NON-INTERVENTIONAL (NI) STUDY PROTOCOL

Study Information

Title	A retrospective chart validation study evaluating the performance of machine learning algorithm (ML) to predict the clinical diagnosis of wild-type transthyretin amyloid cardiomyopathy (ATTRwt-CM) and non-amyloid heart failure among patients with heart failure (HF)
Protocol number	B3461111
Protocol version identifier	1.0
Date	14 July 2023
Research question and objectives	The goal of this study is to evaluate the performance of a ATTRwt-CM ML algorithm to predict the clinical diagnosis of ATTRwt-CM among HF patients against chart-review-based ground truth with a prespecified performance thresholds in sensitivity and specificity.
Author	PPD MPH PPD PPD

1. TABLE OF CONTENTS

2. LIST OF ABBREVIATIONS

Abbreviation	Definition
AE	Adverse Event
AEM	Adverse Event Monitoring
AL	Light Chain Amyloid
ATTR	Transthyretin Amyloid
ATTR-CM	Transthyretin Amyloid Cardiomyopathy
ATTRwt-CM	Wild-type Transthyretin Amyloid Cardiomyopathy
CPT	Clinical Procedural Terminology, Fourth Edition
DrPH	Doctor of Public Health
EHR	Electronic Healthcare Records
FASA	Fellow of American Statistical Association
FDA	Food and Drug Administration
FN	False Negative
FP	False Positive
GPP	Guidelines for Good Pharmacoepidemiology Practices
GRACE	Good ReseArch for Comparative Effectiveness
hATTR-CM	Hereditary or Familiar Transthyretin Amyloid Cardiomyopathy
HCPCS	Healthcare Common Procedure Coding System
HF	Heart Failure

Abbreviation	Definition
HIPPA	Health Insurance Portability and Accountability Act
ICD-10	International Classification of Diseases, Tenth Revision
ICD-10-CM	International Classification of Diseases, Tenth Revision, Clinical Modification
ICD-10-PCS	International Classification of Diseases, Tenth Revision, Procedural Classification System
ICMJE	International Committee of Medical Journal Editors
IDN	Integrated Delivery Network
IEC	Independent Ethics Committee
IRB	Institutional Review Board
ISPE	International Society for Pharmacoepidemiology
ISPOR	International Society for Pharmacoeconomics and Outcomes Research
MA	Master of Arts
MBA	Master of Business Administration
MD	Doctor of Medicine
ML	Machine Learning
MPH	Master of Public Health
MS	Master of Science
NDC	National Drug Code
NIM	Non-inferiority Margin

Abbreviation	Definition
NIS	Non-interventional Study
NLP	Natural Language Processing
PharmD	Doctor of Pharmacy
PhD	Doctor of Philosophy
QC	Quality Control
RA	Rheumatoid Arthritis
RWE	Real World Evidence
TN	True Negative
TP	True Positive
TTR	Transthyretin
wtATTR-CM	Wild-Type Transthyretin Amyloid Cardiomyopathy
US	United States

3. RESPONSIBLE PARTIES

Principal Investigator of the Protocol

Name, degree(s)	Job Title	Affiliation	Address
Janice H. Chung, MPH	Director, Rare Disease RWE Scientist, Evidence Generation Platform RWE	Pfizer Inc.	66 Hudson Blvd East New York New York 10001

Collaborators

Name, degree(s)	Job Title	Affiliation	Address
PPD DrPH	PPD	Pfizer Inc.	66 Hudson Blvd East New York New York 10001
PPD PharmD, MBA		Pfizer Inc.	66 Hudson Blvd East New York New York 10001
PPD PhD		Pfizer Inc.	500 Arcola Rd Collegeville, PA 19426
PPD , PharmD, PhD		Pfizer Inc.	Avda Europa 20 B 28.108 Alcobendas. Madrid
PPD , MS, FASA		Pfizer Inc.	66 Hudson Blvd East New York New York 10001

Name, degree(s)	Job Title	Affiliation	Address
PPD MD, MS	PPD	Pfizer Inc.	66 Hudson Blvd East New York New York 10001

4. ABSTRACT

Title: A retrospective validation study evaluating the performance of machine learning (ML) algorithm to identify chart-based clinical diagnosis of wild-type transthyretin amyloid cardiomyopathy (ATTRwt-CM) and non-amyloid heart failure among patients with heart failure (HF)

Version 1.0, 27 July 2023

PPD MPH, Pfizer, Inc.

