# AI-based Auto Positioning

## Contents

# Introduction

AI technology is making progress rapidly and being used in various industries. Medical devices are not an exception and vendors apply these technologies for development. Regarding computed tomography, which is one of the major medical imaging devices, the scan time was improved dramatically with wide coverage detector technology. On the other hand, there were no major innovations for patient positioning in the past and the operator spent a lot of time, impacting the entire workflow. Another challenge is a variance by operators. The accuracy of patient positioning impacts image quality and radiation dose and it requires operator's skills. To resolve those clinical challenges, GE Healthcare developed and commercialized the Auto Positioning function, using AI technology and a 3D camera. The purpose of this report is to describe the technology of Auto Positioning.

# Clinical Benefits

The combination of AI technology and a high accuracy 3D camera provides the following clinical benefits:

- Workflow improvement enabled by Auto Positioning with AI-based technology
- Consistent image quality by auto centering with 3D camera
- Medical error avoidance by patient orientation identification
- Patient safety by patient-gantry collision check

### Workflow Improvement Enabled by Auto Positioning with AI-based Technology

The Auto Positioning function automatically detects an anatomical landmark by deep learning algorithms and allows minimizing positioning action into a single click operation, which is an 80% reduction of click numbers compared to the manual operation. With this technology, it is a high possibility for it to reduce the time for patient positioning significantly compared to traditional manual operations.

### Consistent Image Quality by Auto Centering with a 3D Camera

The 3D camera detects depth information of patients and is calculated the required table elevation to set the centering in the selected protocol. The auto centering function optimizes the radiation dose and image quality without regard to the operator's skill variance. Also, auto centering supports reducing re-scout scanning related with mis-centering since the required centering information input is before scout scanning. Generally, two direction's scouts are used for scan range planning. The AP (anterior to posterior) direction is for scan range of reconstruction and the lateral is for accurate reconstruction in center of display FOV. With auto centering, the lateral direction of scout is not required, and it reduces the time.

### Medical Error Avoidance by Patient Orientation Identification

The patient orientation is determined by the relative positions of these eight identified anatomical landmarks. If a mismatch is detected, a warning is displayed on the touchscreens. This is an important safety enhancement because a patient orientation mismatch will lead to incorrect image annotation. In the existing manual workflow, when Auto Positioning is not used, the operator is relied upon to visually check the patient orientation without prompting or guidance.

### Patient Safety by Patient-gantry Collision Check

Auto Positioning deterministically produces a 3D surface contour map of the entire patient body from the depth information from the camera. This data in combination with the known system geometries, determines if a part of the patient would come in contact with the gantry if the table/cradle are moved to the scout start position. The collision evaluation is performed around the entire 360 degree of the gantry bore and includes a 2 cm consideration for tolerance stack up.

# Features and Technical Overview

Auto Positioning uses a fixed, ceiling mounted, off the shelf, 2D/3D video camera that is capable of determining distances to points in its field of view. The 3D camera specification is shown in Table 1. It displays standard RGB video images on the CT system's existing gantry-mounted touchscreens (Figure 1). Information from the standard output of the camera is used, along with precise spatial information of the individual CT system's gantry and table installation geometry, to determine the anatomical landmark location and the start and end locations for the scout scan(s).

| Function | Specifications |
|---|---|
| Depth image and RGB image resolution | 1280 x 720 |
| Frame rate | 30 fps |
| Horizontal field of view for Depth images | 65 ±2 degree |
| Horizontal field of view for RGB images | 69 ±1 degree |
| Depth data accuracy | ±20 mm at working distance range of 1000 – 1800 mm |

*Table 1: 3D camera specification for Auto Positioning and centering.*

GE's protocol structure contains a field for the anatomical reference. GE reference protocols contain one of eight anatomical references and these are the only ones that Auto Positioning supports. Almost all users, when creating their own protocols use one of these eight. Auto Positioning is allowed to function if one of the eight supported anatomical references is used in the protocol. These eight anatomical references each directly correlate with the landmark that will be derived. The eight supported anatomical references are: Orbital Meatal baseline (OM), Sternoclavicular Notch (SN), Xyphoid (XY), Iliac Crest (IC), Left and Right Knee (KN), Left and Right Ankle Joint (AJ), as shown in Figure 2.

GE reference and user created protocols also contain the scout range. When the user selects their desired protocol, Auto Positioning uses the anatomical reference and the scout range information to determine the landmark and the scout's start and stop locations.

