EVALUATION OF AUTOMATIC CLASS III DESIGNATION FOR
OSTEODETECT

DECISION SUMMARY

A. De Novo Number:

DEN180005

B. Purpose for Submission:

De novo request for evaluation of automatic class III designation for Imagen Technologies, OsteoDetect

C. Applicant:

Imagen Technologies

D. Proprietary and Established Names:

OsteoDetect

E. Regulatory Information:

1. Regulation section:

   21 CFR 892.2090

2. Classification:

   Class II Special Controls

3 Product code:

   QBS

4. Panel:

   90 (Radiology)

F. Indications for Use:

1. Indications for Use:

   OsteoDetect analyzes wrist radiographs using machine learning techniques to identify

and highlight distal radius fractures during the review of posterior-anterior (PA) and lateral (LAT) radiographs of adult wrists.

2. Special Conditions for Use Statement(s):

   For prescription use only

3. Warnings, precautions, and limitations:

   Clinicians should review OsteoDetect annotated images concurrently with original images before making a final determination on a case. OsteoDetect is an adjunct tool and does not replace the role of the clinician. Clinicians must not use the CAD generated output as the primary interpretation.

   OsteoDetect is not designed to detect fractures other than distal radius fractures. Clinicians should review original images for all suspected pathologies.

   OsteoDetect may not detect a distal radius fracture in both the PA and LAT views. Clinicians should follow standard clinical procedures in assessing PA and LAT views.

   All images must have proper DICOM tags. OsteoDetect performance may be reduced on images that do not have correct DICOM tags.

G. Device Description:

OsteoDetect is a software device designed to assist clinicians in detecting distal radius fractures during the review of posterior-anterior (PA) and lateral (LAT) radiographs of adult wrists. The software uses deep learning techniques to analyze wrist radiographs (PA and LAT views) for distal radius fracture in adult patients.

Intended User Population:

The intended users of OsteoDetect are clinicians in various settings including primary care (e.g., family practice, internal medicine), emergency medicine, urgent care, and specialty care (e.g., orthopedics), as well as radiologists who review radiographs across settings.

When a clinician accesses the patient radiographs in a picture archiving and communication system (PACS) client, both the OsteoDetect annotated radiographs and the original, unaltered radiographs are available in the same patient study, allowing for concurrent reading.

OsteoDetect is an adjunct tool and is not intended to replace a clinician's review of the radiograph or his or her clinical judgment. Clinicians must not use the CAD generated output as the primary interpretation.

Intended Patient Population:

The target population of the device is adults (patients age 22 and older) suspected of having a distal radius fracture.

Compatible Radiological Data Sources:

The radiological data used as input for OsteoDetect are PA and LAT wrist radiographs acquired with Philips Medical Systems DigitalDiagnost, Carestream Health DRX-1, and GE Discovery XR656.

Hardware Requirements:

OsteoDetect requires a dedicated physical machine with the following minimum specifications:

- Intel Xeon E3-1225 v5 or comparable (with at least 2 CPU cores)
- 16GB RAM or more
- 100 GB or more free space
- 100 Mbps or faster Ethernet interface to your institution's DICOM network

It is strongly recommended that the machine is certified compatible with ESXi.

Software Requirements:

OsteoDetect should be installed on top of the vSphere Hypervisor, ESXi 6.0 Update 2.

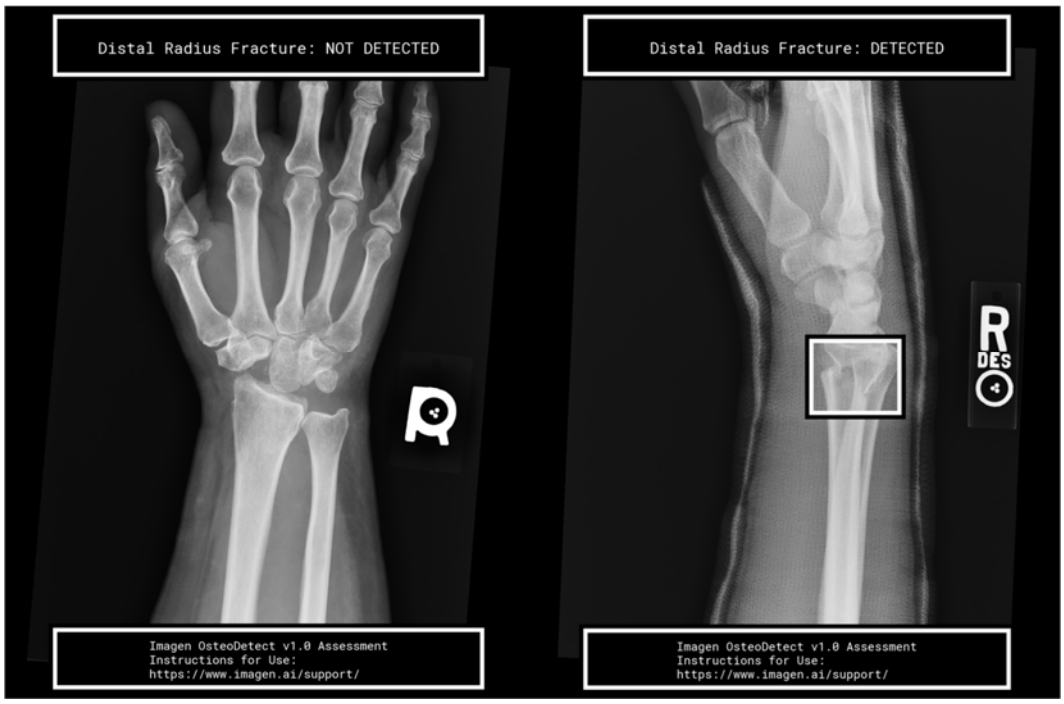
1. Description of user interface and outputs

   The end user interacts with the output of OsteoDetect in the facility's PACS software. Annotated DICOM images processed by OsteoDetect are to be read concurrently by the clinician along with the original, unannotated radiographs.

   If a distal radius fracture is detected by OsteoDetect, the software annotates the image by creating a bounding box surrounding the fracture with the label "Distal Radius Fracture: DETECTED." If none is detected, the message label "Distal Radius Fracture: NOT DETECTED" appears. Whether or not the software detects a distal radius fracture, when OsteoDetect analyzes an image it annotates the radiograph to include a message stating that the radiograph was analyzed by OsteoDetect, and includes instructions for users to access device labeling.

