

Article

Not peer-reviewed version

Statistical and Diagnostic Properties of pRR30, pRR3.25% and Asymmetrical Entropy Descriptors in Atrial Fibrillation Detection

[Bartosz Biczuk](#)^{*}, [Szymon Buś](#), [Sebastian Żurek](#), [Jarosław Piskorski](#), [Przemysław Guzik](#)

Posted Date: 26 January 2024

doi: 10.20944/preprints202401.1922.v1

Keywords: atrial fibrillation; cardiac arrhythmia; electrocardiography; heart rate variability; entropy



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Statistical and Diagnostic Properties of pRR30, pRR3.25% and Asymmetrical Entropy Descriptors in Atrial Fibrillation Detection

Bartosz Biczuk ^{1,2,*}, Szymon Bus ^{3,†}, Sebastian Żurek ¹, Jarosław Piskorski ¹ and Przemysław Guzik ^{4,5}

¹ Institute of Physics, University of Zielona Góra, Zielona Góra, Poland; bbiczuk@gmail.com

² The Doctoral School of Exact and Technical Sciences, University of Zielona Góra

³ Institute of Electronic Systems, Faculty of Electronics and Information Technology, Warsaw University of Technology, Warszawa, Poland; szymon.bus.dokt@pw.edu.pl

⁴ Department of Cardiology - Intensive Therapy, Poznan University of Medical Sciences, Poznań, Poland; pguzik@ptkardio.pl

⁵ University Centre for Sports and Medical Studies, Poznan University of Medical Sciences, Poznań, Poland

* Correspondence: bbiczuk@gmail.com

† Both authors have equal contributions and should be treated as first authors.

Abstract: Background: Early detection of atrial fibrillation (AF) is essential to prevent stroke and other cardiac and embolic complications. We compared the diagnostic properties for AF detection of the percentage of successive RR interval differences greater than or equal to 30 ms or 3.25% of the previous RR interval (pRR30 and pRR3.25%, respectively), and asymmetric entropy descriptors of RR intervals. Previously, both pRR30 and pRR3.25% outperformed many other heart rate variability (HRV) parameters in distinguishing AF from sinus rhythm (SR) in 60-second electrocardiograms (ECGs). Methods: The 60-s segments with RR intervals were extracted from the publicly available Physionet Long-Term Atrial Fibrillation Database (84 recording, 24-hour Holter ECG). There were 31753 60-s segments of AF and 32073 60-s segments of SR. The diagnostic properties of all parameters were analysed with the receiver operator curve analysis, confusion matrix and logistic regression. The best model with pRR30, pRR3.25% and total entropic features (H) had the largest area under the curve (AUC) – 0.98 compared to 0.959 for pRR30 - and 0.972 for pRR3.25%. However, the differences in AUC between pRR30, pRR3.25% alone and the combined model were negligible from a practical point of view. Moreover, combining pRR30, pRR3.25% with H significantly increased the number of false-negative cases by more than threefold. Conclusions: Asymmetric entropy has some potential in differentiating AF from SR in the 60-s RR interval time series, but the addition of these parameters does not seem to make a relevant difference compared to pRR30 and especially pRR3.25%.

Keywords: atrial fibrillation; cardiac arrhythmia; electrocardiography; heart rate variability; entropy

1. Introduction

Atrial fibrillation (AF) is a cardiac arrhythmia with irregular heartbeats. It can lead to the development of tachyarrhythmias, heart failure, dementia and arterial emboli, with ischaemic cerebral stroke being the most common complication [1]. AF also increases the risk of dying prematurely.

AF is often asymptomatic, especially in men and older people [2] and its complications, e.g. ischaemic stroke, may be the first presentation of this arrhythmia. The incidence of AF increases with age and is usually associated with many diseases or risk factors like hypertension, obesity, smoking, coronary artery disease, valvular heart disease, lung disease and hyperthyroidism [3]. However, AF is not rare in otherwise healthy people, e.g. those involved in long-term endurance sports [4]. AF can also be caused by alcohol consumption and the abuse of illicit substances such as methamphetamine, cocaine, opiates and cannabis [5].

An increasing incidence of AF and its clinical importance require effective and timely diagnosis. Recently, parameters derived from heart rate variability (HRV) analysis have become more commonly applied for AF detection in ECGs [6]. Several HRV descriptors have been proposed, such as standard deviation of the interbeat intervals (SDRR), root mean square of successive differences between normal heartbeats (rMSSD), percentage of differences higher than 50 ms (pRR50) [7].

pRR50 is the most studied but specific example of a parameter from the pRRx family (percentage of successive RR interval differences greater than or equal to x ms). Buś et al. [8] explored different parameters from the pRRx family and found that the optimal diagnostic properties for AF detection were found for the threshold $x = 31$ ms (AUC = 0.958, sensitivity = 95.35%, specificity = 90.47). The study was performed on a dataset of over 60 thousands of 1-minute segments of RR intervals. Using the same data, they reported that the pRRx% parameters (defined like pRRx but with a threshold $x\%$ relative to the previous RR interval) have even better diagnostic properties with pRR3.25% (AUC = 0.972, sensitivity = 97.16%, specificity = 93.75%) outperforming other pRRx% and pRRx parameters. Both the pRRx and pRRx% families count occurrences of unusually large successive differences in RR intervals and are generally much greater in AF than in normal sinus rhythm (SR).

Asymmetric entropy based on monotonic runs, introduced by Piskorski and Guzik [9,10], is another set of features used in HRV analysis. Unlike other entropy measures, asymmetric entropy separately examines information derived from monotonic runs consisting of heart rate accelerations, decelerations, or consecutive RR intervals that do not change (neutral monotonic runs).

We hypothesized that heart rate entropy derived from deceleration, acceleration, and neutral runs might differ between ECGs derived from sinus rhythm and atrial fibrillation. If so, the measurement of asymmetric entropy may be useful in distinguishing SR and AF segments of the ECG. This study compared the asymmetric entropy-based descriptors with pRR30 and pRR3.25% and their possible combinations for AF detection in 1-minute segments of RR intervals.