The start and stop locations are in relation to the landmark location determined by Auto Positioning. The user is able to review and adjust these locations via the touchscreen. Following the confirmation of the scout locations, under normal conditions, the system's existing "Move to Scan" functionality, that both elevates the table and moves the cradle into the gantry so that the patient is located at the proper scout scan start location, is initiated. The table height is determined, using the data from the camera, such that the patient is centered vertically in the gantry bore. Proper patient centering is important for the automatic exposure control to function optimally, which in turn results in image quality and dose optimization.

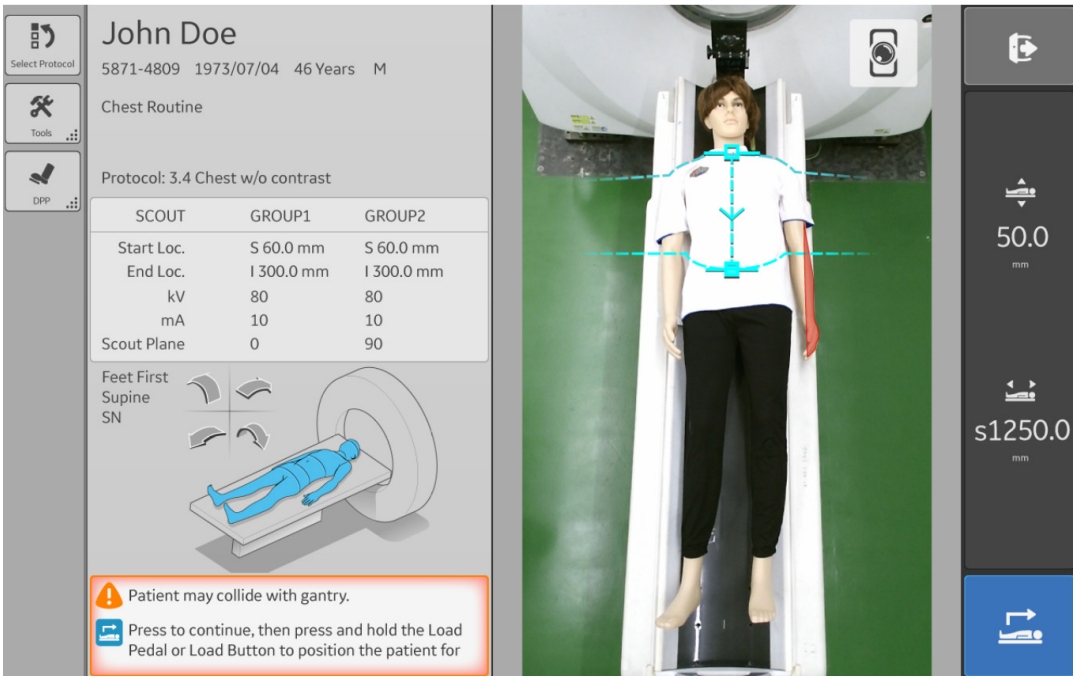



*Figure1: Gantry mounted touchscreen – Xtream Tablet.*

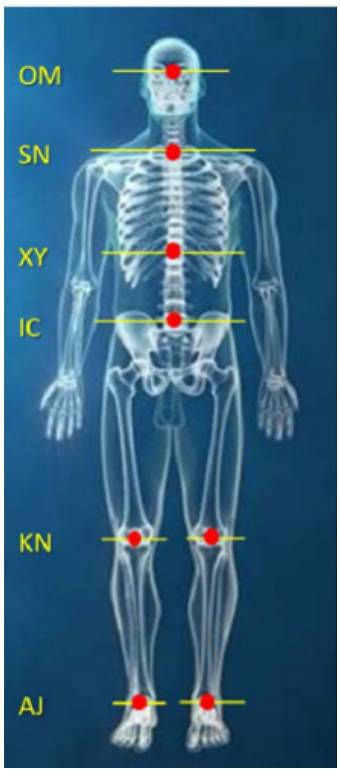



*Figure 2: The eight supported anatomical references/landmarks.*

# Auto Positioning Algorithms

Auto Positioning uses two deep learning algorithms with different inputs that produce comparable outputs to identify all eight of the anatomical landmarks, shown in Figure 2, on the patient's body. All eight of these identified landmarks are used to determine the patient orientation (head or feet first).

