Check for updates

Received: 2 December 2022 | Revised: 14 January 2023 | Accepted: 26 January 2023

DOI: 10.1002/rcs.2503

REVIEW ARTICLE

The International Journal of Medical Robotics and Computer Assisted Surgery

WILEY

# The accuracy of statistical shape models in predicting bone shape: A systematic review

Amogh Patil[1] | Krishan Kulkarni[2] | Shuqiao Xie[3] | Anthony M. J. Bull[3] | Gareth G. Jones[1]

[1]The MSk Lab, Imperial College London, London, UK

[2]Department of Trauma and Orthopaedics, East Lancashire Hospitals NHS Trust, Blackburn, UK

[3]Department of Bioengineering, Imperial College London, London, UK

**Correspondence**
Amogh Patil.
Email: amogh.patil1@nhs.net

## Abstract

**Background:** This systematic review aims to ascertain how accurately 3D models can be predicted from two-dimensional (2D) imaging utilising statistical shape modelling.

**Methods:** A systematic search of published literature was conducted in September 2022. All papers which assessed the accuracy of 3D models predicted from 2D imaging utilising statistical shape models and which validated the models against the ground truth were eligible.

**Results:** 2127 papers were screened and a total of 34 studies were included for final data extraction. The best overall achievable accuracy was 0.45 mm (root mean square error) and 0.16 mm (average error).

**Conclusion:** Statistical shape modelling can predict detailed 3D anatomical models from minimal 2D imaging. Future studies should report the intended application domain of the model, the level of accuracy required, the underlying demographics of subjects, and the method in which accuracy was calculated, with root mean square error recommended if appropriate.

**KEYWORDS**
3D imaging, bone, joints, modelling, orthopaedic, statistical shape modelling

## 1 | INTRODUCTION

Medical imaging plays a vital role in orthopaedic surgery and a more recent development is the use of three-dimensional (3D) imaging to better appreciate anatomy, assist with surgical planning, guide assistive technology such as robotics, and even to develop patient specific implants.[1,2]

The current gold-standard for producing 3D bone models is through segmentation of images derived from Computed Tomography (CT) or Magnetic Resonance Imaging (MRI) scans.[3,4] CT is generally preferred when focussing on bone shape as it displays a lower margin of error when compared to MRI derived models, which tend to be more accurate when focussing on soft tissue.[4–6] Currently, CT-based models have been shown to have an average error of 0.15 mm and MRI-based models 0.23 mm, with both modalities displaying extremely high 3D geometric accuracy.[4] When applied to surgery, Livyatan et al. suggest a model which displays reconstruction accuracy of 1–1.5 mm is 'desirable' whereas 2–3 mm is considered 'acceptable'.[7]

The widespread adoption of these imaging techniques has been restricted due to cost and time implications, and for CT imaging there is the issue of additional radiation exposure for the patient.[8,9] Statistical Shape Models (SSMs) offer a potential solution to this

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2023 The Authors. The International Journal of Medical Robotics and Computer Assisted Surgery published by John Wiley & Sons Ltd.

problem—they can utilise common 2D imaging modalities such as radiographs to predict 3D models of a patient, thus eliminating the need for CT or MRI scans.[10] The required accuracy of these SSMs will depend on their clinical application, and for their potential use in surgical planning or patient-specific implants, the accuracy will need to be within degrees (for surgical planning cuts/angles) and millimetres (in implant design, for example) of the currently accepted gold-standard. It follows that before the widespread adoption of SSMs in clinical practice, their accuracy in reconstructing bone shape must be evaluated.

Accuracy is assessed by comparing the SSM predicted bone shape with a ground truth model. Two commonly employed numerical methods exist to calculate these differences: root-mean-square-error (RMSE) and average error. RMSE is a quadratic scoring rule, with the difference between each forecast and corresponding value squared and then averaged between samples. The outcome is scale dependent.[11] Average error is a broad term describing the average magnitude of errors between forecast and corresponding values. Larger errors are given a higher weight in RMSE as the errors are squared before they are averaged. In average error, the forecast and its corresponding values can be found between points or between points to the surface. As these differences are small, the methods in which correspondence is calculated will not be considered in this review.

This systematic review set-out to assess whether SSMs designed to predict 3D bone shape can deliver a level of accuracy suitable for clinical application.

## 2 | MATERIAL AND METHODS

### 2.1 | Search strategy

An electronic Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) compliant search was conducted in May 2021 and then repeated in September 2022.[12] The search utilised a combination of medical subject heading terms (MeSH) and key words (Table 1). The databases searched included MEDLINE/PubMed, EMBASE and The Cochrane Library—terms were modified to meet each respective database criteria. Reference lists of included studies were also reviewed for eligibility. Any duplicate manuscripts were removed. For information which was replicated by authors in the form of a paper alongside conference proceedings, the publication with the most complete data set was retrieved.

### 2.2 | Selection criteria

Eligibility was confirmed where authors predicted a 3D bone model through application of 2D radiological images using an SSM and compared the predicted model to a ground truth model. Exclusion criteria included studies which did not assess accuracy, paediatric datasets, and non-English language papers—full inclusion and exclusion criteria are listed in Table 2. There were no date limits applied. In titles where inclusion or exclusion was not obvious, the full text was retrieved and reviewed.

**TABLE 1** Search strategy terms for literature search on MEDLINE (via OVID).

| Population AND→ | Intervention AND → | Intervention/Comparison |
|---|---|---|
| Bone* | SSM* | Exp imaging, three-dimensional |
| Joint exp. Joints | Statistic* shape* model* | Exp diagnostic imaging |
| Hip* OR acetabulum OR femur* or pelvi* | Statistic* adj4 shape* | Exp image processing, computer-assisted |
| Knee* OR tibia OR fibula OR patella | | Exp tomography, X-ray computed |
| Ankle* OR foot | | CT |
| Lower limb | | Computed tomography |
| Exp lower extremity | | Exp magnetic resonance imaging |
| Hand* OR wrist* or radius OR ulna | | Magnetic resonance imag* |
| Elbow* OR olecranon | | MRI |
| Shoulder* OR glenohumeral OR acromioclavicular | | Exp fluoroscopy |
| Clavicle OR humerus | | Exp X-rays |
| Upper limb | | Radiograph* |
| Exp upper extremity | | X-ray* |
| Spine OR spinal OR back OR vertebra* | | Exp absorptiometry, photon |
| Cervical OR thoracic OR lumbar | | Dual energy X-ray absorptiometry |
| | | DEXA |
| | | DXA |

1478596x, 2023, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rcs.2503 by Test, Wiley Online Library on [18/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

**TABLE 2** Inclusion and exclusion criteria.

| | Inclusion | Exclusion |
|---|---|---|
| Population | Human, adult (≥18 years) | Paediatric population (<18 years) |
| | Bone/group of bones | Model unrelated to orthopaedics (e.g., chest wall, maxillofacial) |
| | | Isolated soft tissue models |
| Intervention | 3D statistical shape model | Segmentation technique (e.g., manual vs. automatic segmentation) |
| | Use of 2D imaging to derive predicted 3D model | |
| Comparison | 3D ground truth model | Experiment dataset same as training dataset |
| Outcome | Accuracy assessed through comparison of 3D predicted model to 3D ground-truth model using RMSE or average error | |
| Other | English language | Review articles, letters to the editor, supplementary articles, technique papers |

Abbreviations: 2D, two dimensional; 3D, three dimensional; RMSE, root mean square error.

