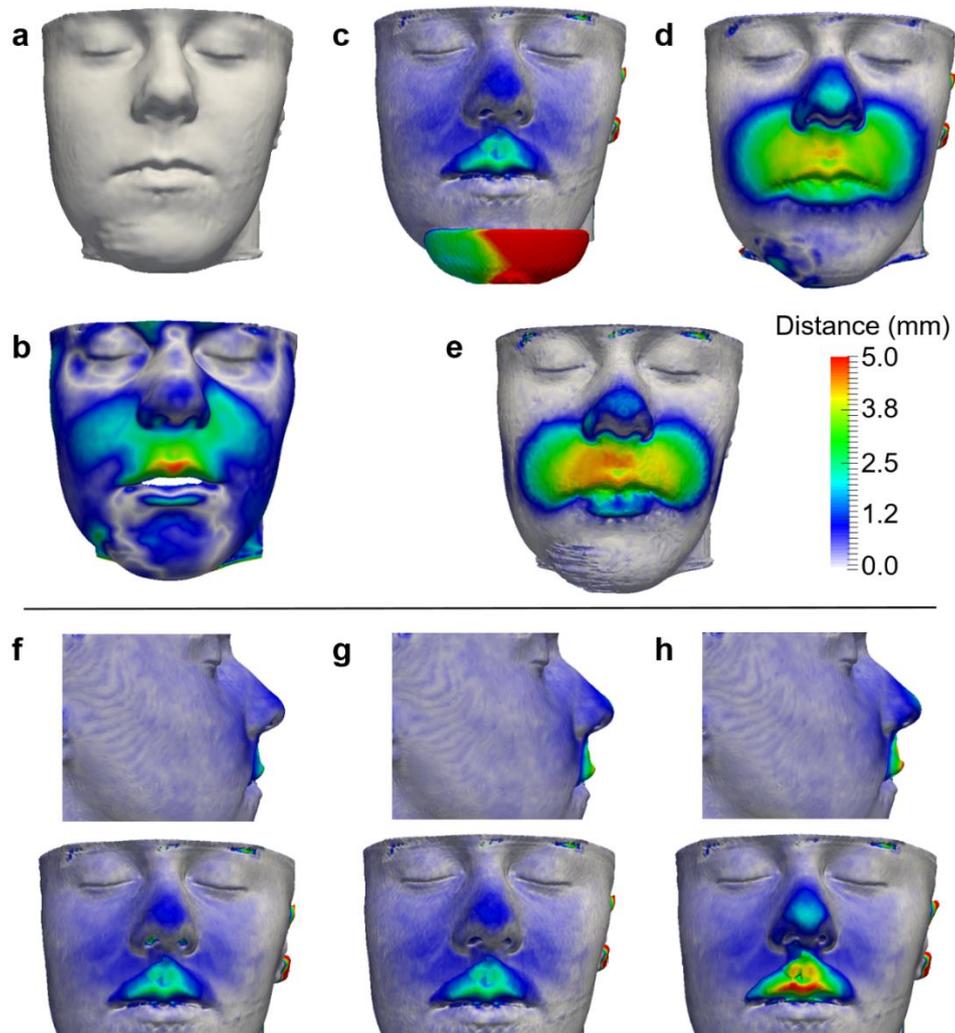


Figure 5.2 Distance colour maps of various soft tissue simulations. Colourmaps are relative to the (a) preoperative CBCT: (b) postoperative CBCT, (c) Dolphin, (d) ProPlan, and (e) PFEM. Patient-specific hard-to-soft tissue ratios were investigated in Dolphin and the corresponding frontal and lateral views are shown for: (f) minimum ratios, (g) default ratios, and (h) maximum ratios. Note that in (c) the chin support of the CBCT scanner is still present; the soft tissues and support have identical grey intensity values and Dolphin does not allow for manual tracing.

The mismatch in maxillary position between the surgical plan and postoperative position on the sagittal plane was investigated (Figure 5.4). For each patient, ProPlan 3D and PFEM provided significantly better results when using the planned position compared to the actual postoperative maxillary position, although the mean differences were small: $RMS_{PROPLAN,POST} = 1.2 \pm 0.4$ mm and $RMS_{PROPLAN,PLAN} = 1.3 \pm 0.4$ mm ($p = 0.002$); $RMS_{PFEM,POST} = 1.3 \pm 0.4$ mm and $RMS_{PFEM,PLAN} = 1.4 \pm 0.4$ mm ($p = 0.002$). The error shows a parabola-like behaviour in which the minimum corresponded to the postoperative position, and a mismatch in the maxilla position therefore renders an error that is larger than the minima. There was no statistical difference in the mean RMS for Dolphin: $RMS_{DOLPHIN,POST} = 1.8 \pm 0.8$ mm and $RMS_{DOLPHIN,PLAN} = 1.8 \pm 0.9$ mm ($p = 0.812$).

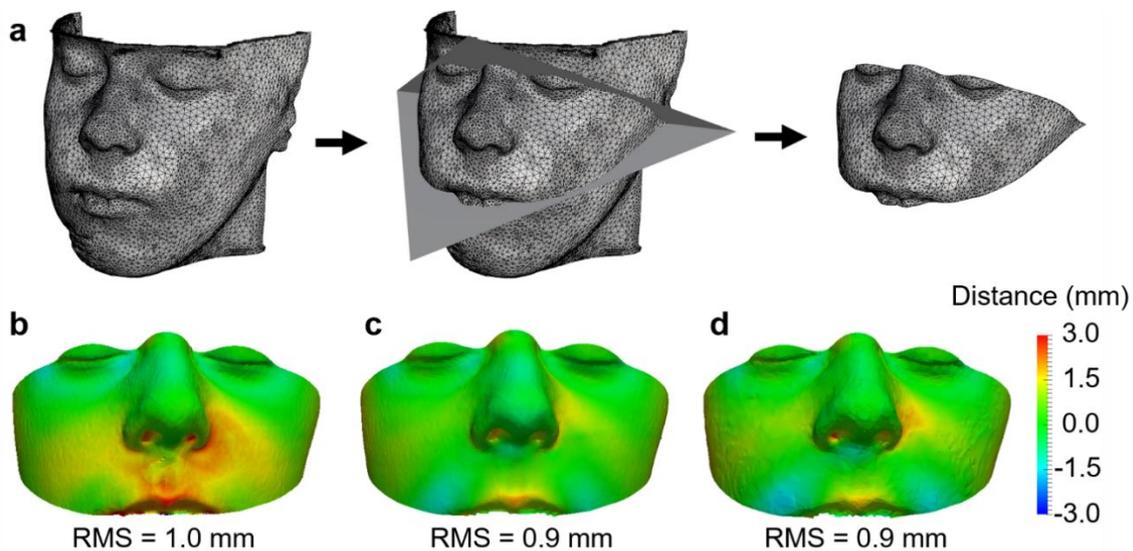


Figure 5.3 Image processing and accuracy of predictions. (a) Cutting plane as defined by the left and right tragus, subnasale, and stomion superior. (b) Comparison of soft tissue simulation with postoperative CBCT for Dolphin, (c) ProPlan CMF, and (d) PFEM. The colourmap demonstrates the overestimation (blue) and underestimation (red) of the simulation compared to the postoperative CBCT.

Table 5.2 Root mean square distance and percentage <2 mm of the soft tissue prediction compared to the postoperative CBCT. Mismatch is the difference in maxillary position in the sagittal plane at the A-point between the postoperative CBCT and the planning. RMS: root mean square, P: percentage.

| Patient | Mis-match (mm) | Dolphin | | ProPlan CMF | | PFEM | |
|-----------|----------------|-----------|-------------|-------------|-------------|-----------|-------------|
| | | RMS (mm) | P <2 mm (%) | RMS (mm) | P <2 mm (%) | RMS (mm) | P <2 mm (%) |
| 1 | -1.0 | 1.6 | 93 | 0.9 | 98 | 0.9 | 98 |
| 2 | -1.7 | 1.0 | 97 | 0.8 | 100 | 0.8 | 99 |
| 3 | -0.2 | 1.9 | 73 | 1.4 | 86 | 1.5 | 81 |
| 4 | -0.2 | 1.9 | 74 | 1.5 | 83 | 1.7 | 76 |
| 5 | +0.5 | 1.4 | 89 | 1.1 | 94 | 1.1 | 91 |
| 6 | +0.3 | 1.5 | 86 | 0.9 | 99 | 1.0 | 96 |
| 7 | +1.0 | 3.4 | 66 | 1.7 | 78 | 1.8 | 76 |
| Mean ± SD | 0.7 ± 0.6 | 1.8 ± 0.8 | 83 ± 12 | 1.2 ± 0.4 | 91 ± 9 | 1.3 ± 0.4 | 88 ± 10 |

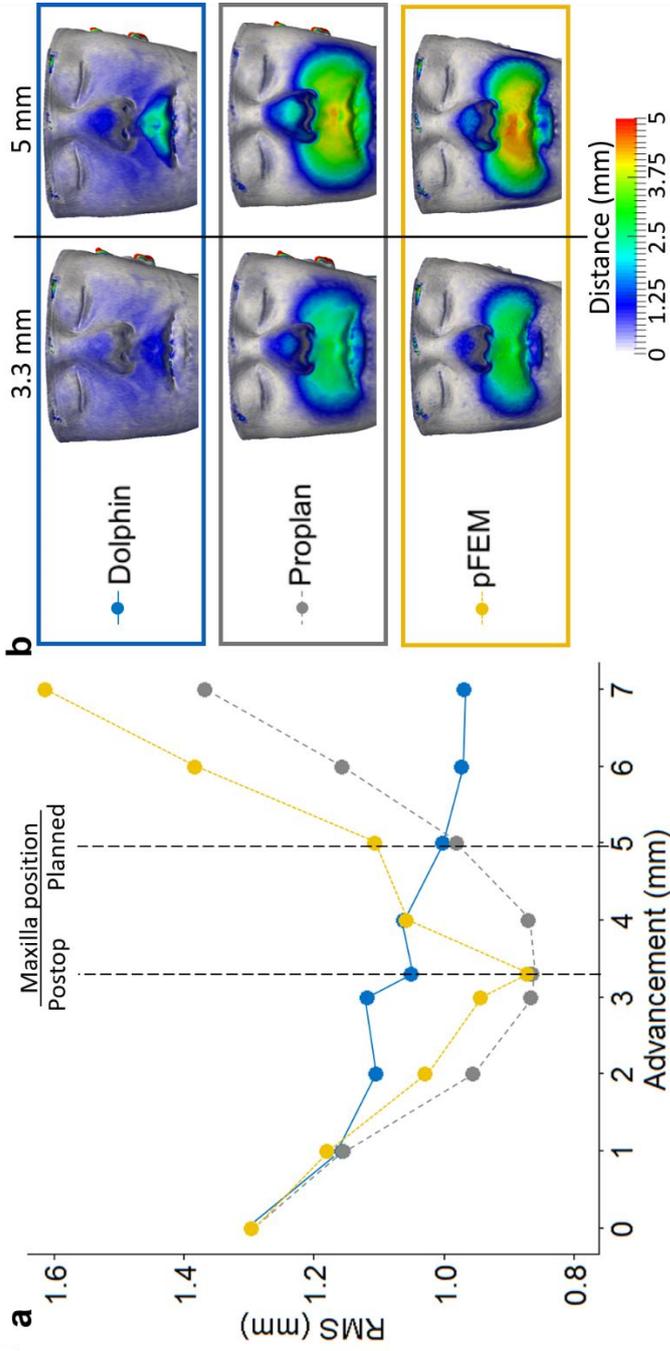


Figure 5.4 Range of advancements and the associated error evaluation. (a) RMS indicating the difference between the soft tissue prediction and the postoperative soft tissues. Each dot represents a simulation made with for a range of horizontal advancements (0-7 mm). The planned maxillary position was 5 mm and postoperative maxillary position was 3.3 mm at the A-point (dotted vertical lines). Soft tissue predictions based on the postoperative maxillary position correspond to the lowest RMS for PFEM and ProPlan CMF, whilst soft tissue predictions at 7 mm correspond to the lowest RMS for Dolphin. **(b)** The colourmaps illustrate how the predicted soft tissue changes between the planned and postoperative position, relative to the preoperative CBCT.

5.4 Discussion

The purpose of this study was to compare 3D soft tissue predictions made with Dolphin, ProPlan, and PFEM to reconstructed soft tissue surfaces from the postoperative CBCT. Strict inclusion criteria were used to minimise intra-patient variation by the inclusion of patients who had CBCT scans taken three months preoperatively and one-year postoperatively (as swelling can be present up to 6 to 12 months postoperatively (Van Der Vlis *et al.*, 2014)), patients who had an isolated Le Fort I osteotomy and who had no orthodontic appliances in place at one-year follow-up. There was a good agreement between the predicted soft tissue and postoperative CBCT for all three methods. Although the accuracy was better for ProPlan CMF and PFEM compared to Dolphin 3D and this difference was statistically significant, it is important to consider how those predictions differ topologically and to assess the clinical significance.

Soft tissue predictions in ProPlan CMF and PFEM were significantly more accurate than in Dolphin 3D. Specifically, Dolphin 3D mainly simulated changes on the 2D midline and in the upper lip region, and limited displacement in the paranasal region. This is due to its sparse landmark-based morphing algorithm, and these findings are in line with a previous pilot study of the same cohort that assessed various landmarks on the midline (average error of 1.7 mm) and off the midline (2.9 mm) (Resnick *et al.*, 2017). On the contrary, predictions with ProPlan CMF and PFEM showed continuous displacements in the nose, upper lip and paranasal region. There was no significant difference between PFEM and ProPlan CMF and their high accuracy renders them suitable for pre-operative assessment and patient communication. However, it is important to consider other differences between the two: ProPlan CMF is designed to be intuitive for the clinician being a commercial ad-hoc software, but it does not allow the user to

manually change any patient-specific parameters, which could limit its use for more extreme patients that fall outside of the normal patient population. PFEM is a numerical methodology that could be implemented in any FEM package, which however requires knowledge of the underlying numerical algorithm. Its computational flexibility allows for creation of a range of predicted outcomes based on patient-specific or population-based parameters, which may improve patient communication and properly inform on the range of possible outcomes. Moreover, it better suits the evaluation of unique cases that do not conform to a standard set of rules.

The mismatch between the planned and postoperative maxilla position in the sagittal plane was relatively small (mean absolute difference = 0.6 ± 0.5 mm), yet it had a statistically significant influence. Unsurprisingly, for ProPlan CMF and PFEM, simulations with the postoperative position were significantly more accurate than simulations with the planned position, compared to the postoperative CBCT, although the clinical significance may well be negligible. Specifically, the error showed a parabola-like behaviour which suggests that the physical model accurately captured the biomechanical behaviour, as most accurate predictions were observed with the true postoperative maxilla position. On the contrary, for Dolphin, there was no significant difference between the planned and postoperative maxilla position, which indicates a lack of biomechanical behaviour of its empirical model. There are two specific scenarios when users should be cautious with Dolphin 3D: the lack of continuity worsens with large advancements and errors increase when the postoperative maxilla advancement is larger than the planned advancement.

Some limitations of this study must be noted. Seven patients were retrospectively included due to the strict inclusion criteria to minimise intra-cohort variability, yet there was

still some lack of homogeneity amongst the various diagnosis and type of advancement procedures; a large prospective study with a uniform cohort of patients would be desirable to verify the findings in this chapter. Whilst all patients had an isolated Le Fort I procedure, the maxillary repositioning was different due to the indications of vertical hyper and hypoplasia; some patients had superior maxillary repositioning and others had inferior maxillary repositioning. Despite these differences, all patients had maxillary advancement and the variability within this single procedure reflects the fact that each patient receives a bespoke surgical plan rather than a set procedure. Additionally, the patients had orthodontic appliances in place during preoperative and not during postoperative CBCT which might have introduced a small error in the presence of the appliances as well as the orthodontic movements that occurred postoperatively. To refine measurement of the surgical changes from a CT scan, the skeletal movements should be assessed by placing landmarks onto the bone instead of the teeth, and CT scans should be taken postoperatively before orthodontic treatment commences. The downside of this approach is that the lack of landmarks on the maxilla may affect the reproducibility of landmark localisation.

Moreover, the surface distances were computed using the closest-point difference, without any point-to-point correspondence, which may result in error underestimation, especially in the coronal plane (Badiali *et al.*, 2015). Also, only the A-point was used to assess the postoperative maxilla position, which does not capture the overall 3D movements, although the small mean differences in RMS between using the planned position and the postoperative position suggest the influence of bone position on overall RMS is limited. No standardised protocol exists to evaluate the mismatch in maxillary position; however, an automated tool has recently been proposed (Baan *et al.*, 2016). Another limitation is that the area of the face used

for RMS includes also lateral parts of the face that are minimally affected by maxillary advancement; this may have camouflaged larger error values. However, this area is defined by four landmarks only and fittingly captures the region of interest. Last, the position of the mandible and the lower lip was omitted from this analysis, although a change in mandible position due to autorotation following maxillary repositioning, including impaction of the maxilla, may influence the shape of the soft tissues in the region of interest in the nose, upper lip and paranasal regions.

5.5 Summary

In conclusion, Dolphin 3D, ProPlan CMF, and PFEM were evaluated for soft tissue simulation in orthognathic surgery and this chapter showed that clinically useful predictions can be obtained with each method when considering the overall RMS and the percentage of surface points of the 3D prediction that is accurate within 2 mm. However, it is crucial to be aware of the underlying computational model and the resulting topological soft tissue prediction. The results confirmed that the physical models in ProPlan CMF and PFEM are superior compared to the empirical model in Dolphin 3D. Moreover, the results demonstrated that it is essential to appreciate the intricacies of the method of choice, as each software can provide clinically meaningful simulations for a wide range of perioperative applications. The landmark-based model in Dolphin 3D failed to provide continuous 3D predictions which limits its use to 2D lateral simulations. Comparing the two physical models, the differences in accuracy were not statistically or clinically significant. The benefit of ProPlan CMF lies in the fact that it is an ad-hoc commercial software, relatively easy to use, whereas PFEM is a numerical methodology that requires deeper understanding of computational modelling.

Conversely, ProPlan CMF has no manual settings whilst PFEM is highly customisable, which implies that for certain extreme medical cases – not uncommon in craniomaxillofacial surgery – PFEM might be preferable.

Chapter 6 **3D MORPHABLE MODELS FOR AUTOMATED COMPUTER- ASSISTED DIAGNOSIS AND SURGICAL PLANNING**

Part of the work described in this chapter is under review with *Scientific Reports*

- Knoops PGM, Papaioannou A, Borghi A, Breakey RWF, Wilson A, Jeelani NUO, Zaferiou S, Steinbacher D, Padwa BL, Dunaway DJ, Schievano S. A machine learning framework for automated computer-assisted diagnosis and planning in plastic and reconstructive surgery.

The algorithms in this chapter were developed and implemented by Dr Athanasios Papaioannou (section 6.2.2).

After the description of advanced medical imaging methods in chapter 3 and the physical modelling approach in chapters 4 and 5, this chapter provides an alternative approach to surgical planning based on statistical shape modelling. Although the method proposed in chapter 4 provided substantial improvements over previously described physical models, some intrinsic limitations remain relating to the computational cost and the explorative nature, and ultimately the time and resources needed to deploy physical computer-assisted surgical planning. The practical consequence is that adoption of such planning tools is limited to specialised hospitals (Gøthesen *et al.*, 2013), despite long-term cost benefits (Xia *et al.*, 2006). Therefore, in this chapter a fully-automated statistical modelling framework is described for surgical planning based on a large-scale clinical 3DMM.

6.1 Introduction

Over 200,000 maxillofacial procedures are performed in the USA every year with the purpose of ameliorating orthognathic (jaw) deformities (American Society of Plastic Surgeons, 2017). For these operations, vast quantities of patient data are collected (Kanevsky *et al.*, 2016), thus providing a great opportunity for the development of machine-learning-based methods, for use in clinical decision-making and to enable automated personalised medicine approaches (Bennett and Hauser, 2013; Mirnezami and Ahmed, 2018). Although the application of statistical shape models in plastic and reconstructive surgery is not new – in orthognathic surgery it has been used to elucidate how syndromes affect skull growth (Maas *et al.*, 2018), to quantify (Rodriguez-Florez *et al.*, 2017) or to predict (Crombag *et al.*, 2014) the corrective effect of surgical techniques on skull deformities, and for outcome evaluation (Meulstee *et al.*, 2015) – its clinical usefulness has been limited due to the low number of samples, absence of

automated processing methods, and lack of state-of-the-art mathematical models. Therefore, a machine-learning-based framework was developed involving a large number of data points and a fully automated processing for diagnosis and clinical decision-making in maxillofacial surgery.

Machine-learning-based models, including statistical shape models, have been proposed to streamline and automate processes in computer-assisted surgical planning (Zachow, 2015), thereby making this a more accessible technology. However, accurate statistical modelling of face shape features is a challenging task due to the large anatomical variation in the human population; and to build a statistical model that can truthfully represent each given face, a large collection of high-quality 3D images is required from a population diverse in age, gender, and ethnicity (Paysan *et al.*, 2009; Booth *et al.*, 2016; Huber *et al.*, 2016; Dai *et al.*, 2017). State-of-the-art computer vision algorithms are required to automatically process these 3D images and construct a high-dimensional statistical model. A popular machine learning approach, originally used to reconstruct accurate and complete 3D representations from single 2D images and for photo-realistic manipulation (Blanz and Vetter, 1999), involves 3DMM – statistical models of face shape and texture. Current applications of 3DMM include facial recognition (Blanz, 2006), expression normalisation (Amberg, Knothe and Vetter, 2008), and face reconstruction from videos (Kittler *et al.*, 2016), but no models exist for surgical planning.

In this chapter, a fully-automated large-scale clinical 3DMM is described; a machine-learning-based framework involving supervised learning for diagnostics, risk stratification, and treatment simulation, constructed with databases comprising 5,000 3D face scans of healthy volunteers and patients admitted for orthognathic surgery.