The anatomical reference in the selected protocol will determine which one of the eight identified landmarks will be used in the determination of the scout's start and stop locations. The start and stop locations are determined by applying the scout's scan range in the selected protocol to the associated identified landmark.

The RGBLandmarkNet network uses 2D video images as inputs and outputs, all eight of the predefined landmark locations in X and Z. In parallel, the DepthLandmarkNet network uses the 3D depth data from the camera to also produce all eight of the predefined landmark locations.

There are not any special clothing requirements for the patient when Auto Positioning is used. Auto Positioning also operates per design if the patient is covered with a sheet.

The scout's start and stop locations that are displayed as an overlay on the touchscreen image are determined from the 2D video.

The visual display of the scout's start and stop locations must be confirmed by the operator before proceeding. The location accuracy of both the scout start and stop locations is ±10 mm.

The 3D depth images are used to generate a "point cloud" on a mesh of points on the patient surface contour as determined from the depth information. The point cloud is then segmented to produce the body contour. The body contour is used to deterministically calculate the vertical geometric center of the patient. The center point location is then used to calculate the required table elevation for patient centering. The body contour is also used to calculate if a part of the patient would come in contact with the gantry and if the table/cradle are moved into the required scanning position determined by Auto Positioning. Additionally, the body contour is used to highlight the area(s) of the patient where the contact would occur. The overall processing workflow is described in Figure 3 below.