   The annotated radiographs are available in the same patient study as the unaltered radiographs, allowing for a concurrent read. The OsteoDetect device is not intended to be used as a primary read for the diagnosis of distal radius fractures; annotated images should always be read concurrently with the original, unannotated radiographs according

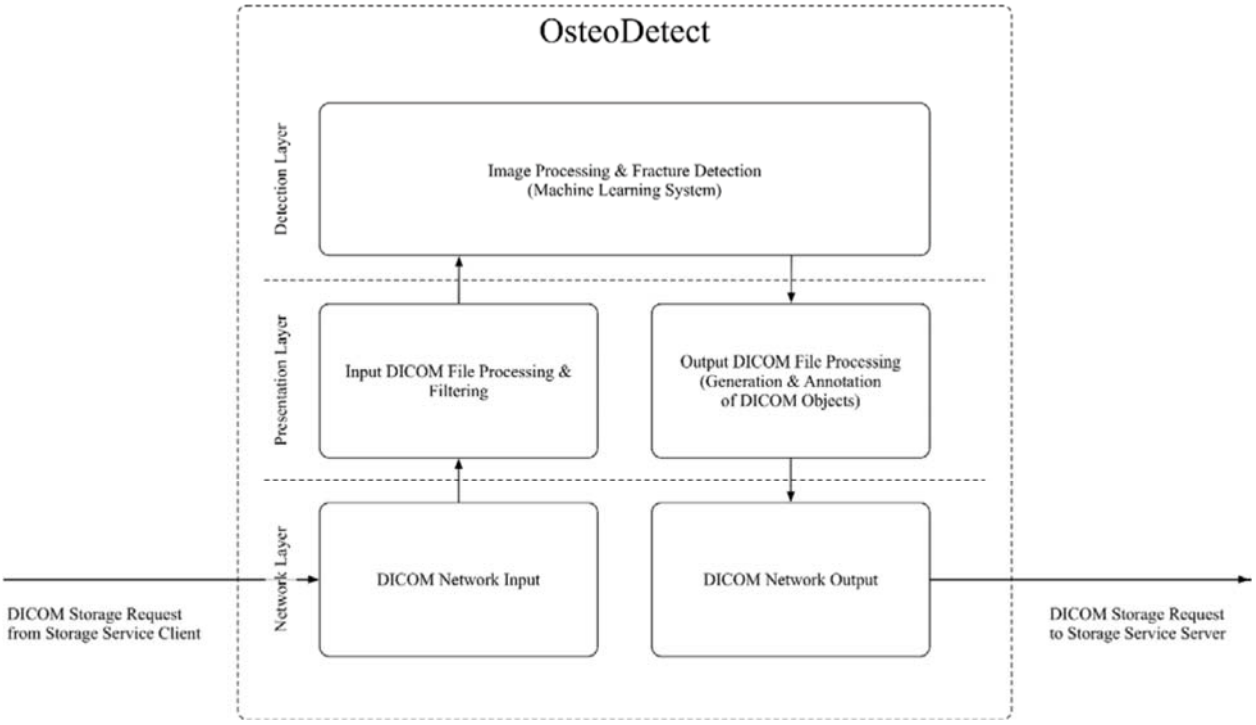
to standard of practice.



The images above are an example of the OsteoDetect annotated images after a wrist radiograph is analyzed by the algorithm (left when no fracture is detected, and right when a fracture is detected and a bounding box is generated on the image).

2. Technological Characteristics and Software Algorithm Summary

OsteoDetect is a software-only device which operates in three layers - a Network Layer, a Presentation Layer, and a Decision Layer (as described in the data flow diagram below):



Workflow of the device begins and ends at the Network layer. The Network Input component accepts Computed Radiography and Digital Radiography (CR and DR) DICOM images and passes them to the Input DICOM File Processing & Filtering.

Upon receiving a DICOM object, the Input DICOM File Processing & Filtering component applies a number of filtering rules based on image characteristics and DICOM tags (age, modality, anatomy, contrast) to ensure that only eligible images are analyzed by the algorithm.

The Image Processing & Fracture Detection (Machine Learning System) processes eligible DICOM objects and analyses for the presence or absence of distal radius fractures. It is comprised of three modules:

- Image Preprocessor module – the image undergoes image preprocessing.
- Fracture Detector module – this module analyzes the processed image and generates two outputs: a confidence score (confidence for presence of distal radius fracture), and a conditional probability map encoding the location of the fracture (if present).
- Image Postprocessor module – determines whether a fracture is present based on the value of the confidence score and (if present) generates the coordinates of the fracture bounding box.

The Output DICOM File Processing component creates a DICOM image containing the original radiograph with a fracture determination, a message stating that the image was analyzed by OsteoDetect (with link to instructions/labeling), and a bounding box containing the suspected fracture (if detected).

The DICOM Network Output sends the OsteoDetect-annotated DICOM images back to the DICOM Storage Server.

H. Standard/Guidance Document Referenced

The following device-specific FDA guidance documents are relevant to OsteoDetect and were followed in the preparation of this submission:

- Computer-Assisted Detection Devices Applied to Radiology Images and Radiology Device Data – Premarket Notification [510(k)] Submissions" (July 3, 2012)
- "Clinical Performance Assessment: Considerations for Computer-Assisted Detection Devices Applied to Radiology Images and Radiology Device Data – Premarket Approval (PMA) and Premarket Notification [510(k)] Submissions" (July 3, 2012)

I. Performance Characteristics:

The device is a software-only device. Some common performance characteristics for other device types are included below with a note that these characteristics are not applicable to this type of software-only device.

1. Biocompatibility/Materials

   Not applicable

2. Shelf Life/Sterility

   Not applicable

3. Electromagnetic Compatibility and Electrical Safety

   Not applicable

4. Magnetic Resonance (MR) Compatibility

   Not applicable

Nonclinical performance data were provided to address the following areas:

5. Software

   The device is a software only device.

   Imagen Technologies provided software documentation at a Moderate Level of Concern according to the "Guidance for the Content of Premarket Submissions for Software Contained in Medical Devices" (May 11, 2005).