## 2.3 | Data extraction and critical appraisal

Search results were exported to a reference manager, Covidence (Veritas Health Innovation Ltd.). Two independent reviewers (AP and KK) analysed the titles and abstracts to identify studies meeting the inclusion criteria. Any conflicts between author decisions were discussed at a consensus meeting. A third reviewer (GJ) was available if an agreement could not be reached, however, there were no papers to which this applied.

The same reviewers utilised a two-part standardised form to extract data based on the Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies (CHARMS) tool.[13] The first form identified general study characteristics such as the publication year, study authors and anatomical area of interest. Details of the SSM were also collected, including population type (cadaver, human, plastic model), number of specimens, general demographics of specimens and the type of 3D imaging. The second form focussed on the validation experiment designed to assess accuracy—population type, number of subjects the model was tested on, method of 2D imaging, derivation of ground truth model.

Finally, the numerical value of accuracy alongside the method in which it was calculated was recorded. These were then classified as either 'RMSE' or 'average error', as described by the two equations below (Figure 1). In studies where both methods were tested, the best value was taken. Any conflicts were discussed at a consensus meeting and a third reviewer (GJ) was available if an agreement was not reached, however, there were no papers to which this applied.

## 2.4 | Risk of bias

Risk of Bias of included studies was assessed by two independent reviewers (AP and KK), utilising the Prediction model Risk of Bias Assessment Tool (PROBAST) tool.[14] PROBAST contains a total of 20 signalling questions, divided into four key domains relevant to prediction models: Participants, Predictors, Outcomes, Analysis. Each domain is rated for risk of bias (low, high or unclear) based on the signalling questions. The applicability of these domains from each model is also assessed in relation to the review question.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(predicted_i - ground\ truth_i)^2}$$

$$Average\ error = \frac{1}{n}\sum_{i=1}^{n} abs(predicted_i - ground\ truth_i)$$

**FIGURE 1** $Predicted_i$ = 3D model predicted from SSM, $Ground\ truth_i$ = 3D model derived from CT/MRI imaging. CT, computed tomography; MRI, magnetic resonance imaging; SSM, statistical shape model.

# 3 | RESULTS

## 3.1 | Search results, study and subject characteristics

A total of 2127 citations were retrieved. After the initial screening of abstracts, the full text of 75 records were analysed (Figure 2). Following full-text review, 34 studies were included for final data extraction.[15–48]

The 34 studies originated from several countries, with the majority of authors based in Europe. A total of 2233 images were utilised to develop 31 different SSMs—three study groups applied the same model in two separate published papers; this work is included as one model. 26 of these models (83.9%) used one scan per subject whereas three models (9.7%) utilised contralateral scans taken from the same subject. In two models it was unclear if each scan was from separate subjects. Almost all models were developed

1478596x, 2023, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rcs.2503 by Test, Wiley Online Library on [18/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

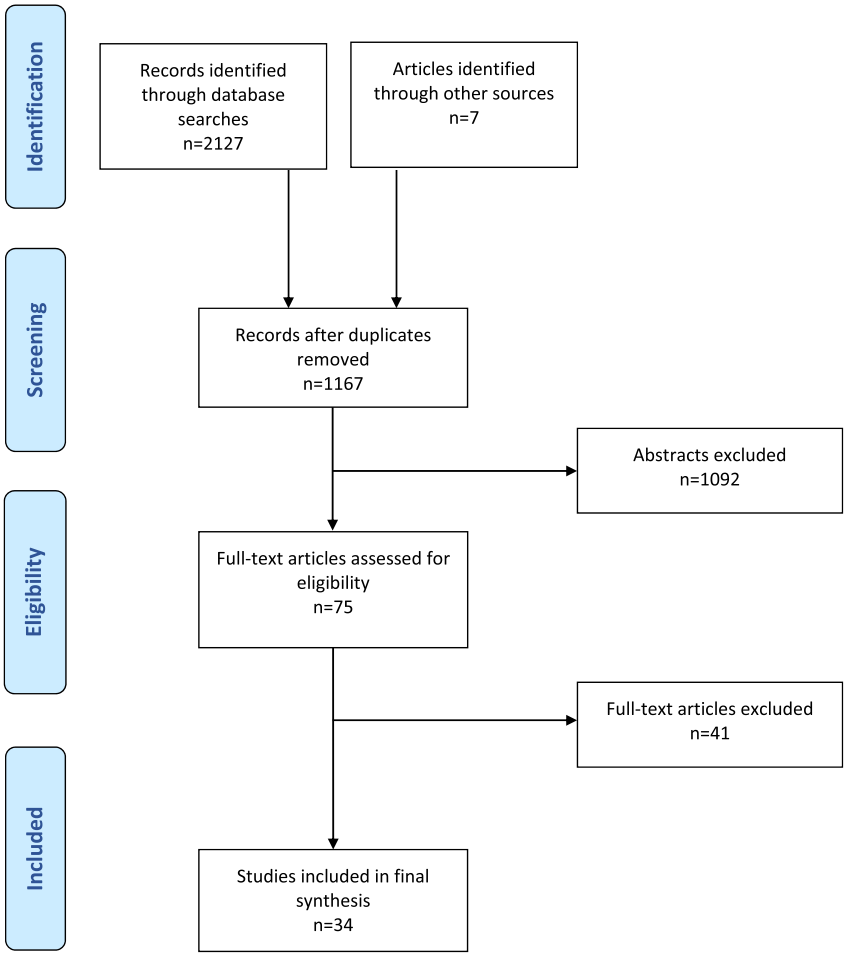


**FIGURE 2** PRISMA diagram of articles selected for review.

using patients (58.1%) or cadaveric specimens (16.1%) or a combination of both (16.1%). A single study (3.2%) made use of dry bones as well as patient images to develop their SSM (Table 3). In two models (6.5%) it was unclear where images were obtained to develop their model.

Fourteen models (45.1%) specified details the gender of subjects included, with 573 males and 465 females. 14 models (45.1%) provided details of the underlying age and only one study (3.2%) commented on ethnic background of the subjects used to develop the SSM. The models described six different bone shapes: Scapula, Lumbar Vertebra, Pelvis, Femur, Patella, Tibia (Table 3).

## 3.2 | Assessment of accuracy

Sample sizes for validation experiments varied between 1 and 180 subjects. In total, the models were validated on 740 subjects—these consisted of patients, cadavers and artificial saw bones (Table 2). The choice of 2D imaging applied to the models varied between plain radiographs, fluoroscopy, DXA or digitally reconstructed radiographs (DRRs). The ground-truth models were derived solely from CT scans in 30 (88.2%) studies. Two papers (5.9%) used a combination of CT scan and laser-scan reconstruction. The remaining two papers (5.9%) utilised MRI scans (Table 4).