2. Materials and Methods

2.1. Data

This study used de-identified data from the Long-Term Atrial Fibrillation Database (LTAfDB) [10,11]. The LTAfDB consists of 84 extended 24-hour Holter electrocardiographic (ECG) recordings sampled at 128 Hz. The database contains information on R-wave locations and corresponding heart rhythms (normal, supraventricular, ventricular, atrial and technical artefacts) in patients with paroxysmal AF and other arrhythmias. For our analysis, we selected only continuous ECG fragments with either AF or sinus rhythm (SR) lasting at least 60 seconds, discarding segments labelled as other rhythms.

The used data pre-processing method is shown in Figure 1. The RR interval time series were divided into 60-second contiguous segments. Each RR interval within the segment had to originate from SR to label the whole segment as SR. If not, the segment was excluded from further analysis. Similarly, for AF segments, each beat within the segment had to be AF, and segments containing ventricular beats were also excluded. To minimize the number of potentially unidentified technical artefacts, we removed RR intervals shorter than 240 ms or longer than 3000 ms for both SR and AF. In addition, we removed RR intervals corresponding to ventricular premature beats from both SR and AF ECGs. For SR only, supraventricular premature beats were also removed. Segments where the total length of excluded RR intervals exceeded 3% (1.8 s) of the segment length were excluded from the analysis. After preprocessing, the total 60-second RR series was 63,636 (31,919 SR, 31,717 AF). ECG segments where the total length of RR intervals after filtering was less than 58 seconds were also discarded.

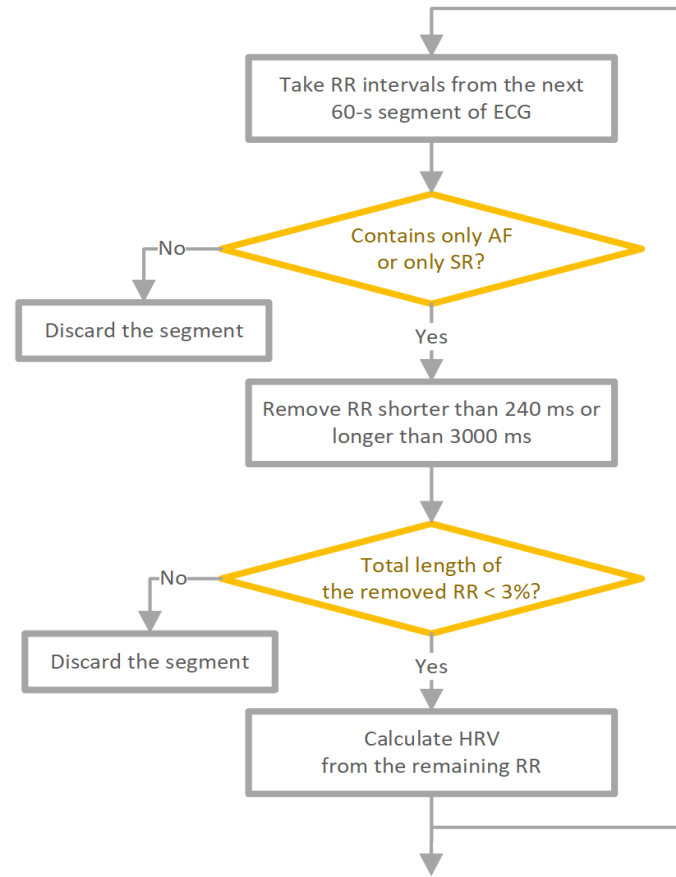


Figure 1. Data preprocessing scheme [8].

2.2. Software

We used Python programming language (version 3.9, Python Software Foundation, Wilmington, DE, USA) for statistical analysis. Part of runs-based asymmetrical calculation was done thanks to free GPL3 software written in Python, HRAexplorer, which can be reviewed and downloaded at <https://github.com/jaropis/HRAexplorer>. An interactive online version of this software in the R programming language may be found at <https://hraexplorer.com/>. Data preprocessing and calculating pRRx and pRRx% parameters were done using the code from https://github.com/simonbus/prrx_af.

2.3. Asymmetrical entropy

Asymmetrical entropy is an approach presented in [9]. In summary, the RR series is partitioned into monotonic runs, as shown in Figure 1. If we define a series of RR intervals as:

$$RR_N \equiv (RR_1, RR_2, \dots, RR_N)$$

we can use this to build a series of differences. Starting from the second element, we subtract $(i-1)$ -th element from i -th element like this:

$$\Delta \equiv (\delta_1, \delta_2, \dots, \delta_{N-1})$$

Where single difference looks like this:

$$\delta_i = RR_{i+1} - RR_i$$

We can use this series to construct a symbolic series of signs of those differences:

$$\text{sgn}(\delta_i) = \{+, \delta_i > 0, -, \delta_i < 0, 0, \delta_i = 0\}$$

Using this, we can write definitions for acceleration, deceleration and neutral series:

Definition 1. A deceleration series of length i (DR $_i$) is an uninterrupted series of i decelerations (sign -) that starts and ends with acceleration (sign +) or neutral difference (sign 0).

Definition 2. An acceleration series of length i (AR i) is an uninterrupted series of i accelerations (sign -) that starts and ends with deceleration (sign +) or neutral difference (sign 0).

Definition 3. A neutral series of length i (NR i) is an uninterrupted series of i neutral differences (sign 0) that starts and ends with acceleration (sign -) or deceleration (sign +).

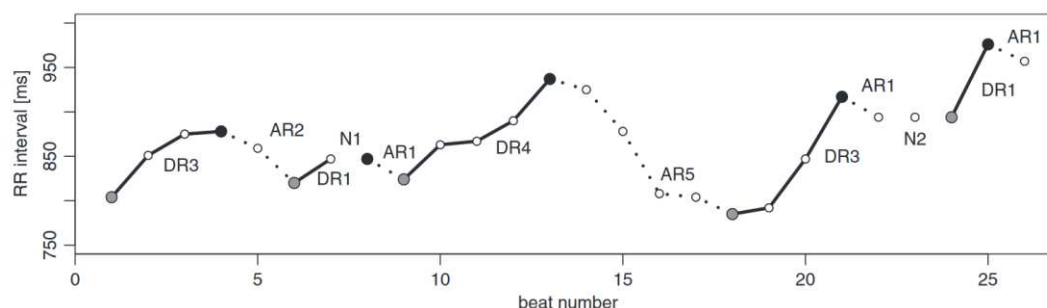


Figure 2. Example of runs-based partitioning of a short tachogram. The runs can be classified as deceleration and acceleration runs of length i (DR i and AR i , respectively), with neutral runs represented by the symbol Ni, which can interrupt the deceleration/acceleration runs. A full grey circle indicates the start of a deceleration run. A full black circle marks the start of an acceleration run. These circles can be used as reference points for the respective runs [9].