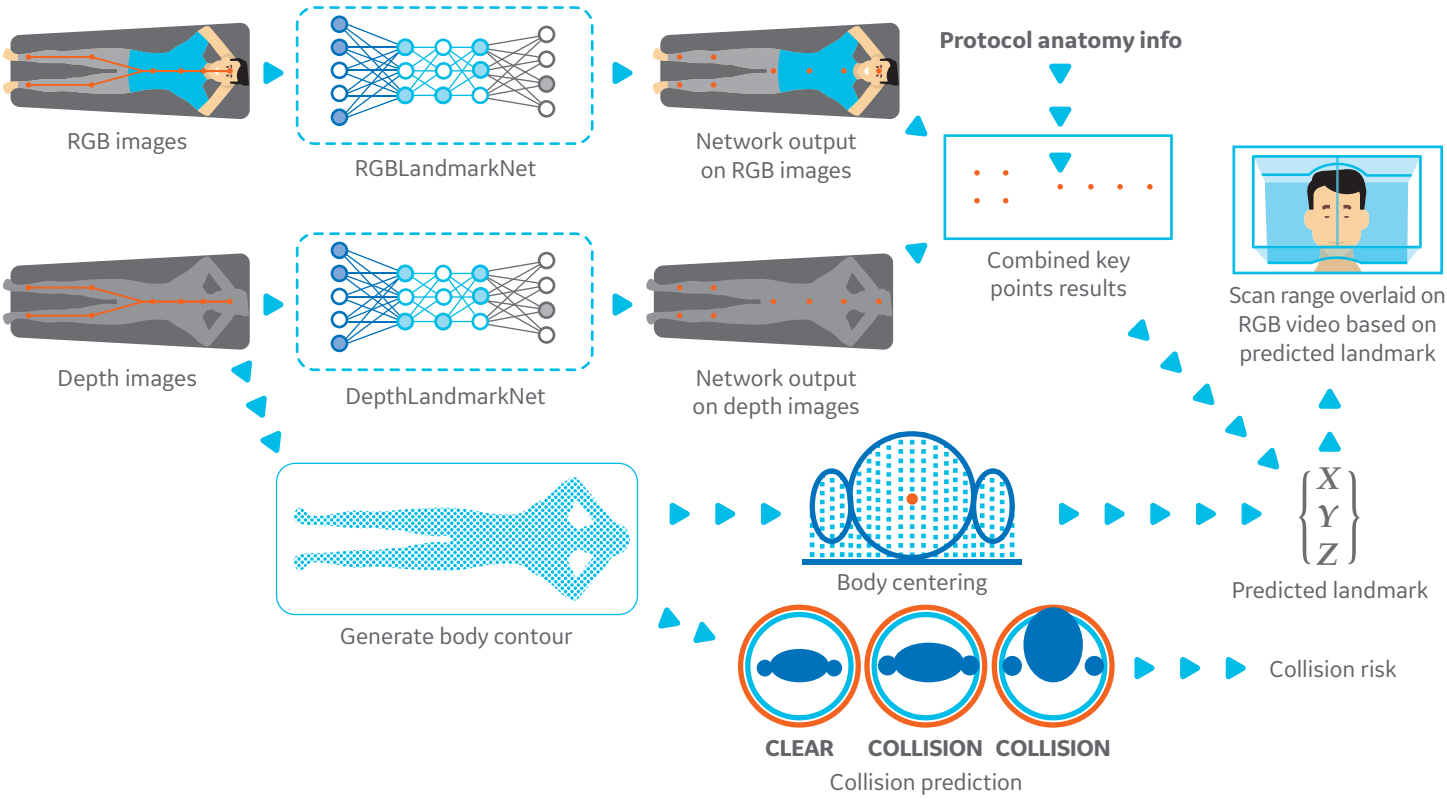



*Figure 3: Deep Learning Algorithm Processing Workflow.*

# Algorithm Function

The goal of each deep learning network DepthLandmarkNet and RGBLandmarkNet is to estimate eight anatomical landmarks on the patient. Figure 4 to the right, describes how the networks take the depth image and the RGB image as inputs and output the predicted landmark locations.

### Depth Pre-process and RGB Pre-process:

The depth and RGB image's field of view extends beyond the patient. The "extra" area (data) in the images can be thought of as "environmental noise." It is desirable to eliminate this irrelevant data prior to being fed into the networks for inference. This is accomplished by using the edge of the table in the images to determine "bounding boxes" that are used to eliminate the irrelevant data.

### DepthLandmarkNet and RGBLandmarkNet Inference:

The network structure for DepthLandmarkNet and RGBLandmarkNet are the same, although they were trained with different data – one with depth images and the other with RGB images. The network is a ResNet-like CNN, which takes a pre-processed image and returns eight channel heatmaps, as illustrated in Figure 5 showing the example of RGBLandmarkNet inference. These heatmaps are used for eight anatomical landmark location regressions in next step.

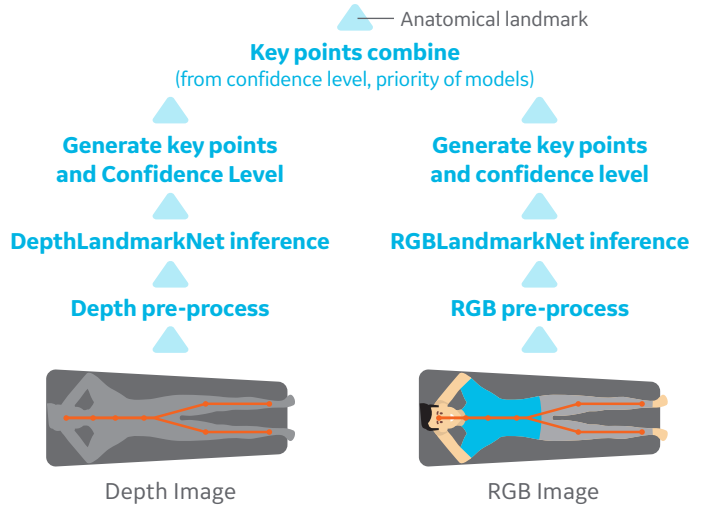



*Figure 4: Overview of algorithms.*

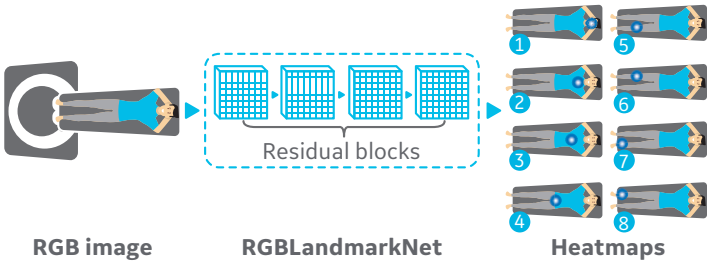



*Figure 5: Example of RGBLandmarkNet inference.*

# Dataset for Training, Validation and Test

The training, validation, and test datasets for DepthLandmarkNet and RGBLandmarkNet were able to be acquired from volunteers.

The data was separated into three groups: training data, validation data, and testing data, where the three groups of data contain depth and RGB images. The three datasets were kept separate from each other. The training and validation datasets contained a total of approximately 22,000 images of which about 20% were segregated as the validation dataset. The test dataset contained approximately 3,500 images.

