

ORIGINAL RESEARCH

Outcomes of Echocardiography-Detected Rheumatic Heart Disease: Validating a Simplified Score in Cohorts From Different Countries

Bruno R. Nascimento , MD, MSc, PhD; Maria Carmo P. Nunes , MD, PhD; Emily M. Lima, BSc, PhD; Amy E. Sanyahumbi , MD; Nigel Wilson , MBChB; Elizabeth Tilton, MD; Marc G. W. Rémond , PhD; Graeme P. Maguire , MD; Antonio Luiz P. Ribeiro , MD, PhD; Peter N. Kazembe , MBChB; Craig Sable , MD; Andrea Z. Beaton , MD; on behalf of the Programa de Rastreamento da Valvopatia Reumática (PROVAR) Investigators*

BACKGROUND: The natural history of latent rheumatic heart disease (RHD) detected by echocardiography remains unclear. We aimed to assess the accuracy of a simplified score based on the 2012 World Heart Federation criteria in predicting mid-term RHD echocardiography outcomes in children from 4 different countries.

METHODS AND RESULTS: Patient-level baseline and follow-up data of children with latent RHD from 4 countries (Australia, n=62; Brazil, n=197; Malawi, n=40; New Zealand, n=94) were combined. A simplified echocardiographic scoring system previously developed from Brazilian and Ugandan cohorts, consisting of 5 point-based variables with respective weights, was applied: mitral valve anterior leaflet thickening (weight=3), excessive leaflet tip motion (3), regurgitation jet length ≥ 2 cm (6), aortic valve focal thickening (4), and any regurgitation (5). Unfavorable outcome was defined as worsening diagnostic category, persistent definite RHD or development/worsening of valve regurgitation/stenosis. The score model was updated using methods for recalibration. 393 patients (314 borderline, 79 definite RHD) with median follow-up of 36 (interquartile range, 25–48) months were included. Median age was 14 (interquartile range, 11–16) years and secondary prophylaxis was prescribed to 16%. The echocardiographic score model applied to this external population showed significant association with unfavorable outcome (hazard ratio, 1.10; 95% CI, 1.04–1.16; $P=0.001$). Unfavorable outcome rates in low (≤ 5 points), intermediate (6–9), and high-risk (≥ 10) children at 3-year follow-up were 14.3%, 20.8%, and 38.5% respectively ($P<0.001$). The updated score model showed good performance in predicting unfavorable outcome.

CONCLUSIONS: The echocardiographic score model for predicting RHD outcome was updated and validated for different latent RHD populations. It has potential utility in the clinical and screening setting for risk stratification of latent RHD.

Key Words: echocardiography ■ follow-up ■ prognosis ■ rheumatic heart disease ■ screening

Globally, rheumatic heart disease (RHD) is the most prevalent acquired cardiovascular disease in children and young adults, affecting 40.5 million

people and causing 306 000 deaths annually.^{1,2} RHD is strongly associated with poor socioeconomic conditions, resulting in an uneven global distribution. The

Correspondence to: Bruno Ramos Nascimento, MD, MSc, PhD, FACC, FESC, Hospital das Clínicas da Universidade Federal de Minas Gerais, Rua Muzambinho 710/802, Serra, Belo Horizonte, Minas Gerais, Brazil, CEP 30.210-530. E-mail: ramosnas@gmail.com

*A complete list of the PROVAR Investigators can be found in the Appendix at the end of the manuscript.

Presented in part at the American College of Cardiology Scientific Sessions, Chicago, IL, March 28–30, 2020.

Supplementary Material for this article is available at <https://www.ahajournals.org/doi/suppl/10.1161/JAHA.121.021622>

For Sources of Funding and Disclosures, see page 10.

© 2021 The Authors. Published on behalf of the American Heart Association, Inc., by Wiley. This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

JAHA is available at: www.ahajournals.org/journal/jaha

CLINICAL PERSPECTIVE

What Is New?

- A simplified echocardiography score developed from Brazilian and Ugandan cohorts, consisting of 5 point-based variables, seem to predict unfavorable echocardiographic outcome in schoolchildren with latent rheumatic heart disease.
- When applied to populations from 4 different countries, the simplified score accurately predicted unfavorable echocardiography outcome at 3 years (hazard ratio, 1.10; 95% CI, 1.04–1.16; $P=0.001$), with good discrimination of those at higher risk.
- The area under the receiver operating characteristic curve (C-statistic) was 0.70, 95% CI, 0.64–0.76.

What Are the Clinical Implications?

- This simplified echocardiographic risk score has a potential application as a risk stratification tool following rheumatic heart disease screening in 4 distinct populations.
- It can be an additional tool to guide clinical monitoring and secondary prophylaxis, especially in underserved areas.
- More research is necessary for additional validation of the score in different settings and populations, to allow for its broad application.

Nonstandard Abbreviations and Acronyms

RHD	rheumatic heart disease
WHF	World Heart Federation

highest burden of disease is seen in low- and middle-income countries, where access to health resources is frequently suboptimal. In such settings, RHD diagnosis is frequently delayed so that at time of diagnosis the patient has established late sequelae and markedly symptomatic valve disease.² RHD is a preventable disease and the role of primary prevention and secondary prophylaxis following a diagnosis of acute rheumatic fever is standard. Nonetheless, approaches for the management of latent RHD – RHD diagnosed by echocardiographic screening absent a history of acute rheumatic fever – remains controversial.^{3,4}

Latent RHD, is a relatively new diagnostic category.^{5–8} Diagnostic definitions were initially heterogeneous. However, in 2012 the World Heart Federation (WHF) published standardised criteria for the diagnosis of latent RHD (ie, a diagnosis of RHD based on echocardiographic

findings in patients that were asymptomatic without a known history of acute rheumatic fever). The WHF criteria define 3 diagnostic categories (definite RHD, borderline RHD, and normal),⁹ based on functional and morphological findings of the left-sided valves.

Although diagnostic standardization enabled meaningful comparisons between RHD screening studies (a crucial step for research purposes) the prognosis of latent RHD and the impact of interventions such as secondary prophylaxis and intensive clinical follow-up are still under investigation.¹⁰ Considering that disease progression, particularly from borderline to definite RHD, is considered a serious adverse health outcome,^{10–12} efforts have been made to identify subpopulations at highest risk of progression so as to target them with secondary prevention initiatives. For example, in Africa a subdivision of the definite category has been proposed based on the hazard for echocardiographic and clinical progression in those with different valvular manifestations of definite RHD.¹² In addition, Nunes and colleagues developed a simplified echocardiographic score derived from screening studies in Brazil and Uganda which showed good discrimination of risk of progression in 2.3 years in a separate Ugandan population subset.¹¹ Although follow-up in this study was short-term, the simplified echocardiography score holds promise as a tool for guiding patient care. However, a wider application of the score requires robust external validation in different settings.

In this study, we aimed to assess the accuracy of the simplified scoring system developed by Nunes et al,¹¹ consisting of 5 components of the WHF criteria, in predicting mid-term RHD echocardiographic outcomes in a cohort of children from 4 different countries who were previously diagnosed with subclinical RHD on screening echocardiography.

METHODS

Upon reasonable request to the corresponding author, study data and analytical methods may be made available to other researchers for the purposes of reproducing the results or replicating the procedures of this study.