It is common practice [9] [12] [13] to remove neutral runs from the record by adding white noise with a sufficiently small standard deviation. However, we will include neutral runs in our calculation and model building.

We calculate the Shannon entropy for each type of the monotonic run. If we define the value p as a probability estimate of how likely it is to find an interval in the RR recording that belongs to a given run type of a given length:

$$p_{i,k} = \frac{(\text{number of } r_i^k) \times i}{n}$$

where i is run length and k is run type and can be AR, DR, NR, the Shannon entropy for this estimator is as follows:

$$H_k = - \sum_{i=1}^{\max(i)_k} p_{i,k} \cdot \ln p_{i,k}$$

Mathematical details can be found [9]

2.4. pRR30 and pRR3.25%

One of the key parameter in HRV analysis is pRR50, a special case of a pRR x parameter, which measures the percentage of successive RR intervals differing by at least $x = 50$ ms. Recent study [8] has explored a whole family of pRR x parameters and searched for a candidate for AF detection. They also explored pRR x % parameters (percentage of relative RR interval differences of at least x % of the previous RR interval).

In these study we found pRR30 and pRR3.25% to be promising candidates for AF detection, outperforming pRR50 in 60 seconds long RR time series.

2.5. Statistical analysis

Data distribution was not normal [Guzik P, Więckowska B. Data distribution analysis – a preliminary approach to quantitative data in biomedical research. JMS. 2023 Jun. 27;92(2):e869. DOI: 10.20883/medical.e869]. Non-parametric Spearman correlation analyzed associations between

parameters. Receiver operating characteristic (ROC) [13] was used to investigate which parameters discriminated between 60-second segments of sinus rhythm (SR) and atrial fibrillation (AF). For each HRV parameter studied, Youden's criterion was used to calculate the optimal cutoff from the ROC curve [14].

We used a 2:1 ratio to divide the final set of 63,826 recordings into the training dataset (42,763 samples) and the test dataset (21,063 samples). Univariate and multivariate logistic regression examined individual parameters or their combinations with either pRR30 or pRR3.25% or both pRR30 and pRR3.25%. A non-parametric bootstrap with 1,000 samples was used to estimate classification metrics with a 95% confidence interval. Accuracy, sensitivity and specificity were compared between features using a paired t-test, and the diagnostic odds ratio (DOR) [15] was studied by the Wilcoxon test [16]. Only $p < 0.05$ was considered statistically significant.

3. Results

Our data consisted of 64738 segments, from which 912 were discarded after filtration, leaving 63826 segments. 31753 of them were classified as AF and 32073 as SR. Mean RR is 698 ms for AF and 858 ms for SR.

3.1. Entropy distribution

Figure 4 presents histograms for all studied types of asymmetrical entropy: total entropy (H), entropy of acceleration runs (HAR), entropy of deceleration runs (HDR) and entropy of neutral runs (HNR). All forms of entropy have narrower and higher histograms for AF than for SR. The peaks of histograms for H overlap for AF and SR, but are separated for HAR, HDR and HNR. Median HAR and HDR are higher for AF than for SR. In contrast, the median H and HNR is lower in AF.

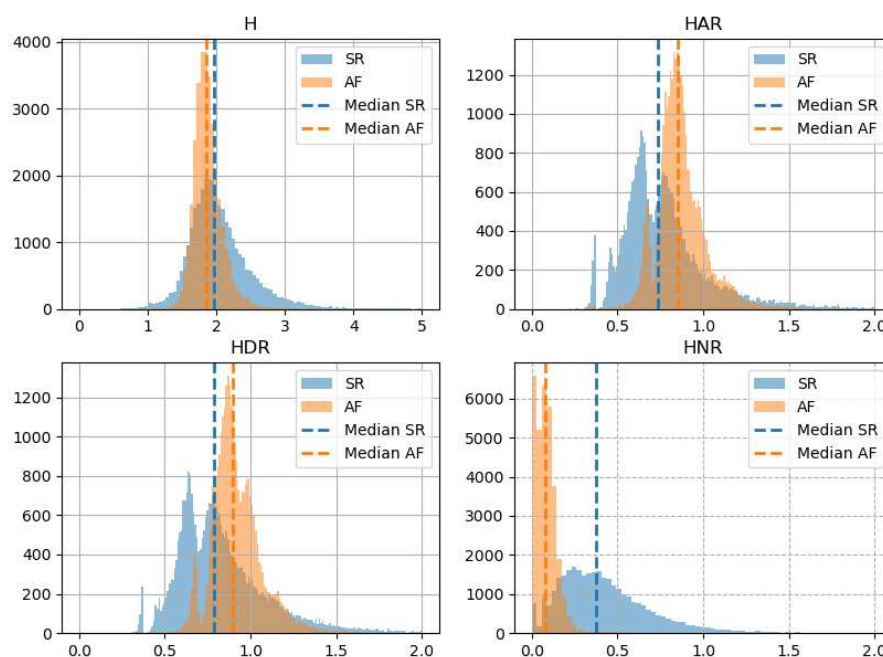


Figure 4. Histograms of different types of asymmetrical entropy for AF and SR.

Figure 5 shows how the histograms of the pRR30 and pRR3.25% differ in AF and SR. For SR, the most common values of pRR30 and pRR3.25% are close to 0%. The rates gradually decrease for higher values of pRR30 and pRR3.25%. In contrast, the distributions are shifted to the right for AF and the most common values of pRR30 and pRR3.25% are around 85% and 90%, respectively.

