International Journal of Cardiology 291 (2019) 112–118

Contents lists available at ScienceDirect

International Journal of Cardiology

journal homepage: www.elsevier.com/locate/ijcard

The World Heart Federation criteria raise the threshold of diagnosis for mild rheumatic heart disease: Three reviewers are better than one

Nicola Culliford-Semmens [a,1], Ross Nicholson [b,1], Elizabeth Tilton [a,1], John Stirling [a,1], Karishma Sidhu [a,1], Rachel Webb [a,1], Nigel Wilson [a,*,1]

[a] Green Lane Paediatric and Congenital Cardiac Services, Starship Children's Hospital, Auckland, New Zealand
[b] Department of Paediatrics and Child Health, Kidz First Children's Hospital, Auckland, New Zealand

ARTICLE INFO

Article history:
Received 13 March 2018
Received in revised form 18 February 2019
Accepted 25 February 2019
Available online 2 March 2019

Keywords:
Rheumatic heart disease
Echocardiography
Screening

ABSTRACT

Background: The World Heart Federation (WHF) criteria, published in 2012, provided an evidence-based guideline for the minimal diagnosis of echocardiographically-detected RHD.
Primary aim of the study was to determine whether use of the WHF criteria altered the threshold for the diagnosis of echocardiographically-detected RHD compared with the previous WHO/NIH criteria. A secondary aim was to explore the utility of a three reviewer reporting system compared to a single or two reviewer reporting structure.
Methods: 144 de-identified echocardiograms (RHD, congenital valvar abnormality, physiological valvar regurgitation) were independently reported using the WHF criteria by two reviewers blinded to the previous WHO/NIH diagnosis. If there was discordance between the two reviewers, a third cardiologist independently performed a tie-breaker review.
Results: There was a 21% reduction of cases classified as RHD using the WHF criteria compared to the modified WHO/NIH criteria (68 cases compared to 86, $p = 0.04$). There was a 60% consensus across the different diagnostic categories with 2 reviewers, 89% majority agreement with 3 reviewers. 11% required an open label discussion. There was moderate agreement between 2 reviewers for any RHD, kappa 0.57 (CI 0.44–0.70), with no significant difference in agreement between the different categories.
Conclusion: The WHF criteria have raised the threshold for the diagnosis of RHD compared to the WHO/NIH criteria. However, inter-reporter variability of the WHF criteria is high. A three reviewer system is likely more accurate than a single or two reporter system for the diagnosis of mild RHD. This has resource implications for echocardiographic screening programmes.

1. Introduction

Rheumatic heart disease (RHD) is an important global health problem. Revised Global Burden of Diseases estimates are 33 million prevalent cases of RHD causing 9 million Disability-Adjusted Life Years lost and 270,000 deaths per year [1]. Since 2007, numerous echocardiography based case finding programmes for subclinical RHD have been undertaken in endemic regions. Several of the first reports used World Health Organisation/National Institute of Health (WHO/NIH) criteria, which had been established by a joint WHO/NIH working party at a meeting in Cairns in 2005, and were initially only available on a web site, and not formally published until 2010 [2]. Use of these criteria often resulted in very high prevalence rates of RHD, 38.4/1000 in Tonga [3], 48/1000 in Nicaragua [4], 51/1000 in India [5], 55/1000 [6] and 43.7/1000 [7] in New Zealand [2]. These high rates of previously undetected RHD, up to 5% of children in high prevalence regions, led some to question whether there was over diagnosis of RHD using echocardiography [8].

The World Heart Federation (WHF) criteria, published in 2012, provided the first evidence-based international consensus guideline for the diagnosis of echocardiographically-detected RHD. These criteria are based on the surgical, pathological and echocardiographic descriptions of RHD [9]. The WHF criteria aimed to define the minimal criteria to diagnose RHD as opposed to advanced RHD where the echocardiographic features are usually unequivocally diagnostic [9]. Although the WHF guidelines have been widely adopted since 2012, it is uncertain if their use has altered the threshold for the diagnosis of RHD compared to previous criteria.

The primary aim of this study was to determine whether the use of the WHF criteria has changed the threshold for the diagnosis

* Corresponding author at: Green Lane Paediatric and Congenital Cardiac Services, Starship Children's Hospital, Private Bag 92024, Victoria Street West, Auckland 1142, New Zealand.
E-mail address: NigelW@adhb.govt.nz (N. Wilson).
[1] These authors take responsibility for all aspects of the reliability and freedom from bias of the data presented and their discussed interpretation.

https://doi.org/10.1016/j.ijcard.2019.02.058
0167-5273/

Table 1
Modified WHO/NIH criteria. [2]