Version: OsteoDetect 1.0
Level of Concern: Moderate
Software Description: Imagen provided a general description of the features in the software documentation and in the device description. The description of the software is consistent with the device functionality described in the device description.
Device Hazard Analysis: Imagen provided separate analyses of the device and cybersecurity concerns. The content of the hazard analysis is sufficient and assesses pre- and post- mitigation risks. The device hazard analysis includes: • identification of the hazardous event • severity of the hazard • probability of the hazard • cause(s) of the hazard • method of control or mitigation • corrective measures taken, including an explanation of the aspects of the device design/requirements, that eliminate, reduce, or warn of a hazardous event • verification of the control implementation, which is traceable through the enumerated traceability matrix
Software Requirement Specifications (SRS): The SRS includes user, engineering, algorithmic, cybersecurity, and various other types of requirements that give a full description of the functionality of the device. The SRS is consistent with the device description and software description.
Architecture Design Chart: The architecture design chart provides the software overview and includes flow diagrams representative of process flow for various features of the OsteoDetect software.
Software Design Specification (SDS): The SDS is traceable to the SRS and demonstrates how individual requirements are implemented in the software design and includes appropriate linkages to predefined verification testing.
Traceability Analysis/Matrix: Imagen provided traceability between all documents including the SRS, SDS, and subsequent verification and validation. Hazards mitigations are traceable throughout all documents.
Software Development Environment: Imagen outlined the software development environment and the processes/procedures used for medical device software development. The content is consistent with expected quality system norms.
Verification and Validation Testing: The validation and system level verifications procedures are based upon the requirements with clearly defined test procedures and pass/fail criteria. All tests passed. Unit level test procedures, actual, and expected results are included for all design specifications.
Revision Level History: This is the initial version 1.0 for the device. This is the version described in the standalone and reader performance testing reports.
Unresolved Anomalies: Imagen stated that there are no unresolved anomalies.
Cybersecurity: The cybersecurity documentation is consistent with the recommendations for information that should be included in premarket submissions outlined in the FDA guidance document “Content of Premarket Submissions for

Management of Cybersecurity in Medical Devices: Guidance for Industry and Food and Drug Administration Staff" (issued October 2, 2014). Information related to cybersecurity reviewed included: Hazard analysis related to cybersecurity risks; traceability documentation linking cybersecurity controls to risks considered; summary plan for validating software updates and patches throughout the lifecycle of the medical device; summary describing controls in place to ensure that the medical device will maintain its integrity; and device instructions for use and product specifications related to recommended cybersecurity controls appropriate for the intended use of the device.

6. Standalone Performance Testing Protocols and Results

Imagen designed and executed a standalone performance assessment to evaluate OsteoDetect under a variety of conditions, characterizing the following parameters:

- Fracture detection accuracy;
- Localization accuracy; and
- Generalizability across potential confounders pertaining to image acquisition and image attributes.

a. Data Characteristics

The study data population is an independent dataset not used for model development, consisting of 1000 images (500 PA, 500 LAT) which were randomly sampled from an existing validation database of consecutively collected images from patients receiving wrist radiographs at the (b) (4) from November 1, 2016 to April 30, 2017. The study population included 35.0% males, 65.0% females, and 6.1% individuals over age 80. 33.4% of cases were post-surgical radiographs.

Study data were acquired with the following compatible X-ray imaging devices: Carestream Health DRX-1 (48.2%), GE Discovery XR656 (24.8%), and Philips Medical Systems DigitalDiagnost (27.0%).

Inclusion Criteria:

- De-identified radiograph
- PA or LAT wrist view
- Adult patient, minimum of 22 years of age
- Image acquired by one of these devices: Philips Medical System DigitalDiagnost, Carestream Health DRX-1, GE Discovery XR656

Exclusion Criteria:

- Images used during model development
- Same image exclusion criteria of the final finished software in clinical use

Additional image selection criteria:

- At least 30 images from each supported device to be included in the testing image set

b. Test Protocol

A pre-production release candidate of OsteoDetect with finalized models and Core Image Processing code was set up and run in a production-like environment. Standalone performance testing was assessed on an independent dataset not used for OsteoDetect model development. Individuals involved with model development were not involved in the study execution or analysis.

Ground truth for each case was determined by three US board certified orthopedic hand surgeons who independently interpreted images using the standard clinical definition of a distal radius fracture. Ground truth for the presence/absence of distal radius fracture is defined as the majority opinion of at least 2 of the 3 clinicians participating in the truthing process.

To determine localization accuracy, each clinician both determines whether a distal radius fracture is present and, if so, draws a bounding box of minimal size to enclose the entire fracture. The fracture localization ground truth is then defined as the union of the bounding box of each clinician identifying the fracture ("ground truth bounding box").

Training of the reviewing clinicians was performed by a member of the Clinical Team before the truthing process, and included reviewing the Ground Truth Labeling process which instructs the user to "draw the smallest possible box that encloses the entire pathology."

Scoring and Statistical Analysis:
Each image analyzed by OsteoDetect is grouped into one of four categories:

- False Positive: The model produces an output bounding box, and either it does not overlap with the ground truth bounding box (Type A) or the ground truth bounding box is empty (Type B). Note that Type A is considered by the sponsor as both a location-specific False Positive (bounding box does not contain a fracture) and as a non-location specific True Positive (OsteoDetect accurately reports presence of a fracture).
- False Negative: The model does not produce a bounding box, and the ground truth bounding box is not empty.
- True Positive: The model produces a bounding box with overlaps with the ground truth bounding box. Overlapping is defined by at least one pixel of overlap.
- True Negative: The model does not produce a bounding box, and the ground truth bounding box is empty.

c. Endpoints

- **Detection Accuracy Evaluation**: The primary endpoint of the standalone study is to characterize the detection accuracy of OsteoDetect for detecting distal radius fractures in wrist radiographs of adult patients. Detection accuracy was evaluated through the Area Under the Curve (AUC) of the Receiver Operating Characteristic curve (ROC) and agreement (sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV), as appropriate). The AUC of the ROC was estimated using scores from the Wilcoxon rank-sum test (PROC NPAR1WAY as implemented in SAS (SASR version 9.4)). The relationship between the AUC and the Wilcoxon rank-sum test statistics is:

  $$AUC = (W-W_0)/(N_1*N_0)+0.5$$

  where N_1 and N_0 are the frequencies of class 1 and 0 (fracture present/fracture absent, respectively) and W0 is the Expected Sum of Ranks under a null hypothesis of randomly ordered data, and W is the Wilcoxon rank-sums. Bootstrap re-sampling was employed to estimate both the 90% and 95% confidence intervals for the AUC (1,000 bootstrap samples).