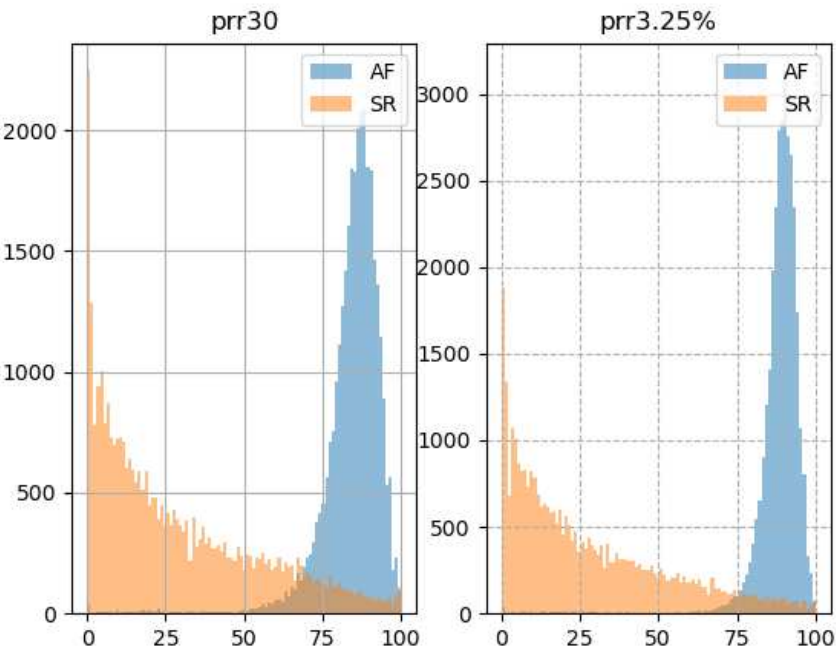
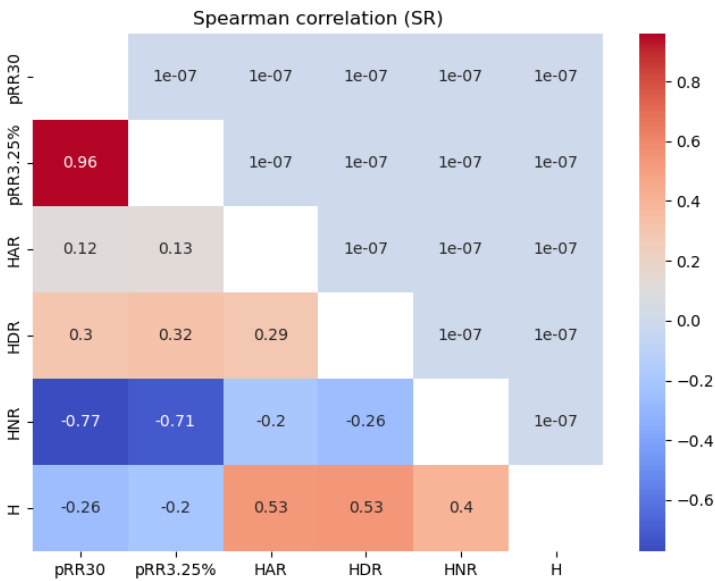


Figure 5. Histograms of prr30 and prr3.25% for sinus rhythms and atrial fibrillation recordings.

Figure 6 shows the Spearman correlation values between all the features examined for the SR and AF data separately. For SR, HNR correlated strongly and negatively (dark blue) with pRR30 (-0.77) and pRR3.25% (-0.71). The correlations between HNR and pRR30 and pRR3.25% were weaker for AF, with the most pronounced correlations for HNR and pRR30 (-0.59).



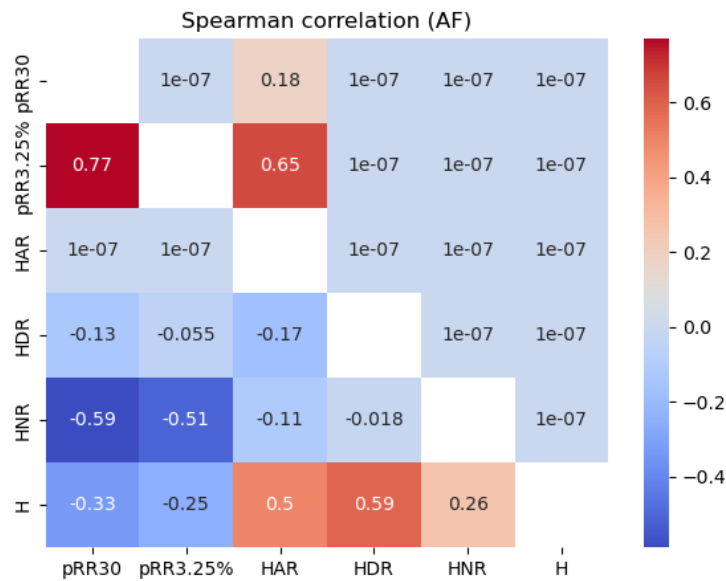


Figure 6. Spearman correlation matrix for HRV parameters for SR and AF data. Lower triangle: correlation coefficients, upper triangle: p-values.

3.3. Diagnostic properties of single HRV parameters

Table 1 characterizes the diagnostic properties of all investigated parameters to differentiate AF from SR for the 1-minute ECG segments. As pRR3.25% has the highest AUC, it is considered the reference value for comparisons with other parameters and is included in all multivariate diagnostic models. AUC for the discrimination of AF from SR for the 1-minute RR interval ECG segments were all significantly different from 0.5. (p at least $< 1 \times 10^{-114}$). The top three strongest candidates for AF detection are pRR3.25% (AUC = 0.9727) pRR30 (0.9596), and HNR (0.9315).

Table 1. Area under the ROC curve (AUC) and optimal cutoff values for the training set. Median values of classification metrics (from the confusion matrix) for AF detection in the test set using single HRV parameters with the optimal cutoff.

Parameter	AUC	Cutoff for AF	Accuracy	Sensitivity	Specificity	PPV	NPV	DOR
pRR3.25%	0.9727	<72.3684%	95.75	99.60	92.25	92.12	99.60	2931.82
pRR30	0.9596	<66.8874%	91.00	95.62	86.80	86.82	95.62	143.70
HNR	0.9315	<0.1884	84.86	95.45	75.26	77.83	94.79	63.82
HAR	0.6951	<0.7546	67.85	86.83	50.59	61.52	80.86	6.76
HDR	0.6686	<0.7726	67.68	91.49	46.00	60.66	85.61	9.16
H	0.6186	<2.0647	60.03	85.68	36.69	55.18	73.80	3.47

Figure 7 shows the classification metrics for AF detection on the testing dataset using the optimal cutoff method for single HRV parameters. For every metric, the best score achieved three parameters: pRR3.25%, pRR30 and HNR. The NPV difference between pRR30 and HNR is the smallest. The whiskers on the plots represent 95% confidence intervals from the bootstrap.