6.2 Methodology¹

6.2.1 Data sources

Two face databases were used, one containing faces from the general public and one propriety patient database. A sample of 3D face scans from the general public were collected from the Large Scale Facial Model (LSFM) (Booth *et al.*, 2018) database which is available under a non-commercial licence for academic use (Zafeiriou and Dunaway, 2018). LSFM comprises 9,663 3D scans from volunteers taken with a 3dMDface system under controlled conditions at the Science Museum in London. Volunteers were excluded based on their age – to age-match the patient database – and those with incomplete records. This provided a database with the following demographics (Table 6.1): mean age = 22.2 ± 3.7 years, 55% male and 45% female, and ethnicity = 83% white, 8% Asian, 4% mixed heritage, 3% black, and 2% other. The patient database included 274 3D surface scans taken with the Vectra M3 system (Canfield Scientific, Parsippany, NJ, USA), collected from 151 patients who underwent orthognathic surgery at Boston Children’s Hospital and Yale-New Haven Hospital between December 2010 and September 2017 (Table 6.1), with the following demographics: mean age = 18.4 ± 2.4 years, 44% male and 56% female, and ethnicity = 76% white, 10% Asian, 10% mixed heritage/other, and 8% black. Additional information included scan date, operation date, surgical procedure, indication for surgery, and syndromic diagnosis (Figure 6.1).

Patient data for this study were retrospectively retrieved from electronic medical records after receiving approval from the Institutional Review Board at Boston Children’s

¹ The Python implementation of the 3DMM pipeline and experiments was done by Dr Athanasios Papaioannou

Hospital (#00019505) and the Human Investigations Committee at Yale-New Haven Hospital (HIC #110100793).

Table 6.1 Face database characteristics. *n* is the number of individuals for whom that measurement was available. SD: standard deviation.

| Characteristics | Dataset | | |
|---------------------------|--|--|--|
| | Patient | LSFM (all ages) | LSFM (14-28yr) |
| Number of subjects | 151 | 10,619 | 3,943 |
| Number of images | 273 | 10,619 | 3,943 |
| Age: mean (SD), years | 18.4 (2.4), <i>n</i> =151 | 24.5 (14.7), <i>n</i> =9,460 | 22.2 (3.7) , <i>n</i> =3,943 |
| Age: range, years | 14 – 28, <i>n</i> =151 | 0 – 85, <i>n</i> =9,460 | 14 – 28, <i>n</i> =3,943 |
| Gender (% male/female) | 44%/56%, <i>n</i> =151 | 47.7%/52.2%, <i>n</i> =9,582 | 55.5%/44.5%, <i>n</i> =3,942 |
| Ethnicity | 72% White, 10% Asian, 10% Mixed Heritage/Other, 8% | 82% White, 9% Asian, 5% Mixed Heritage, 3% Black, 1% Other, <i>n</i> =9,554 | 83% White, 8% Asian, 4% Mixed Heritage, 3% Black, 2% Other, <i>n</i> =3,928 |

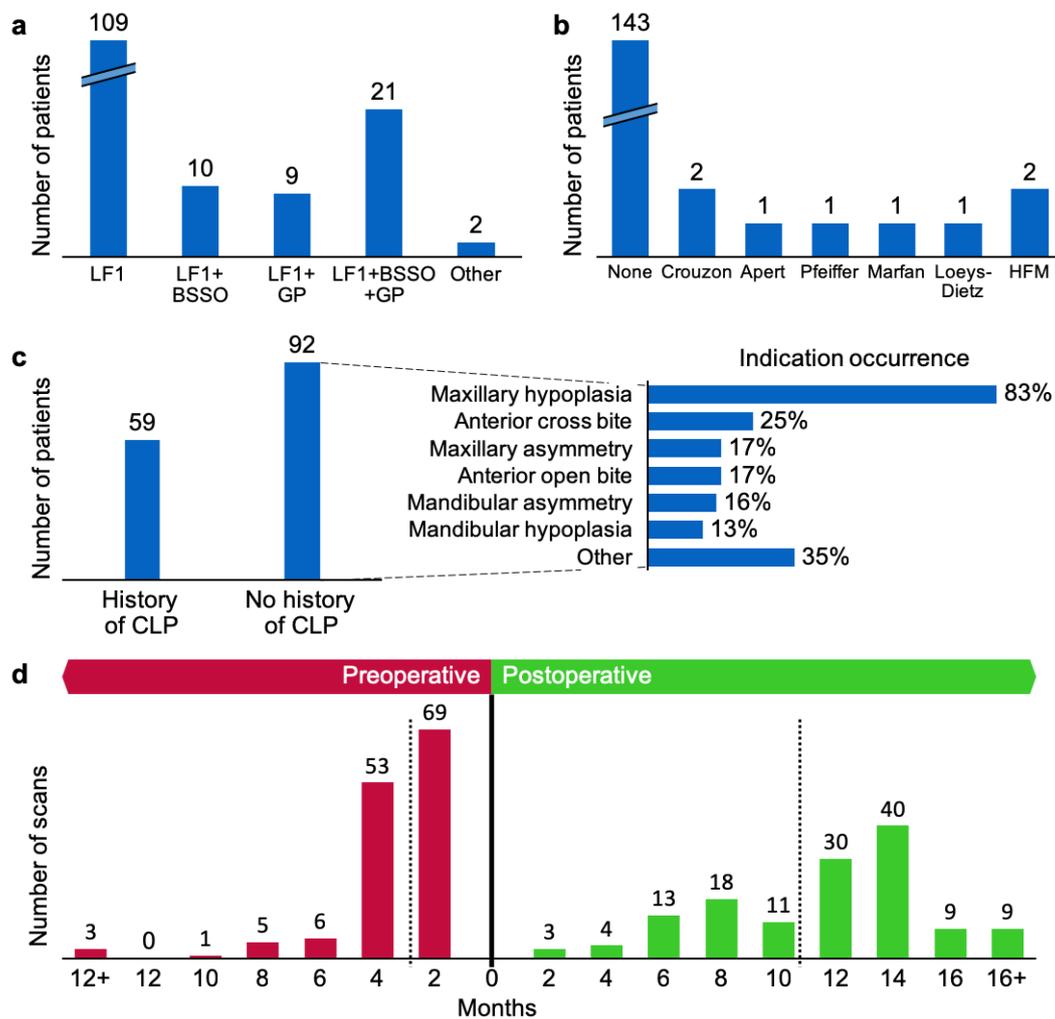


Figure 6.1 Patient characteristics. (a) Number of patients per surgical procedures, LF1: Le Fort 1 osteotomy, BSSO: bilateral sagittal split osteotomy, GP: genioplasty. (b) Number of patients with syndromic diagnosis, HFM: hemifacial microsomia. (c) Number of patients according to the main indication for surgery: history or no history cleft lip and/or palate (CLP). For those without a history of CLP, often multiple indications are present. (d) Timing of preoperative and postoperative scans: average preoperative scan was 73 ± 86 days (red, dotted line) and the average postoperative scan 326 ± 123 days (green, dotted line).

6.2.2 3D morphable model construction

The 3DMM training pipeline, based on the approach proposed in (Booth *et al.*, 2016), was used to create three models: a global model, a bespoke preoperative model, and a bespoke postoperative model. The global model consisted of all LSFM faces as well as 113 pairs of preoperative and postoperative patient faces. The bespoke preoperative and postoperative models, respectively, included all preoperative and postoperative scans available. The pipeline comprises four main functional blocks:

1. Automatic annotation – each 3D mesh was rendered from a number of virtual cameras positioned around the subject into 2D images. A landmark localisation algorithm – an active appearance model (AAM) – was applied to find the 2D landmarks on the rendered images, and each 2D landmark set was projected onto the 3D surface, rendering the 3D landmarks.
2. Alignment and statistical modelling – the collection of scans was brought into the same space by removing similarity effects (rotation, translation, scale) via generalised Procrustes analysis (GPA); leaving only shape information.
3. Dense correspondence – the aligned collection of 3D scans was registered into a form where each scan had the same number of points joined into a triangulation shared across all scans. Dense correspondence was accomplished via non-rigid ICP (NICP) (Amberg, Romdhani and Vetter, 2007), using the LSFM mean face as a template. This process deforms the template mesh to the shape of each patient face to obtain a set of deformed templates.
4. Statistical analysis – the 3DMM model was built by applying principal component analysis (PCA) on the corresponding meshes and finding the eigenvectors-bases with the greatest variance. Any 3D face shape can be defined as a linear combination of these bases, which makes up the 3DMM.

6.2.2.1 Intrinsic model characterisation

The models were characterised and validated with the following intrinsic metrics (Styner *et al.*, 2003): compactness, generalisation, and specificity. Additionally, the performance of the models was benchmarked to the (LSFM), a state-of-the-art 3DMM constructed with 9,663 scans. Compactness is a measure of the cumulative variance in the data that is retained with a certain number of principal components, which was extracted from the model construction. Generalisation describes how well a face unknown to the 3DMM can be approximated by the existing model. Specifically, leave-one-out cross-validation was used for all patient faces in all three models (LSFM, global, bespoke preoperative): a model was constructed for all faces but one, and then is fitted to the excluded face. This was repeated for all patient faces. A large generalisation error suggesting overfitting – the inability of a model to represent previously unseen faces. Specificity measures how well synthesised faces can be approximated by ground-truth images. Specifically, faces (n=10,000) were randomly synthesised for each model, and the specificity error was computed as the lowest average error of all vertices between a synthesised face and the closest ground-truth neighbour.

6.2.2.2 Manifold visualisation

T-distributed stochastic neighbour embedding (t-SNE) was used as a dimensionality reduction technique (Van Der Maaten and Hinton, 2008) to visualise a high-dimensional manifold onto a 2D space. Various hyper-parameters (perplexity = 2 to 100; iterations = 1,000 to 5,000) were tested as well as different numbers of randomly sampled LSFM faces (n=200 to n=3000) together with all preoperative patient faces (n=119) and postoperative patient faces (n=127).

6.2.2.3 Classification for diagnosis

Classification was performed using a subgroup of randomly selected faces ($n=300$) from the LSFM database and faces from pre-operative patients ($n=119$). The whole dataset was split in a stratified manner with various proportions between training and test set (80-20%, 60-40%, and 50-50%). Thus, for the 80-20% case, patient ($n=95$) and LSFM ($n=240$) faces were in the training set and patient ($n=24$) and LSFM ($n=60$) faces in the test set.

For the classifier, a Support Vector Machine (SVM) (Cortes and Vapnik, 1995) with linear kernel $\langle x, x' \rangle$ was used as SVMs are powerful tools for small sample size problems. Additionally, the *scikit* (Pedregosa *et al.*, 2011) implementation of SVM with “one-vs-the-rest” multi-class strategy was employed with default values for the penalty parameter ($C = 1.0$) and gamma. To calculate the mean accuracy, training and test sets were created according to a Monte-Carlo cross-validation scheme by randomly selecting the training and test set 1,000 times.

6.2.2.4 Regression for surgical simulation

To automatically predict face shape outcomes based on the preoperative scan, linear regression (LR), ridge regression (RR), least-angle regression (LARS), and least absolute shrinkage and selection operator regression (LASSO) were tested, using their *scikit* implementation (Pedregosa *et al.*, 2011). For RR and LASSO, the alpha parameter that defines the strength of regularization term was set to 0.5 and 0.1, respectively. For LARS, the number of nonzero coefficients was set to 1 and the default values for all the other parameters. The global model was used to perform regression, which included LSFM ($n=3,664$), preoperative ($n=113$) and postoperative ($n=113$) faces. Specifically, the preoperative and postoperative

scans came from the same unique patients, and regression was performed using the leave-one-out scheme. A design matrix was learnt between the components of the preoperative and postoperative patients, which was then used to map the preoperative components to the postoperative components.

6.2.3 Error quantification

The error was quantified using the average Euclidean distance (AED), calculated from the per-vertex distance between two meshes (6.1):

$$AED = \frac{\sum_{i=1}^n \sqrt{(x_{i,A} - x_{i,B})^2 + (y_{i,A} - y_{i,B})^2 + (z_{i,A} - z_{i,B})^2}}{n} \quad (6.1)$$

where x , y , and z corresponded to the Cartesian coordinates of mesh A and B , and n was the number of vertices per mesh. Each mesh was in dense correspondence and therefore vertex i represented the same anatomical location in both meshes. To compute signed errors for colourmaps, the error was positive if the reference mesh had a larger value on the z -axis.

6.3 Results

6.3.1 Model validation

Compactness showed that 81.8% and 91.6% of the variance are respectively described by the first 10 and 20 principal components for the bespoke preoperative model, 79.6% and 89.3% for the global model, and 79.9% and 89.3% for LSFM (Figure 6.2a). The generalisation error demonstrated the ability to describe patient faces that were not used for training. At 100 components, the global (0.3 mm) and bespoke preoperative (0.4 mm) models outperformed LSFM (1.4 mm), due to lack of patient data in the latter (Figure 6.2b). Additionally, the bespoke preoperative model initially outperformed the global model, but after 48 components this trend reversed, as the bespoke model ran out of statistical variance sooner due to a lower number of samples. For specificity, faces were synthesised ($n=10,000$) and compared to their closest real neighbour. Values of in the range of 0.3-0.4 mm quantitatively indicated good agreement with real faces (Figure 6.2c,d).

6.3.2 Qualitative and quantitative shape analysis

To investigate how the three models differ, shape and variance were qualitatively and quantitatively evaluated. Specifically, the mean shape and first five principal components with standard deviation of +3SD and -3SD were computed, and the differences between the mean shapes were computed. In the LSFM face (Figure 6.3a) and in the postoperative face (Figure 6.3c), lengthening-widening (component 1) and concavity-convexity (component 2) captured most variance whilst in the mean preoperative face (Figure 6.3b), a component of underdevelopment of the upper and lower jaw (component 2) was present. These differences

were confirmed by a direct comparison of the mean preoperative face to the mean LSFM face (Figure 6.4), revealing maxillary hypoplasia (underdevelopment of the upper jaw) and mandibular hyperplasia (overdevelopment of the lower jaw) preoperatively. Although surgery successfully ameliorates the jaw discrepancy, a difference in nose shape remained.

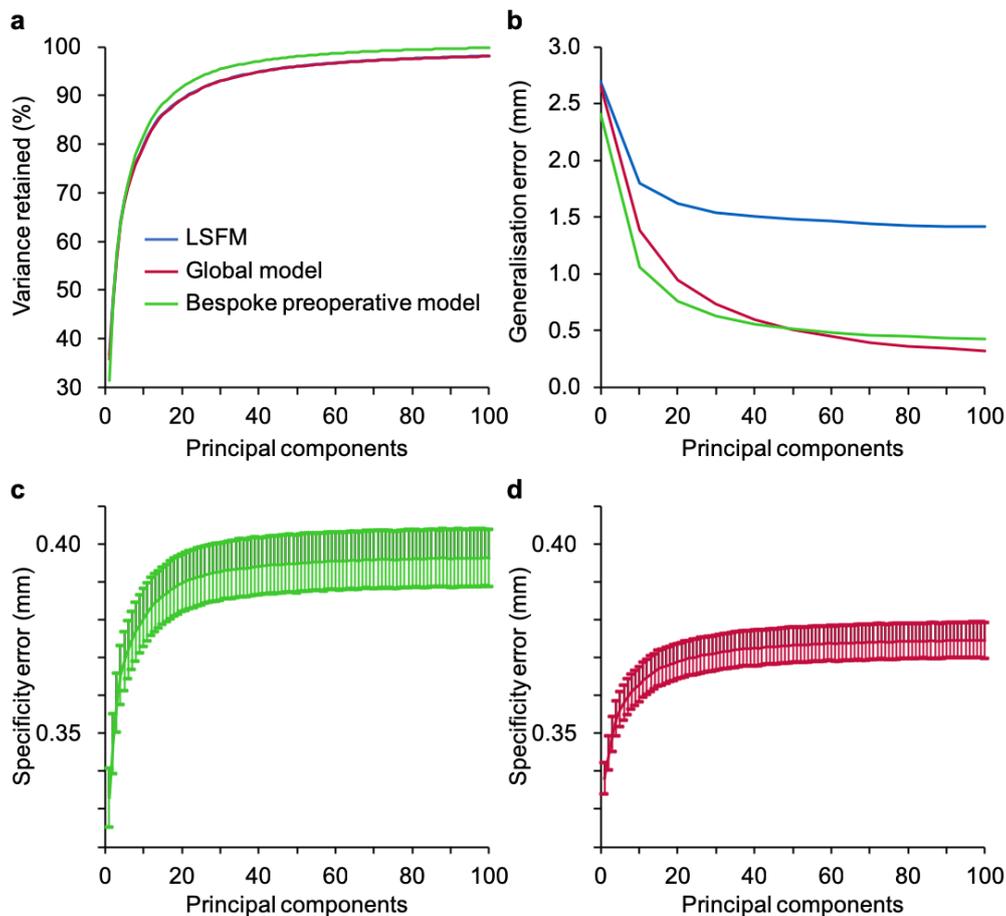


Figure 6.2 Compactness, generalisation, and specificity for the three models.

Characterisation of the three 3DMM compared to LSFM. **(a)** Compactness, the amount of variance retained for a certain number of principal components, is 79.6% at 10 components for the global model (red), 81.8% for the bespoke preoperative model (green) and 79.9% for LSFM (blue). **(b)** Generalisation demonstrates the ability to describe faces that were not used to construct the original model, and at 100 components is 0.3 mm, 0.4 mm, and 1.4 mm for the global model, bespoke preoperative model, and LSFM, respectively. Specificity measures how well synthetic faces resemble real faces; **(c)** the bespoke preoperative model showed errors of 0.40 ± 0.01 mm, and **(d)** the global model 0.37 ± 0.01 mm.

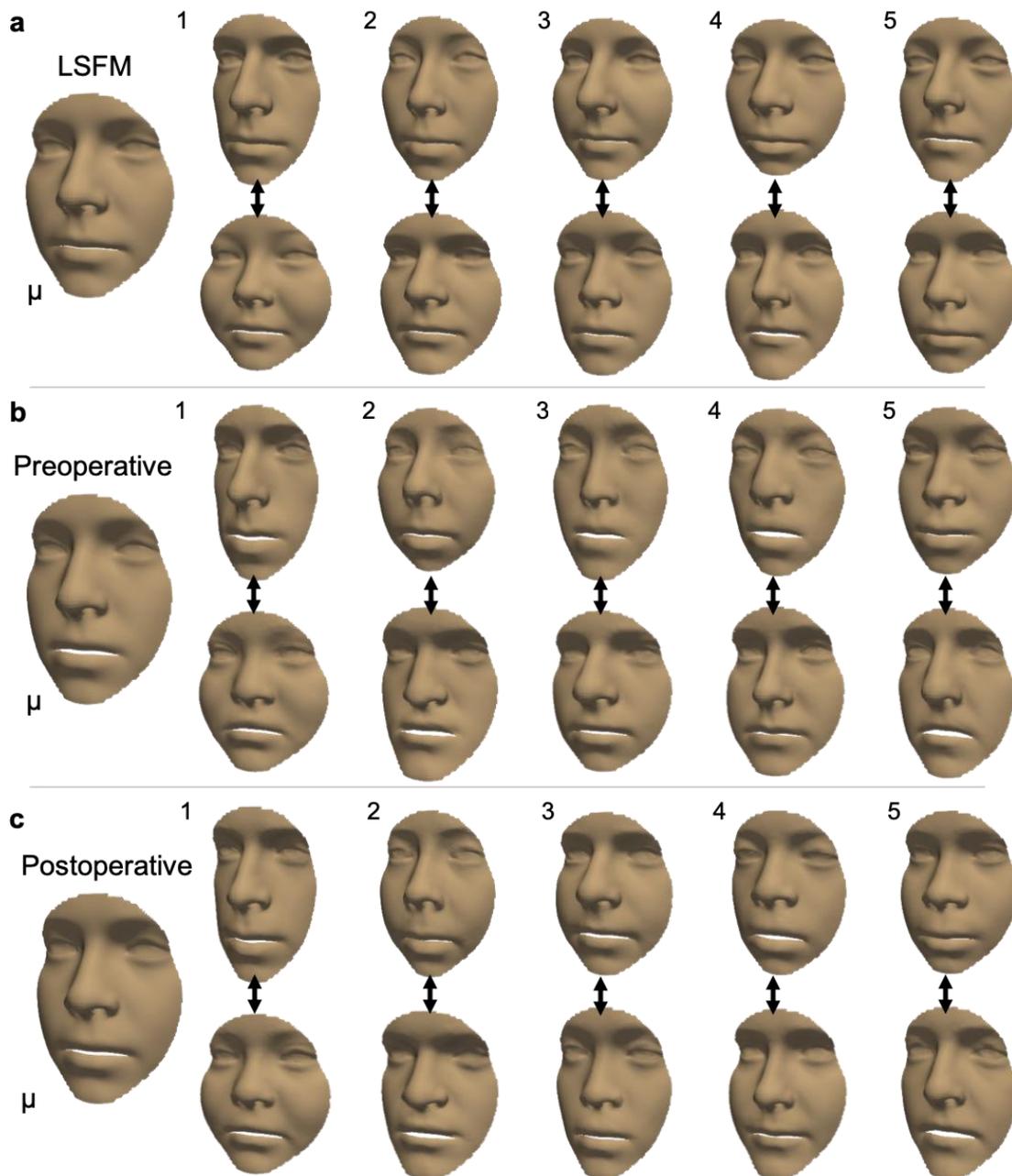


Figure 6.3 Visualisation of the mean shape and variation for a face from the general public, a preoperative face, and a postoperative face. Following generation of the (a) global LSFM, (b) bespoke preoperative and (c) bespoke postoperative shape model, mean shape (μ) and first five shape eigenvectors are displayed, with weights for the standard deviation of +3SD (top row) and -3SD (bottom row).