Rationale and background: To identify suspected patients with ATTRwt-CM, Pfizer developed and assessed a ML model using a large, nationally representative database with subsequent performance evaluations in multiple different claim- and EHR-based databases.[1,12-15] This model used patient's medical claims data to differentiate potential ATTRwt-CM patients from those who suffered from other causes of HF among HF patients 50 years of age or older. CCI [redacted]. The care provider can determine if the identified patient requires further evaluation and/or referral to other appropriate healthcare providers is needed. This tool may facilitate early and accurate diagnosis of ATTRwt-CM. CCI [redacted].

The ML model was developed using an operating definition of suspicious ATTRwt-CM patients with the presence of E85.82, an ICD-10-CM (International Classification of Diseases, Tenth Revision, Clinical Modification) diagnosis code. However, the ICD-10-CM code is an administrative code used for billing purpose. Therefore, an additional clinical study will be required to validate that the ML algorithm can accurately identify patients with clinically confirmed ATTRwt-CM among HF patients 50 years of age or older.

Study objective: This study aims to evaluate the performance of the ATTRwt-CM ML model to predict the clinical diagnosis of ATTRwt-CM among HF patients against chart-review-based ground truth with a prespecified performance thresholds in sensitivity and specificity.

Study design: This is an observational, retrospective non-inferiority study with a study sample from a large national EHR database integrated with administrative claims.

Population: This study will include HF patients ≥50 years old with clinical diagnosis of ATTRwt-CM or non-amyloid HF ascertained by charts. Patients will be required to have ≥12 months of continuous activity in the EHR or claims prior to the Index Date. There will be no minimal follow-up period for patients in this study.

Data source: Optum's Electronic Health Record and Integrated Claims-Clinical database that combines adjudicated claims data (where available) with Optum's EHR data, are the data sources for this study. In addition, this study will use data available in the unstructured for a subset of ATTRwt-CM and non-amyloid heart failure among patients with HF that have clinical patient notes and other data that cannot be mapped into the larger structured database. Optum's longitudinal repository is derived from dozens of healthcare provider organizations in the US, that include more than 57 contributing sources and 111K sites of care: treating more than 106 million patients receiving care in the US.

Sample size: Using various keyword search, feasibility assessment was performed. In the Optum EHR data, there were 422 potential ATTRwt-CM records available; yet, before reviewing the unstructured notes, one is unable to determine the total number of clinically diagnosed ATTRwt-CM patients. Given that ATTRwt-CM is a rare disease, the team will attempt to maximize the number of potential cases by performing note review for all 422 charts. Similar number of control charts will be reviewed to balance the number of cases and controls. We estimate that the sample sizes will likely provide 90% power to satisfy the non-inferiority hypotheses.

Data analysis: The ML model will identify patients with HF as suspected ATTRwt-CM and the remainder patients as not suspected for ATTRwt-CM/non-amyloid heart failure. The patients with HF will be entered in the ML algorithm along with their diagnosis codes listed in all their records. The model will classify patients based on the presence and absence of cardiac and non-cardiac features and categorized as suspected (at-risk) or not suspected with ATTRwt-CM. The model performance metrics (sensitivity and specificity) will be calculated based on the concordance between case ascertainment of ATTRwt-CM using diagnostic codes with ascertainment using clinical patient notes review, the latter using a variety of criteria to define ATTRwt-CM.

5. AMENDMENTS AND UPDATES

None.

6. MILESTONES

Milestone	Planned date
Start of data collection	September 1, 2023
End of data collection	October 31, 2023
Final study report	November 30, 2023

7. RATIONALE AND BACKGROUND

Transthyretin amyloid cardiomyopathy (ATTR-CM) is a rare condition that is an unrecognized cause of heart failure. The typical clinical presentation includes signs and symptoms of heart failure. The existing natural history data of the disease is poor, with ATTRv-CM and ATTRwt-CM having an estimated median survival of 2.5 years and 3.6 years after diagnosis respectively, and typically resulting from progressive heart failure or sudden death.[2-4] Because the symptoms are heterogeneous, including both cardiac and non-cardiac involvement, the diagnosis of ATTR-CM, including ATTRwt-CM, is often delayed, or missed.[2,5] Early diagnosis of ATTR-CM is key as progression worsens rapidly; however, diagnostic delay and misdiagnosis of ATTR-CM is common, with a reported mean time to diagnosis of 6.1 years in ATTRwt-CM.[6] Current diagnostic tools include speckle-tracking echocardiography, cardiac magnetic resonance imaging with contrast, bone scintigraphy, and/or blood tests for amyloidogenic light chain amyloid. Recent expert consensus recommendations and clinical "red flag" clues have been established to help raise suspicion of ATTR-CM and have been classified into cardiac features (i.e., heart failure, atrial fibrillation) and extracardiac features (i.e., history of carpal tunnel syndrome).[7-11] A screening tool that raises greater clinical awareness of these "red flags" may help raise suspicion and identification of the ATTRwt-CM.