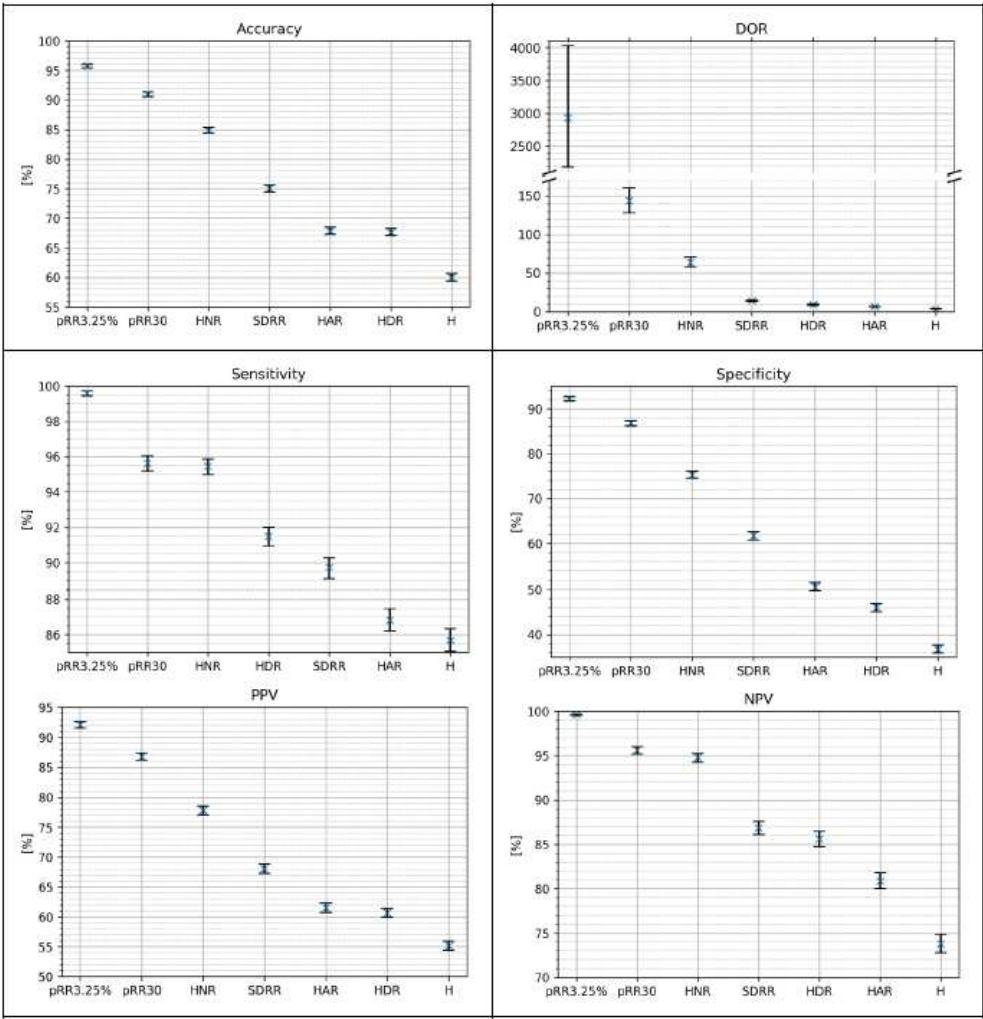


Figure 7. Classification metrics of AF detection on the test dataset using the optimal cutoff method. DOR - diagnostic odds ratio, PPV - positive predictive value, NPV - negative predictive value.

3.4. Diagnostic values of models built using pRR30, pRR3.25% and asymmetric entropy indices

Table 2 presents results of AF detection using univariate or multivariate logistic regression models on testing dataset. The highest AUC was achieved by model with three parameters: pRR30, pRR3.25% and H while pRR30, pRR3.25%, HAR has the highest DOR. From univariate models pRR3.25% has highest AUC and all other metrics.

Table 2. Comparison of different metrics for AF detection using univariate and multivariate logistic regression models. The three top metrics with the highest AUC are bolded. pRR3.25% stands out as the most powerful individual metric, exhibiting the highest AUC. Notably, it consistently appears in the top-performing metric pairs and triplets, consistently demonstrating superior performance in distinguishing AF from SR.

Feature	AUC	Accuracy	DOR	FP+FN [%]	FP [%]	FN [%]
pRR3.25%	0.972	0.9513	643.0	4.87	3.98	0.88
pRR30	0.959	0.9277	233.5	7.23	5.57	1.65
H	0.613	0.6118	2.6	38.82	24.3	14.52
HAR	0.7	0.6536	3.6	34.64	16.11	18.53
HDR	0.67	0.6155	2.6	38.45	16.91	21.54

HNR	0.93	0.8842	84.5	11.58	09.03	2.55
pRR3.25 & pRR30%	0.978	0.955	687.1	4.5	3.57	0.93
pRR3.25% & H	0.986	0.9566	705.6	4.34	3.38	0.96
pRR3.25% & HAR	0.982	0.9549	693.0	4.51	3.59	0.92
pRR3.25% & HDR	0.977	0.9517	623.3	4.83	3.9	0.93
pRR3.25% & HNR	0.974	0.9516	632.3	4.84	3.93	0.91
pRR30 & H	0.973	0.9327	252.3	6.73	05.01	1.72
pRR30 & HAR	0.969	0.9312	240.5	6.88	5.13	1.76
pRR30 & HDR	0.963	0.9279	235.8	7.21	5.57	1.64
pRR30 & HNR	0.96	0.9284	232.9	7.16	5.47	1.69
pRR30 & pRR3.25%, & H	0.988	0.9577	695.5	4.23	3.2	01.03
pRR30 & pRR3.25%, & HAR	0.985	0.9581	732.4	4.19	3.22	0.97
pRR30 & pRR3.25%, & HDR	0.982	0.9541	633.7	4.59	3.58	1
pRR30 & pRR3.25%, & HNR	0.979	0.9537	631.3	4.63	3.65	0.99
HNR & HAR, & HDR	0.935	0.8838	82.8	11.62	09.01	2.61

From Table 2 we selected three models for further validation. The best (in AUC) single feature model which is pRR3.25%, best two-features model which is pRR3.25%& pRR30 and best three features model which is pRR3.25%,&pRR30& H. A 1000-resampled bootstrap test revealed a statistically significant difference in AUC (p-value = 0.02) between the one-feature and two-feature models. Furthermore, the comparisons between the one-feature versus three-feature and two-feature versus three-feature models yielded p-values less than 10⁻⁶, indicating highly significant differences. The same bootstrap methodology was employed to assess the performance of the models in terms of accuracy, sensitivity, specificity, DOR, FN, and FP. In all cases, the differences between the models were statistically significant, with p-values consistently below 10⁻⁵. Figure provides histograms with a visual representation of these findings.