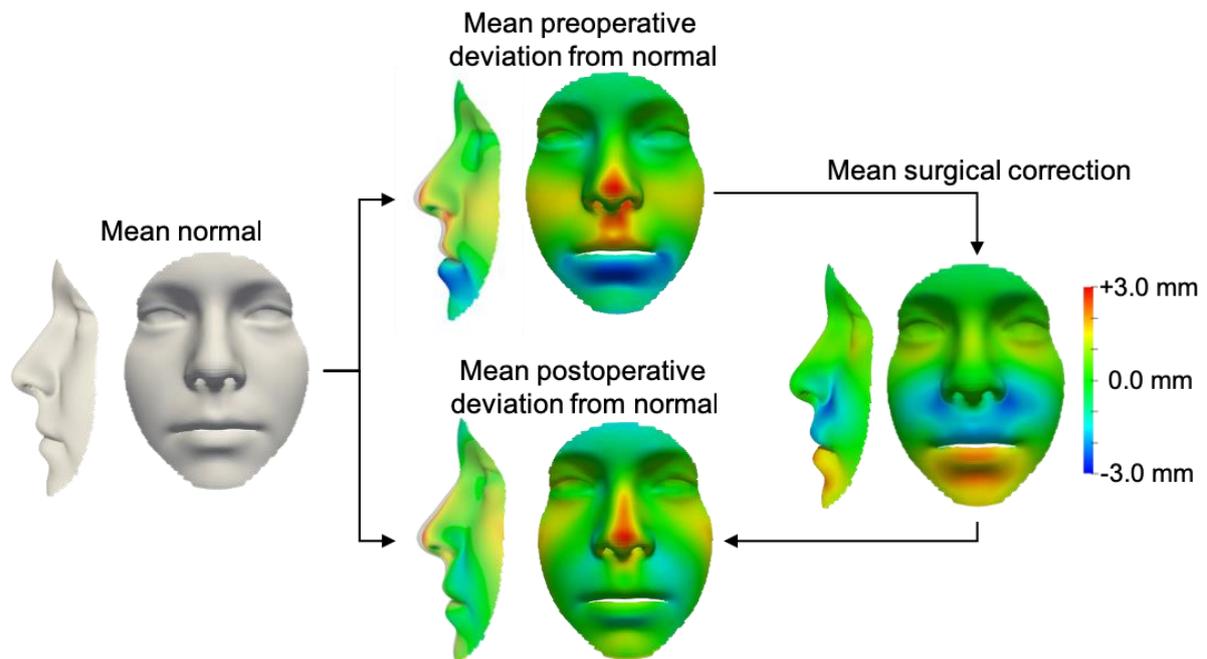


Figure 6.4 Comparison of the mean LSFM, preoperative, and postoperative face. Colourmaps illustrate deviations from the mean LSFM face. The mean preoperative face colourmap is consistent with indications for orthognathic surgery – the cohort of orthognathic patients shows upper jaw underdevelopment (red) and lower jaw overdevelopment (blue). The mean surgical correction appropriately ameliorates these jaw malformations; however, the mean postoperative face retains nose abnormality.

6.3.3 Manifold visualisation

To test the diagnostic potential, first t-SNE (Van Der Maaten and Hinton, 2008) was used for dimensionality reduction of the high dimensional shape vectors, and to visualise the global manifold in two dimensions. With labels for LSFM, preoperative, and postoperative, no distinct groups were uncovered (Figure 6.5a), although the majority of patient faces, preoperatively and postoperatively, appear to populate the perimeter of the t-SNE embedding which demonstrates substantial shape similarity amongst the groups.

To elucidate the complex relationship amongst neighbouring face shapes, a patient's face (Figure 6.5b) is displayed that is close to two LFSM faces (Figure 6.5c,d) in the t-SNE embedding, showing resemblance in the facial profile and particularly the upper lip area. Patient faces that often appear normal in the classification, as detailed in the next paragraph, are also displayed (Figure 6.5e-g).

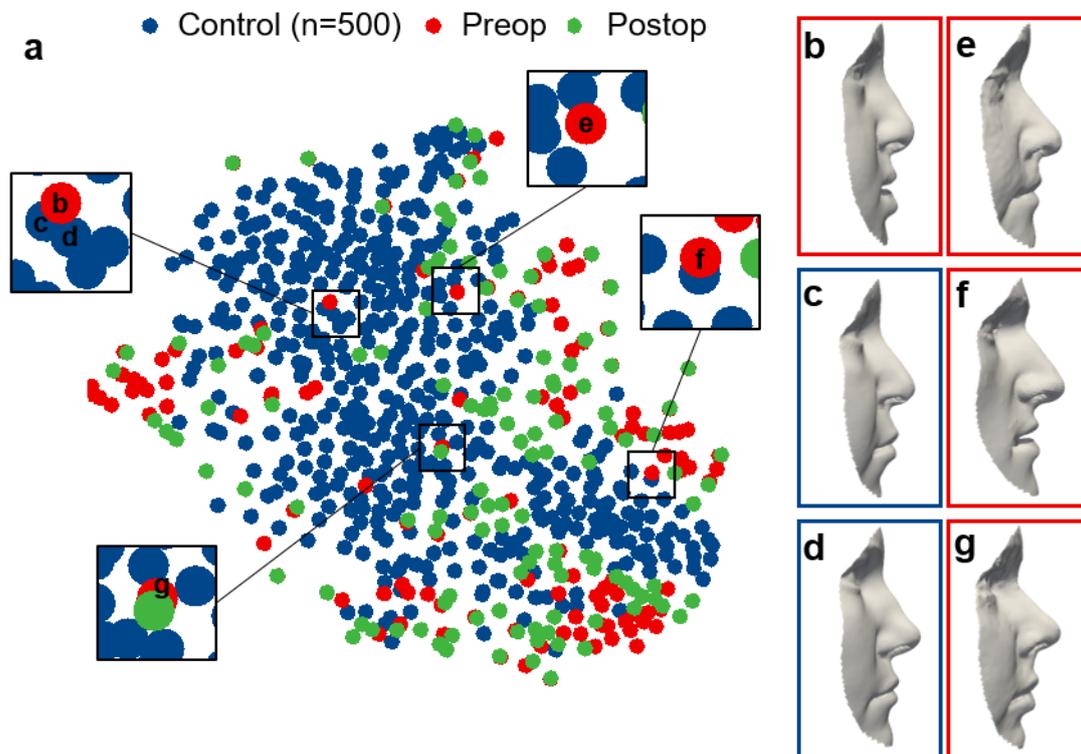


Figure 6.5 t-SNE embedding of the high-dimensional face manifold. (a) The t-SNE embedding in two dimensions was generated with randomly sampled LFSM faces, for visualisation purposes, and labelled according to LFSM (blue, $n=500$), preoperative patient (red, $n=119$), and postoperative patient (green, $n=127$) faces. Lateral views of (b) a patient and (c), (d) two close neighbours to illustrate shape similarity within the t-SNE embedding, particularly in the upper lip angle; (e-f) Faces that correspond to false negatives in the classification experiment.

6.3.4 Classification for diagnosis

Classification with a split of 80%-20% between training and testing data provided overall classification accuracy of 95.4% (Figure 6.6a). Patient faces were diagnosed with 95.5% sensitivity and 95.2% specificity and a positive and negative predictive value of 87.5% and 98.3%, respectively (Figure 6.6b). Considering false negatives – patient faces incorrectly labelled as being from the non-patient LSFM population (Figure 6.5e-g, Figure 6.7a) – 4 patients were incorrectly classified in more than 200 out of 1,000 iterations. Regarding false positives – LSFM faces incorrectly labelled as patients (Figure 6.7b) – 3 patients were incorrectly classified in more than 150 out of 1,000 iterations.

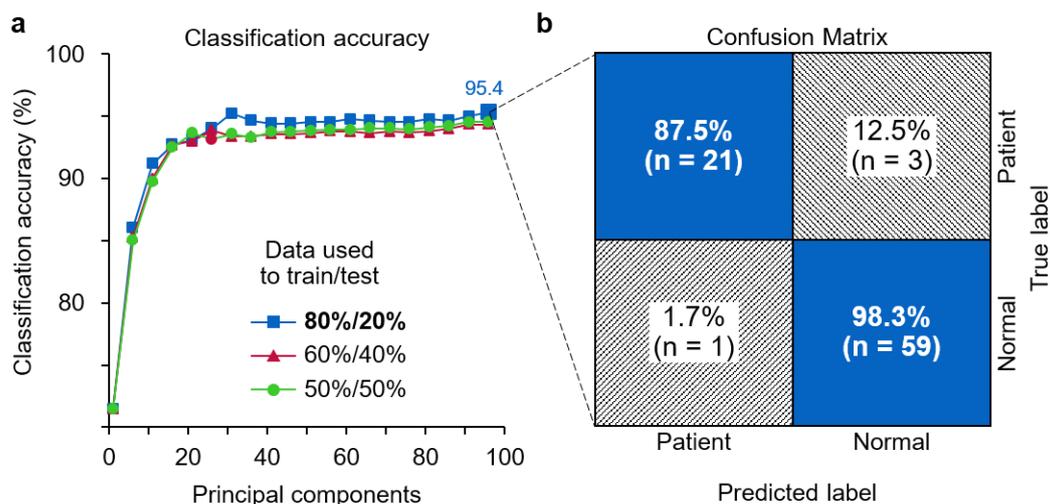


Figure 6.6 Classification of preoperative patient and non-patient LSFM faces. (a) A split of 80-20% provides 95.4% classification accuracy at 96 principal components, superior to other splits. **(b)** Average confusion matrix, obtained from classification using preoperative patient scans (n=140) and randomly selected LSFM scans (n=280), representing the average of 1,000 iterations. With an 80-20% split, patient (n=112) and LSFM (n=224) scans were used for training and patient (n=28) and LSFM (n=56) for testing. Shape abnormality was diagnosed with 95.5% sensitivity and 95.2% specificity and with a positive and negative predictive value of 87.5% and 98.3%, respectively.

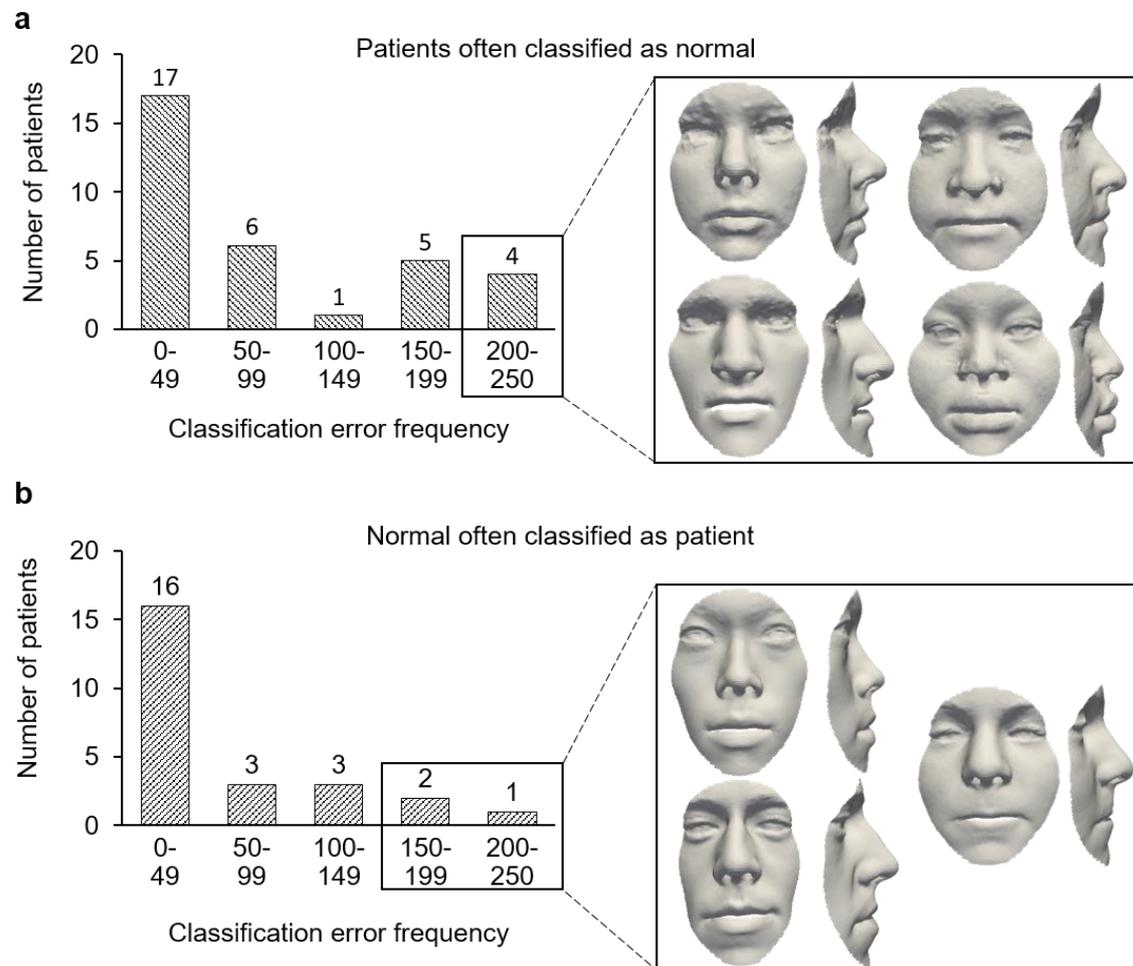


Figure 6.7 False negatives and false positives. (a) For 1,000 iterations, 33 unique patient scans were classified as a false positive and 4 of those very often (more than 200 out of 1,000), signifying orthognathic surgery are not driven by aesthetic indications alone, and (b) 25 unique LFSM scans were classified as false negatives, of which 3 often (more than 150 out of 1,000), suggestive of mild jaw malformation or untreated patients in the non-patient sample.

6.3.5 Regression for surgical simulation

To demonstrate the automated simulation of the postoperative face shape, linear regression (LR), ridge regression (RR), least-angle regression (LARS), and least absolute

shrinkage and selection operator regression (LASSO) were tested on the global model. The average error between the predicted shape and the ground-truth postoperative shape, at 100 components, was lowest with LARS (1.1 ± 0.3 mm) and RR (1.1 ± 0.3 mm), followed by LASSO (1.3 ± 0.3 mm) and LR (3.0 ± 1.2 mm) (Figure 6.8a). Using more than 40 components, LR exhibits overfitting which reduces its generalization beyond the training data. To demonstrate the quality of the patient-specific predictions, the differences between preoperative, postoperative, simulated, were visualised and quantified (Figure 6.8b). To check that predictions were indeed patient-specific rather than mimicking the population mean, all simulated faces (n=113) were additionally compared to the mean global face and mean bespoke postoperative face (Figure 6.9). At 100 components, the difference between RR simulations is much smaller compared to the postoperative 3D scan (1.1 mm, see above) than compared to the mean global face (1.8 mm) and the mean bespoke postoperative face (1.6 mm).

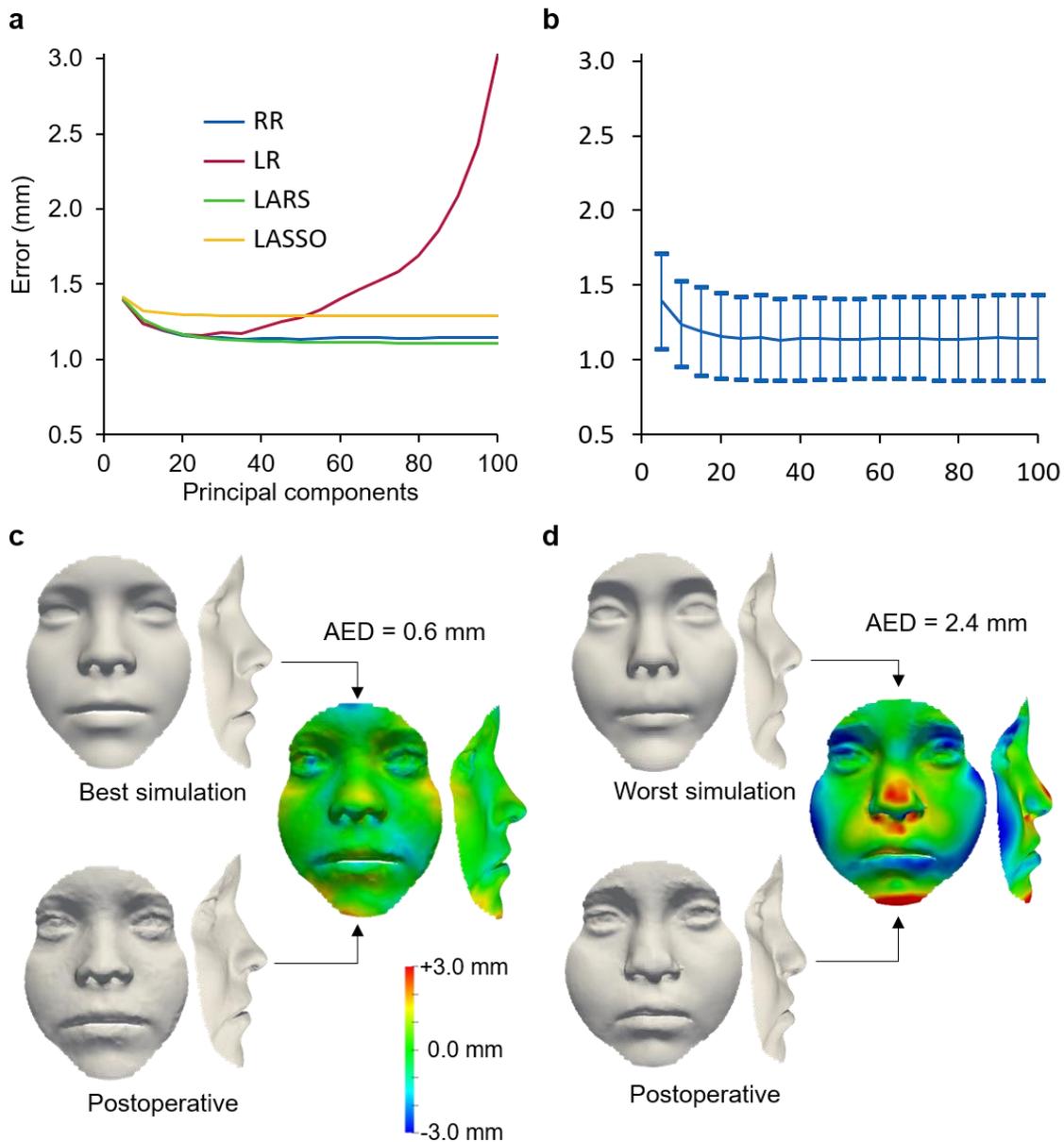


Figure 6.8 Regression for postoperative face shape simulation. (a) Overall error of the ground-truth postoperative face shape compared to the simulated shape using ridge regression (RR), linear regression (LR), least-angle regression (LARS), and least absolute shrinkage and selection operator regression (LASSO). (b) Mean and standard deviation for ridge regression. (c) The best simulated face, with an error of 0.6 mm between the simulated shape and the postoperative face shape, and (d) the worst simulated face, with an error of 2.4 mm.

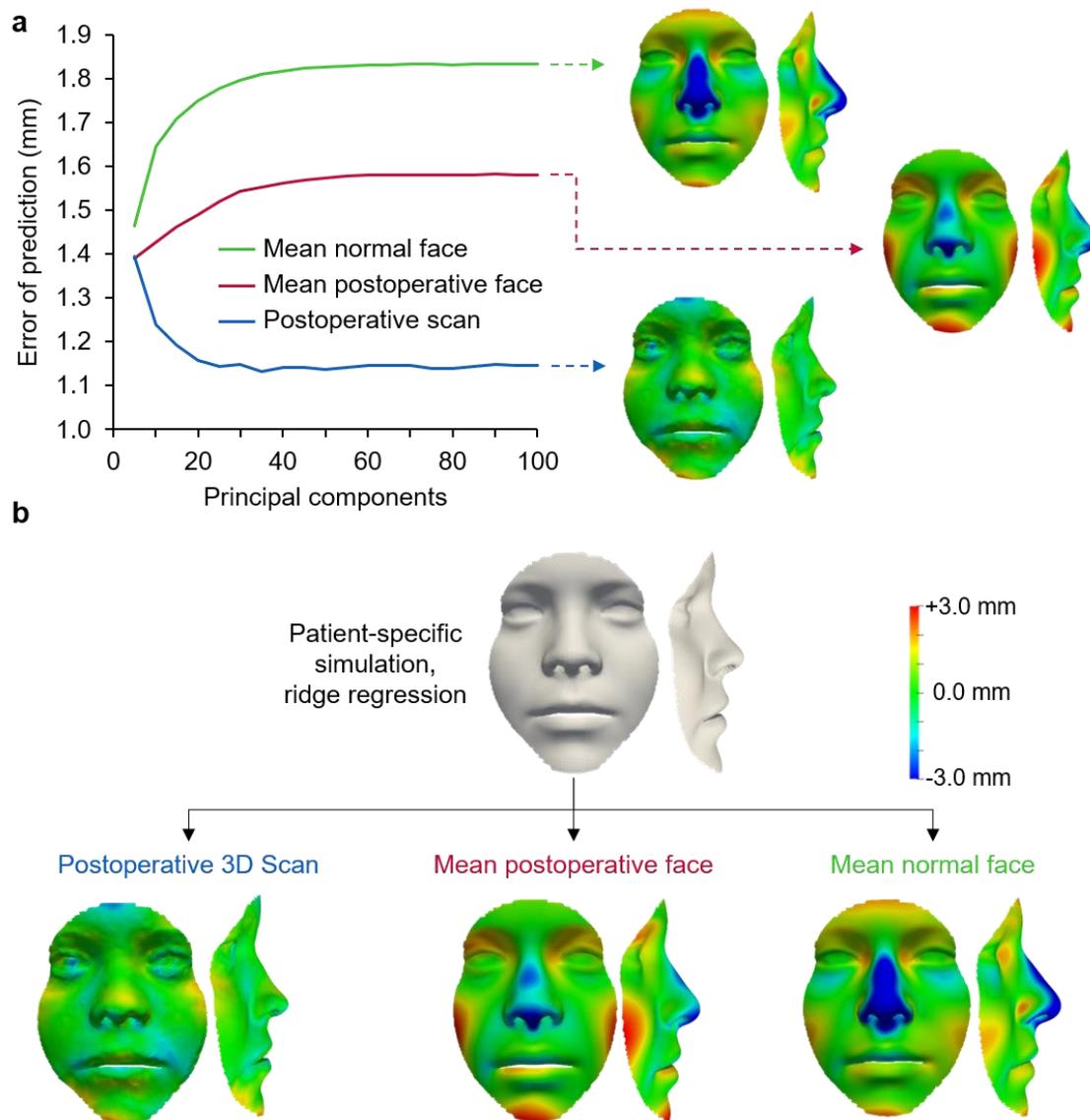


Figure 6.9 Regression error evaluation. (a) Colourmaps illustrate that the patient-specific simulations are better than predicting a face closer to the mean LSFM or the mean postoperative face; a comparison of error between the postoperative 3D scan ($n=127$) and patient-specific ridge regression simulations (1.1 ± 0.3 mm), the average postoperative face (1.6 ± 0.5 mm), and the average LSFM face (1.8 ± 0.5 mm). (b) Colourmaps elucidate how the best patient-specific simulation compares to the ground-truth postoperative face, the mean postoperative face, and the mean LSFM face.

