Cardiovascular risk assessment using ASCVD risk score in fibromyalgia: a single-centre, retrospective study using “traditional” case control methodology and “novel” machine learning

Sandeep Surendran1, C. B. Mithun1, Merlin Moni2, Arun Tiwari1 and Manu Pradeep3*

Abstract

Background: In autoimmune inflammatory rheumatological diseases, routine cardiovascular risk assessment is becoming more important. As an increased cardiovascular disease (CVD) risk is recognized in patients with fibromyalgia (FM), a combination of traditional CVD risk assessment tool with Machine Learning (ML) predictive model could help to identify non-traditional CVD risk factors.

Methods: This study was a retrospective case–control study conducted at a quaternary care center in India. Female patients diagnosed with FM as per 2016 modified American College of Rheumatology 2010/2011 diagnostic criteria were enrolled; healthy age and gender-matched controls were obtained from Non-communicable disease Initiatives and Research at AMrita (NIRAM) study database. Firstly, FM cases and healthy controls were age-stratified into three categories of 18–39 years, 40–59 years, and ≥ 60 years. A 10 year and lifetime CVD risk was calculated in both cases and controls using the ASCVD calculator. Pearson chi-square test and Fisher's exact were used to compare the ASCVD risk scores of FM patients and controls across the age categories. Secondly, ML predictive models of CVD risk in FM patients were developed. A random forest algorithm was used to develop the predictive models with ASCVD 10 years and lifetime risk as target measures. Model predictive accuracy of the ML models was assessed by accuracy, f1-score, and Area Under 'receiver operating Curve' (AUC). From the final predictive models, we assessed risk factors that had the highest weightage for CVD risk in FM.

Results: A total of 139 FM cases and 1820 controls were enrolled in the study. FM patients in the age group 40–59 years had increased lifetime CVD risk compared to the control group (OR = 1.56, p = 0.043). However, CVD risk was not associated with FM disease severity and disease duration as per the conventional statistical analysis. ML model for 10-year ASCVD risk had an accuracy of 95% with an f1-score of 0.67 and AUC of 0.825. ML model for the lifetime ASCVD risk had an accuracy of 72% with an f1-score of 0.79 and AUC of 0.713. In addition to the traditional risk factors for CVD, FM disease severity parameters were important contributors in the ML predictive models.

© The Author(s) 2021. Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.
Introduction
Fibromyalgia (FM) is a debilitating condition that presents with generalized widespread pain and is often associated with sleep disturbances, fatigue, and cognitive dysfunction [1]. The prevalence rates of FM range from 1 to 3% in the general population [2].

While there is no singular pathological process that explains the development of FM; three key mechanisms include increased central nervous system response to peripheral pain stimulation via amplification of signaling, abnormal ascending pain pathways, and small fiber neuropathy [3, 4]. It is well known that other diseases with altered neuronal pathways like depression are more likely to develop cardiovascular disease (CVD). Major depression is more prevalent in people who have suffered an acute myocardial infarction, with up to 18–20% of patients meeting the criteria for depression patients using structured interviews [5]. However, this is not just a cause-effect relationship where the occurrence of a major cardiac event leads to the secondary development of depression. Patients with depression are more likely to have a cardiac event across the world. A cohort study in the South-Asian cohort population had reported that patients with depression were at an increased risk for developing CVD by 41% for men and 48% for women [6]. Western populations also showed similar results where a combination of depressive symptoms and stress in low-income persons was associated with an increased risk of incident CVD [7].

While depression and CVD are well studied, the inter-relationship between FM and CVD is less described in the literature. A causal association is lent weight to by findings that FM is associated with dysfunction of the autonomic nervous system and the inflammatory system; both of which play a part in the pathophysiology of CVD [8].

A meta-analysis of the key studies on cytokines in FM has shown elevated interleukin-6 (IL-6) levels in FM patients, which is similar to a state of chronic mild inflammation [9]. IL-6 has a downstream knockdown effect on lipid metabolism and anti-inflammatory therapy with canakinumab (human anti-IL-1β monoclonal antibody) has been studied in a randomized controlled trial for CVD prevention [10, 11]. Thus, it is also possible that the neuro-immunological changes linked with FM may itself be a risk factor for CVD. CVD risk assessment in a Taiwanese cohort has shown findings that are suggestive of the same [12]. To the best of our knowledge, regional data regarding this is lacking from India.