Based on Livyatan et al.'s criteria, 24 (77.4%) of the included models demonstrated desirable accuracy and five (16.1%) demonstrated acceptable accuracy.[7] One study utilised an SSM which displayed desirable accuracy in femur reconstruction but acceptable accuracy in the tibia.[27] One study reported an unacceptable error of 3.2 mm.[17] The three study groups which applied the same model in different papers showed desirable accuracy in both publications.[15,16,25,36,37,48] Overall, accuracy in studies using RMSE ranged between 0.45 and 1.95 mm. Studies utilising mean errors ranged between 0.16 and 3.2 mm (Table 5).

## 3.3 | Risk of bias (Table 6)

Risk of bias was generally low with regards to Participants and Outcomes. Certain studies provided limited data on methods of participant selection and therefore the risk of bias was unclear. There was a high risk of bias in the Predictors and Analysis domains due to minimal information being provided with regards to the demographics of the training dataset, which is known to have a substantial effect on the accuracy of the SSM, as well as the use of small sample sizes.[49] Applicability of the models, due to the strict inclusion and exclusion criteria, was generally marked as high where datasets were complete.

1478596x, 2023, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rcs.2503 by Test, Wiley Online Library on [18/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

**TABLE 3** Details of statistical shape model development.

| Study n | Model n | Lead author, year | Country | Anatomical location | Statistical shape model | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Imaging | Anatomical specimens (*n*) | Use of contralateral anatomy | Population type | Males/females (*n*) | Age (mean ± SD or range) |
| 1 | 1 | Mutsvanga, 2017[17] | South Africa | Scapula | CT | 27 | No | Cadaver | - | - |
| 2 | 2 | Picazo, 2018[18] | Spain | Lumbar spine | qCT | 90 | No | Adult–alive | 66/24 | 55.2 ± 11.3 |
| 3 | 3 | Whitmarsh, 2013[30] | Spain | Lumbar spine | CT | 66 | No | Adult–Alive | 0/66 | 54 ± 11 |
| 4 | 4 | Zheng, 2010 (C)[35] | Switzerland | Lumbar spine | CT | 39 (L1 = 3, L2 = 5, L3 = 9, L4 = 14, L5 = 8) | No | Adult–Alive | - | - |
| 5 | 5 | Zheng, 2011[34] | Switzerland | Lumbar spine | CT | 50 | No | Adult–Alive | - | - |
| 6 | 6 | Sadowsky, 2007[19] | United States | Pelvis | CT | 110 | No | Adult–Alive | 110/0 | - |
| 7 | 7 | Zheng, 2010 (B)[32] | Switzerland | Pelvis | CT | 24 | No | Dry bones<br>Adult–Alive | - | - |
| 8 | 8 | Schumann, 2013[21] | Switzerland | Femoral head<br>Acetabulum | CT | 30 | Yes | Adult–Alive | - | - |
| 9 | 9 | Klima, 2016[46] | Czech republic | Femoral shaft | CT | 22 | No | Unclear | - | - |
| 10 | 10 | Whitmarsh, 2010[29] | Spain | Proximal femur | CT | 60 | No | Cadaver | 28/32 | 79 ± 11 |
| 11 | 11 | Hurvitz, 2008[45] | Israel | Proximal femur | CT | 17 | No | Cadaver | - | - |
| 12 | 12 | Vaanenen, 2015[26] | Finland | Proximal femur | CT | 69 | Yes | Cadaver<br>Adult–Alive | 57/12 | 54 |
| 13 | 13 | Kurazume, 2009[47] | Japan | Proximal femur | CT | 56 | No | Adult–Alive | - | - |
| 14 | 14 | Dong, 2009[43] | Switzerland | Proximal femur | CT | 13 | No | Unclear | - | - |
| 15 | 15 | Spinelli, 2012[23] | Spain | Proximal femur | CT | 115 | No | Adult–Alive | 42/73 | 55 ± 12 |
| 16 | 16 | Humbert, 2017[44] | Spain | Proximal femur | qCT | 111 | No | Adult–Alive | 30/81 | 56.2 ± 12.1 |
| 17 | 17 | Whitmarsh, 2011[28] | Spain | Proximal femur | qCT | 85 | No | Adult–Alive | 27/58 | 55 ± 12 |
| 18 | 18 | Zheng, 2009 (B)[36] | Switzerland | Proximal femur | CT | 30 | No | Adult–Alive | - | - |
| 19 | 19 | Zheng, 2009 (A)[33] | Switzerland | Proximal femur | CT | 30 | No | Adult–Alive | - | - |
| 20 | 19 | Zheng, 2010 (A)[37] | Switzerland | Proximal femur | CT | 30 | No | Adult–Alive | - | - |
| 21 | 20 | Grassi, 2021[40] | Sweden | Proximal femur | CT | 59 | No | Adult–alive | 40/19 | 58 (18–188) |
| 22 | 21 | Tang, 2005[24] | Canada | Distal femur | CT | 40 | Yes | Cadaver | - | - |

(Continues)



1478596x, 2023, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rcs.2503 by Test, Wiley Online Library on [18/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 3 (Continued)

| Study n | Model n | Lead author, year | Country | Anatomical location | Statistical shape model | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Imaging | Anatomical specimens (*n*) | Use of contralateral anatomy | Population type | Males/females (*n*) | Age (mean ± SD or range) |
| 23 | 22 | Zhu, 2011[39] | United States | Distal femur | MRI | 40 | No | Adult–Alive | 24/16 | 29.9 |
| 24 | 23 | Cerveri, 2018[38] | Italy | Distal femur | CT | 80 | No | Adult–Alive | - | 67 ± 10 |
| 25 | 24 | Baka, 2011[16] | Netherlands | Distal femur | CT angiogram/CT | 43 | No | Cadaver<br>Adult–Alive | 30/13 | 23–92 |
| 26 | 24 | Baka, 2012[15] | Netherlands | Distal femur | CT angiogram/CT | 43 | No | Cadaver<br>Adult–Alive | 30/13 | 23–92 |
| 27 | 25 | Wu, 2021[31] | United States | Distal femur | CT | 300 | Unclear | Adult–Alive | - | - |
| 28 | 26 | Schumann, 2016[20] | Switzerland | Femur<br>Tibia | CT | 129–Femur<br>17–Tibia | No | Adult–Alive | - | - |
| 29 | 27 | Gajny, 2022[41] | France | Femur<br>Tibia | CT | 158 | Unclear | Cadaver | - | - |
| 30 | 28 | Baka, 2014[27] | Netherlands | Distal femur<br>Proximal tibia<br>Proximal fibula | CT | 61 | No | Cadaver<br>Adult–Alive | - | 21–66 |
| 31 | 29 | Li, 2014[48] | United States | Distal femur<br>Proximal tibia | CT | 152 | Yes | Adult–Alive | 34/46 | M 51.1 ± 6.6<br>F 43.5 ± 6.2 |
| 32 | 29 | Tsai, 2015[25] | United States | Distal femur<br>Proximal tibia<br>Patella | CT | 152 | Yes | Adult–Alive | 34/46 | M 51.1 ± 6.6<br>F 43.5 ± 6.2 |
| 33 | 30 | Lu, 2021[42] | Taiwan | Distal femur, proximal tibia | CT | 60 | No | Adult–alive | 60/0 | 22.89 ± 2 |
| 34 | 31 | Smoger, 2017[22] | United States | Patella | MRI | 50 | No | Cadaver<br>Adult–alive | 35/25 | 64 (44–87) |

Abbreviations: CT, computerised tomography; F, female; M, male; MRI, magnetic resonance imaging.

