

1 **Title:** Improving prenatal detection of congenital heart disease with a scalable
2 composite analysis of six fetal cardiac ultrasound biometrics

3
4 **Authors:** Aneela Reddy MD, Sara Rizvi MS, Anita J. Moon-Grady MD, Rima Arnaout
5 MD

6
7 **ABSTRACT**

8
9 Although screening of prenatal congenital heart disease (CHD) has improved over the
10 last decade, the diagnosis rate can still be as low as 40%. The axial 4 chamber (A4C) is
11 the most reliably obtained cardiac view in the fetal screening ultrasound but alone only
12 has a maximum clinical sensitivity of 50-60%, particularly in large multicenter studies in
13 low-risk populations. Standard biometrics, like cardiac axis (CA), cardiothoracic ratio
14 (CTR) and cardiac chamber fractional area change (FAC), have individually been
15 shown to be useful for CHD screening and can all be obtained from A4C alone.
16 However, these biometrics are vastly underutilized because they are time-consuming to
17 extract and difficult to interpret all at once. We hypothesized that using six standard
18 biometrics in combination can improve complex CHD screening versus any one
19 biometric alone. K-means clustering was performed to segregate the patterns of heart
20 measurements into clusters. Sensitivity and specificity for CHD was 87% and 75%,
21 respectively. Here, we demonstrate that a composite of six standard biometric has
22 better sensitivity and accuracy for CHD than any one biometric alone and better than
23 A4C visual assessment.
24

25 MAIN TEXT

26
27 Although screening of prenatal congenital heart disease (CHD) has improved over the
28 last decade, the diagnosis rate can still be as low as 40%¹. The axial 4 chamber (A4C)
29 is the most reliably obtained cardiac view in the fetal screening ultrasound but alone
30 only has a maximum clinical sensitivity of 50-60%, particularly in large multicenter
31 studies in low-risk populations.² Standard biometrics, like cardiac axis (CA),
32 cardiothoracic ratio (CTR) and cardiac chamber fractional area change (FAC), have
33 individually been shown to be useful for CHD screening and can all be obtained from
34 A4C alone.^{3,4} However, these biometrics are vastly underutilized because they are time-
35 consuming to extract and difficult to interpret all at once.

36 We hypothesized that using six standard biometrics in combination can improve
37 complex CHD screening versus any one biometric alone. We included 105 fetal
38 echocardiograms (20 normal, 85 abnormal comprising 12 different CHD lesions). The
39 chambers, thorax, and spine from the A4C view were segmented by an expert reader,
40 and CA, CTR, right ventricular (RV) FAC, left ventricular (LV) FAC, right atrium:left
41 atrium area ratio, and RV:LV area ratio were automatically calculated (Fig 1). K-means
42 clustering (scikit-learn.org/) was performed to segregate the patterns of heart
43 measurements into clusters.

44 The optimal number of clusters was four (based on silhouette score), with RV:LV ratio
45 and CTR as the most important features distinguishing clusters. Cluster 1 was
46 predominantly normal hearts. Cluster 2 consisted of lesions with either typically normal
47 A4C views (aortic coarctation, d-transposition of great arteries (dTGA)) or more subtle
48 abnormalities in the A4C view (tetralogy of Fallot (TOF), atrioventricular canal defect
49 (AVSD), aortic stenosis (AS), truncus arteriosus, total anomalous pulmonary venous
50 return (TAPVR)). Clusters 3 and 4 included severe CHD lesions (single ventricle,
51 Ebstein's anomaly, and hypoplastic left heart syndrome (HLHS)).