6.4 Discussion

Although there has been great interest in the use of machine learning in plastic and reconstructive surgery, a lack of data and complex interpretability currently limit its adoption in routine clinical practice (Hashimoto *et al.*, 2018). A novel approach involving 3DMM was trained with 4,216 3D face scans – two orders of magnitude larger than previously developed clinical models (Meulstee *et al.*, 2015). Using a state-of-the-art computer vision framework (Booth *et al.*, 2018), a model was designed that comprehensively integrates high-quality 3D scans to automatically classify orthognathic patient faces and non-patient faces – an indication if someone should be seen by a specialist based on their aesthetics – and to automatically predict the patient-specific postoperative outcome. This approach can help the objective assessment of preoperative and postoperative face shape, which could be used to inform patients better during a medical consultation. Additionally, this approach provides a goal-driven surgical planning approach for the surgeon.

The machine learning approach has several important advantages over computer-assisted surgical planning with traditional methods. Conventional surgical simulation is a time-consuming explorative process in which the surgeon manually tests various procedural approaches and assesses the optimal osteotomy and bone position. The machine learning model accurately and automatically predicts the postoperative face shape and reduces the planning process to a single step. However, it does leave the surgeon to decide on the appropriate surgical procedure that delivers the simulated face shape, and an automatic method is needed that deduces the bone movements necessary to achieve a given soft tissue shape (Lubkoll, Schiela and Weiser, 2014). Alternatively, a combined soft tissue-skeletal model can be implemented

(Madsen *et al.*, 2018), however, these models would require a large number of head CT or MR images, thus renouncing the advantages 3D surface imaging has over volumetric imaging methods.

The average models showed an interesting difference between the mean normal LSFM face shape and the postoperative face shape. Additionally, the t-SNE embedding (Figure 6.5) of all 100 components showed no distinct groups which may suggest that orthognathic patient faces are similar to faces from the general population when weighing all shape features in such a face manifold. However, classification demonstrated that machine learning algorithms are able to learn and identify combinations of components that correspond to patient faces with high the sensitivity and specificity.

Whilst the operation successfully ameliorated the jaw discrepancy, some nose malformation remained postoperatively which is in line with clinical outcomes in Le Fort I advancement (Metzler *et al.*, 2014; DeSesa *et al.*, 2016; Sawh-Martinez *et al.*, 2018). Looking at classification, the false negative rate of 12.5% is undesirable as real patients would potentially be missed by the model. This can be partly attributed to the multifactorial indications for surgery, as previous studies have shown that aesthetics is the primary driver for surgery in only 71% of patients (Rivera *et al.*, 2000). Thus, shape alone may not be the main consideration for at least 12.5% of orthognathic surgery patients. Indeed, the 3DMM should be used as a machine-learning-based support tool in clinical decision-making, not to fully replace human assessment.

Whilst the false positive rate of 1.7% is low, it is substantially larger than the incidence of craniofacial anomalies, including jaw malformation (1 in 1,600 live births (Mossey and

Castilla, 2001)) and the incidence of cleft lip and palate (1 in 700 live births, with about 20% requiring an operation later in life (Good, Mulliken and Padwa, 2007)). Faces in the LSFM database were collected from the general population where subjects were not excluded for facial anomalies, mild or untreated jaw malformations. Moreover, some of the samples may comprise people that have previously had orthognathic surgery, but this data was not available. It should also be considered that orthognathic surgery has become such routine practice that some subjects may choose to undergo surgery for mild anomalies that are present within the general population, and thus no perfect binary classifier exists.

The reported accuracy of computer-assisted orthognathic surgical simulation ranges from 0.5 mm to 2.0 mm, depending on the software used (Schendel, Jacobson and Khalessi, 2013; Resnick *et al.*, 2017; De Riu *et al.*, 2018). Clinically meaningful predictions can be obtained with most commercial software, but intrinsic limitations limit its use in doctor-patient communication (chapter 5). The 3DMM performed within the range of traditional programs whilst being fully-automated, and, although the model shows high sensitivity and specificity, the results also suggest there is scope for further improvements. Recently, in computer vision and machine learning, new modelling frameworks and algorithms have been developed that achieve remarkable success in various applications. Deep Neural Networks (DNNs) (Esteva *et al.*, 2017), including Convolutional Neural Networks (CNNs) (Tewari *et al.*, 2017) and Generative Adversarial Networks (GANs) have greatly impacted and increased the performance of automatic systems designed for speech recognition, visual object detection, scene recognition, and face recognition. Whilst it is likely that these models will be used for clinical application in the near future, including computer-assisted surgical planning, their big data requirement limits its current use, and approaches that handle 3D data are still relatively

poor compared to more traditional methods like 3DMM. Therefore, the first step to improve further the performance of the model is to increase the number of scans. Generally, machine learning and artificial intelligence rely on big data for their success, but for rare diseases there are limited resources and it is often difficult to obtain access to high-quality, standardised data (Hashimoto *et al.*, 2018). Cloud-based platforms have been proposed to integrate data collection, ultimately to improve the quality of care for rare diseases (Long *et al.*, 2017). Alternatively, unsupervised methods have been tested (Genova *et al.*, 2018), although their use has not been demonstrated on patient populations. In addition, with the projection of 7.2 billion smartphone subscriptions globally in 2023 (Cerwall *et al.*, 2018), and the potential of using smartphones to capture high-quality photos and 3D scans (Muratov *et al.*, 2016), mobile devices equipped with diagnostic algorithms will play an increasingly important role in low-cost universal care (Esteva *et al.*, 2017). The 3DMM, constructed with non-ionising 3D scans, can help to accelerate this development and pave the way for shape analysis in other parts of surgery, including craniofacial and aesthetic surgery, and to replace applications that rely on CT scans (Mendoza *et al.*, 2014). A second way of increasing the performance of the model would be the integration of shape data and electronic medical records to create a multimodal machine-learning approach. This could help improve the understanding of how functional and aesthetic indications correlate to various standardised patient outcomes (Porter, Larsson and Lee, 2016) and for phenotype-genotype correlations (Tassabehji *et al.*, 2005).

6.5 Summary

In summary, this chapter described the first large-scale clinical 3DMM, a statistical model involving supervised learning, relying only on 3D surface scans. Automated image processing enabled classification, which could be used to provide a binary output whether or not someone should be referred to a specialist based on their face shape features. Second, a specialist could use these tools to automatically simulate the postoperative face shape, thus reducing the computer-assisted planning process to a single step. The striking performance of both classification and regression supports the paradigm of machine learning in clinical decision-making and computer-assisted surgical planning. Future validation of the model in larger patient cohorts and multimodal models where shape models are combined with electronic medical records could lead to a valuable new diagnostic and planning tool, ultimately facilitating low-cost care, objective treatment planning and evaluation, and safer and more precise surgery.

Compared to physical models, the significant benefits of this methodology are the fully-automated nature of the framework, which leaves the surgeon to only decide on the appropriate surgical procedure that delivers the simulated face shape, and the subjective outcome assessment and diagnosis. The main limitation relates to the use of 3D scans: whilst being a non-ionising imaging method, they carry no information on the bone anatomy. As stated, future research should focus on implementing a large-scale clinical 3DMM of bone and soft tissue, but in light of the current limitations a combined statistical-physical model is introduced in the next chapter.

**Chapter 7 COMBINED STATISTICAL
AND PHYSICAL MODELLING**

In this penultimate chapter, an integrated framework combining physical and statistical modelling is presented. The statistical model, described in chapter 6, demonstrated significant benefits over physical models; however, the use of 3D scans resulted in simulations of only the soft tissue surface, leaving the necessary bone changes to be determined. A statistical model of bone and soft tissue could be constructed, but this would require a large number of CT or MRI scans as input. Currently at GOSH and our collaborating hospitals, the ionising radiation prevents frequent scanning of patients and volunteers, and thus large databases of bone and soft tissue are not readily available. This chapter describes a framework that uses the statistical and the physical model in succession to automatically obtain a patient-specific soft tissue simulation and infer the bone movements.

7.1 Introduction

Computer-assisted surgical planning is a time-consuming technology, currently limited to specialised hospitals. This can partially be explained by the mathematical models used in commercial software – physical models simulate the soft tissue response from changes to the underlying bone, but the surgical plan needs to be manually defined. Despite long-term cost-effectiveness (Xia *et al.*, 2006), the initial barriers including time and cost investment prevent the widespread integration of such systems into clinical practice. Statistical modelling tools can help automate large parts of the computer-assisted surgical planning pipeline to address some of the above-mentioned constraints. Although statistical models originate from computer vision (Blanz and Vetter, 1999), some clinical applications have recently been postulated (Staal *et al.*, 2015; Ibrahim *et al.*, 2016; Kaya *et al.*, 2018). However, their clinical utility is limited due to manual landmark identification methods, a limited number of samples, and sole reliance

on outdated methods such as PCA. In the previous chapter, a statistical model for clinical application was developed with an automated framework, a large number of samples, and state-of-the-art algorithms. Due to the use of 3D surface data, however, only the soft tissue was simulated, and the necessary bone changes remained undetermined. Therefore, this chapter describes a proof-of-concept study of combined physical and statistical modelling in orthognathic surgery.

7.2 Methodology

7.2.1 Patient characteristics

An 18-year-old female patient who had orthognathic surgery with an isolated Le Fort I osteotomy was retrospectively selected for this proof of concept study. Indications for surgery included maxillary hypoplasia and anterior crossbite. CBCT and 3D photography were taken 92 days preoperatively and 580 days postoperatively.

7.2.2 Combined modelling framework

The combined modelling framework consisted of the statistical model developed in chapter 6, followed by the inverse application of the PFEM from chapter 4 (Figure 7.1). Specifically, after diagnosis (see chapter 6), the ridge regression is used to automatically simulate the postoperative shape outcome and PFEM is utilised in succession to predict the necessary bone changes by proving a range of bone movements. These steps ensure automation of part of the modelling pipeline whilst retaining modelling flexibility. The simulated soft tissue from 3DMM and the bone positions from PFEM were compared to the actual postoperative scans.

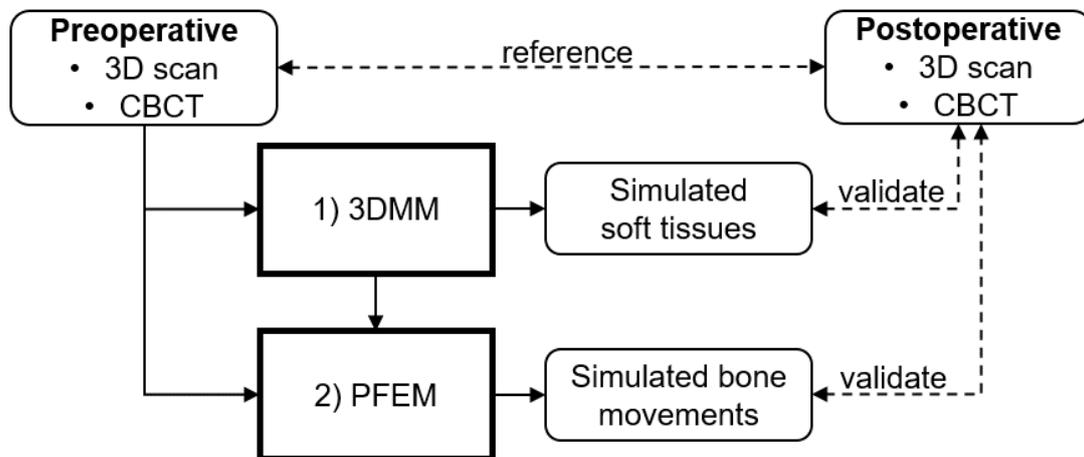


Figure 7.1 Schematic overview of the combined modelling framework. The framework comprised two steps: statistical modelling and physical modelling. In the first step, the 3DMM automatically simulated the face shape outcome. In the second step, the bone movements were obtained via inverse application of PFEM. A comparison with preoperative and postoperative CBCT and 3D scans validated the results.

The PFEM approach was set up as described in chapter 4, but rather than examining the propagation of uncertainties in the model parameters on the soft tissue outcome, a DOE optimisation with varying bone position was set up to find the best match for the 3DMM soft tissue simulation. The modelling parameters for this patient included a range of translations: horizontal (3.0, 8.0) mm, vertical (-2.0, +2.0) mm, and lateral (-4.0, +4.0) mm – based on the range of translations seen in orthognathic surgery (Table 4.1, Table 5.1). Rotations were assumed to be negligible. Additional parameter settings included population-specific variables (see chapter 4): soft tissue Young’s Modulus = 0.157 MPa, soft tissue Poisson’s ratio = 0.465, soft tissue viscoelastic relaxation = 75%, nasal cartilage Young’s modulus = 1.20 MPa, nasal cartilage Poisson’s ratio = 0.32, bone Young’s modulus = 10 GPa, bone Poisson’s ratio = 0.3. A total of 14 unique simulations were included in the DOE.

Two comparisons were made to validate the combined modelling approach: the 3DMM soft tissue simulation was compared to the postoperative 3D scan, and the simulated movements from the PFEM simulation were compared to the actual bone movements (Figure 7.1). To determine the actual bone movements, the preoperative and postoperative CBCT were compared using a custom algorithm implemented in Matlab based on the methodology of (Baan *et al.*, 2016) (Figure 7.2): bone STL files were generated from preoperative and postoperative CBCT in Simpleware ScanIP (Figure 7.2a). In Matlab, ICP was used to register the postoperative to the preoperative STL using the skull base as unaffected by the surgery (Figure 7.2b). A plane was defined on each maxilla using three landmarks: the upper incisor, mesial cusp 16 and mesial cusp 26 (Figure 7.2c). GPA was performed to determine the translation and rotation matrix that puts in correspondence the postoperative maxilla plane to the preoperative maxilla plane (Figure 7.2d). This defines the translation and rotation of the maxilla achieved in surgery.

7.2.3 Data analysis

The differences between preoperative and postoperative scans as well as the simulated face shapes from the 3DMM and the PFEM were visualised using colourmaps. The surface comparisons were computed as closest-point distance vectors, as in chapter 3 using VMTK (Antiga *et al.*, 2008) in Matlab and visualised in Paraview (Ahrens, Geveci and Law, 2005). RMS was used as quantitative measure and contour plots were used to illustrate the RMS between various simulations for a range of parameters.

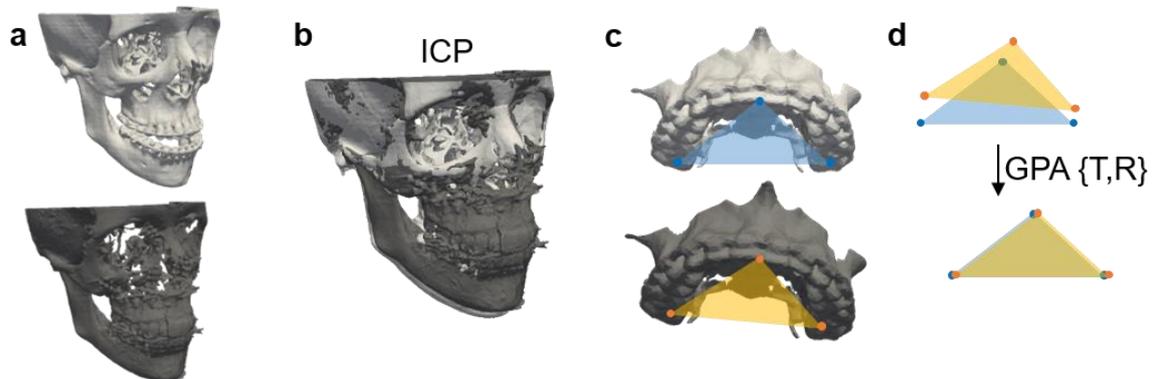


Figure 7.2 Pipeline to determine postoperative maxilla translation and rotation. (a) The preoperative (light grey) and postoperative (dark grey) STL files were imported into Matlab. (b) An iterative closest-point (ICP) algorithm was used to register the postoperative STL to the preoperative STL using the skull base as matching region. (c) Three landmarks were placed onto each maxilla to form a triangle: the upper incisor, mesial cusp 16 and mesial cusp 26; and (d) generalised Procrustes analysis (GPA) was used to place the two triangles into correspondence from which the translation (T) and rotation (R) were determined.

7.3 Results

7.3.1 Postoperative maxilla changes

Maxillary hypoplasia and anterior open bite were present in the preoperative CBCT (Figure 7.3a). Postoperatively, these issues were ameliorated, and soft tissue changes were observed in the upper lip, paranasal region and cheeks (Figure 7.3b). Translation of the maxilla was 9.1 mm anteriorly, 1.2 mm inferiorly, and 4.2 mm laterally (Figure 7.3c), and rotation was 0.5 degrees yaw, 2.0 degrees pitch, and 0.6 degrees roll (Figure 7.3d).

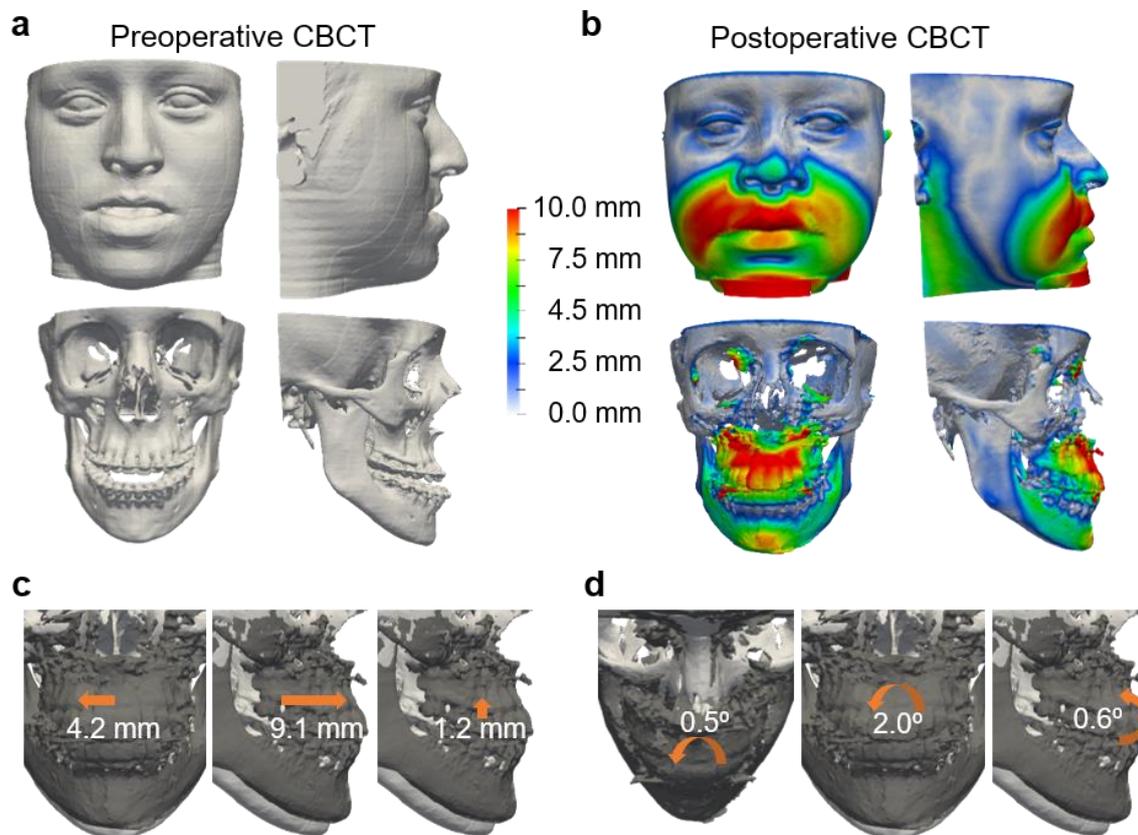


Figure 7.3 Preoperative and postoperative CBCT. (a) Preoperative CBCT, (b) colourmap of the postoperative CBCT with distances relative to the preoperative CBCT, red indicating differences up to 10 mm. (c) Translation and (d) rotation of the maxilla (dark grey).

7.3.2 Statistical model simulation

Based on the preoperative 3D scan, the statistical model automatically simulated the soft tissue outcome, with $RMS = 0.9$ mm (Figure 7.4). The colourmap illustrates how the statistical model correctly simulated the maxillary hypoplasia correction, but some errors in the nose, cheek, and lower lip regions remained.

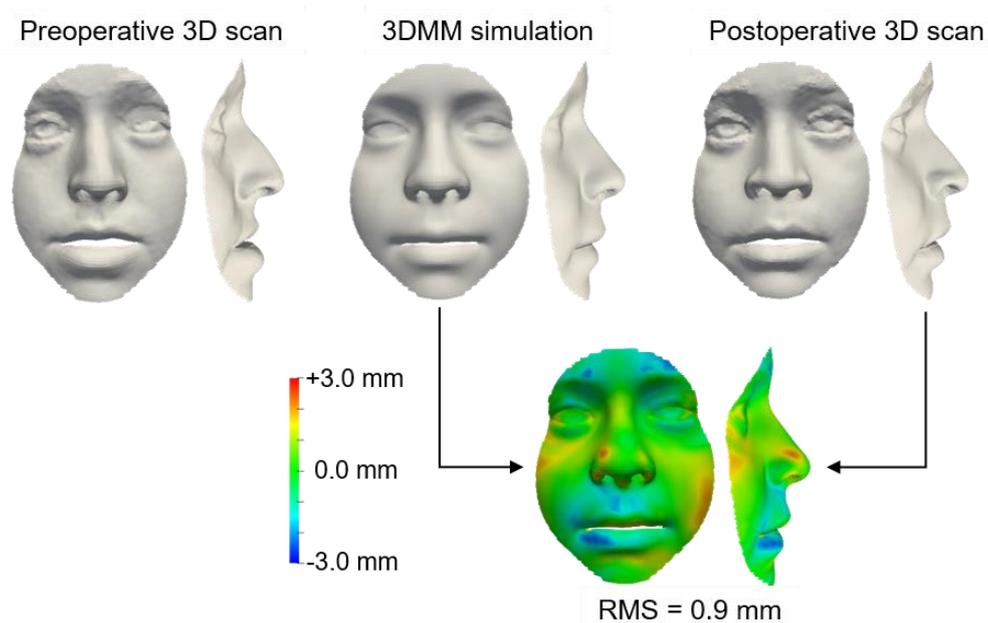


Figure 7.4 Accuracy of the 3DMM simulation. The 3DMM simulation was generated using ridge regression from the preoperative 3D scan. The colourmap illustrates the agreement between the simulation and the postoperative 3D scan with RMS = 0.9 mm.

