



Universitat de Girona

ENFOQUES METODOLÓGICOS PARA MEJORAR LA PRECISIÓN Y FIABILIDAD DE LOS SISTEMAS DE MONITORIZACIÓN CONTINUA DE GLUCOSA EN PACIENTES CON DIABETES TIPO I Y EN PACIENTES CRÍTICOS MEDIANTE TÉCNICAS LINEALES Y NO LINEALES

Yenny Teresa LEAL MONCADA

Dipòsit legal: Gi. 1271-2013

<http://hdl.handle.net/10803/120377>

ADVERTIMENT. L'accés als continguts d'aquesta tesi doctoral i la seva utilització ha de respectar els drets de la persona autora. Pot ser utilitzada per a consulta o estudi personal, així com en activitats o materials d'investigació i docència en els termes establerts a l'art. 32 del Text Refós de la Llei de Propietat Intel·lectual (RDL 1/1996). Per altres utilitzacions es requereix l'autorització prèvia i expressa de la persona autora. En qualsevol cas, en la utilització dels seus continguts caldrà indicar de forma clara el nom i cognoms de la persona autora i el títol de la tesi doctoral. No s'autoritza la seva reproducció o altres formes d'explotació efectuades amb finalitats de lucre ni la seva comunicació pública des d'un lloc aliè al servei TDX. Tampoc s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX (framing). Aquesta reserva de drets afecta tant als continguts de la tesi com als seus resums i índexs.

ADVERTENCIA. El acceso a los contenidos de esta tesis doctoral y su utilización debe respetar los derechos de la persona autora. Puede ser utilizada para consulta o estudio personal, así como en actividades o materiales de investigación y docencia en los términos establecidos en el art. 32 del Texto Refundido de la Ley de Propiedad Intelectual (RDL 1/1996). Para otros usos se requiere la autorización previa y expresa de la persona autora. En cualquier caso, en la utilización de sus contenidos se deberá indicar de forma clara el nombre y apellidos de la persona autora y el título de la tesis doctoral. No se autoriza su reproducción u otras formas de explotación efectuadas con fines lucrativos ni su comunicación pública desde un sitio ajeno al servicio TDR. Tampoco se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR (framing). Esta reserva de derechos afecta tanto al contenido de la tesis como a sus resúmenes e índices.

WARNING. Access to the contents of this doctoral thesis and its use must respect the rights of the author. It can be used for reference or private study, as well as research and learning activities or materials in the terms established by the 32nd article of the Spanish Consolidated Copyright Act (RDL 1/1996). Express and previous authorization of the author is required for any other uses. In any case, when using its content, full name of the author and title of the thesis must be clearly indicated. Reproduction or other forms of for profit use or public communication from outside TDX service is not allowed. Presentation of its content in a window or frame external to TDX (framing) is not authorized either. These rights affect both the content of the thesis and its abstracts and indexes.



Universitat de Girona

Tesis Doctoral

Enfoques metodológicos para mejorar la precisión y fiabilidad de los sistemas de monitorización continua de glucosa en pacientes con diabetes tipo 1 y en pacientes críticos mediante técnicas lineales y no lineales

Yenny Teresa Leal Moncada
2013





Universitat de Girona

TESIS DOCTORAL

**Enfoques metodológicos para mejorar la
precisión y fiabilidad de los sistemas de
monitorización continua de glucosa en
pacientes con diabetes tipo 1 y en pacientes
críticos mediante técnicas lineales y no
lineales**

Yenny Teresa Leal Moncada

2013

Programa Doctoral en Tecnología

Director y Co-directora:
Dr. Josep Vehí y Dra. Magda L. Ruíz O.

Tesis presentada para optar al título de Doctora por la Universitat de Girona

Lista de Publicaciones

Esta es la lista de las publicaciones en revistas y conferencias derivadas de esta tesis:

Artículos de revistas

Leal, Y., Garcia-Gabin, W., Bondia, J., Esteve, E., Ricart, W., Fernández-Real, J.M., Vehí, J. (2010) “Real-time glucose estimation algorithm for continuous glucose monitoring using autoregressive models.” *Journal of Diabetes Science and Technology*. 4(2):391-403. PMID: 20307401 [PubMed - indexado en MEDLINE].

Leal, Y., Garcia-Gabin, W., Bondia, J., Esteve, E., Ricart, W., Fernández-Real J.M., Vehí, J. (2010) “Enhanced algorithm for glucose estimation using the continuous glucose monitoring system.” *Medical Science Monitor*. 16(6):MT51-8. IF: 1.699 (2010) Q3 [indexado en JCR]. PMID: 20512100 [PubMed - indexado en MEDLINE].

Lorenzo, C., **Leal, Y.**, Bonet, A., Bondia, J., Palerm, C.C., Tache, A., Sirvent, J.M., Vehi, J. (2012) “Real-time continuous glucose monitoring in an intensive care unit: better accuracy in patients with septic shock.” *Diabetes Technology and Therapeutics*, 14(7):568-75. IF: 2.205 (2012) Q3 [indexado en JCR]. PMID: 22512288 [PubMed - indexado en MEDLINE].