Non-invasive CVD risk assessment tools are well described and validated in many cohorts. The most recent of these scales put forward by the American Heart Association (AHA) and American College of Cardiology is the Atherosclerotic Cardiovascular (ASCVD) risk assessment tool. Unlike the more traditionally used Framingham Risk Score, the ASCVD risk score can be applied to other races also and predicts both the 10-year and lifetime ASCVD risk of an individual [13]. This tool—ASCVD risk calculator, helps to predict future risk of CVD and is in fact, currently recommended to be used as a guide in the initiation of statin therapy [13]. As there is a paucity of studies assessing CVD risk in FM patients, this study has also adopted this tool to quantify the CVD risk.

As FM is a disease that is more common in females, an increased CVD risk in the affected females would be a relevant observation [14]. The identification of any elevated CVD risk as compared to the general population would help in stratifying such patients for interventions and primary prevention strategies.

Machine learning (ML) is the process by which the large computing capacity of modern computers is applied to analyse the varied statistical relations within a data set for an outcome prediction [15]. In the recent years, many studies have shown ability of ML to predict accurate predictions in medicine [16].

There are two main techniques of machine learning. The first is “supervised” learning when the goal is predicting a known output or target using classified/sorted data. The other type is called unsupervised learning where there are no outputs to predict; rather the aim is to find naturally occurring patterns or groupings within the unsorted data using massive statistical computing power. Both types have found applications in medicine with supervised machine learning methods being the more commonly used currently. This is largely because it is easier to create datasets that allow communication between researching physicians and data analysts. One of the best-known ML projects launched by the Microsoft

Conclusion: FM patients of the 40–59 years age group had increased lifetime CVD risk in our study. Although FM disease severity was not associated with high CVD risk as per the conventional statistical analysis of the data, it was among the highest contributor to ML predictive model for CVD risk in FM patients. This also highlights that ML can potentially help to bridge the gap of non-linear risk factor identification.

Keywords: Fibromyalgia, Heart disease risk factors, Machine learning, Cardiovascular diseases
Corporation (USA) is Project Hanover which uses supervised ML and aims to help in drug research in oncology [17].

Machine learning (ML) tools are increasingly used to understand the complex interplay of the various factors associated with diseases. Recently ML tools were used to develop an algorithm to aid diagnosis of systemic lupus erythematosus by Adamichou et al. [18].

The primary objective of our study was to assess the CVD risk in female FM patients using the ASCVD risk score calculator and compare it with that of the general population.

A secondary objective of our study was to develop a supervised machine learning (ML) algorithm for studying the interactions between characteristics of the FM patients and ASCVD risk categories.

Methods

Study design (Fig. 1)

This was a retrospective case–control study done in a setting of Rheumatology outpatient department (OPD) of a quaternary care center in the southern Indian state of Kerala. Electronic medical records of FM patients were accessed after the study was approved by the institutional ethics committee.

We included the female FM patients diagnosed as per 2016 modified 2010/2011 ACR criteria, who had documented parameters needed for risk score calculation with ASCVD risk calculator (done within 6 months of diagnosis), baseline average pain score on a visual analog scale more than 5, and visited rheumatology OPD from April 2017 to April 2020 [1]. We excluded those FM patients who had any associated autoimmune disease, primary psychiatric diagnosis, breastfeeding, and patients taking any concomitant medications which could affect the fasting lipid profile.

Healthy control participants data was obtained after age and gender matching from the datasets of a population-based cross-sectional study conducted at the same quaternary center by Menon et al.- Non-communicable disease Initiatives and Research at AMrita (NIRAM) study [19].

The ASCVD risk score was calculated in both FM patients and controls. The details collected for ASCVD risk score calculation included age (in years), height (in meters), weight (in kilograms), and co-morbid conditions like hypertension, diabetes mellitus, and dyslipidemia. The FM disease severity was assessed using Fibromyalgia Impact Questionnaire-Revised (FIQR), the average pain score on the Brief Pain Inventory (BPI) form. Patients with a FIQR score of more than 60 at baseline were considered to have severe disease. The fasting lipid profile was measured using AU2700 plus Beckman Coulter (USA).

