CurveBeam AI

ADVANCING ACCURACY IN ORTHOPEDIC IMAGING

WHITE PAPER

Clinical Accuracy of Formus Labs' Atlas for CBCT Segmentation

Executive Overview

This study evaluated the accuracy of Formus Atlas automated segmentation service against expert manual annotation on cone beam CT (CBCT), and assessed whether a higher-dose protocol offers any clinically meaningful advantage over the standard-dose protocol.

Formus (Auckland, New Zealand) Atlas software is an AI-driven image processing solution that automatically segments medical imaging data performing image segmentation, landmark identification, and quantitative measurements to rapidly generate accurate, high-fidelity 3D anatomical models for patient-specific orthopedic workflows. *Atlas is in developmment and is not available for clinical use.*

Results: Across four representative orthopedic specimens, Atlas demonstrated sub-millimetric mean surface distance errors relative to manual annotation, as well as protocol-agnostic performance between higher and standard dose for most cases. In the most anatomically complex case, the higher dose protocol provided clearly better agreement with manual segmentation, suggesting a role as a targeted high-fidelity option when additional geometric certainty is warranted despite the higher radiation exposure.

Methods

Four cadaveric specimens (C1–C4) underwent bilateral CBCT and multi-detector CT (MDCT) imaging with two protocols: comparing segmentations of CBCT vs MDCT (Group 1) and comparing segmentations of standard dose (CBCT-STD) vs higher dose (CBCT-LRG) settings for CBCT (Group 2). Manual segmentations for femur, tibia, fibula, patella, and pelvis served as the reference standard.

The manual segmentations were performed by undergraduate biomedical engineering students using 3D Slicer. The results of this investigation were published in a previous whitepaper.

Atlas was then used to generate automated segmentations on the same DICOM scans of the four specimens.

Two core comparisons were made. Accuracy was quantified as mean surface distance (mm) between paired STL models, with associated standard deviation (SD) as a measure of variability.

Group 3: CBCT-LRG vs CBCT-STD assessed whether protocol choice altered Atlas outputs when both scans were segmented automatically.

Results:

· C1: mean 0.164 mm, SD 0.74 mm

· C2: mean −0.073 mm, SD 0.31 mm

· C3: mean 0.000 mm, SD 0.67 mm

· C4: mean −0.011 mm, SD 0.34 mm

Group 4: Atlas vs manual segmentation to evaluate algorithmic accuracy for CBCT-STD and CBCT-LRG protocols.

Results:

· CBCT-STD, pooled mean ≈ 0.29 mm (per-sample 0.06, −0.06, 0.36, 0.80 mm)

· CBCT-LRG, pooled mean ≈ 0.15 mm (per-sample 0.04, −0.00, 0.36, 0.20 mm)

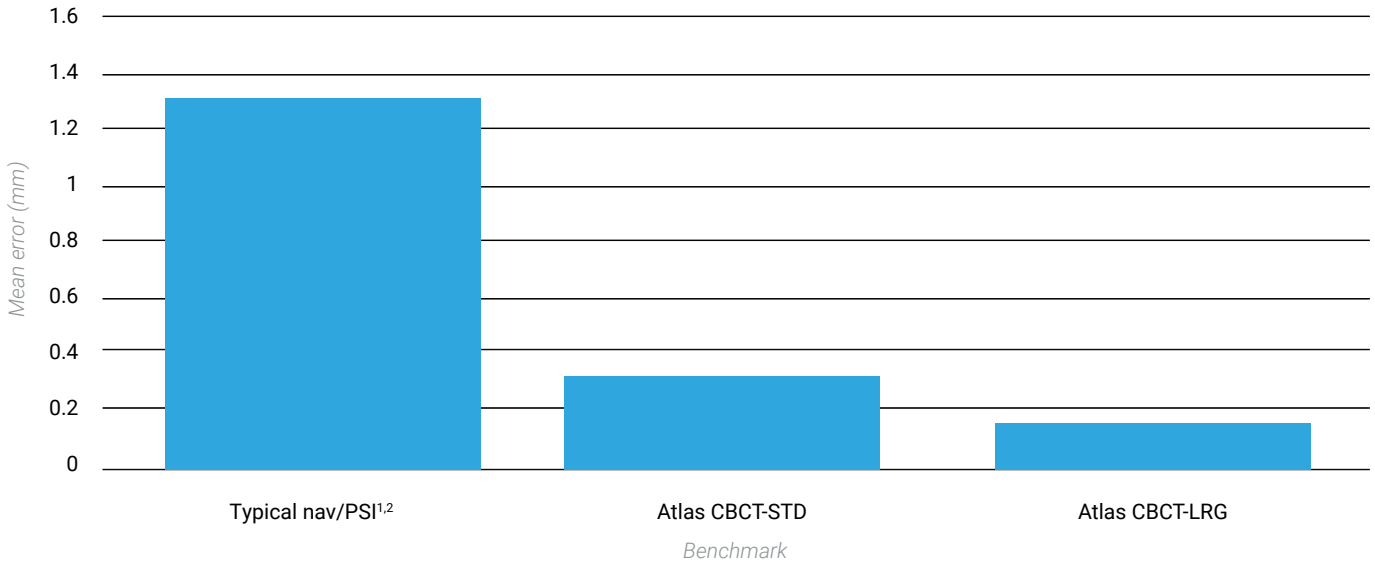
Detailed bone-level tables and overlay images illustrate performance per bone and per side.