  Sensitivity, specificity, positive predictive value, and negative predictive value are reported at OsteoDetect's predetermined single operating point. This operating point was set at the level estimated to yield a 95% sensitivity on a randomly withheld subset of the model's training data. Both 90% and 95% confidence intervals were calculated using Wilson's method (when appropriate) or generated using bootstrap re-sampling.

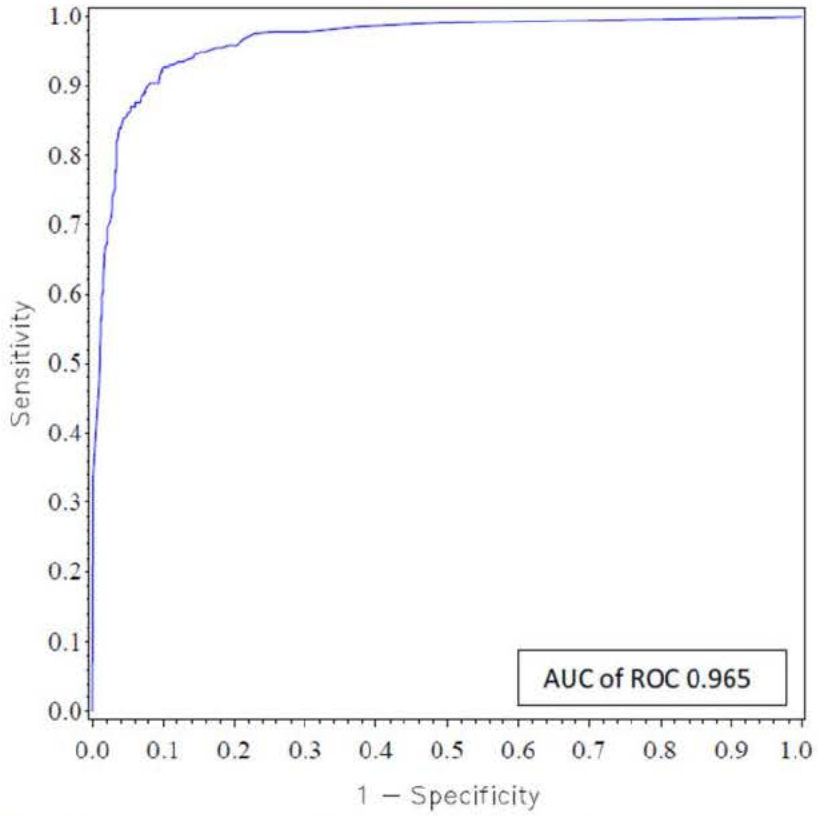
- **Localization Accuracy:** Localization accuracy testing measures the average distance between the centroid of the ground truth bounding box to the centroid of OsteoDetect's bounding box. Localization accuracy against the union of the truther bounding boxes and against each of the individual bounding boxes from the three truthers was reported.

  The distance was measured in pixels, where the distance between two centroids is the Euclidean distance measure based on row and column pixel location. This is only reported for the true positive radiographs because they are the only ones with bounding boxes for at least two out of the three truthers. For localization accuracy, the distances between centroids of the OsteoDetect model bounding box and the truther bounding box are reported.

- **Generalizability Testing:** Generalizability testing is performed to demonstrate OsteoDetect's ability to perform across different confounding variables. Confounders considered as factors in a multivariate logistic model were device, contrast, age group, sex, and presence/absence of a cast. In addition, AUC, sensitivity, specificity, positive predictive value, and negative predictive value was calculated for all radiographs in total and for both the PA and LAT views across the following subgroups: elderly status (i.e., 80 years old) and post-surgical status.

d. Results

Detection Accuracy: The AUC of the ROC is 0.965 (95% confidence interval: 0.953, 0.976). The sensitivity, specificity, PPV, and NPV of OsteoDetect are reported in the table below.



OsteoDetect Standalone Evaluation: Sensitivity, Specificity, PPV, and NPV (90% and 95% with Wilson's Confidence Intervals:

		Wilson's Confidence Intervals	
Performance Metric	**Estimate**	**90%**	**95%**
Sensitivity	0.921	(0.892, 0.942)	(0.886, 0.946)
Specificity	0.902	(0.882, 0.919)	(0.877, 0.922)
PPV	0.813	(0.777, 0.844)	(0.769, 0.850)
NPV	0.961	(0.946, 0.972)	(0.943, 0.973)

Localization Accuracy:
The average number of pixels between the centroids of the OsteoDetect model's predicted bounding box and that of the reference standard bounding box was 33.52

(standard deviation of 30.03). The average image size was 1,663 x 1,109 pixels (area of 1,844,267 pixels^2) and the average area of the bounding boxes was 30,164 pixels^2 for the Reference Standard and 34,924 for OsteoDetect. These localization accuracy results indicate that OsteoDetect will typically draw boxes close to the site of detected distal radius fractures.

Generalizability:

Potential observable confounders such as sex, age, or post-surgical imaging were not found to statistically affect performance estimates. When evaluated individually by radiograph and by each subgroup, the AUC of the ROC for each subgroup remained high. The lowest subgroup, post-surgical radiographs, has an AUC of 0.926.

Sensitivity, specificity, PPV, and NPV was high for all subgroups, including both views and all reported characteristics:

Subgroup		Performance Metric (90% Wilson's CI)			
View	**Characteristic**	**Sensitivity**	**Specificity**	**PPV**	**NPV**
PA	Post-Surgical: No	0.904 (0.832, 0.947)	0.934 (0.902, 0.955)	0.805 (0.724, 0.867)	0.970 (0.945, 0.984)
	Post-Surgical: Yes	0.920 (0.864, 0.954)	0.779 (0.698, 0.844)	0.829 (0.762, 0.880)	0.893 (0.820, 0.939)
	Age < 80	0.908 (0.863, 0.940)	0.897 (0.866, 0.922)	0.813 (0.759, 0.857)	0.952 (0.928, 0.969)
	Age ;: 80	0.950 (0.804, 0.989)	0.800 (0.591, 0.917)	0.864 (0.703, 0.944)	0.923 (0.718, 0.983)
LAT	Post-Surgical: No	0.940 (0.874, 0.973)	0.951 (0.925, 0.968)	0.818 (0.736, 0.879)	0.985 (0.968, 0.993)
	Post-Surgical: Yes	0.921 (0.855, 0.959)	0.750 (0.658, 0.824)	0.795 (0.716, 0.857)	0.900 (0.818, 0.947)
	Age < 80	0.925 (0.878, 0.955)	0.912 (0.883, 0.934)	0.804 (0.746, 0.851)	0.969 (0.949, 0.981)
	Age ;: 80	1.000 (0.787, 1.000)	0.875 (0.684, 0.958)	0.833 (0.601, 0.943)	1.000 (0.838, 1.000)

e. Conclusion

The observed results of standalone testing demonstrate that OsteoDetect by itself, in the absence of any interaction with a clinician, can detect and localize distal radius fractures in wrist radiographs with high sensitivity, specificity, AUC, PPV, NPV, and localization accuracy.