This is an external validation study using primary data from pre-existing RHD screening cohorts that included systematic echocardiographic follow-up of patients initially diagnosed with latent RHD. Lead investigators from different research groups were contacted and invited to participate in the pooled analysis. After acceptance, patient-level baseline and follow-up data of children (aged 5 to 18 years) diagnosed with latent RHD during screening studies conducted in 4 countries (Australia,¹³ $n=62$; Brazil [sample not included in the original derivation study],¹⁴ $n=197$; Malawi,¹⁵ $n=40$; New Zealand,¹⁶ $n=94$) were provided. The data included

echocardiographic findings, derived from studies with portable or handheld devices, for each of the variables used in making a diagnosis under the WHF criteria.⁹ According to these criteria, borderline RHD is defined as a single type of finding (i.e., abnormal mitral morphology, or isolated pathological mitral/aortic regurgitation), while definite RHD includes combinations of these, or more advanced findings (i.e., abnormal mitral morphology plus pathological regurgitation, or mitral stenosis, or abnormal aortic morphology plus pathological regurgitation, or borderline findings at both valves).⁹

Ethics approval for inclusion of each individual cohort data set into the pooled set was obtained from the local institutional review boards, and all patients included in this analysis signed an informed consent form before enrollment.

Data sets were merged into a single spreadsheet (Excel, Microsoft, Redmond, VA, USA), and data verification and standardization was performed by 2 investigators (B.R.N. and M.C.P.N.) before statistical analysis. If needed, original investigators were contacted by email to provide clarifications or additional/missing data. Baseline and follow-up diagnoses by the readers from the original study teams were used; echocardiographic images were not centrally re-analyzed for the purposes of this study, aiming to maintain the original observations.

The simplified echocardiographic score proposed by Nunes et al, consisting of 5 point-based variables with respective weights (mitral valve anterior leaflet thickening [weight=3], excessive leaflet tip motion [3], regurgitation jet length ≥ 2 cm [6], aortic valve focal thickening [4], and any regurgitation [5])¹¹ was applied to the pooled population. In the original derivation study, the risk of unfavorable outcome of latent RHD over 3 years was estimated for 3 groups: low-risk (0–6 points, 7% risk), intermediate risk (7–9 points, 30%) and high-risk (≥ 10 points, 53%). Details of the derivation and validation methodology used to generate the simplified echocardiographic score have been described elsewhere.¹¹ Summarily, 3 cohorts of children were used for: (1) score derivation (echocardiographic screening study in Minas Gerais, Brazil; $n=9501$), (2) score validation (echocardiographic screening study in Gulu, Uganda; $n=7312$), and (3) outcomes prediction (longitudinal echocardiography cohort in Kampala, Uganda; $n=227$). In derivation, variables independently associated with definite RHD were assigned points proportional to their regression coefficients.¹¹

An unfavorable outcome was defined as a worsening in diagnostic category (borderline to definite), remaining with definite RHD, a worsening in the grade of mitral or aortic regurgitation, or the development or worsening grade of mitral stenosis. A favorable outcome was defined as disease regression (ie, an improvement in diagnostic category), a reduction in the

severity of valvular regurgitation, or remaining with stable borderline RHD (Figure S1).

Patient and Public Involvement

Patients and public were not involved in the design and conduct of this research.

Statistical Analysis

Statistical analysis was performed using SPSS® software version 23.0 for Mac OSX (SPSS Inc., Chicago, Illinois) and R for Statistical Computing version 4.0.3 (R Foundation, Vienna, Austria). As we used pre-existing cohorts, no pre-specified sample size calculation was performed, and we included all the schoolchildren enrolled in the 4 primary studies. Categorical variables, expressed as numbers and percentages, were compared between each cohort included in the pooled set and between the 3 outcome groups (progression from borderline to definite RHD versus remaining with definite RHD versus remaining borderline RHD or regression from definite to borderline RHD) using Fisher exact test. Continuous data, expressed as mean ($\pm$ SD) or median (interquartile range [IQR]), were compared using the Student unpaired *t*-test or the Mann-Whitney *U* test, as appropriate. External validation was performed by applying the original echocardiographic score model to an independent population of schoolchildren with latent RHD to assess its discrimination and calibration in predicting mid-term unfavorable outcome (Figure S1). Additionally, the score model was updated by recalibrating the intercept and slope^{17,18} and its performance was assessed in the external validation data set using different methods. The overall model performance was assessed using the Brier score, calibration using the Hosmer–Lemeshow test and calibration plots, and discrimination using C-statistic. For validation of the updated model, a bootstrapping procedure with 200 samples was used. A separate analysis was performed excluding individuals under prophylaxis.

Kaplan–Meier survival analysis was undertaken to assess unfavorable outcome -free survival (times-to-an unfavorable outcome, considering the follow-up intervals) in the 3 risk categories defined by the original score (low-risk [0–6 points]), intermediate risk (7–9 points), and high-risk (≥ 10 points). Differences between these groups were assessed by the non-parametric log-rank test. A 2-tailed significance level of 0.05 was considered statistically significant.

RESULTS

At total, 393 patients (314 with borderline and 79 with definite RHD) were included in the final analysis, with median follow-up of 36 (IQR, 25–48) months. Median

age was 14 (IQR, 11–16) years and penicillin prophylaxis was prescribed to 16%. Detailed characteristics of the individual cohorts included in the pooled data set are presented in Table 1. Of note, the longest median follow-up intervals were observed in New Zealand and Malawi (58 [IQR, 37–77] and 48 [IQR, 24–48] months, respectively). These countries also had the highest proportions of definite RHD cases at baseline (31.9% and 27.5%). The highest proportion of patients in the high-risk group, according to the original score was observed in Australia. This group also had the highest rate of echocardiographic unfavorable outcome (Table 1). Overall, the number of patients in each risk strata was: low 182 (46.3%), intermediate 120 (30.5%), and high 91 (23.2%).

Among children with borderline RHD, 42 (13%) progressed to definite, 114 (36%) remained stable, and 151 (48%) regressed to normal. For those with definite RHD, 44 (56%) remained definite, while 15 (19%) regressed to borderline, and 16 (20%) to normal (Table 1). Among 63 patients receiving penicillin prophylaxis, 22 (35%) had unfavorable outcome. The baseline echocardiographic findings for patients relating to variables used in the WHF criteria (stratified by favorable/unfavorable RHD outcome at follow-up) are detailed in Table 2. Additional stratification according to progression/stabilization/regression is provided in Table S1.

Validation of the Simplified Echocardiographic Score Model

The original echocardiographic score model applied to the independent population of schoolchildren showed poor calibration (Hosmer–Lemeshow test, $P=0.059$), and the calibration plot showed that predicted values were lower than observed values (Figure S2). The discrimination was modest (C-statistic of 0.67, 95% CI, 0.61–0.74). The score model with the updated intercept and calibration slope (with the coefficients intercept=−0.724 and slope=0.241) resulted in improvement in overall model performance (Hosmer–Lemeshow test of 4.19, $P=0.651$) and the calibration plot improved considerably (Figure 1A). The discrimination also improved with C-statistic of 0.70 (95% CI, 0.64–0.76) (Figure 2). The Brier score (0.156) indicated a good overall performance of the updated model. Bootstrap validation demonstrated that the model was well calibrated with the Hosmer–Lemeshow test of 11.228 ($P=0.129$) and reasonable discrimination (C-statistic of 0.70 [95% CI, 0.59–0.80]). Applying the updated echocardiographic score model (Table 3) to the validation population, the predicted unfavorable outcome rates in the low, intermediate, and high-risk groups were 13.5%, 23.2%, and 36.9%, respectively. The observed rates in these groups were 14.3%, 20.8%, and 38.5% respectively (Figure 1B, Figure S3).