	RHD			Murmur	Doppler	Morphological
	Definite	Probable	Possible			
Pathological murmur of mitral and/or aortic regurgitation	+	−	−	+	−	−
Echocardiographic features of rheumatic heart disease: Pathological grade mitral and/or aortic regurgitation; and valve morphological changes of mitral and/or aortic valve consistent with rheumatic heart disease; or mitral stenosis – mean gradient of 4 or more millimetres of mercury	+	−	−	−	−	+
Acute rheumatic fever with echocardiographic features of rheumatic heart disease.	+	−	−	−	−	−
No murmur[a]	−	+	−	−	−	−
Echocardiographic features of rheumatic heart disease: Pathological grade mitral and/or aortic regurgitation; and morphological changes of mitral and/or aortic valve consistent with rheumatic heart disease	−	+	−	−	−	−
No murmur	−	−	+	−	−	−
Echocardiographic features of rheumatic heart disease: Pathological grade mitral and/or aortic regurgitation without morphological changes of rheumatic heart disease; or morphological changes of mitral and/or aortic valve consistent with rheumatic heart disease without pathological grade mitral or aortic regurgitation	−	−	+	−	−	−
Pathological grade regurgitation was defined as a colour Doppler jet meeting all the minimum criteria below [8,11]	−	−	+	−	+	−
Mitral regurgitation:	−	−	−	−	+	−
• Substantial colour jet seen in two or more planes						
• Extending >2 cm beyond mitral valve leaflets in at least one plane						
• Holosystolic with well-defined spectral envelope on pulse wave/continuous wave Doppler						
• High velocity >3.5 m per second						
Aortic regurgitation:	−	−	−	−	+	−
• Substantial colour jet seen in two or more planes						
• Extending >1 cm beyond aortic valve leaflets in at least one plane						
• Holodiastolic with well-defined spectral envelope on pulse wave or continuous wave Doppler						
• High velocity >3.5 m per second						
Care was taken to exclude congenital valve pathology such as bicuspid aortic valve, dilated aortic sinuses, and congenital mitral valve prolapse. To fulfil a diagnosis of rheumatic heart disease, at least one of the following morphological features of rheumatic heart disease was required.						
Mitral valve morphological features:	−	−	−	−	−	+
• Thickening** of anterior mitral valve leaflet mid-point or tip						
• Fixed elbow – dog leg – deformity of anterior mitral valve leaflet mid-point or tip						
• Prolapse of anterior mitral valve leaflet – in the absence of clinical features of congenital mitral valve prolapse						
• Thickening** or retraction of posterior mitral valve leaflet						
• Thickening, tethering, retraction, or rupture to chordae of anterior mitral valve leaflet						
Aortic valve morphological features:	−	−	−	−	−	+
• Thickening of aortic valve leaflets or closure line in parasternal short-axis views						
• Rolled aortic valve leaflet edges						
• Overt prolapse of aortic valve leaflets						
• Coaptation defect of aortic valve leaflets						

2012 WHF criteria for echocardiographic diagnosis of RHD for individuals aged≤20 years [9] (*Reproduced with permission from Remenyi B, et al. Nat Rev Cardiol. 2012;9(5):297-309 - Box 1.*)

Definite RHD (**either A**, **B**, **C**, **or D**):
A) Pathological MR and at least two morphological features of RHD of the MV
B) MS mean gradient ≥4 mm Hg[a]
C) Pathological AR and at least two morphological features of RHD of the AV[b]
D) Borderline disease of both the AV and MV[c]

Borderline RHD (**either A**, **B**, **or C**):
A) At least two morphological features of RHD of the MV without pathological MR or MS
B) Pathological MR
C) Pathological AR

Normal echocardiographic findings (**all of A**, **B**, **C**, **and D**):
A) MR that does not meet all four Doppler echocardiographic criteria (physiological MR)
B) AR that does not meet all four Doppler echocardiographic criteria (physiological AR)
C) An isolated morphological feature of RHD of the MV (for example, valvar thickening) without any associated pathological stenosis or regurgitation
D) Morphological feature of RHD of the AV (for example, valvar thickening) without any associated pathological stenosis or regurgitation

Pathological mitral regurgitation (**All four Doppler criteria must be met**)	**Pathological aortic regurgitation** (**All four Doppler criteria must be met**)
Seen in two views	Seen in two views
In at least one view, jet length ≥2 cm	In at least one view, jet length ≥1 cm
Velocity ≥3 m/s for one complete envelope	Velocity ≥3 m/s in early diastole
Pan-systolic jet in at least one envelope	Pan-diastolic jet in at least one envelope
Morphologic features of the MV	**Morphologic features of the AV**
AMVL thickening ≥3 mm	Irregular or focal thickening
Chordal thickening	Coaptation defect
Restricted leaflet motion	Restricted leaflet motion
Excessive leaflet tip motion during systole	Prolapse

[a] Congenital MV anomalies must be excluded.
[b] Bicuspid AV, dilated aortic root, and hypertension must be excluded.
[c] Combined AR and MR in high prevalence regions and in the absence of congenital heart disease is regarded as rheumatic. Abbreviations: AR, aortic regurgitation; AV, aortic valve; MR, mitral regurgitation; MS, mitral stenosis; MV, mitral valve; RHD, rheumatic heart disease; WHF, World Heart Federation.

Table 2
WHO/NIH classification vs. WHF classification.