7.3.3 Inverse physical model simulation

The DOE scheme with 14 unique simulations found that horizontal advancement of about 7.6 mm, vertical repositioning of 0 mm, and lateral repositioning of 0.5 mm provided the best match to the 3DMM simulation (Figure 7.5). RMS between the 3DMM and PFEM was 1.6 mm (Figure 7.5). The comparison between the predicted and the postoperative bone position showed good agreement horizontally and vertically, and moderate agreement laterally (Table 7.1).

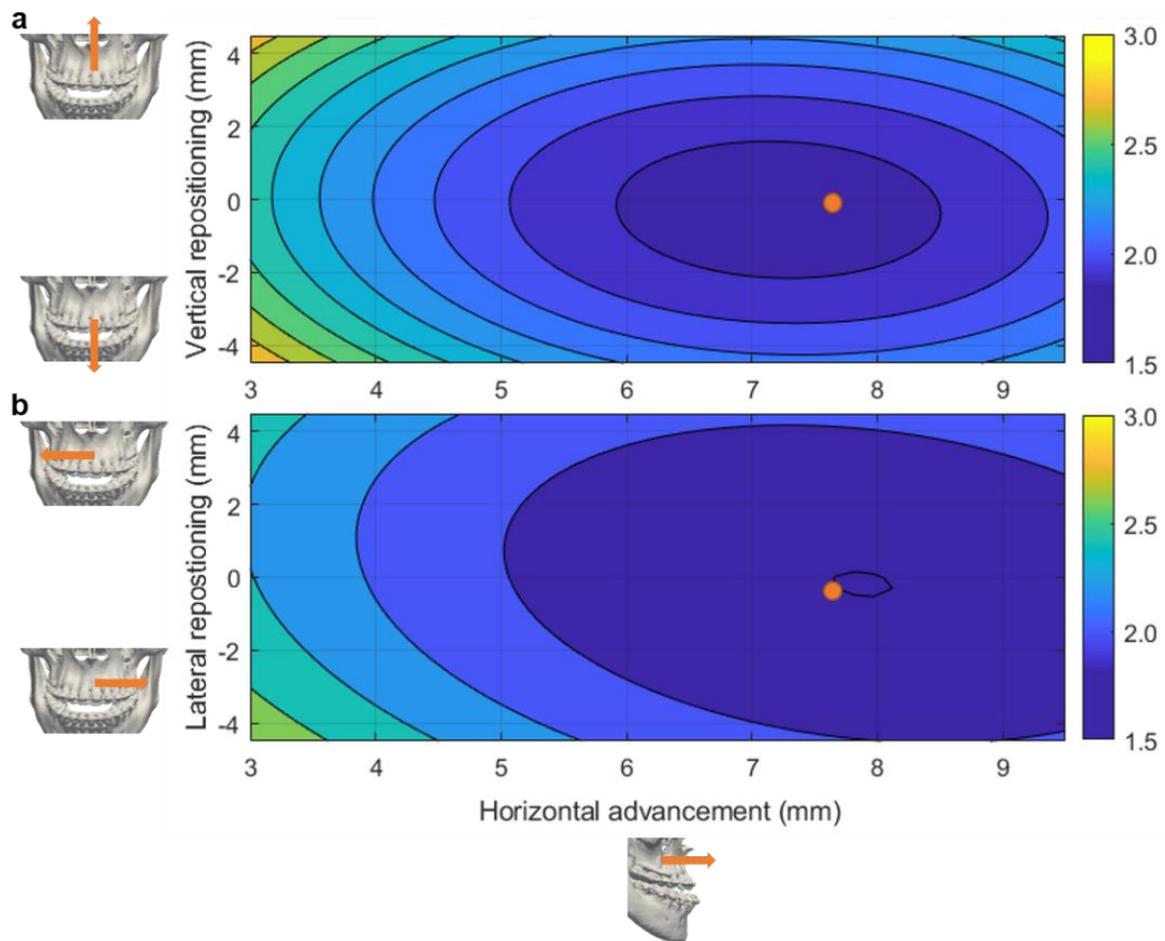


Figure 7.5 Contour plots of the agreement between PFEM and 3DMM simulations. The contour plots illustrate the overall agreement in soft tissue surface between 14 PFEM simulations and the 3DMM. **(a)** Horizontal advancement and vertical repositioning, and **(b)** horizontal advancement and lateral repositioning. The best simulation result was obtained at 7.6 mm horizontal advancement, 0 mm vertical repositioning, and 0.5 mm lateral repositioning (orange dot). The contour plot is represented by a polynomial fit between the 14 simulations, and the colour indicates the magnitude of the RMS from 1.5 mm (blue) to 3.0 mm (yellow).

Table 7.1 Comparison of the rotation and translation from postoperative CBCT and simulation.

| | Translation (mm) | | | Rotation (degrees) | | |
|-------------------|------------------|------------|------------|--------------------|------------|------------|
| | Horizontal | Vertical | Lateral | Jaw | Pitch | Roll |
| Postoperative | 9.1 | 1.2 | 4.2 | 0.5 | 2.0 | 0.6 |
| Simulation | 7.6 | 0.0 | 0.5 | 0.0* | 0.0* | 0* |
| Difference | 1.5 | 1.2 | 3.7 | 0.5 | 2.0 | 0.6 |

*Rotations were not included in the simulation and therefore corresponded to 0.0 degrees.

7.4 Discussion

Surface scans used to construct the statistical model of chapter 6 lack bone information, and large databases of CT or MR images that capture both soft and bone tissue information are not readily available. Therefore, the relationship between bone displacements and face shape changes cannot be inferred from such surface scan based statistical models alone. The proof-of-concept study presented in this chapter demonstrated how the statistical model can be employed in conjunction with PFEM to automatically obtain patient-specific soft tissue outcomes and also compute the required bone movements to achieve such soft tissue change. The framework demonstrated clinically acceptable accuracy in the horizontal and vertical translation, and moderate agreement in the lateral translation.

Some limitations must be considered. First of all, this proof-of-concept study was tested and demonstrated on a single patient; the results need to be validated in a large cohort of patients. The postoperative CBCT was taken 580 days postoperatively which ensured no swelling was present but postoperative orthodontics may have altered the dentition and thus the face shape. As the postoperative orthodontic protocol was not available, this is a limitation

of the study and future validation should consider the surgical movements and orthodontics separately. Moreover, the quantification of the postoperative changes was performed using landmarks on the teeth (Figure 7.2) – in the case of substantial postoperative orthodontic treatment this measure may not accurately represent the surgical changes of the maxilla. The lateral movement of the maxilla of 4.2 mm is much larger than previously reported values (Baan *et al.*, 2016), which may be explained by the above difference between maxillary movement and dental movements following orthodontic treatment. A postoperative CBCT scan prior to postoperative orthodontic movement would be required to accurately assess these changes.

The error between the PFEM and 3DMM, when selecting the optimal bone position, may be underestimated due to lack of sensitivity in the closest-point analysis without point-to-point correspondence (Badiali *et al.*, 2015), as also noted in chapter 5. The lack of sensitivity to the parallel movement of two surfaces mainly affects the coronal and sagittal planes and may partly explain the large error between the suggested lateral bone position and the postoperative position. Additionally, the 3DMM used a cropped oval face mask as region to calculate RMS which further limits the sensitivity in the overall face.

The combined modelling approach is shape driven – the 3DMM provides the desired shape outcome and the PFEM finds the bone movements necessary to achieve it. Whilst appearance is the main consideration for surgery in 71% of patients (Rivera *et al.*, 2000), for other patients function is the primary driver and a shape-driven approach might lead to suboptimal results. To improve the accuracy of the proposed combined mode, the individual parts should be further refined. The 3DMM could be improved by further optimising the machine learning algorithms and by including more patient data; the PFEM by mimicking the

surgical variables better and improving the material properties and the anatomical detail in the mesh. In the current model, the maxilla rotation was assumed to be negligible, which was acceptable in this case considering the rotations measured from the postoperative CBCT, but this may not be the case for many orthognathic patients considering rotation and impaction of the maxilla is often needed to achieve a satisfactory occlusion postoperatively (Steinhäuser *et al.*, 2008). Additionally, the exact location of the osteotomy was fixed, and the type of osteotomy predetermined; if this tool was to be further optimised, these should be included as parameters of the PFEM. Also, rotation of the mandible is often observed after repositioning of the maxilla but was not included in the PFEM. Moreover, potential errors in the 3DMM simulation were propagated throughout the combined model the soft tissue simulation acted as a benchmark for PFEM. Even if the PFEM simulation perfectly matched the 3DMM, an error in the bone position can be expected in the same order of magnitude as the 3DMM error. An alternative to the integration of two separate models would be a single large-scale model of bone and soft tissue, for example based on the methodology as proposed by (Madsen *et al.*, 2018). Their approach involves co-registration of two separate shape models: one for the soft tissue and one for the skeleton. The benefit is that the whole modelling framework could be automated, and that 3D surface data may be used to build a large-scale soft tissue model, but this implies the use of large datasets of CT or MRI data, yet not available.

7.5 Summary

In summary, this penultimate chapter showed the integration of the statistical model and the physical model to generate a patient-specific soft tissue simulation and suggest required bone movements. It demonstrated the high accuracy of the 3DMM to simulate a postoperative

face shape, and the PFEM was successfully applied to find the required bone movements with acceptable clinical accuracy. Future validation of this approach is necessary, as this proof-of-concept study was only demonstrated on a single patient. Moreover, additional modelling variables are to be included, such as the type of osteotomy, the osteotomy location, all six degrees of freedom, autorotation of the mandible, and secondary surgery.

In the future, in parallel to the suggested improvements above, a large-scale clinical 3DMM of bone and soft tissue should be developed but, as noted previously, large databases of CT data are not readily available because ionising radiation prevents frequent scanning of patients and volunteers. This exemplifies that further multicentre collaboration is essential for clinical validation of various computer-assisted surgical planning tools, support the development of novel algorithms, and for the wider research effort in craniomaxillofacial surgery.

Chapter 8 **CONCLUSIONS**

This chapter summarises the main findings and outcomes of this thesis, drawing from all previous chapters, and describes in detail how the application of engineering principles has contributed to personalised approaches in craniomaxillofacial surgery. The limitations are also discussed, as well as leads for future research and final remarks.

8.1 Overview

The aim of this thesis was to develop, apply and validate novel 3D computational models, and, combined with large medical image datasets, to enhance surgical planning and simulation – ultimately to improve clinical decision-making and patients’ quality of life. The first objective was to evaluate the range of 3D imaging tools at GOSH to establish which data sources can and cannot be used in surgical planning, measured by the accuracy of various imaging systems. The second objective was to improve on deterministic prediction models, which are currently the gold-standard in craniomaxillofacial surgical planning, by developing a PFEM that weighs the uncertainties in mechanical, geometric and loading properties. The third and final objective was to construct and validate a computer-assisted surgical methodology based on a statistical model, specifically a large-scale 3DMM, and demonstrate its potential in diagnosis and surgical planning.

8.2 Detailed outcomes

Chapter 3 – 3D Craniomaxillofacial imaging

Different 3D scanning systems were compared including static and hand-held systems, and systems of high and low cost. In the context of craniomaxillofacial 3D imaging, all systems

but the Structure Sensor – which lacks accuracy in areas of high curvature – can be used to capture face shape to a clinically acceptable standard. These findings imply that scans taken with different high-quality systems can be combined to form large datasets as long as the acquisition procedure is standardised, which is important for collaboration between hospitals with different scanning systems, for example for research into rare craniomaxillofacial disorders where the number of cases treated at individual hospitals is small. Whilst the Structure Sensor, the cheapest system in the comparison, was not able to capture the face shape accurately, it may be used for the analysis of head circumference and cephalic index, and the user-friendliness suggested that such systems hold great promise for the future (Beaumont *et al.*, 2017). To summarise:

1. Several different 3D scanners can be used to capture 3D face shape data, as long as they have sufficient imaging accuracy and a standardised imaging protocol is used.
2. Images from multiple 3D scanners can be used to form large datasets as long as the above requirements are met.

Chapter 4 – Probabilistic finite element modelling for surgical simulation

In chapter 4, a new probabilistic methodology was proposed for computer-assisted surgical simulation to overcome limitations of traditional deterministic models, thereby making the outcome predictions more clinically relevant. A PFEM was developed which involved a DOE approach, entailing a series of deterministic FEM with a distribution of input variables and simulations with confidence intervals. The probabilistic analysis provided insight into how various modelling parameters affect the surgical results, including the mismatch between the planned and actual bone position, and patient-specific and population-specific material

properties. The PFEM framework was developed and validated using clinical patient data and demonstrated clinically acceptable accuracy. The main outcome of this study was:

3. Probabilistic finite element modelling provides patient-specific and clinically accurate soft tissue simulations, and the range of potential outcomes can be used to aid patient communication and support the clinical decision-making process.

Chapter 5 – Clinical comparison of physical models

In chapter 5, following the introduction and proof-of-concept of PFEM in the previous chapter, the methodology was compared to commercial computer-assisted surgical planning tools, including Dolphin 3D and ProPlan CMF. Dolphin 3D makes use of a landmark-based empirical model that allows for patient-specific hard-to-soft tissue ratios. Although it can provide correct lateral (2D) simulations, the 3D simulations lack accuracy and realism. ProPlan CMF incorporates a physical model that provides accurate simulations – significantly more accurate than those made with Dolphin 3D – and the software easy to use. Despite the accurate simulations, ProPlan CMF only allows the surgeon to define the osteotomy and bone repositioning but it lacks manual settings for patient-specific material properties, which makes it less useful in the planning of extreme cases. PFEM is highly customisable and allows for patient-specific material properties to be introduced, but requires a deeper understanding of computational modelling. In the analysed cohort of patients, which had a limited sample size, PFEM simulations provided results equally accurate to those made with ProPlan CMF.

The surgical planning tools in the comparison were limited to those available at GOSH and our collaborating hospitals, but other programs are available on the market to simulate and plan surgery. This chapter demonstrated the importance of understanding how the software

works and assessing the strengths and limitations, especially if such simulations are used in patient communication or influence the clinical decision-making process. In summary, this chapter concluded that:

4. Dolphin has limited 3D soft tissue simulation validity due to its landmark-based algorithm, thus its clinical use should be limited to 2D simulation only.
5. ProPlan CMF is the best-in-class commercial simulation software; however, care must be taken when considering unusual or complex craniomaxillofacial cases, as it does not allow for bespoke parameter modelling.
6. PFEM demonstrates clinical accuracy and allows for great customisation including patient-specific parameters, but it also requires expert modelling knowledge and computational power.

Chapter 6 – 3D morphable models for automated computer-assisted diagnosis and surgical planning

Chapter 6 described the first large-scale clinical 3DMM for computer-assisted surgical simulation and diagnosis. A large number of 3D surface scans of patient and faces from the general population were included to train and validate a diagnostic classification model, which provided a binary output whether an individual should be referred to a specialist or not. Additionally, the postoperative face shape was automatically simulated via regression, reducing the computer-assisted planning process to a single step and leaving the surgeon only to decide on the appropriate surgical procedure that can deliver the simulated face shape. The striking performance of classification and regression supports the paradigm of machine learning in clinical decision-making, and may ultimately facilitate low-cost care, objective treatment planning and evaluation, and safer and more precise surgery. In summary:

7. A 3DMM – a machine-learning based model – can fully-automatically diagnose orthognathic patients based on their face shape features and generate patient-specific surgical plans as accurately as conventional computer-assisted surgical planning tools.
8. Automation and speed are the main benefits of the statistical modelling pipeline without compromising on accuracy. The model is easily scalable and adaptable to an ever-increasing amount of clinical patient data.

Chapter 7 – Combined physical and statistical modelling

Chapter 7 showed the integration of the statistical model and the physical model to generate a patient-specific soft tissue simulation and advise the surgeon on the required bone movements. This shape driven modelling approach has great flexibility as the PFEM can be modified according to specific surgical and functional requirements, whilst the goal-driven 3DMM provides an optimised outcome that acts as a benchmark for the PFEM. The results were promising and clinically accurate, but this proof-of-concept analysis was demonstrated only on a single patient and requires validation in a larger number of cases. Moreover, this chapter exemplifies that collaboration between hospitals is essential for clinical validation of various computer-assisted surgical planning approaches and to support the development of the next generation surgical planning methods. To summarise the findings in this chapter:

9. A combined 3DMM and PFEM model can generate a patient-specific surgical plan including the soft tissue outcome and the bone movements.
10. The combined approach retains the advantages of each individual framework: goal-driven automation from the 3DMM and flexibility in modifications and patient-specific optimisation through PFEM.

11. Large-scale clinical databases of bone and soft tissue are necessary to further improve and validate statistical and physical models.

8.3 Limitations and future directions

8.3.1 Sample size and data

One of the major limitations of this thesis is the relatively small sample size in the PFEM methodology (chapter 4), clinical validation of PFEM (chapter 5), and the proof-of-concept study of the combined model (chapter 7) – prospective large-cohort studies confirming the findings of these preliminary works are needed. Data from various imaging sources can be combined (chapter 3) to facilitate such large-scale validation studies and for the development of methodologies relying on large databases (chapter 6).

The small sample sizes in this thesis reflects the conservative imaging protocol at GOSH and other hospitals, where patients' exposure to ionising radiation is kept to a minimum. CT scans are taken preoperatively to aid surgical assessment and planning, but in the follow-up period CT scans are not normally taken unless complications arise. To increase the number of images without exposing patients to ionising radiation, MRI and 3D surface scanning could be more frequently used. Novel MRI acquisition sequences have been suggested as a replacement for bone CT imaging (Eley *et al.*, 2014), but the long image acquisition time, remaining technical constraints, and the need for general anaesthesia in young patients restrict the use. 3D surface scans have a fast acquisition time and high patient compliance, however, the lack of tomographic bone data limits its use to simulation of the soft tissues alone (chapter 6).

Due to the limited sample size, the variability in the patient characteristics including the type of surgery must be considered. Whilst all patients in this thesis had orthognathic surgery, the PFEM methodology (chapter 4) was developed on a cohort of bimaxillary patients who had a Le Fort I osteotomy and BSSO whilst the comparison study (chapter 5) was performed on a cohort of patients with a single-jaw Le Fort I osteotomy. The PFEM was developed without modelling the BSSO which may have induced errors in the simulated soft tissues, however, the high accuracy of PFEM in chapter 5 suggests that this simplification did not substantially affect the accuracy of the methodology. As stated previously, the PFEM should be further validated and optimised for specific types of operations and patient subsets, for which larger patient datasets are needed.

Lastly, only basic patient characteristics such as age and gender were included. In the future, multimodal models should be created that combine shape models with clinically relevant information from electronic medical records and patient reported outcome measures (PROMs) including, morbidity, quality of life, speech function and cost. Such multimodal models may help elucidate the relationship between 3D shape features and clinically relevant psychosocial and functional outcomes, and ultimately support the value based healthcare paradigm (Porter, Larsson and Lee, 2016).

8.3.2 Error evaluation, clinical accuracy and subjectivity

Throughout this thesis, the error has been evaluated as the agreement between two surfaces. The closest point distances were calculated, and RMS or the percentage within 2 mm was used to determine the accuracy of a simulation. However, as noted previously, the usefulness of using such an average measure can be limited for the assessment specific

landmarks (Kouchi *et al.*, 2012). Another issue with RMS without point-to-point correspondence, as noted in chapter 5 and 7, is the lack of sensitivity to the parallel movements of two surfaces in the coronal and sagittal plane, which may lead to error underestimation (Badiali *et al.*, 2015). It remains computationally challenging to develop a methodology that puts complex meshes into dense correspondence – this would be a substantial piece of work in its own right. Moreover, the 2 mm value used to indicate clinical accuracy (Aung, Ngim and Lee, 1995) is an empirical norm and, to the best of my knowledge, there are no published studies that have investigated the influence of small shape differences on the perception of face shape and texture. The cut-off value does not account for the fact that, clinically, a change of 2 mm in one area of the face may have a lesser impact than 2 mm in another area of the face, and, in particular to the untrained eye, shape changes much larger than 2 mm may go unnoticed. Thus, in the context of craniomaxillofacial surgery, studies on the relationship between face shape changes and perception are needed to remove some of the subjectivity. The difference between clinical significance and statistical significance was demonstrated in chapter 5, where the landmark-based model in Dolphin 3D failed to provide 3D simulations that were clinically relevant, due to the landmark-based algorithm. On the contrary, PFEM and ProPlan CMF provided clinically useful 3D simulations; however, the differences in accuracy with Dolphin were not statistically significant. This is a limitation of using an average error value and highlights the importance of using quantitative and qualitative measures in conjunction. Moreover, this may help explain why clinical consensus on the accuracy of computer-assisted surgical tools has not yet been reached. Thus, besides perception studies, standardised protocols are needed to evaluate and compare outcomes, such as those proposed by (Baan *et al.*, 2016).

8.4 Conclusions

This thesis presented a multidisciplinary approach to computer-assisted surgical planning involving clinical patient data, medical image analysis, engineering methods, and state-of-the-art machine learning and computer vision algorithms. The methodologies developed and validated in this thesis are relevant for clinical, academic, and commercial applications. Clinically, these methods could be extended beyond planning of orthognathic procedures: commercial software is already used in other domains within plastic and reconstructive surgery and in orthopaedic surgery. Future lines of research were suggested, including the prospective use of the probabilistic finite element method as well as further improvements to the model. Moreover, the 3DMM has opened up a plethora of applications – commercially and academically – and this thesis showed that machine-learning-based methods not only hold great promise, but also can readily be developed, and have an impact in clinical decision-making and surgical planning.