1478596x, 2023, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rcs.2503 by Test, Wiley Online Library on [18/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

**TABLE 4** Validation experiments of included studies.

| | | Validation experiment | | | |
|---|---|---|---|---|---|
| **Lead author, year** | **Anatomical location** | **Anatomical specimens (*n*)** | **Population type** | **2D imaging** | **Ground truth imaging** |
| Mutsvanga, 2017[17] | Scapula | 1 | Cadaver | XR | CT |
| Picazo, 2018[18] | Lumbar spine | 180 | Adult (alive) | DXA | qCT |
| Whitmarsh, 2013[30] | Lumbar spine | 30 | Adult (alive) | DXA | qCT |
| Zheng, 2010 (C)[30] | Lumbar spine | 12 | Cadaver | Fluoroscopy | CT |
| Zheng, 2011[34] | Lumbar spine | 4 | Cadaver | Fluoroscopy | CT |
| Sadowsky, 2007[19] | Pelvis | 11 | Adult (alive) | DRR | CT |
| Zheng, 2010 (B)[32] | Pelvis | 14 | Plastic<br>Cadaver<br>Adult (alive) | Radiograph | CT |
| Schumann, 2013[21] | Femoral head<br>Acetabulum | 7 | Cadaver | Radiograph | CT |
| Klima, 2016[46] | Femoral shaft | 96 | Adult (alive) | DRR | CT |
| Whitmarsh, 2010[29] | Proximal femur | 30 | Cadaver | DRR | qCT |
| Hurvitz, 2008[45] | Proximal femur | 1 | Cadaver | DRR | qCT |
| Vaanenen, 2015[26] | Proximal femur | 12 | Adult (alive) | DXA | CT |
| Kurazume, 2009[47] | Proximal femur | 4 | Adult (alive) | Fluoroscopy | CT |
| Dong, 2009[43] | Proximal femur | 5 | Cadaver | Radiograph | CT |
| Spinelli, 2012[23] | Proximal femur | 2 | Cadaver | DXA | qCT |
| Humbert, 2017[44] | Proximal femur | 157 | Adult (alive) | DXA | qCT |
| Whitmarsh, 2011[28] | Proximal femur | 30 | Adult (alive) | DXA | qCT |
| Zheng, 2009 (B)[36] | Proximal femur | 23 | Plastic<br>Cadaver | Radiograph | CT<br>T-scan (laser reconstruction) |
| Zheng, 2009 (A)[33] | Proximal femur | 1 | Plastic bone | Radiograph | CT |
| Zheng, 2010 (A)[37] | Proximal femur | 11 | Cadaver | Fluoroscopy | CT<br>T-scan (laser reconstruction) |
| Grassi, 2021[40] | Proximal femur | 11 | Cadaver | DXA | CT |
| Tang, 2005[24] | Distal femur | 2 | Cadaver | DRR | CT |
| Zhu, 2011[39] | Distal femur | 10 | Adult (alive) | Fluoroscopy | MRI |
| Cerveri, 2018[38] | Distal femur | 20 | Adult (alive) | DRR | CT |
| Baka, 2011[16] | Distal femur | 2 | Cadaver | Radiograph | CT |
| Baka, 2012[15] | Distal femur | 10 | Adult (alive) | Fluoroscopy | CT |
| Wu, 2021[31] | Distal femur | 5 | Adult (alive) | Fluoroscopy | CT |
| Schumann, 2016[20] | Femur<br>Tibia | 36 | Cadaver | Radiograph | CT |
| Gajny, 2022[41] | Femur<br>Tibia | 29 | Cadaver | DRR | CT |
| Baka, 2014[27] | Distal femur<br>Proximal tibia<br>Proximal fibula | 16 | Cadaver adult (alive) | Fluoroscopy | CT |

(Continues)

1478596x, 2023, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rcs.2503 by Test, Wiley Online Library on [18/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 4 (Continued)

| | | Validation experiment | | | |
|---|---|---|---|---|---|
| Lead author, year | Anatomical location | Anatomical specimens (*n*) | Population type | 2D imaging | Ground truth imaging |
| Li, 2014[48] | Distal femur | 3 | Adult (alive) | Fluoroscopy | CT |
| | Proximal tibia | | | | |
| Tsai, 2015[25] | Distal femur | 2 | Adult (alive) | Fluoroscopy | CT |
| | Proximal tibia | | | | |
| | Patella | | | | |
| Lu, 2021[42] | Distal femur | 10 | Adult (alive) | Fluroscopy | CT |
| | Proximal tibia | | | | |
| Smoger, 2017[22] | Patella | 49 | Cadaver | Radiograph | MRI |

Abbreviations: CT, computerised tomogrophy; DRR, digitally reconstructed radiograph; DXA, dual-energy x-ray absorptiometry; MRI, magenetic resonance imaging.

TABLE 5 Best achievable accuracies categorised as root mean square error (RMSE) or average error.

| Study | Anatomical location | Best accuracy (mm ± SD or range) |
|---|---|---|
| | | RMSE |
| Wu, 2021[31] | Distal femur | 1.04 ± 0.33 |
| | Proximal tibia | 1.03 ± 0.19 |
| Klima, 2016[46] | Femoral shaft | 1.74 |
| Li, 2014[48] | Distal femur | 1.16 |
| | Proximal tibia | 1.4 |
| Baka, 2014[27] | Distal femur | 1.18 |
| | Proximal tibia and fibula | 1.56 |
| Baka, 2012[15] | Distal femur | 1.48 ± 0.41 |
| Cerveri, 2018[38] | Distal femur | 0.75 |
| Smoger, 2017[22] | Patella | 0.45 ± 0.007 |
| Baka, 2011[16] | Distal femur | 0.68 |
| Zheng, 2009 (A)[33] | Proximal femur | 1.3 |
| Tang, 2005[24] | Distal femur | 1.71 |
| Lu, 2021[42] | Distal femur | 0.64 ± 0.084 |
| | Proximal tibia | 0.69 ± 0.069 |
| | | Average error |
| Tsai, 2015[25] | Distal femur | 0.30 ± 0.81 |
| | Proximal tibia | 0.34 ± 0.79 |
| | Patella | 0.36 ± 0.59 |
| Gajny, 2022[41] | Femur | 1 |
| | Tibia | 1.1 |
| Schumann, 2013[21] | Acetabulum | 1.06 ± 0.14 |
| | Femoral head | 1.01 ± 0.16 |

