

**DE NOVO CLASSIFICATION REQUEST FOR
FIBRESOLVE**

REGULATORY INFORMATION

FDA identifies this generic type of device as:

Radiology software for referral of findings related to fibrotic lung disease.

Radiology software for referral of findings related to fibrotic lung disease is a prescription image processing device that analyzes computed tomography images to suggest the presence of disease or of an imaging finding suggestive of disease. The output of this device is intended to be used as adjunctive information as part of a referral pathway in the overall diagnostic assessment process.

NEW REGULATION NUMBER: 21 CFR 892.2085

CLASSIFICATION: Class II

PRODUCT CODE: QWO

BACKGROUND

DEVICE NAME: Fibresolve

SUBMISSION NUMBER: DEN220040

DATE DE NOVO RECEIVED: June 29, 2022

SPONSOR INFORMATION:

Imvaria, Inc.
2930 Domingo Ave, #1496
Berkeley, CA 94705

INDICATIONS FOR USE

Fibresolve is a software-only device that receives and analyzes lung computed tomography (CT) imaging data in order to provide a diagnostic subtype classification in suspected cases of interstitial lung disease (ILD). The device supplements the standard-of-care workflow by providing a qualitative, diagnostic classification output of imaging findings based on machine learning pattern recognition, in order to provide adjunctive information as part of a referral pathway to an appropriate Multidisciplinary Discussion (MDD) or as part of an MDD. Specifically, the tool is used to serve as an adjunct in the diagnosis of idiopathic pulmonary fibrosis (IPF) prior to invasive testing. The results of Fibresolve are intended to be used only by clinicians qualified in the care of lung disease,

specifically in caring for patients with ILD, in conjunction with the patient's clinical history, symptoms, and other diagnostic tests, as well as the clinician's professional judgment.

The input to Fibresolve is a DICOM-compliant lung CT scan. Clinical case eligibility includes the following criteria:

Age > 22 years old.

Pulmonary symptoms suggestive of possible ILD including IPF.

LIMITATIONS

- The sale, distribution, and use of the Fibresolve are restricted to prescription use in accordance with 21 CFR 801.109.
- Fibresolve cannot be used to rule out IPF.
- Always ensure results should be used in conjunction with other clinical and diagnostic findings, consistent with professional standards of practice, including information obtained by alternative methods, and clinical evaluation, as appropriate. The output should not be solely relied upon as the sole determinant for referral and adequate patient follow-up or management.
- Series slice thickness must be $\leq 3\text{mm}$.

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

DEVICE DESCRIPTION

Fibresolve is a Software as a Medical Device (SaMD) for the qualitative disease assessment of DICOM-compliant chest computed tomography (CT) imaging for the detection of image content consistent with patterns found in patients with idiopathic pulmonary fibrosis (IPF). Fibresolve uses a deep learning algorithm that assesses CT images for patterns consistent with specific disease diagnosis, identifying patterns consistent with IPF among cases of Interstitial Lung Disease (ILD). Fibresolve produces a binary output ("Suggestive of IPF" or "Inconclusive"). Fibresolve does not provide any visual aid to the clinician to assist in interpreting the image.

The device consists of the following 3 components: (1) Image Receiver application programming interface (API) for image acquisition in the cloud; (2) Ingestion Pipeline and Analysis System for image processing and analysis; and (3) Output API for device output transmission.

(1) The Image Receiver API is accessed via any DICOM-compliant system (e.g., PACS). Images are submitted through the API by the hospital or clinic, or by the manufacturer. The API passes the images to the Ingestion Pipeline and Analysis System (2). The input to the device is a single stack of axial slice, DICOM-compliant, 3 mm thickness or less lung CT images from a list of validated device manufacturers.

(2) (a) The Ingestion Pipeline and (b) Analysis System accepts the images, selects cases appropriate for processing, processes the images for analyses, analyzes the images, and stores the

images. This Analysis System includes the analysis algorithm that generates the assessment for the case. The device outputs a report which includes identifying information and technical details about the case data and a binary result stating whether the data are determined to be suggestive for the target disease state.

- The Ingestion Pipeline identifies applicable CT imaging series from the case and verifies that the series is valid, completes quality checks, and confirms adequacy for analysis.
- The analysis algorithm is a 3D deep learning model developed and trained using images from multiple facilities. No segmentation is performed as part of the Analysis Algorithm.

(3) The Output API transmits the report data for the clinician to review. The Output API is either integrated into the hospital or clinic notification software (e.g., electronic health record) for electronic transmission or the device manufacturer transmits the Report in human-readable format directly (e.g., via fax). The clinician then incorporates the device Report as part of diagnostic decision-making.