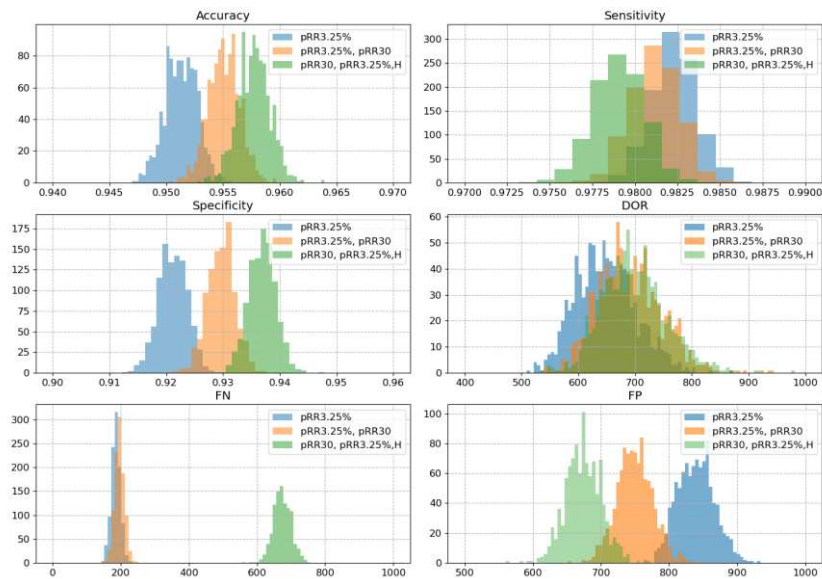


Figure 8. Histograms presenting 1000 times bootstrapped results of model built from pRR3.25% vs pRR30, pRR3.25% versus model built from pRR30,pRR3.25%,HAR. The biggest difference is in FN where adding H to the model increases the number of FN over three times compared to models without H.

It is worth noting that the scales in Figure 8 for accuracy, sensitivity and specificity on the respective x-axes have been adapted. In this way it is easier to see the small differences. In practice, however, the range of these differences is negligible. Optimizing the models for AUC is at the expense of undetection of AF. The median FN for pRR3.25% and pRR3.25% & pRR30 increases from less than 200 to almost 700 undetected AF for the model with pRR3.25%, pRR30 and H.

4. Discussion

We found that both pRR3.25% and pRR30 outperformed all asymmetric entropy parameters in their diagnostic value for detecting AF in 1-minute ECGs. For AF detection, the best single parameter was pRR3.25% with an AUC of 0.972 and a DOR of 643, the second best was pRR30 with an AUC of 0.959 and a DOR of 233.5, and the third best was HNR with an AUC of 0.93 and a DOR of 84.5. All combinations that had either a higher AUC or DOR than a single pRR3.25% contained pRR3.25%. The best combination when measured by AUC was pRR3.25%, pRR30 and H, while when measured by DOR it was pRR3.25% and H. However, the combined models with higher AUC or DOR had increased false negative values. Using pRR3.25% as a reference, the observed increases in AUC were of negligible clinical value as they were up to 0.006, which is less than 1% improvement. In addition, using the best combination, i.e. pRR3.25%, pRR30 and H, was associated with a risk of increasing false-negative AF diagnoses - the number of undiagnosed AF episodes increased almost fourfold.

This study is not the first one to demonstrate entropy based features for AF detection. In [23] Richman JS and Moorman JR have introduced sample entropy, a novel method for comparing two physiological time series. Liu C et al. and Zhao L et al [21,22] have compared different entropic measures and proposed a novel one, called normalized fuzzy entropy for AF detection. They have achieved 0.86 accuracy for 60 beats long RR segments. Their approach and other tested features was based on sample entropy [23].

Runs-based asymmetrical entropy is very different from sample entropy. It does not depend on parameters which represent matching radius for determining similarity between subsequences of data points within a time series and can break relative consistency which means it does not always provide consistent results when sample size is increasing [24]. Run-based asymmetric entropy has a more straightforward interpretation. It quantitatively separates the contributions of accelerating, decelerating and neutral runs to the overall asymmetry of the time series (Figure 2).

We found a negative correlation between the asymmetric entropy of NR and pRR30. In addition, the HNR distribution differs between AF and SR. The AF classifier built from HNR alone has an accuracy of over 0.88, the best among other run-based asymmetric entropy parameters. Neutral runs are either ignored, removed from the analysis or replaced by the addition of white noise [9,12,13]. However, neutral runs may contain valuable information, especially in the context of AF detection. Typically, their contribution to HRV is either zero (to short-term HRV) or random and depends on the ECG sampling frequency. However, their lower rate in AF provides interesting information.

The nature of AF makes RR intervals highly random with low stability. This results in a low probability of finding two or more consecutive RR intervals of the same duration. It also explains why there are fewer neutral runs in AF than in SR. Changes in RR interval duration in SR are gradual. They are controlled by many physiological mechanisms that modulate sinus node depolarisation. In AF, fast fibrillation waves above 350 depolarisations/minute reset and quieten the sinus node. Changes between successive RR intervals in AF are less well controlled and therefore more dramatic. We have previously reported higher values of pRRx and pRRx% in AF than in SR [8]. Therefore, the probability of small changes between consecutive RR intervals is low in AF but higher in SR. Here we show that information derived from neutral runs has some diagnostic properties. However, it is uncertain whether this information should be incorporated into AF detection algorithms. All run-based asymmetric entropy measures, either as single parameters or in combination with pRR3.25% or pRR30, are not particularly effective in discriminating AF from SR. They are also more complex to calculate than pRR3.25% or pRR30.

