The Automated Deep Learning Detection of Mitral Valve Prolapse and Mitral Annular Disjunction at Scale: A UK Biobank CMR Study

Tugba Sanver[1], Hao Xu[1], Hannah Nicholls[2], Nay Aung[2], Pier Giorgio Masci[1], Alistair Young[1]

[1]King's College London, London, United Kingdom
[2]Queen Mary University of London, London, United Kingdom

Abstract

Mitral valve prolapse (MVP) and its structural variant, mitral annular disjunction (MAD), are associated with mitral regurgitation and increased risk of arrhythmias but remain underdiagnosed due to reliance on manual image review. We developed a two-stage nnU-Net deep learning pipeline on three-chamber end-systolic CMR views: landmark detection to localise the mitral valve, followed by leaflet segmentation, with screen-positivity defined as leaflet displacement beyond the annular plane exceeding 2 mm. Trained on 440 cases (98 MVP, 342 controls) and applied to 60,895 UK Biobank participants, the pipeline flagged 9.1% as screen-positive. Against expert adjudication (n=6,966), it achieved 0.999 sensitivity and 0.575 specificity for MVP/MAD. These findings support the pipeline's potential to enable population-scale phenotyping of valvular disease that would be infeasible with manual review alone.

1. Introduction

Mitral valve prolapse (MVP) occurs when one or both mitral leaflets billow into the left atrium (LA) during systole. The condition is present in roughly 2–3% of the population and carries an elevated risk of ventricular arrhythmia and sudden cardiac death, even among young individuals with no other apparent cardiac disease [1]. A related structural finding, mitral annular disjunction (MAD), involves an abnormal gap between the posterior mitral annulus and the inferolateral wall of the left ventricle (LV). MAD is commonly observed alongside MVP and has independently been linked to arrhythmic risk [2]. However, reported prevalence figures for MAD are inconsistent across the literature, largely because studies differ in imaging technique, diagnostic thresholds, patient selection, and the severity of accompanying mitral regurgitation. Among MVP patients, a meta-analytic estimate places the prevalence of MAD at 32.6% [3]. Distinguishing the two conditions can be challenging, since MAD may produce leaflet displacement that closely resembles true prolapse, and their respective contributions to valvular morphology and clinical outcomes are not yet fully understood.

Standard practice for identifying MVP and MAD relies on manual review of cine cardiovascular magnetic resonance (CMR) images in the three-chamber view. However, this approach is not scalable to large cohorts. Automated segmentation methods based on deep learning, such as nnU-Net, provide a more scalable solution: They can flag potential MVP/MAD cases for targeted expert review, thereby significantly reducing the manual effort necessary for population-level screening. nnU-Net has previously demonstrated strong out-of-the-box performance across a range of biomedical segmentation tasks without manual architecture tuning, making it a natural candidate for scalable, population-level cardiac phenotyping of this kind [4].

2. Methods

2.1. Study Population and Data

A total of 98 patients with MVP and 342 healthy participants from the UK Biobank were utilised to develop the network [5]. Trained networks were subsequently applied to 60,895 participants from the UK Biobank (under the UK Biobank access application 2964).

2.2. Image Acquisition and End-Systolic Frame Selection

Cine imaging in the three-chamber (3CH) long-axis orientation was performed on 1.5T Siemens scanners (Magnetom Aera, Siemens Healthcare, Germany) with breath-holding near expiration for both the training and testing cohorts. An automated algorithm identified the end-systolic (ES) frame, defined as the point of maximal mitral leaflet excursion. To confirm complete valve closure, the two frames immediately prior to this ES frame were examined.

2.3. Automated Deep Learning Pipeline

We built a two-stage pipeline based on nnU-Net [4], a self-configuring segmentation framework that automatically tailors its architecture and preprocessing steps to the provided dataset. The first stage used a landmark detection network to locate four anatomical reference points: the anterior and posterior margins of the mitral annulus, the anterior margin of the aortic valve, and the LV apex, which together defined a region of interest (ROI) around the mitral valve [6]. The second stage applied a segmentation network within this ROI to outline the area enclosed by the anterior leaflet, posterior leaflet, and annular plane.

Training labels were generated through manual annotation of ES images in cvi42 (Circle Cardiovascular Imaging, Canada), marking the mitral valve plane and leaflet boundaries (Fig. 1). The full dataset of 440 cases was divided into four outer folds of 110 cases each, with each fold using 110 cases for testing and the remaining 330 for training. Within each fold, nnU-Net's built-in five-fold cross-validation produced an ensemble of five sub-networks, whose outputs were averaged to yield a single per-case segmentation, resulting in four independent ensemble predictions per case.

