

Clasificación automática de pacientes con Infarto de Miocardio.

RESUMEN

En el presente trabajo se propone una nueva técnica multivariable basada en el procesamiento de las señales Electrocardiográficas (ECG) y Vectorcardiográficas (VCG), con el fin de detectar pacientes con Infarto de Miocardio (IM) y poder diferenciarlos de sujetos sanos. Con este objetivo se proponen y analizan estadísticamente 7 parámetros

(2 VCG, 5 ECG): a) Volumen del bucle QRS, b) Área planar del bucle QRS, c) Vector diferencia en el intervalo ST-T, d) Vector diferencia de la onda T y e-g) Área de la onda T en las derivaciones ortogonales X, Y, Z. Los resultados muestran que estos índices presentan diferencias significativas (p -valor < 0.05) entre las dos poblaciones. Además, con la combinación de ellos es posible clasificar o diferenciar la población de sujetos sanos de la de con IM con una Sensibilidad = 93.81%, una Especificidad = 93.67% y una Eficacia = 93.74%. Se concluye que la técnica propuesta podría ser de gran utilidad, por su bajo costo y sencillez, como una técnica de diagnóstico alternativa en unidades de emergencias y/o de cuidados intensivos.

1. INTRODUCCIÓN

El Infarto de Miocardio (IM) es la necrosis de una porción de tejido cardíaco provocada, normalmente, por el bloqueo total o parcial de los vasos sanguíneos que lo irrigan. Luego, la entrega insuficiente de oxígeno y nutrientes a las células miocárdicas, o Isquemia Cardíaca (IC), provoca la depleción de las reservas de Adenosina Trifosfato (ATP) celular. Si esta es sostenida en el tiempo causa la necrosis o infarto de las células miocárdicas. Asimismo,

R Correa¹

P Arini²

L Correa³

M Valentinuzzi⁴

E Laciari⁵

¹ Gabinete de Tecnología Médica
E-mail: rcorrea@gateme.unsj.edu.ar
Facultad de Ingeniería
Universidad Nacional de San Juan
CONICET

² Instituto Argentino de Matemática.
E-mail: pedro.arini@conicet.gov.ar
CONICET.
Instituto de Ingeniería Biomédica
Facultad de Ingeniería
Universidad de Buenos Aires

³ Gabinete de Tecnología Médica
E-mail: lcorrea@gateme.unsj.edu.ar
Facultad de Ingeniería
Universidad Nacional de San Juan
CONICET

⁴ Instituto de Ingeniería Biomédica
E-mail: maxvalentinuzzi@arnet.com.ar
Facultad de Ingeniería
Universidad de Buenos Aires

⁵ Gabinete de Tecnología Médica
E-mail: laciari@gateme.unsj.edu.ar
Facultad de Ingeniería
Universidad Nacional de San Juan
CONICET

esta disminución abrupta de la concentración de ATP celular produce una disociación electromecánica capaz de generar arritmias y hasta la fibrilación ventricular. Esto, asociado al tamaño y la localización de la zona afectada por el IM, puede desencadenar la muerte súbita (o aguda) del sujeto que la padece. Así, el IM es la principal causa de muerte de hombres y mujeres en todo el mundo[1]

Además, la disminución del ATP celular provoca cambios en los potenciales transmembrana dependientes de energía, que se registran como cambios en el Electrocardiograma (ECG) y Vectorcardiograma (VCG) de superficie durante la repolarización ventricular. Así se producen Desplazamientos del Segmento ST, Dispersión de la Repolarización Ventricular y Alternancias de Onda T. De la misma manera, otros investigadores han demostrado que también se registran modificaciones en estas señales durante la Despolarización, tales como el intervalo QT, la forma y tamaño de la onda Q [1-3]

Asimismo, apoyándose en el empleo de nuevas tecnologías y avances en el análisis digital del ECG, se han desarrollado, en los últimos años, varias técnicas que utilizan el VCG construido a partir del ECG para el monitoreo de pacientes con IC [4,5] y antes o después de un IM [6,7]. De la misma manera, varios investigadores han utilizado técnicas basadas en el ECG y el VCG para la detección, determinación del tamaño de la zona afectada y localización del IM[6,8,9].