This performance was observed for both PA and LAT views across elderly status and post-surgical status, and it was observed on radiographs representative of OsteoDetect's intended use. Other potential confounders did not significantly affect performance.

8. Animal and/or Cadaver Testing

None provided.

J. Summary of Clinical Information

1. Reader Study Testing Protocols and Results

Imagen performed a fully crossed, multiple reader multiple case (MRMC) retrospective study to assess performance of OsteoDetect when used with concurrent reading (consistent with the device's intended use) to assist clinicians in detecting distal radius fractures during interpretation of PA and LAT radiographs of adult wrists. OsteoDetect's clinical performance was compared to conventional, unaided clinical interpretations.

a. Data Characteristics

Readers evaluated 200 cases under aided and unaided conditions. The cases were randomly sampled from the same validation database used for the standalone performance study. The study cases included 65% female participants and 35% male participants. 44.5% cases contained fractures, and 36.5% cases were post-surgical.

The image selection process was designed to ensure that at least 25% of cases have distal radius fracture, and that at least 5% cases correspond to patients 81 years of age and older. Otherwise, the same inclusion/exclusion criteria were used as in the standalone performance study.

Prospective readers must meet one of the following criteria in order to be eligible for the study:

- Medical resident or fellow participating in an Accreditation Council for Graduate Medical Education (ACGME)-accredited program, or
- U.S. board-eligible/board-certified and licensed Medical Doctor, Doctor of Osteopathy, Nurse Practitioner, or Physician's Assistant.

b. Test Protocol

Study Readers:

24 clinical readers were enrolled in the clinical performance evaluation reader study. Selected readers represented the intended use population and included radiologists (n=4), orthopedic surgeons (n=4), emergency medicine physicians (n=4), emergency medicine physician assistants (n=4), internal medicine physicians (n=4), and family practice physicians (n=4), with a mean 15.3 years of experience (range: 3–35 years). Each of the 24 readers completed all 400 reads, 200 OsteoDetect-Aided and 200 OsteoDetect-Unaided.

Each image was assigned a ground truth label (presence or absence of distal radius fracture), determined by majority opinion of three U.S. board-certified orthopedic hand surgeons. Each case is comprised of the PA, LAT, and (when available) oblique images for a patient visit. Each case was assigned ground truth based on presence or absence of distal radius fracture in any of the case's images. (Note that this ground truth definition is made on a complete case basis, whereas in the standalone testing ground truth was defined on an individual image basis, as OsteoDetect performs its assessment on an individual image basis while the end user diagnoses on a per-case basis).

A sample size calculation determined that 24 readers and 200 cases were needed to achieve greater than 80% power to detect a difference between aided and unaided reads.

Study duration: 4 months

All readers independently read all cases, both OsteoDetect-aided and OsteoDetect-unaided. Reading sessions were separated by a minimum one-month washout period.

For each case, each reader is asked to 1) determine presence or absence of distal radius fracture, making a binary yes/no response, and 2) report confidence in their binary decision on a 0 – 100% scale.

c. Endpoints

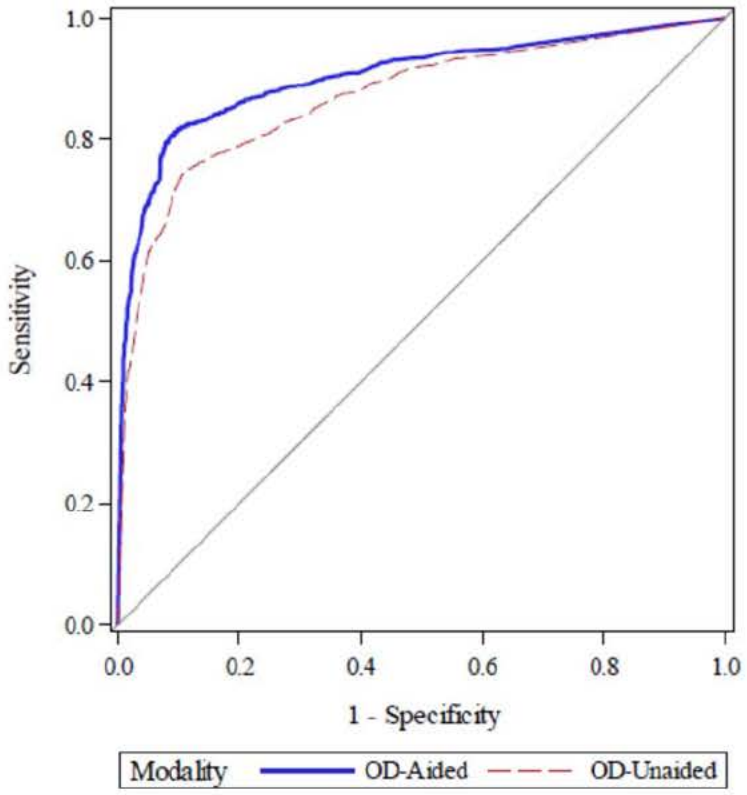
Primary objective: The primary objective of the clinical reader study is to determine whether the diagnostic accuracy of readers aided by OsteoDetect is superior to reader accuracy when unaided by OsteoDetect, as determined by the AUC of the ROC curve:
H_0: $AUC_{aided} - AUC_{unaided} \leq 0$
H_1: $AUC_{aided} - AUC_{unaided} > 0$

Secondary objective: The secondary objective is to report the population-average sensitivity and specificity of the OsteoDetect-aided and unaided reads.

d. Test Results

Primary objective:

The diagnostic accuracy of OsteoDetect-aided reads was found to be superior to OsteoDetect-unaided reads for the intended use population (p=0.0056). The least squares mean difference between the AUC for OsteoDetect-aided and OsteoDetect-unaided reads is 0.049 (95% CI, (0.019, 0.080)).