Leal, Y., Gonzalez-Abril, L., Lorenzo, C., Bondia, J., Vehí, J. (2013) “Detection of correct and incorrect measurements in real-time continuous glucose monitoring systems by applying a post-processing support vector machine.” *IEEE Transactions on Biomedical Engineering*. *En impresión*, (DOI: 10.1109/TBME.2013.2244092). IF: 2.348 (2012) Q2 [indexado en JCR].

Leal, Y., Ruiz, M., Lorenzo, C., Bondia, J., Mujica, L, Vehí, J. (2013) “Principal component analysis in combination with case-based reasoning for detecting therapeutically correct and incorrect measurements in continuous glucose monitoring systems.” *Biomedical Signal Processing and Control*. *En impresión*, (DOI: 10.1016/j.bspc.2013.05.008). IF: 1.074 (2012) Q3 [indexado en JCR].

Artículos de Conferencias

Lorenzo, C., **Leal, Y.T.**, Bonet, A., Vehí, J., Bondia, J., Taché, A., Sirvent, J.M. (2010) “Accuracy assessment of real-time continuous glucose monitors in critically ill patients: Preliminary results.” En *23rd European Society of Intensive Care Medicine Annual Congress*. 9-13 de Octubre, Barcelona, España.

Leal, Y., Lorenzo, C., Bondia, J., Bonet, A., Vehí, J. (2011) “Accuracy assessment of real-time continuous glucose monitoring system in critically-ill patients.” En *Workshop on Control, Dynamics, Monitoring and Applications*. 7-9 de Febrero, Caldes de Montbui, España.

Lorencio, C., **Leal, Y.**, Bonet, A., Vehí, J., Bondia, J., Palerm, C.C., Taché, A., Sirvent, J.M. (2011) “Accuracy assessment of subcutaneous continuous glucose monitoring in an intensive care unit according to the septic shock state.” En *24th European Society of Intensive Care Medicine Annual Congress*. 1-5 de Octubre, Berlín, Alemania.

Leal, Y., Gonzalez-Abril, L., Lorencio, C., Ruiz, M., Bondia, J., Vehí, J. (2012) “Un nuevo enfoque para detectar medidas de glucosa erróneas en los Sistemas de Monitorización Continuos de Glucosa.” En *XIV Jornadas sobre Sistemas cualitativos y sus Aplicaciones en Diagnosis, Robótica e Inteligencia Ambiental (JARCA)*. 24-28 de Junio, Tarragona, España.

Leal, Y., Gonzalez-Abril, L., Lorencio, C., Vehí, J. (2013) “Detección de medidas incorrectas en los SMCs aplicando una nueva estrategia de post-procesamiento de SVM para ajustar el sesgo en función de la sensibilidad.” En *XV Jornadas sobre Sistemas cualitativos y sus Aplicaciones en Diagnosis, Robótica e Inteligencia Ambiental (JARCA)*. 24-27 de Junio, Murcia, España.

Pósters

Leal, Y., Garcia-Gabin, W., Bondia, J., Vehí, J. (2009) “Real-time glucose estimation algorithm for continuous glucose monitoring using autoregressive models.” Póster, En *9th Diabetes Technology Meeting (DTM)*. 5-7 de Noviembre, San Francisco, EE. UU.

Leal, Y., Lorencio, C., Bondia, J., Palerm, C.C., Bonet, A., Vehí, J. (2010) “Accuracy assessment of real-time continuous glucose monitoring systems in an intensive care unit: A pilot study in medical and surgical patients.” Póster, En *10th Diabetes Technology Meeting (DTM)*. 11-13 de Noviembre, Bethesda, EE. UU.

Leal, Y., Lorencio, C., Bondia, J., Palerm, C.C., Sirvent, J.M., Bonet, A., Vehí, J. (2011) “Numerical-accuracy assessment of a real-time continuous glucose monitoring system in medical, surgical, and trauma critically-ill patients.” Póster, En *4th International Conference on Advanced Technologies and Treatments for Diabetes (ATTD)*. 16-19 de Febrero, Londres, Reino Unido.

Lorencio, C., **Leal, Y.**, Bonet, A., Vehí, J., Bondia, J., Taché, A., Sirvent, J.M., Palerm, C.C. (2011) “Valoración de la precisión de la monitorización continua de la glucosa intersticial en pacientes críticos.” Póster, En *XXVI Congreso Nacional de la Sociedad Española de Nutrición Parenteral y Enteral (SENPE)*. 24-27 de Mayo, Gerona, España.

Lorencio, C., **Leal, Y.**, Bonet, A., Vehí, J., Bondia, J., Taché, A., Palerm, C.C., Sirvent, J.M. (2011) “Variabilidad de la precisión de un sistema de monitorización continua de glucosa en pacientes críticos dependiendo de su situación hemodinámica.” Póster, En *XLVI Congreso Nacional de la Sociedad Española de Medicina Intensiva, Crítica y Unidades Coronarias (SEMICYUC)*. 12-15 de Junio, Bilbao, España.

