

# Uma análise estatística de eletrocardiograma como suporte no diagnóstico de insuficiência cardíaca congestiva

## A statistical analysis of electrocardiogram as support in the diagnosis of congestive heart failure

---

**Leonardo Silva Sousa**

*Universidade Ceuma, Departamento de Educação, São Luís, MA, Brasil.*

**Carlos Eduardo Ferreira Junior**

*Universidade Ceuma, Departamento de Educação, São Luís, MA, Brasil.*

**Afonso Jansen de Melo Farias**

*Universidade Ceuma, Departamento de Educação, São Luís, MA, Brasil.*

**Gabriel Mendes Mouta**

*Universidade Ceuma, Departamento de Educação, São Luís, MA, Brasil.*

**Mariane Soares dos Santos**

*Universidade Ceuma, Departamento de Educação, São Luís, MA, Brasil.*

**Jonathan Araújo Queiroz**

*Universidade Ceuma, Departamento de Educação, São Luís, MA, Brasil*

DOI: 10.47573/aya.5379.2.55.29

## RESUMO

Cada vez mais se torna comum problemas cardiovasculares, no Brasil infelizmente é uma doença que afeta milhares de pessoas, segundo a OMS (Organização Mundial da Saúde), a taxa de mortalidade por insuficiência cardíaca aumentou bastante nas últimas décadas, sendo mais de 17 milhões de mortes anuais no mundo causadas pela síndrome. Esse estudo propõe que por meio da análise estatística do sinal de ECG podemos fazer o diagnóstico do ICC de maneira precisa. Fazendo uso de uma base de dados com ECG de pessoas saudáveis e com ICC e utilizando nesses sinais métodos estatísticos como a média, mediana, assimetria e o desvio padrão, obtemos as características de cada sinal, nos permitindo prever e diagnosticar prováveis casos de insuficiência cardíaca congestiva. O resultado alcançado pelo método proposto com essa metodologia foi de uma acurácia de mais de 90%, em algumas das métricas propostas, o modelo chegou a uma acurácia de 100%. Com esses resultados o método tem potencial de ser utilizado no auxílio no diagnóstico da ICC. Além da alta acurácia, o método é relativamente simples e utiliza pouco poder computacional para rodar as métricas.

**Palavras-chave:** insuficiência cardíaca. congestiva. análise. eletrocardiograma.

## ABSTRACT

Cardiovascular problems become more and more common, in Brazil unfortunately it is a disease that affects thousands of people, according to the WHO (World Health Organization), the mortality rate from heart failure has increased significantly in recent decades, with more than 17 million annual deaths worldwide caused by the syndrome. This study proposes that through statistical analysis of the ECG signal we can accurately diagnose CHF. Using an ECG database of healthy people with CHF and using statistical methods such as mean, median, asymmetry and standard deviation on these signs, we obtain the characteristics of each sign, allowing us to predict and diagnose probable cases of heart failure congestive. The result achieved by the proposed method with this methodology was an accuracy of more than 90%, in some of the proposed metrics, the model reached an accuracy of 100%. With these results, the method has the potential to be used to aid in the diagnosis of CHF. In addition to its high accuracy, the method is relatively simple and uses little computational power to run the metrics.

**Keywords:** heart failure. congestive. analysis. electrocardiogram.

## INTRODUÇÃO

A insuficiência cardíaca congestiva (ICC) é caracterizada pela incapacidade do coração de bombear sangue adequadamente para as demais regiões do corpo, geralmente causado pela fraqueza ou rigidez do músculo responsável pela contração e relaxamento do mesmo, a mesma causa má funcionalmente dos rins, problemas pulmonares e etc.

A doença tem uma incidência maior em pessoas a partir dos 50 anos, estudos apontam que a prevalência de casos nesta idade aumentou 0,9% e passou a aumentar 17,4% em pessoas com idade  $\geq 85$  anos, com incidência de ataque cardíaco de 14,4/1000 e há estudos que indicam que ela pode aumentar para 40% devido a estimativa de vida estar crescendo e devido aos hábitos alimentares da nova geração.