The updated score model was significantly associated with unfavorable echocardiographic outcome in continuous format (hazard ratio, 1.101; 95% CI, 1.042–1.164; $P=0.001$, for each point increase). Survival rates (free of unfavorable outcomes) in low-risk children at 1, 2, and 3 years of follow-up were 96%, 89%, and 71%, respectively, compared with 92%, 88%, and 77% in the intermediate group and 94%, 87%, and 60% in the high-risk group (log-rank test, $P=0.028$, Figure 3). The high-risk group was associated with unfavorable outcomes (hazard ratio, 1.851; 95% CI, 1.108–3.094; $P=0.019$), whereas no differences was found comparing intermediate and low risk (hazard ratio, 1.062; 95% CI, 0.611–1.846; $P=0.831$). The Kaplan–Meier curves revealed a significant gradual decrease in progression-free rates after 3-year follow-up in the high-risk compared with low and intermediate groups (Figure 3).

Excluding children under secondary prophylaxis, the score's association with unfavorable outcome remained unchanged: hazard ratio of 1.100 (95% CI, 1.031–1.172; $P=0.004$), with a slightly lower C-statistic of 0.66 (95% CI, 0.58–0.73).

DISCUSSION

Our study, using pooled patient-level data from RHD screening cohorts originating in 4 different countries, revealed that a simplified risk score previously developed by Nunes et al. based on the WHF criteria^{9,11} accurately predicted the risk of an echocardiographic unfavorable outcome in individuals previously diagnosed with latent RHD. The model was updated considering differences in disease progression and prevalence among the countries. Our results indicate that the model of Nunes et al¹¹ may be a potential tool for guiding medical care of children diagnosed with latent RHD following screening programs.

In the decade since the publication of the WHF criteria,⁹ which standardized echocardiographic-diagnosis of latent RHD, the production of large data sets from across the globe using these criteria have made it possible to assess the performance and weight of its individual components. The initial derivation and validation of a risk score by Nunes and colleagues demonstrated that discrimination between borderline and definite echo-detected RHD in the derivation cohort was near-perfect, and that in an external validation cohort from Africa the score was consistently able to predict disease outcome.¹¹ One limitation of this study, however, was that a similar group of experts analyzed all the echocardiographic images collected for participants, which could have favored a more homogeneous interpretation. Subsequently, an external validation was performed in a Brazilian cohort, with good – although more modest – discrimination of those at risk of echo

Table 1. Baseline Characteristics of Patients With Borderline and Definite Rheumatic Heart Disease

Variable	Australia (n=62)	Brazil (n=197)	Malawi (n=40)	New Zealand (n=164)
Screening setting	Remote communities	Public and private schools, primary care	Schools/surrounding communities	Urban and rural schools
Age range (for inclusion, y)	8–18	5–18	5–16	5–17
Screening personnel	Accredited echocardiographers	Physicians and non-physicians	Pediatric cardiologist, ultrasonographer	Accredited echocardiographers
Echocardiography adjudication	1 expert pediatric cardiologist	2 independent cardiologists and 1 tie-breaker	2 independent cardiologists and 1 tie-breaker	Pediatric Cardiology Department
Echocardiography machines	GE Vivid e/f	GE Vivid Q, VScan	Phillips CX50	GE Vivid e
BPG recommendation	At the discretion of cardiologist	At the discretion of cardiologist	Definite cases	Definite or probable cases
Age (median, IQR, y)	14.0 (11.9–15.6)	14.0 (11.7–16.0)	11.1 (9.4–13.3)	N/A
Follow-up time (mo, median, IQR, y)	44 (40–48)	31 (22–36)	48 (24–48)	58 (37–77)
Baseline diagnosis (n, %)				
Borderline	51 (82.3)	170 (86.3)	29 (72.5)	64 (68.1)
Definite	11 (17.7)	27 (13.7)	11 (27.5)	30 (31.9)
Follow-up diagnosis (n, %)				
Normal	20 (32.3)	91 (46.2)	15 (37.5)	41 (43.6)
Borderline	23 (37.1)	54 (27.4)	17 (42.5)	35 (37.2)
Definite	19 (30.6)	42 (21.3)	8 (20.0)	17 (18.1)
Score categories (n, %)				
Low	19 (30.6)	121 (61.4)	10 (25.0)	32 (34)
Intermediate	9 (14.5)	49 (24.9)	21 (52.5)	41 (43.6)
High	34 (54.8)	27 (13.7)	9 (22.5)	21 (22.3)
Unfavorable echocardiographic outcome* (n, %)	19 (30.6)	42 (21.3)	8 (20.0)	17 (18.1)

AR indicates aortic regurgitation; AV, aortic valve; BPG, benzathine penicillin G; GE, General Electric; IQR, interquartile range (Q1–Q3); MR, mitral regurgitation; MS, mitral stenosis; MV, mitral valve; N/A, not available; and RHD, rheumatic heart disease.

*Unfavorable echocardiographic outcome: worsening in diagnostic category (borderline to definite), remaining with definite rheumatic heart disease, a worsening in the grade of mitral or aortic regurgitation, or the development or worsening grade of mitral stenosis.

Table 2. Number and Percent of Patients With and Without Each Echocardiographic Variable Who Presented Unfavorable Outcome at 2-Year Follow-Up

Valve	Variable (n=present)	Unfavorable outcome (n=86)	
		Variable present (+)	Variable absent (-)
Mitral valve, n (%) [*]	Anterior leaflet thickening [†] (n=165)	50 (30.3)	36 (15.8)
	Chordal thickening (n=21)	13 (61.9)	73 (19.6)
	Restricted leaflet motion (n=46)	25 (54.3)	61 (17.6)
	Excessive leaflet tip motion (n=86)	31 (36.0)	55 (17.9)
	Mitral stenosis (n=3)	3 (100)	83 (21.3)
	Any regurgitation (n=338)	82 (24.2)	4 (7.3)
	Regurgitation seen in 2 views (n=156)	42 (26.9)	44 (18.6)
	Jet length ≥ 2 cm [‡] (n=298)	73 (24.5)	16 (16.8)
	Velocity ≥ 3 m/s for 1 envelope [§] (n=181)	49 (27.1)	37 (17.5)
	Pansystolic jet (color Doppler) (n=122)	23 (18.9)	63 (23.3)
Aortic valve, n (%) [*]	Irregular or focal thickening (n=17)	7 (41.2)	79 (21.0)
	Coaptation defect (n=4)	2 (50.0)	84 (21.6)
	Restricted leaflet motion (n=1)	1 (100)	85 (21.7)
	Leaflet prolapse (n=4)	2 (50)	84 (21.6)
	Any regurgitation (n=92)	17 (18.5)	69 (22.9)
	Regurgitation seen in 2 views (n=86)	16 (18.6)	70 (22.8)
	Jet length ≥ 1 cm [‡] (n=61)	12 (19.7)	74 (22.3)
	Velocity ≥ 3 m/s in early diastole [§] (n=48)	10 (20.8)	76 (22.0)
	Pandistolic jet (color Doppler) (n=22)	3 (13.6)	83 (22.4)

^{*}Congenital mitral valve or aortic valve abnormalities were excluded.

[†]Abnormal thickening of the anterior mitral valve leaflet ≥ 3 or >4 mm using harmonic imaging.

[‡]In at least 1 view.

[§]Measurements available with the Vivid-Q exams.

progression.¹⁴ Now, for the first time, the use of multinational longitudinal data, managed by different research groups, has allowed for a more robust assessment of the power of the score to predict outcomes.