Leal, Y., Ruiz, M., Lorencio, C., Bondia, J., Vehí, J. (2012) “Statistical methods and expert systems for the detection of therapeutically correct and incorrect measurements in RTCGMS.” Póster, En *5th International Conference on Advanced Technologies and Treatments for Diabetes*

(*ATTD*). 8-11 de Febrero, Barcelona, España.

Leal, Y., Gonzalez-Abril, L., Ruiz, M., Lorencio, C., Bondia, J., Ortega, JA, Vehí, J. (2012) “Applying new support vector machines post-processing strategies for the detection of correct and incorrect measurements in CGMS.” Póster, En *12th Diabetes Technology Meeting (DTM)*. 8-10 de Noviembre, Bethesda, EE. UU.

Patente

Bondia, J., Barceló-Rico, F., Díez, J.L., Rossetti, P., Vehí, J., **Leal, Y.** (2012). *Sistema y método de estimación de glucosa en plasma*. Número de publicación: WO2012156568 A1 (22 de Noviembre 2012). Número de solicitud: PCT/ES2012/070358 (21 de Mayo 2012).

Glosario

Los siguientes acrónimos, abreviaturas y variables se utilizan en esta tesis.

Acrónimos y abreviaturas

2SR	Recuperación en dos pasos	Two-steps retrieval
3SR	Recuperación en tres pasos	Three-steps retrieval
A-Gold™	Algoritmo del CGMS® Gold™	CGMS® Gold™ Algorithm
A-Guardian®	Algoritmo del Guardian® RTCGMS	Guardian® REAL-Time CGMS Algorithm
RTCGMS		
ABG	Glucosa en plasma arterial	Arterial Blood Glucose
ABMAR-SMCG	Algoritmo basado en modelos autorregresivos para el SMCG	Algorithm for Continuous Glucose Monitoring Using Autoregressive Models
AD	Diferencia absoluta	Absolute Difference
ADA	Asociación Americana de Diabetes	American Diabetes Association
AP	Páncreas artificial	Artificial Pancreas
APACHE	Sistema de Clasificación de Severidad de Enfermedad APACHE II	Acute Physiology and Chronic Health Evaluation
AR	Auto-regresivo	AutoRegressive
ARMAX Models	Modelos autorregresivos de media móvil con entrada exógena	AutoRegressive Moving Average models with eXogeneous input
ARP-SMCG	Algoritmo retrospectivo propuesto para el SMCG	Proposed Retrospective Algorithm
ARX Models	Modelos autorregresivos con entrada exógena	AutoRegressive models with eXogeneous input
BC	Base de Casos	Case Base
BG	Glucosa en plasma o en sangre	Blood Glucose
BJ Models	Modelos Box-Jenkins	Box-Jenkins models
CBR	Razonamiento basado en casos	Case-Based Reasoning
CG-EGA	Rejilla de análisis de errores de glucosa continua	Continuous Glucose Error Grid Analysis
CIII	Insulinoterapia endovenosa en perfusión continua	Continuous Intravenous Insulin Infusion
Clarke EGA	Rejilla de análisis de errores de Clarke	Clarke Error Grid Analysis
Consensus EGA	Rejilla de análisis de errores consensuada	Consensus Error Grid Analysis
Criterios ISO	Criterios definidos en la norma ISO 15197	International Organization for Standardization criteria
CSII	Infusión continua de insulina subcutánea	Continuous Subcutaneous Insulin Infusion
DCCT	Estudio sobre Control de la Diabetes y sus Complicaciones	Diabetes Control and Complications Trial
DM	Diabetes Mellitus	Diabetes Mellitus
DM-1	Diabetes mellitus tipo 1	Type 1 Diabetes Mellitus
DM-2	Diabetes mellitus tipo 2	Type 2 Diabetes Mellitus
EGA	Rejilla de análisis de errores	Error Grid Analysis
FDA	Agencia de Drogas y Alimentos	Food and Drug Administration
Gold standard	Medida de glucosa en sangre de referencia	Gold standard
GOx	Glucosa-Oxidasa	Glucose-Oxidase
IDF	Federación Internacional de Diabetes	International Diabetes Federation
IG	Glucosa intersticial	Interstitial Glucose
IQR	Rango intercuartílico	Interquartile Range
ISF	Fluido intersticial	Interstitial Fluid
JDRF	Fundación para la Investigación de la Diabetes Juvenil	Juvenile Diabetes Research Foundation
MA	Media móvil	Moving Average
MCG	Monitorización continua de glucosa	Continuous Glucose Monitoring
MDI	Múltiples inyecciones diarias	Multiple Daily Injection
MLRSR	Razón de la Sensibilidad de la Regresión Lineal Modificada	Modified Linear Regression Sensitivity Ratio
Modelos OE	Modelos del error en la salida	Output Error models
NIH	Instituto Nacional de Salud en los EE. UU	National Institutes of Health