The volunteers were chosen to reflect a wide range of body sizes, ages, and heights. The volunteers were placed in different possible orientations indifferent environmental conditions such as: cradle position, table elevation, clothing, lighting, table models, imaging accessories, camera installation position, and devices and operators positioned around the table. The training and validation data were used in the training process of the two Auto Positioning networks, while the testing data was used to test and verify the performance of the networks.

### Ground Truth for Test Dataset

The location of all eight of the anatomical landmarks were marked (labeled) on the test's dataset images by at least two experienced application specialist and three development engineers, to capture expected randomness of human identification. The three development engineers were trained by clinical application specialists. The locations of the eight anatomical landmarks were marked on the RGB images. The landmark locations were not directly marked on the 3D depth images by the humans because the unique appearance of these images would make the task unrealistic. Rather, the human-identified landmarks from the RGB image were geometrically transposed to the 3D depth image using the known positional relationship between the RGB sensor and the depth sensor of the 3D camera.

### Challenging Cases in the Test Set

The following conditions/variables were captured in the test set and include in the test results.

- Patient covered with a sheet, patient wearing a hospital gown, patients wearing various types/style/colors of street clothing
- Small height children (< 1.2 m), tall patients (> 2.0 m), obese patients
- CT imaging accessories including patient positioning straps, head holder, cradle extender, knee support pad
- Operators in random positions around the patient, patient occlusion by the operator, patient occlusion by clinical equipment
- Light levels (dim, medium, strong) in the scan room, oblique lighting from windows

| Variables | Training/ Validation dataset | Test dataset |
|---|---|---|
| Dataset size | ~22,000 images of 235 people | ~3500 images of 118 people |
| Volunteer height | 120 cm to 190 cm | 80 cm to 210 cm |
| Volunteer weight | 35 kg to 100 kg | 15 kg to 180 kg |
| Volunteer age | 10 to 55 years old | 2 to 55 years |
| Volunteer orientation | Head first, feet first | Head first, feet first |
| Volunteer posture | Supine, lateral, prone, cardiac posture for scan<br>Patient hand positions: over the head, alongside of the body, placed on chest | Supine, lateral, prone, cardiac posture for scan<br>Patient hand positions: over the head, alongside of the body, placed on chest |
| Volunteer clothing | Normal summer and winter clothes | Normal summer and winter clothes, and hospital gowns |
| Height from the Table top to scan plane isocenter | -100 mm, -150 mm, -250 mm, -300 mm, -500 mm | -100 mm, -150 mm, -250 mm, -300 mm, -400 mm, -500 mm |
| Height of camera above the floor | 2.3 m, 2,4 m, 2,5 m, 2.6 m, 2.7 m | 2.2 m, 2.3 m, 2,4 m, 2,5 m, 2.6 m, 2.7 m, 2.8 m |
| Camera installation angle | 0°, 7.5°, 15°, 25° | 0°, 7.5°, 15° |
| Table models | Lite table, GT1700 table, GT 1700 MidV table, GT 2000 table, no CT table | Lite table, GT1700 table, GT 1700 MidV table, GT 2000 table, no CT table |
| Light conditions | **Dim** (all lights off), **Medium** (half of the lights on), **Strong** (all lights on), **Natural light from windows** (provides an oblique light source) | **Dim** (all lights off), **Medium** (half of the lights on), **Strong** (all lights on), **Natural light from windows** (provides an oblique light source) |
| Other environmental conditions | w/wo sheet<br>w/wo random operator around<br>w/wo head holder<br>w/wo cradle extender | w/wo sheet<br>w/wo patient positioning straps<br>w/wo knee support pad<br>w/wo random operator around<br>w/wo head holder<br>w/wo cradle extender |

*Table 2: Variables considered in the datasets.*


*Table 2: Variables considered in the datasets.*

## Test Methods

The testing was performed following the below steps:

1. **Generate Acceptance Area from Labels**
   On each image, for each of the eight anatomical landmarks identified on the RGB image by the humans for the determination of the ground truth (and their corresponding locations on the depth image), a circular acceptance area was determined based on the variation of the human-identified labels. The average two-dimensional position of human labels is the center of the circular area as calculated using the formula in Figure 7. $(X, Y)$ are the coordinates of the center point. $xi$ and $yi$ are the coordinates for each label.