TABLE 5 (Continued)

| Study | Anatomical location | Best accuracy (mm ± SD or range) |
|---|---|---|
| Picazo, 2018[18] | Lumbar vertebrae (L1–L4) | 1.51 |
| Humbert, 2017[44] | Proximal femur | 0.93 |
| Mutsvanga, 2017[17] | Scapula | 3.2 |
| Schumann, 2016[20] | Proximal femur | 1.12 ± 0.89 |
| | Distal femur | 1.13 ± 0.84 |
| | Proximal tibia | 0.88 ± 0.63 |
| | Distal tibia | 0.78 ± 0.56 |
| Zheng, 2011[34] | Lumbar vertebrae | 0.77 |
| Zhu, 2011[39] | Distal femur | 0.16 ± 1.16 |
| Zheng, 2010 (B)[37] | Pelvis | 1.6 |
| Zheng, 2009 (B)[36] | Proximal femur | 0.95 ± 0.22 |
| Zheng, 2010 (A)[37] | Proximal femur | 1 |
| Zheng, 2010 (C)[35] | Lumbar vertebrae | 1 |
| Sadowsky, 2007[19] | Pelvis | 1.42 |
| Whitmarsh, 2013[30] | Lumbar vertebrae–L2 | 1 |
| | Lumbar vertebrae–L3 | 0.93 |
| | Lumbar vertebrae–L4 | 1.34 |
| Whitmarsh, 2010[29] | Proximal femur | 1.2 |
| Hurvitz, 2008[45] | Proximal femur | 1.44 |
| Vaanenen, 2015[26] | Proximal femur | 1.41 |
| Kurazume, 2009[47] | Proximal femur | 0.8–1.1 |
| Dong, 2009[43] | Proximal femur | 1.6 |
| Spinelli, 2012[23] | Proximal femur | 1.2 |
| Whitmarsh, 2011[30] | Proximal femur | 1.1 |
| Grassi, 2021[40] | Proximal femur | 1.02 |

1478596x, 2023, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rcs.2503 by Test, Wiley Online Library on [18/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

**TABLE 6** Risk of bias (RoB) assessment using the PROBAST tool in order from low RoB to high RoB.

| | Domain 1—Participants | | | | Domain 2—Predictors | | | | Domain 3—Outcome | | | | Domain 4—Analysis | | Overall Assessment | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | RoB | | App | | RoB | | App | | RoB | | App | | RoB | | RoB | App |
| Lead Author, Year | DEV | VAL | DEV | VAL | DEV | VAL | DEV | VAL | DEV | VAL | DEV | VAL | DEV | VAL | | |
| Picazo, 2018 | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| Whitmarsh, 2010 | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| Vaanenen, 2015 | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| Spinelli, 2012 | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| Humbert, 2017 | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| Whitmarsh, 2011 | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| Grassi, 2021 | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| Zhu, 2011 | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| Cerveri, 2017 | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| Baka, 2011 | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| Baka, 2012 | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| Baka, 2013 | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| Li, 2014 | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| Tsai, 2015 | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| Smoger, 2017 | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| Lu, 2021 | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| Whitmarsh, 2013 | + | + | + | + | - | - | - | - | + | + | + | + | + | + | + | + |
| Sadowsky, 2007 | - | - | + | + | - | - | - | - | + | + | + | + | - | - | - | - |
| Klima, 2016 | + | + | + | + | - | - | - | - | + | + | + | + | - | - | - | - |
| Zheng, 2009A | + | + | + | + | - | - | - | - | + | + | + | + | - | - | - | - |
| Hurvitz, 2008 | + | + | + | + | - | - | - | - | + | + | + | + | - | - | - | - |
| Zheng, 2008 | + | + | + | + | - | - | - | - | + | + | + | + | - | - | - | - |
| Zheng, 2009B | + | + | + | + | - | - | - | - | + | + | + | + | - | - | - | - |
| Mutsvanga, 2017 | ? | + | + | + | - | - | - | - | + | + | + | + | - | - | - | - |
| Zheng, 2010B | ? | + | + | + | - | - | - | - | + | + | + | + | - | - | - | - |
| Kurazume, 2009 | ? | + | + | + | - | - | - | - | + | + | + | + | - | - | - | - |
| Tang, 2005 | ? | + | + | + | - | - | - | - | + | + | + | + | - | - | - | - |
| Zheng, 2010A | ? | + | ? | + | - | - | - | - | + | + | + | + | - | - | - | - |
| Zheng, 2011 | ? | + | ? | + | - | - | - | - | + | + | + | + | - | - | - | - |
| Schumann, 2013 | ? | + | ? | + | - | - | - | - | + | + | + | + | - | - | - | - |
| Dong, 2009 | ? | + | ? | + | - | - | - | - | + | + | + | + | - | - | - | - |
| Schumann, 2016 | ? | + | ? | + | - | - | - | - | + | + | + | + | - | - | - | - |
| Jing, 2021 | ? | + | ? | + | - | - | - | - | + | + | + | + | - | - | - | - |
| Gajny, 2022 | + | + | + | + | - | - | - | - | + | + | + | + | - | - | - | - |

*Note*: -, high RoB/high concern of applicability; ?, unclear; +, low RoB/low concern of applicability.

Abbreviations: App, applicability; DEV, development; VAL, validation.

1478596x, 2023, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rcs.2503 by Test, Wiley Online Library on [18/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

## 4 | DISCUSSION

The results of this review provide evidence that accurate and validated 3D anatomical models suitable for clinical use have been predicted from 2D imaging through statistical shape modelling techniques. To our knowledge, this is the first piece of work to systematically review the accuracy of SSMs.

All measurements of accuracy were calculated by comparing a reconstructed model to a ground-truth model derived from CT or MRI, which is the current gold-standard in clinical practice. Furthermore, all test subjects were independent to the training dataset. The best achievable accuracies ranged between 0.45–1.74 mm (RMSE) and 0.16–3.2 mm (Average Error). Based on Livyatan et al.'s description of acceptable clinical accuracy, only one model failed to meet the level required to be 'desirable' or acceptable'.[7] Currently, the gold standard methods of CT and MRI demonstrate accuracy within 0.5 mm when predicting 3D bone shape.[4] The majority of SSMs included in our review do not exceed this error margin.

The calculation of accuracy in this review was understood to be the difference between the predicted 3D model and the ground-truth model, calculated as RMSE or Average Error. However, it must be noted the target accuracy of a model will depend on its intended use, and surface reconstruction error might not be the only valid measure of prediction accuracy. For example, Zheng's research group assessed the accuracy of their SSMs (based on a single AP radiograph to reconstruct the 3D anatomy of the pelvis following total hip replacement) by measuring the angular anteversion of the acetabular cup. They concluded that their SSM method was more accurate than a plain radiograph but inferior to CT-based measurements.[50–52] Whereas Nolte et al. evaluated the accuracy of their SSM femoral bone model using a number of different metrics such as femoral head radius, neck angle, bow angle, and angle between femoral mechanical and anatomical axis, and then combined this with grey scale values to compare the mechanical properties (stress and strain) of the bone with the ground truth.[53] Other examples of different outcome measurements include kinematic studies where accuracy is measured in terms of rotation or translation of a particular bone and implant design where surface point accuracy is key.[15,48,54] This highlights the importance of acknowledging different measurements of accuracy depending on the intended use of the prediction model, as well as the need for standardisation of reporting measures wherever possible. For example, a large bone (femur) or joint (hip) is likely to have a larger acceptable margin of error than a smaller bone (carpal) or joint (facet). A range of measurements which would define acceptable accuracy depending on anatomical location and intended clinical use could be derived from expert opinions and would be of significant benefit to those developing these models.