As previously shown, the pRRx and pRRx% families have interesting diagnostic properties for distinguishing ECG segments with AF from SR. First, we demonstrated that pRR30 outperformed

other commonly used HRV indices including SDNN, SD1, SD2, SD2/SD1, coefficient of variation, for the same purpose [25]. Second, we showed that pRR30 is outperformed by the pRRx% family, especially pRR3.25%. In this study, we report that both pRR30 and pRR3.25 have better diagnostic properties for AF detection than any of the run-based asymmetric entropy descriptors. Adding asymmetric entropy-derived descriptors to models with pRR30 or pRR3.25 slightly improved their AUC or DOR, but at the cost of increasing FN.

The impact of false positive and false negative results can vary depending on the specific condition being tested for and the context in which the test is performed. For example, a false positive result for a serious condition may lead to unnecessary treatment or further testing. Conversely, a false negative result may delay necessary treatment. In AF, a false-negative diagnosis appears to be more harmful than a positive diagnosis. A patient with undiagnosed AF will not receive antithrombotic treatment and will remain at increased risk of ischaemic stroke. In contrast, a false-positive result leads to further testing and has no immediate negative consequences [19] [20].

Our study had several limitations. The analysis was based on RR intervals from 1-minute ECGs only, potentially missing longer-term cardiac dynamics that contribute to entropy. We cannot extrapolate our findings to ECGs of shorter or longer duration. The sampling frequency of our data was only 128 Hz, which increases the rate of neutral RR intervals for technical rather than physiological reasons. In this study we tested the ability to detect AF using classic simple logistic regression models. Although we achieved a high accuracy rate of 0.96, there is potential for even higher accuracy rates using more complex machine learning models. ML algorithms [21] may perform better with random forest than with logistic regression. Whether this applies to our results is uncertain. It is noteworthy that the AUC for AF detection by pRR3.25% is already close to 1 and there is not much room for improvement. However, it may be worth exploring a wider range of machine learning models, such as XGBoost or Random Forest, to determine their ability to detect AF from the asymmetric entropy of neutral runs. These models may be able to capture more complex relationships between the features and the target variable.

The novelty of our study lies in the comprehensive investigation of AF detection using a novel combination of runs-based asymmetric entropy and pRR30 with pRR3.25% and pRR30. Runs based asymmetric entropy has not previously been tested as a candidate for AF detection. Another novelty was not to remove neutral runs but to include them and analyze their characteristics, which seems to be useful.

5. Conclusions

pRR3.25% stands out as a robust and reliable choice for effectively identifying AF episodes, even in comparison with runs-based asymmetrical entropy. While it is observed that more complex models such as pRR3.25%, pRR30 and H exhibit slightly enhanced performance on a general scale, this marginal improvement is often overshadowed by the substantial drawback of a nearly fourfold increase in false negatives. Notably, it is worth emphasizing that every superior model inherently encompasses the pRR3.25% parameter. This underscores the foundational importance of pRR3.25% in contributing to the diagnostic accuracy of any enhanced model. In practical clinical applications, the single-parameter model with pRR3.25% emerges as the optimal choice, delivering a commendable balance between detection precision and false-negative rates.

Our results suggest that HNR may be a helpful parameter for AF detection, and future studies should investigate its potential further.

Author Contributions: Conceptualization, Jarosław Piskorski and Przemysław Guzik; Data curation, Szymon Buś; Formal analysis, Bartosz Biczuk and Szymon Buś; Investigation, Bartosz Biczuk; Methodology, Bartosz Biczuk, Szymon Buś and Sebastian Żurek; Project administration, Przemysław Guzik; Resources, Przemysław Guzik; Software, Bartosz Biczuk, Szymon Buś and Jarosław Piskorski; Supervision, Jarosław Piskorski and Przemysław Guzik; Visualization, Bartosz Biczuk and Szymon Buś; Writing – original draft, Bartosz Biczuk and

Szymon Buś; Writing – review & editing, Bartosz Biczuk, Szymon Buś, Sebastian Żurek, Jarosław Piskorski and Przemysław Guzik.

Funding: This research received no external funding

Conflicts of Interest: The authors declare no conflict of interest.

References

1. G. Hindricks, T. Potpara, N. Dagres, E. Arbelo, J. Bax, C. Blomström-Lundqvist, G. Boriani, M. Castella, G. Dan and P. Dilaveris, "2020 ESC Guidelines for the diagnosis and management of atrial fibrillation developed in collaboration with the European Association for Cardio-Thoracic Surgery (EACTS): The Task Force for the diagnosis and management of atrial fibrillation of the Europea," *European heart journal*, Vols. 42,5, pp. 373-498, 2021.[CrossRef]
2. Ble M, Benito B, Cuadrado-Godia E, Pérez-Fernández S, Gómez M, Mas-Stachurska A, Tizón-Marcos H, Molina L, Martí-Almor J, Cladellas M. Left Atrium Assessment by Speckle Tracking Echocardiography in Cryptogenic Stroke: Seeking Silent Atrial Fibrillation. *J Clin Med*. 2021 Aug 9;10(16):3501. doi: 10.3390/jcm10163501. PMID: 34441797; PMCID: PMC8397042. [CrossRef]
3. Roten L, Goulouti E, Lam A, Elchinova E, Nozica N, Spirito A, Wittmer S, Branca M, Servatius H, Noti F, Seiler J, Baldinger SH, Haeberlin A, de Marchi S, Asatryan B, Rodondi N, Donzé J, Aujesky D, Tanner H, Reichlin T, Jüni P. Age and Sex Specific Prevalence of Clinical and Screen-Detected Atrial Fibrillation in Hospitalized Patients. *J Clin Med*. 2021 Oct 22;10(21):4871. doi: 10.3390/jcm10214871. PMID: 34768391; PMCID: PMC8584962.
4. [CrossRef]
5. Turagam MK, Flaker GC, Velagapudi P, Vadali S, Alpert MA. Atrial Fibrillation In Athletes: Pathophysiology, Clinical Presentation, Evaluation and Management. *J Atr Fibrillation*. 2015;8(4):1309. Published 2015 Dec 31. doi:10.4022/jafib.1309 [PubMed]
6. Lin AL, Nah G, Tang JJ, Vittinghoff E, Dewland TA, Marcus GM. Cannabis, cocaine, methamphetamine, and opiates increase the risk of incident atrial fibrillation. *Eur Heart J*. 2022;43(47):4933-4942. doi:10.1093/eurheartj/ehac558 [CrossRef]
- A. Rizwan, A. Zoha, I. Mabrouk, H. Sabbour, A. Al-Sumaiti, A. Alomainy, M. Imran and Q. Abbasi, "A Review on the State of the Art in Atrial Fibrillation Detection Enabled by Machine Learning," *IEEE Reviews in Biomedical Engineering*, vol. 14, pp. 219-239, 2021. [CrossRef]
7. Khan AA, Lip GYH, Shantsila A. Heart rate variability in atrial fibrillation: The balance between sympathetic and parasympathetic nervous system. *Eur J Clin Invest*. 2019 Nov;49(11):e13174. doi: 10.1111/eci.13174. PMID: 31560809. [CrossRef]
8. Buś S, Jędrzejewski K, Guzik P. Statistical and Diagnostic Properties of pRRx Parameters in Atrial Fibrillation Detection. *J Clin Med*. 2022 Sep 27;11(19):5702. doi: 10.3390/jcm11195702. PMID: 36233572; PMCID: PMC9572524. [CrossRef]
9. Piskorski J, Guzik P. The structure of heart rate asymmetry: deceleration and acceleration runs. *Physiol Meas*. 2011 Aug;32(8):1011-23. doi: 10.1088/0967-3334/32/8/002. Epub 2011 Jun 7. PMID: 21654026. [CrossRef]
10. Goldberger AL, Amaral LA, Glass L, Hausdorff JM, Ivanov PC, Mark RG, Mietus JE, Moody GB, Peng CK, Stanley HE. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation*. 2000 Jun 13;101(23):E215-20. doi: 10.1161/01.cir.101.23.e215. PMID: 10851218. [CrossRef]