   The standard deviation of the difference between the label's position and the center of the acceptance area is the radius of acceptance area as calculated using the formula in Figure 8.

   Figure 9 illustrates the calculation acceptance area. This method for determination of the acceptance area is utilized to eliminate the variance and subjective understanding of labelers for the landmark points.

2. **AI Prediction of the Landmark Points**
   Locations of the eight anatomical landmarks (key points) on both of the paired RGB depth images is estimated by the deep learning networks. The RGBLandmarkNet network outputs the eight estimated key points on the RGB image. The DepthLandmarkNet network outputs the eight estimated key points on the depth image.

   All sixteen key points and their confidence level are generated from a post-processing step for each of the deep learning networks, with eight channel heatmaps. The peak of the heatmap for each channel is regressed as the pixel location of the corresponding anatomical landmarks. A confidence level is also generated according to the height of the peaks in each heatmap. With the input of the anatomy reference information from the user-selected protocol, one of the key points and its corresponding confidence level, from each network, is chosen. These two key points and their confidence levels are compared and one of them is selected for use as the anatomical landmark for the scout according to some logics.

3. **Check Prediction Results vs. Acceptance Area**
   The final outputted estimated landmark from Step 2 on previous page is compared against the corresponding landmark's acceptance area from Step 1 in order to calculate the success rate. If the estimated landmark's position lies within its corresponding acceptance area, it is recorded as a success, otherwise it is recorded as a failure, as illustrated in Figure 10.

## Test Results

Figure 11, below, shows the results of test set evaluation. The results are broken down by anatomical landmark. The results are also shown for the standalone performance of each of the two networks and for the final, combined output used by Auto Positioning. The overall success rate is 91.67%.

| | RGBLandmarkNet Output | DepthLandmarkNet Output | Final Key Points Output |
|---|---|---|---|
| OM | 97% | 74% | 97% |
| SN | 96% | 99% | 99% |
| XY | 74% | 90% | 90% |
| IC | 82% | 99% | 99% |
| KN | 71% | 78% | 78% |
| AJ | 87% | 64% | 64% |

*Figure 11: Landmark points accuracy test result.*

**Generate Acceptance Area from labels**

▼

**AI prediction of landmark points**

▼

**Check prediction result in Acceptance Area**

*Figure 6: Steps for landmark points prediction accuracy*

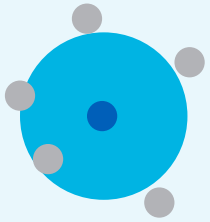
$$X = \frac{\sum_{t=1}^{n} xi}{N}$$

$$Y = \frac{\sum_{t=1}^{n} yi}{N}$$

*Figure 7: Method for calculation of the center of the acceptance area.*

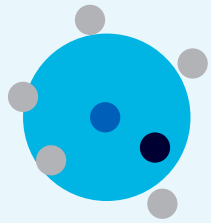
$$R = \sqrt{\frac{\sum_{t=1}^{n} [(X-x_t)^2+(Y-y_t)^2]}{N}}$$

*Figure 8: Method for calculation of the radius of the acceptance area.*



Acceptance area
Original labels
Averaged label

*Figure 9: Acceptance area for each landmark point.*



Acceptance area
Original labels
Averaged label
Estimated position

*Figure 10: Estimated landmark point location in acceptance area.*

## Clinical Testing

The clinical external evaluation testing was performed in a hospital environment on sequential patients, without regard to the type of the CT examination or patient condition. The CT technologists involved were those on the hospital's normal scheduling and did not receive coaching during the exams. The actions of the technologist with regard to the use of Auto Positioning were observed and recorded for the evaluation, and two evaluation tests were performed: Auto Positioning accuracy and auto centering accuracy.

1. **Auto Positioning Accuracy**
   The testing was performed following the below steps:

**AI-predicted landmarks** ▶ **Generate scan range** ▶ **Operator review and confirm scan range** ▶ **Scout scan** ▶ **Check scout image for axial scan**

*Figure 12: Steps used during the clinical evaluation testing.*

   a. **AI Predicted Landmark:** As described in the previous section, the location of the eight landmark points is estimated for each patient. The anatomical reference in the selected protocol is used to select one of the landmark points as the landmark for the scan.

   b. **Generate Scout Scan Range:** The selected protocol's scout scan range is applied to the selected landmark's position from Step 1 to produce the scout's start and end locations. These locations are overlaid in real time on the RGB video display.