En el presente trabajo se desarrolló una nueva técnica de procesamiento de la señales ECG y VCG con el fin de identificar pacientes con IM y poder diferenciarlos de sujetos sanos. La hipótesis de este trabajo es que a partir del análisis discriminante lineal de un conjunto de parámetros ECG y VCG es posible detectar con gran sensibilidad y especificidad cuándo un paciente ha padecido un IM. La gran ventaja de la técnica propuesta, basada en el ECG y VCG, es que es no-invasiva, de muy bajo costo y que se la puede realizar de manera repetitiva sin causar ningún daño al paciente.

2. MATERIALES

En este estudio se utilizaron los registros electrocardiográficos de 97 pacientes con IM (71 varones y 26 mujeres, con edades de 58 +/- 10 y 64 +/- 12 años respectivamente) obtenidos durante la semana posterior al infarto. Como grupo de control se utilizaron los registros de 52 sujetos sanos (39 varones y 13 mujeres, con edades de 42 +/- 14 y 48 +/- 19 años respectivamente). Todos estos registros fueron obtenidos de la base de datos Physikalisch-Technische Bundesanstalt (PTB) adquirida por el Instituto Nacional de Metrología de Alemania¹. Cada registro ECG contiene 15 derivaciones obtenidas de manera simultánea: las 12 derivaciones estándar (I, II, III, aVR, aVL, aVF, V1-V6) más las 3 derivaciones ortogonales de Frank (X, Y, Z). Cada señal fue digitalizada con una frecuencia de muestreo de 1000Hz y con una resolución en amplitud de 16bits [10]. Si bien la base de datos provee estas 15 derivaciones para cada registro ECG, en este trabajo solo se utilizaron las derivaciones ortogonales (X, Y, Z).

La región necrótica del corazón en los registros ECG de los pacientes con IM fue: anterior (n=14), antero-lateral (n=10), antero-septal (n=24), antero-septo-lateral (n=1), inferior (n=23), ínfero-lateral (n=17), ínfero-postero-lateral (n=6), lateral (n=1), posterior (n=1).

3. METODOLOGÍA

Todos los registros ECG (X, Y, Z) provenientes de la base de datos antes descripta fueron preprocesados de la siguiente manera: a) se realizó un filtrado, pasa banda (Butterworth, 4to orden, 0.2-100 Hz, bidireccional) para atenuar los ruidos alta y baja frecuencia, además de un filtro ranura (Butterworth, 2do orden, 50/60 Hz, bidireccional) para atenuar la interferencia de línea; b) se aplicó un filtro basado en interpolación cúbica de splines, para atenuar aún más los artefactos respiratorios y la deriva de línea; c) una vez terminada la etapa de filtrado se aplicó un de-

¹ <http://www.physionet.org/physiobank/database/ptbdb/>

lineador que determina el punto de máxima amplitud, el comienzo y fin de cada onda del ECG [5]; d) luego se excluyeron los latidos excesivamente ruidosos (con un nivel de ruido $RMS > 40 \mu V$, medidos en una ventana de 40 ms localizada a 2/3 de cada intervalo RR); e) por último, los latidos ectópicos también fueron excluidos por medio de una comparación latido a latido contra uno de referencia representativo de la señal a través de la técnica de correlación cruzada. En este estudio se eligió visualmente para cada ECG un latido sinusal de bajo ruido como latido de referencia.

Una vez finalizada la etapa de adaptación (o pre-procesado) de las señales ECG se procedió con el análisis de las mismas. En la Fig. 1 se ilustra un esquema del mismo.