Model	Modality	AUC	Std. Error	OD Aided – OD Unaided	Std. Error	p-value	95% Lower CL	95% Upper CL
ROC AUC Full Model	OD - Aided	0.889	0.029	0.049	0.013	0.0056	0.019	0.080
	OD - Unaided	0.840	0.029	-	-	-	-	-

H1 was not rejected.

Secondary objective:

The sensitivity, specificity, PPV, and NPV for the OsteoDetect-aided and OsteoDetect-unaided Reads are reported below with 95% Wilson's Confidence Interval:

	Estimate (95% Wilson's CI)			
Modality	**Sensitivity**	**Specificity**	**PPV**	**NPV**
OD-Aided	0.803 (0.785, 0.819)	0.914 (0.903, 0.924)	0.883 (0.868, 0.896)	0.853 (0.839, 0.865)
OD-Unaided	0.747 (0.728, 0.765)	0.889 (0.876, 0.900)	0.844 (0.826, 0.859)	0.814 (0.800, 0.828)

The number of true positive, true negative, false positive, and false negative cases for OsteoDetect-aided and OsteoDetect-unaided reads were reported as follows:

Modality	TP	TN	FP	FN
OD-Aided	1715	2436	228	421
OD-Unaided	1596	2368	296	540
Total	3311	4804	524	961

Assessment of Aided and Unaided reader performance based on important cohorts

Imagen reported the AUC of the ROC, OsteoDetect-aided and OsteoDetect-unaided, broken down for each reader type, which included Emergency Medicine Physicians, Emergency Medicine Physician Assistants, Family Practice Physicians, Internal Medicine Physicians, Orthopedic Surgeons, and Radiologists. Improvement in the performance of OsteoDetect-aided reads was seen across all reader types as compared with OsteoDetect-unaided reads.

The improvement in sensitivity, specificity, positive predictive value, and negative predictive value was observed consistently whether or not oblique (OBL) radiographic views are available in a case:

		Estimate (95% Wilson's CI)			
Modality	**Number (Type) of Views**	**Sensitivity**	**Specificity**	**PPV**	**NPV**
OD - Aided	2 (PA + LAT)	0.810 (0.782, 0.835)	0.915 (0.897, 0.930)	0.871 (0.845, 0.893)	0.872 (0.852, 0.889)
OD - Unaided	2 (PA + LAT)	0.732 (0.700, 0.761)	0.898 (0.879, 0.914)	0.835 (0.806, 0.860)	0.825 (0.803, 0.845)
OD - Aided	3 (PA + LAT + OBL)	0.798 (0.776, 0.819)	0.914 (0.899, 0.927)	0.890 (0.871, 0.907)	0.839 (0.820, 0.856)
OD - Unaided	3 (PA + LAT + OBL)	0.757 (0.733, 0.779)	0.882 (0.865, 0.898)	0.849 (0.827, 0.868)	0.806 (0.786, 0.824)

Imagen also assessed and reported the OsteoDetect-aided and OsteoDetect-unaided AUC (and sensitivity, specificity, PPV, NPV) results broken down by relevant confounders (patient demographics and presence or absence of a cast, post-surgical images, imaging device used to acquire radiographs). The study demonstrated consistent improvement of performance across metrics and confounders in OsteoDetect-aided reads as compared to OsteoDetect-unaided reads.

e. Conclusion

The primary endpoint was met, demonstrating improvement in performance of readers when aided by OsteoDetect as measured by the AUC of the ROC. The full DBM model described in the study protocol shows a statistically significant higher AUC in OsteoDetect-aided reads than in the OsteoDetect-unaided reads (p=0.0056).

The secondary endpoint demonstrated improvement of sensitivity, specificity, positive predictive value, and negative predictive value of readers aided by OsteoDetect as compared to when unaided by OsteoDetect. For each above specified performance metric, there was no overlap in the 95% confidence intervals of the OsteoDetect-aided and the OsteoDetect-unaided reads.

K. Proposed Labeling:

The labeling is sufficient and it satisfies the requirements of 21 CFR Part 801, including 21 CFR Part 801.109 for prescription devices, and the special controls for this device type. The OsteoDetect User Manual provides the detailed instructions for use. Other elements of the required labeling for OsteoDetect for this device type are noted below.

1. Indicated patient population

   The target population for the device is adults (age 22 or older) who are suspected of having a distal radius fracture.

2. Indicated user population

   The intended users of OsteoDetect are clinicians in various settings including primary care (e.g., family practice, internal medicine), emergency medicine, urgent care, and specialty care (e.g., orthopedics), as well as radiologists who review radiographs across settings.
   When a clinician accesses the patient radiographs in a PACS client, both the OsteoDetect-annotated radiographs and the original, unaltered radiographs are available in the same patient study, allowing for concurrent reading.

3. Intended Reader Protocol and Device Inputs and Outputs

   The end user interacts with the output of OsteoDetect in the facility's PACS software.

Annotated DICOM images processed by OsteoDetect are to be read concurrently by the clinician with the original, unannotated radiographs.

If a distal radius fracture is detected by OsteoDetect, the software annotates the image by creating a bounding box surrounding the fracture with the label "Distal Radius Fracture: DETECTED." If none is detected, the message label "Distal Radius Fracture: NOT DETECTED" appears. Whether or not the software detects a distal radius fracture, when OsteoDetect analyzes an image it annotates the radiograph to include a message stating that the radiograph was analyzed by OsteoDetect, and includes instructions for users to access device labeling.

The annotated radiographs are available in the same patient study as the unaltered radiographs, allowing for a concurrent read. Annotated images should always be read concurrently with the original, unannotated radiographs according to standard of practice.

4. Device Limitations

Warnings:

- Clinicians should review OsteoDetect annotated images concurrently with original images before making a final determination on a case. OsteoDetect is an adjunct tool and does not replace the role of the clinician. Clinicians must not use the CAD generated output as the primary interpretation.
- OsteoDetect is not designed to detect fractures other than distal radius fractures. Clinicians should review original images for all suspected pathologies.
- OsteoDetect may not detect a distal radius fracture in both the PA and LAT views. Clinicians should follow standard clinical procedures in assessing PA and LAT views.
- All images must have proper DICOM tags. OsteoDetect performance may be reduced on images that do not have correct DICOM tags.