The system does not include an image viewer or produce visual output for diagnostic use.

Design Limitation: The device does not interpret images according to established clinical radiological features; the device cannot be used to infer the presence or absence of radiological features associated with the disease or condition named in the indications for use.

SUMMARY OF NON-CLINICAL/BENCH STUDIES

Software

The Fibresolve software documentation and testing provided demonstrate that the device meets all requirements outlined in the FDA guidance document, “*Guidance for the Content of Premarket Submissions for Software Contained in Medical Devices*” for software of moderate concern.

SUMMARY OF CLINICAL INFORMATION

Performance Testing

The device’s standalone performance as well as the performance in comparison to a clinical comparator were provided. The data used in all performance testing was sourced from the Lung Tissue Research Consortium (LTRC), with the dataset collected from sites independent from those used in device development. In total, 300 ILD cases were included, with primary analysis centered on a subset of 137 cases with CT reconstruction slice thickness ≤ 3 mm. Information on patient age, sex, race, ethnicity, smoking status, as well as CT manufacturer, site, and slice thickness were provided.

Table 1: Age, sex, and CT manufacturer information for all 300 cases and in the slice thickness ≤ 3 mm subset.

		Full Dataset (n=300) [%]	Slice Thickness ≤ 3 mm (n=137) [%]
Age	≤ 40	5.0	4.3
-	41-50	10.0	9.5
-	51-60	28.0	25.5
-	61-70	41.3	43.8
-	> 70	15.7	16.8
Sex	Female	49.7	46.7
-	Male	50.3	53.2
Manufacturer	GE Medical Systems	27.0	10.9
-	Philips	7.7	5.1
-	Siemens	63.0	81.8
-	Toshiba	2.3	2.1

Study Limitation: The test dataset acquisition period ranges from 2005 until 2018. The study results may not generalize for cases acquired with more modern equipment and imaging protocols. Special controls (including postmarket data collection) are employed to mitigate the risk associated with uncertain device performance for modern images.

The primary clinical comparator was an Expert Panel (EP, n=5), composed of expert pulmonologists and thoracic radiologists, who performed a chart review blinded to the diagnostic outcome and to all invasive testing, but given access to chart information such as demographics; relevant histories such as medical, smoking, and environmental exposures; medications, pulmonary function tests; rheumatic serologies; and CT imaging. The panel was instructed to return a binary result based on the 2018 ATS IPF Guidelines for the non-invasive diagnosis of IPF, which require a definite usual interstitial pneumonia (UIP) pattern by high resolution CT in conjunction with certain other chart information (Raghu et al. Am J Respir Crit Care Med. 2018; 198(5):e44-e68). Experts were instructed to not identify probable UIP radiographic criteria as IPF.

Study Limitation: Since completion of the study, Guidelines for the non-invasive diagnosis of IPF have been revised (Raghu et al. Am J Respir Crit Care Med. 2022; 205(9):e18-e47). These revisions broaden the CT requirements needed for the non-invasive diagnosis to also include the Probable UIP pattern category, when considered alongside other chart information. Therefore, the Expert Panel in the study operated according to more stringent decision criteria than is currently recommended, and may have exhibited a higher specificity and lower sensitivity than would be expected in current clinical practice as a result.

The LTRC data included reference standard diagnoses, used to establish device and EP performance. The reference standard diagnosis was assigned to each case by an MDD from the LTRC, assessing all clinical, imaging, laboratory, demographic, and pathologic data with an average of approximately 1-year of clinical information for patients in the registry. Final clinical diagnosis was provided. From the total 300-case dataset, 83 IPF cases were identified, with 217 non-IPF, consisting of a variety of other diagnoses including unclassifiable interstitial lung disease (18%), chronic hypersensitivity pneumonitis (17%), and nonspecific interstitial pneumonia (11%). In the subset of cases with slice thickness ≤ 3 mm, the radiological UIP pattern category was further retrospectively assigned by panel consensus, which was separately composed of an expert pulmonologist and radiologist pair.

Study Limitation: The reference standard diagnoses in the test dataset were established from 2005 until 2018. Imaging practices and interpretation guidelines for patients suspected of having an ILD have changed significantly in the past two decades, raising the uncertainty that the reference standard diagnoses used in performance testing reflect current clinical diagnosis practices. Special controls (including postmarket data collection) are employed to mitigate the risk associated with uncertain device performance arising from uncertain reference standard determination.

The primary endpoints of the study were that the device specificity surpass a pre-specified value of 80%, and that the device sensitivity be non-inferior to the EP.