Within the context of our study, it is important to evaluate the history and intent of the 2012 WHF consensus guideline⁹: the criteria were developed to enable international standardization of the echocardiographic diagnosis of subclinical RHD, thereby allowing for a direct between-study comparison on prevalence and severity of RHD in different settings. However, there are a number of recognized limitations about the utility and applicability of the variables comprising the WHF criteria as a single diagnostic tool. In particular, there are issues on the utility of the WHF criteria in lower-resourced screening environments where echocardiography providers usually undergo limited training such that they may not be able to capture all images required to measure all variables included in the WHF criteria. In addition, the WHF criteria lack adequate prognostic value, noticeably for clinical outcomes associated with the natural history of RHD.^{10,19} Furthermore, application of the complete WHF criteria requires fully functional echocardiography devices with spectral Doppler and such devices are frequently not available in settings where RHD is endemic. Considerable research efforts

have been made to address these limitations, leading to the development of simplified screening protocols with diagnosis based solely on valvular regurgitation and single-view acquisition.^{5,20,21} However practical these approaches may be, they exclude morphological abnormalities that may present early in the natural history of RHD, even if there is no existent valvular regurgitation. Aiming to improve the practical utility of the WHF criteria, some investigators regrouped the original set of variables, according to risk. For example, Beaton and colleagues stratified the original definite RHD category (which included prominent functional and morphological findings) into moderate and severe definite RHD, and demonstrated that these subcategories were prospectively associated with higher risk of progression.¹²

The simplified approach proposed in our study addresses some of these inherent limitations of the WHF criteria. The 2 functional and 3 morphological components included in the model are easily identified by the substantially less expensive handheld echocardiography machines, even in the absence of spectral Doppler.¹¹ Furthermore, it proposes a systematization of the risk subsets proposed by Beaton et al.¹² Finally, the reduced set of diagnostic variables also favor the point-of-care identification of individuals

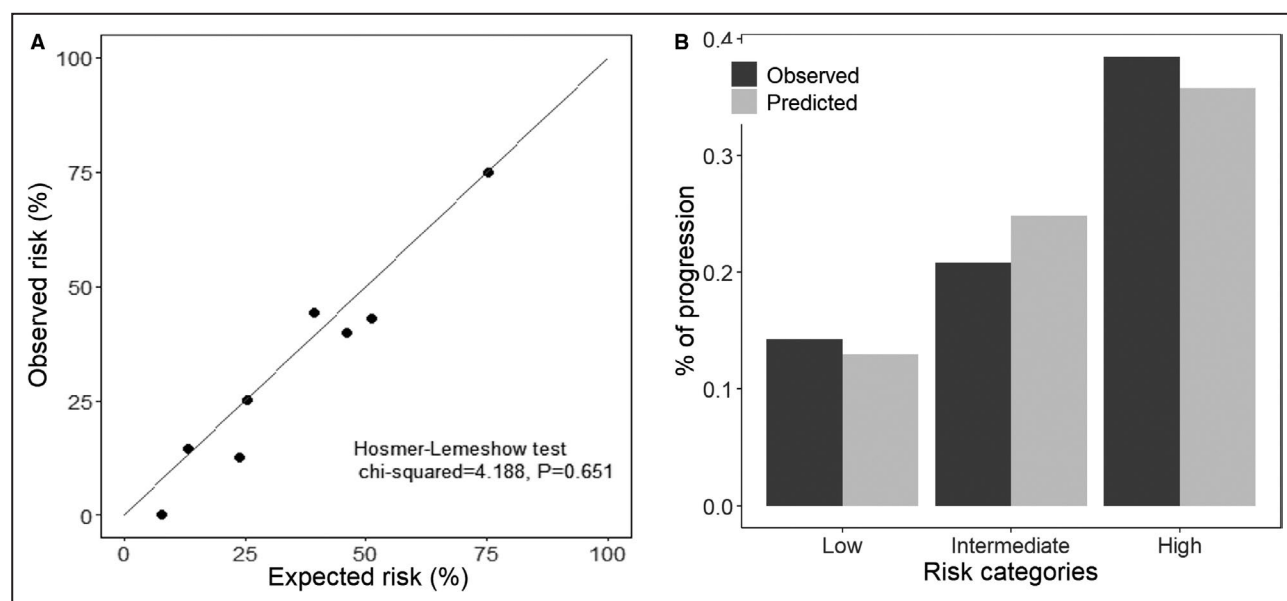


Figure 1. Plotted predicted-to-observed adverse outcome for the (A) total sample and for (B) each risk score group in the pooled population after recalibration.

with latent RHD by non-cardiologists with limited training. While for the overall set of criteria studies have reported intrareviewer agreement of 89% and between 71.4% and 94.1% in Brazil and Uganda, respectively, and inter-rater agreement of 92% and from 66.7% to 82.8%,^{6,22} when handheld devices and simplified criteria are used, diagnostic agreement between experts and trained non-experts is still considerable.^{23,24} Thus, the risk score classification proposed in this study has the potential to be even more reproducible.

Our pooled data highlight the differences in nature and severity of RHD between screening cohorts worldwide. There were major differences in baseline characteristics between cohorts, such as the higher proportions of definite cases in New Zealand¹⁶ and Malawi,¹⁵ while the cohort from Australia¹³ demonstrated worse risk and progression profiles in sharp contrast to the milder phenotypes observed in the Brazilian RHD cohort.¹⁴ There are, in addition, known epidemiological differences as well as screening approaches in these different settings. In Oceania, for example, there is a disproportionately high prevalence of RHD with a trend to severe phenotypes in indigenous and rural populations where screening programs are focused.¹³ This contrasts with the more general school and primary care-based screening approaches, guided by socioeconomic indexes, in Brazil and Malawi.^{7,14,15} This heterogeneity possibly led to a lower predictive accuracy of the echocardiographic score model when tested in new schoolchildren compared with that found in the development study, indicating the need for updating the model to adjust for the setting of the validation sample.^{17,18}

External Validation of the Echocardiographic Score Model

There are no generally accepted rules of how external validation should be performed, but it requires documentation of predictors and outcome values in new individuals. Model validation is not simply repeating the analytical steps applied in the development study in other individuals to assess whether the same predictors and weights are found. The purpose of validation is to

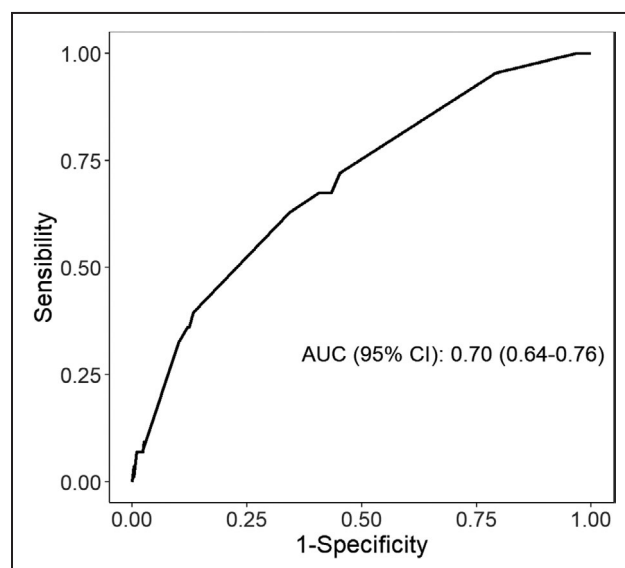


Figure 2. Receiver operator characteristic curve for echocardiographic score showing predicted probability of unfavorable echo outcome from the model (area under the curve=0.70). AUC indicates area under the curve.