It is noteworthy that the distal femur was a commonly modelled bone, with 12 separate publications.[15,16,20,24,25,27,31,38,39,41,42,48] A likely reason for this is the role of patient specific instrumentation (PSI) in knee replacement (TKR) surgery. PSI is a technique designed to improve overall surgical accuracy during prosthesis insertion.[55] This innovation involves designing patient-specific cutting blocks from pre-operative CT or MRI imaging alongside a full-length standing AP radiograph.[55] Intra-operatively these guides are designed to fit the patient's own bones like a glove, and once secured, guide the surgeon, via slots, to perform the bony cuts according to a previously agreed 3D plan.[38,55] If this could be achieved without the need for CT or MRI, then this approach would become more efficient and cost-effective. Five of the studies describe sub-millimetre accuracy of the distal femur SSI model, with one study (Tsai et al.) reporting sub-millimetre accuracy for the entire knee joint.[25] These results are promising and suggest SSMs may play a significant role in future PSI work. Indeed, it is interesting to note that Zimmer Biomet has recently introduced the world's first CE marked, x-ray based PSI for TKR surgery.[56]

### 4.1 | Limitations

Morphology of joints and bones are known to vary according to ethnicity and gender.[57,58] The underlying ethnicity of the study participants was only mentioned in one of the included studies and gender was only provided in 16 papers. This key information may affect the model's accuracy depending on the patient cohort is it being applied to and resulted in several papers being categorised as 'at risk of bias'. It should be noted that one possible reason for not providing demographics may be related to anonymous cadaveric specimens being donated by medical facilities.[17,36] However, it will be important for future studies, particularly those using patients or images from joint registries, to include this essential information to ensure that applicability for all patient populations can be assessed.

Although all included measurements were a comparison between the predicted 3D model and the ground-truth model, a meta-analysis of model accuracy was not possible due to the inconsistent use of RMSE and average error. Moreover, the average error was calculated using varying methods across the studies, with some using medians and others using means. This variation in reporting the overall accuracy meant that direct statistical comparison between values is not possible. Again, for future studies, it would be useful to standardise the method of reporting to facilitate model comparisons, and a recommended set of reporting guidelines is proposed (Figure 3). All papers aiming to validate a statistical shape model would ideally have a standardised method of calculating accuracy, such as the RMSE of average error, allowing for direct comparison. We recommend this be reported alongside other commonly used clinical bone metrics, such as neck/shaft angle, femoral mechanical and anatomical axis angle or acetabular anteversion depending on the intended use.

Also of note were the number of studies which validated their SSM using a small number of bone samples (<50) and had a prediction error above 1 mm ($n = 21$), which makes interpretation difficult. Another limitation to be aware of is that this review focussed on a highly specific area of research which meant that several of the included papers were from the same research groups, with a resultant risk of publication bias.

1478596x, 2023, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rcs.2503 by Test, Wiley Online Library on [18/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

**FIGURE 3** Proposed set of reporting guidelines for future studies.

**Statistical Shape Model**

- Anatomical Location
- Imaging Modality *(including slice thickness, resolution)*
- Number of Subjects
- Number of Anatomical Specimens *(ie were contralateral images used)*
- Population Type *(Cadaver, Patient, Dry Bone)*
- Gender Distribution
- Age (*Range)*
- Ethnicity *(and stated if not known)*
- Underlying Pathology
- Method of Segmentation
- Review of Imaging (*Clinician/Scientist as well as level of expertise)*

**Validation Experiment**

- 2D Imaging Modality *(including different types and number of views)*
- Ground-Truth Imaging Modality
- Number of Subjects
- Number of Anatomical Specimens *(ie were contralateral images used)*
- Population Type *(Cadaver, Patient, Dry Bone)*
- Gender Distribution
- Age (*Range)*
- Underlying Pathology
- SSM application domain *(ie surgical planning, custom implant design)*
- Measure of accuracy *(RMSE alongside the relevant application variables – ie. neck/shaft angle, acetabular cup placement)*

## 5 | CONCLUSION

This systematic review provides evidence that statistical shape modelling has the potential to accurately reconstruct detailed 3D anatomical models from standard 2D imaging. Prior to acceptance in healthcare, clinical validation studies are required with sufficient sample sizes and varying populations. A standard set of reporting guidelines has been proposed, to facilitate the analysis and comparison of SSMs in future studies, particularly with regards to the detailed demographics of the training and validation sets, the required variables necessary for the application, and the method of reporting model accuracy.

## AUTHOR CONTRIBUTIONS

Gareth G. Jones, Amogh Patil and Shuqiao Xie contributed to study conception and design. Data collection was performed by Amogh Patil and Krishan Kulkarni. Data analysis was performed by Amogh Patil, Krishan Kulkarni, Shuqiao Xie, Anthony M. J. Bull and Gareth G. Jones. Original manuscript was prepared by Amogh Patil. The final manuscript was reviewed and edited by Krishan Kulkarni, Shuqiao Xie, Anthony M. J. Bull and Gareth G. Jones.

## ACKNOWLEDGEMENTS

We would like to thank the MSk Lab, Department of Surgery and Cancer and Department of Bioengineering at Imperial College London for their support in producing this manuscript. No funding was sought or received in the production of this manuscript.

## CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interests.

## DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the published references provided within the text. Any other data that is required may be requested from the corresponding author upon reasonable request.

## ORCID

*Amogh Patil* https://orcid.org/0000-0003-3737-8448
*Anthony M. J. Bull* https://orcid.org/0000-0002-4473-8264

## REFERENCES

1. Fadero PE, Shah M. Three dimensional (3D) modelling and surgical planning in trauma and orthopaedics. *Surgeon*. 2014;12:328-333.
2. Galvez M, Asahi T, Baar A, et al Use of three-dimensional printing in orthopaedic surgical planning. *JAAOS Glob Res Rev*. 2018;2(5):e071. https://doi.org/10.5435/jaaosglobal-d-17-00071
3. Neubert A, Wilson KJ, Engstrom C, et al Comparison of 3D bone models of the knee joint derived from CT and 3T MR imaging. *Eur J Radiol*. 2017;93:178-184. https://doi.org/10.1016/j.ejrad.2017.05.042
4. Rathnayaka K, Momot KI, Noser H, et al Quantification of the accuracy of MRI generated 3D models of long bones compared to CT generated 3D models. *Med Eng Phys*. 2012;34(3):357-363. https://doi.org/10.1016/j.medengphy.2011.07.027
5. Kang D.-G, Kim K.-I, Bae J.-K. MRI-based or CT-based patient-specific instrumentation in total knee arthroplasty: how do the