				WHF (2012)			
				RHD (n = 68)*		Non RHD (n = 76)	
				Borderline	Definite	Normal	Congenital
WHO / NIH (2006)	RHD (n=86)	Possible	53	21	8	23	1
		Probable	22	12	3	4	3
		Definite	11	5	5	1	
	Non RHD (n=58)	Normal	37	8	1	28	
		Congenital	21	4	1	5	11
			144	50	18	61	15

*Number and type of valve disease ($n = 68$)
Mitral borderline (38) definite (9).
Aortic borderline (12) definite (2).
Mitral and aortic definite (7).
Physiologic mitral regurgitation normal (49).
Physiologic aortic regurgitation normal (4).
Both aortic and mitral physiologic regurgitation (8).
Congenital mitral disease (4).
Congenital aortic disease (10) – 7 bicuspid valves and 2 dilated aortic roots, 1 not further specified.
Both congenital aortic and mitral disease (1).

of echocardiographically detected RHD compared with the WHO/NIH criteria.

A secondary aim was to assess the reproducibility of the WHF criteria using a three reviewer reporting system compared to a single reviewer reporting system for the diagnosis of mild RHD. The latter has been the norm for RHD echo case detection programmes in many resource-limited settings, unlike New Zealand where a multi-reviewer reporting structure has been utilised [6,7,10].

2. Methods

144 abnormal echocardiograms were selected from children who had undergone school-based echocardiography screening in two regions of New Zealand from 2007 to 2009. In this time period a total of 1827 students, 1142 from the South Auckland region [6] and 685 from the Tairawhiti region [7], underwent echocardiography. Only two children reported a past history of sore joints that could have been an episode of previous ARF.

Of the 144 cases, 108 had mitral valve abnormalities, 29 had aortic valve abnormalities and 7 had both mitral and aortic valve abnormalities. 69 (48%) of the 144 were male, median age 12, range 6–17 years.

Table 1 shows the WHO-NIH classification and definitions of RHD [2,6]. 86 of 144 cases had been classified as RHD by the WHO-NIH classification, comprising 53 possible RHD cases, 22 probable RHD cases and 11 definite RHD cases. By definition, the 11 definite cases had a typical murmur of mitral regurgitation. 95% (82/86) of those with RHD were mild cases, and 5% had moderate or severe RHD.

21 of 144 cases had been diagnosed as a congenital abnormality of the mitral or aortic valve. The remaining 37 had mitral regurgitation (MR) or aortic regurgitation (AR) deemed to be at upper limit of the physiological range by the earlier WHO/NIH analysis. At that time, the final diagnosis had been recorded by the reporting cardiologists, but the individual colour-Doppler features that constituted pathological or physiological valvar regurgitation were not recorded.

In the current study these 144 echocardiograms were re-analysed according to the WHF criteria (Table 1) [9]. The 144 echocardiograms were de-identified and reviewers were blinded to the original WHO/NIH classification. Physiological mitral regurgitation is reported with structurally normal valves in approximately 15% of children and in a higher proportion in adults, whereas physiological aortic regurgitation occurs in only 1% of children [12]. There can be difficulty discriminating the upper limit physiological MR or AR from those with pathological regurgitation. Thus, a number of cases previously labelled upper limit physiological MR or AR were included as a spiked 'normal' cohort rather than including cases without any MR or AR as the latter do not pose diagnostic uncertainty. Each echocardiogram was independently reviewed by two paediatric cardiologists. Echocardiograms were reviewed using the Prosolv viewing application to enable reviewers to perform their own measurements of anterior mitral valve leaflet thickness and regurgitant jet lengths. Reviewers completed a standardised report form for each echocardiogram, reporting the presence or absence of pathological mitral and aortic regurgitation (including the colour - Doppler features that constituted pathological or physiological valvar regurgitation for that diagnosis), the presence or absence of morphological features of RHD as well as the overall classification (definite RHD/borderline RHD/congenital/normal).

If there was discordance between the two reviewers regarding the overall diagnostic category for each echocardiogram, a third reviewer then independently, without knowledge of the other two reviewers' conclusion, reviewed the echocardiogram and completed the standardised report to determine a majority agreement. If, after review by three cardiologists, a majority opinion was not reached as each of the 3 reviewers had assigned a separate category of the four available categories, a combined open review of the echocardiogram was undertaken by the three reviewers in order to reach majority agreement. During this process the three reviewers remained blinded to the original WHO/NIH diagnosis.

RHD Prevalence rates were re-calculated for the WHF criteria using the total number of screened cases as the denominator.