In an effort to identify suspected patients with ATTRwt-CM, Pfizer developed and validated a random forest-based ML risk prediction model using a large, nationally representative database with subsequent validation and performance in multiple additional cohorts.[1,12-15]his model identified patient characteristics that could be used to differentiate ATTRwt-CM from other causes of HF among HF patients 50 years of age or older. This model showed an association between clinical features that were associated with ATTRwt-CM using US medical claims. CCI [redacted]

In developing the ML algorithm, the only clinical validation cohort confirmed the diagnosis of suspicious ATTRwt-CM patients identified by an algorithm using the presence of an ICD-10 (International Classification of Diseases, Tenth Revision) diagnosis code, E85.82. ICD-10 code is an administrative code used for billing purpose and it likely may not be readily used as a gold standard to confirm a clinical diagnosis. The presence of a diagnosis code may mean that the patient has a clinical diagnosis; however, it also often means that the clinician may screen the patient for the disease, which requires further clinical information to confirm the diagnosis. Given this, an additional clinical validation study is vital to ensure that the ML algorithm can predict the risk of clinically confirmed ATTRwt-CM among patients with HF.

8. RESEARCH QUESTION AND OBJECTIVES

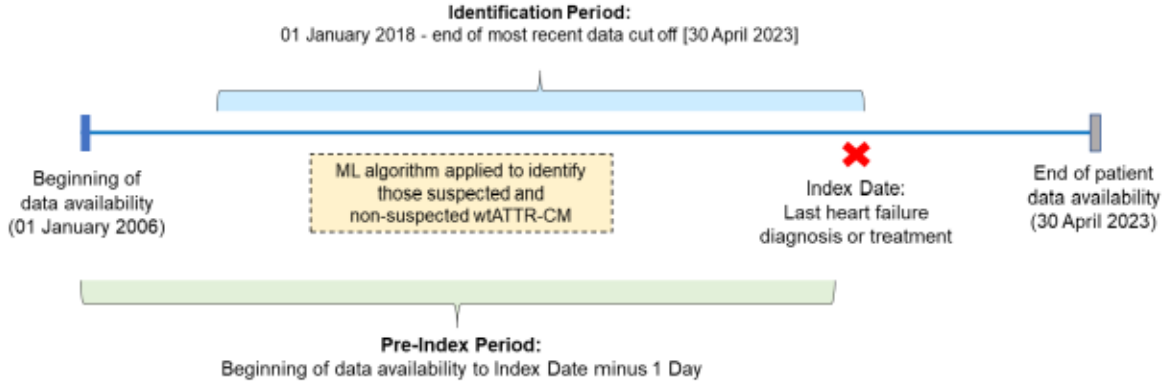
This study aims to further evaluate the performance of the ATTRwt-CM ML model to predict the clinical diagnosis of ATTRwt-CM among HF patients against chart-review-based ground truth with a prespecified performance thresholds in sensitivity and specificity.