   c. **Operator Review and Confirmation of the Scout Scan Range:** The CT technologist reviews the displayed scan range. The technologist is are able to adjust the lines of scout's start and end location directly on UI if they are not satisfied with the locations generated by Auto Positioning. The technologist must confirm the locations before being able to proceed in the Auto Positioning workflow.

   d. **Scout Scan:** After confirmation of the scout's start and end locations, the patient is moved into the gantry and the protocol's scout scan(s) are performed.

   e. **Check of the Scout Image(s) for the Protocol's Diagnostic Axial/Helical Scan(s):** Identical to the normal workflow, the CT technologist uses the scout image(s) to determine and set the diagnostic scan(s)' stop and start locations.

### Criteria Used to Determine Success/Failure of Auto Positioning's Scout Scan Range

If one of the below situations occur for a patient's exam, a failure of Auto Positioning is recorded, otherwise Auto Positioning's performance is recorded as a success for the exam.

- If the CT technologist adjusts both the displayed scout start and end locations in the same direction in Step 3 on previous page, this is indicative of an inaccurate landmark determined in Step 1 and is recorded as a failure of Auto Positioning
- If the scout image(s) acquired in Step 4 do not have sufficient coverage to enable the technologist to set up the diagnostic scan(s) in Step 5 and the patient needs to be re-scouted, a failure of Auto Positioning is recorded. This is conservative because the lack of sufficient coverage may be the result of a mistake by the technologist and not a failure of Auto Positioning

### Test Results

The combined performance of Auto Positioning in 566 of the 576 exams (98.26%) performed at evaluation site was recorded as successful.

The distribution of the CT exam conditions in the external evaluation is shown in Figure 13.

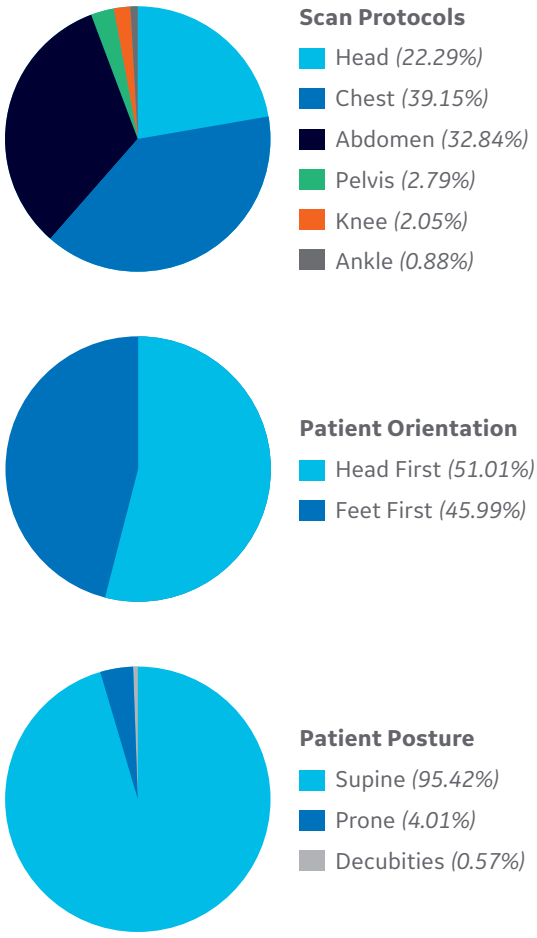



*Figure 13: Combined distribution of CT exam conditions from both evaluation sites.*

## 2. Auto Centering Accuracy

### Test Method

In the Auto Positioning accuracy test, 2 scout images (0˚ and 90˚) are acquired. The lateral scout images (90˚) can be used to calculate the patient centering based on scout attenuation, shown in Figure 14. This centering results will be compared with the gantry geometry center in elevation direction. The difference between them (patient centering based on scout attenuation minus gantry geometry center in elevation direction) is defined as off-centering offset. Positive offset means patient is positioned higher and negative offset means patient is positioned lower than geometry center. Ideally, this off-centering offset should be 0, which means patient is centered perfectly in elevation direction.

### Criteria Used to Determine the Performance of Auto Centering

If off-centering offset based on lateral scout images is within 20 mm, patient centering is good enough for following diagnostic scans. Clinical data including Head, Chest, and Abdomen/Pelvis are used to demonstrate the auto centering performance.