Based on the age, the study participants were categorized into three age group strata of 18–39 years, 40–59 years, and 60 years or older. The 10-year and lifetime ASCVD cardiovascular risk score was calculated for each study participant in the age group of 41–60 years. For participants less than 40 years of age, only lifetime risk could be calculated whereas for the participants of more than 60 years of age, only a 10-year risk could be calculated. The ASCVD lifetime and 10-year risk scores were further categorized into low risk and high risk based on the standard cut-off. The 10-year CVD risk of less than 7.5% was considered as low while 7.5% or more was considered high. Lifetime CVD risk less than 39% was considered low, 39% or more was considered high. The FM patients were also stratified based on disease severity and disease duration to study the association of these parameters with ASCVD risk scores.

To address the secondary objective of the study, a supervised ML algorithm was developed to study the interactions between FM disease characteristics and ASCVD risk scores. A random forest classifier ML algorithm was used for the development and training of the model to help predict the ASCVD risk category of FM patients using baseline clinical features and biochemical parameters. Baseline characteristics of higher importance for high CVD risk in FM patients were identified using this ML algorithm.

Statistical methods

We used SPSS Version 20.0 (IBM Corporation, USA) for statistical analysis. All continuous variables were summarized using mean (SD) or median (IQR). Kolmogorov–Smirnov test was used to ascertain whether the data were normally distributed. Categorical variables were expressed in counts (%). We used the Pearson Chi-Square test or Fisher’s exact test for categorical variables and independent sample t-test or Mann–Whitney test for continuous variables. Anaconda-Python 3.0 (USA) software was used for ML analysis.

Machine learning (ML) model preparation and evaluation of the model

The 10 years and Lifetime ASCVD risk status were the target variables (high risk versus low risk) that were to be predicted in their respective models. The inputs were all other variables mentioned above and cumulatively, these included 21 parameters for each patient—age, height, weight, body mass index (BMI), presence of co-morbidities (hypothyroidism, diabetes mellitus, hypertension, and dyslipidemia), systolic-diastolic blood pressures, smoking status, BPI (total and average) score, FIQR
Cases: Review of electronic medical records of FM patients from April 2017- April 2020 (n=275)

Inclusion Criteria:
1. Female patients satisfying 2016 modification of ACR 2010/2011 fibromyalgia criteria
2. Baseline VAS pain score > 5

Exclusion criteria:
1. Associated autoimmune disease or primary psychiatric diagnosis
2. Fasting lipid profile details not available within six months of diagnosis.
3. Breastfeeding
4. Concomitant medicines use which could affect the fasting lipid profile

FM cases enrolled to study (n=139)

Controls: NIRAM study database (n=4507)
Population-based cross-sectional study studying CVD risk in a general population at the same centre

Age and gender matched healthy controls selected

Healthy controls enrolled to study (n = 1820)

ASCVD risk score calculation and categorization into high and low CVD risk

Comparison of ASCVD risk scores between FM cases and controls across age categories of 18-39 years, 40-59 years, and more than 60

Input variables: Age, height, weight, BMI, blood pressure, smoking status, FIQR score, BPI score, ESR, CRP, and lipid profile

10-year ASCVD risk model
Lifetime ASCVD risk model

Extraction of feature importance
(To identify baseline characteristics of higher importance for high CVD risk)

Fig. 1 Study design and methodology

Abbreviations:
FM - fibromyalgia, NIRAM - Non-communicable disease Initiatives and Research at Amrita, CVD - Cardiovascular disease, ACR - American college of rheumatology, VAS - Visual analogue scale, ASCVD - Atherosclerotic cardiovascular disease, ML - Machine learning, BMI - Body mass index, FIQR - Revised fibromyalgia impact questionnaire, BPI - Brief pain inventory, ESR - Erythrocyte sedimentation rate, and CRP - C-reactive protein.
score, ESR, CRP levels and lipid profile (total cholesterol, LDL, HDL levels and triglyceride). To apply the ML risk algorithms, the data was imported using Python 3.0 using JupyterLab coding environment. For the ML modeling, the Scikit-learn python library, which has been specifically developed for data science and ML, was used. The FM patients were split into the data set into training and validation subgroups. The training subgroup was derived from a random sampling of 80% of the FM patients, and the ‘validation’ cohort comprised the remaining 20% which was imported into a random forest classifier algorithm to help derive an ML model for the 10-year CVD risk and lifetime CVD risk, respectively. Then validation subgroup data was used by these trained algorithms to test the performance of the model. Each model’s hyper-parameters were determined manually with the n-estimators and using fivefold cross-validation on the training subgroup to determine the values which led to the best performance. The model with the best performance was then represented using accuracy score, f-score, area under the curve (A.U.C) for both ML models. In both the 10-year risk random classifier model and lifetime risk random classifier model in addition to model training, the feature importance and their weightage were calculated.