9. RESEARCH METHODS

9.1. Study Design

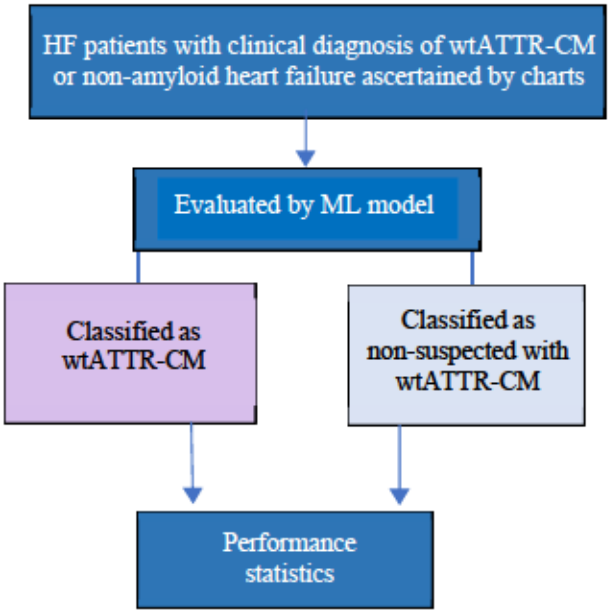
This is a retrospective cohort study. Patients with HF who meet inclusion and exclusion criteria will be identified from the Optum EHR database between 01 January 2018 and 30 April 2023. A subset of patients with clinical diagnosis of ATTRwt-CM or non-amyloid heart failure ascertained by charts will be included in this study (study population). Detailed methodology of clinical diagnosis ascertainment is found in Appendix A. Patients will be indexed at their latest heart failure diagnosis or treatment date and will be required to have ≥12 months of activity before the index date. There will be no minimal follow-up period for patients in this study.

Figure 1. Study Design



The ML algorithm will calculate the predicted probability of ATTRwt-CM for these heart failure patients based on the presence and absence of features (demographics, clinical diagnosis, etc.,) during the pre-index period. Detailed ML algorithm methodology can be found in Huda et al.[1] Based on a pre-specified predicted probability threshold of 0.5, we will further classify these patients as suspected ATTRwt-CM (predicted probability ≥0.5) and not suspected with ATTRwt-CM/non-amyloid heart failure (predicted probability <0.5). The model performance metrics (sensitivity and specificity) will be calculated by cross-tabulating diagnosis based on ML algorithm and diagnosis based on chart ascertainment.

Figure 2. Evaluating ML Model Performance



9.2. Setting

This study will use the US-based Optum's deidentified EHR data, with supplemental linkage to Optum's Integrated Claims-Clinical adjudicated claims dataset for those patients in the EHR who also have claims data available. For a subset of patients in the EHR, data exist including patient electronic clinical notes which will allow manual clinician review at Optum to confirm clinical diagnosis of ATTRwt-CM or non-amyloid heart failure based on available data that align with the Pfizer clinician specified specification. Optum's role, processes, and data is described in more detail in Section 9.4 and Section 9.6.

9.2.1. Inclusion Criteria

Primary Analysis

This study will include HF patients (defined as having ≥1 claim for HF or HF treatment) ≥50 years old with clinical diagnosis of ATTRwt-CM or non-amyloid HF ascertained by charts. Patients will be required to have ≥12 months of continuous activity in the EHR or claims prior to the Index Date.

Sensitivity analysis (pending feasibility assessment)

1. The earliest HF diagnosis or treatment date as the index date
2. Patients without HF diagnosis or treatment (pending feasibility assessment)
3. High Risk Population:

- Age 65+
- Male gender
- Presence of carpal tunnel syndrome

9.2.2. Exclusion Criteria

Patients with any of the following diagnoses:

- Light chain (AL) amyloidosis
- Intracranial hemorrhage
- Cerebral amyloid angiopathy
- End stage renal disease
- Blood cancer

9.3. Derived Variables

Variable	Role	Operational definition
HF	Index	• *Latest claim/encounter for HF or HF treatment (see Appendix B for the code list)*
ATTRwt-CM	Exposure	• ***Confirmed:*** *Ground truth ascertained by charts* • ***Suspected*** *ATTRwt-CM: Identified by the ML algorithm (predicted probability ≥ 0.5)* • ***Not Suspected with*** *ATTRwt-CM/non-amyloid heart failure*: Identified by the ML algorithm (predicted probability < 0.5)

9.4. Data Sources

Optum's EHR and Integrated Claims-Clinical dataset, which combines adjudicated claims data (where available) with Optum's EHR data, are the data sources for this study. Optum integrates EHR data with claims, prescribing, dispensing, and practice management data by partnering directly with several multi-specialty medical groups, IDNs and hospital chains to extract data from their EHR and various information technology systems in the US. By normalizing, validating, and aggregating the de-identified data, Optum generates a longitudinal view of patient care.