Study Limitation: The Agency disagrees with these endpoints. Splitting sensitivity and specificity primary endpoints to different comparators is not statistically sound as it favors the device in both comparisons, by comparing the sensitivity to a control with a high decision threshold and the specificity to a control with a lower decision threshold. Nevertheless, the totality of evidence regarding device accuracy suggests that it offers a small probable benefit when considered as part of a referral pathway to an MDD or as part of an MDD. The current challenges of diagnosis of IPF were given weight in this assessment.

In the total 300-case dataset, both endpoints were met, with device sensitivity 41% (95% CI: 30-52%) and specificity 87% (95% CI: 81-91%); and EP sensitivity 13% (95% CI 6.8-23) and specificity 96% (95% CI: 93-98%). For the relevant ≤ 3 mm slice thickness subgroups:

Table 2: Device Performance per slice thickness

Slice Group	n	Sensitivity	Specificity	PPV
< 1.5 mm	55	56% [CI 31-78%]	68% [CI 50-82%]	45% [CI: 24-68%]
≥ 1.5 mm, < 3 mm	42	50% [CI: 23-77%]	93% [CI: 76-99%]	78% [CI: 40-97%]
3 mm	40	63% [CI: 24-91%]	91% [CI: 75-98%]	63% [CI: 24-91%]
≤ 3 mm	137	55% [CI: 39-71%]	82% [CI: 73-89%]	56% [CI: 40-72%]

Table 3: EP Performance per slice thickness

Slice Group	n	Sensitivity	Specificity	PPV
< 1.5 mm	55	11% [CI 1-35%]	92% [CI 78-98%]	40% [CI: 5-85%]
≥ 1.5 mm, < 3 mm	42	21% [CI: 5-51%]	96% [CI: 82-99%]	75% [CI: 19-99%]
3 mm	40	38% [CI: 9-76%]	97% [CI: 84-99%]	75% [CI: 19-99%]
≤ 3 mm	137	20% [CI: 9-36%]	95% [CI: 88-98%]	62% [CI: 32-86%]

In on-label ≤ 3 mm slice thicknesses, the device's sensitivity and specificity were 55% (95% CI: 39-71%) and specificity 82% (95% CI: 73-89%); and EP sensitivity 20% (95% CI 9-36%) and specificity 95% (95% CI: 88-98%). The PPV (positive predictive value) was similar between the device and the EP. The device appears to identify true positive IPF patients, at the expense of a high rate of false positives (specificity lower bound of 73%). This tradeoff offers a positive benefit-risk determination, considering the benefits provided by an early IPF diagnosis. This is

provided that the risks associated with additional workup and treatment that a false positive non-IPF patient would be exposed to are mitigated by appropriate MDD review.

Study Limitation: The device specificity was lower than 80% in relevant ≤ 3 mm slice groups, with confidence intervals extending down to 73% for the total ≤ 3 mm slice group. Appropriate labeling places device interpretation firmly within a referral pathway to an MDD or as part of an MDD to mitigate the risk that a false positive could be inappropriately treated for IPF or that a false positive could potentially forego potentially curative treatment for another fibrotic lung disease. The device's performance suggests that there is a small probable benefit for the device, and special controls (including postmarket data collection) are employed to reduce the risk associated with uncertain device performance.

Table 4: Device performance within radiologically defined UIP categories in ≤ 3 mm slices

Fibresolve Performance in Diagnosing IPF	n	Sensitivity	Specificity	PPV
Fibresolve in Definite UIP	19	67% [CI: 38-88%]	25.0% [CI: 1-81%]	77% [CI: 46-95%]
Fibresolve in Probable UIP	27	64% [CI: 35-87%]	62% [CI: 32-86%]	64% [CI: 35-87%]
Fibresolve in Indeterminate UIP	5	N/A	80.0% [CI: 28-99%]	N/A
Fibresolve in Alternative Diagnosis	86	27% [CI: 6-61%]	89% [CI: 80-95%]	27% [CI: 6-61%]

Table 5: EP performance within radiologically defined UIP categories in ≤ 3 mm slices

ECP Performance in Diagnosing IPF	n	Sensitivity	Specificity	PPV
ECP in Definite UIP	19	33% [CI: 12-62%]	25% [CI: 1-81%]	63% [CI: 24-91%]
ECP in Probable UIP	27	21% [CI: 5-51%]	100% [CI: 75-100%]	100% [CI: 29-100%]
ECP in Indeterminate UIP	5	N/A	100% [CI: 48-100%]	N/A
ECP in Alternative Diagnosis	86	0% [CI: 0-28%]	97% [CI: 91-99%]	0% [CI: 0-84%]

The device and EP performance in relevant radiological UIP categories shows potential differences according to those categories, although subgroups were not powered so these differences are not known to be statistically significant. The device may identify more patients as probable IPF than human readers, as suggested by differences in sensitivity between the device and EP. Some of these patients may especially benefit from consideration by an MDD, which may otherwise be delayed or foregone. However, the reported specificity of the device points to a high false positive rate. Interpretation of results by an MDD is critical to ensure that the benefits of the device for true positive IPF patients outweigh the risks of the device for false positive patients with other ILDs that have very different prognoses and treatment paths.