CurveBeam AI

ADVANCING ACCURACY IN ORTHOPEDIC IMAGING

WHITE PAPER

Graph 1. Atlas Segmentation Error vs Typical Tolerances (mm)



Results

Protocol Effect (Group 3)

Across all four specimens, the mean difference between CBCT-LRG and CBCT-STD segmentations produced by Atlas remained very close to zero, with pooled mean approximately 0.02 mm and SD around 0.51 mm. Individual samples ranged from −0.073 to 0.164 mm, with no consistent directional bias and narrow variability for C2–C4.

Bone-wise tables show:

· Femur, tibia, fibula, and patella: mean differences typically within ±0.3 mm.

· Pelvis: larger SDs in some samples, but mean differences still small and centered near zero.

For surgeons and radiologists, this implies that, from the perspective of Atlas' geometric output, CBCT-STD and CBCT-LRG produce practically interchangeable segmentations in the majority of cases. This supports using the standard-dose protocol as a default, reserving the higher-dose CBCT-LRG for specific indications rather than as a routine requirement.

Overall algorithmic accuracy

When comparing Atlas to manual annotation on CBCT-STD, mean surface differences by specimen were 0.061 mm (C1), −0.058 mm (C2), 0.361 mm (C3), and 0.797 mm (C4), with pooled mean around 0.29 mm and SD 1.14 mm. On CBCT-LRG, the corresponding means were 0.042 mm, −0.005 mm, 0.358 mm, and 0.200 mm, with pooled mean about 0.15 mm and SD 0.98 mm.

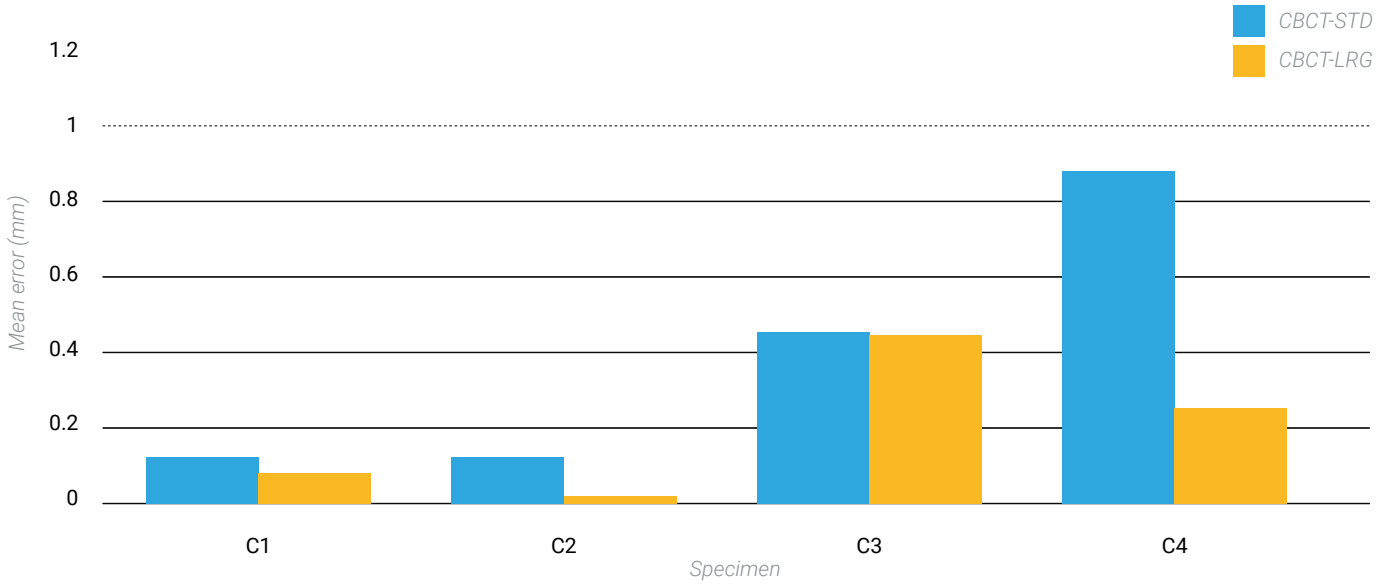
These mean surface differences are well below accuracy magnitudes commonly reported for orthopedic surgical navigation and patient-specific cutting guide workflows (approximately 1.0–1.5 mm) and are within, or smaller than, published inter-observer variability for manual orthopedic bone segmentation. In practical

CurveBeam AI

ADVANCING ACCURACY IN ORTHOPEDIC IMAGING

WHITE PAPER

Graph 2. Atlas vs Manual Mean Surface Distance by Specimen (mm)



terms, Atlas behaves like a highly consistent "second expert," with average discrepancies well below 0.5 mm.[1,2]

Sample-level complexity and C4 outlier

The most difficult case was C4, which represents a patient with a BMI of 40 and significant bone remodeling, with higher variance in both manual and automated measurements. In this specimen:

· CBCT-STD: Atlas vs manual mean ≈ 0.80 mm, SD ≈ 1.17 mm.

