

**DE NOVO CLASSIFICATION REQUEST FOR
CAPTION INTERPRETATION AUTOMATED EJECTION FRACTION SOFTWARE**

REGULATORY INFORMATION

FDA identifies this generic type of device as:

Radiological machine learning-based quantitative imaging software with predetermined change control plan. A radiological machine learning based quantitative imaging software with predetermined change control plan is a software-only device which employs machine learning algorithms on radiological images to provide quantitative imaging outputs. The device includes functions to support outputs such as view selection, segmentation and landmarking. The design specifications include planned modifications that may be made to the device consistent with an established predetermined change control plan.

NEW REGULATION NUMBER: 21 CFR 892.2055

CLASSIFICATION: Class II

PRODUCT CODE: QVD

BACKGROUND

DEVICE NAME: Caption Interpretation Automated Ejection Fraction Software

SUBMISSION NUMBER: DEN220063

DATE DE NOVO RECEIVED: September 28, 2022

SPONSOR INFORMATION:

Caption Health, Inc.
2000 Sierra Point Pkwy, 8th Floor
Brisbane, CA 94005

INDICATIONS FOR USE

The Caption Interpretation Automated Ejection Fraction software is used to process previously acquired transthoracic cardiac ultrasound images, to store images, and to manipulate and make measurements on images using an ultrasound device, personal computer, or a compatible DICOM-compliant PACS system in order to provide automated estimation of left ventricular ejection fraction. This measurement can be used to assist the clinician in a cardiac evaluation.

The Caption Interpretation Automated Ejection Fraction Software is indicated for use in adult patients.

LIMITATIONS

Prescription Use only: The sale, distribution, and use of this device are restricted to prescription use in accordance with 21 CFR 801.109.

Warnings:

- Product users are responsible for image quality and diagnosis. A qualified medical professional must inspect the data being used for analysis and diagnosis, and ensure that the data is sufficient and appropriate in anatomical correctness and both spatial and temporal resolution for the measurement being employed.
- The Caption Interpretation Automated Ejection Fraction Software is not intended for transesophageal echocardiograms (TEE).
- This device must be operated by or under the direction and supervision of a licensed physician.

PLEASE REFER TO THE LABELING FOR A COMPLETE LIST OF WARNINGS, PRECAUTIONS AND CONTRAINDICATIONS.

DEVICE DESCRIPTION

The Caption Interpretation Automated Ejection Fraction Software (“AutoEF”) applies machine learning algorithms to process two-dimensional transthoracic echocardiography images for calculating left ventricular ejection fraction.

The current implementation of the device adds a predetermined change control plan (PCCP) to the device cleared under K210747, which allows future modifications to the device.

The version of Caption Interpretation AutoEF cleared under K210747 performs left ventricular ejection fraction estimation using apical four chamber (A4C), apical two chamber (A2C) and the parasternal long-axis (PLAX) cardiac ultrasound images.

The software uses an algorithm that was derived through use of deep learning and locked prior to validation. The product operates as an add-in to a DICOM PACS system, ultrasound device, or personal computer. Caption Interpretation receives imaging data either directly from an ultrasound system or from a module in a PACS system.

The device includes the following main components:

1. **Clip Annotation and Selection:** The AutoEF software includes a function that processes video clips in a study to automatically classify clips that are PLAX, AP4, and AP2 views. This view selection is based on a convolutional neural network. It also includes a function, Image Quality Score (IQS), that allows selection of best available PLAX, AP4, and AP2 clips within the study or provide an indication to the user if there are no clips with sufficient quality to estimate ejection fraction, based on prespecified IQS thresholds.

2. **Ejection Fraction Estimator and Confidence Metric:** The automated ejection fraction estimation is performed using a machine learning model trained on apical and parasternal long-axis views. The model is trained with a dataset from a large number of patients, representative of the intended patient population and variety of contemporary cardiac ultrasound scanners. The EF calculation can be performed on a 3-view combination (PLAX, AP4 and AP2), 2-view combinations (AP4 and AP2, AP2 and PLAX, AP4 and PLAX), or single views (AP4, AP2). The confidence metric provides expected error in left-ventricle ejection fraction estimation and is based on IQS.