Adverse Effects:

OsteoDetect is designed to assist clinicians in their review of posterior-anterior (PA) and lateral (LAT) wrist radiographs. It does not modify the existing image acquisition process nor alter the native images. Additionally, the use of the software does not directly involve the patient. Therefore, there is no known direct safety or health risk caused by, or related to, the use of the device. The indirect risks associated with the use of OsteoDetect are:

1. False positive. When reviewing radiographs with the aid of OsteoDetect, there is a chance that a false positive may occur, potentially leading to additional diagnostic work-up.
2. False negative. When reviewing radiographs with the aid of OsteoDetect, there is a chance that users may rely too heavily on the absence of OsteoDetect findings without sufficiently assessing the native image. This may result in missing fractures that may have otherwise been found.

3. Device misuse. When reviewing radiographs with the aid of OsteoDetect, there is a chance that the device could be misused to analyze images from an unintended patient population or on images acquired with incompatible imaging hardware or incompatible image acquisition parameters, potentially leading to inappropriate information regarding the presence/location of a fracture being provided to the end user.
4. Device failure. When reviewing radiographs with the aid of OsteoDetect, the device could fail and lead to absence of results, delay of results, or incorrect results. This may result in delayed or inaccurate patient diagnosis

These potential risks are mitigated by the fact that OsteoDetect is only intended to be used as an adjunct tool and is not intended as a replacement for the clinician's review of the radiograph or his or her clinical judgment, in addition to the mandatory special controls assigned to this device type, along with general controls. The original, unaltered radiographs are available for concurrent reading with the OsteoDetect-annotated radiographs.

5. Compatible Hardware

OsteoDetect performs computer-aided detection on radiographs produced by the following systems:

- Philips Medical Systems DigitalDiagnost
- Carestream Health DRX-1
- GE Discovery XR656

OsteoDetect does not support any other imaging system.

OsteoDetect requires a dedicated physical machine with the following minimum specifications:

- Intel Xeon E3-1225 v5 or comparable (with at least 2 CPU cores)
- 16GB RAM or more
- 100 GB or more free space
- 100 Mbps or faster Ethernet interface to the institution’s DICOM network

It is strongly recommended that the machine is certified compatible with ESXi.

OsteoDetect should be installed on top of the vSphere Hypervisor, ESXi 6.0 Update 2.

6. Device Instructions

Device instructions are provided in the user manual for OsteoDetect.

The user manual specifies radiology technologist training requirements to ensure that the appropriate metadata tags are applied to all cases so that OsteoDetect can correctly identify images to be analyzed by the algorithm. Technologists are required to:

- Identify the intended use of OsteoDetect;
- Identify OsteoDetect’s warnings.

The user manual specifies that clinicians using OsteoDetect are required to:

- Identify the intended use of OsteoDetect;
- Identify OsteoDetect's warnings;
- Identify the key performance characteristics of OsteoDetect:
- Be aware of the OsteoDetect outputs on an image.

7. Performance Testing Summary

A summary of the clinical performance assessment, including its primary and secondary objectives, study design, and results are included in the user manual. The results reported in this section include the ROC curves for OsteoDetect-aided and -unaided reads, and sensitivity, specificity, positive predictive value and negative predictive value results for aided and unaided reads. The breakdown of performance results included performance when 2 radiograph views (PA + LAT) or 3 radiograph views (PAT + LAT + OBL) were available for the reader to review.

A summary of the standalone performance assessment is included in the End User Guide, Section 6 (Summary of Standalone Performance Assessment). This report includes a discussion of the objectives, study design, and an overview of the study results. The results reported in this section include the ROC curve, sensitivity, specificity, PPV and NPV, localization accuracy, and generalizability results across several potentially confounding variables (age > 80 years, PA vs. LAT view, post-surgical imaging).

L. Identified Risks to Health and Identified Mitigations

Identified Risks	Identified Mitigation Measures
False positive results	General controls and special controls (1) and (2).
False negative results	General controls and special controls (1) and (2).
Device misuse (analyzing images from unintended patient population or of an unintended anatomical site; or images acquired with an unintended modality, incompatible imaging hardware, or incompatible image acquisition parameters) resulting in lower device performance (inappropriate detection/diagnosis information being displayed to the end user)	General controls and special controls (1) and (2).
Device failure could lead to absence of results, delay of results, or incorrect results, which can lead to delayed or inaccurate patient diagnosis	General controls and special controls (1) and (2).

M. Benefit/Risk Determination

Summary	
Summary of the Benefit(s)	The clinical MRMC study demonstrated a statistically significant improvement in reader performance in detecting distal radius fracture in adult patients (as measured by the primary endpoint of the ROC Area Under the Curve) when aided with OsteoDetect as compared to performance at the same task without OsteoDetect, according to clinical standard of care. AUCaided - AUCunaided = 0.889 - 0.840= 0.049 (two sided 95% confidence level [0.019,0.080]) OsteoDetect-aided read performance also showed statistically significant improvement as measured by sensitivity, specificity, PPV, and NPV as compared with the unaided read performance. Specifically, Imagen's study demonstrated a device-aided sensitivity of 80.3% (two sided 95% confidence interval for the mean: [78.5%,81.9%]) and device-aided specificity of 91.4% (two sided 95% confidence interval for the mean: [90.3%,92.4%]). By comparison, the study demonstrated non-aided sensitivity and specificity of 74.7% [72.8%,76.5%] and 88.9% [87.6%,90.0%], respectively. Earlier detection of a distal radial fracture will allow earlier intervention, potentially allowing closed reduction and casting instead of open reduction, minimizing the risk of delayed pain and post-traumatic arthritis.
Summary of the Risk(s)	There are minimal potential risks associated with use of the device, including: • The device could provide false positive results, which could contribute to the end user using this information to make a false positive diagnosis. Such a false positive diagnosis can result in unnecessary patient treatment or followup. • The device could provide false negative results, which could contribute to the end user using this information to make a false negative diagnosis. A false negative diagnosis could lead to delays in diagnosis and treatment of the fracture and increase the likelihood of negative outcomes such as incomplete fracture healing. • The device could be misused to analyze images from an unintended patient population or on images acquired with