### Table 1 Baseline characteristics of study participants

|                          | FM cases (n = 139) | Healthy controls (n = 1820) |
|---------------------------|--------------------|-----------------------------|
| **A: Baseline characteristics of FM cases and healthy control group** |                     |                             |
| Age in years (mean ± SD)  | 46.85 ± 10.15      | 48.94 ± 10.7                |
| BMI kg/m² (mean ± SD)     | 27.34 ± 4.58       | 26.97 ± 3.88                |
| Type 2 diabetes mellitus  | 19 (13.66%)        | 397 (21.81%)                |
| Hypertension, no. (%)     | 30 (21.58%)        | 397 (21.81%)                |
| Dyslipidaemia, no. (%)    | 89 (64.49%)        | 310 (39.40%)                |
| **B: Clinical characteristics of FM patients** |                     |                             |
| Disease duration (in years) (mean ± SD) | 5.04 ± 4.51 |                      |
| Baseline FIQR score (IQR) | 55.33 (26.32)      |                             |
| Baseline BPI average pain (IQR) | 6.66 (3)    |                             |
| CRP (mg/L) (mean ± SD)    | 4.44 ± 4.39       |                             |
| ESR (mm/h) (mean ± SD)    | 23.31 ± 12.50     |                             |

FM, fibromyalgia; SD, standard deviation; BMI, body mass index; FIQR, revised fibromyalgia impact questionnaire; IQR, interquartile range; CRP, C-reactive protein; ESR, erythrocyte sedimentation rate

### Table 2 Age stratified distribution of ASCVD risk scores in FM cases and healthy controls

| Age (in years) | ASCVD risk score† | Lifetime risk |
|----------------|-------------------|---------------|
|                | 10-year risk      |               |
|                | <7.5% (low)       | ≥7.5% (high)  | OR (CI) | p-value | <39% (low) | ≥39% (high) | OR (CI) | p-value |
| 18–39 years    |                   |               |         |         |     |         |         |         |         |
| FM cases (n = 29) | –                | –              | –      | 19      | 10  | 0.649   | 0.267   |
| Healthy controls (n = 428) | –             | –              | –      | 235     | 193 | (0.29–1.41) |         |         |
| 40–59 years    |                   |               |         |         |     |         |         |         |         |
| FM cases (n = 97) | 92               | 5              | 2.27   | 0.093   | 62  | 35      | 1.56    | 0.044   |
| Healthy Control (n = 1071) | 1046          | 25             | (0.85–6.08) | 787 | 284 | (1.01–2.42) |         |         |
| ≥ 60 years     |                   |               |         |         |     |         |         |         |         |
| FM cases (n = 13) | 8                | 5              | 0.56   | 0.307   | –   | –       | –       |         |
| Healthy control (n = 321) | 151            | 170            | (0.18–1.73) | –   | –   | –       |         |         |

FM, fibromyalgia; ASCVD, Atherosclerotic cardiovascular disease; OR, odd’s ratio; CI, 95% confidence interval; Emboldened summary estimate is statistically significant at p < 0.05.

† 10-year ASCVD risk calculated only for patients 40 years and more; Lifetime ASCVD risk calculated for patients who are less than 60 years of age
Results
A total of 139 FM cases and 1820 age and gender-matched healthy controls were enrolled in the study. Table 1 shows the baseline characteristics of study participants.

Table 2 shows the age-stratified distribution of ASCVD risk scores between the FM cases and healthy controls. The lifetime risk of CVD is significantly higher in the FM cases of age group 40–59 years than the age and gender-matched healthy controls ($OR = 1.56$, $95\% CI = 1.01–2.42$, $p = 0.043$).

Table 3 shows ASCVD risk scores of FM patients stratified for FM disease severity and disease duration. There is no difference noted in CVD risk with FM disease severity and duration ($p > 0.05$).