1478596x, 2023, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rcs.2503 by Test, Wiley Online Library on [18/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

two systems compare? *Arthroplasty*. 2020;2(1):1. https://doi.org/10.1186/s42836-019-0020-6
6. White D, Chelule KL, Seedhom BB. Accuracy of MRI vs CT imaging with particular reference to patient specific templates for total knee replacement surgery. *Int J Med Robot Comput Assist Surg*. 2008;4(3):224-231. https://doi.org/10.1002/rcs.201
7. Livyatan H, Yaniv Z, Joskowicz L. Gradient-based 2-D/3-D rigid registration of fluoroscopic X-ray to CT. *IEEE Trans Med Imag*. 2003;22(11):1395-1406. https://doi.org/10.1109/tmi.2003.819288
8. Garvey CJ, Hanlon R. Computed tomography in clinical practice. *Br Med J*. 2002;324:1077-1080.
9. Grover VPB, Tognarelli JM, Crossey MME, Cox IJ, Taylor-Robinson SD, McPhail MJW. Magnetic resonance imaging: principles and techniques: lessons for clinicians. *J Clin Exp Hepatol*. 2015;5:246-255.
10. Markelj P, Tomaževič D, Likar B, Pernuš F. A review of 3D/2D registration methods for image-guided interventions. *Med Image Anal*. 2012;16(3):642-661. https://doi.org/10.1016/j.media.2010.03.005
11. Hyndman RJ, Koehler AB. Another look at measures of forecast accuracy. *Int J Forecast*. 2006;22(4):679-688. https://doi.org/10.1016/j.ijforecast.2006.03.001
12. Liberati A, Altman DG, Tetzlaff J, et al The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ*. 2009;339(jul21 1):b2700. https://doi.org/10.1136/bmj.b2700
13. Moons KGM, de Groot JAH, Bouwmeester W, et al Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the CHARMS checklist. *PLoS Med*. 2014;11(10):e1001744. https://doi.org/10.1371/journal.pmed.1001744
14. Wolff RF, Moons KGM, Riley RD, et al PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. *Ann Intern Med*. 2019;170(1):51. https://doi.org/10.7326/M18-1376
15. Baka N, De Bruijne M, Van Walsum T, et al Statistical shape model-based femur kinematics from biplane fluoroscopy. *IEEE Trans Med Imag*. 2012;31(8):1573-1583. https://doi.org/10.1109/TMI.2012.2195783
16. Baka N, Kaptein BL, de Bruijne M, et al 2D-3D shape reconstruction of the distal femur from stereo X-ray imaging using statistical shape models. *Med Image Anal*. 2011;15(6):840-850. https://doi.org/10.1016/j.media.2011.04.001
17. Mutsvangwa T, Wasswa W, Burdin V, Borotikar B, Douglas TS. Interactive patient-specific 3D approximation of scapula bone shape from 2D X-ray images using landmark-constrained statistical shape model fitting. In: *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society EMBS*; 2017:1816-1819.
18. Picazo ML, Baro AM, Del Rió Barquero LM, et al 3-D subject-specific shape and density estimation of the lumbar spine from a single anteroposterior DXA image including assessment of cortical and trabecular bone. *IEEE Trans Med Imag*. 2018;37(12):2651-2662. https://doi.org/10.1109/TMI.2018.2845909
19. Sadowsky O, Chintalapani G, Taylor RH. Deformable 2D-3D registration of the pelvis with a limited field of view, using shape statistics. *Lect Notes Comput Sci*. 2007;4792 LNCS(PART 2):519-526.
20. Schumann S, Bieck R, Bader R, Heverhagen J, Nolte LP, Zheng G. Radiographic reconstruction of lower-extremity bone fragments: a first trial. *Int J Comput Assist Radiol Surg*. 2016;11(12):2241-2251. https://doi.org/10.1007/s11548-016-1427-y
21. Schumann S, Liu L, Tannast M, Bergmann M, Nolte LP, Zheng G. An integrated system for 3D hip joint reconstruction from 2D X-rays: a preliminary validation study. *Ann Biomed Eng*. 2013;41(10):2077-2087. https://doi.org/10.1007/s10439-013-0822-6
22. Smoger LM, Shelburne KB, Cyr AJ, Rullkoetter PJ, Laz PJ. Statistical shape modeling predicts patellar bone geometry to enable stereoradiographic kinematic tracking. *J Biomech*. 2017;58:187-194. https://doi.org/10.1016/j.jbiomech.2017.05.009
23. Spinelli M, Taddei F, Humbert L, et al Fe prediction of strains for patient-specific proximal femurs generated from a single dxa. *J Biomech*. 2012;45(3):S538. https://doi.org/10.1016/S0021-9290(12)70539-0
24. Tang TSY, Ellis RE. 2D/3D deformable registration using a hybrid atlas. *Lect Notes Comput Sci*. 2005;3750 LNCS:223-230.
25. Tsai TY, Li JS, Wang S, Li P, Kwon YM, Li G. Principal component analysis in construction of 3D human knee joint models using a statistical shape model method. *Comput Methods Biomech Biomed Eng*. 2015;18(7):721-729. https://doi.org/10.1080/10255842.2013.843676
26. Väänänen SP, Grassi L, Flivik G, Jurvelin JS, Isaksson H. Generation of 3D shape, density, cortical thickness and finite element mesh of proximal femur from a DXA image. *Med Image Anal*. 2015;24(1):125-134. https://doi.org/10.1016/j.media.2015.06.001
27. Baka N, Kaptein BL, Giphart JE, et al Evaluation of automated statistical shape model based knee kinematics from biplane fluoroscopy. *J Biomech*. 2014;47(1):122-129. https://doi.org/10.1016/j.jbiomech.2013.09.022
28. Whitmarsh T, Humbert L, De Craene M, Del Rio Barquero LM, Frangi AF. Reconstructing the 3D shape and bone mineral density distribution of the proximal femur from dual-energy x-ray absorptiometry. *IEEE Trans Med Imag*. 2011;30(12):2101-2114. https://doi.org/10.1109/tmi.2011.2163074
29. Whitmarsh T, Humbert L, De Craene M, et al 3D bone mineral density distribution and shape reconstruction of the proximal femur from a single simulated DXA image: an in vitro study. *Med Imaging 2010 Image Process*. 2010;7623:76234U. https://doi.org/10.1117/12.844110
30. Whitmarsh T, Humbert L, Del Río Barquero LM, Di Gregorio S, Frangi AF. 3D reconstruction of the lumbar vertebrae from anteroposterior and lateral dual-energy X-ray absorptiometry. *Med Image Anal*. 2013;17(4):475-487. https://doi.org/10.1016/j.media.2013.02.002
31. Wu J, Mahfouz MR. Reconstruction of knee anatomy from single-plane fluoroscopic x-ray based on a nonlinear statistical shape model. *J Med Imaging*. 2021;8(01):1-21. https://doi.org/10.1117/1.jmi.8.1.016001
32. Zheng G. Statistical shape model-based reconstruction of a scaled, patient-specific surface model of the pelvis from a single standard AP x-ray radiograph. *Med Phys*. 2010;37(4):1424-1439. https://doi.org/10.1118/1.3327453
33. Zheng G, Gollmer S, Schumann S, Dong X, Feilkas T, González Ballester MA. A 2D/3D correspondence building method for reconstruction of a patient-specific 3D bone surface model using point distribution models and calibrated X-ray images. *Med Image Anal*. 2009;13(6):883-899. https://doi.org/10.1016/j.media.2008.12.003
34. Zheng G, Nolte LP, Ferguson SJ. Scaled, patient-specific 3D vertebral model reconstruction based on 2D lateral fluoroscopy. *Int J Comput Assist Radiol Surg*. 2011;6(3):351-366. https://doi.org/10.1007/s11548-010-0515-7
35. Zheng G, Nolte LP, Ferguson SJ. 2D/3D reconstruction of a scaled lumbar vertebral model from a single fluoroscopic image. In: *Annual International Conference of the IEEE Engineering in Medicine and Biology Society EMBS'10*; 2010:4395-4398.
36. Zheng G, Schumann S. 3D reconstruction of a patient-specific surface model of the proximal femur from calibrated x-ray radiographs: a validation study. *Med Phys*. 2009;36(4):1155-1166. https://doi.org/10.1118/1.3089423
37. Zheng G, Schumann S, González Ballester MA. An integrated approach for reconstructing a surface model of the proximal femur