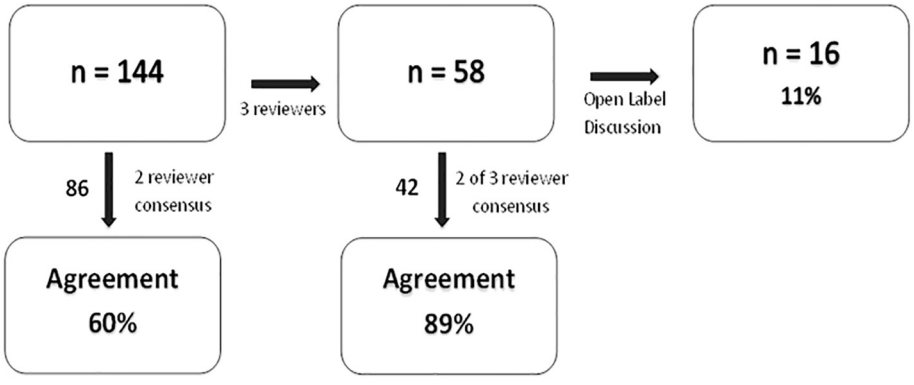
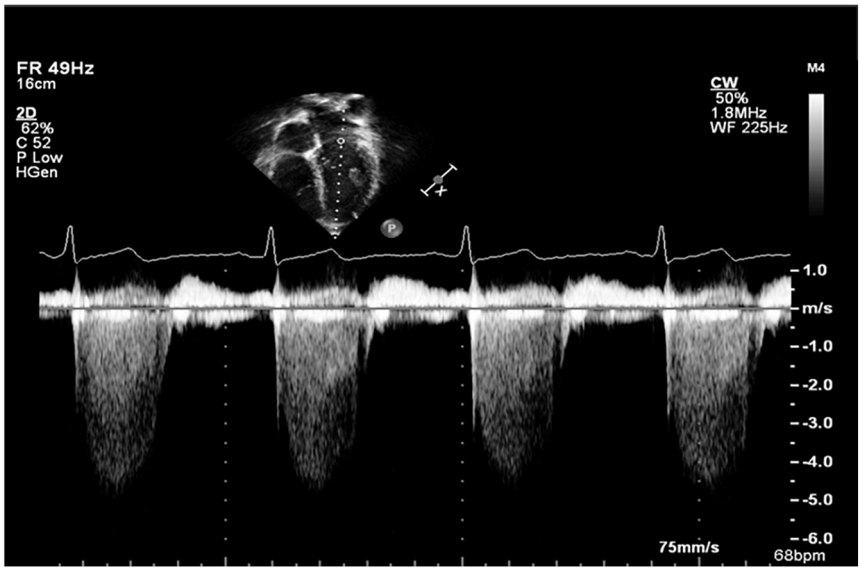
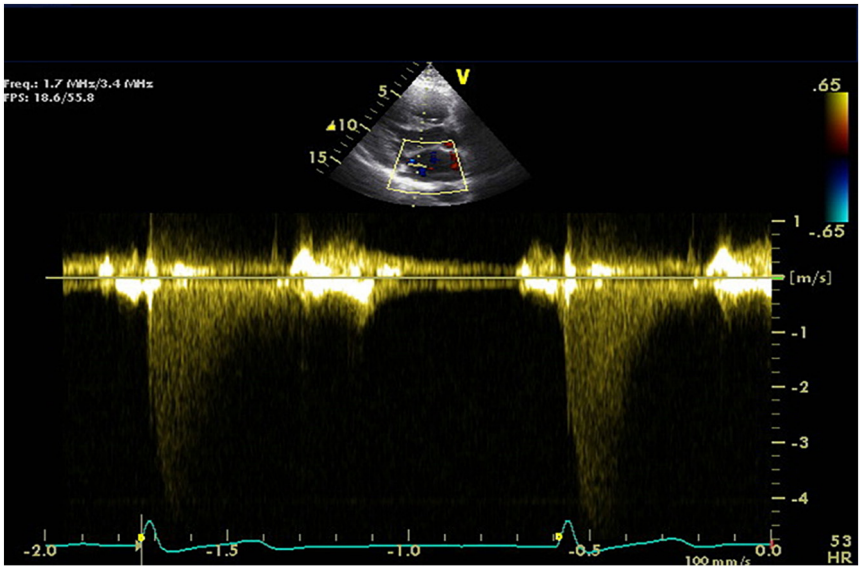


Fig. 1. Reviewer agreement of the WHF criteria with a 3 reviewer system.

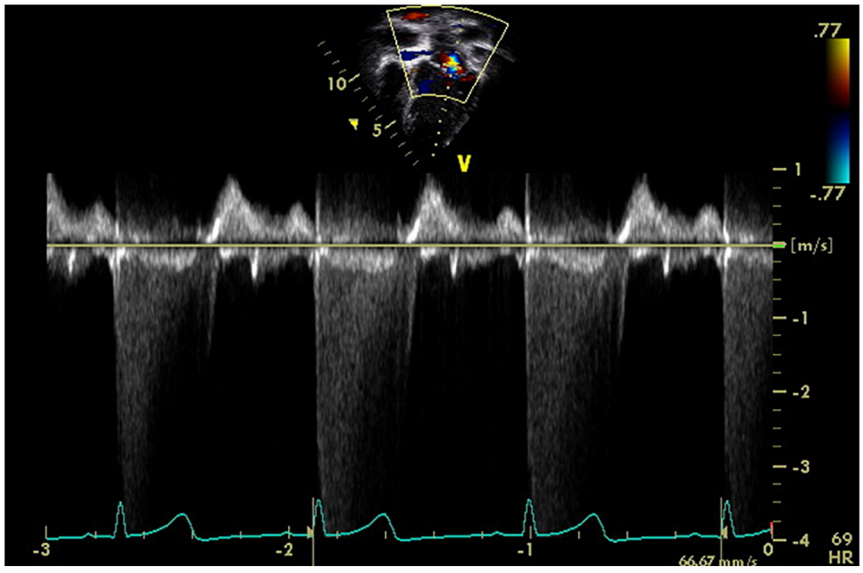
A


FR 49Hz
16cm
2D
62%
C 52
P Low
HGen
CW
50%
1.8MHz
WF 225Hz
M4
1.0
m/s
-1.0
-2.0
-3.0
-4.0
-5.0
-6.0
75mm/s
68bpm