The outputs of the software are presented to the clinician. The outputs are as follows:

- EF calculation stated as a percentage (e.g., “59%”);
- Predicted standard error based on image quality (e.g., “±6%”);
- Qualitative characterization of the EF estimate mapped to guidelines from the American Society of Echocardiography (e.g., “Normal”);
Predicted likelihood of the qualitative bin based on the standard error listed above (e.g., “80% likelihood”);
- Information about which clip(s) and view(s) were used as the basis of the EF calculation;
- Single-view EF calculations of views incorporated into overall EF calculation (e.g., “PLAX: 55%; AP4: 63%; AP2: 58%”).

The clinician can edit the estimation, if desired, and must decide whether to accept or reject it for use in the final report.

PREDETERMINED CHANGE CONTROL PLAN (PCCP)

Caption Interpretation Auto Ejection Fraction Software includes an FDA-authorized predetermined change control plan (PCCP). The authorized PCCP includes detailed description of planned device modifications, and the associated methodology to develop, validate and implement those modifications in a manner that ensures the continued safety and effectiveness of the Caption Interpretation Automated Ejection Fraction Software. The plan includes the following four modification categories:

- Modification #1-Training on Additional Data
- Modification #2-Incorporation of additional 2D TTE views
- Modification #3-Optimization of implementation of core algorithm(s) in software
- Modification #4-Improving Algorithm Operating Speed

Verification and Validation Activities

PCCP describes specifications and methods to achieve appropriate risk control for anticipated changes. Risk Analysis will be performed to identify any new hazards created by algorithm changes, and to trace mitigation activity and completeness.

The evaluation of the modification will be performed with verification at the algorithm, subsystem, and end-to-end system level. Functional and end-to-end verification tests confirm proper operation of revised algorithms in the product. Validation will be performed at the end-to-end system level with expert-labeled ground truth data. Training and test datasets will be independent. The performance targets must meet the prespecified criteria as detailed in the

PCCP. The specific verification and validation details for each individual modification are described in the PCCP.

The PCCP specifies a minimum percentage of new data for the test datasets, and that test datasets may not be used more than a specified number of times for the same purpose.

The key acceptance criteria included, but were not limited to, the following:

1. 80% positive predictive value and 80% sensitivity for correct mode, view, and minimum number of frames.
2. 80% of clips meet expert criteria for suitability for EF estimation.
3. Overall (all views) and all combined views Auto EF is within 9.2% RMSD of expert EF.
4. New views superior to 11.024% RMSD individually.
5. For each standard view, the Confidence Metric must meet the equivalence to expert EF criteria as defined in PCCP.

Implementation of Modifications and Control

The machine learning algorithms used in Caption Interpretation Auto Ejection Fraction Software are trained, tuned, and “locked” prior to release i.e., they do not continuously learn in the field. Caption Health follows its design controls as well as the test methods and acceptance criteria defined in the PCCP to verify and validate new release versions of Caption Interpretation AutoEF. Caption Health intends to deploy a single version of the software within the United States, e.g., there will not be site specific versions of the software.

Communication and transparency to users

If a new release version of the software is available, due to implementation of modification made in accordance with the PCCP, Caption Health notifies all existing customers that an update is available and requests that they upgrade to the latest version of the software. Caption Health continues to follow-up with each customer to ensure that they are on the latest version of the software. Version numbering of the Caption Interpretation software will be updated in the software interface and in the User Manual. User manual will be updated to reflect the associated performance, associated inputs, supporting evidence, version history and other necessary information about the modification(s) made pursuant to the PCCP.