Acrónimos y abreviaturas

P-EGA	Rejilla de análisis de errores puntuales	Point-Error Grid Analysis
PC	Componente principal	Principal Component
PCA	Análisis de componentes principales	Principal Component Analysis
PCA-CBR	PCA en combinación con CBR	PCA in combination with CBR
PCA-CBR-2SR	PCA en combinación con CBR usando 2SR	PCA in combination with CBR using 2SR
PCA-CBR-3SR	PCA en combinación con CBR usando 3SR	PCA in combination with CBR using 3SR
PLS-DA	Análisis discriminante por mínimos cuadrados parciales	Partial Least Squares Discriminant Analysis
POLY	Función Kernel polinomial	Polynomial Function Kernel
R-EGA	Rejilla de análisis de errores en la tasa de cambio	Rate-Error Grid Analysis
RAD	Diferencia absoluta relativa	Relative Absolute Difference
RBF	Función Kernel de base radial Gaussiana	Radial Basis Function Kernel
RT	Tiempo real	Real-Time
RTCGMS	Sistemas de monitorización continua de glucosa en tiempo real	Continuous Glucose Monitoring Systems in REAL-Time
SC	Subcutáneo	Subcutaneous
SD	Desviación estándar	Standard Deviation
SIMCA	Modelado blando independiente de analogías entre clases	Soft Independent Modeling of Class Analogy
SMBG	Sistemas de auto-monitorización de glucosa en plasma	Self-Monitoring Blood Glucose
SMCG	Sistemas de monitorización continua de glucosa	Continuous Glucose Monitoring Systems
SOFA	Evaluación de Fallo Orgánico Secuencial	Sepsis-related Organ Failure Assessment
SVM-1	Estrategia de post-procesamiento de SVM para ajustar el sesgo en función de la G_{media}	
SVM-2	Estrategia de post-procesamiento de SVM para ajustar el sesgo en función de la <i>Sensibilidad</i>	
SVMs	Máquinas de soporte vectorial	Support Vector Machines
TC	Terapéuticamente correctas	Therapeutically correct
TI	Terapéuticamente incorrectas	Therapeutically incorrect
UCI	Unidad de cuidados intensivos	Intensive Care Unit

Variables

ABG	Concentración de BG arterial (mg/dL)
C	Parámetro de penalización o regularización
d	Grado del polinomio de la función Kernel polinomial
D -statistic	Estadístico T^2 o estadístico de Hotelling
d_Q	Q -distance
d_t	score-distance
G_{AR}	Concentración de BG estimada por el ABMAR-SMCG (mg/dL)
$G_{Proposed}$	Concentración de BG estimada por el ARP-SMCG (mg/dL)
$G_{Capillary}$	Concentración de BG capilar (mg/dL)
G_{Gold}	Concentración de BG estimada por el A-Gold TM (mg/dL)
G_{RTCGMS}	Concentración de BG estimada por el A-Guardian [®] RTCGMS (mg/dL)
HbA1 _c	Hemoglobina glicada A1 _c
<i>Insulin</i>	Dosis de CIII
<i>ISIG</i>	Corriente eléctrica medida por el sensor de glucosa (nA)
Matriz E	Matriz residual
Matriz Λ	Matriz de valores propios (diagonal)
Matriz P	Matriz de loadings
Matriz S	Matriz de covarianza
Matriz T	Matriz de scores
Mean RAD	Media de la diferencia absoluta relativa (%)
Median RAD	Mediana de la diferencia absoluta relativa (%)
<i>Offset</i>	Valor de <i>Offset</i>
Q -statistic	Estadístico Q o error de predicción cuadrático
<i>Sepsis</i>	Estado séptico de los pacientes
σ	Ancho de la función de base radial Gaussiana
<i>Temperature</i>	Temperatura corporal axilar
<i>Valid ISIG</i>	Corriente eléctrica medida por el sensor de glucosa adelantada 10 minutos (nA)
VBG	Concentración de BG venosa (mg/dL)