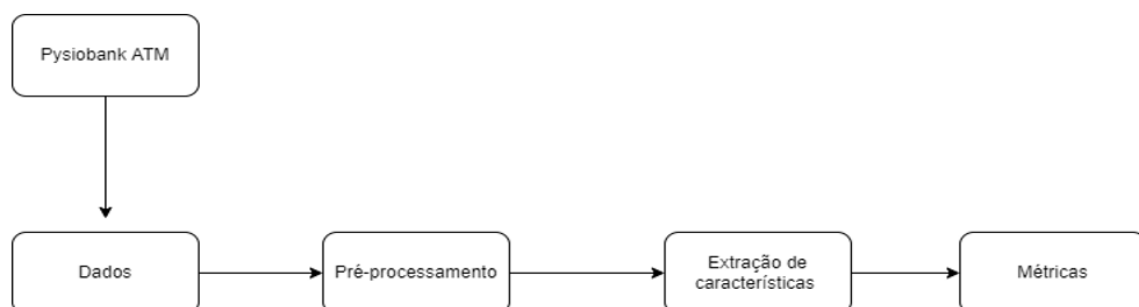
Ela tem proporções epidêmicas, em todo o mundo, por volta de 26 milhões de pessoas são afetadas e ao mesmo tempo que é frequente também é muito mortal, após diagnóstico 30 a 40% das pessoas morrem após um ano, e embora seja mais recorrente em pessoas de idades mais elevadas, ela pode ocorrer em pessoas que não possuem o perfil de uma pessoa propensa à doença.

Devido ao grau de seriedade da doença, busca-se constantemente agregar dados e informações auxiliares ao diagnóstico, este estudo tem o objetivo de apresentar informações de como os dados coletados se comportam fazendo uso de métodos estatísticos para demonstrá-los, a fim de que a informação obtida possa auxiliar profissionais e conseqüentemente os pacientes, para o profissional conhecer ainda mais pontos de vistas sobre como funcionam os dados para que sejam ponderados mais fatores ao se fazer o diagnóstico, e no auxílio deste processo, por meio de sinais cardíacos, pretende-se comparar informações que foram separadas em 2 grupos: um com pacientes com ICC e outro com pacientes saudáveis, e assim ajudar na detecção e auxiliar a identificar as peculiaridades do sinal de pessoas com a doença e assim resultar na diminuição de erros na análise do órgão com o uso do ECG.

## METODOLOGIA

O estudo foi categorizado em quatro partes, sendo elas: Dados, pré-processamento, extração de características e métricas, como mostrado no Fluxograma abaixo.

Figura 1- Fluxograma do método proposto, 1 versão



## BANCO DE DADOS

Foi utilizada duas bases de dados para este estudo, foram a BIDMC Congestive Heart Failure Database (Banco de dados para Insuficiência Cardíaca Congestiva), e também o MIT-BIH Normal Sinus Rhythm Database (Banco de dados de Pacientes Saudáveis). Foram coletados dados de 18 indivíduos com ritmos Sinus (pacientes saudáveis), e 15 com indivíduos com insuficiência cardíaca. E com esses dados foram utilizados para cada procedimento que for mostrado ao decorrer do estudo.

## PRÉ – PROCESSAMENTO

Nesse estágio, fizemos o tratamento dos dados com o objetivo de diminuir qualquer distorção possível nos resultados, e deixar o sinal com uma boa qualidade.

Primeiro igualamos as bases de dados de saudáveis e CCI deixando as com o mesmo tamanho, e retiramos os períodos iniciais e finais dos sinais para eliminar a chance de erro na análise. A parte inicial e final do sinal é um intervalo de tempo que pode ter muitos ruídos que podem causar uma irregularidade na análise, e isso pode afetar na etapa de classificação do sinal.