SUMMARY OF NON-CLINICAL/BENCH STUDIES

Software

The only change in this device compared to the Caption Interpretation Auto Ejection Fraction Software cleared in K210747 was the addition of PCCP. The software documentation was, therefore, referenced from K210747. The device was identified as having a moderate level of concern as defined in the FDA guidance document “*Guidance for the Content of Premarket Submissions for Software Contained in Medical Devices.*” The software documentation included:

1. Software/Firmware Description
2. Device Hazard Analysis
3. Software Requirement Specifications

4. Architecture Design Chart
5. Software Design Specifications
6. Traceability
7. Software Development Environment Description
8. Verification and Validation Documentation
9. Revision Level History
10. Unresolved Anomalies
11. Cybersecurity

Software documentation and testing provided demonstrate that the device meets all the recommendations in the guidance document.

SUMMARY OF CLINICAL INFORMATION

Algorithm Performance Testing

The only component of this device which is new compared to the previously cleared device in K210747 is the addition of a PCCP. No new performance testing was provided for the current device and performance testing used for clearance of K210747 also applies to this device and is summarized below.

Caption Health AutoEF performance, including the integrated operation of the Clip Selector, was validated on a dataset of 186 patient studies in a retrospective multicenter study. These studies were representative of patient conditions, including range of ejection fraction values, and a significant portion of high body mass index patients as is typically associated with technically difficult studies.

Clip Annotation Performance

A Clip Annotator study verified the ability of the software to receive, annotate and select clips for Auto EF calculation. Results of the Clip Annotator were compared to evaluation by a panel of expert readers. The study-level positive predictive value (PPV) and sensitivity were computed, and each was tested for whether it exceeded the pre-specified performance goals of 80% PPV and 80% sensitivity. That study met the pre-defined acceptance criteria and found that the observed PPV point estimates for the Clip Annotator were greater than 97% for identification of the imaging mode and the view. Similarly, observed sensitivity point estimates were greater than 96% across views and imaging mode.

Clinical Validation of AutoEF

Caption Interpretation AutoEF was also the subject of a pivotal clinical investigation to validate successful performance of the EF calculation in comparison to conventional EF calculation methods. The study comprised a dataset of 186 patient studies acquired from three sites across the US. This study was conducted to support the clearance of K210747.

The dataset demographics and distribution of ultrasound scanners are tabulated below:

Table 1a: Dataset Demographics - Sex

Sex	Studies <i>n</i>, %
Male	104 (56%)
Female	82 (44%)
Total	186 (100%)

Table 1b: Dataset Demographics - BMI

BMI	Studies <i>n</i>, %
BMI < 25 kg/m ²	68 (37%)
25 ≤ BMI < 30 kg/m ²	64 (34%)
BMI ≥ 30 kg/m ²	54 (29%)
Total	186 (100%)

Table 1c: Dataset Demographics – Ejection Fraction

LVEF	Studies <i>n</i>, %
EF < 30%	33 (18%)
30% ≤ EF < 53%	80 (43%)
EF ≥ 53%	73 (39%)
Total	186 (100%)

Table 1d: Dataset Demographics – Institution

Institution	Studies <i>n</i>, %
Minneapolis Heart Institute	60 (32%)
Duke University	66 (36%)
Northwestern University	60 (32%)
Total	186 (100%)

Table 1e: Distribution of Ultrasound System in Dataset

Ultrasound System	Studies <i>n</i>, %
Philips/CX50	4 (2%)
Acuson/SEQUO IA	23 (12%)
Philips/EPIQ 7C	2 (1%)
GE/Vivid i	20 (11%)

GE/Vivid E9	9 (5%)
GE/Vivid E95	24 (13%)
GE/Vivid7	8 (5%)
Siemens/ ACUSON SC2000	6 (3%)
Philips/iE33	90 (48%)
Total	186 (100%)

Reference Standard: The reference *standard* for the ejection fraction was established based on by 2D echo using the biplane Modified Simpson's method of disks, by expert cardiologists. The test compared the Caption Health AutoEF to the expert produced and reported biplane Modified Simpson's ejection fraction.

Study Endpoints: The primary endpoint was the root mean square deviation (RMSD) between AutoEF derived values based on the best available view and the reference standard. The acceptance criterion was defined as RMSD superior to 9.2% for view combinations. Likewise, the acceptance criteria for the combination of views were also defined as RMSD superior to 9.2%. The performance criterion for single views was defined as RMSD superior to 11.024%.