### Test Results

The auto centering performance for Head, Chest, and Abdomen/Pelvis are shown in Figure 15.

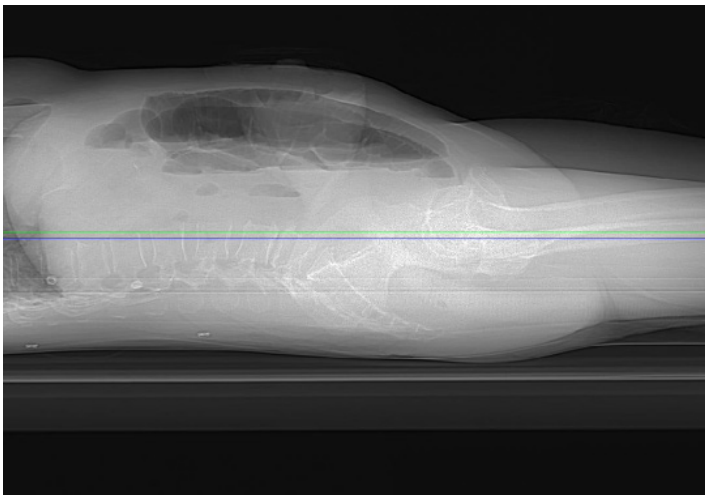


*Figure 14: Auto centering testing (Green line: gantry geometry center in elevation direction; Blue line: patient centering based on lateral scout attenuation).*

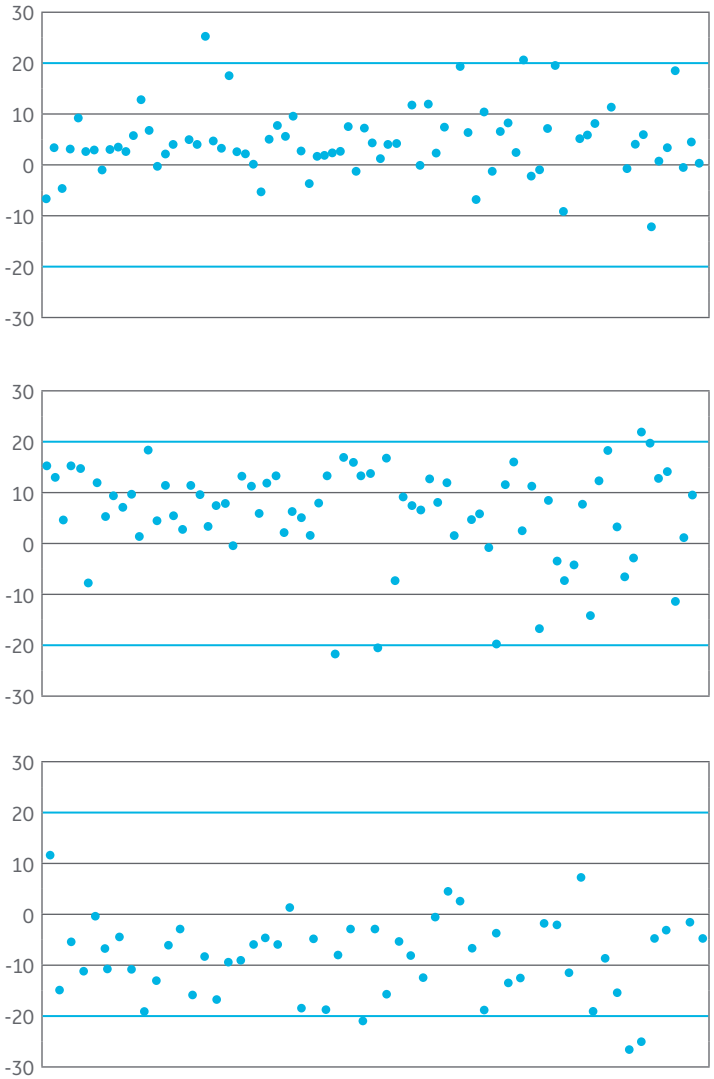



*Figure 15: Auto centering testing results. (a) Head (b) Chest (c) Abdomen/Pelvis.*

| | Within 20 mm |
|---|---|
| Head | 97.4% |
| Chest | 94.7% |
| Abdomen/Pelvis | 94.3% |

*Table 3: Auto centering performance.*



February 2021
JB0000000