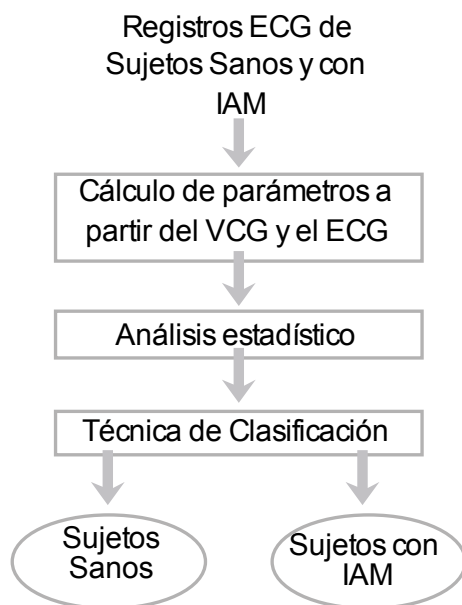


Figura 1. Esquema General de la técnica propuesta.

3.1. Cálculo de los parámetros

En este estudio se calcularon 7 parámetros vectorcardiográficos: 2 medidos durante la despolarización ventricular y 5 sobre la repolarización ventricular. Cabe destacar, que algunos de estos parámetros fueron propuestos con anterioridad por este grupo de investigadores para el estudio de los cambios dinámicos registrados en el ECG y VCG durante un episodio de isquemia cardíaca aguda[5,7].

De esta manera, los índices analizados y evaluados en este trabajo se describen como:

- Volumen del bucle QRS (QRSV) [mV3]: consiste en estimar primeramente el conjunto de puntos que producen el volumen convexo mínimo que encierra todos los puntos del bucle QRS, utilizando el algoritmo Convex Hull [11]. Luego, es evaluado el volumen convexo del bucle QRS. El fin de este parámetro es cuantificar la planaridad y morfología de este bucle en el espacio 3-D [5].

- Área planar del bucle QRS (QRSAP) [mV2]: es la estimación del área de la superficie interna del bucle QRS calculada en la proyección de este sobre el plano de mejor ajuste obtenido por la técnica de mínimos cuadrados. Se estima que esta área está relacionada con anomalías hemodinámicas en las enfermedades cardíacas [5] (figura 2-a).

- Vector diferencia en el intervalo ST-T (ST-TVD) [mVs]: Se define como el área diferencial entre la señal ECG en el intervalo ST-T corriente (desde cada punto J hasta el fin de la onda T) y el de referencia de la señal ECG (también en el intervalo ST-T) evaluado durante los primeros 30s de cada registro ECG. El objetivo de este parámetro es evaluar de una forma general los cambios ocurridos durante la repolarización ventricular [7] (figura 2-b).

- Vector diferencia de la onda T (TVD) [mVs]: Se define como el área diferencial entre cada onda T de la señal ECG y la de referencia evaluado durante los primeros 30s de cada registro ECG [7] (figura 2-c). El objetivo de este parámetro es estimar los cambios producidos durante la parte media y final la repolarización ventricular.

- Área de la onda T en las derivaciones X, Y, Z (aTX, aTY, aTZ) [mVs]: es área encerrada entre el complejo QRS y el eje de las abscisas en cada una de las derivaciones ortogonales. Es una estimación de la energía de la despolarización ventricular (figura 2-d).

3.2. Análisis Estadístico

Con el conjunto de parámetros evaluados en las dos poblaciones (sujetos sanos y Pacientes con IM) se realizó un análisis descriptivo de cada uno de ellos comparando las medias y la desviación estándar.