Índice de figuras

2.1. Producción de insulina en el páncreas humano	20
2.2. Proyecciones globales y regionales de la IDF sobre el número de personas con diabetes entre los 20-79 años, 2011 y 2030.	21
2.3. Panorama mundial de la diabetes por región en el 2011.	22
2.4. Bucle de control de glucosa en lazo cerrado empleando la ruta <i>s.c-s.c.</i>	27
2.5. Tecnologías de MCG desde el punto de vista de su capacidad invasiva.	30
2.6. El proceso de estimación de BG en los SMCGs mínimamente invasivos.	32
2.7. El CGMS [®] Gold [™]	36
2.8. El Guardian [®] RTCGMS	37
2.9. Transporte de glucosa plasma-ISF y procesos fisiológicos implicados.	38
2.10. Gráfico que ilustra los criterios ISO.	48
2.11. Rejilla Clarke EGA.	50
2.12. Rejilla Consensus EGA.	50
2.13. Rejilla CG-EGA.	51
3.1. Ejemplo de aplicar cuatro diferentes métodos de interpolación para muestrear las determinaciones de <i>VBG</i> cada cinco minutos.	67
3.2. Rejillas de Clarke EGA y Consensus EGA para evaluar el A-Gold [™] y el ARP-SMCG.	69
3.3. Ejemplos de perfiles glucémicos hospitalarios estimados por el A-Gold [™] y el ARP-SMCG: Estimación de episodios hipoglucémicos.	70
3.4. Ejemplos de perfiles glucémicos hospitalarios estimados por el A-Gold [™] y el ARP-SMCG: Estimación de episodios hiperglucémicos.	70
3.5. Ejemplos de perfiles glucémicos domiciliarios estimados por el A-Gold [™] y el ARP-SMCG.	71
3.6. Esquema de bloques general de un modelo AR.	74
3.7. Ciclo de identificación de un modelo AR.	76
3.8. Rejillas de Clarke EGA y Consensus EGA para evaluar el ABMAR-SMCG y el A-Gold [™]	87
3.9. Ejemplos de perfiles glucémicos hospitalarios estimados por el A-Gold [™] y el ABMAR-SMCG: Estimación de episodios hipoglucémicos.	88
3.10. Otros ejemplos de perfiles glucémicos hospitalarios estimados por el A-Gold [™] y el ABMAR-SMCG: Estimación de episodios hipoglucémicos.	89
3.11. Ejemplos de perfiles glucémicos hospitalarios estimados por el A-Gold [™] y el ABMAR-SMCG: Estimación de episodios hiperglucémicos y normoglucémicos.	89
3.12. Ejemplos de perfiles glucémicos hospitalarios estimados por el ABMAR-SMCG con y sin corregir las dos estimaciones de G_{AR} anteriores a la introducción de la calibración.	91

3.13. Ejemplos de perfiles glucémicos hospitalarios estimados por el A-Gold TM y el ABMAR-SMCG introduciendo solamente puntos de calibración antes de cada ingesta: Estimación de episodios hipoglucémicos.	92
3.14. Ejemplos de perfiles glucémicos hospitalarios estimados por el A-Gold TM y el ABMAR-SMCG introduciendo solamente puntos de calibración antes de cada ingesta: Estimación de episodios hiperglucémicos y normoglucémicos.	92
4.1. Ejemplo del perfil glucémico estimado por el A-Guardian [®] RTCGMS en un paciente crítico.	104
4.2. Gráficos de Bland-Altman para evaluar el A-Guardian [®] RTCGMS.	108
4.3. Rejillas de Clarke EGA para evaluar el A-Guardian [®] RTCGMS.	112
5.1. Distribución de los datos de acuerdo con Clarke EGA	121
5.2. Ejemplo de una porción del conjunto de datos utilizado para construir la metodología PCA-CBR	121
5.3. Representación gráfica de PCA	124
5.4. Valores propios	125
5.5. <i>Scores</i> , <i>D-statistic</i> y <i>Q-statistic</i> para los subconjuntos de entrenamiento y validación.	127
5.6. <i>D-statistic</i> versus <i>Q-statistic</i>	127
5.7. Ciclo de las 4-Rs	128
5.8. Descriptores que constituyen una estructura de caso	129
5.9. Distribución de los casos en BC y conjunto de validación	130
5.10. Diagrama metodología PCA-CBR	135
5.11. Resultados comparativos de PCA-CBR-2SR, SIMCA y PLS-DA (Enfoque 1).	139
5.12. Resultados comparativos de PCA-CBR-3SR, SIMCA y PLS-DA (Enfoque 2).	140
5.13. Clasificación de los pares de datos ABG/G_{RTCGMS} de dos pacientes representativos al comparar la metodología PCA-CBR con la clasificación de acuerdo a Clarke EGA.	143
5.14. Perfil de un paciente médico obtenido al aplicar la metodología PCA-CBR-3SR usando la BC No. 9.	144
5.15. Perfil de un paciente post-quirúrgico obtenido al aplicar la metodología PCA-CBR-2SR usando la BC No. 1.	145
6.1. Distribución de los datos de acuerdo con los criterios ISO	151
6.2. Representación gráfica de un ejemplo de las propiedades de las métricas <i>sensibilidad</i> y <i>especificidad</i>	158
6.3. Resultados de la clasificación de tres pacientes representativos de acuerdo con el estado séptico, utilizando SVM-1 y SVM-2.	168
6.4. Clasificación de los pares de datos ABG/G_{RTCGMS} de tres pacientes representativos contrastando los resultados obtenidos al aplicar la metodología SVM-1 con los criterios ISO.	170
6.5. Perfil de un paciente representativo en shock séptico obtenido al aplicar la metodología basada en SVM-1.	172
6.6. Perfil de un paciente representativo con sepsis obtenido al aplicar la metodología basada en SVM-1.	173
6.7. Perfil de un paciente representativo no séptico obtenido al aplicar la metodología basada en SVM-1.	174