## EXTRAÇÃO DE CARACTERÍSTICAS

Para fazer a extração de características utilizamos as seguintes técnicas de estatística: média, mediana, assimetria e desvio padrão. As equações para cada variável são descritas abaixo:

Figura 2 - Representação da fórmula da média aritmética

$$\bar{x} = \frac{\sum x_i}{n}$$

Onde:

X - Média aritmética

$\sum x_i$  - Soma dos valores X da amostra

n - Tamanho da amostra

Figura 3 - Representação da Fórmula da Mediana

$$Mediana = l_i + \left[ \frac{\left( \frac{n}{2} - f_{ac\ ant} \right)}{f_i} \right] \times h$$

Onde:

li - Limite inferior da classe da mediana.

Fant ant - Frequência acumulada da classe que da classe anterior.

Fi – Frequência.

h - Amplitude da classe da mediana.

Figura 4 - Representação da Fórmula da Assimetria

$$A = \frac{3(\bar{X} - Md)}{S}$$

Onde:

A - Assimetria

$\bar{X}$  - Média

Md - Mediana

S - Desvio padrão amostral

Figura 5 - Representação da Fórmula do Desvio Padrão

$$DP = \sqrt{\frac{\sum_{i=1}^n (x_i - M_A)^2}{n}}$$

onde:

$\sum$  - símbolo de somatório. Indica que temos que somar todos os termos, desde a primeira posição (i=1) até a posição n

$x_i$  - valor na posição i no conjunto de dados

$M_A$  - Média aritmética dos dados

n - quantidade de dados

## MÉTRICAS

Após a extração das características de cada sinal, os separamos em duas características:

Verdadeiro positivo (VP), em que um registro com ICC é classificado com ICC.

Verdadeiro negativo (VN), em que um registro saudável é classificado como saudável.

E para se obter uma acurácia do modelo utilizamos a soma do VN (Verdadeiro negativo) e VP (Verdadeiro Positivo) o resultado foi dividido pelo número total de amostras.

Acurácia é dada por:

$$Ac = (VP+VN) / n$$

n - o número total de registros do banco de dados.

Para fazermos o cálculo da acurácia contamos a quantidade de VN que se encontra dentro da região de aceitação, qualquer VN fora dessa região é ignorado, fazemos o mesmo com os VP, somamos todos que se encontram fora da região de aceitação, e os que estão dentro da

região de aceitação e ignorado na soma.

## RESULTADOS E DISCUSSÃO

Utilizando as métricas propostas conseguimos uma grande precisão na classificação dos sinais principalmente na média e mediana que a acurácia chegou a 100% de precisão, com resultados expressos em relação a faixa da região de aceitação.

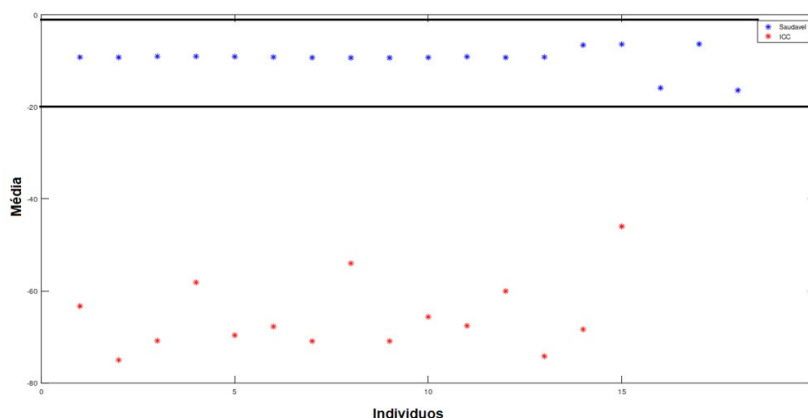
Também analisamos os sinais utilizando a assimetria e o desvio padrão como mostra a tabela 1, na assimetria conseguimos uma acurácia que ficou abaixo das métricas anteriores ficando em torno de 90%. E o desvio padrão ficou em torno de 70% mostrando que essa métrica é ineficaz na extração das características do sinal.