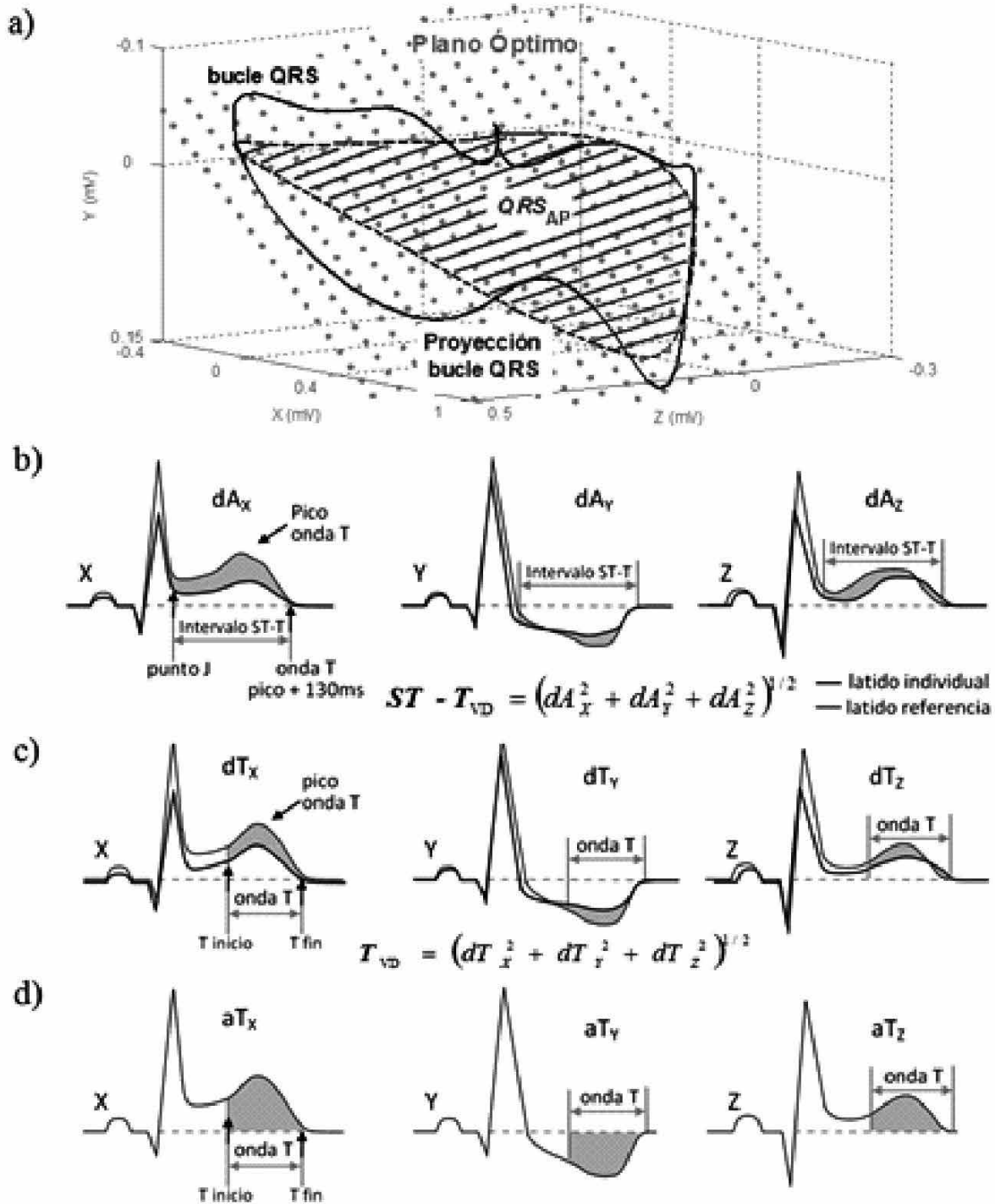


Figura 2. Cálculo de parámetros se puede observar en: a) un bucle del Complejo QRS donde se estiman el QRS_V y el QRS_{AP} ; b) las áreas sombreadas dentro del intervalo ST-T utilizadas para en cálculo del $ST-T_{VD}$; c y d) las áreas bajo la onda T el usadas en el computo del T_{VD} en c) y aT_x , aT_y y aT_z en d).

3.3. Técnica de Clasificación

Luego de realizar el análisis descriptivo, se aplicó la técnica de Análisis Discriminante Lineal (ADL) [12] a los parámetros obtenidos para cada registro ECG de cada paciente. Con el fin de determinar si a partir de los mismos es posible clasificar los diferentes registros ECG (de cada paciente) como Sujetos Sanos o como pacientes con IM.