Pipeline performance was assessed against expert adjudication using sensitivity (recall), precision (PPV), specificity, NPV, and F1 score, for two separate tasks: identifying any positive finding (MVP or MAD) and identifying MVP specifically, regardless of MAD status.

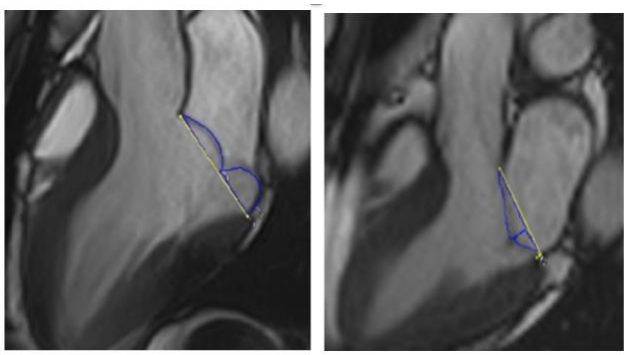

Figure 1. Manual annotations on 3CH end-systolic frames. Left: Bileaflet positive case (both leaflets, blue contours, prolapse into the left atrium in systole surpassing the annular plane, yellow straight line). Right: MVP-negative case. Yellow line = mitral valve plane; blue contours = leaflet boundaries.

2.4. MVP and MAD Detection Criterion

The algorithm measured the maximal distance of anterior or posterior mitral leaflets from the annular plane for each case. Consistent with the recognised CMR diagnostic standard for MVP [7], if leaflet displacement exceeded 2 mm towards the LA in at least one of the four ensemble predictions for a case, it was categorised as screen-positive. Given that the segmentation network is unable to distinguish between mitral annular disjunction (MAD) and mitral valve prolapse (MVP), and algorithmically positive cases could potentially indicate MVP, MAD, or both, definitive classification necessitates manual expert review.

2.5. Manual Expert Evaluation

All algorithmically screen-positive cases in the UK Biobank (n = 60,895) underwent manual adjudication by eight experts. These readers assessed both the automated segmentation output and the complete cine animation of the cardiac cycle. The cine review was selected as the definitive evaluation because it captures dynamic leaflet motion throughout systole and enables differentiation between MVP and other causes of apparent leaflet displacement. 565 cases were excluded due to poor image quality.

3. Results

3.1. Population Screening

The trained pipeline was then run on 60,895 UK Biobank participants. 300 cases were dropped from analysis due to poor image resolution or segmentation failure. Within the remaining cohort, 5,540 cases (9.1%) were flagged as screen-positive by at least one of the ensemble predictions. This rate is well above the expected 2–3% population prevalence of MVP, confirming the deliberately conservative over-detection built into the 2 mm threshold and underscoring the need for systematic manual review. Sample end-systolic frames illustrating each adjudication category are provided in Fig. 2.

3.2. Manual Evaluation Results

Table 1. Manual adjudication results (N=6,966). Prevalences are within the adjudicated (screen-positive enriched) dataset and are not representative of population prevalence.

Category	n	Prevalence
MAD+, MVP+	1400	20.1
MAD−, MVP+	895	12.8
MAD+, MVP−	1213	17.4
MAD−, MVP−	3458	49.6
Total	6966	100