Os gráficos 1 a 4 ilustram a região de aceitação dos dados quando usamos a média (gráfico 1), mediana (gráfico 2) a assimetria (gráfico 3) e o desvio padrão (gráfico 4). Além disso, apresentamos a combinação das três características em um único gráfico 3D, sendo elas media, moda e assimetria (gráfico 5).

**Tabela 1- As Métricas e seus percentuais de Acurácia**

Característica	Acurácia
Média	100%
Mediana	100%,
Assimetria	91%
Desvio Padrão	70%

**Gráfico 1 -representação com os dados da média**



**Gráfico 2 - representação com os dados da mediana**

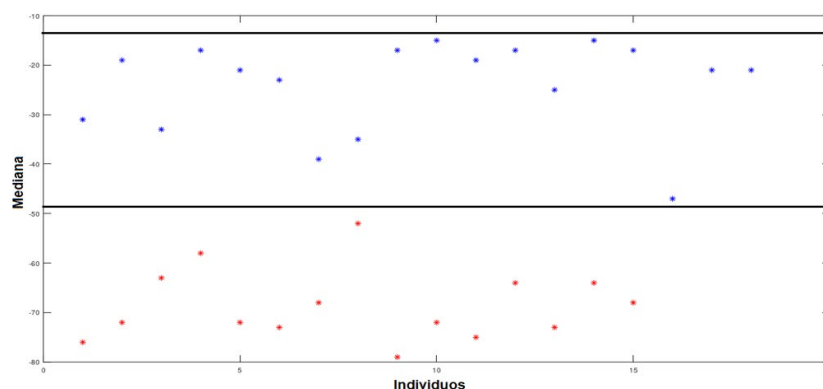


Gráfico 3 - representação com os dados da Assimetria

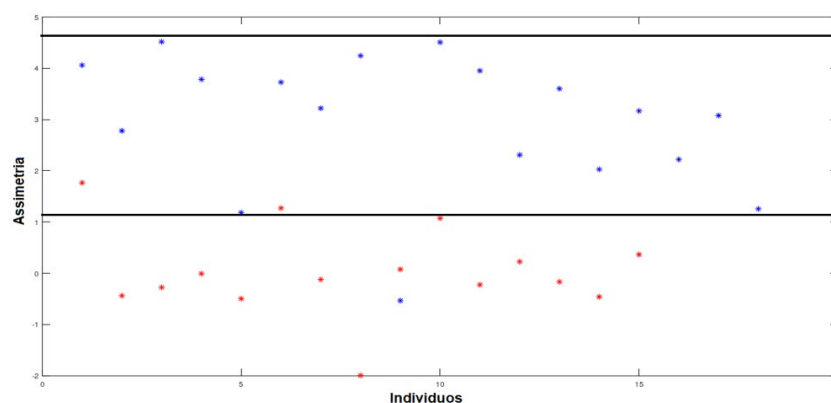


Gráfico 4 - representação com os dados do desvio padrão

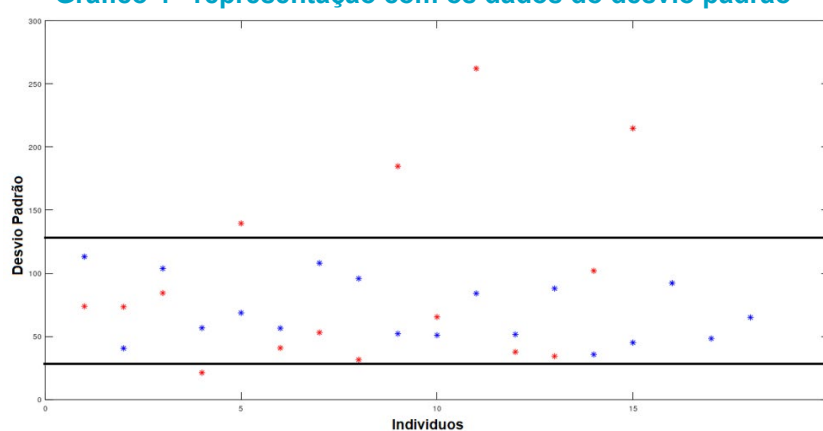
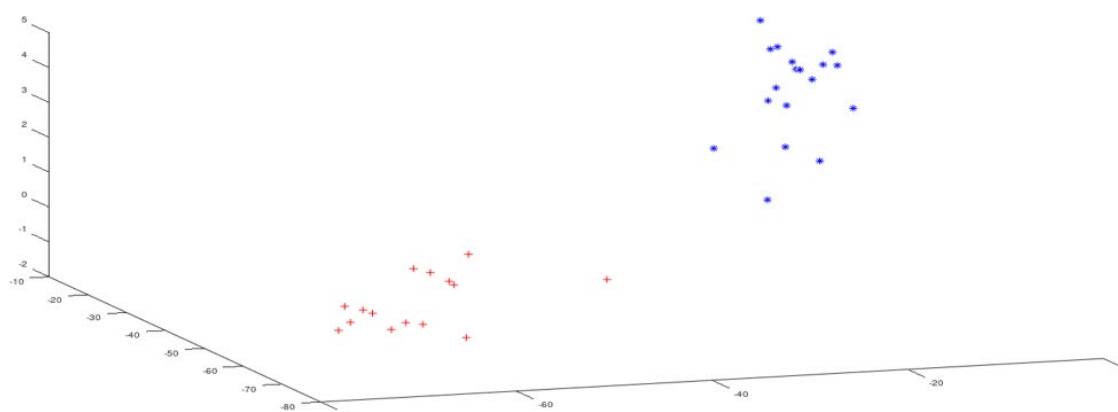


Gráfico 5- Representação dos dados da média, mediana e assimetria no gráfico 3D



## CONSIDERAÇÕES FINAIS

Utilizando a metodologia proposta neste estudo foi possível fazer a extração das características dos sinais de ECG das pessoas saudáveis e com ICC utilizando-se de análise estatística.

Através dos resultados obtidos conseguimos catalogar e separar os resultados verda-

deiro positivo e verdadeiro negativo, delimitamos uma região de aceitação para determinar a eficiência das métricas, com essa separação nos permitiu calcular a acurácia do modelo.