Básicamente, un clasificador ADL es una combinación lineal de las variables discriminantes, esto es:

$$y = \mu_0 + \mu_1 X_1 + \mu_2 X_2 + \dots + \mu_p X_p \quad (1)$$

donde y es el valor de salida de la función discriminante;

μ_n (con $n=1, \dots, p$) son los coeficientes de las variables discriminantes.

X_n son las variables discriminantes a las que se le asignan los valores de los parámetros propuestos y p es el número de variables en el análisis.

La función discriminante puede ser utilizada para clasificar cada registro ECG como perteneciente a un grupo determinado (Sano o con IM).

La técnica ADL estima los coeficientes de la función discriminante a partir de un subconjunto de registros ECG cuyo grupo es conocido (denominado de entrenamiento), para lo que se utiliza un grupo de 52 sujetos sano y otro del mismo tamaño de pacientes con IM.

Estos últimos seleccionados aleatoriamente de los 97 disponibles en la base de datos. Asimismo, como esta selección modifica el desempeño del clasificador se repitió 200 veces, de manera que los resultados obtenidos son el promedio de estas iteraciones del clasificador.

Luego, se aplicó el método de validación cruzada denominado dejando uno afuera (o loo-cv de sus siglas en inglés: leave one out cross validation). Este método utiliza todo el grupo de características ECG-VCG seleccionadas para la estimación de la función discriminante, menos una que se utiliza para la posterior validación de la función estimada. Así, se repite este algoritmo hasta que todos los conjuntos de características han sido utilizados (dejados afuera) para la validación.

4. RESULTADOS

Los resultados del análisis descriptivo se muestran en la Tabla I en la que se pueden observar los valores medios y desviaciones estándar calculados para cada registro ECG de Sujetos Sanos y de los con IM. En esta tabla también se incluyen los valores de la significancia estadística (p-valor) de la comparación de las medias entre estos grupos. En los pacientes que presentan más de un registro ECG durante la primer semana se promediaron los valores de sus parámetros, de esta manera a cada paciente le correspondió un solo valor de cada parámetro.

Tabla 1 Valores medios (VM) y desviaciones estándar (DE) de cada parámetro para ambas poblaciones y significancia estadística (p-valor)

Sujetos		QRS_V	QRS_{AP}	$ST-T_{VD}$	T_{VD}	aT_x	aT_y	aT_z
		[mV ³]	[mV ²]	[mVs]	[mVs]	[mVs]	[mVs]	[mVs]
Sanos	VM	0,15	1,22	4,43	4,36	40,25	21,93	-22,11
	DE	0,16	0,65	1,49	3,18	20,69	11,92	17,96
IM	VM	0,06	0,55	5,11	15,38	9,56	10,51	-8,06
	DE	0,04	0,29	2,57	23,66	16,57	18,15	24,46
p-valor		<,001	<,001	,091	<,001	<,001	<,001	,001

Con el fin de evaluar el desempeño de la clasificación de cada uno de los registros ECG del grupo de Validación, se ha calculado 3 índices que comparan los resultados predichos por el algoritmo con la clasificación otorgada por la Base de datos. Estos índices son Sensibilidad (Sen), Especificidad (Esp) y Eficacia (Ef) o exactitud. Así, en la Tabla II se pueden observar los resultados de la clasificación utilizando la técnica de ADL para cada uno de los parámetros estudiados y también

para la mejor combinación de ellos (MC). En este estudio se determinó la mejor combinación de los parámetros utilizando la técnica de minimización del Lambda de Wilks, que permite determinar el menor conjunto variables determinantes que más contribuyen a la diferenciación de los grupos. De esta manera, se estableció que la mejor combinación es QRS_V , QRS_{AP} , $ST-T_{VD}$, T_{VD} y aT_X .