Study Limitation: The device specificity was lower than 80% in relevant radiological UIP pattern groups, with wide confidence intervals. Furthermore, there was low representation of the "Indeterminate UIP" category (n=5), which makes it difficult to draw any substantive conclusions regarding the device or EP performance in that group. Appropriate labeling places

device interpretation firmly within a referral pathway to an MDD or as part of an MDD to mitigate the risk that a false positive could be inappropriately treated for IPF or that a false positive could potentially forego potentially curative treatment for another disease. In this context, the device's performance suggests that there is a small probable benefit for the device, and special controls (including postmarket data collection) are employed to reduce the risk associated with uncertain device performance.

Pediatric Extrapolation

In this De Novo request, existing clinical data were not leveraged to support the use of the device in a pediatric patient population.

POSTMARKET SURVEILLANCE

The results of device performance testing suggest that it offers a small probable benefit within the intended context of use. However, as discussed in Study Limitation notes above, there were several issues that increased the uncertainty associated with these results. Considering the challenges associated with assembling data pertaining to the diagnosis of this uncommon disease, FDA is requiring a postmarket surveillance study be employed to further elucidate device performance, using modern imaging data for the on-label slice thicknesses. The results of this study are intended to be used to update device labeling in the future to narrow the uncertainty associated with the device performance in important clinical categories, such that the user can better understand and incorporate the adjunctive information provided by the device into their clinical decision making.

LABELING

The labeling meets the requirements of 21 CFR 801.109 for prescription devices and includes information on 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 summaries. Placement of the device as part of a referral pathway to an appropriate MDD or as part of an MDD is clearly stated, and there are warnings that clinical decisions should not be based solely upon the device's output and that the device cannot be used to rule out IPF. Postmarket data collection and its purpose is acknowledged in labeling.

RISKS TO HEALTH

The table below identifies the risks to health that may be associated with use of a radiology software for referral of findings related to fibrotic lung disease and the measures necessary to mitigate these risks.

Identified Risks to Health	Mitigation Measures
False positive findings leading to harmful incorrect and/or delayed management of a patient with an alternative underlying fibrotic lung disease.	Clinical performance testing Postmarket surveillance Labeling
False negative findings leading to harmful incorrect and/or delayed management of a patient with a fibrotic lung disease	Clinical performance testing Postmarket surveillance Labeling
Incorrect and/or delayed patient management due to the	Clinical performance testing

device being misused to analyze images from an unintended patient population or from images acquired with incompatible imaging hardware or image acquisition parameters	Postmarket surveillance Labeling
Incorrect and/or delayed patient management due to misinterpretation of device output or overreliance on device output for radiological image interpretation	Labeling
Device failure leading to the absence of results, delay of results, or incorrect results, leading to inaccurate or delayed patient assessment	Software verification, validation, and hazard analysis

SPECIAL CONTROLS

In combination with the general controls of the FD&C Act, radiology software for referral of findings related to fibrotic lung disease is subject to the following special controls:

- (1) Data obtained from premarket clinical performance validation testing and postmarket surveillance acquired under anticipated conditions of use must demonstrate that the device performs as intended when used to analyze data from the intended patient population, unless FDA determines based on the totality of the information provided for premarket review that data from postmarket surveillance is not required. The following must be met:
 - (i) Validation report(s) must include a detailed description of the data, criteria, and methods used to define the reference standard that were used to evaluate device performance. The reference standard used in the clinical validation must be justified for the condition named in the device's indications for use.
 - (ii) The performance of the device must be compared to an appropriate clinical control, e.g. the performance or agreement of clinicians performing the same task.
 - (iii) The performance assessment must be based on pre-specified diagnostic accuracy measures (e.g., receiver operator characteristic plot, sensitivity, specificity, positive/negative predictive values, and diagnostic likelihood ratios).
 - (iv) Test datasets must contain a sufficient number of cases from important cohorts (i.e., subsets defined by clinically relevant demographics, confounders, effect modifiers, concomitant diseases, challenging cases such as early 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.
- (2) Software verification, validation, and hazard analysis must be performed. Software documentation must include a detailed technical description of all image analysis algorithms, including the model inputs and outputs, each major component or block, and any model limitations.
- (3) Labeling must include:
 - (i) A detailed description of the clinical environment and context of use, including information on interpretation of outputs within the intended workflow;
 - (ii) A detailed description of compatible imaging hardware and imaging protocols;
 - (iii) A detailed summary of the performance testing for each device output, including: test methods, dataset characteristics, testing environment, results (with confidence