Optum's longitudinal EHR repository is derived from dozens of healthcare provider organizations in the US, that include more than 57 contributing sources and 111K sites of care: treating more than 106 million patients receiving care in the US. The data is certified as de-identified by an independent statistical expert following HIPAA statistical de-identification rules and managed according to Optum customer data use

agreement.[1,2]linical, claims and other medical administrative data is obtained from both Inpatient and Ambulatory EHRs, practice management systems and numerous other internal systems. Information is processed, normalized, and standardized across the continuum of care from both acute inpatient stays and outpatient visits. Optum data elements include demographics, medications prescribed and administered, immunizations, allergies, lab results (including microbiology), vital signs and other observable measurements, clinical and inpatient stay administrative data and coded diagnoses and procedures. In addition, Optum uses natural language processing (NLP) computing technology to transform critical facts from physician notes into usable datasets. The NLP data provides detailed information regarding signs and symptoms, family history, disease related scores (ie, RAPID3 for RA, or CHADS2 for stroke risk), genetic testing, medication changes, and physician rationale behind prescribing decisions that might never be recorded in the EHR.

Diagnosis data, laboratory data, and surgical procedure data for the study period of interest will be first be obtained from structured data (via International Classification of Diseases, Ninth Revision/Tenth Revision, Clinical Modification [ICD-9-CM/ICD-10-CM], International Classification of Diseases, Ninth Revision/Tenth Revision, Procedure Classification System [ICD-9- PCS/ICD-10-PCS], or Current Procedure Terminology (CPT) codes where applicable. Drug treatment data may also be pulled from prescription written, medication administration, and procedure tables when appropriate (via ICD-9-CM/ICD-10-CM, National Drug Center [NDC], CPT, and Healthcare Common Procedure Coding System [HCPCS] codes where applicable). In addition, Optum will use data available in the medical records for a subset of patients with heart failure that have clinical patient notes and other data that cannot be mapped into the larger structured database. These data may contain verbatim medical data, including text-based descriptions of medical information, such as medical records, physician notes, neurological scans, X-rays, or narrative fields in a database. When possible/appropriate, for validation of ATTRwt-CM or nonamyloid heart failure diagnoses and analyses of risk factors and background epidemiology, lab data will also be used to augment the structured electronic health data.

9.5. Study size

The sample size formula follows the binomial approach for binary outcomes where:

$$N^{0.5} = (Z_{1-\alpha} + Z_{1-\beta})\,(p(1-p))^{0.5}/NIM$$

For 90% power, the formula results in 78 individuals for sensitivity and 128 individuals for specificity. Using various keyword search, feasibility assessment was performed. In the Optum EHR data, after applying inclusion and exclusion criteria in Section 9.2.1 and Section 9.2.2 , there were 422 potential ATTRwt-CM records available; yet, before reviewing the unstructured notes, one is unable to determine the total number of clinically diagnosed ATTRwt-CM patients. Given that ATTRwt-CM is a rare disease for which the ICD-10 code has only been available recently and the use of which clinicians are still not quite familiar with or comfortable using, the team will attempt to maximize the number of

potential cases by performing note review for all 422 charts. Similar number of control charts will be reviewed to balance the number of cases and controls. Any increase in sample size will result in more precise 95% CIs.

We estimate that the sample sizes will likely provide 90% power to satisfy the non-inferiority hypotheses listed below.

The non-inferiority hypotheses are:

H0: Sensitivity ≤ 0.80

H1: Sensitivity ≥ 0.90

and

H0: Specificity ≤ 0.72

H1: Specificity ≥ 0.82

Sensitivity and Specificity estimates were evaluated based on the validation and performance metrics from the random forest-based ML risk prediction model receiver operating characteristics curve in Huda et al.[1] Sensitivity is defined as the probability that an individual with the disease in the population will screen positive for disease by the algorithm (True Positives). Specificity is defined as the probability that an individual without the disease will be correctly identified as not having the disease by the algorithm (True Negatives).

The distance from H0 to H1 is the non-inferiority margin (NIM), which is 0.10 for both sensitivity and specificity.

For sensitivity, we wish to reject H0 (inferiority to 0.90) in favor of H1 (non-inferiority to 0.90) while for specificity we wish to reject H0 (inferiority to 0.82) in favor of H1 (non-inferiority to 0.82).

9.6. Data Management

Two major data sources for this study are described in greater detail in various sections throughout the protocol. Their collection, retrieval, preparation, and storage are summarized as follows:

1. **Optum's EHR database:** The EHR Database is a longitudinally linked structured data source. It has been formally de-identified by an independent statistical expert following HIPAA statistical de-identification rules and managed according to Optum customer data use agreements.[1,2] The EHR includes structured fields rendered by NLP technology, wherein Optum data experts mine provider notes and then normalize, validate, and integrate them into the electronic database.