	incompatible imaging hardware or incompatible image acquisition parameters, leading to inappropriate information regarding the presence/location of a fracture being provided to the end user. • The device could fail and lead to absence of results, delay of results, or incorrect results, which can lead to delayed or inaccurate patient diagnosis However, based on the performance data and the application of general controls and special controls established for this device type, use of this device is unlikely to increase the rate of false negative or false positive diagnoses of distal radius fracture as compared with the current clinical standard of practice. Further, possible misuse of the device does not present additional risks compared with the misuse of other types of radiological image processing devices.
Summary of Other Factors	Currently, injuries to the hand and wrist account for approximately 20% of visits to the emergency department, with a total of approximately 3.5 million hand and wrist injuries noted in 2009, for an incidence of 1130 injuries per 100,000 persons per year. (Hand (NY) 2012; 7:18-22) By detecting a potentially occult distal radial fracture earlier, pain, post-traumatic arthritis, and possible disability are alleviated, resulting in a potential significant improvement in US public health. This software tool brings the expertise of musculoskeletal radiologists and orthopaedic surgeons specializing in hand surgery to emergency medicine physicians and physician assistants, family practice physicians, and internal medicine physicians. Therefore, subspecialty expertise is potentially available to emergency rooms across the country with the use of this software device.
Conclusions Do the probable benefits outweigh the probable risks?	The benefited population would include adult patients with clinically suspected distal radius fracture. This software potentially brings subspecialty expertise of musculoskeletal radiologists and orthopaedic surgeons specializing in hand and wrist surgery to physicians and physician assistants working in emergency rooms across the country. The underlying statistics as discussed above justify that the performance of the device has been clinically validated in a fully crossed design. The 4 types of clinical providers (emergency medicine physicians, emergency medicine physician assistants, family medicine physicians, and internal medicine physicians) most likely to staff emergency rooms were separately evaluated, and all showed improvement in the detection of a distal radial fracture aided by the device. Given that the standard of care in 2018 is to cast and have a follow-up X-ray in 10-14 days or perform an MRI in cases where a distal radial fracture is clinically suspected in

	the face of a normal hand/wrist X-ray series, a software device that detects distal radial fracture earlier is a potential significant benefit to public health. Accordingly, the probable benefits of the device outweigh the probable risks, mainly a false positive diagnosis, given the combination of general controls and special controls established for this device type.. The hand and wrist are the most commonly injured part of the body and 1/6 of fractures presenting to emergency rooms are distal radial fractures. This software device has the potential to result in significant improvement in public health while introducing minimal additional clinical risks of extra follow-up, treatment, and imaging.

Patient Perspectives

This submission did not include specific information on patient perspectives for this device.

N. Conclusion

The information provided in this de novo submission is sufficient to classify this device into class II under regulation 21 CFR 892.2090. FDA believes that the stated special controls, in combination with the general controls, provide a reasonable assurance of the safety and effectiveness of the device type. The device is classified under the following:

Product Code: QBS

Device Type: Radiological Computer Assisted Detection and Diagnosis Software

Class: II

Regulation: 21 CFR 892.2090

(a) Identification. A radiological computer assisted detection and diagnostic software is an image processing device intended to aid in the detection, localization, and characterization of fracture, lesions, or other disease specific findings on acquired medical images (e.g. radiography, MR, CT). The device detects, identifies and characterizes findings based on features or information extracted from images, and provides information about the presence, location, and characteristics of the findings to the user. The analysis is intended to inform the primary diagnostic and patient management decisions that are made by the clinical user. The device is not intended as a replacement for a complete clinician's review or their clinical judgment that takes into account other relevant information from the image or patient history.

Classification. Class II (special controls). A radiological computer assisted detection and diagnosis software must comply with the following special controls:

1. Design verification and validation must include:
    i. A detailed description of the image analysis algorithm, including but not limited to a description of the algorithm inputs and outputs, each major component or block, how the algorithm and output affects or relates to clinical practice or patient care, and any algorithm limitations.
    ii. A detailed description of pre-specified performance testing protocols and dataset(s) used to assess whether the device will provide improved assisted-read detection and diagnostic performance as intended in the indicated user population(s), and to characterize the standalone device performance for labeling. Performance testing includes standalone test(s), side-by-side comparison(s), and/or a reader study, as applicable.
    iii. Results from standalone performance testing used to characterize the independent performance of the device separate from aided user performance. The performance assessment must be based on appropriate diagnostic accuracy measures (e.g., receiver operator characteristic plot, sensitivity, specificity, positive and negative predictive values, and diagnostic likelihood ratio). Devices with localization output must include localization accuracy testing as a component of standalone testing. The test dataset must be representative of the typical patient population with enrichment made only to ensure that the test dataset contain a sufficient number of cases from important cohorts (e.g., subsets defined by clinically relevant confounders, effect modifiers, concomitant disease, and subsets defined by image acquisition characteristics) such that the performance estimates and confidence intervals of the device for these individual subsets can be characterized for the intended use population and imaging equipment.
    iv. Results from performance testing that demonstrate that the device provides improved assisted-read detection and/or diagnostic performance as intended in the indicated user population(s) when used in accordance with the instructions for use. The reader population must be comprised of the intended user population in terms of but not limited to clinical training, certification, and years of experience. The performance assessment must be based on appropriate diagnostic accuracy measures (e.g., receiver operator characteristic plot, sensitivity, specificity, positive and negative predictive values, and diagnostic likelihood ratio). Test datasets must meet the requirements described in 1(iii) above.
    v. Appropriate software documentation, including device hazard analysis, software requirements specification document, software design specification document, traceability analysis, system level test protocol, pass/fail criteria, testing results, and cybersecurity measures.

2. Labeling must include the following:

i. A detailed description of the patient population for which the device is indicated for use.
ii. A detailed description of the device instructions for use, including the intended reading protocol and how the user should interpret the device output.
iii. A detailed description of the intended user, and any user training materials as programs that addresses appropriate reading protocols for the device to ensure that the end user is fully aware of how to interpret and apply the device output.
iv. A detailed description of the device inputs and outputs.
v. A detailed description of compatible imaging hardware and imaging protocols.
vi. Warnings, precautions, and limitations must include situations in which the device may fail or may not operate at its expected performance level (e.g., poor image quality or for certain subpopulations), as applicable.
vii. A detailed summary of the performance testing, including: test methods, dataset characteristics, results, and a summary of sub-analyses on case distributions stratified by relevant confounders, such as anatomical characteristics, patient demographics and medical history, user experience, and imaging equipment.