Índice de tablas

2.1.	Complicaciones de la DM a largo plazo.	24
2.2.	Comparación de diferentes dispositivos para MCG desde el punto de vista de su capacidad invasiva.	31
2.3.	Características de los SMCGs mínimamente invasivos más utilizados en la literatura.	34
2.4.	Matriz de confusión para un problema de clasificación binario.	52
3.1.	Características basales y condiciones primarias de los pacientes con DM-1 del Hospital Josep Trueta utilizados en el desarrollo del ARP-SMCG.	57
3.2.	Resultados de evaluar la precisión analítica del A-Gold TM y del ARP-SMCG de acuerdo con el estado glucémico.	65
3.3.	Resultados de evaluar la precisión analítica del A-Gold TM y del ARP-SMCG de acuerdo con la tasa de cambio de la glucemia.	66
3.4.	Resultados de evaluar la precisión clínica del A-Gold TM y del ARP-SMCG.	68
3.5.	Resultados de algunos criterios de precisión analítica del A-Gold TM reportados en la literatura.	72
3.6.	Características basales y condiciones primarias de los pacientes con DM-1 del Hospital Josep Trueta utilizados en el desarrollo del ABMAR-SMCG.	77
3.7.	Resultados de la mediana de la métrica <i>Best Fit</i> obtenidos aplicando una estrategia de validación cruzada dejando dos fuera.	79
3.8.	Resultados de la mediana de la métrica <i>Best Fit</i> en cada uno de los pacientes con DM-1, obtenidos con una estructura BJ con polinomios de tercer orden.	81
3.9.	Parámetros del modelo poblacional BJ de tercer orden.	82
3.10.	Resultados de evaluar la precisión analítica del A-Gold TM y del ABMAR-SMCG.	85
3.11.	Resultados de evaluar la precisión clínica del A-Gold TM y del ABMAR-SMCG.	85
3.12.	Resultados de algunos criterios de precisión analítica para algoritmos de calibración en RT reportados en la literatura.	93
4.1.	Características basales y condiciones primarias de los pacientes de la UCI del Hospital Josep Trueta incluidos en el estudio clínico.	101
4.2.	Rangos glucémicos de la población mixta y de cada una de las cohortes de pacientes de acuerdo con su estado séptico.	107
4.3.	Resultados de evaluar la precisión analítica del A-Guardian [®] RTCGMS de acuerdo con su estado séptico.	109
4.4.	Resultados de evaluar la precisión analítica del A-Guardian [®] RTCGMS de acuerdo con otros factores clínicos.	110
4.5.	Resultados de evaluar la precisión clínica del A-Guardian [®] RTCGMS.	111
5.1.	Características basales y condiciones primarias de los pacientes de UCI utilizados en el desarrollo de la metodología PCA-CBR.	119

5.2. Rangos glucémicos de la población mixta y de cada una de las cohortes de pacientes utilizados.	120
5.3. Varianza capturada por el modelo PCA.	126
5.4. Porcentaje de explicación de cada variable para cada PC.	126
5.5. BCs y sus correspondientes conjuntos de validación.	130
5.6. Resultados de las métricas <i>proporción de aciertos, especificidad, sensibilidad</i> y G_{media} obtenidos al evaluar la metodología PCA-CBR-2SR y PCA-CBR-3SR. . .	133
5.7. Resultados de las métricas <i>proporción de aciertos, especificidad, sensibilidad</i> y G_{media} obtenidos al evaluar la metodología PCA-CBR-2SR y PCA-CBR-3SR después de remover los datos que no son fiables.	134
5.8. Resultados de aplicar SIMCA.	137
5.9. Resultados de aplicar PLS-DA.	138
5.10. Resultados de evaluar la metodología PCA-CBR en dos pacientes representativos.	141
6.1. Resultados promedios de las métricas <i>proporción de aciertos, especificidad, sensibilidad</i> y G_{media} para la validación cruzada de cinco capas.	162
6.2. Resultados de la evaluación de la metodología basada en SVM-1 de acuerdo con el estado séptico de los pacientes.	164
6.3. Resultados de la evaluación de la metodología basada en SVM-2 de acuerdo con el estado séptico de los pacientes.	166



Universitat de Girona

Enfoques metodológicos para mejorar la precisión y fiabilidad de los sistemas de monitorización continua de glucosa en pacientes con diabetes tipo 1 y en pacientes críticos mediante técnicas lineales y no lineales

Programa Doctoral en Tecnología.

Tesis presentada para optar al título de Doctora por la Universitat de Girona.

Firma

Fecha

Yenny Teresa Leal Moncada
Autor

Dr. Josep Vehí
Director

Dra. Magda L. Ruíz O.
Co-directora

Girona, España
2013

A Iván, mi ángel, quien durante todos estos años me ha acompañado en mis sueños. Su presencia espiritual, su amor y su recuerdo guían cada uno de mis pasos y me animan a continuar en los momentos adversos.

A mis padres Jairo y María del Carmen, mi luz, mi fortaleza, mi vida.

Agradecimientos

Esta tesis doctoral es el resultado no sólo de mi esfuerzo y mi trabajo, sino también del apoyo de muchas personas que directa o indirectamente han colaborado conmigo. Sin embargo, en poco tiempo y después de varios años, es muy difícil recordarlos a todos ellos. Por lo tanto, pido disculpas de antemano por no incluir en estas líneas a personas que también se lo merecen y de quienes por la prisa me he olvidado.