In addition to these structured data, Optum has clinical notes available from some EHR systems and can use technology to search the verbatim text for phrases of interest and extract a small portion of those notes for review and clinical assessment. All data elements from this source are stored on Optum's firewalled, password-protected database. These data can only be extracted by approved Optum study personnel using standard and commercially available software (eg, SAS, SQL, Python).

2. **Optum's Integrated Claims-Clinical database:** The EHR will be supplementally linked to patients in Optum's Integrated Claims-Clinical adjudicated claims database. For a subset of these patients, electronic clinical notes will undergo manual clinician review to confirm myocarditis cases.

For the validation process using the electronic clinical notes (which will be converted to deidentified structured data for analyses), the Optum Natural Language Programming (NLP) team will determine if clinical notes are available for all qualifying HF patients ≥50 years old with a clinical diagnosis of ATTRwt-CM or non-amyloid HF. Optum will perform a series of "enhanced search" queries on the patient notes to determine key term content. Next, the Optum NLP team will use the key terms to extract note snippets into a file for review by the Optum Clinical team. These notes snippets will then be used to validate ATTRwt-CM or non-amyloid HF per Pfizer clinical team specification that were identified in the EHR using the inclusion and exclusion criteria specified in Section 9.2.1 and Section 9.2.2. After clinical review of the notes involving two clinicians from Optum, and categorization of the content for analysis, the resulting table of criteria will be created. The ATTRwt-CM (cases) or non-amyloid HF (controls) will then be classified as validated or not validated, and relevant information entered into a spreadsheet that will serve as the data collection tool. Finally, the Optum clinical review team will provide these results of the case validation to the Pfizer Core Study Team so that the ML algorithm could be run using the structured data to calculate the sensitivity and specificity of the algorithm.

Data will be handled and analyzed using Python software, Version 3.6 or later.

9.7. Data Analysis

9.7.1. Statistical Methods

The ML model will identify patients with HF as suspected ATTRwt-CM and the remainder patients as not suspected for ATTRwt-CM/non-amyloid HF. The patients with HF will be entered in the ML algorithm along with their diagnosis codes listed in all their records. The model will classify patients based on the presence and absence of cardiomyopathy features and categorized as suspected (at-risk) or not suspected with ATTRwt-CM. The model performance metrics (sensitivity and specificity) will be calculated by cross-tabulating between diagnosis based on ML algorithm and diagnosis based on chart ascertainment.

The analyses for noninferiority of sensitivity and specificity are straightforward. The calculations for these measures are shown below.

Table 1. Sensitivity and Specificity

ML classification	Cohort (Ascertained by Charts)		
	ATTRwt-CM	Non-amyloid HF	
+	True Positive (TP)	False Positive (FP)	TP+FP
-	False Negative (FN)	True Negative (TN)	TN+FN
	TP+FN	TN+FP	

Sensitivity = TP / (TP + FN)

Specificity = TN / (TN + FP)

We would then calculate the one-sided 95% confidence interval (CI) for sensitivity, assuming 90% power:

$$95\%\ CI = sensitivity - 1.645(SE_{sensitivity})$$

where $SE_{sensitivity}$ = square root $[sensitivity - (1\text{-}sensitivity)]/n_{sensitivity}$)

The 95% CI for sensitivity must equal or exceed a cut point (CP)value between 0.80 and 0.90 where:

CP =0.90 - NIM Z1- β / (Z1-α + Z1-β)

When $\alpha = 0.05$ and power = 0.90 we reject inferiority for all values $\geq$ CP = 0.8561.

Similarly, for specificity, we reject inferiority for all values $\geq$ CP = 0.7761.

9.8. Quality Control

All codes and/or code algorithms will be reviewed by the clinical lead from Pfizer, as well as the protocol authors. The study team will be responsible for upholding high levels of quality control throughout the data analysis and reporting processes. Quality control (QC) checks will be performed on all Python programming, data tables, and reports generated in the course of this research. QC findings and documentation of remedial action will be maintained. Storage of programming, data, and reports will be carried out per standard procedures.

In the validation study, after the Optum NLP team creates key term tags for use in the snippet review for the two Optum clinicians who will review patient notes with these key terms mentioned, the two clinicians will record their findings from the notes regarding whether the notes support a ATTRwt-CM or non-amyloid HF. These results will be provided to the Pfizer study team for review and any clarifications or questions will be managed between the two groups until resolved.