Results:

The primary endpoint for the subject device was met with RMSD EF%[95% CIs]: 7.21[6.62, 7.74] for the best available view.

Likewise, the results for evaluating the performance of the algorithm against its objective performance goal for combinations of views was also met. The results for the combination views are summarized in the table:

#	View Combination	RMSD EF% [95% CI]	p-value
1	AP4 and AP2	7.27 [6.55, 7.92]	< 0.0001
2	AP4 and PLAX	7.50 [6.85, 8.09]	< 0.0001
3	AP4, AP2, and PLAX	7.24 [6.64, 7.80]	< 0.0001
4	AP2 and PLAX	8.04 [7.32, 8.7]	0.0002
5	Best available	7.21 [6.62, 7.74]	< 0.0001

Results for testing of performance for individual views also met the acceptance criterion:

#	View	RMSD EF% [95% CI]	p-value
1	AP4 only	7.76 [7.01, 8.45]	< 0.0001
2	AP2 only	8.27 [7.44, 9.03]	< 0.0001

Finally, testing of the confidence metric functionality verified successful performance of the Confidence Metric in estimating the error range of the EF estimates around the reference EF using equivalence criteria and the evidence that the difference between the estimated EF and the reference EF is normally distributed.

Subgroup analyses were provided for range of EF, BMI, gender, institution, and ultrasound systems with statistically insignificant differences in RMSD across the subgroups.

PLEASE REFER TO THE LABELING FOR COMPLETE RESULTS OF THE SUBGROUP ANALYSIS.

LABELING

The labeling supports the decision, including information on detailed description of device inputs and outputs, instructions for use, intended patient population and intended users of the device, adequate warnings and precautions as well as detailed performance testing summary. Additionally, the labeling mentions that the device has a PCCP and provides high-level details of the modifications agreed to in the PCCP. The labeling also includes information on how the modifications will be implemented and how the users will be informed about the device modifications made in accordance with the PCCP. Additionally, version history of the device is also included in the labeling. All the necessary information to grant the De Novo request for this device is available in the labeling.

RISKS TO HEALTH

The table below identifies the risks to health that may be associated with use of a radiological machine learning-based quantitative imaging software with predetermined change control plan and the measures necessary to mitigate these risks.

Identified Risks to Health	Mitigation Measures
Inaccurate device output leading to patient receiving incomplete or suboptimal treatment/diagnosis	Design verification and validation activities identified in special control (1); Certain labeling information identified in special control (4)
Implementation of modifications agreed in the authorized PCCP leads to algorithm producing inaccurate output, including:	Special controls (2)-(3) and 4(vii) Certain activities identified in special controls (1)

• Performance related to existing specifications at the time of clearance • Performance related to planned additional device capabilities and associated specifications	
Misunderstanding of changes to the device input criteria, output performance, or other aspects of the design as changes are implemented under the PCCP, leading to misuse and incorrect treatment/diagnosis	Special control (2)-(3); Labeling information identified in special control (4)(vii)

SPECIAL CONTROLS

In combination with the general controls of the FD&C Act, the radiological machine learning-based quantitative imaging software with predetermined change control plan is subject to the following special controls:

- (1) Design verification and validation must include:
 - (i) A detailed description of the image postprocessing algorithms, including a detailed description of the algorithm inputs and outputs, each major component or block, and algorithm limitations.
 - (ii) Detailed description of training data including detailed annotation methods and important cohorts (e.g., subsets defined by patient demographics, clinically relevant confounders, and subsets defined by image acquisition characteristics).
 - (iii) Performance testing protocols and results that demonstrate that the underlying algorithms function as intended. The performance assessment must be based on objective performance measures (e.g., error metrics, Bland-Altman plots, dice similarity coefficient (DSC), Hausdorff distance, sensitivity, specificity, predictive value). The test dataset must be independent from data used in training/development and contain sufficient numbers of cases from important cohorts (e.g., subsets defined by clinically relevant confounders, effect modifiers, concomitant diseases, 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) Software verification, validation, and hazard analysis.
- (2) As part of the design verification and validation activities, you must document the planned device modifications of the quantitative imaging software, and the associated methodology for the development, verification, and validation of modifications made consistent with the performance requirements in the plan.
- (3) As part of the risk management activities, you must identify and assess the risks of the planned modification(s) and identify corresponding risk mitigations.
- (4) Labeling must include:

- (i) A detailed description of the patient population for which the device was validated;
- (ii) A description of the intended user and expertise needed for safe use of the device;
- (iii) A detailed description of the device inputs and outputs;
- (iv) A detailed description of compatible imaging hardware and imaging protocols;
- (v) A detailed summary of the current performance of the device and a summary of the performance testing conducted to support safe and effective use of the device including test methods, dataset characteristics (including demographics), testing environment, results (with confidence intervals), and a summary of sub-analyses on case distributions stratified by relevant confounders;
- (vi) A description of 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) Labeling related to the predetermined change control plan (PCCP), including:
 - (A) A statement that the device has a PCCP;
 - (B) A description of modification(s) implemented for quantitative imaging and supporting algorithms, including a summary of current performance, associated inputs, validation requirements, and related evidence; and
 - (C) A version history, a description of how device modification(s) will be implemented, and a description of how users will be informed of device modification(s) made in accordance with the PCCP.

BENEFIT/RISK DETERMINATION

Left Ventricle Ejection Fraction (LVEF) is an important parameter for assessing the severity of decrease in the systolic function of heart and in directing patient management for various cardiovascular diseases. Caption interpretation automated ejection fraction software, based on machine learning applied to 2D transthoracic cardiac ultrasound images, provides automated estimation of LVEF. The device performance demonstrated using a clinical dataset, covering the intended patient population and representative ultrasound systems, met the pre-specified performance criteria, which translate to clinically significant probable benefits.

The probable benefits of machine learning-based devices include the capability to *improve* their performance through iterative modifications, including learning from real-world data. Caption interpretation automated ejection fraction software is a machine learning based device. It includes a predetermined change control plan (PCCP) which supports performance *improvement* through specific device modifications. The authorized PCCP includes detailed description of planned device modifications, and the associated methodology to develop validate and implement those modifications in a manner that ensures the continued safety and effectiveness of the Caption interpretation automated ejection fraction software.

The probable risks for this device are related to inaccurate results which could impact patient management decisions; it may be exacerbated by transparency of the communication to the end users of the device about intended use and performance. Causes may include:

- Inaccuracies in the initial algorithm, including problems with the generalizability of the device performance due to its underlying models being based on machine learning.
- Inaccuracies after implementation of modification(s) agreed in the authorized PCCP
- Lack of transparency to end user about changes to device made through PCCP.

The probable benefits of this machine learning-based device include plan to implement predetermined modifications that would improve the overall performance. These benefits outweigh the risks, given the mitigations employed including the special controls in combination with the general controls. In particular, the implementation plan for the modifications authorized in the PCCP would allow enhancing the generalizability of the device and improving the performance for quantitative imaging applications while ensuring continued safety and effectiveness. The implementation of modifications in the authorized PCCP, made according to a verification and validation plan developed in conjunction with FDA input and agreed to by FDA, would allow health care professionals quicker access to improved tools for patient management, without necessitating additional marketing submissions.

Patient Perspectives

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

Benefit/Risk Conclusion

In conclusion, given the available information above, the data support that for the intended use and indications for use statements above, the probable benefits outweigh the probable risks for the Caption Interpretation Automated Ejection Fraction Software. The device provides benefits and the risks can be mitigated by the use of general controls and the identified special controls.

CONCLUSION

The De Novo request for the Caption Interpretation Automated Ejection Fraction Software device is granted and the device is classified under the following:

Product Code: QVD

Device Type: Radiological machine learning-based quantitative imaging software with predetermined change control plan

Class: II

Regulation: 21 CFR 892.2055