Machine learning (ML) analysis and features importance
Out of a total of 139 FM patients enrolled, 110 patients were of more than 40 years of age whose 10-year ASCVD risk was calculated. An ML model was derived using the SciKit Random Classifier algorithm. The best model was selected after hyperparameter tuning (manually with the n-estimators, and using fivefold cross-validation) and the precision of 96%, recall of 95% with an F1 score of 0.95 was achieved. The AUC of this 10-year ASCVD ML model was 0.825 and is shown in Fig. 2A.

Table 3 ASCVD risk scores stratified for FM disease severity and duration

| ASCVD risk score | 10-year risk | Lifetime risk |
|-----------------|--------------|---------------|
|                 | Low | High | $p$-value† | Low | High | $p$-value† |
| FM disease severity§ |    |      |          |    |      |          |
| Severe          | 49  | 5    | 0.564    | 20  | 34   | 0.135    |
| Mild/moderate   | 80  | 5    |          | 40  | 45   |          |
| FM disease duration |    |      |          |    |      |          |
| > 2 years       | 88  | 7    | 0.914    | 39  | 56   | 0.546    |
| < 2 years       | 41  | 3    |          | 21  | 23   |          |

ASCVD, atherosclerotic cardiovascular disease; FM, fibromyalgia

† $p$-value was calculated using Pearson Chi-square and Fisher’s exact test, a value of $< 0.05$ was considered significant
§ FM disease severity defined by revised fibromyalgia impact questionnaire (FIQUR); FIQUR score $> 60$ was considered severe FM and FIQUR score $\leq 60$ considered mild-moderate FM
A lifetime ASCVD risk calculated for 126 FM patients of less than 60 years age. Using the SciKit Random Classifier algorithm again, an ML model was derived. The best model developed has a precision of 72% recall of 72% with an F1 score of 0.79 was achieved. The AUC of the lifetime ASCVD ML model was 0.713 and is shown in Fig. 2B. The feature importance in the individual models are shown in Table 4.

**Discussion**

In our study, the conventional case–control design showed that female FM patients of age group 40–59 years had increased lifetime ASCVD risk. Nevertheless, ASCVD risk was not increased in female FM patients of the age groups of 18–39 years and more than 60 years. Also, this conventional study design failed to show any association of ASCVD risk with FM disease severity and disease duration.

These results are in concordance with the study done in Taiwan which also shows that FM is a risk factor for CVD [12]. This finding of increased CVD risk in patients with FM is homologous to CVD risk in another related condition—major depressive disorder [20]. Depression has been demonstrated to be an independent risk factor for the onset of a wide range of cardiovascular disorders in a meta-analysis [21].

As FM pathogenesis is multifaceted and we are continuing to learn with every research, evidence to support the role of inflammation in FM etiopathogenesis is increasing. Pro-inflammatory cytokines, reactive oxygen species and inflammatory factors are considered to play a distinct role in inflammatory response in FM patients [22]. Serum CRP levels in FM patients are reported to be higher than control population which is known to improve by exercise and dietary changes [23]. This low-grade inflammation could be the common link between FM and CVD.

The CVD risk is multi-factorial but amongst the parameters used in the ASCVD risk calculator, dyslipidemia is strikingly high in fibromyalgia patients in comparison to the control population in this study (Table 1A). This is finding is consistent with various studies across the globe [24, 25]. A low-grade subclinical inflammatory process with an elevation of inflammatory cytokine (interleukin-6) is well described in fibromyalgia patients [26]. The role of IL-6 in lipid metabolism is well described and this could be the common mechanism that links the two.