Tabla 2 Resultados de la clasificación utilizando cada uno de los parámetros estudiados y la Mejor Combinación (MC)

	QRS_V	QRS_{AP}	$ST-T_{VD}$	T_{VD}	aT_X	aT_Y	aT_Z	MC
Sen (%)	69,79	90,76	56,69	71,34	79,45	65,62	73,93	93,81
Esp (%)	76,81	73,69	61,94	79,63	84,00	72,50	54,00	93,67
Ef (%)	72,29	84,69	58,56	74,29	81,07	68,07	66,84	93,74

5. DISCUSIÓN Y CONCLUSIONES

Durante estos últimos años varios investigadores han propuesto diferentes técnicas de clasificación para identificar pacientes con IM, basadas en el ECG de superficie. Entre ellas podemos citar a Bakul et al.[14] quienes propusieron un conjunto de características denominadas Relative Frequency Band Coefficient para la identificación automática del riesgo de infarto de cardíaco, alcanzando una $Sen=85,57\%$, $Esp=83,97\%$ y una $Ef=85,23\%$. Asimismo, Keshtkar et al. [9] propusieron la evaluación de un conjunto de coeficientes Wavelet calculados sobre el ECG de señal promediada, a través de Redes neuronales como índices para detectar el IM logrando una $Sen=93\%$, una $Esp=86\%$ y una $Ef=89,5$. Por su parte, Maharaj and Alonso [15] utilizaron un clasificador discriminante multivariable basado en la descomposición multiescala wavelet de la señal ECG obteniendo una Sen entre el 80% al 90% y una Esp de 90%. No obstante, aunque todas estas técnicas tienen sus ventajas y desventajas, ninguna hace referencia a los datos de los pacientes utilizados (edad, sexo, tiempo transcurrido desde

que sufrió el IM, etc.). Tampoco se informa cuántos registros de cada paciente fueron utilizados. Finalmente, la mayoría clasifican un conjunto de latidos de cada registro ECG sin relacionarlo con el paciente del cual fueron extraídos.

En este trabajo, en cambio, se propone una nueva técnica multivariable para la identificación de pacientes con IM, utilizando para cada paciente solo los registros ECG obtenidos durante la primera semana posterior al IM. Para ello se propusieron 7 parámetros: 2 medidos durante la despolarización ventricular QRS_V y QRS_{AP} y 5 medidos durante la repolarización ventricular $ST-T_{VD}$, T_{VD} y aT (calculado en X, Y, Z).

En la tabla I se exponen los Valores medios (VM) y desviaciones estándar (DE) de cada parámetro propuesto para ambas poblaciones. En ella se observa que los parámetros tienen una elevada dispersión ya que su varianza a menudo tiene valores próximos a su media. No obstante, al comparar las medias poblacionales se pueden ver diferencias significativas (p -valor < 0.05) en todos los parámetros excepto $ST-T_{VD}$ lo que sugiere que los

mismos pueden ser utilizados para diferenciar ambas poblaciones.

A partir de esta característica de los índices estudiados, es decir que permiten diferenciar entre la población de sujetos sano y la de pacientes con IM, en la Tabla II se resumen los valores de los índices de desempeño (Sen, Esp y Ef) del clasificador propuesto utilizado para evaluar el poder discriminante cada uno de los parámetros estudiados y la mejor combinación de ellos. En esta tabla se puede ver que los parámetros con mayor poder discriminante individual son QRS_{AP} y aT_x y que cuando los combinamos con el QRS_V , $ST-T_{VD}$ y T_{VD} se pueden clasificar los sujetos sanos y con IM con una Sen = 93.81%, Esp= 93.67% y una Ef= 93.74%. Lo que demuestra el excelente desempeño de la técnica de clasificación propuesta cuando es comparada con las propuestas por otros investigadores [9,14,15].

Se concluye que la nueva técnica multivariable de identificación de pacientes con IM basada en índices ECG y VCG logra diferenciar con un excelente desempeño las poblaciones de sujetos sanos y con IM, y que podría ser utilizada como una técnica de diagnóstico alternativa en una unidad de emergencias médicas.