B


Freq.: 1.7 MHz/3.4 MHz
FPS: 18.6/55.8
V
5
10
15
.65
-.65
1
[m/s]
-1
-2
-3
-4
-2.0
-1.5
-1.0
-0.5
0.0
100 mm/s
53
HR

C


10
5
V
.77
-.77
1
[m/s]
-1
-2
-3
-4
-3
-2
-1
0
66.67 mm/s
69
HR

Ethical approval for these studies was obtained from the New Zealand Health and Disability Ethics Committee *NTY/06/12/139* with amendments.

2.1. Statistical analysis

Proportions of categories of RHD were compared by chi-squared analysis. Levels of agreement of the WHF criteria between 2 reviewers were analysed by Cohen's κ statistic [13]. Pre-defined kappa agreement interpretations are <0 no agreement, 0–0.2 slight agreement, 0.21–0.4 fair agreement, 0.41–0.6 moderate agreement, 0.61–0.8 substantial agreement and >0.81 almost perfect agreement.

3. Results

The categories of final diagnosis for the respective criteria are shown in Table 2.

There was a significantly lower number of cases classified as RHD using the WHF criteria compared to the modified WHO/NIH criteria (68 cases compared to 86 $p = 0.04$, Table 2).

The derived prevalence of RHD for the screened population was 37/1000 (95% CL 29–47) using the WHF criteria as compared to 47/1000 (95% CL 38–58) using the WHO/NIH criteria, $p = 0.13$.

There was a significantly higher number of cases classified as normal using the WHF criteria compared to the modified WHO/NIH criteria (61/144 vs 37/144 $p = 0.003$).

There were a total of 21 cases classified as congenital valve abnormality (12 Aortic and 9 Mitral Valve anomalies) at the time of the original reporting using WHO/NIH criteria and 15 (9 Aortic and 6 Mitral) when the WHF criteria were used. Overall 11 cases were consistently diagnosed as congenital (9 Aortic, 2 Mitral) (Table 2).

After the initial review by 2 reviewers there was agreement for 60% ($n = 86$) of cases leading to need of a third reviewer in 40% of cases (Fig. 1). With the inclusion of the third reviewer a majority agreement (2 of 3 reviewers) was reached in 89% of cases. Thus, open label discussion was required in 11% of cases. There was no significant difference in agreement between the different categories at the two reviewer stage: 56% (28 of 50 cases) agreement for borderline cases, 67% (12 of 18 cases) agreement for definite cases, 59% (36 of 61) for normal cases and 67% (10 of 15 cases) for congenital cases (p 0.81). Kappa values varied between categories at the two reviewer stage: borderline RHD kappa 0.27 (CI 0.11–0.42) and definite RHD kappa 0.4 (CI 0.20–0.59). There was moderate agreement for any RHD, kappa 0.57 (CI 0.44–0.70). After three reviewers there was consensus agreement in 89% or 128 cases, again with no significant difference in proportions of agreement between categories: 82% for borderline cases, 89% agreement for definite cases, 95% for normal cases and 87% for congenital cases (p 0.14).

The agreement of diagnosis of reviewer A with the final diagnosis for any RHD was 'almost perfect' with kappa 0.83 (CI 0.74–0.92), and the agreement for reviewer B for any RHD was 'substantial' kappa 0.65 (CI 0.53–0.77).

There was 71% agreement by two reviewers for the diagnosis of borderline and definite aortic valve disease ($n = 14$). There was a 49% agreement by two reviewers for borderline and definite mitral valve disease ($n = 47$). There was 100% agreement when both mitral and aortic RHD was present ($n = 7$).

4. Discussion

This study addresses two important aspects of the echocardiographic diagnosis of mild RHD. Firstly the study objectively shows that the WHF criteria have raised the threshold for the diagnosis of RHD compared to the earlier WHO/NIH criteria. Secondly, the study supports the use of a three-reviewer system rather than relying on a single reviewer reporting cases.