As métricas propostas nos retornaram um alto valor de acurácia, mostrando ser bastante promissor na detecção da ICC. Além disso, as técnicas utilizadas são relativamente simples, e utiliza pouco recurso computacional para poder ser executado.

A simples abordagem do método faz se de grande importância em ambientes com poucos recursos, pois o mesmo só necessita do sinal de ECG. Mostrando que mesmo pode ser utilizado como mais um auxílio para a detecção da ICC. Ajudando a equipe médica na tomada de decisões cabíveis.

## REFERÊNCIAS

Nowell M. fine, MD Insuficiência cardíaca (IC) Última revisão/alteração completa nov 2020. Disponível:<https://www.msdmanuals.com/pt-br/casa/dist%C3%BArbios-do-cora%C3%A7%C3%A3o-e-dos-vasos-sangu%C3%ADneos/insufici%C3%A2ncia-card%C3%ADaca/insufici%C3%A2ncia-card%C3%ADaca-ic>. acesso em 05 dezembro de 2021.

G. A. A. L. G. L. H. J. I. P. C. MR; “The MIT-BIH Normal Sinus Rhythm Database”. PhysioBank, PhysioToolkit, and PhysioNet. Acedido em 24 de novembro de 2021, em: <https://archive.physionet.org/physiobank/database/nsrdb/>

G. A. A. L. G. L. H. J. I. P. C. MR; “The BIDMC Congestive Heart Failure Database”. PhysioBank, PhysioToolkit, and PhysioNet. Acedido em 24 de novembro de 2021, em: <https://archive.physionet.org/physiobank/database/chfdb/>

DA CONSOLAÇÃO V. MOREIRA, maria. Insuficiência cardíaca na era moderna: das melhores evidências para a prática clínica. Revista de medicina de minas gerais. Belo Horizonte, Minas Gerais. data não divulgada