- intervals), and a summary of sub-analyses on case distributions stratified by relevant confounders;
- (iv) According to the timeframe included in any postmarket surveillance protocol approved by FDA to satisfy the requirements paragraph (1) of this section, a detailed summary of the postmarket surveillance data must be provided, including updates to the labeling to accurately reflect device performance based upon data collected during the postmarket surveillance experience; and
 - (v) Limiting statements that indicate:
 - (A) A description of situations in which the device may fail or may not operate at its expected performance level (e.g., the impact of poor image quality on device performance or degraded performance in certain subpopulations), as applicable, including any limitations in the dataset used to train, tune, and test the algorithm during device development;
 - (B) A discussion of what the device detects in the context of diagnosing fibrotic lung disease; and
 - (C) A warning that users should use the device in conjunction with other clinical and diagnostic findings, including information obtained by alternative methods and clinical evaluation, as appropriate.

BENEFIT/RISK DETERMINATION

To arrive at a subtype classification for ILD, a physician combines information from a variety of sources (patient history, imaging, pathology, pulmonary function testing, labs) which each contribute positively or negatively to affect the overall certainty for a given diagnosis. Fibresolve analyzes CT images based on no human-derived metrics associated with IPF and produces a binary output: “Suggestive of IPF” or “Inconclusive.”

The major risks of this device are derived from incorrect (false positive / false negative) results: A false positive result could lead to an incorrect diagnosis of IPF, with omission of curative medication for the true underlying ILD resulting in progression of disease and death; A false negative may result in missed or delayed diagnosis which is associated with worse outcomes in IPF, increased chance of unnecessary invasive biopsy in IPF when noninvasive diagnosis was feasible, or harmful treatment for a mistakenly diagnosed alternate ILD. These risks are partially mitigated by placing use of the device as part of an MDD or in a referral pathway to an MDD, and through clear labeling that a positive result is only “Suggestive of IPF” and that a negative result is “inconclusive” and does not rule out IPF.

The benefits of the device are ease of use, improved accessibility, and improved detection of IPF, which will need to be interpreted in light of a significant false positive rate by an appropriate MDD. Uncertainty is introduced into the study results by a substandard comparator arm, inappropriate statistical methods, and questions regarding how well the data represents the target population. The level of uncertainty for the probable benefits of the device is high. However, postmarket data can be employed to further elucidate device performance in the on-label slice thicknesses and in a transparent demonstration of device binary output performance in the four SOC radiological UIP categories. These postmarket activities will increase transparency for users

regarding the performance of the device in US populations according to the SOC imaging analysis as described by clinical society guidelines.

Patient Perspectives

This submission did not include specific information on patient perspectives for this device; however, these conclusions were informed by activities (engagement with a Network of Experts and a Patient Panel Discussion) performed while the file was under review.

Benefit/Risk Conclusion

In conclusion, given the available information above, for the following indication statement:

Fibresolve is a software-only device that receives and analyzes lung computed tomography (CT) imaging data in order to provide a diagnostic subtype classification in suspected cases of interstitial lung disease (ILD). The device supplements the standard-of-care workflow by providing a qualitative, diagnostic classification output of imaging findings based on machine learning pattern recognition, in order to provide adjunctive information as part of a referral pathway to an appropriate Multidisciplinary Discussion (MDD) or as part of an MDD. Specifically, the tool is used to serve as an adjunct in the diagnosis of idiopathic pulmonary fibrosis (IPF) prior to invasive testing. The results of Fibresolve are intended to be used only by clinicians qualified in the care of lung disease, specifically in caring for patients with ILD, in conjunction with the patient's clinical history, symptoms, and other diagnostic tests, as well as the clinician's professional judgment.

The input to Fibresolve is a DICOM-compliant lung CT scan. Clinical case eligibility includes the following criteria:

Age > 22 years old.

Pulmonary symptoms suggestive of possible ILD including IPF.

The probable benefits outweigh the probable risks for the Fibresolve. 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 Fibresolve is granted and the device is classified under the following:

Product Code: QWO

Device Type: Radiology software for referral of findings related to fibrotic lung disease

Class: II

Regulation: 21 CFR 892.2085