To conclude, this thesis described the development and validation of novel computer-assisted surgical planning methodologies based on probabilistic finite element modelling and on 3D morphable models. These computational tools could lead to future improvements in surgical planning software by making the simulations more insightful and clinically relevant, and by automating and streamlining the process. These developments will make such technology more accessible and provide intelligent clinical decision-making support, ultimately to aid surgeons to improve patients' quality of life.

REFERENCES

- Aboul-Hosn Centenero, S. and Hernández-Alfaro, F. (2012) '3D planning in orthognathic surgery: CAD/CAM surgical splints and prediction of the soft and hard tissues results - Our experience in 16 cases', *Journal of Cranio-Maxillofacial Surgery*, 40(2), pp. 162–168. doi: 10.1016/j.jcms.2011.03.014.
- Ahrens, J., Geveci, B. and Law, C. (2005) 'ParaView: An end-user tool for large-data visualization', in *Visualization Handbook*. doi: 10.1016/B978-012387582-2/50038-1.
- Almukhtar, A., Khambay, B., Ju, X., McDonald, J. and Ayoub, A. (2017) 'Accuracy of generic mesh conformation: The future of facial morphological analysis', *JPRAS Open*, 14, pp. 39–48. doi: 10.1016/j.jpra.2017.08.003.
- Amberg, B., Knothe, R. and Vetter, T. (2008) 'Expression invariant 3D face recognition with a morphable model', in *2008 8th IEEE International Conference on Automatic Face and Gesture Recognition, FG 2008*. doi: 10.1109/AFGR.2008.4813376.
- Amberg, B., Romdhani, S. and Vetter, T. (2007) 'Optimal step nonrigid ICP algorithms for surface registration', in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. doi: 10.1109/CVPR.2007.383165.
- American Society of Plastic Surgeons (2017) 'Reconstructive plastic surgery statistics', in *Plastic Surgery Statistics Report*, p. 10.
- Antiga, L., Piccinelli, M., Botti, L., Ene-Iordache, B., Remuzzi, A. and Steinman, D. A. (2008) 'An image-based modeling framework for patient-specific computational hemodynamics', *Medical and Biological Engineering and Computing*, 46(11), pp. 1097–112. doi: 10.1007/s11517-008-0420-1.
- Anwar, M. and Harris, M. (1990) 'Model surgery for orthognathic planning', *British Journal of Oral and Maxillofacial Surgery*, 28(6), pp. 393–397. doi: 10.1016/0266-4356(90)90037-L.
- Armstrong, R. A. (2014) 'When to use the Bonferroni correction', *Ophthalmic and Physiological Optics*, 34(5), pp. 502–508. doi: 10.1111/opo.12131.
- Aung, S. C., Ngim, R. C. K. and Lee, S. T. (1995) 'Evaluation of the laser scanner as a surface

- measuring tool and its accuracy compared with direct facial anthropometric measurements', *British Journal of Plastic Surgery*, 48(8), pp. 551–558. doi: 10.1016/0007-1226(95)90043-8.
- Aynechi, N., Larson, B. E., Leon-Salazar, V. and Beiraghi, S. (2011) 'Accuracy and precision of a 3D anthropometric facial analysis with and without landmark labeling before image acquisition', *Angle Orthodontist*, 81(2), pp. 245–252. doi: 10.2319/041810-210.1.
- Baan, F., Liebrechts, J., Xi, T., Schreurs, R., De Koning, M., Bergé, S. and Maal, T. (2016) 'A new 3D tool for assessing the accuracy of bimaxillary surgery: The OrthoGnathicanAlyser', *PLoS ONE*, 11(2), p. e0149625. doi: 10.1371/journal.pone.0149625.
- Badiali, G., Roncari, A., Bianchi, A., Taddei, F., Marchetti, C. and Schileo, E. (2015) 'Navigation in Orthognathic Surgery: 3D Accuracy', *Facial Plastic Surgery*, 31(5), pp. 463–473. doi: 10.1055/s-0035-1564716.
- Barbarino, G. G., Jabareen, M. and Mazza, E. (2009) 'Experimental and numerical study of the relaxation behavior of facial soft tissue', *Proceedings in Applied Mathematics and Mechanics*, 9(1), pp. 87–90. doi: 10.1002/pamm.200910023.
- Barbarino, G. G., Jabareen, M., Trzewik, J., Nkengne, A., Stamatias, G. and Mazza, E. (2009) 'Development and Validation of a Three-Dimensional Finite Element Model of the Face', *Journal of Biomechanical Engineering*, 131(4), p. 041006. doi: 10.1115/1.3049857.
- Barbarino, G., Jabareen, M. and Mazza, E. (2011) 'Experimental and numerical study on the mechanical behavior of the superficial layers of the face', *Skin Research and Technology*, 17(4), pp. 434–444. doi: 10.1111/j.1600-0846.2011.00515.x.
- Barel, A. O., Lambrecht, R. and Clarys, P. (2004) 'Mechanical Function of the Skin: State of the Art', in *Skin Bioengineering*, pp. 69–83. doi: 10.1159/000060577.
- Bathe, K.-J. (2014) *Finite Element Procedures*. K.J. Bathe, Watertown, MA.
- Beaumont, C. A. A., Knoops, P. G. M., Borghi, A., Jeelani, N. U. O., Koudstaal, M. J., Schievano, S., Dunaway, D. J. and Rodriguez-Florez, N. (2017) 'Three-dimensional

- surface scanners compared with standard anthropometric measurements for head shape', *Journal of Cranio-Maxillofacial Surgery*, 45(6), pp. 921–927. doi: 10.1016/j.jcms.2017.03.003.
- Beltran, D. and Basañez, L. (2014) 'A comparison between active and passive 3D vision sensors: BumblebeeXB3 and microsoft kinect', in *Advances in Intelligent Systems and Computing*, pp. 725–734. doi: 10.1007/978-3-319-03413-3_54.
- Bennett, C. C. and Hauser, K. (2013) 'Artificial intelligence framework for simulating clinical decision-making: A Markov decision process approach', *Artificial Intelligence in Medicine*, 57(1), pp. 9–19. doi: 10.1016/j.artmed.2012.12.003.
- Benson, B. W., Flint, D. J., Liang, H. and Opatowsky, M. J. (2014) 'Advances in Diagnostic Imaging for Pathologic Conditions of the Jaws', *Head and Neck Pathology*, 8(4), pp. 383–391. doi: 10.1007/s12105-014-0575-z.
- Berthaume, M. a., Dechow, P. C., Iriarte-Díaz, J., Ross, C. F., Strait, D. S., Wang, Q. and Grosse, I. R. (2012) 'Probabilistic finite element analysis of a craniofacial finite element model', *Journal of Theoretical Biology*, 300, pp. 242–253. doi: 10.1016/j.jtbi.2012.01.031.
- Blanz, V. (2006) 'Face recognition based on a 3D morphable model', in *FGR 2006: Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition*. doi: 10.1109/FGR.2006.42.
- Blanz, V. and Vetter, T. (1999) 'A morphable model for the synthesis of 3D faces', in *Proceedings of the 26th annual conference on Computer graphics and interactive techniques - SIGGRAPH '99*. doi: 10.1145/311535.311556.
- Bolkart, T. and Wuhler, S. (2015) 'A groupwise multilinear correspondence optimization for 3D faces', in *Proceedings of the IEEE International Conference on Computer Vision*. doi: 10.1109/ICCV.2015.411.
- Booth, J., Roussos, A., Ponniah, A., Dunaway, D. and Zafeiriou, S. (2018) 'Large Scale 3D Morphable Models', *International Journal of Computer Vision*, 126(2–4), pp. 233–254.

doi: 10.1007/s11263-017-1009-7.

- Booth, J., Roussos, A., Zafeiriou, S., Ponniah, A. and Dunaway, D. (2016) ‘A 3D Morphable Model Learnt from 10,000 Faces’, in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. doi: 10.1109/CVPR.2016.598.
- Burke, P. H., Banks, P., Beard, L. F. H., Tee, J. E. and Hughes, C. (1983) ‘Stereophotographic measurement of change in facial soft tissue morphology following surgery’, *British Journal of Oral Surgery*, 21(4), pp. 237–245. doi: 10.1016/0007-117X(83)90012-4.
- Cerwall, P., Lundvall, A., Jonsson, P., Carson, S. and Moller, R. (2018) *Ericsson Mobility Report 2018*. Available at: <https://www.ericsson.com/assets/local/mobility-report/documents/2018/ericsson-mobility-report-june-2018.pdf>.
- Cevidane, L. H. C., Tucker, S., Styner, M., Kim, H., Chapuis, J., Reyes, M., Proffit, W., Turvey, T. and Jaskolka, M. (2010) ‘Three-dimensional surgical simulation’, *American Journal of Orthodontics and Dentofacial Orthopedics*, 138(3), pp. 361–371. doi: 10.1016/j.ajodo.2009.08.026.
- Chabanas, M., Luboz, V. and Payan, Y. (2003) ‘Patient specific finite element model of the face soft tissues for computer-assisted maxillofacial surgery’, *Medical Image Analysis*, 7(2), pp. 131–151. doi: 10.1016/S1361-8415(02)00108-1.
- Chabanas, M., Marécaux, C., Payan, Y. and Boutault, F. (2002) ‘Computer aided planning for orthognatic surgery’, in *CARS 2002 Computer Assisted Radiology and Surgery*, pp. 988–993. doi: 10.1007/978-3-642-56168-9_165.
- Chabanas, M. and Payan, Y. (2000) ‘A 3D Finite Element Model of the Face for Simulation in Plastic and Maxillo-Facial Surgery’, in *MICCAI 2000*, pp. 1068–1075. doi: 10.1007/978-3-540-40899-4_111.
- Cortes, C. and Vapnik, V. (1995) ‘Support-Vector Networks’, *Machine Learning*, 20(3), pp. 273–297. doi: 10.1023/A:1022627411411.
- Cotin, S., Delingette, H. and Ayache, N. (1999) ‘Real-time elastic deformations of soft tissues for surgery simulation’, *IEEE Transactions on Visualization and Computer Graphics*,

5(1), pp. 62–73. doi: 10.1109/2945.764872.

Cotin, S., Delingette, H. and Ayache, N. (2000) ‘Hybrid elastic model for real-time cutting, deformations, and force feedback for surgery training and simulation’, *Visual Computer*. doi: 10.1007/PL00007215.

Courant, R. (1943) ‘Variational methods for the solution of problems of equilibrium and vibrations’, *Bulletin of the American Mathematical Society*, 49(1), pp. 1–23. doi: 10.1090/S0002-9904-1943-07818-4.

Crombag, G. A. J. C., Verdoorn, M. H. A. S., Nikkhah, D., Ponniah, A. J. T., Ruff, C. and Dunaway, D. (2014) ‘Assessing the corrective effects of facial bipartition distraction in Apert syndrome using geometric morphometrics’, *Journal of Plastic, Reconstructive and Aesthetic Surgery*, 67(6), pp. e151–e161. doi: 10.1016/j.bjps.2014.02.019.

Dai, H., Pears, N., Smith, W. and Duncan, C. (2017) ‘A 3D Morphable Model of Craniofacial Shape and Texture Variation’, in *Proceedings of the IEEE International Conference on Computer Vision*. doi: 10.1109/ICCV.2017.335.

Delingette, H. (1998) ‘Toward realistic soft-tissue modeling in medical simulation’, *Proceedings of the IEEE*, 86(3), pp. 512–523. doi: 10.1109/5.662876.

DeSesa, C. R., Metzler, P., Sawh-Martinez, R. and Steinbacher, D. M. (2016) ‘Three-dimensional Nasolabial Morphologic Alterations Following Le Fort I’, *Plastic and Reconstructive Surgery Global Open*, 4(8), p. e848. doi: 10.1097/GOX.0000000000000685.

Deuffhard, P., Weiser, M. and Zachow, S. (2006) ‘Mathematics in Facial Surgery’, *Notes of the AMS*, 53(9), pp. 1012–1016.

Dvortsin, D. P., Sandham, A., Pruijm, G. J. and Dijkstra, P. U. (2008) ‘A comparison of the reproducibility of manual tracing and on-screen digitization for cephalometric profile variables’, *European Journal of Orthodontics*, 30(6), pp. 586–591. doi: 10.1093/ejo/cjn041.

Edler, R., Abd Rahim, M., Wertheim, D. and Greenhill, D. (2010) ‘The use of facial

- anthropometrics in aesthetic assessment’, *Cleft Palate-Craniofacial Journal*, 47(1), pp. 48–57. doi: 10.1597/08-218.1.
- Eley, K. A., Watt-Smith, S. R., Sheerin, F. and Golding, S. J. (2014) “‘Black Bone’ MRI: a potential alternative to CT with three-dimensional reconstruction of the craniofacial skeleton in the diagnosis of craniosynostosis’, *European Radiology*, 24(10), pp. 2417–26. doi: 10.1007/s00330-014-3286-7.
- Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M. and Thrun, S. (2017) ‘Dermatologist-level classification of skin cancer with deep neural networks’, *Nature*, 542, pp. 115–8. doi: 10.1038/nature21056.
- Farkas, L. G. (1996) ‘Accuracy of Anthropometric Measurements: Past, Present, and Future’, *The Cleft Palate-Craniofacial Journal*, 33(1), pp. 10–22. doi: 10.1597/1545-1569_1996_033_0010_aoampp_2.3.co_2.
- Fonseca, C. M. and Fleming, P. J. (1993) ‘Genetic Algorithms for Multiobjective Optimization: Formulation, Discussion and Generalization’, in *Proceedings of the Fifth International Conference on Genetic Algorithms*, pp. 416–423.
- Geng, Z. J. (1996) ‘Rainbow three-dimensional camera: New concept of high-speed three-dimensional vision systems’, *Optical Engineering*, 35(2), pp. 376–383. doi: 10.1117/1.601023.
- Genova, K., Cole, F., Maschinot, A., Sarna, A., Vlastic, D. and Freeman, W. T. (2018) ‘Unsupervised Training for 3D Morphable Model Regression’, in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 8377–8386. doi: 10.1109/CVPR.2018.00874.
- Gilardino, M. S., Chen, E. and Bartlett, S. P. (2009) ‘Choice of internal rigid fixation materials in the treatment of facial fractures.’, *Craniofacial trauma & reconstruction*, 2(1), pp. 49–60. doi: 10.1055/s-0029-1202591.
- Gilchrist, M. D., Murphy, J. G., Parnell, W. and Pierrat, B. (2014) ‘Modelling the slight compressibility of anisotropic soft tissue’, *International Journal of Solids and Structures*,

51(23–24), pp. 3857–3865. doi: 10.1016/j.ijsostr.2014.06.018.

Giovanoli, P., Tzou, C. H. J., Ploner, M. and Frey, M. (2003) ‘Three-dimensional video-analysis of facial movements in healthy volunteers’, *British Journal of Plastic Surgery*, 56(7), pp. 644–652. doi: 10.1016/S0007-1226(03)00277-7.

Gladilin, E., Ivanov, A. and Roginsky, V. (2004) ‘Generic approach for biomechanical simulation of typical boundary value problems in cranio-maxillofacial surgery planning’, in *Medical Image Computing and Computer Assisted Intervention–MICCAI 2004*, pp. 380–388.

Gladilin, E., Zachow, S., Deuflhard, P. and Hege, H.-C. (2002) ‘Adaptive nonlinear elastic FEM for realistic prediction of soft tissue in craniofacial surgery simulations’, in *Proceedings of SPIE - The International Society for Optical Engineering*, pp. 1–8. doi: 10.1117/12.466906.

Gladilin, E., Zachow, S., Deuflhard, P. and Hege, H.-C. (2003) ‘Realistic prediction of individual facial emotion expressions for craniofacial surgery simulations’, in *Proceedings of SPIE - The International Society for Optical Engineering*, pp. 520–527. doi: 10.1117/12.479584.

Good, P. M., Mulliken, J. B. and Padwa, B. L. (2007) ‘Frequency of Le Fort I osteotomy after repaired cleft lip and palate or cleft palate’, *Cleft Palate-Craniofacial Journal*, 44(4), pp. 396–401. doi: 10.1597/06-075.1.

Gøthesen, Ø., Slover, J., Havelin, L., Askildsen, J. E., Malchau, H. and Furnes, O. (2013) ‘An economic model to evaluate cost-effectiveness of computer assisted knee replacement surgery in Norway’, *BMC Musculoskeletal Disorders*, 14(202). doi: 10.1186/1471-2474-14-202.

Grellmann, W., Berghaus, A., Haberland, E. J., Jamali, Y., Holweg, K., Reincke, K. and Bierögel, C. (2006) ‘Determination of strength and deformation behavior of human cartilage for the definition of significant parameters’, *Journal of Biomedical Materials Research - Part A*, 78(1), pp. 168–174. doi: 10.1002/jbm.a.30625.

-
- Hammond, P. (2007) 'The use of 3D face shape modelling in dysmorphology', *Archives of Disease in Childhood*, 92(12), pp. 1120–1126. doi: 10.1136/adc.2006.103507.
- Hashimoto, D. A., Rosman, G., Rus, D. and Meireles, O. R. (2018) 'Artificial Intelligence in Surgery: Promises and Perils', *Annals of Surgery*, 268(1), pp. 70–6. doi: 10.1097/SLA.0000000000002693.
- Heike, C. L., Cunningham, M. L., Hing, A. V., Stuhaug, E. and Starr, J. R. (2009) 'Picture perfect? Reliability of craniofacial anthropometry using three-dimensional digital stereophotogrammetry', *Plastic and Reconstructive Surgery*, 124(4), pp. 1261–1272. doi: 10.1097/PRS.0b013e3181b454bd.
- Heike, C. L., Upson, K., Stuhaug, E. and Weinberg, S. M. (2010) '3D digital stereophotogrammetry: a practical guide to facial image acquisition', *Head & Face Medicine*, 6(18). doi: 10.1186/1746-160X-6-18.
- Honrado, C. P. and Larrabee, W. F. (2004) 'Update in three-dimensional imaging in facial plastic surgery', *Current Opinion in Otolaryngology & Head and Neck Surgery*, 12(4), pp. 327–331. doi: 10.1097/01.moo.0000130578.12441.99.
- Horgan, T. J. and Gilchrist, M. D. (2003) 'The creation of three-dimensional finite element models for simulating head impact biomechanics', *International Journal of Crashworthiness*, 8(4), pp. 353–366. doi: 10.1533/ijcr.2003.0243.
- Hrennikoff, A. (1941) 'Solution of problems of elasticity by framework method', *Journal of Applied Mechanics*, 8(4), pp. 169–175.
- Huber, P., Hu, G., Tena, R., Mortazavian, P., Koppen, W. P., Christmas, W. J., Rättsch, M. and Kittler, J. (2016) 'A Multiresolution 3D Morphable Face Model and Fitting Framework', in *Proceedings of the 11th Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*. doi: 10.5220/0005669500790086.
- Ibrahim, A., Suttie, M., Bulstrode, N. W., Britto, J. A., Dunaway, D., Hammond, P. and Ferretti, P. (2016) 'Combined soft and skeletal tissue modelling of normal and dysmorphic midface postnatal development', *Journal of Cranio-Maxillofacial Surgery*, 44(11), pp.

1777–85. doi: 10.1016/j.jcms.2016.08.020.

Jachowicz, J., McMullen, R. and Prettypaul, D. (2007) ‘Indentometric analysis of in vivo skin and comparison with artificial skin models’, *Skin Research and Technology*, 13(3), pp. 299–309. doi: 10.1111/j.1600-0846.2007.00229.x.

Janakiraman, N., Feinberg, M., Vishwanath, M., Nalaka Jayaratne, Y. S., Steinbacher, D. M., Nanda, R. and Uribe, F. (2015) ‘Integration of 3-dimensional surgical and orthodontic technologies with orthognathic “surgery-first” approach in the management of unilateral condylar hyperplasia’, *American Journal of Orthodontics and Dentofacial Orthopedics*, 148(6), pp. 1054–1066. doi: 10.1016/j.ajodo.2015.08.012.

Johnson, C. (1987) *Numerical solutions of partial differential equations by the finite element method*. Cambridge University Press.

Kanevsky, J., Corban, J., Gaster, R., Kanevsky, A., Lin, S. and Gilardino, M. (2016) ‘Big Data and Machine Learning in Plastic Surgery: A New Frontier in Surgical Innovation’, *Plastic and Reconstructive Surgery*, 137(5), p. 890e–897e. doi: 10.1097/PRS.0000000000002088.

Kau, C. H., Cronin, A. J. and Richmond, S. (2007) ‘A three-dimensional evaluation of postoperative swelling following orthognathic surgery at 6 months’, *Plastic and Reconstructive Surgery*, 119(7), pp. 2192–2199. doi: 10.1097/01.prs.0000260707.99001.79.

Kaya, O., Pluijmers, B. I., Staal, F., Ruff, C., Padwa, B. L., Koudstaal, M. J. and Dunaway, D. J. (2018) ‘Describing the mandible in patients with craniofacial microsomia based on principal component analysis and thin plate spline video analysis’, *International Journal of Oral and Maxillofacial Surgery*, 48(3), pp. 302–308. doi: 10.1016/j.ijom.2018.08.015.

Keeve, E., Girod, S., Kikinis, R. and Girod, B. (1998) ‘Deformable modeling of facial tissue for craniofacial surgery simulation’, *Computer Aided Surgery*, 3(5), pp. 228–238. doi: 10.3109/10929089809149844.