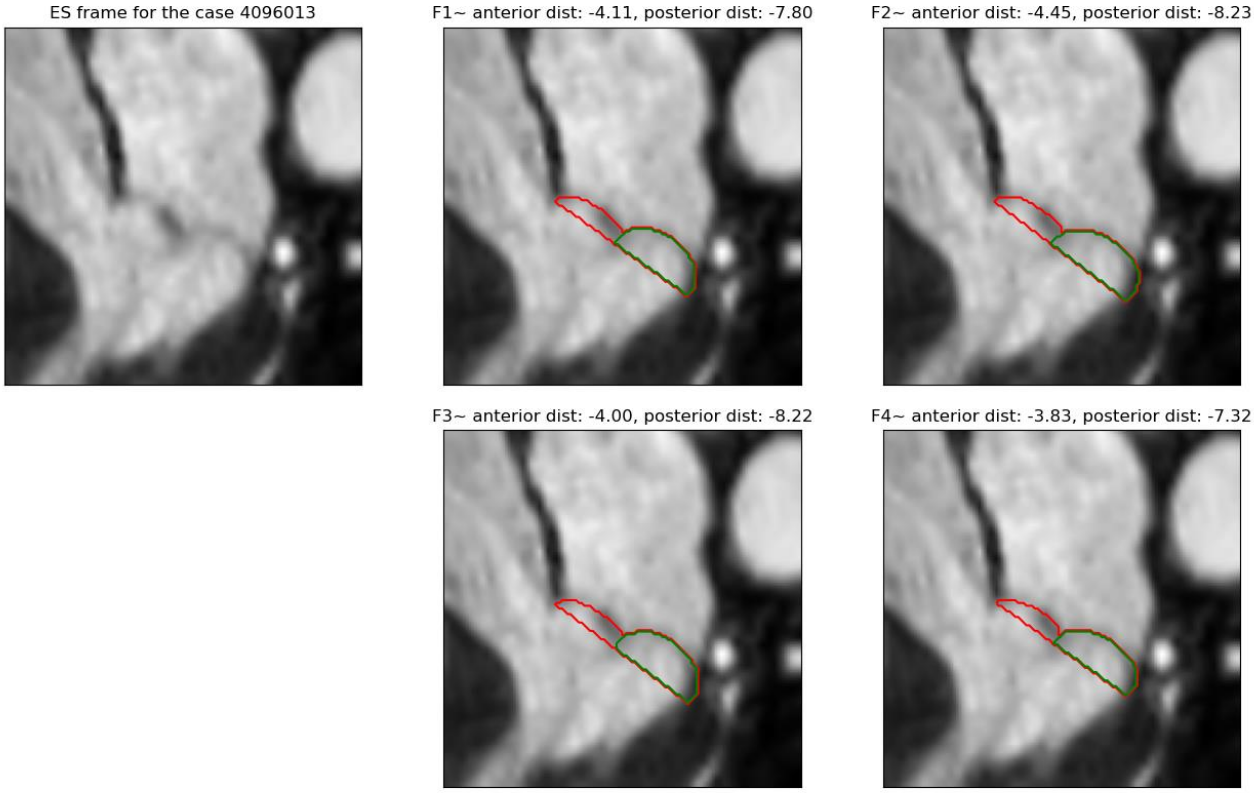

Figure 2. Representative end-systolic 3CH frames and automated segmentation outputs across the four ensemble predictions (F1–F4) for an MVP-positive case. Red = anterior leaflet; green = posterior leaflet. Negative displacement values indicate protrusion above the annulus (MVP). Displacement is measured independently for each ensemble member (F1–F4); a case is classified as screen-positive if any one of the four exceeds the 2 mm threshold, which explains the small variation in measured distances seen across panels for the same case.

For adjudication, a random subset of 1,991 algorithmically negative cases was combined with all 5,540 algorithmically positive cases, giving 7,531 cases in total. Of these, 565 were excluded due to inadequate image quality, leaving 6,966 cases with a final assessment. Each was independently classified as positive or negative for MVP and MAD. Within this adjudicated set, MVP, with or without co-occurring MAD, was identified in 32.9% of cases (n=2,295) (Table 1).

3.3. Diagnostic Performance

Compared against expert adjudication, the pipeline showed near-perfect sensitivity (0.999) for detecting either MVP or MAD, with a precision (PPV) of 0.705 and an F1 score of 0.827. When restricted to MVP detection alone, sensitivity reached 1.000, with precision at 0.461 and F1 at 0.631. The correspondingly lower specificity values (0.575 for MVP/MAD combined, 0.426 for MVP alone) are consistent with the intended overcalling at the 2 mm threshold and reflect the enrichment strategy used to select the manually reviewed cohort.

4. Discussion

This pipeline achieved near-perfect sensitivity (0.999–1.000) for MVP/MAD screening, at the cost of specificity; a deliberate trade-off given the 2 mm threshold and ensemble decision rule, which prioritise capturing true positives over minimising false alarms. This design is appropriate for population screening, where downstream expert adjudication can efficiently filter false positives, but missed cases cannot be recovered. This trade-off is particularly important for MVP and MAD, given their established link to ventricular arrhythmia and sudden cardiac death [1,2]; a screening approach that minimises false negatives is therefore preferable even at the expense of requiring additional expert adjudication.

The gap between the screen-positive rate (9.1%) and the adjudicated prevalence reflects known limitations of threshold-based displacement measurement, including sensitivity to imaging artefacts and off-axis slice positioning, which the segmentation network cannot resolve on its own. This underscores the pipeline's role as a screening tool rather than a standalone diagnostic

classifier.