**Table 4** Feature importance of ML predictive models of 10 years and lifetime ASCVD risk (random forest algorithm)

| Feature importance in 10-year ASCVD risk ML predictive model | Feature importance in lifetime ASCVD risk ML predictive model |
|---------------------------------------------------------------|---------------------------------------------------------------|
| Age 0.150 | 0.074 |
| Height 0.063 | 0.051 |
| Weight 0.078 | 0.035 |
| BMI 0.065 | 0.039 |
| Systolic blood pressure 0.043 | 0.069 |
| Diastolic blood pressure 0.030 | 0.037 |
| Diabetes mellitus 0.041 | 0.007 |
| Hypertension 0.006 | 0.015 |
| Dyslipidaemia 0.012 | 0.009 |
| Hypothyroidism 0.023 | 0.013 |
| Smoker 0.000 | 0.000 |
| Disease duration 0.036 | 0.043 |
| FIQR score 0.071 | 0.062 |
| BPI total score 0.038 | 0.029 |
| BPI average score 0.066 | 0.047 |
| ESR 0.036 | 0.044 |
| CRP 0.040 | 0.050 |
| Serum total cholesterol 0.077 | 0.179 |
| Serum LDL 0.041 | 0.083 |
| Serum HDL 0.022 | 0.051 |
| Serum triglycerides 0.070 | 0.055 |

Feature importance is the degree of contribution in the random forest ML predictive model. Seven parameters of high feature importance are emboldened in each ML classifier model.

ML, Machine learning; ASCVD, atherosclerotic cardiovascular disease; BMI, body mass index; FIQR, revised fibromyalgia impact questionnaire; BPI, brief pain inventory; ESR, erythrocyte sedimentation rate; CRP, C-reactive protein; LDL, low-density lipoprotein; HDL, high-density lipoprotein.
diseases explaining our observation [10]. FM patients have lower physical fitness that could be another reason for increased CVD risk as suggested in the al-Ándalus Project [27].

Since the conventional statistical analysis showed a significantly high ASCVD risk in FM patients of age 40–59 years but failed to show any association of ASCVD risk with FM disease severity and disease duration; ML algorithms were developed and trained to understand the non-linear relationships of CVD risk in FM patients. The application of ML models showed some interesting relationships with all the possible risk factors assessed, which were overlooked by the traditional study design. Notably, the ML models, displayed the importance of FM disease severity with ASCVD risk status, in addition to other traditional CVD risk features like dyslipidemia. The high importance feature of both the 10-year ASCVD risk score and lifetime ASCVD risk ML models were age, weight, total cholesterol, FIQR score, and triglyceride level. Other features included BMI and BPI (average pain) score for the 10-year ASCVD ML model while the lifetime ASCVD risk ML model had revealed height and low-density lipoprotein levels, in addition to the aforementioned five parameters. Both models showed excellent validity after training. This demonstrates the ability of ML in identifying the non-linear relationships between the disease datasets (FM characteristics) and the target variable (ASCVD risk score) and improving prediction models.

The application of ML algorithms for the prediction of cardiovascular diseases is not novel. Data from a large cohort of 378,000 patients from a family practice database in England were analyzed using machine learning and was shown to significantly improve the accuracy of cardiovascular risk prediction when compared to 10-year ASCVD risk scores [28].

While the use of ML for data analysis from the general population for CVD risk prediction is increasingly studied; our study explores the novel application of ML in FM for studying the role of non-linear predictors of ASCVD risk (Table 4). However, in addition to the high feature importance of traditional risk factors (age, BMI, and fasting lipid profile) our ML model was able to identify that FM disease severity also had high feature importance for increasing CVD risk in female FM patients which is an interesting finding.

Given the widespread prevalence of heart disease, the importance of characterizing the "at-risk" population for CVD cannot be understated. In addition to the ASCVD risk score, carotid intimal medial thickness could be used as a surrogate marker for predicting CVD risk in future studies, as recommended by a meta-analysis report by Willeit et al. [29].

The external validity of our study could be limited as it was an observational study from a single center. Further prospective studies are required to establish causation, but as this study shows excellent internal validity, it can be used to guide future research. This study highlights a probable novel non-linear FM disease severity association with the ASCVD risk score that could help better stratify an individual FM patient for CVD risk.

Conclusion
Traditional risk factors like obesity, diabetes mellitus, hypertension, and dyslipidemia are the main pillars of CVD risk. In our study, the conventional case–control design showed that premenopausal female patients with FM could represent a new CVD risk group. The novel supervised ML study design identified FM disease severity as a high contributor feature importance for CVD risk, which was missed by the conventional study design. More definitive evidence can be gathered by future prospective studies employing non-invasive surrogate markers of CVD risk like carotid intimal medial thickness and coronary artery calcium scoring.

Acknowledgements
The authors would like to extend their note of gratitude to the clinical and administrative staff of Amrita Institute of Medical Sciences, Kochi, India for facilitating the data collection process.