11. Petrutiu S, Sahakian AV, Swiryn S. Abrupt changes in fibrillatory wave characteristics at the termination of paroxysmal atrial fibrillation in humans. *Europace*. 2007 Jul;9(7):466-70. doi: 10.1093/europace/eum096. Epub 2007 May 31. PMID: 17540663. [CrossRef]
12. Y. Mina-Paz, V. Santana-García, L. Tafur-Tascon, M. Cabrera-Hernández, A. Pliego-Carrillo and J. Reyes-Lagos, "Analysis of Short-Term Heart Rate Asymmetry in High-Performance Athletes and Non-Athletes.," *Symmetry*, vol. 14, no. 1229, 2022. [CrossRef]
13. Sibrecht G, Piskorski J, Krauze T, Guzik P. Heart Rate Asymmetry, Its Compensation, and Heart Rate Variability in Healthy Adults during 48-h Holter ECG Recordings. *J Clin Med*. 2023 Feb 3;12(3):1219. doi: 10.3390/jcm12031219. PMID: 36769867; PMCID: PMC9917705. [CrossRef]
14. W. B. Guzik P, "Data distribution analysis – a preliminary approach to quantitative data in biomedical research.," *JMS [Internet]*, 2023. [CrossRef]
15. T. Fawcett, "An introduction to ROC analysis," *Pattern Recognit. Lett.*, vol. 27, p. 861–874, 2006. [CrossRef]
16. W. Youden, "Index for rating diagnostic tests.," *Cancer*, vol. 3, pp. 32–35;2-3, 1950. [CrossRef]
- A. Glas, J. Lijmer, M. Prins, G. Bonsel and P. Bossuyt, "The diagnostic odds ratio: A single indicator of test performance.," *J. Clin. Epidemiol.*, vol. 56, pp. 1129–1135, 2003. [CrossRef]
17. F. Wilcoxon, "Individual Comparisons by Ranking Methods.," *Biom. Bull*, vol. 1, pp. 80-83., 1945. [CrossRef]
18. Renshaw AA, Gould EW. Reducing false-negative and false-positive diagnoses in anatomic pathology consultation material. *Arch Pathol Lab Med*. 2013 Dec;137(12):1770-3. doi: 10.5858/arpa.2013-0012-OA. PMID: 24283857. [CrossRef]
19. Klinkman MS, Coyne JC, Gallo S, Schwenk TL. False positives, false negatives, and the validity of the diagnosis of major depression in primary care. *Arch Fam Med*. 1998 Sep-Oct;7(5):451-61. doi: 10.1001/archfami.7.5.451. PMID: 9755738. [CrossRef]
20. Liu C, Oster J, Reinertsen E, Li Q, Zhao L, Nemati S, Clifford GD. A comparison of entropy approaches for AF discrimination. *Physiol Meas*. 2018 Jul 6;39(7):074002. doi: 10.1088/1361-6579/aacc48. PMID: 29897343; PMCID: PMC8522195. [CrossRef]
21. Zhao L, Liu C, Wei S, Shen Q, Zhou F, Li J. A New Entropy-Based Atrial Fibrillation Detection Method for Scanning Wearable ECG Recordings. *Entropy (Basel)*. 2018 Nov 26;20(12):904. doi: 10.3390/e20120904. PMID: 33266628; PMCID: PMC7512487. [CrossRef]
22. Richman JS, Moorman JR. Physiological time-series analysis using approximate entropy and sample entropy. *Am J Physiol Heart Circ Physiol*. 2000 Jun;278(6):H2039-49. doi: 10.1152/ajpheart.2000.278.6.H2039. PMID: 10843903. [CrossRef]
23. Żurek S, Grabowski W, Wojtiuk K, Szewczak D, Guzik P, Piskorski J. Relative Consistency of Sample Entropy Is Not Preserved in MIX Processes. *Entropy (Basel)*. 2020 Jun 21;22(6):694. doi: 10.3390/e22060694. PMID: 33286466; PMCID: PMC7517231. [CrossRef]
24. Buś S, Jędrzejewski K, Guzik P. Using Minimum Redundancy Maximum Relevance Algorithm to Select Minimal Sets of Heart Rate Variability Parameters for Atrial Fibrillation Detection. *J Clin Med*. 2022 Jul 11;11(14):4004. doi: 10.3390/jcm11144004. PMID: 35887768; PMCID: PMC9318370 [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.