Kharbanda, O. P. and Wadhawan, N. (2013) ‘Orthodontic Diagnosis and Treatment Planning:

- Collaborating with Medical and Other Dental Specialists’, in *Integrated Clinical Orthodontics*. Blackwell Publishing Ltd. doi: 10.1002/9781118702901.ch3.
- Kim, H., Jürgens, P. and Reyes, M. (2011) ‘Soft-Tissue Simulation for Cranio-Maxillofacial Surgery: Clinical Needs and Technical Aspects’, in *Studies in Mechanobiology, Tissue Engineering and Biomaterials*, pp. 413–440. doi: 10.1007/8415_2011_105.
- Kittler, J., Huber, P., Feng, Z. H., Hu, G. and Christmas, W. (2016) ‘3D morphable face models and their applications’, in *Lecture Notes in Computer Science*. doi: 10.1007/978-3-319-41778-3_19.
- Ko, J. M. (2016) ‘Genetic syndromes associated with craniosynostosis’, *Journal of Korean Neurosurgical Society*, 59(3), pp. 187–191. doi: 10.3340/jkns.2016.59.3.187.
- Koch, R. M., Gross, M. H., Carls, F. R., von Büren, D. F., Fankhauser, G. and Parish, Y. I. H. (1996) ‘Simulating facial surgery using finite element models’, in *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques - SIGGRAPH '96*. doi: 10.1145/237170.237281.
- Kouchi, M., Mochimaru, M., Bradtmiller, B., Daanen, H., Li, P., Nacher, B. and Nam, Y. (2012) ‘A protocol for evaluating the accuracy of 3D body scanners’, *Work*, 41(Suppl 1), pp. 4010–4017. doi: 10.3233/WOR-2012-0064-4010.
- Kovacs, L., Zimmermann, A., Brockmann, G., Baurecht, H., Schwenzer-Zimmerer, K., Papadopoulos, N. A., Papadopoulos, M. A., Sader, R., Biemer, E. and Zeilhofer, H. F. (2006) ‘Accuracy and precision of the three-dimensional assessment of the facial surface using a 3-D laser scanner’, *IEEE Transactions on Medical Imaging*, 25(6), pp. 742–754. doi: 10.1109/TMI.2006.873624.
- Kretschmer, W. B., Zoder, W., Baciut, G., Bacuit, M. and Wangerin, K. (2009) ‘Accuracy of maxillary positioning in bimaxillary surgery’, *British Journal of Oral and Maxillofacial Surgery*, 47(6), pp. 446–469. doi: 10.1016/j.bjoms.2009.06.004.
- Krimmel, M., Schuck, N., Bacher, M. and Reinert, S. (2011) ‘Facial surface changes after cleft alveolar bone grafting’, *Journal of Oral and Maxillofacial Surgery*, 69(1), pp. 80–83. doi:

10.1016/j.joms.2010.03.009.

Kuijpers-Jagtman, A. M. (2013) 'Cleft Lip and Palate: Role of the Orthodontist in the Interdisciplinary Management Team', in *Integrated Clinical Orthodontics*. Blackwell Publishing Ltd, pp. 153–167. doi: 10.1002/9781118702901.ch9.

Lee, H.-J., Suh, H.-Y., Lee, Y.-S., Lee, S.-J., Donatelli, R. E., Dolce, C. and Wheeler, T. T. (2014) 'A better statistical method of predicting postsurgery soft tissue response in Class II patients', *The Angle Orthodontist*, 84(4), pp. 322–328. doi: 10.2319/050313-338.1.

Lee, S. (2004) 'Three-dimensional photography and its application to facial plastic surgery', *Archives of Facial Plastic Surgery*, 6(6), pp. 410–414. doi: 10.1001/archfaci.6.6.410.

Lee, S. J., Liong, K., Tse, K. M. and Lee, H. P. (2010) 'Biomechanics of the deformity of septal L-struts', *Laryngoscope*, 120(8), pp. 1508–1515. doi: 10.1002/lary.20976.

Lee, Y., Terzopoulos, D. and Walters, K. (1995) 'Realistic modeling for facial animation', in *Proceedings of the 22nd annual conference on Computer graphics and interactive techniques - SIGGRAPH '95*, pp. 55–62. doi: 10.1145/218380.218407.

Lekakis, G., Claes, P., Hamilton, G. S. and Hellings, P. W. (2016) 'Three-Dimensional Surface Imaging and the Continuous Evolution of Preoperative and Postoperative Assessment in Rhinoplasty', *Facial Plastic Surgery*. doi: 10.1055/s-0035-1570122.

Liebrechts, J., Xi, T., Timmermans, M., De Koning, M., Bergé, S., Hoppenreijns, T. and Maal, T. (2015) 'Accuracy of three-dimensional soft tissue simulation in bimaxillary osteotomies', *Journal of Cranio-Maxillofacial Surgery*, 43(3), pp. 329–335. doi: 10.1016/j.jcms.2014.12.012.

Long, E., Lin, H., Liu, Z., Wu, X., Wang, L., Jiang, J., An, Y., Lin, Z., Li, X., Chen, J., Li, J., Cao, Q., Wang, D., Liu, X., Chen, W. and Liu, Y. (2017) 'An artificial intelligence platform for the multihospital collaborative management of congenital cataracts', *Nature Biomedical Engineering*, 1(2), p. 0024. doi: 10.1038/s41551-016-0024.

Van Loon, B., Van Heerbeek, N., Bierenbroodspot, F., Verhamme, L., Xi, T., De Koning, M. J. J., Ingels, K. J. A. O., Bergé, S. J. and Maal, T. J. J. (2015) 'Three-dimensional changes

- in nose and upper lip volume after orthognathic surgery’, *International Journal of Oral and Maxillofacial Surgery*, 44(1), pp. 83–89. doi: 10.1016/j.ijom.2014.08.001.
- Van Loon, B., Maal, T. J., Plooij, J. M., Ingels, K. J., Borstlap, W. A., Kuijpers-Jagtman, A. M., Spauwen, P. H. and Bergé, S. J. (2010) ‘3D Stereophotogrammetric assessment of pre- and postoperative volumetric changes in the cleft lip and palate nose’, *International Journal of Oral and Maxillofacial Surgery*, 39(6), pp. 534–540. doi: 10.1016/j.ijom.2010.03.022.
- Lübbers, H. T., Medinger, L., Kruse, A., Grätz, K. W. and Matthews, F. (2010) ‘Precision and accuracy of the 3dmd photogrammetric system in craniomaxillofacial application’, *Journal of Craniofacial Surgery*, 21(3), pp. 763–767. doi: 10.1097/SCS.0b013e3181d841f7.
- Lubkoll, L., Schiela, A. and Weiser, M. (2014) ‘An Optimal Control Problem in Polyconvex Hyperelasticity’, *SIAM Journal on Control and Optimization*, 52(3), pp. 1403–22. doi: 10.1137/120876629.
- Luboz, V., Chabanas, M., Swider, P. and Payan, Y. (2005) ‘Orbital and maxillofacial computer aided surgery: Patient-specific finite element models to predict surgical outcomes’, *Computer Methods in Biomechanics and Biomedical Engineering*, 8(4), pp. 259–265. doi: 10.1080/10255840500289921.
- Luboz, V., Promayon, E. and Payan, Y. (2014) ‘Linear Elastic Properties of the Facial Soft Tissues Using an Aspiration Device: Towards Patient Specific Characterization’, *Annals of Biomedical Engineering*, 42(11), pp. 2369–2378. doi: 10.1007/s10439-014-1098-1.
- Maas, B. D. P. J., Pluijmers, B. I., Knoops, P. G. M., Ruff, C., Koudstaal, M. J. and Dunaway, D. (2018) ‘Using principal component analysis to describe the midfacial deformities in patients with craniofacial microsomia’, *Journal of Cranio-Maxillofacial Surgery*, 46(12), pp. 2032–2041. doi: 10.1016/j.jcms.2018.09.019.
- Van Der Maaten, L. and Hinton, G. (2008) ‘Visualizing Data using t-SNE’, *Journal of Machine Learning Research*, 9, pp. 2579–2605. doi: 10.1007/s10479-011-0841-3.

- Mada, S. K., Smith, M. L., Smith, L. N. and Midha, P. S. (2003) 'Overview of passive and active vision techniques for hand-held 3D data acquisition', in *Opto-Ireland 2002: Optical Metrology, Imaging, and Machine Vision*, pp. 16–27. doi: 10.1117/12.463773.
- Madsen, D., Lüthi, M., Schneider, A. and Vetter, T. (2018) 'Probabilistic Joint Face-Skull Modelling for Facial Reconstruction', *cvpr*. doi: 10.1109/CVPR.2018.00555.
- Mangado, N., Piella, G., Noailly, J., Pons-Prats, J. and Ballester, M. Á. G. (2016) 'Analysis of Uncertainty and Variability in Finite Element Computational Models for Biomedical Engineering: Characterization and Propagation', *Frontiers in Bioengineering and Biotechnology*, 4(85). doi: 10.3389/fbioe.2016.00085.
- Marchetti, C., Bianchi, A., Muyldermans, L., Di Martino, M., Lancellotti, L. and Sarti, A. (2011) 'Validation of new soft tissue software in orthognathic surgery planning', *International Journal of Oral and Maxillofacial Surgery*, 40(1), pp. 26–32. doi: 10.1016/j.ijom.2010.09.004.
- Mazza, E. and Barbarino, G. G. (2011) '3D Mechanical Modeling of Facial Soft Tissue for Surgery Simulation', *Facial Plastic Surgery Clinics of North America*, 19(4), pp. 623–637. doi: 10.1016/j.fsc.2011.07.006.
- Mazzoni, S., Badiali, G., Lancellotti, L., Babbi, L., Bianchi, A. and Marchetti, C. (2010) 'Simulation-guided navigation: A new approach to improve intraoperative three-dimensional reproducibility during orthognathic surgery', *Journal of Craniofacial Surgery*, 21(6), pp. 1698–1705. doi: 10.1097/SCS.0b013e3181f3c6a8.
- Mazzoni, S., Bianchi, A., Schiariti, G., Badiali, G. and Marchetti, C. (2015) 'Computer-aided design and computer-aided manufacturing cutting guides and customized titanium plates are useful in upper maxilla waferless repositioning', *Journal of Oral and Maxillofacial Surgery*, 73(4), pp. 701–707. doi: 10.1016/j.joms.2014.10.028.
- Mendoza, C. S., Safdar, N., Okada, K., Myers, E., Rogers, G. F. and Linguraru, M. G. (2014) 'Personalized assessment of craniosynostosis via statistical shape modeling', *Medical Image Analysis*, 18(4), pp. 635–46. doi: 10.1016/j.media.2014.02.008.

-
- Metzler, P., Geiger, E. J., Chang, C. C., Sirisoontorn, I. and Steinbacher, D. M. (2014) 'Assessment of three-dimensional nasolabial response to le Fort I advancement', *Journal of Plastic, Reconstructive and Aesthetic Surgery*, 67(6), pp. 756–63. doi: 10.1016/j.bjps.2014.03.023.
- Meulstee, J., Liebregts, J., Xi, T., Vos, F., De Koning, M., Bergé, S. and Maal, T. (2015) 'A new 3D approach to evaluate facial profile changes following BSSO', *Journal of Cranio-Maxillofacial Surgery*, 43(10), pp. 1994–9. doi: 10.1016/j.jcms.2015.08.007.
- Miettinen, K. (2008) 'Introduction to multiobjective optimization: Noninteractive approaches', in *Lecture Notes in Computer Science*, pp. 1–26. doi: 10.1007/978-3-540-88908-3-1.
- Miettinen, K. M. (1998) *Nonlinear Multiobjective Optimization*. Kluwer Academic Publishers, Boston.
- Mirnezami, R. and Ahmed, A. (2018) 'Surgery 3.0, artificial intelligence and the next-generation surgeon', *British Journal of Surgery*, 105(5), pp. 463–465. doi: 10.1002/bjs.10860.
- Mollard, B., Lavallée, S. and Bettega, G. (1998) 'Computer assisted orthognathic surgery', in *Medical Image Computing and Computer-Assisted Intervention - MICCAI '98*, pp. 21–28.
- Mollemans, W. (2003) 'Tetrahedral mass spring model for fast soft tissue deformation', in *Surgery Simulation and Tissue Modeling*, pp. 145–154. doi: 10.1007/3-540-45015-7_14.
- Mollemans, W., Schutyser, F., Nadjmi, N., Maes, F. and Suetens, P. (2007) 'Predicting soft tissue deformations for a maxillofacial surgery planning system: From computational strategies to a complete clinical validation', *Medical Image Analysis*, 11(3), pp. 282–301. doi: 10.1016/j.media.2007.02.003.
- Mossey, P. and Castilla, E. (2001) 'Craniofacial anomalies and associated birth defects', in *Global registry and database on craniofacial anomalies*, pp. 15–33. Available at: <http://apps.who.int/iris/bitstream/handle/10665/42840/9241591102.pdf?sequence=1&isAllowed=y>.

- Mundluru, T., Almkhatar, A., Ju, X. and Ayoub, A. (2017) 'The accuracy of three-dimensional prediction of soft tissue changes following the surgical correction of facial asymmetry: An innovative concept', *International Journal of Oral and Maxillofacial Surgery*, 46(11), pp. 1517–1524. doi: 10.1016/j.ijom.2017.04.017.
- Muratov, O., Slynko, Y., Chernov, V., Lyubimtseva, M., Shamsuarov, A. and Bucha, V. (2016) '3DCapture: 3D Reconstruction for a Smartphone', in *IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 893–900. doi: 10.1109/CVPRW.2016.116.
- Nadjmi, N., Tehranchi, A., Azami, N., Saedi, B. and Mollemans, W. (2013) 'Comparison of soft-tissue profiles in le Fort I osteotomy patients with Dolphin and Maxilim softwares', *American Journal of Orthodontics and Dentofacial Orthopedics*, 144(5), pp. 654–662. doi: 10.1016/j.ajodo.2013.06.019.
- Naini, F. B. (2016) 'Historical Evolution of Orthognathic Surgery', in *Orthognathic Surgery*, pp. 23–82. doi: 10.1002/9781119004370.ch2.
- Naini, F. B. (2017) 'Hugo L. Obwegeser (1920–2017) – The father of modern orthognathic surgery', *Journal of Orthodontics*, 44(4), pp. 317–319. doi: 10.1080/14653125.2017.1392064.
- Nikkhah, D., Ponniah, A., Ruff, C. and Dunaway, D. (2013) 'Planning surgical reconstruction in treacher-collins syndrome using virtual simulation', *Plastic and Reconstructive Surgery*, 132(5), p. 790e–805e. doi: 10.1097/PRS.0b013e3182a48d33.
- Obwegeser, H. L. (2016) 'Introduction: Orthognathic Surgery – A Life ' s Work', in *Orthognathic Surgery: Principles, Planning and Practice*, pp. 1–20.
- Occipital Inc. (2016) *Structure Sensor - Capture the world in 3D*. Available at: <http://structure.io/> (Accessed: 6 July 2016).
- Olate, S., Zaror, C. and Mommaerts, M. Y. (2017) 'A systematic review of soft-to-hard tissue ratios in orthognathic surgery. Part IV: 3D analysis – Is there evidence?', *Journal of Cranio-Maxillofacial Surgery*. doi: 10.1016/j.jcms.2017.05.013.

-
- Pan, B., Zhang, G., Xia, J. J., Yuan, P., Ip, H. H. S., He, Q., Lee, P. K. M., Chow, B. and Zhou, X. (2016) ‘Prediction of soft tissue deformations after CMF surgery with incremental kernel ridge regression’, *Computers in Biology and Medicine*, 75, pp. 1–9. doi: 10.1016/j.compbiomed.2016.04.020.
- Papadopoulos, M. A., Christou, P. K., Athanasiou, A. E., Boettcher, P., Zeilhofer, H. F., Sader, R. and Papadopoulos, N. A. (2002) ‘Three-dimensional craniofacial reconstruction imaging’, *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontics*, 93(4), pp. 382–393. doi: 10.1067/moe.2002.121385.
- Pataky, T. C., Koseki, M. and Cox, P. G. (2016) ‘Probabilistic biomechanical finite element simulations: whole-model classical hypothesis testing based on upcrossing geometry’, *PeerJ Computer Science*, 2, p. e96. doi: 10.7717/peerj-cs.96.
- Patete, P., Eder, M., Raith, S., Volf, A., Kovacs, L. and Baroni, G. (2013) ‘Comparative assessment of 3D surface scanning systems in breast plastic and reconstructive surgery’, *Surgical Innovation*, 20(5), pp. 509–515. doi: 10.1177/1553350612463443.
- Payan, Y. (2017) ‘Soft tissue finite element modeling and calibration of the material properties in the context of computer-assisted medical interventions’, in *CISM International Centre for Mechanical Sciences, Courses and Lectures*, pp. 133–144. doi: 10.1007/978-3-319-45071-1_6.
- Paysan, P., Knothe, R., Amberg, B., Romdhani, S. and Vetter, T. (2009) ‘A 3D Face Model for Pose and Illumination Invariant Face Recognition’, in *2009 Sixth IEEE International Conference on Advanced Video and Signal Based Surveillance*, pp. 296–301. doi: 10.1109/AVSS.2009.58.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Pedregosa, F., Gramfort, A., Thirion, B., Prettenhofer, P., Vanderplas, J., Brucher, M., Perrot, M. and Duchesnay, E. (2011) ‘Scikit-learn: Machine Learning in Python’, *Journal of Machine Learning Research*, 12, pp. 2825–2830. doi: 10.1007/s13398-014-0173-7.2.
- Peiró, J. and Sherwin, S. (2005) ‘Finite Difference , Finite Element and Finite Volume Methods

for Partial Differential’, in *Handbook of Materials Modeling*, pp. 2415–2446.

Peterman, R. J., Jiang, S., Johe, R. and Mukherjee, P. M. (2016) ‘Accuracy of Dolphin visual treatment objective (VTO) prediction software on class III patients treated with maxillary advancement and mandibular setback’, *Progress in Orthodontics*, 17, p. 19. doi: 10.1186/s40510-016-0132-2.

Picinbono, G., Delingette, H. and Ayache, N. (2000) ‘Real-time large displacement elasticity for surgery simulation: Non-linear tensor-mass model’, in *Medical Image Computing and Computer-Assisted Intervention - MICCAI*, pp. 643–652. doi: 10.1007/978-3-540-40899-4_66.

Pieper, S. D., Laub, D. R. and Rosen, J. M. (1995) ‘A finite-element facial model for simulating plastic surgery’, *Plastic and Reconstructive Surgery*, 96(5), pp. 1100–1105. doi: 10.1097/00006534-199510000-00014.

Pieper, S., Rosen, J. and Zeltzer, D. (1992) ‘Interactive graphics for plastic surgery: A task-level analysis and implementation’, in *Proceedings of the 1992 symposium on Interactive 3D graphics*, pp. 127–134. doi: <http://doi.acm.org/10.1145/147156.147180>.

Pluijmers, B. I., Ponniah, A. J. T., Ruff, C. and Dunaway, D. (2012) ‘Using principal component analysis to describe the Apert skull deformity and simulate its correction’, *Journal of Plastic, Reconstructive and Aesthetic Surgery*, 65(12), pp. 1750–1752. doi: 10.1016/j.bjps.2012.07.007.

Porter, M. E., Larsson, S. and Lee, T. H. (2016) ‘Standardizing Patient Outcomes Measurement’, *New England Journal of Medicine*, 374(6), pp. 504–6. doi: 10.1056/NEJMp1511701.

Reh, S., Beley, J.-D., Mukherjee, S. and Khor, E. H. (2006) ‘Probabilistic finite element analysis using ANSYS’, *Structural Safety*, 28(1–2), pp. 17–43. doi: 10.1016/j.strusafe.2005.03.010.

Resnick, C. M., Dang, R. R., Glick, S. J. and Padwa, B. L. (2017) ‘Accuracy of three-dimensional soft tissue prediction for Le Fort I osteotomy using Dolphin 3D software: a

- pilot study', *International Journal of Oral and Maxillofacial Surgery*. International Association of Oral and Maxillofacial Surgeons, 46(3), pp. 289–295. doi: 10.1016/j.ijom.2016.10.016.
- Richmon, J. D., Sage, A., Wong, V. W., Chen, A. C., Sah, R. L. and Watson, D. (2006) 'Compressive biomechanical properties of human nasal septal cartilage', *American Journal of Rhinology*, 20(5), pp. 496–501. doi: 10.2500/ajr.2006.20.2932.
- Ricketts, R. M. (1972) 'The value of cephalometrics and computerized technology', *Angle Orthodontist*, 42(3), pp. 179–199. doi: 10.1043/0003-3219(1972)042<0179:TVOCAC>2.0.CO;2.
- De Riu, G., Viridis, P. I., Meloni, S. M., Lumbau, A. and Vaira, L. A. (2018) 'Accuracy of computer-assisted orthognathic surgery', *Journal of Cranio-Maxillofacial Surgery*, 46(2), pp. 293–298. doi: 10.1016/j.jcms.2017.11.023.
- Rivera, S. M., Hatch, J. P., Dolce, C., Bays, R. A., Van Sickels, J. E. and Rugh, J. D. (2000) 'Patients' own reasons and patient-perceived recommendations for orthognathic surgery', *American Journal of Orthodontics and Dentofacial Orthopedics*, 118(2), pp. 134–141. doi: 10.1067/mod.2000.107010.
- Rodriguez-Florez, N., Göktekin, Ö. K., Bruse, J. L., Borghi, A., Angullia, F., Knoops, P. G. M., Tenhagen, M., O'Hara, J. L., Koudstaal, M. J., Schievano, S., Jeelani, N. U. O., James, G. and Dunaway, D. J. (2017) 'Quantifying the effect of corrective surgery for trigonocephaly: A non-invasive, non-ionizing method using three-dimensional handheld scanning and statistical shape modelling', *Journal of Cranio-Maxillofacial Surgery*, 45(3), pp. 387–394. doi: 10.1016/j.jcms.2017.01.002.
- Rupin, F., Saïed, A., Dalmas, D., Peyrin, F., Hauptert, S., Raum, K., Barthel, E., Boivin, G. and Laugier, P. (2009) 'Assessment of microelastic properties of bone using scanning acoustic microscopy: A face-to-face comparison with nanoindentation', *Japanese Journal of Applied Physics*, 48(7S), p. 07GK01. doi: 10.1143/JJAP.48.07GK01.
- San Miguel Moragas, J., Van Cauteren, W. and Mommaerts, M. Y. (2014) 'A systematic review on soft-to-hard tissue ratios in orthognathic surgery part I: Maxillary repositioning