The use of the earlier WHO/NIH criteria had previously resulted in 86 cases diagnosed as having any RHD, reducing to 68 cases diagnosed as RHD using the WHF criteria. Thus, the use of the WHF criteria has likely influenced the generally lower prevalence of RHD found in echocardiographic screening conducted since 2012 [8,14–18]. Colquhoun et al. previously reported lower prevalence rates with the WHF compared to WHO/NIH criteria using a single reviewer reporting system [17]. The RHD screening community can confidently conclude that the use of the WHF criteria had increased specificity and lowered the sensitivity (raised threshold) for RHD compared to the WHO/NIH criteria. There are now 3 studies that further address the specificity of the WHF criteria: no cases of definite RHD were found in low risk populations of RHD in Australia [8], the USA [18], and New Zealand [12]. These data, and the current study, support the on-going use of the WHF criteria for the echocardiographic diagnosis of subclinical RHD globally. The global burden of RHD has been recently been updated [19]. Echocardiographic prevalence of RHD is a key component of these estimates [20].

It has to be recognised that the two sets of criteria are not directly comparable as the WHO/NIH had a clinical component with the inclusion of a pathological heart murmur. Auscultation has been proven to be less specific and sensitive for the detection of RHD [3,6,17] and was not the focus of this study. However the 'possible RHD' category of the earlier WHO/NIH classification is made on colour-Doppler and should be closely equivalent to the WHF borderline category although there are subtle differences in wording. The WHO/NIH criteria includes "a substantial" colour jet for mitral regurgitation which is subjective and the term "well-defined spectral envelope on pulse wave/continuous wave Doppler" which can also be subjective. Our study results show a poor correlation between those diagnosed as possible RHD and borderline RHD. The WHF analysis required the reporter to strictly meet all 4 colour Doppler criteria to sign off the case as borderline. The cardiologists reporting the studies using the WHO/NIH criteria graded cases as possible RHD if they judged the regurgitation was pathological based on the 4 Doppler characteristics being met but had recorded the conclusion and not the individual Doppler components on their report. Fig. 2 illustrates that it can be subjective whether a CW Doppler jet profile is a "complete envelope" or not as per the WHF criteria. It has to be acknowledged that cut offs for colour Doppler characteristics that make the MR or AR pathological can be equivocal as these parameters are part of a biological continuum.

To our knowledge, there has been only one study which reclassified echocardiograms using WHF criteria after the original diagnosis was made using WHO/NIH criteria [15]. Their interpretation of the WHO/NIH criteria as published was slightly different than that used in New Zealand [6] but their study showed a closer correlation of the two sets of criteria.

4.1. Congenital mitral and aortic valve anomalies

Features of congenital mitral valve prolapse, with prolapse of the body of the leaflets rather than the tips of the leaflet (rheumatic), are highlighted in the WHF 2012 guidelines [9]. However, more subtle mitral valve congenital anomalies exist [21]. We found variation of cases being classified as congenital in the two analyses illustrating the difficulty of assigning subtle structural mitral valve variations to be of rheumatic or congenital aetiology. There was more consistency of agreement for congenital aortic valve anomalies than mitral valve anomalies between the two reviews. The clinical implications are

Fig. 2. Continuous wave Doppler of mitral regurgitation. Panel A - Consensus of the reviewers that the Doppler was a high velocity pansystolic jet of pathological mitral regurgitation. Panel B - Consensus of reviewers that this Doppler profile was not high velocity pansystolic jet: consistent with physiological mitral regurgitation. Panel C - No consensus between reviewers as to whether the Doppler profile was pansystolic or not.

significant for an individual who has a mitral valve showing mild regurgitation: is the aetiology rheumatic, congenital or a valve showing the upper limit of physiological MR.

The presence of any aortic regurgitation in a child, found in only 1% of normal children [12], should always lead one to look carefully at the valve morphology for a congenital lesion. A bicuspid aortic valve or dilated aortic root is usually easily identifiable but there can be other more subtle congenital valve variants. Once congenital aortic valve anomalies are excluded, then the probability that mild AR is rheumatic, is high.

4.2. One, two and three-reviewer systems

Most RHD screening studies using the WHF criteria have used a single-reviewer system, albeit some of these studies have reported their intra-observer readings [8,15–17,22,23].

Our study reveals the potential inaccuracy of a single or even two reviewers reporting abnormal RHD echocardiograms in a screening context. We found consensus in 60% of cases only. On the other hand there was 89% consensus agreement with a third reviewer used as to determine a majority opinion. An alternative would be for an open discussion when here is disagreement by two reviewers [24,25]. There have been other studies using a three-reviewer system [26]. Such systems have been used in other public health screening domains [27,28].

Our Kappa results showing moderate agreement are in the same range as others using a second reviewer to report abnormal RHD echocardiograms using the WHF criteria. Zuhlke et al. reported definite RHD K 0.77 and Borderline RHD K 0.69 for 2 reviewers for a cohort of 99 cases [25]. Engleman et al. reported K 0.59 for definite RHD and K 0.56 for any RHD for a cohort of 196 cases [16]. Bertaina et al. reported K 0.63 for borderline RHD in 50 cases [24]. Roberts et al. [8] reported the comparison of a single paediatric cardiologist using the WHF criteria to a group of cardiologists reading the echocardiograms some years earlier and found kappa 0.4–0.6 for the question 'is the mitral/aortic valve normal?' and 'is there significant mitral/aortic regurgitation?' Agreement was even lower to the questions 'is there any pathology?' k 0.4 and 'is there pathology RHD?' k 0.3.