9.9. Limitations of the Research Methods

There are some important limitations to consider for this study:

- Patients may have received health care outside of the network of providers and healthcare organizations that contribute to the Optum databases, or prior to having the index diagnosis and we may not be able to exclude the possibility that a patient was diagnosed with HF, as well as other risk factors, prior to entry into a healthcare system where Optum has EHR access.

- This study utilizes EHR data that have inherent limitations such as the possibility for entry errors and missing information which could lead to extreme or incorrect values. As a result, there is a potential for bias due to a lack of quality control and consistency for data collection and reporting within the database. Study samples obtained from EHR data analyses are not randomly selected and unobserved factors may introduce bias in subject selection.

9.10. Other Aspects

Not applicable.

10. PROTECTION OF HUMAN SUBJECTS

All parties will ensure the protection of patient personal data and will not include patient names on any sponsor forms, reports, publications, or in any other disclosures, except where required by laws. In case of data transfer, Pfizer will maintain high standards of confidentiality and protection of patient personal data.

10.1. Patient Information

This study involves data that exist in anonymized structured format and contain no patient personal information.

10.2. Patient Consent

As this study involves anonymized structured data, which according to applicable legal requirements do not contain data subject to privacy laws, obtaining informed consent from patients by Pfizer is not required.

10.3. Institutional Review Board (IRB)/Independent Ethics Committee (IEC)

There is no requirement for review since this study only uses commercially available de-identified secondary data sources and is considered exempt from the requirements for "human subjects research" in the US.

10.4. Ethical Conduct of the Study

The study will be conducted in accordance with legal and regulatory requirements, as well as with scientific purpose, value and rigor and follow generally accepted research practices described in in Guidelines for Good Pharmacoepidemiology Practices (GPP) issued by the International Society for Pharmacoepidemiology (ISPE), Good Practices for Outcomes Research issued by the International Society for Pharmacoeconomics and Outcomes Research (ISPOR), and the GRACE (Good ReseArch for Comparative Effectiveness) Principles.

11. MANAGEMENT AND REPORTING OF ADVERSE EVENTS/ADVERSE REACTIONS

11.1. Structured Data Analysis

This study involves data that exist as structured data by the time of study start. In these data sources, it is not possible to link (ie, identify a potential association between) a particular product and medical event for any individual. Thus, the minimum criteria for reporting an AE (ie, identifiable patient, identifiable reporter, a suspect product, and event) cannot be met.

12. PLANS FOR DISSEMINATING AND COMMUNICATING STUDY RESULTS

One or more abstracts may be developed and submitted to relevant scientific conference(s) and one or more manuscripts may be developed and submitted to relevant peer-reviewed medical journals. Authorship will follow the guidelines proposed by the International Committee of Medical Journal Editors (ICMJE; www.icmje.org). All authors should meet the criteria for authorship, and all people who meet the criteria should be authors. Any potential conflicts of interest will be disclosed.

13. REFERENCES

1. Huda A, Castaño A, Niyogi A, et al. A machine learning model for identifying patients at risk for wild-type transthyretin amyloid cardiomyopathy. *Nat Commun*. 2021;12(1):2725. doi:10.1038/s41467-021-22876-9.

2. Maurer MS, Bokhari S, Damy T, et al. Expert consensus recommendations for the suspicion and diagnosis of transthyretin cardiac amyloidosis. Circ Heart Fail. 2019;12(9):e006075.

3. Grogan M, Scott CG, Kyle RA, et al. Natural history of wild-type transthyretin cardiac amyloidosis and risk stratification using a novel staging system. J Am Coll Cardiol. 2016;68(10):1014-1020. doi:10.1016/j.jacc.2016.06.033

4. Gillmore JD, Damy T, FontanaM, et al. A new staging system for cardiac transthyretin amyloidosis. Eur Heart J. 2018;39(30):2799-2806. doi:10.1093/eurheartj/ehx589

5. Dorbala S, Ando Y, Bokhari S, et al. ASNC/AHA/ASE/EANM/HFSA/ISA/SCMR/SNMMI expert consensus recommendations for multimodality imaging in cardiac amyloidosis: part 2 of 2-diagnostic criteria and appropriate utilization. J Nucl Cardiol. 2020;27(2):659–673.