Table 3. Echocardiographic Variables Used for Score After Recalibration of Intercept and Slope

Variable	β coefficient	SE	Z value	P value	Points
Mitral valve					
Anterior leaflet thickening	2.941	0.597	4.922	<0.0001	3
Excessive leaflet tip motion	3.102	0.543	5.716	<0.0001	3
Regurgitation jet length ≥ 2 cm	5.601	0.705	7.941	<0.0001	6
Aortic valve					
Irregular or focal thickening	4.460	0.970	4.597	<0.0001	4
Any regurgitation	4.794	0.718	6.679	<0.0001	5
Intercept	-10.221	0.893	-1.448	<0.001	...

combine the information captured in the original model with information from new individuals, which likely have improved transportability to other individuals.^{25,26}

In the present study, we aimed to perform a geographical validation of a previously developed model in other countries, and also in Brazil, including a population of schoolchildren that was not involved in the development study. Therefore, we applied the original model with its predictors and assigned weights (eg, regression coefficients), as estimated in the original study.¹⁷ Subsequently, we measured the predictors and outcome values in the new sample, and the model's predictive performance was quantified. As expected, we found a low predictive accuracy of the score model when tested in this population, what is likely to occur in geographical validations, in which a more stringent form of validation is used.²⁶ Then, the model was updated to adjust to the local circumstances of the validation sample (eg, environmental exposure, genetics, age at diagnosis, ethnicity, socioeconomic backgrounds, health system organization). After updating, the model's performance improved as measured by its discrimination and calibration. Therefore, rather than developing another prediction model, our aim was to evaluate whether the previously developed score predicts latent RHD outcome in distinct populations. A much better alternative to redeveloping new models for each new patient sample is to update existing prediction models and adjust or recalibrate them to local circumstances or setting of the validation sample at hand.²⁶ It is noteworthy that when a lower predictive accuracy is found, investigators tend to simply reject that model and develop or fit a new one, sometimes by completely repeating the entire selection of predictors. This leads to a loss of previous scientific information captured in the development study, which is counterintuitive to the notion that inferences and guidelines to enhance evidence-based medicine should be based on as much information as possible.^{17,25,26}

Several methods for updating prediction models have been proposed and evaluated.²⁵ Recalibration methods are attractive because of their stability,

which is related to the fact that few parameters are estimated. In our study, the intercept and the overall calibration slope were recalibrated, considering differences in RHD outcomes frequency between samples, and potential overfitting of regression coefficients in the original model. We used the simple updating methods with further improvement in the calibration of the model.¹⁸

Progression of Latent RHD

Our findings reinforce the variable evolution of latent RHD. In all our cohorts, over one third of cases regressed to normal, mostly comprising those with borderline disease. This was slightly lower than rates of regression reported in previous global reports (47%–67%).^{10,12} Within those individuals with borderline RHD, rates of progression to definite were overall low (13%) and this finding was impacted by the largest sample size from New Zealand. This finding is comparable with rates reported in a recent meta-analysis.¹⁹ Given these outcomes, in the absence of morphological findings or aortic involvement, the category of borderline RHD may be cautiously evaluated as an overlap with “normal” echocardiography. In contrast, for individuals with definite RHD, over one half remained in this category, in accordance with published pooled data,^{10,19} suggesting that this is a more clinically relevant category, as the persistence of associated morphological and functional abnormalities in 2 observations may be a surrogate for a continuing pathological process. For this reason, we considered remaining definite as an unfavorable echocardiographic outcome and, as previously suggested,^{11,12,14} higher risk scores were strongly associated with persistence over time.

It is important to note that we were unable to determine whether there was any association between use of secondary prophylaxis and progression or stability of disease as only a small proportion of our pooled cohort was prescribed prophylaxis (according to different local guidelines), and adherence was not systematically captured in some studies. For example, in Malawi,¹⁵ penicillin was recommended for all definite

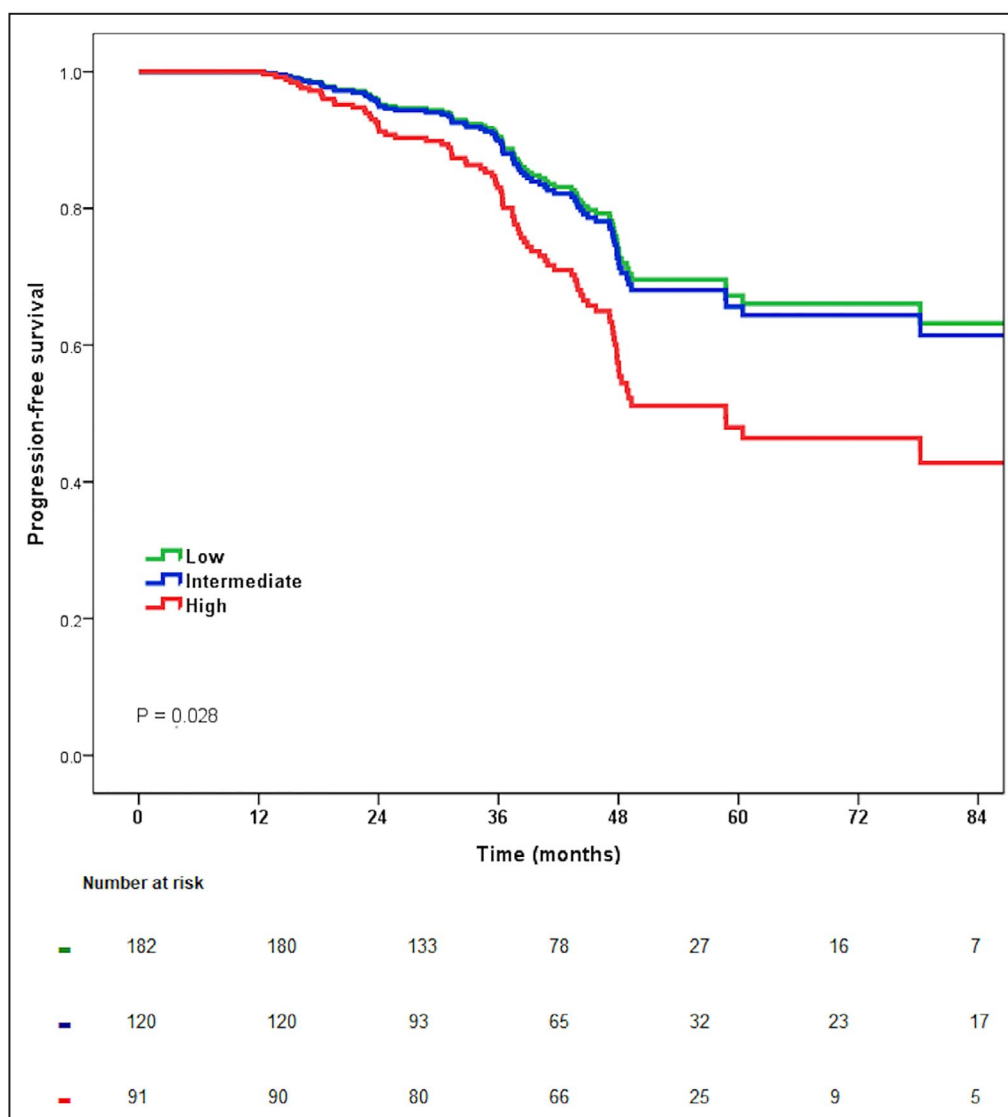


Figure 3. Cumulative survival free of unfavorable outcome in children with echocardiography-detected rheumatic heart disease according to 3 risk categories: low-risk (0–6 points), intermediate risk (7–9 points), and high-risk (≥ 10 points) of the simplified score, with results of the Cox proportional analysis.

cases, as it is in Australian RHD guidelines, whereas in Brazil it was left to the discretion of the attending cardiologist.¹⁴ Given the low rates of secondary prophylaxis, our outcomes can be practically interpreted as being representative of the “natural” evolution of latent RHD. Currently, a large randomized clinical trial is underway in Uganda examining the impact of penicillin prophylaxis in children and adolescents with screen-detected RHD (*Gwoko Adunu pa Lutino*; ClinicalTrials.gov no. NCT03346525). This study should bring definite answers in relationship to the association between secondary prophylaxis and progression of latent RHD.