A larger validation study of reproducibility of the WHF criteria has been recently reported by Remenyi et al. [29] 16 cardiologists reported 200 echocardiograms from RHD population-based surveys. Their study's reference gold standard was defined by a consensus two-thirds majority in 71% of the 200 cases and by reference panel adjudication in the remaining 29%. Overall agreement for the diagnosis of RHD was substantial, kappa 0.70 (95% CI 0.67–0.72) with sensitivity and specificity of 0.89 and 0.80. A diagnosis of RHD or not is of more clinical relevance from a patient's perspective than the individual categories of RHD. The diagnosis of pathologic mitral regurgitation was reliable and almost perfect, kappa 0.87. Agreement for morphologic changes of RHD was variable.

It is important to emphasize that these level of agreements using Kappa methodology apply to cohorts with abnormal echocardiograms with RHD, not to population screening cohorts where approximately 95% of cases are completely normal. All the cases in the present study were selected as abnormal or questionably abnormal (those selected with upper limit of physiological MR or AR) i.e. the sample was intentionally spiked with abnormal cases. When normal echocardiograms are included as in population screening analysis of agreement the Kappa is much higher as the proportion of normal cases is much higher. For example Aliku et al. reported 33 abnormal cases from a cohort of 455 cases [26]. With 2 reviewers the Kappa agreement for all RHD was 0.86 and for definite RHD ($n = 10$) the kappa was 1.0.

4.3. Limitations

After the decoding process, it was not possible for the reviewers to be certain whether harmonics were on or off. Harmonic imaging improves imaging resolution, is associated with fewer artifacts, but can make structures appear thicker than conventional (fundamental) imaging [11].

We acknowledge that there was no gold standard test for RHD as a reference for the final echocardiographic diagnosis. The methodology used the three reviewer system to create the gold standard, similar to that used in other public health screening domains [27,28].

5. Conclusion

We have highlighted some of the challenges of the echocardiographic diagnosis of mild RHD. Use of the WHF criteria has raised the threshold for a diagnosis of echocardiographically-detected RHD compared to earlier WHO/NIH criteria. There are limitations of reproducibility of the WHF criteria for mild RHD. The study highlights that at the lower end of structural and functional changes of valvular disease there is significant variance of up to 40% in diagnosis of the aetiology of the changes as to whether they are RHD, congenital or normal.

The final screening diagnosis and the subsequent implications for clinical management are of huge importance for the patient. In particular, being told they may have RHD (and may need BPG prophylaxis) is significant. We found 90% diagnostic agreement using a three reviewer system. The study therefore supports the use of a 3 reviewer echo reporting system as opposed to a 1 reviewer system in the diagnosis of mild RHD.

In New Zealand, a middle income country, resources exist for a three reviewer system but many RHD echocardiography studies are undertaken in low income countries where logistics may not easily allow for more than one reviewer.

Acknowledgements

We would like to thank Charlene Nell, Desktop Support Administrator, for preparing the manuscript and for excellent secretarial assistance.

Conflicts of interest

No competing interests.

Acknowledgement of grant support

The study was funded by the Health Research Council of New Zealand, the Ministry of Health, New Zealand, Cure Kids New Zealand, Te Puni Kokiri and the National Heart Foundation of New Zealand HRC ref # 13/365.

References

[1] J.R. Carapetis, A. Beaton, M.W. Cunningham, et al., Acute rheumatic fever and rheumatic heart disease, Nat. Rev. Dis. Primers. 2 (2016), 15084.
[2] Carapetis JR, Paar JA, Cherian T. Standardization of epidemiologic protocols for surveillance of post-streptococcal sequelae: acute rheumatic fever, rheumatic heart disease and acute post-streptococcal glomerulonephritis. NIH: National Institute of Allergy and Infectious Diseases. http://www.niaid.nih.gov/topics/strepThroat/Documents/groupasequelae.pdf2006 Aug 2010: [1–32 pp.].
[3] J.R. Carapetis, M. Hardy, T. Fakakovikaetau, et al., Evaluation of a screening protocol using auscultation and portable echocardiography to detect asymptomatic rheumatic heart disease in Tongan schoolchildren, Nat. Clin. Pract. Cardiovasc. Med. 5 (7) (2008) 411–417.
[4] J.A. Paar, N.M. Berrios, J.D. Rose, et al., Prevalence of rheumatic heart disease in children and young adults in Nicaragua, Am. J. Cardiol. 105 (12) (2010) 1809–1814.
[5] M. Bhaya, S. Panwar, R. Beniwal, R.B. Panwar, High prevalence of rheumatic heart disease detected by echocardiography in school children, Echocardiography 27 (4) (2010) 448–453.
[6] R.H. Webb, N.J. Wilson, D.R. Lennon, et al., Optimising echocardiographic screening for rheumatic heart disease in New Zealand: not all valve disease is rheumatic, Cardiol. Young 21 (4) (2011) 436–443.
[7] G. Cramp, M. Stonehouse, R. Webb, et al., Undetected rheumatic heart disease revealed using portable echocardiography in a population of school students in Tairawhiti, New Zealand. N Z Med J. 125 (1363) (2012) 53–64.
[8] K. Roberts, G. Maguire, A. Brown, et al., Echocardiographic screening for rheumatic heart disease in high and low risk Australian children, Circulation 129 (19) (2014) 1953–1961.