6. Rozenbaum MH, Large S, Bhambri R, Stewart M, Whelan J, van Doornewaard A, Dasgupta N, Masri A, Nativi-Nicolau J. Impact of Delayed Diagnosis and Misdiagnosis for Patients with Transthyretin Amyloid Cardiomyopathy (ATTR-CM): A Targeted Literature Review. Cardiol Ther. 2021 Jun;10(1):141-159.

7. Kittleson MM, Ruberg FL, Ambardekar AV, Brannagan TH, Cheng RK, Clarke JO, Dember LM, Frantz JG, Hershberger RE, Maurer MS, Nativi-Nicolau J, Sanchorawala V, Sheikh FH. 2023 ACC Expert Consensus Decision Pathway on Comprehensive Multidisciplinary Care for the Patient With Cardiac Amyloidosis: A Report of the American College of Cardiology Solution Set Oversight Committee. J Am Coll Cardiol. 2023 Mar 21;81(11):1076-1126.

8. Witteles RM, Bokhari S, Damy T, et al. Screening for transthyretin amyloid cardiomyopathy in everyday practice. JACC Heart Fail. 2019;7(8):709–716.

9. Nativi-Nicolau JN, Karam C, Khella S, Maurer MS. Screening for ATTR amyloidosis in the clinic: overlapping disorders, misdiagnosis, and multiorgan awareness. Heart Fail Rev 2022;27(3):785-793. https://doi.org/10.1007/s10741-021-10080-2.

10. Maurer MS, Elliott P, Comenzo R, Semigran M, Rapezzi C. Addressing common questions encountered in the diagnosis and management of cardiac amyloidosis. Circulation 2017;135(14):1357-1377.

11. Rapezzi C, Lorenzini M, Longhi S, et al. Cardiac amyloidosis: the great pretender. Heart Fail Rev 2015;20(2):117-124. https://doi.org/10.1007/s10741-015-9480-0.

12. Bruno M, Sheer R, Reed C, Schepart A, Radhika N, and Simmons, JD. Clinical characteristics and health care resource use of patients at risk for wild-type transthyretin amyloid cardiomyopathy identified by machine learning model. Journal of Managed Care and Specialty Pharmacy. 2023;29(5):530-540. doi: 10.18553/jmcp.2023.29.5.530.

13. Heitner S, Elman M, Masri A, Aphinyanaphongs Y, Reyentovich A, Ateya M, Emir B, Fowler R, Mills JR, Nolen KD, Sohn A, Huda A, Castano A, Bruno M. Performance evaluation of a machine learning model for systematic identification of wild-type transthyretin amyloid cardiomyopathy at two academic medical centers. Heart Failure Society of America - 24th Annual Scientific Meeting. 2020.

14. Heitner S, Elman M, Masri A, Aphinyanaphongs Y, Reyentovich A, Ateya M, Emir B, Fowler R, Mills JR, Nolen KD, Sohn A, Huda A, Castano A, Bruno M. Performance evaluation of a machine learning model for systematic identification of wild-type transthyretin amyloid cardiomyopathy at two academic medical centers. Heart Failure Society of America - 24th Annual Scientific Meeting. 2020.

15. Heitner S, Elman M, Masri A, Aphinyanaphongs Y, Reyentovich A, Ateya M, Emir B, Fowler R, Mills JR, Nolen KD, Sohn A, Huda A, Castano A, Bruno M. Performance evaluation of a machine learning model for systematic identification of wild-type transthyretin amyloid cardiomyopathy at two academic medical centers. Heart Failure Society of America - 24th Annual Scientific Meeting. 2020.

16. Real-World Data: Assessing Electronic Health Records and Medical Claims Data To Support Regulatory Decision Making for Drug and Biological Products Guidance for Industry.

17. Guidance for Industry and FDA Staff Best Practices for Conducting and Reporting Pharmacoepidemiologic Safety Studies Using Electronic Healthcare Data.

14. LIST OF TABLES

15. LIST OF FIGURES

16. ANNEX 1. LIST OF STAND-ALONE DOCUMENTS

Appendix A. Chart Review Specification

Double click on the icon to view the chart specification:

Chart Review Specification

Appendix B. Code Lists Used in Identification of Case and Controls

Double Click on the icon to view the code lists:

Code List for Case and Control Identifica

17. ANNEX 2. ADDITIONAL INFORMATION

Not Applicable.

Document Approval Record

Document Name:	B3461111 Protocol v1.0
Document Title:	B3461111 Protocol v1.0

Signed By:	Date(GMT)	Signing Capacity
PPD	30-Aug-2023 20:44:20	Author Approval