Me gustaría expresar mi agradecimiento a mi director, el Prof. Josep Vehí, por su apoyo académico y financiero y su orientación a través de este largo proceso. De igual forma, quiero dar las gracias a mi co-directora, la Dra. Magda Ruíz, por su soporte en conocimiento y experiencia para lograr mis objetivos de investigación.

Quiero agradecer a la Dra. Carol Lorencio y al Prof. Luis González. A Carol por creer en mí, por enseñarme valiosos conocimientos médicos, por la confianza, dedicación, compromiso, amistad y ánimo en todas las etapas de nuestro trabajo. A Luis por el interés y compromiso que mostró hacia mi trabajo y por sus valiosos aportes académicos y personales.

Además, quiero expresar mis agradecimientos:

A todos los miembros del grupo de investigación MICELab por su amistad y soporte durante mi estancia allí, en particular, a la Dra. Remei Calm, quién me brindó sus consejos y su soporte personal en momentos claves.

A la Universidad de Girona por el apoyo económico durante todo este periodo a través de la Beca de Investigación “Beca de Recerca UdG BR08/20”. Este trabajo también fue financiado parcialmente por el Ministerio de Ciencia e Innovación Español a través de los proyectos coordinados de investigación DPI-2007-66728 y DPI-2010-20764-C02, y por la Generalitat de Cataluña a través de la subvención 2009 SGR 523.

A los proyectos Insulaid2 y CLOSEDLOOP4MEAL y a su coordinador el Dr. Jorge Bondia.

Al Dr. Josep María Sirvent y al Dr. Alfonso Bonet por darme la oportunidad de desarrollar esta investigación en la UCI del Hospital Josep Trueta, así como al personal médico y de enfermería, por su trabajo y por hacerme sentir un miembro más de su equipo.

Al equipo médico y de enfermería de la Unidad de Endocrinología del Hospital Josep Trueta por llevar a cabo el estudio clínico en los pacientes con DM-1.

A los pacientes que voluntariamente participaron en los estudios clínicos que soportaron este trabajo.

Al Dr. César Palerm y al Dr. Luis Eduardo Mujica por sus comentarios y sugerencias sobre mi investigación.

A Medtronic, Inc., por proporcionar algunos de los dispositivos utilizados en este estudio.

Mi gratitud por sus valiosos y certeros aportes durante todos estos años para Xavier Berjaga, de quién he aprendido muchas cosas que serán útiles para mi vida profesional. Xavi gracias por la paciencia, por escuchar mis ideas y proporcionarme numerosas pistas que me ayudaron a conseguir logros exitosos en mi investigación, por ser mi asesor, pero también mi amigo.

El tiempo que he pasado en Girona en el transcurso de mis estudios de máster y doctorado ha sido sin duda una de las mayores experiencias de mi vida. Quiero expresar mi gratitud a mis amigos de Girona, a quienes han estado a mi lado en momentos decisivos y con quienes he compartido momentos increíbles. Sólo tengo palabras de agradecimiento para ellos, especialmente para Nico, Jeimy Vélez, Sam, Sebas, Luca, Eduard, Josep, Lien y para

la Familia Mujica Ruíz por recibirme en su familia. En especial, quiero agradecer a mi familia en Girona: Maira, Fabián, Joan, Silvia y Mauricio. Muchas gracias Maira por la invaluable amistad, el apoyo personal y académico, y el afecto que me has brindado durante todos estos años. A Fabián gracias por la amistad, la confianza y el ánimo en momentos claves. A Joan gracias por llegar para alegrarnos la vida. A Silvia por ser tan tolerante conmigo, por las confidencias, el cariño, el estímulo en los momentos difíciles. Pero si hay alguien que me conoce íntegramente y me ha soportado de todas las formas posibles durante estos años ha sido Mauricio. No tengo palabras para agradecer todo el apoyo que me has brindado. Has estado desde el primer día en los buenos momentos, pero sobre todo has sido incondicional cuando las cosas se ponían tan difíciles. Gracias por ser mi compañero de viajes, de fines de semana, de aventuras, de tristezas y alegrías, por ser mi amigo, mi familia y mi mayor apoyo en esta etapa.

Yo he tenido la suerte de tener calidad de amigos a lo largo de mi vida, testigos de mis triunfos y soportes esenciales en las situaciones adversas, a quienes considero miembros de mi familia. A ellos quiero agradecer porque a pesar de la distancia, me acompañaron en este largo proceso, nunca se olvidaron de mí y estuvieron firmes animándome a continuar y a finalizar esta etapa. Mil gracias Lina, Yaz, Sandra, Anyuri, Erika, Rafa, Javier y Darwin por su cariño sincero y su apoyo incondicional.

