To whom it may concern:

First of all, we would like to show our appreciation for considering to publish our manuscript in your decent journal. We hope that the reviewers could be satisfied to our answers. At last, thank you for these great comments and not only we learn a lot from them but also they improve the quality of our manuscript a lot.

Yours Sincerely
Chung-Yu Lin, M.D.

Reviewer #1:

1. **Study Design: This is a cross-sectional study, not a longitudinal one. Cross-sectional studies can only observe data at a specific point in time and cannot determine causality. This may limit the understanding of the relationship between risk factors and abnormal MPS. Longitudinal studies can track the same group of participants over a period of time, thereby more accurately determining causal relationships. Therefore, it is suggested to consider using a longitudinal study design in future research.**

Response:

We completely agree with the reviewer's comment. However, at present, this study is still ongoing. We continue to collect more participants. Right now, participants with two consecutive Thallium scan with a period 3 years still are too few, only 125 participants at present. In the future, when we finish the enrollment, we will repeat the study with same methods. More solid results will be obtained.

2. **Sample Size: The sample size of the study is relatively small (556 T2D patients). This may affect the statistical power and generalizability of the research results. It is suggested to increase the sample size, or provide some reasons to explain why this sample size is sufficient.**

Response:

As reviewer's concern, in the present study, the n number was not too large. Since Thallium scan is an expensive tool for evaluating coronary artery perfusion. To increase the number would be difficult. However, since we used four different models, the selection bias could be reduced by integrating these methods. In the same time, many other studies used machine learning even with smaller n number. For instance, the study done by Roy-Cardinal et al only included 66 patients (Marie-Helene Roy-Cardinal, Francois Destrempes, Gilles Soulez, Guy Cloutier, Assessment of Carotid Artery Plaque Components With Machine Learning Classification Using Homodyned-K Parametric Maps and Elastograms, IEEE Trans Ultrason Ferroelectr Freq Control, 2019 Mar;66(3):493-504.). In another study done by Latha et al., the n number was 361 (S Latha, P Muthu, Samiappan Dhanalakshmi, R Kumar, Khin Wee Lai 3, Xiang Wu, Emerging Feature Extraction Techniques for Machine Learning-Based Classification of Carotid Artery Ultrasound Images , Comput Intell Neurosci. 2022 May 12;2022:1847981). Moreover, in the book published by Vinaya et al., the authors pointed out that 200 participants would be enough for machine learning analysis (https://link.springer.com/chapter/10.1007/978-3-031-28183-9_26). We put the reviewer's concern in our limitation.

3. **Machine Learning Methods: The paper uses multiple machine learning methods to analyze the data, which is a promising approach. However, when describing these methods, it is suggested to provide more details, including the selection and adjustment process of hyperparameters, and how to deal with potential overfitting problems.**

Response:

We have added the following paragraph to respond to the reviewer's suggestion.

According to the proposed scheme, for modeling effective RF、SGB、NB and XGBoost models, use 10-fold cross-validation hyperparameters of each method are tuned and evaluated. The MLR method without hyperparameter tuning, the baseline method, was constructed by using the proposed scheme. The values of hyperparameters which generate the best RF、SGB、NB and XGBoost models are listed in the following table.

Summary of the values of the hyperparameters for the best RF、CART, NB and XGBoost models are shown in .

Methods	Hyperparameters	Best Value	Meaning
RF	mtry	8	The number of random features used in each tree.
	ntree	500	The number of trees in forest.
CART	minispilt	20	The minimum number of observations required to attempt a split in a node.

	minibucket	7	The minimum number of observations in a terminal node.
	maxdepth	10	The maximum depth of any node in the final tree.
	xval	10	Number of cross-validations.
	cp	0.03588	Complexity parameter: The minimum improvement required in the model at each node.
XGBoost	nrounds	100	The number of tree model iterations.
	max_depth	3	The maximum depth of a tree.
	eta	0.4	Shrinkage coefficient of tree.
	gamma	0	The minimum loss reduction.
	subsample	0.75	Subsample ratio of columns when building each tree.
	colsample_bytree	0.8	Subsample ratio of columns when constructing each tree.
	rate_drop	0.5	Rate of trees dropped.
	skip_drop	0.05	Probability of skipping the dropout procedure during a boosting iteration.
	min_child_weight	1	The minimum sum of instance weight.
NB	fL	0	Adjustment of Laplace smoother.
	usekernel	TRUE	Using kernel density estimate for continuous variable versus a Gaussian density estimate.
	adjust	1	Adjust the bandwidth of the kernel density.

RF: random forest; CART: classification and regression tree; SGB: stochastic gradient boosting; NB: Naïve Byer's classifier; XGBoost: eXtreme gradient boosting.

As for the overfitting issue, we utilized nested cross-validation (Nest-CV) to address and present the generalization and robustness of our result. Nest-CV is a variation of traditional CV with a straightforward concept. Under the structure of Nest-CV, two loops are required (inner-loop and outer-loop). The inner-loop is used for hyper-parameter tuning (which is equal to k-fold CV) whereas the outer-loop is used for model evaluation with the best found hyper-parameter set in inner-loop. Moreover, the concern of overfitting problem can be addressed utilizing Nest-CV which can also be found in several studies. However, the details of the Nest-CV is beyond the scope of the present study, we did not put these context in the manuscript. Please understand our rationale. (Parvandeh, S., Yeh, H. W., Paulus, M. P., & McKinney, B. A. (2020). Consensus features nested cross-validation. Bioinformatics, 36(10), 3093-3098. Tsamardinos, I., Rakhshani, A., & Lagani, V. (2015). Performance-estimation properties of crossvalidation-based protocols with simultaneous hyper-parameter optimization. International Journal on Artificial Intelligence Tools, 24(05), 1540023. Wei, L., Owen, D., Rosen, B., et al. (2021). A deep survival interpretable radiomics model of hepatocellular carcinoma patients. Physica Medica, 82, 295-305. Vabalas, A., Gowen, E., Poliakoff, E., & Casson, A. J. (2019). Machine learning algorithm validation with a limited sample size. PloS one, 14(11), e0224365.)