Among MVP-positive cases, MAD was co-present in 61.0% (1400/2295), compared with 26.0% among MVP-negative cases, indicating a substantially elevated co-occurrence consistent with prior reports that MAD is frequently found alongside MVP [3]. This overlap presents a methodological limitation: because the segmentation network cannot itself distinguish MAD from MVP based on static leaflet displacement alone, definitive classification currently depends on manual cine review, which limits full automation of the pipeline.

Future work could incorporate dynamic annular motion features, such as the systolic curling motion, a distinguishing characteristic of MAD [8], which may enable algorithmic differentiation between the two entities without requiring expert adjudication. Other directions include threshold tuning or loss-function modifications to better balance sensitivity and specificity, expanding the training cohort with external datasets, incorporating additional long-axis views to capture non-central prolapse, and evaluating whether quantitative leaflet displacement measures can support downstream arrhythmic risk stratification.

5. Conclusion

This pipeline enables scalable, high-sensitivity automated MVP/MAD screening in population CMR cohorts, achieving near-perfect sensitivity while substantially reducing the manual review burden associated with large-scale imaging analysis. By combining automated pre-screening with targeted expert adjudication, this approach supports large-scale prevalence estimation and lays the groundwork for future cardiovascular risk stratification research linking valve morphology to arrhythmic outcomes.

Acknowledgments

This study was conducted in the UK Biobank resource under access application 2964. We thank the participants for their generous donations as well as the staff for their meticulous planning, collection, analysis and provision of data. TS was supported by the Centre for Doctoral Training in Digital Twins for Healthcare, King's College London. Part-funded by EPSRC grant Z533762 (AAY).

References

[1] B. Essayagh, A. Sabbag, C. Antoine, G. Benfari, L. T. Yang, J. Maalouf, S. Asirvatham, H. Michelena, and M. Enriquez-Sarano, "Presentation and outcome of arrhythmic mitral valve prolapse," J. Am. Coll. Cardiol., vol. 76, no. 6, pp. 637–649, Aug. 2020.

[2] D. Zugwitz, K. Fung, N. Aung, E. Rauseo, C. McCracken, J. Cooper, S. El Messaoudi, R. H. Anderson, S. K. Piechnik, S. Neubauer, et al., "Mitral annular disjunction assessed using CMR imaging: Insights from the UK Biobank population study," JACC Cardiovasc. Imaging, vol. 15, pp. 1856–1866, Nov. 2022.

[3] A. Sabbag et al., "EHRA expert consensus statement on arrhythmic mitral valve prolapse and mitral annular disjunction complex," Europace, vol. 24, pp. 1981–2003, Dec. 2022.

[4] F. Isensee, P. F. Jaeger, S. A. A. Kohl, J. Petersen, and K. H. Maier-Hein, "nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation," Nat. Methods, vol. 18, pp. 203–211, Feb. 2021.

[5] S. E. Petersen et al., "UK Biobank's cardiovascular magnetic resonance protocol," J. Cardiovasc. Magn. Reson., vol. 18, Art. no. 8, Feb. 2016.

[6] M. Sahota et al., "Machine learning evaluation of LV outflow obstruction in hypertrophic cardiomyopathy using three-chamber cardiovascular magnetic resonance," Int. J. Cardiovasc. Imaging, vol. 38, pp. 2695–2705, Dec. 2022.

[7] M. Spartera, G. Pessoa-Amorim, A. Stracquadanio, F. von Knobelsdorff-Brenkenhoff, J. V. Monmeneu, G. Miñana, R. Eiros, R. Fernández-Jiménez, I. Garcia-Lunar, B. Ibanez, J. Schulz-Menger, S. G. Myerson, S. Neubauer, S. E. Petersen, V. M. Ferreira, N. A. B. Ntusi, and S. K. Piechnik, "Cardiovascular magnetic resonance characterization of mitral valve prolapse," JACC Cardiovasc. Imaging, vol. 1, no. 2, pp. 134–143, Mar. 2008.

[8] S. V. Wifstad, G. Monopoli, H. A. Kildahl, et al., "Automated multimodal measurements of curling motion in patients with mitral valve disorders," Res. Square, preprint, 2026.

Address for correspondence:

Tugba Sanver
School of Biomedical Engineering and Imaging Sciences, Becket House 5th floor, 1 Lambeth Palace Road, SE1 7EU, United Kingdom
tugba.sanver@kcl.ac.uk