AGRADECIMIENTOS

Todos los autores son financiados por el Consejo Nacional de Investigaciones Científicas y Técnicas. Asimismo, este trabajo es parte de los proyectos de la Agencia Nacional de Promoción Científica y Tecnológica (ANPCYT – PICT-O-UNSJ 0027), de la Universidad Nacional de San Juan (CICICT-CA-UNSJ I972).

REFERENCE

1. Thygesen K, Alpert JS, Jaffe AS, Simoons ML, Chaitman BR, et al. (2012) Third Universal Definition of Myocardial Infarction. *Journal of the American College of Cardiology* 60: 1581–1598. Available: <http://www.sciencedirect.com/science/article/pii/S0735109712028963>.
2. Luna A Bayés de (2007) Nueva terminología de las paredes del corazón y nueva cla-

sificación electrocardiográfica de los infartos con onda Q basada en la correlación con la resonancia magnética. *Revista Española de Cardiología* 60: 683–689.

3. MEMBERS* WC, O’Gara PT, Kushner FG, Ascheim DD, Casey DE, et al. (2013) 2013 ACCF/AHA Guideline for the Management of ST-Elevation Myocardial Infarction: A Report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. *Circulation* 127: e362–e425. Available: <http://circ.ahajournals.org/content/127/4/e362.short>.

4. Correa R, Arini PD, Correa L, Valentinuzzi ME, Laciari E (2013) Acute myocardial ischemia monitoring before and during angioplasty by a novel vectorcardiographic parameter set. *Journal of electrocardiology* 46: 635–643.

5. Correa R, Arini PD, Valentinuzzi ME, Laciari E (2013) Novel set of vectorcardiographic parameters for the identification of ischemic patients. *Medical Engineering & Physics* 35: 16–22. Available: <http://www.sciencedirect.com/science/article/pii/S1350453312000537>.

6. Arif M, Malagore IA, Afsar FA (2012) Detection and localization of myocardial infarction using K-nearest neighbor classifier. *Journal of medical systems* 36: 279–289.

7. Correa R, Arini PD, Correa LS, Valentinuzzi M, Laciari E (2014) Novel technique for ST-T interval characterization in patients with acute myocardial ischemia. *Computers in biology and medicine* 50: 49–55.

8. Hui Yang STSB, Komanduri R (2012) Spatiotemporal representation of cardiac vectorcardiogram (VCG) signals. *BioMedical Engineering OnLine* 11:16: 1–15.

9. Keshtkar A, Seyedarabi H, Sheikhzadeh P, others (2013) discriminant analysis between myocardial infarction patients and healthy subjects using Wavelet Transformed signal averaged electrocardiogram and probabilistic neural network. *Journal of Medical Signals and Sensors* 3.

10. Boussejot R, Kreiseler D, Schnabel A (1995) Nutzung der EKG-Signal daten bank CARDIODAT der PTB über das Internet. *Bio-*

medizinische Technik, Band 40, Ergänzungsband 1: 317.

11. Barber CB, Dobkin DP, Huhdanpaa H (1996) The Quickhull Algorithm for Convex Hulls. *ACM Trans Math Softw* 22: 469–483. Available: <http://doi.acm.org/10.1145/235815.235821>.

12. Flores J, Jiménez E, G G (2001) Análisis Discriminante. Editorial La Muralla, S.A.; Edición: 1.

13. Altman D (n.d.) Practical statistics for medical research. 1991, editor Chapman and Hall/CRC; 1st ed edition.

14. Bakul G, Tiwary U (2010) Automated risk identification of myocardial infarction using relative frequency band coefficient (RFBC) features from ECG. *The open biomedical engineering journal* 4: 217.

15. Maharaj EA, Alonso AM (2012) Discriminant analysis of multivariate time series using Wavelets. *JEL Classification Statistics and Econometrics Series*: C38; C22.