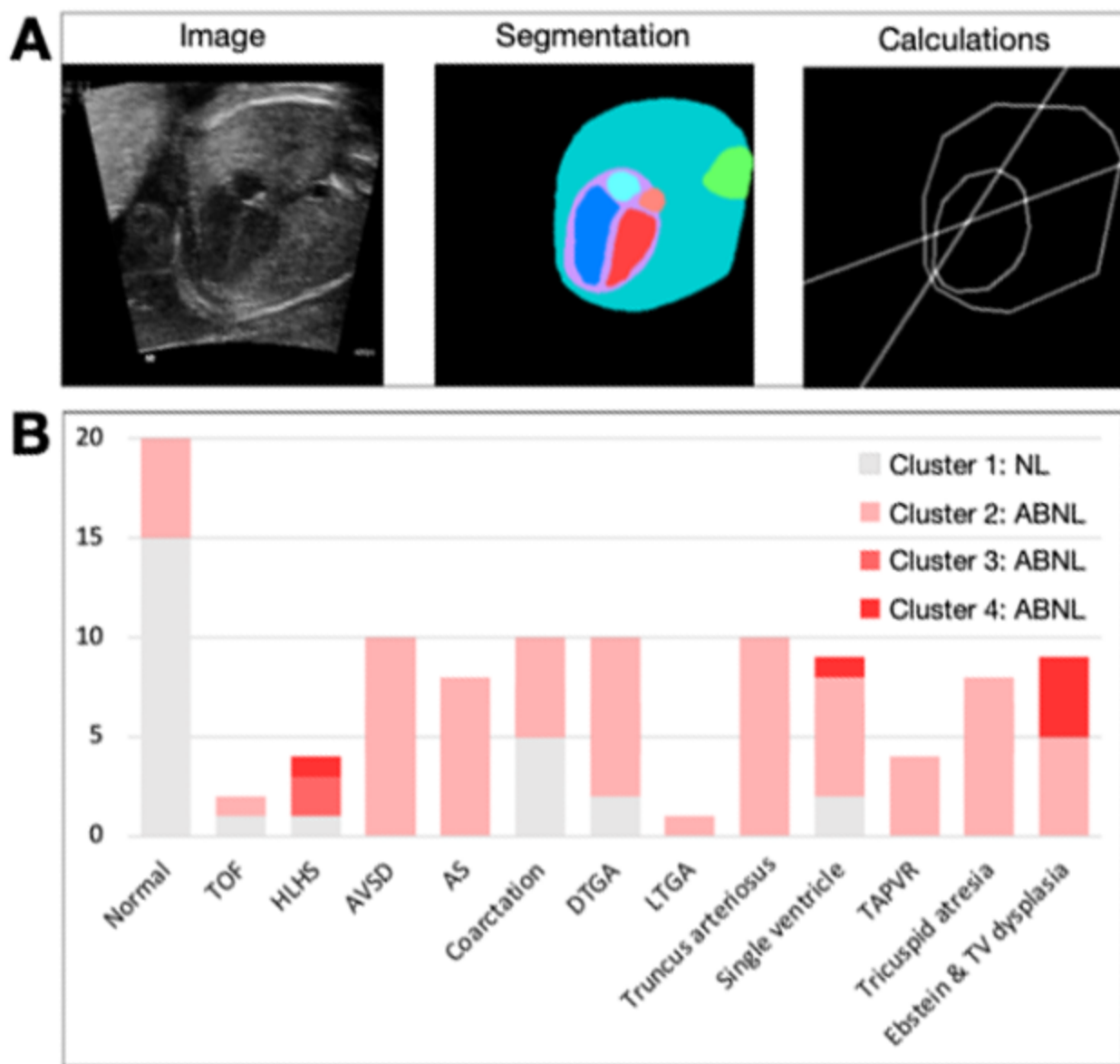
52 We binarized these clusters into normal (1) vs CHD (2-4) and calculated sensitivity,
53 specificity, and accuracy of distinguishing normal vs CHD hearts. Sensitivity and
54 specificity for CHD was 87% and 75%, respectively (Table 1).

55 Of the 11 false negatives, 8 belonged to lesions that often have a normal A4C
56 (coarctation, TOF, dTGA). Other false negatives included two cases of single ventricle
57 (SV) and one HLHS. For these SV cases, the ratio of the putative LV and RV segments
58 were normal, even though there was no complete interventricular septum. Similarly, the
59 false-negative HLHS was one where the LV was in fact dilated rather than hypoplastic,
60 which again resulted in a normal RV:LV ratio.

61 Here, we demonstrate that a composite of six standard biometric has better sensitivity
62 and accuracy for CHD than any one biometric alone and better than A4C visual
63 assessment (Table 1)^{2,5}. We focused on the biometrics described in order to test
64 whether increased diagnostic accuracy is possible using measurements that are simple