Our model was better refined to predict echocardiographic outcomes in the highest risk category. In

contrast to the original outcomes-prediction study,¹¹ for our cohort there was overlap between the low and intermediate risk categories on the Kaplan–Meier curve. While this finding limits refinement, and may possibly arise because of the data heterogeneity, the ability of our model to identify the subgroup at highest risk of unfavorable outcome points towards its usefulness as a public health tool, and urges further investigation. In addition, a simpler and more accurate prediction of disease progression may improve the cost-effectiveness of population screening, as the prevalence of latent RHD detected in echocardiographic studies is >7 times higher than that detected by auscultation,²⁷ even though a smaller number, with high-risk profile, will benefit from

closer clinical monitoring and, presumably, prophylaxis to halt disease progression. Thus, accurate risk stratification at time of diagnosis potentially implies not only on redefinition of latent RHD epidemiology, but also on the approach of children and families.

Limitations

Our study has several limitations. First, follow-up was relatively short to determine clinical outcomes, while advanced sequelae of RHD may take decades to develop. Thus, the score must be further assessed as longer follow-ups are available. However, the investigation of early echocardiographic predictors of disease progression is valuable for risk stratification and planning of clinical follow-up. Second, we used screening diagnoses determined by the original echocardiography readers, without reassessment of images. Although all studies that contributed to our pooled data applied the WHF criteria, inter-reader variability may account for some heterogeneity of our data, especially in relationship to specific morphological findings. Nevertheless, we aimed to assess the application of the score in real life programs, in which central reads are rarely available. Third, imaging acquisition strategies differed among the different included studies. For example, some used handheld echocardiographic devices while others used standard portable devices. Furthermore, in some studies baseline scans were performed by experts while in others they were performed by non-physicians. For this reason, the original score consisted of variables that could be measured without spectral Doppler and, again here, task-shifting is a frequent approach in real-life screening studies. Fourth, the original intercept and slope for the risk model required recalibration, possibly because of the above-mentioned interstudy heterogeneity. This suggests that a fine tuning of the model may be necessary for specific populations and regions, as it is not possible to infer whether the score would be equally effective in different countries. Fifth, no clinical variables – such as sex or age – were incorporated into the model, once we opted to validate the original score, composed exclusively by echocardiographic findings. Lastly, no inferences about the impact of penicillin on subclinical RHD can be drawn from our data. Despite these limitations, to the best of our knowledge, this is the only study evaluating this novel simplified score in one of the largest cohorts of subclinical RHD. A good accuracy to identify a high-risk subgroup suggests that the score may be a useful stratification tool with potential impact of patient management on secondary prophylaxis.

CONCLUSIONS

The simple risk score developed by Nunes and colleagues provided an accurate prediction of RHD status

at 3-year follow-up in a pooled multinational cohort. The updated score showed good performance in different screening populations worldwide, and has potential value for risk stratification of latent RHD.

APPENDIX

PROVAR (Programa de Rastreamento da Valvopatia Reumática) investigators:

Alison T. Reese, Andrea Z. Beaton, Antonio Luiz P. Ribeiro, Breno D. F. Rezende, Bruno R. Nascimento, Clara L. Fraga, Craig Sable, Domingos Sávio G. Ferreira-Filho, Frederico V. B. Macedo, Juliane Franco, Kaciane K. B. Oliveira, Luiza P. A. Santos, Márcia M. Barbosa, Maria Carmo P. Nunes, Rodrigo T. L. Rocha, Sanny Cristina C. Faria.

ARTICLE INFORMATION

Received March 12, 2021; accepted August 13, 2021.

Affiliations

Serviço de Cardiologia e Cirurgia Cardiovascular e Centro de Telessaúde, Hospital das Clínicas da Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil (B.R.N., M.C.P.N., E.M.L., A.L.P.R.); Departamento de Clínica Médica, Faculdade de Medicina da Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil (B.R.N., M.C.P.N., A.L.P.R.); Texas Children's Hospital, Houston, TX (A.E.S., P.N.K.); Department of Paediatric and Congenital Cardiac Services, Starship Children's Hospital, Auckland, New Zealand (N.W., E.T.); Faculty of Health and Medicine, University of Newcastle, Callaghan, New South Wales, Australia (M.G.W.R., G.P.M.); Western Clinical School, University of Melbourne, Melbourne, Australia (G.P.M.); Cardiology, Children's National Health System, Washington, DC (C.S.); and The Heart Institute, Cincinnati Children's Hospital Medical Center, University of Cincinnati School of Medicine, Cincinnati, OH (A.Z.B.).

Acknowledgments

Conception and design of the research: Nascimento, BR, Sable, C, Beaton, AZ, Nunes, MCP, Ribeiro, AL; Acquisition of data: Nascimento, BR, Nunes, MCP, Sanyahumbi, AE, Wilson, N, Tilton, E, Rémond, MGW, Maguire, GP, Kazembe, PN; Analysis and interpretation of data: Nascimento, BR, Nunes, MCP, Beaton, AZ, Malveira, E; Statistical analysis: Malveira, E, Nascimento, BR, Nunes, MCP; Obtaining financing: Beaton, AZ, Sable, C, Nascimento, BR; Writing of the manuscript: Nascimento, BR, Sable, C, Nunes, MCP; Beaton, AZ. Critical revision of the manuscript for intellectual content: All authors; Authors responsible for the overall content as guarantors: Nascimento, BR, Beaton, AZ, Sable, C, Nunes, MCP.

Sources of Funding

The PROVAR+ investigators would like to thank Edwards Lifesciences Foundation® for supporting and funding the primary care screening program (PROVAR+) in Brazil, General Electric Healthcare® for providing echocardiography equipment and WiRed Health Resources for providing online curriculum on heart disease and echocardiography. The Telehealth Network of Minas Gerais was funded by the State Government of Minas Gerais, by its Health Department (Secretaria de Estado da Saúde de Minas Gerais) and FAPEMIG (Fundação de Amparo à Pesquisa de Minas Gerais), and by the Brazilian Government, including the Health Ministry and the Science and Technology Ministry and its research and innovation agencies, CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) e FINEP (Financiadora de Estudos e Projetos). Dr Ribeiro was supported in part by CNPq (Bolsa de produtividade em pesquisa, 310679/2016-8) and by FAPEMIG (Programa Pesquisador Mineiro, PPM-00428-17). Dr Nascimento was supported in part by CNPq (Bolsa de produtividade em pesquisa, 312382/2019-7), by the Edwards Lifesciences Foundation (Every Heartbeat Matters Program 2020) and by FAPEMIG (grant APQ-000627-20). Medical students received scholarships from the National Institute of Science and Technology for Health Technology Assessment (IATS, project: 465518/2014-1).

Disclosures

None.