1478596x, 2023, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rcs.2503 by Test, Wiley Online Library on [18/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

from sparse input data and a multi-resolution point distribution model: an in vitro study. *Int J Comput Assist Radiol Surg*. 2010;5(1):99-107. https://doi.org/10.1007/s11548-009-0386-y
38. Cerveri P, Sacco C, Olgiati G, Manzotti A, Baroni G. 2D/3D reconstruction of the distal femur using statistical shape models addressing personalized surgical instruments in knee arthroplasty: a feasibility analysis. *Int J Med Robot Comput Assist Surg*. 2017;13(4):1-13. https://doi.org/10.1002/rcs.1823
39. Zhu Z, Li G. Construction of 3D human distal femoral surface models using a 3D statistical deformable model. *J Biomech*. 2011;44(13):2362-2368. https://doi.org/10.1016/j.jbiomech.2011.07.006
40. Grassi L, Fleps I, Sahlstedt H, et al Validation of 3D finite element models from simulated DXA images for biofidelic simulations of sideways fall impact to the hip. *Bone*. 2021;142(July 2020):115678. https://doi.org/10.1016/j.bone.2020.115678
41. Gajny L, Girinon F, Bayoud W, et al Fast quasi-automated 3D reconstruction of lower limbs from low dose biplanar radiographs using statistical shape models and contour matching. *Med Eng Phys*. 2022;101(December 2021):103769. https://doi.org/10.1016/j.medengphy.2022.103769
42. Lu HY, Shih KS, Lin CC, et al Three-dimensional subject-specific knee shape reconstruction with asynchronous fluoroscopy images using statistical shape modeling. *Front Bioeng Biotechnol*. 2021;9:1-13. https://doi.org/10.3389/fbioe.2021.736420
43. Dong X, Zheng G. Automatic extraction of proximal femur contours from calibrated X-ray images using 3D statistical models: an in vitro study. *Int J Med Robot Comput Assist Surg*. 2009;5(2):213-222. https://doi.org/10.1002/rcs.253
44. Humbert L, Martelli Y, Fonolla R, et al 3D-DXA: assessing the femoral shape, the trabecular macrostructure and the cortex in 3D from DXA images. *IEEE Trans Med Imag*. 2017;36(1):27-39. https://doi.org/10.1109/TMI.2016.2593346
45. Hurvitz A, Joskowicz L. Registration of a CT-like atlas to fluoroscopic X-ray images using intensity correspondences. *Int J Comput Assist Radiol Surg*. 2008;3(6):493-504. https://doi.org/10.1007/s11548-008-0264-z
46. Klima O, Kleparnik P, Spanel M, Zemcik P. Intensity-based femoral atlas 2D/3D registration using Levenberg-Marquardt optimisation. *Med Imaging 2016 Biomed Appl Mol Struct Funct Imaging*. 2016;9788:97880F.
47. Kurazume R, Nakamura K, Okada T, et al 3D reconstruction of a femoral shape using a parametric model and two 2D fluoroscopic images. *Comput Vis Image Understand*. 2009;113(2):202-211. https://doi.org/10.1016/j.cviu.2008.08.012
48. Li JS, Tsai TY, Wang S, et al Prediction of in vivo knee joint kinematics using a combined dual fluoroscopy imaging and statistical shape modeling technique. *J Biomech Eng*. 2014;136(12). https://doi.org/10.1115/1.4028819
49. Hoaglund FT, Low WD. Anatomy of the femoral neck and head, with comparative data from Caucasians and Hong Kong Chinese. *Clin Orthop Relat Res*. 1980;152:10-16.
50. Craiovan B, Renkawitz T, Weber M, Grifka J, Nolte L, Zheng G. Is the acetabular cup orientation after total hip arthroplasty on a two dimension or three dimension model accurate? *Int Orthop*. 2014;38(10):2009-2015.
51. Zheng G, Von Recum J, Nolte LP, Grützner PA, Steppacher SD, Franke J. Validation of a statistical shape model-based 2D/3D reconstruction method for determination of cup orientation after THA. *Int J Comput Assist Radiol Surg*. 2012;7(2):225-231. https://doi.org/10.1007/s11548-011-0644-7
52. Zheng G. Statistically deformable 2D/3D registration for estimating post-operative cup orientation from a single standard AP X-ray radiograph. *Ann Biomed Eng*. 2010;38(9):2910-2927. https://doi.org/10.1007/s10439-010-0060-0
53. Nolte D, Bull AMJ. Femur finite element model instantiation from partial anatomies using statistical shape and appearance models. *Med Eng Phys*. 2019;67:55-65. https://doi.org/10.1016/j.medengphy.2019.03.007
54. Nolte D, Ko ST, Bull AMJ, Kedgley AE. Reconstruction of the lower limb bones from digitised anatomical landmarks using statistical shape modelling. *Gait Posture*. 2020;77:269-275. https://doi.org/10.1016/j.gaitpost.2020.02.010
55. Ast MP, Nam D, Haas SB. Patient-specific instrumentation for total knee arthroplasty: a review. In: *Orthopedic Clinics of North America*. Vol 43. Elsevier; 2012:e17-e22.
56. Zimmer Biomet Introduces the World's First CE Marked, X-Ray-Based Patient Specific Instrument System for Total Knee Replacement Surgery [Internet]. [cited 2022 Apr 30]. https://www.prnewswire.com/news-releases/zimmer-biomet-introduces-the-worlds-first-ce-marked-x-ray-based-patient-specific-instrument-system-for-total-knee-replacement-surgery-300481390.html
57. Asseln M, Hänisch C, Schick F, Radermacher K. Gender differences in knee morphology and the prospects for implant design in total knee replacement. *Knee*. 2018;25(4):545-558. https://doi.org/10.1016/j.knee.2018.04.005
58. Pavlova AV, Saunders FR, Muthuri SG, et al Statistical shape modelling of hip and lumbar spine morphology and their relationship in the MRC National Survey of Health and Development. *J Anat*. 2017;231(2):248-259. https://doi.org/10.1111/joa.12631

**How to cite this article:** Patil A, Kulkarni K, Xie S, Bull AMJ, Jones GG. The accuracy of statistical shape models in predicting bone shape: a systematic review. *Int J Med Robot*. 2023;19(3):e2503. https://doi.org/10.1002/rcs.2503

1478596x, 2023, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rcs.2503 by Test, Wiley Online Library on [18/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License