65 and already widely accepted, using the view most reliably obtained during screening.
66 We used expert segmentation to remove the possibility that automatic segmentation
67 (e.g., from a deep learning model⁷) could confound the findings.
68 Interestingly, there is improved accuracy even for lesions that traditionally have a
69 normal A4C view. About 40% of CHD lesions are typically only appreciated in the
70 outflow tract or 3-vessel views.⁶ 31 of our 105 fetal echocardiograms are lesions that
71 typically have a normal A4C view. With our composite biometric, the sensitivity,
72 specificity and accuracy for these lesions were 74%, 75% and 75% respectively.
73 Therefore, with composite analysis of all six biometrics, we can increase the utility of the
74 A4C view even for these lesions. Given that A4C is the most reliably obtained cardiac
75 view in the fetal screening ultrasound, the improvement in utility of the A4C view has
76 profound screening implications. We do not expect for the composite biometric to
77 replace overall clinical assessment, but rather to serve as an aid, similar to how single
78 biometrics are currently used.
79 Despite its value, this study has several limitations. The size of the dataset analyzed is
80 relatively small; however, it did include a comprehensive number of CHD lesions. Also,
81 while there is improved diagnostic accuracy when using a composite of the six
82 biometrics, this method is not perfect. The 11 false negative hearts had biometric
83 features that calculated as normal with respect to the RV:LV ratio and CTR—which
84 were important distinguishing features among clusters—despite being abnormal lesions.
85 This shortcoming in fact raises the exciting possibilities that there may be additional,
86 non-traditional biometrics that can be incorporated into the model for greater
87 performance; these may come from the A4C view or, may incorporate other views such
88 as the LVOT view. Testing additional non-standard biometrics within our model would
89 be an interesting avenue for further study.
90 Now that we have shown the value-add of analyzing multiple biometrics within a
91 composite model, advances in automation make multiple-biometric approaches
92 increasingly practical. End-to-end automation (i.e., including automatic deep learning-
93 based segmentation⁷) would eliminate manual tracing and allow interpretation of
94 multiple biometrics at once in large data sets, thereby decreasing clinician burden and
95 bringing enhanced biometric screening into practice. We look forward to further
96 automating, refining, and scaling testing to prove clinical utility of composite biometrics
97 in a larger population.
98

99 **Figure 1. Segmentation and automatic calculation of biometrics.** (A) Axial 4-
100 chamber images were segmented (LV (red), RV (blue), LA (light red), RA (light blue),
101 heart (purple), thorax (teal), and spine (green)). Chamber area ratios, cardiothoracic
102 ratio, and cardiac axis measurements were automatically calculated from these
103 segmentations; chamber fractional area changes were calculated using similar
104 segmentations across the cardiac cycle. Calculations for fractional area change, RA:LA
105 and RV:LV ratio were all derived from pixel areas. (B) K-means clustering of the
106 resulting measurements is used as a composite biometric. Clustering result by lesion is
107 shown.
108



110 **Table 1. Performance of a composite biometric in distinguishing normal vs CHD**
111 **hearts, compared to each individual biometric and to visual assessment as**
112 **established by current literature**

113

Biometric	Sensitivity	Specificity	Accuracy
Cardiothoracic ratio	27%	100%	41%
Cardiac axis	6%	100%	4%
LV FAC	85%	50%	78%
RV FAC	75%	70%	74%
RA:LA ratio	84%	20%	71%
RV:LV ratio	62%	45%	59%
Six-biometric composite	87%	75%	85%
A4C visual assessment*	49%	100%	—

114 LV, left ventricle. RV, right ventricle. LA, left atrium. RA, right atrium. FAC, fractional
115 area change. A4C, axial 4-chamber. *^{2,5}

116 **References**

117

- 118 1. Donofrio MT, Moon-Grady AJ, Hornberger LK, et al. Diagnosis and treatment of fetal
119 cardiac disease: a scientific statement from the American Heart Association.
120 *Circulation* 2014;129(21):2183-2242.
- 121 2. Oggè G, Gaglioti P, Maccanti S, et al. Prenatal screening for congenital heart disease
122 with four-chamber and outflow-tract views: a multicenter study. *Ultrasound Obstet*
123 *Gynecol Off J Int Soc Ultrasound Obstet Gynecol* 2006;28(6):779-784.
- 124 3. Shipp TD, Bromley B, Hornberger LK, et al. Levorotation of the fetal cardiac axis: a
125 clue for the presence of congenital heart disease. *Obstet Gynecol* 1995;85(1):97-102.
- 126 4. DeVore GR, Satou G, Sklansky M. Area of the fetal heart's four-chamber view: a
127 practical screening tool to improve detection of cardiac abnormalities in a low-risk
128 population. *Prenat Diagn* 2017;37(2):151-155.
- 129 5. Zhang YF, Zeng XL, Zhao EF, et al. Diagnostic Value of Fetal Echocardiography for
130 Congenital Heart Disease. *Medicine* 2015;94(42):e1759.
- 131 6. Gonçalves LF, Bronsteen R, Lee W. Fetal Heart: A 4-chamber View Is Not Enough.
132 *Clin Obstet Gynecol* 2012;55(1):266-280.
- 133 7. Arnaout R, Curran L, Zhao Y, et al. An ensemble of neural networks provides expert-
134 level prenatal detection of complex congenital heart disease. *Nat Med*
135 2021;27(5):882-891.