[9] B. Remenyi, N. Wilson, A. Steer, et al., World Heart Federation criteria for echocardiographic diagnosis of rheumatic heart disease—an evidence-based guideline, Nat. Rev. Cardiol. 9 (5) (2012) 297–309.
[10] F. Perelini, N. Blair, N. Wilson, A. Farrell, A. Aitken, Family acceptability of school-based echocardiographic screening for rheumatic heart disease in a high-risk population in New Zealand, J. Paediatr. Child Health 51 (7) (2015) 682–688.
[11] N.J. Wilson, J.M. Neutze, Echocardiographic diagnosis of subclinical carditis in acute rheumatic fever, Int. J. Cardiol. 50 (1) (1995) 1–6.
[12] R.H. Webb, T.L. Gentles, J.W. Stirling, et al., Valvular regurgitation using portable echocardiography in a healthy student population: implications for rheumatic heart disease screening, J. Am. Soc. Echocardiogr. 28 (8) (2015) 981–988.
[13] J.R. Landis, G.G. Koch, The measurement of observer agreement for categorical data, Biometrics 33 (1) (1977) 159–174.
[14] M. Allen, L. Allen, F. Marumatakimanu, et al., Rheumatic rescue: prospective application of the World Heart Federation echo screening guidelines in Samoa, Glob. Heart 9 (1) (2014) e333–PW61.
[15] A. Beaton, E. Okello, T. Aliku, et al., Latent rheumatic heart disease: outcomes 2 years after echocardiographic detection, Pediatr. Cardiol. 35 (7) (2014) 1259–1267.
[16] D. Engelman, G.R. Wheaton, R.L. Mataika, et al., Screening-detected rheumatic heart disease can progress to severe disease, Heart Asia 8 (2) (2016) 67–73.
[17] S.M. Colquhoun, J.H. Kado, B. Remenyi, et al., Echocardiographic screening in a resource poor setting: borderline rheumatic heart disease could be a normal variant, Int. J. Cardiol. 173 (2) (2014) 284–289.
[18] B.C. Clark, A. Krishnan, R. McCarter, et al., Using a low-risk population to estimate the specificity of the World Heart Federation criteria for the diagnosis of rheumatic heart disease, J. Am. Soc. Echocardiogr. 29 (3) (2016) 253–258.
[19] D.A. Watkins, C.O. Johnson, S.M. Colquhoun, et al., Global, regional, and national burden of rheumatic heart disease, 1990–2015, N. Engl. J. Med. 377 (8) (2017) 713–722.
[20] E. Marijon, D.S. Celermajer, X. Jouven, Rheumatic heart disease - an iceberg in tropical waters, N. Engl. J. Med. 377 (8) (2017) 780–781.
[21] B. Remenyi, T.L. Gentles, Congenital mitral valve lesions: correlation between morphology and imaging, Ann. Pediatr. Cardiol. 5 (1) (2012) 3–12.
[22] M. Mirabel, R. Bacquelin, M. Tafflet, et al., Screening for rheumatic heart disease: evaluation of a focused cardiac ultrasound approach, Circ. Cardiovasc. Imaging 8 (1) (2015), e002324.
[23] A. Beaton, J.C. Lu, T. Aliku, et al., The utility of handheld echocardiography for early rheumatic heart disease diagnosis: a field study, Eur. Heart J. Cardiovasc. Imaging 16 (5) (2015) 475–482.
[24] G. Bertaina, B. Rouchon, B. Huon, et al., Outcomes of borderline rheumatic heart disease: a prospective cohort study, Int. J. Cardiol. 228 (2017) 661–665.
[25] L. Zuhlke, M.E. Engel, C.E. Lemmer, et al., The natural history of latent rheumatic heart disease in a 5 year follow-up study: a prospective observational study, BMC Cardiovasc. Disord. 16 (46) (2016) 1–6.
[26] T. Aliku, C. Sable, A. Scheel, et al., Targeted echocardiographic screening for latent rheumatic heart disease in Northern Uganda: evaluating familial risk following identification of an index case, PLoS Negl. Trop. Dis. 10 (6) (2016), e0004727.
[27] M.H. Stoler, M. Schiffman, Interobserver reproducibility of cervical cytologic and histologic interpretations: realistic estimates from the ASCUS-LSIL Triage Study, JAMA 285 (11) (2001) 1500–1505.
[28] E.L. Thurfjell, K.A. Lernevall, A.A. Taube, Benefit of independent double reading in a population-based mammography screening program, Radiology 191 (1) (1994) 241–244.
[29] B. Remenyi, J. Carapetis, K. Sidhu, N. Wilson, Validation of the World Heart Federation evidence-based echocardiographic criteria for rheumatic heart disease, Glob. Heart 11 (2 (Supplement)) (2016) (e3-OC02-8).