Authors’ contributions
SS: conceptualization, data curation, formal analysis, methodology, software, validation, visualization, writing—original draft, writing—editing. CBM: conceptualization, project administration, data curation, formal analysis, methodology, writing—review. MM: conceptualization, data curation, formal analysis, methodology, supervision, validation, visualization, writing—review. AT: conceptualization, data curation, formal analysis, methodology, validation, visualization, writing—review. MP: data curation, formal analysis, methodology, validation, visualization, writing—review. The authors read and approved the final manuscript.

Funding
Not applicable.

Availability of data and materials
The patient medical records data used to support the findings of this study have been deposited in the Mendeley Data repository (http://dx.doi.org/10.17632/hvmshtzjm9.1).

Declarations
Ethics approval and consent to participate
The present study, including access to the patients’ biological and registry data from the hospital information system, was approved by the Institutional Ethics Committee (ECR/129/Inst/KL/2013/RR-16). India. Since this is a retrospective study, formal consent from patients could not be obtained.

Consent for publication
Not applicable.

Competing interests
The authors declare that they no competing interests.
Author details
1 Department of Rheumatology, Amrita Institute of Medical Sciences, Amrita Vishwa Vidyapeetham, Kochi, Kerala, India. 2 Department of General Medicine, Amrita Institute of Medical Sciences, Amrita Vishwa Vidyapeetham, Kochi, Kerala, India. 3 Department of Pharmacology, Amrita Institute of Medical Sciences, Amrita Vishwa Vidyapeetham, Kochi, Kerala, India.

Received: 18 June 2021   Accepted: 12 November 2021
Published online: 27 November 2021