Supplementary Material

Table S1

Figures S1–S3

REFERENCES

- James SL, Abate D, Abate KH, Abay SM, Abbafati C, Abbasi N, Abbastabar H, Abd-Allah F, Abdela J, Abdelalim A, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet*. 2018;392:1789–1858. doi: 10.1016/S0140-6736(18)32279-7
- Roth GA, Mensah GA, Johnson CO, Addolorato G, Ammirati E, Baddour LM, Barengo NC, Beaton AZ, Benjamin EJ, Benziger CP, et al. Global burden of cardiovascular diseases and risk factors, 1990–2019: update from the GBD 2019 study. *J Am Coll Cardiol*. 2020;76:2982–3021. doi: 10.1016/j.jacc.2020.11.010
- Kotit S, Said K, ElFaramawy A, Mahmoud H, Phillips DIW, Yacoub MH. Prevalence and prognostic value of echocardiographic screening for rheumatic heart disease. *Open Heart*. 2017;4:e000702. doi: 10.1136/openhrt-2017-000702
- Watkins DA, Beaton AZ, Carapetis JR, Karthikeyan G, Mayosi BM, Wyber R, Yacoub MH, Zuhlke LJ. Rheumatic heart disease worldwide: JACC scientific expert panel. *J Am Coll Cardiol*. 2018;72:1397–1416. doi: 10.1016/j.jacc.2018.06.063
- Diamantino A, Beaton A, Aliku T, Oliveira K, Oliveira C, Xavier L, Perlman L, Okello E, Nascimento B, Ribeiro ARP, et al. A focussed single-view hand-held echocardiography protocol for the detection of rheumatic heart disease. *Cardiol Young*. 2018;28:108–117. doi: 10.1017/S1047951117001676
- Nascimento BR, Beaton AZ, Nunes MCP, Diamantino AC, Carmo GAL, Oliveira KKB, Oliveira CM, Meira ZMA, Castilho SRT, Lopes ELV, et al. Echocardiographic prevalence of rheumatic heart disease in Brazilian schoolchildren: data from the PROVAR study. *Int J Cardiol*. 2016;219:439–445. doi: 10.1016/j.ijcard.2016.06.088
- Nascimento BR, Sable C, Nunes MCP, Diamantino AC, Oliveira KKB, Oliveira CM, Meira ZMA, Castilho SRT, Santos JPA, Rabelo LMM, et al. Comparison between different strategies of rheumatic heart disease echocardiographic screening in Brazil: data from the PROVAR (Rheumatic Valve Disease Screening Program) Study. *J Am Heart Assoc*. 2018;7:e008039. doi: 10.1161/JAHA.117.008039
- Beaton A, Aliku T, Okello E, Lubega S, McCarter R, Lwabi P, Sable C. The utility of handheld echocardiography for early diagnosis of rheumatic heart disease. *J Am Soc Echocardiogr*. 2014;27:42–49. doi: 10.1016/j.echo.2013.09.013
- Reményi BO, Wilson N, Steer A, Ferreira B, Kado J, Kumar K, Lawrenson J, Maguire G, Marjion E, Mirabel M, et al. World Heart Federation criteria for echocardiographic diagnosis of rheumatic heart disease—an evidence-based guideline. *Nat Rev Cardiol*. 2012;9:297–309. doi: 10.1038/nrcardio.2012.7
- Zühlke L, Engel ME, Lemmer CE, Van de Wall M, Nkepu S, Meiring A, Bestawros M, Mayosi BM. The natural history of latent rheumatic heart disease in a 5 year follow-up study: a prospective observational study. *BMC Cardiovasc Disord*. 2016;16:46. doi: 10.1186/s12872-016-0225-3
- Nunes MCP, Sable C, Nascimento BR, Lima EMD, da Silva JLP, Diamantino AC, Oliveira KKB, Okello E, Aliku T, Lwabi P, et al. Simplified echocardiography screening criteria for diagnosing and predicting progression of latent rheumatic heart disease. *Circ Cardiovasc Imaging*. 2019;12:e007928. doi: 10.1161/CIRCIMAGING.118.007928
- Beaton A, Aliku T, Dewyer A, Jacobs M, Jiang J, Longenecker CT, Lubega S, McCarter R, Mirabel M, Mirembe G, et al. Latent rheumatic heart disease: identifying the children at highest risk of unfavorable outcome. *Circulation*. 2017;136:2233–2244. doi: 10.1161/CIRCULATIONAHA.117.029936
- Remond M, Atkinson D, White A, Brown A, Carapetis J, Remenyi B, Roberts K, Maguire G. Are minor echocardiographic changes associated with an increased risk of acute rheumatic fever or progression to rheumatic heart disease? *Int J Cardiol*. 2015;198:117–122. doi: 10.1016/j.ijcard.2015.07.005
- Bechtluft BM, Nascimento BR, Sable C, Fraga CL, Barbosa MM, Reis SD, Diamantino AC, Meira ZMA, Castilho SRT, Arantes NF, et al. Validation of a simplified score for predicting latent rheumatic heart disease progression using a prospective cohort of Brazilian schoolchildren. *BMJ open*. 2020;10:e036827. doi: 10.1136/bmjopen-2020-036827
- Sanyahumbi A, Beaton A, Guffey D, Hosseiniour MC, Karlsten M, Minard CG, Penny DJ, Sable CA, Kazembe PN. Two-year evolution of latent rheumatic heart disease in Malawi. *Congenit Heart Dis*. 2019;14:614–618. doi: 10.1111/chd.12756
- Cramp G, Stonehouse M, Webb R, Chaffey-Aupouri G, Wilson N. Undetected rheumatic heart disease revealed using portable echocardiography in a population of school students in Tairarwhiti, New Zealand. *N Z Med J*. 2012;125:53–64.
- Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. *Eur Heart J*. 2014;35:1925–1931. doi: 10.1093/eurheartj/ehu207
- Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, Pencina MJ, Kattan MW. Assessing the performance of prediction models: a framework for some traditional and novel measures. *Epidemiology*. 2010;21:128. doi: 10.1097/EDE.0b013e3181c30fb2
- Noubiap JJ, Agbor VN, Bigna JJ, Kaze AD, Nyaga UF, Mayosi BM. Prevalence and progression of rheumatic heart disease: a global systematic review and meta-analysis of population-based echocardiographic studies. *Sci Rep*. 2019;9:17022. doi: 10.1038/s41598-019-53540-4
- Remenyi B, Davis K, Draper A, Bayley N, Paratz E, Reeves B, Appelbe A, Wheaton G, da Silva Almeida IT, dos Santos J, et al. Single parasternal-long-axis-view-sweep screening echocardiographic protocol to detect rheumatic heart disease: a prospective study of diagnostic accuracy. *Heart lung circ*. 2020;29:859–866. doi: 10.1016/j.hlc.2019.02.196
- Engelman D, Kado JH, Remenyi B, Colquhoun SM, Carapetis JR, Donath S, Wilson NJ, Steer AC. Focused cardiac ultrasound screening for rheumatic heart disease by briefly trained health workers: a study of diagnostic accuracy. *Lancet Glob health*. 2016;4:e386–e394. doi: 10.1016/S2214-109X(16)30065-1
- Beaton A, Lu JC, Aliku T, Dean P, Gaur L, Weinberg J, Godown J, Lwabi P, Mirembe G, Okello E, et al. The utility of handheld echocardiography for early rheumatic heart disease diagnosis: a field study. *Eur Heart J Cardiovasc Imaging*. 2015;16:475–482. doi: 10.1093/ehjci/jeu296
- Beaton A, Nascimento BR, Diamantino AC, Pereira GTR, Lopes ELV, Miri CO, Bruno KKO, Chequer G, Ferreira CG, Lafeta LCX, et al. Efficacy of a standardized computer-based training curriculum to teach echocardiographic identification of rheumatic heart disease to nonexpert users. *Am J Cardiol*. 2016;117:1783–1789. doi: 10.1016/j.amjcard.2016.03.006
- Ploutz M, Lu JC, Scheel J, Webb C, Ensing GJ, Aliku T, Lwabi P, Sable C, Beaton A. Handheld echocardiographic screening for rheumatic heart disease by non-experts. *Heart*. 2016;102:35–39. doi: 10.1136/heartjnl-2015-308236
- Debray TP, Vergouwe Y, Koffijberg H, Nieboer D, Steyerberg EW, Moons KG. A new framework to enhance the interpretation of external validation studies of clinical prediction models. *J Clin Epidemiol*. 2015;68:279–289. doi: 10.1016/j.jclinepi.2014.06.018
- Moons KG, Kengne AP, Grobbee DE, Royston P, Vergouwe Y, Altman DG, Woodward M. Risk prediction models: II. External validation, model updating, and impact assessment. *Heart*. 2012;98:691–698. doi: 10.1136/heartjnl-2011-301247
- Rothenbuhler M, O'Sullivan CJ, Stortecky S, Stefanini GG, Spitzer E, Estill J, Shrestha NR, Keiser O, Juni P, Pilgrim T. Active surveillance for rheumatic heart disease in endemic regions: a systematic review and meta-analysis of prevalence among children and adolescents. *Lancet Glob health*. 2014;2:e717–e726. doi: 10.1016/S2214-109X(14)70310-9

SUPPLEMENTAL MATERIAL

Table S1. Baseline echocardiographic findings for patients with progression, stabilization and regression of rheumatic heart disease at 2-year follow-up.