- osteotomy', *Journal of Cranio-Maxillofacial Surgery*, 42(7), pp. 1341–1351. doi: 10.1016/j.jcms.2014.03.024.
- San Miguel Moragas, J., Oth, O., Büttner, M. and Mommaerts, M. Y. (2015) 'A systematic review on soft-to-hard tissue ratios in orthognathic surgery part II: Chin procedures', *Journal of Cranio-Maxillofacial Surgery*, 43(8), pp. 1530–1540. doi: 10.1016/j.jcms.2015.07.032.
- Sarti, A., Gori, R. and Lamberti, C. (1999) 'A physically based model to simulate maxillo-facial surgery from 3D CT images', *Future Generation Computer Systems*, 15(2), pp. 217–221.
- Sarti, A., Lamberti, C., Gori, R., Erbacci, G., Bassani, L., Bianchi, A. and Marchetti, C. (2007) 'Virtual planning of facial reconstructions', *Imaging Decisions MRI*, 11(1), pp. 29–38. doi: 10.1111/j.1617-0830.2007.00089.x.
- Sawh-Martinez, R., Lin, A. M., DeSesa, C. R., Wu, R. T., Gary, C. S. and Steinbacher, D. M. (2018) 'Clockwise and Counterclockwise Le Fort I Movements Influence Nasolabial Morphology Differently', *Plastic and Reconstructive Surgery*, 142(6), pp. 1572–1581. doi: 10.1097/PRS.0000000000004988.
- Schaaf, H., Wilbrand, J. F., Boedeker, R. H. and Howaldt, H. P. (2010) 'Accuracy of photographic assessment compared with standard anthropometric measurements in nonsynostotic cranial deformities', *Cleft Palate-Craniofacial Journal*, 47(5), pp. 447–453. doi: 10.1597/09-026.
- Schendel, S. A. (2015) 'Computer simulation in the daily practice of orthognathic surgery', *International Journal of Oral and Maxillofacial Surgery*, 44(12), pp. 1451–1456. doi: 10.1016/j.ijom.2015.05.022.
- Schendel, S. A., Jacobson, R. and Khalessi, S. (2013) '3-dimensional facial simulation in orthognathic surgery: Is it accurate?', *Journal of Oral and Maxillofacial Surgery*, 71(8), pp. 1406–14. doi: 10.1016/j.joms.2013.02.010.
- Schendel, S. A. and Lane, C. (2009) '3D Orthognathic Surgery Simulation Using Image

- Fusion', *Seminars in Orthodontics*, 15(1), pp. 48–56. doi: 10.1053/j.sodo.2008.09.012.
- Schendel, S., Hazan-Molina, H., Rachmiel, A. and Aizenbud, D. (2012) 'The Future in Craniofacial Surgery: Computer-Assisted Planning', *Rambam Maimonides Medical Journal*, 3(2), p. e0012. doi: 10.5041/RMMJ.10079.
- Da Silveira, A. C., Daw, J. L., Kusnoto, B., Evans, C. and Cohen, M. (2003) 'Craniofacial applications of three-dimensional laser surface scanning', *Journal of Craniofacial Surgery*, 14(4), pp. 449–456. doi: 10.1097/00001665-200307000-00009.
- Soh, C. L. and Narayanan, V. (2013) 'Quality of life assessment in patients with dentofacial deformity undergoing orthognathic surgery — A systematic review', *International Journal of Oral and Maxillofacial Surgery*, 42(8), pp. 974–980. doi: 10.1016/j.ijom.2013.03.023.
- Staal, F. C. R., Ponniah, A. J. T., Angullia, F., Ruff, C., Koudstaal, M. J. and Dunaway, D. (2015) 'Describing Crouzon and Pfeiffer syndrome based on principal component analysis', *Journal of Cranio-Maxillofacial Surgery*, 43(4), pp. 528–36. doi: 10.1016/j.jcms.2015.02.005.
- Stefanou, G. (2009) 'The stochastic finite element method: Past, present and future', *Computer Methods in Applied Mechanics and Engineering*, 198(9–12), pp. 1031–1051. doi: 10.1016/j.cma.2008.11.007.
- Stegmann, M. B. and Delgado, D. (2002) 'A Brief Introduction to Statistical Shape Analysis', *Informatics and Mathematical Modelling*, p. 15. doi: 10.1007/s00464-012-2330-4.
- Steinhäuser, E. W. (1996) 'Historical development of orthognathic surgery.', *Journal of cranio-maxillo-facial surgery: official publication of the European Association for Cranio-Maxillo-Facial Surgery*, 24(4), pp. 195–204. doi: 10.1016/S1010-5182(96)80002-3.
- Steinhäuser, S., Richter, U., Richter, F., Bill, J. and Rudzki-Janson, I. (2008) 'Profile changes following maxillary impaction and autorotation of the mandible', *Journal of Orofacial orthopedics*, 69(1), pp. 31–41. doi: 10.1007/s00056-008-0723-8.

- Stokbro, K., Aagaard, E., Torkov, P., Bell, R. B. and Thygesen, T. (2014) 'Virtual planning in orthognathic surgery', *International Journal of Oral and Maxillofacial Surgery*, 43(8). doi: 10.1016/j.ijom.2014.03.011.
- Styner, M. a, Rajamani, K. T., Nolte, L.-P., Zsemlye, G., Székely, G., Taylor, C. J. and Davies, R. H. (2003) 'Evaluation of 3D correspondence methods for model building.', in Taylor, C. and Noble, J. A. (eds) *Information processing in medical imaging. IPMI 2003. Lecture Notes in Computer Science*. doi: 10.1007/b11820.
- Suh, H.-Y., Lee, S.-J., Lee, Y.-S., Donatelli, R. E., Wheeler, T. T., Kim, S.-H., Eo, S.-H. and Seo, B.-M. (2012) 'A more accurate method of predicting soft tissue changes after mandibular setback surgery', *Journal of Oral and Maxillofacial Surgery*, 70(10), pp. e553–e562. doi: 10.1016/j.joms.2012.06.187.
- Swennen, G. R. J., Mollemans, W. and Schutyser, F. (2009) 'Three-Dimensional Treatment Planning of Orthognathic Surgery in the Era of Virtual Imaging', *Journal of Oral and Maxillofacial Surgery*, 67(10), pp. 2080–2092. doi: 10.1016/j.joms.2009.06.007.
- Tassabehji, M., Hamond, P., Karmiloff-Smith, A., Thompson, P., Thorgeirsson, S. S., Durkin, M. E., Popescu, N. C., Hutton, T., Metcalfe, K., Rucka, A., Stewart, H., Read, A. P., Maconochie, M. and Donnai, D. (2005) 'GTF2IRD1 in Craniofacial Development of Humans and Mice', *Science*, 310(5751), pp. 1184–7. doi: 10.1126/science.1116142.
- Tenhagen, M., Bruse, J. L., Rodriguez-Florez, N., Angullia, F., Borghi, A., Koudstaal, M. J., Schievano, S., Jeelani, O. and Dunaway, D. (2016) 'Three-Dimensional Handheld Scanning to Quantify Head-Shape Changes in Spring-Assisted Surgery for Sagittal Craniosynostosis', *Journal of Craniofacial Surgery*, 27(8), pp. 2117–2123. doi: 10.1097/SCS.00000000000003108.
- Terzic, A., Combescure, C. and Scolozzi, P. (2014) 'Accuracy of computational soft tissue predictions in orthognathic surgery from three-dimensional photographs 6 months after completion of surgery: A preliminary study of 13 patients', *Aesthetic Plastic Surgery*, 38(1), pp. 184–191. doi: 10.1007/s00266-013-0248-4.
- Terzopoulos, D. and Fleischer, K. (1988) 'Modeling Inelastic Deformation: Viscoelasticity,

-
- Plasticity, Fracture’, *ACM SIGGRAPH Computer Graphics*, 22(4), pp. 269–278. doi: 10.1145/378456.378522.
- Terzopoulos, D., Platt, J., Barr, A. and Fleischer, K. (1987) ‘Elastically deformable models’, in *SIGGRAPH ’87 Proceedings of the 14th annual conference on Computer graphics and interactive techniques*, pp. 205–214. doi: 10.1145/37402.37427.
- Teschner, M., Girod, S. and Girod, B. (2000) ‘Direct Computation of Nonlinear Soft-Tissue Deformation’, in *Vision Modeling and Visualization*, pp. 1–8.
- Tewari, A., Zollhöfer, M., Kim, H., Garrido, P., Bernard, F., Pérez, P. and Theobalt, C. (2017) ‘MoFA: Model-Based Deep Convolutional Face Autoencoder for Unsupervised Monocular Reconstruction’, in *IEEE International Conference on Computer Vision Workshops, ICCVW 2017*. doi: 10.1109/ICCVW.2017.153.
- Then, C., Vogl, T. J. and Silber, G. (2012) ‘Method for characterizing viscoelasticity of human gluteal tissue’, *Journal of Biomechanics*, 45(7), pp. 1252–1258. doi: 10.1016/j.jbiomech.2012.01.037.
- Tzou, C. H. J., Artner, N. M., Pona, I., Hold, A., Placheta, E., Kropatsch, W. G. and Frey, M. (2014) ‘Comparison of three-dimensional surface-imaging systems’, *Journal of Plastic, Reconstructive and Aesthetic Surgery*, 67(4), pp. 489–497. doi: 10.1016/j.bjps.2014.01.003.
- Vannier, M. W., Marsh, J. L. and Warren, J. O. (1984) ‘Three dimensional CT reconstruction images for craniofacial surgical planning and evaluation.’, *Radiology*, 150(1), pp. 179–84. doi: 10.1148/radiology.150.1.6689758.
- Van Der Vlis, M., Dentino, K. M., Vervloet, B. and Padwa, B. L. (2014) ‘Postoperative swelling after orthognathic surgery: A prospective volumetric analysis’, *Journal of Oral and Maxillofacial Surgery*, 72(11), pp. 2241–2247. doi: 10.1016/j.joms.2014.04.026.
- Wang, J., Zou, D., Li, Z., Huang, P., Li, D., Shao, Y., Wang, H. and Chen, Y. (2014) ‘Mechanical Properties of Cranial Bones and Sutures in 1–2-Year-Old Infants’, *Medical Science Monitor*, 20, pp. 1808–1813. doi: 10.12659/MSM.892278.

- Westermarck, A., Zachow, S. and Eppley, B. L. (2005) 'Three-dimensional osteotomy planning in maxillofacial surgery including soft tissue prediction', *Journal of Craniofacial Surgery*, 16(1), pp. 100–104. doi: 10.1097/00001665-200501000-00019.
- Willinger, R., Kang, H. S. and Diaw, B. (1999) 'Three-dimensional human head finite-element model validation against two experimental impacts', *Annals of Biomedical Engineering*, 27, pp. 403–410. doi: 10.1114/1.165.
- Wippold, F. J. (2007) 'Head and neck imaging: The role of CT and MRI', *Journal of Magnetic Resonance Imaging*, 25(3), pp. 453–65. doi: 10.1002/jmri.20838.
- Wong, J. Y., Oh, A. K., Ohta, E., Hunt, A. T., Rogers, G. F., Mulliken, J. B. and Deutsch, C. K. (2008) 'Validity and reliability of craniofacial anthropometric measurement of 3D digital photogrammetric images', *Cleft Palate-Craniofacial Journal*, 45(3), pp. 232–239. doi: 10.1597/06-175.
- Wu, J. Z., Cutlip, R. G., Welcome, D. and Dong, R. G. (2006) 'Estimation of the viscous properties of skin and subcutaneous tissue in uniaxial stress relaxation tests', *Bio-Medical Materials & Engineering*, 16(1), pp. 53–66.
- Wu, J. Z., Dong, R. G., Smutz, W. P. and Schopper, A. W. (2003) 'Nonlinear and viscoelastic characteristics of skin under compression: experiment and analysis', *Bio-Medical Materials and Engineering*, 13(4), pp. 373–385.
- Xia, J., Ip, H. H. S., Samman, N., Wong, H. T. F., Gateno, J., Wang, D., Yeung, R. W. K., Kot, C. S. B. and Tideman, H. (2001) 'Three-dimensional virtual-reality surgical planning and soft-tissue prediction for orthognathic surgery', *IEEE Transactions on Information Technology in Biomedicine*, 5(2), pp. 97–107. doi: 10.1109/4233.924800.
- Xia, J. J., Gateno, J., Teichgraber, J. F., Christensen, A. M., Lasky, R. E., Lemoine, J. J. and Liebschner, M. A. K. (2007) 'Accuracy of the Computer-Aided Surgical Simulation (CASS) System in the Treatment of Patients With Complex Craniomaxillofacial Deformity: A Pilot Study', *Journal of Oral and Maxillofacial Surgery*, 65(2), pp. 248–54. doi: 10.1016/j.joms.2006.10.005.

-
- Xia, J. J., Phillips, C. V., Gateno, J., Teichgraeber, J. F., Christensen, A. M., Gliddon, M. J., Lemoine, J. J. and Liebschner, M. A. K. (2006) ‘Cost-Effectiveness Analysis for Computer-Aided Surgical Simulation in Complex Cranio-Maxillofacial Surgery’, *Journal of Oral and Maxillofacial Surgery*. doi: 10.1016/j.joms.2005.12.072.
- Xia, J. J., Shevchenko, L., Gateno, J., Teichgraeber, J. F., Taylor, T. D., Lasky, R. E., English, J. D., Kau, C. H. and McGrory, K. R. (2011) ‘Outcome study of computer-aided surgical simulation in the treatment of patients with craniomaxillofacial deformities’, *Journal of Oral and Maxillofacial Surgery*, 69(7), pp. 2014–2024. doi: 10.1016/j.joms.2011.02.018.
- Yan, W. and Pangestu, O. D. (2011) ‘A modified human head model for the study of impact head injury’, *Computer Methods in Biomechanics and Biomedical Engineering*, 14(12), pp. 1049–1057. doi: 10.1080/10255842.2010.506435.
- Zachow, S. (2015) ‘Computational Planning in Facial Surgery’, *Facial Plastic Surgery*, 31(5), pp. 446–62. doi: 10.1055/s-0035-1564717.
- Zachow, S., Gladilin, E., Zeilhofer, H. F. and Sader, R. (2001) ‘Improved 3D osteotomy planning in cranio-maxillofacial surgery’, in *Lecture Notes in Computer Science*, pp. 473–481. doi: 10.1007/3-540-45468-3_57.
- Zafeiriou, S. and Dunaway, D. (2018) *Large Scale Facial Model (LSFM)*. Available at: <https://xip.uclb.com/i/software/LSFM.html> (Accessed: 29 November 2018).
- Zhang, X., Tang, Z., Liebschner, M. A. K., Kim, D., Shen, S., Chang, C. M., Yuan, P., Zhang, G., Gateno, J., Zhou, X., Zhang, S. X. and Xia, J. J. (2016) ‘An eFace-Template Method for Efficiently Generating Patient-Specific Anatomically-Detailed Facial Soft Tissue FE Models for Craniomaxillofacial Surgery Simulation’, *Annals of Biomedical Engineering*, 44(5), pp. 1656–1671. doi: 10.1007/s10439-015-1480-7.
- Zienkiewicz, O. C., Taylor, R. L. and Zhu, J. Z. (2005) *The Finite Element Method: Its Basis and Fundamentals*. Butterworth-Heinemann. doi: 10.1016/B978-1-85617-633-0.00015-0.

Appendix A **LIST OF PUBLICATIONS**

A.1 Peer reviewed journal articles directly related to this work

1. Knoops PGM, Borghi A, Breakey RWF, Ong J, Jeelani NUO, Bruun R, Schievano S, Dunaway D, Padwa BL. Three-dimensional soft tissue prediction in orthognathic surgery: a clinical comparison of Dolphin, ProPlan CMF, and probabilistic finite element modelling. *Int J Oral Maxillofac Surg* 2018 <https://doi.org/10.1016/j.ijom.2018.10.008>
2. Knoops PGM, Borghi A, Ruggiero F, Badiali G, Bianchi A, Marchetti C, Rodriguez-Florez N, Breakey RWF, Jeelani NUO, Dunaway DJ, Schievano S. A novel soft tissue prediction methodology for orthognathic surgery using probabilistic finite element modelling. *PLoS ONE*, 2018;13(5):e0197209 <dx.doi.org/10.1371/journal.pone.0197209>
3. Knoops PGM, Beaumont CAA, Borghi A, Rodriguez-Florez N, Breakey RWF, Rodgers W, Angulia F, Jeelani NUO, Schievano S, Dunaway DJ. Comparison of 3D scanners for craniomaxillofacial imaging. *J Plast Reconstr Aesthet Surg*, 2017;70(4):441-9, <dx.doi.org/10.1016/j.bjps.2016.12.015>
4. Knoops PGM, Papaioannou A, Borghi A, Breakey RWF, Wilson A, Jeelani NUO, Zafeiriou S, Steinbacher D, Padwa BL, Dunaway DJ, Schievano S. A machine learning framework for automated computer-assisted diagnosis and planning in plastic and reconstructive surgery. *Sci Rep* [under review]

A.2 Other peer reviewed journal articles

5. O'Hara J, Way B, Borghi A, Knoops PGM, Chua D, Hayward R. Introducing the turriccephaly index: a validated method for analyzing the natural history of turriccephaly in Apert syndrome. *J Craniomaxillofac Surg* 2018 <dx.doi.org/10.1016/j.jcms.2018.12.007>

-
6. Maas BDPJ, Ruff C, Pluijmers BI, Knoops PGM, Dunaway DJ, Koudstaal MJ. Using principal component analysis to describe the midfacial deformities in patients with craniofacial microsomia. *J Craniomaxillofac Surg* 2018
[dx.doi.org/10.1016/j.jcms.2018.09.019](https://doi.org/10.1016/j.jcms.2018.09.019)
 7. Breakey RWF, Knoops PGM, Borghi A, Rodriguez-Florez N, James G, Dunaway DJ, Schievano S, Jeelani NUO. Intracranial volume and head circumference in unoperated syndromic craniosynostosis and age matched controls. *Plast Reconstr Surg* 2018;142(5):708e-717e [dx.doi.org/10.1097/PRS.0000000000004843](https://doi.org/10.1097/PRS.0000000000004843)
 8. Breakey RWF, Knoops PGM, Borghi A, Rodriguez-Florez N, Hayward R, Dunaway DJ, Schivano S, Jeelani NUO. Intracranial volume measurement: a systematic review and comparison of different techniques. *J Craniofac Surg*, 2017;28(7):1746-51
[dx.doi.org/10.1097/SCS.0000000000003929](https://doi.org/10.1097/SCS.0000000000003929)
 9. Beaumont CAA, Knoops PGM, Borghi A, Jeelani NUO, Koudstaal MJ, Schievano S, Dunaway DJ, Florez-Rodriguez N. Three-dimensional surface scanners compared with standard anthropometric measurements for head shape. *J Craniomaxillofac Surg*, 2017;45(6):921-27 [dx.doi.org/10.1016/j.jcms.2017.03.003](https://doi.org/10.1016/j.jcms.2017.03.003)
 10. Rodgers W, Glass GE, Schievano S, Borghi A, Rodriguez-Florez N, Tahim A, Angullia F, Breakey W, Knoops PGM, Tenhagen M, O'Hara J, Ponniah A, James G, Dunaway DJ, Jeelani NUO. Spring assisted cranioplasty for the correction of non-syndromic scaphocephaly: a quantitative analysis of 100 consecutive cases. *Plast Reconstr Surg*, 2017;140(1):125-34 [dx.doi.org/10.1097/PRS.0000000000003465](https://doi.org/10.1097/PRS.0000000000003465)
 11. Rodriguez-Florez N, Göktekin ÖK, Bruse JL, Borghi A, Angullia F, Knoops PGM, Tenhagen M, O'Hara JL, Koudstaal MJ, Schievano S, Jeelani NUO, James G, Dunaway

D. Quantifying the effect of corrective surgery for trigonocephaly: a non-invasive non-ionizing method using 3D handheld scanning and statistical shape modelling.

J Craniomaxillofacial Surg, 2017;45(3):387-94 [dx.doi.org/10.1016/j.jcms.2017.01.002](https://doi.org/10.1016/j.jcms.2017.01.002)

12. Knoops PGM, Biglino G, Hughes AD, Parker KH, Xu L, Schievano S, Torii R. A mock circulatory system incorporating a compliant 3D-printed anatomical model to investigate pulmonary haemodynamics. *Artif Organs*, 2017;41(7):637-46, [dx.doi.org/10.1111/aor.12809](https://doi.org/10.1111/aor.12809)

A.3 Peer reviewed conference publications directly related to this work

1. Knoops PGM, Papaioannou A, Borghi A, Breakey RWF, Wilson A, Jeelani NUO, Zafeiriou S, Steinbacher D, Padwa BL, Padwa BL, Dunaway DJ, Schievano S. A machine learning framework for automated diagnosis and computer-assisted surgery planning in orthognathic surgery. *International Society of Craniofacial Surgery*, 2019, Paris, France (Oral presentation)
2. Knoops PGM, Borghi A, Breakey RWF, Ong J, Jeelani NUO, Bruun R, Padwa BL, Schievano S, Dunaway D. Comparison of Dolphin, ProPlan CMF, and probabilistic finite element modelling for 3D soft tissue prediction in craniomaxillofacial surgery. *Biannual Conference of the European Society of Craniofacial Surgery* 2018, Athens, GR. (Oral presentation)
3. Knoops PGM, Borghi A, Dunaway DJ, Schievano S, Padwa BL. Comparison of Dolphin and ProPlan CMF for 3D soft tissue prediction in orthognathic surgery. *24th Congress of*

-
- the European Association for Cranio Maxillo Facial Surgery* 2018, Munich, DE (Oral presentation)
4. Knoops PGM, Borghi A, Padwa BL, Dunaway DJ, Schievano S. Statistical finite element modelling for maxillofacial surgery simulation and clinical decision-making. *8th World Congress of Biomechanics*, 2018, Dublin, IE (Oral presentation)
 5. Knoops PGM, Borghi A, Papaioannou A, Rodriguez-Florez N, Breakey RWF, Schievano S, Dunaway DJ. Statistical modelling in orthognathic surgery. *National Craniofacial Audit Meeting* 2018, London, UK (Oral presentation)
 6. Knoops PGM. 3D imaging & statistical and physical modelling. *Inaugural International Craniofacial Morphometrics Symposium* 2018, Vail, CO, USA (Oral presentation)
 7. Knoops PGM, Borghi A, Ruggiero F, Bianchi A, Schievano S, Dunaway DJ. Uncertainties in soft tissue prediction in maxillofacial surgery: a probabilistic finite element modelling investigation. *International Society of Craniofacial Surgery*, 2017, Cancun, Mexico (Poster)
 8. Knoops PGM, Borghi A, Schievano S, Jeelani NUO, Dunaway DJ. Total monobloc osteotomy: image analysis of the rigid external distractor. *23rd Congress of the European Association of Cranio Maxillo Facial Surgery*, 2016, London, UK (Poster)