References
1. Wolfe F, et al. 2016 Revisions to the 2010/2011 fibromyalgia diagnostic criteria. Semin Arthr Rheum. 2016;46(3):319–29. https://doi.org/10.1016/j.semarthrit.2016.08.012.
2. Walitt B, Nahin RL, Katz RS, Bergman MJ, Wolfe F. The prevalence and characteristics of fibromyalgia in the 2012 National Health Interview Survey. PLoS ONE. 2015;10(9):e0138024. https://doi.org/10.1371/JOURNAL.PONE.0138024.
3. Staud R, Spaeth M. Psychophysical and neurochemical abnormalities of pain processing in fibromyalgia. CNS Spectr. 2008;13(3 Suppl 5):12–7. https://doi.org/10.1016/J.S1916-0155.0527-7.
4. Caro XJ, Winter EF. The role and importance of small fiber neuropathy in fibromyalgia pain. Curr Pain Headache Rep. 2015;19(12):1–7. https://doi.org/10.1007/S11916-015-0527-7.
5. Staud R, Spaeth M. Psychophysical and neurochemical abnormalities of pain processing in fibromyalgia. CNS Spectr. 2008;13(3 Suppl 5):12–7. https://doi.org/10.1016/J.S1916-0155.0527-7.
6. Ho Jee Y, Chang H, Ji Jung K, Ha Jee S. Cohort study on the effects of acupoint acupuncture on the quality of life in patients with fibromyalgia. Medicine. 2016;95(49):e5542–e5542. https://doi.org/10.1097/MD.000000000005542.
7. gravelly RH, Ceko M, Bushnell MC. Fibromyalgia and depression. Pain Res Treat. 2012. https://doi.org/10.1155/2012/486590.
8. van der Kooy K, van Hout H, Marwijk H, Marten H, Stenhouwer C, Beekman A. Depression and the risk for cardiovascular diseases: systematic review and meta analysis. Int J Geriatr Psychiatry. 2007;22(7):613–26. https://doi.org/10.1002/GPS.1723.
9. Benilday IC. Role of inflammation in the pathogenesis and treatment of fibromyalgia. Rheumatol Int. 2019;39(5):781–91. https://doi.org/10.1007/S00296-019-04251-6.
10. Metyas S, Rezk T, Arifeld D, Leptich T. Autoinflammation and immunomodulation in inflammatory fibromyalgia syndrome: a review. Curr Rheumatol Rev. 2017;13(2):98–102. https://doi.org/10.1177/155739711666616091920530.
11. Sircas-Mainar A, Blanca-Tamayo M, Navarro-Arteida R, Rejas-Gutiérrez J. Use of resources and costs profile in patients with fibromyalgia or generalized anxiety disorder in primary care settings. Atene Primaria. 2009;41(2):84. https://doi.org/10.1016/1616-0919/20530.
12. da Cunha Ribeiro RP, et al. Cardiac autonomic impairment and chronotropic incompetence in fibromyalgia. Arthritis Res Ther. 2011;13(6):1–5. https://doi.org/10.1186/AR3519.
13. Uçeyler N, Häuser W, Sommer C. Systematic review with meta-analysis: cytokines in fibromyalgia syndrome. BMC Musculoskeletal Disord. 2011;12:1–15. https://doi.org/10.1186/1471-2474-12-245.
14. Wolfe F, Ross K, Anderson J, Russell IJ, Hebert L. The prevalence and characteristics of fibromyalgia in the general population. Arthritis Rheum. 1995;38(1):19–28. https://doi.org/10.1002/ART.1780380104.
15. Deo RC. Machine learning in medicine. Circulation. 2015;132(20):1920. https://doi.org/10.1161/CIRCULATIONAHA.115.01593.
16. Jordan ML, Mitchell TM. Machine learning: trends, perspectives, and prospects. Science (New York, NY). 2015;349(6245):252–60. https://doi.org/10.1126/SCIENCE.AAA8415.
17. "Project Hanover," Microsoft Research (2016). https://www.microsoft.com/en-us/research/project/project-hanover/. Accessed 24 Sept 2021.
18. Adamichou C, et al. Lupus or not? SLE Risk Probability Index (SLERPI): a simple, clinician-friendly machine learning-based model to assist the diagnosis of systemic lupus erythematosus. Ann Rheum Dis. 2021;80(6):758–66. https://doi.org/10.1136/ANNRHEUMDIS-2020-219069.
19. Menon VP, et al. Assessment of 2013 AHA/ACC ASCVD risk scores with behavioral characteristics of an urban cohort in India: preliminary analysis of noncommunicable disease Initiatives and Research at AMrita (NIRAM) study. Medicine. 2016;95(49):e5542–e5542. https://doi.org/10.1097/MD.0000000000005542.
20. Grazely RH, Ceko M, Bushnell MC. Fibromyalgia and depression. Pain Res Treat. 2012. https://doi.org/10.1155/2012/486590.
21. Benilday IC. Role of inflammation in the pathogenesis and treatment of fibromyalgia. Rheumatol Int. 2019;39(5):781–91. https://doi.org/10.1007/S00296-019-04251-6.
22. Metyas S, Rezk T, Arifeld D, Leptich T. Autoinflammation and immunomodulation in inflammatory fibromyalgia syndrome: a review. Curr Rheumatol Rev. 2017;13(2):98–102. https://doi.org/10.1177/155739711666616091920530.
23. Sircas-Mainar A, Blanca-Tamayo M, Navarro-Arteida R, Rejas-Gutiérrez J. Use of resources and costs profile in patients with fibromyalgia or generalized anxiety disorder in primary care settings. Atene Primaria. 2009;41(2):84. https://doi.org/10.1016/1616-0919/20530.
24. Gurer G, Sendur OF, Ay C. Serum lipid profile in fibromyalgia women. Clin Rheumatol. 2006;25(3):300–3. https://doi.org/10.1007/S00067-005-0004-5.
25. Mendieta D, et al. IL-8 and IL-6 primarily mediate the inflammatory response in fibromyalgia patients. J Neuroimmunol. 2016;290:22–5. https://doi.org/10.1016/j.jneuroim.2015.11.011.
26. Borges-Cosic M, et al. Sedentary time, physical activity, and sleep quality in fibromyalgia: the al-Ándalus project. Scand J Med Sci Sports. 2019;29(2):266–74. https://doi.org/10.1111/SMS.13318.
27. Weng SF, Reps J, Kai J, Garibaldi JM, Qureshi N. Can machine-learning improve cardiovascular risk prediction using routine clinical data? PLoS ONE. 2017;12(4):e0174944. https://doi.org/10.1371/JOURNAL.PONE.0174944.
28. Willeit P, et al. Carotid intima-media thickness progression as surrogate marker for cardiovascular risk. Circulation. 2020;142:621–42. https://doi.org/10.1161/CIRCULATIONAHA.120.046361.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.