Valve:	Variable:	Progressed: Borderline Definite (N=42)	Remained to Definite (N=44)	Regressed / stable (borderline) / other (N=307)
Mitral valve, N (%):	Anterior leaflet thickening	14 (33.3)	36 (81.8)	115 (37.5)
	Chordal thickening	1 (2.4)	12 (27.3)	8 (2.6)
	Restricted leaflet motion	4 (9.5)	21 (47.7)	21 (6.8)
	Excessive leaflet tip motion	7 (16.7)	24 (54.5)	53 (17.3)
	Mitral stenosis	0	3 (6.8)	0
	Any regurgitation	40 (95.2)	42 (95.5)	256 (83.4)
	Regurgitation seen in 2 views	12 (28.6)	30 (68.2)	114 (37.1)
	Jet length ≥ 2 cm†	34 (81.0)	39 (88.6)	225 (73.3)
	Velocity ≥ 3 m/s for 1 envelope§	17 (40.5)	32 (72.7)	132 (43.0)
	Pansystolic jet (color Doppler)	15 (35.7)	8 (18.2)	99 (32.2)
Aortic valve, N (%):	Irregular or focal thickening	1 (2.4)	6 (13.6)	10 (3.3)
	Coaptation defect	0	2 (4.5)	2 (0.7)
	Restricted leaflet motion	0	1 (2.3)	0
	Leaflet Prolapse	0	2 (4.5)	2 (0.7)
	Any regurgitation	6 (14.3)	11 (25.0)	75 (24.4)
	Regurgitation seen in 2 views	6 (14.3)	10 (22.7)	70 (22.8)
	Jet length ≥ 1 cm†	1 (2.4)	11 (25.0)	49 (16.0)
	Velocity ≥ 3 m/s in early diastole§	3 (7.1)	7 (15.9)	38 (12.4)
	Pandistolic jet (color Doppler)	0	3 (6.8)	19 (6.2)

*Congenital mitral valve or aortic valve abnormalities were excluded. †Abnormal thickening of the anterior mitral valve leaflet ≥ 3 or >4 mm using harmonic imaging. ‡In at least 1 view. §Measurements available with the Vivid-Q exams.

Figure S1. Diagram with criteria for definition of favorable and unfavorable outcome for patients with latent Rheumatic Heart Disease.

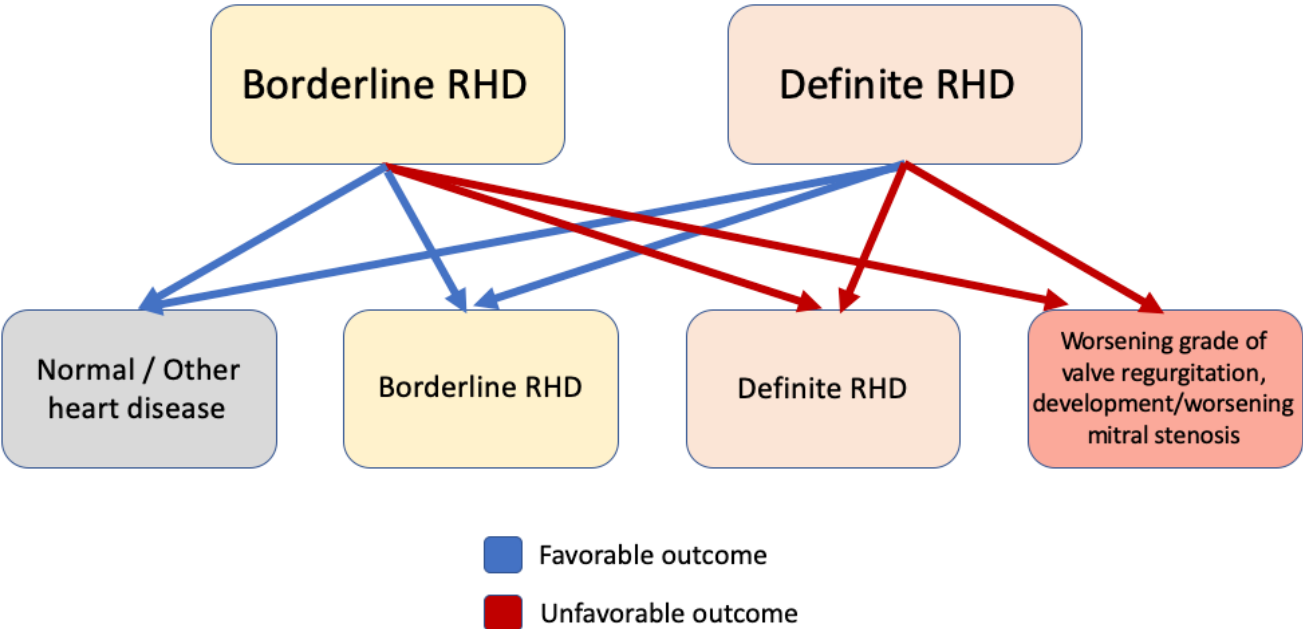


Figure S2. Plotted predicted-to-observed adverse outcome for the total sample in the original score, prior to recalibration.

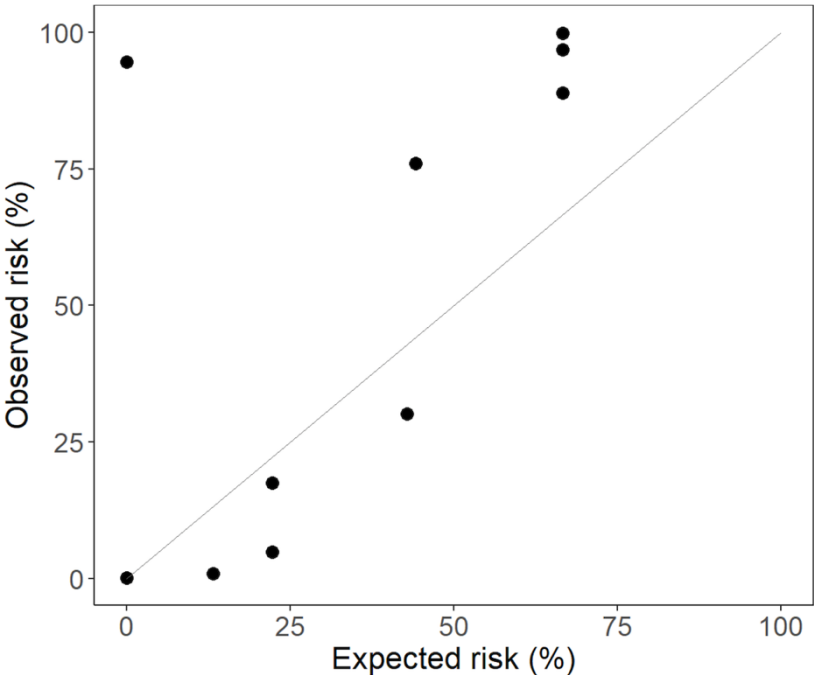


Figure S3. Boxplot with comparison of predicted risks of unfavorable outcome of latent Rheumatic Heart Disease in each risk category in the pooled population.