· CBCT-LRG: Atlas vs manual mean ≈ 0.20 mm, SD ≈ 0.99 mm.

Here, the higher-dose CBCT-LRG protocol yields a four-fold reduction in mean error versus CBCT-STD, suggesting that, for patients of a certain size, the additional radiation can translate into clinically meaningful geometric gains. In the remaining specimens (C1–C3), Atlas performance is similar across both protocols, with mean differences ≤0.36 mm and no strong protocol dependence.

Clinical interpretation for orthopedic practice

For surgeons

· Manual-grade accuracy: Atlas' average differences relative to expert annotation (≈0.15–0.29 mm depending on protocol) are well within accepted error budgets for preoperative planning, patient-specific instrumentation, and 3D printed guides. Surgeons can treat Atlas segmentations as clinically equivalent to well-performed manual work, especially for routine primary arthroplasty and straightforward deformity cases.

· Protocol-agnostic in typical cases: For most hips and lower-extremity reconstructions, CBCT-STD scans provide enough image quality for Atlas to generate accurate models, as confirmed by the near-zero LRG vs STD differences in Group 3 and the close Atlas–manual agreement in C1–C3.

· High-dose protocol for edge cases: In patients with a higher BMI (evidence here was 40 or greater) or tight custom implant fits, the improved accuracy observed with CBCT-LRG in specimen C4 (0.20 mm vs 0.80 mm) supports selectively choosing the higher-dose protocol where added certainty is worth the exposure.

For radiologists

· Default protocol recommendation: CBCT-STD should be the default protocol for Atlas-enabled workflows, achieving high segmentation accuracy with lower radiation and broad applicability.

· Escalation criteria for CBCT-LRG: CBCT-LRG can be defined as a "high-fidelity mode" reserved for:

· Patients with a higher BMI (40).

· Significant bone remodeling.

· Research or early evaluation of new implant systems where maximal geometric fidelity is required.

· Radiation stewardship: Positioning CBCT-LRG this way aligns with As Low As Reasonably Achievable (ALARA) principles. Dose increase is targeted, justified, and limited to cases with demonstrable potential benefit, while routine work remains on a lower-dose protocol.

Practical protocol strategy

A clear, tiered protocol strategy emerges from these data:

Tier 1 – Routine cases (default):

· Imaging: CBCT-STD
· Segmentation: Atlas

· Appropriate for routine lower extremity reconstruction where segmentation accuracy within sub millimetric tolerances is acceptable in patients with a BMI less than 40.

· Rationale: Excellent accuracy with lower radiation and minimal need for escalation.

Tier 2 – High-complexity cases (selective):

· Imaging: CBCT-LRG
· Segmentation: Atlas

· May be considered in cases requiring maximal geometric fidelity, such as complex anatomy or tight tolerance implant implications.

· Rationale: Improved Atlas–manual agreement in complex anatomy (as in C4) can justify higher dose when geometry is critical.

In both tiers, a short, standardized visual review of Atlas output (overlay on CBCT) provides an additional safety net without undermining the time savings.

Conclusion

For orthopedic surgeons and clinical administrators, this study supports three clear messages:

1. Atlas is accurate enough for routine clinical use, with sub-millimetric mean differences to manual annotation on both CBCT-STD and CBCT-LRG across typical cases.

2. CBCT-STD should remain the default protocol for Atlas-based workflows, as protocol-agnostic performance in Group 3 and strong Atlas–manual agreement in C1–C3 indicate that standard-dose images are sufficient in most scenarios.

3. CBCT-LRG, despite higher radiation, is clinically justifiable in selected complex cases, where the observed improvement in segmentation concordance (e.g., 0.20 mm vs 0.80 mm in C4) can meaningfully de-risk surgical planning and implant design.

This tiered approach balances accuracy, dose, and workflow efficiency, and provides a defensible framework for protocol selection in Atlas-enabled orthopedic imaging services.

1. Yeo, Cheng Hong et al. “Assessing the accuracy of bone resection by cutting blocks in patient-specific total knee replacements.” ISRN orthopedics vol. 2012 509750. 20 May. 2012, doi:10.5402/2012/509750
2. Çalbıyık, Murat. “Clinical Outcome of Total Knee Arthroplasty Performed Using Patient-Specific Cutting Guides.” Medical science monitor : international medical journal of experimental and clinical research vol. 23 6168-6173. 29 Dec. 2017, doi:10.12659/msm.908213