Mi gratitud eterna para los miembros de mi familia por desear siempre lo mejor para mí y apoyarme en todos los sentidos durante mi vida. Sé que este logro los llena de sincera alegría. A mis familias adoptivas: Lizcano Villamizar, Cárdenas Quintero, Quintero Duarte y Lizcano Corredor. A mi hermana Carola por compartir mis alegrías, tristezas y mis ganas de vivir. A mi madrina Rita por respaldar desde que era niña mis deseos de aprender y formarme profesional y personalmente. A mis adoradas tías Yolanda, Blanca, Petra y Lidia, a Magaly, Angélica, Lisa, Josep, Pablo, Joel y a mi madrina Carmen por preocuparse por mi bienestar durante todos estos años. Y por supuesto, gracias a mis consentidos, mis sobrinos amados, Gabriela, Andrés, David, Joel, Mariana y Jacob, quienes con su inocencia de niños, sus cartas y mensajes me robaron sonrisas e ilusiones, aún durante la fase más difícil de mi investigación.

Estos últimos párrafos los dedico a los cuatro grandes amores de mi vida, sin cuyo esfuerzo constante no habría sido posible alcanzar mis metas. A Jhon Jairo, gracias por llegar para quedarse y por cuidar y dar amor a papá y mamá durante estos largos años en que he estado ausente físicamente. Mi amor y gratitud infinitos para mis padres María del Carmen y Jairo, quienes me han brindado su amor sincero e incondicional, han creído en mí y en mis capacidades, me han enseñado a soñar, a ser persistente, a luchar; han hecho numerosos sacrificios para formarme como profesional, pero ante todo como ser humano, me han dado un gran ejemplo de fortaleza para seguir adelante en los momentos más adversos de la vida y me han impulsado a venir aquí, a esforzarme y trabajar para alcanzar este objetivo, aunque ello implicara separarnos físicamente. Este triunfo papí y mami está dedicado a ustedes y es tan suyo como mío.

Finalmente, y aunque sé que nunca podrá leer estas líneas, quiero rendir un homenaje a quién me ha brindado la fuerza y la motivación para no dejar este proceso inconcluso, a quién dedico cada uno de los días de mi vida, a mi hermano Iván. Es por amor a su recuerdo, a los sueños que siempre quisimos compartir juntos, que sigo adelante. Él es el ángel que me acompaña en mis sueños, que me cuida y guía cada uno de mis pasos. En mi corazón siempre hay un lugar para ti, durante toda mi vida guardaré una parte de ti conmigo y donde quiera que yo esté, tú estarás ahí.

Gracias a todos!

*Yenny T. Leal Moncada
Girona, España
Julio 2013*

Índice general

Índice de figuras	XIII
Índice de tablas	XV
1. Introducción	7
1.1. Motivación	7
1.2. Problemas abiertos y desafíos	10
1.3. Metodología propuesta	11
1.4. Objetivos	13
1.5. Estructura de la tesis	14
2. Sistemas de Monitorización Continua de Glucosa (SMCGs)	19
2.1. Visión de conjunto sobre la DM	19
2.1.1. Cifras actuales y previsiones futuras	21
2.1.2. Tipos de DM	22
2.1.3. Complicaciones por DM	23
2.1.4. Tratamiento para la DM-1	25
2.1.5. Estado de la investigación actual	26
2.2. Monitorización Continua de la Glucosa (MCG)	28
2.2.1. Tecnología actual para MCG	30
2.2.2. SMCGs mínimamente invasivos	32
2.3. Técnicas propuestas en la literatura para abordar los problemas abiertos en la tecnología de MCG mínimamente invasiva	37
2.3.1. Dinámica entre BG e IG	38
2.3.2. Calibración y retardo del sensor	39
2.3.3. Eliminación de ruido	43
2.3.4. Predicción	43
2.3.5. Fiabilidad	44
2.4. Criterios para evaluar la precisión y la fiabilidad de los SMCGs	46
2.4.1. Criterios para evaluar la precisión de los SMCGs	46
2.4.2. Métricas para evaluar la fiabilidad de los SMCGs	51
2.5. Limitaciones para avanzar en esta área de investigación	53
2.6. Resumen	54
3. Algoritmos de estimación de glucosa en plasma para SMCGs	55
3.1. Estudio observacional y recolección de datos	56
3.1.1. Descripción de los pacientes y del entorno clínico del ensayo	56
3.1.2. MCG en la Unidad de Endocrinología del Hospital Josep Trueta: protocolo de actuación MODAS	57

3.2.	Algoritmo para estimación de BG usando técnicas de regresión lineal: Algoritmo retrospectivo propuesto para el SMCG (ARP-SMCG)	59
3.2.1.	Descripción del A-Gold TM	59
3.2.2.	ARP-SMCG	61
3.2.3.	Evaluación de la precisión del ARP-SMCG usando criterios analíticos	63
3.2.4.	Evaluación de la precisión del ARP-SMCG usando criterios clínicos	66
3.2.5.	Análisis de la capacidad de estimación de hipoglucemias e hiperglucemias del ARP-SMCG	68
3.2.6.	Análisis de los resultados obtenidos	71
3.3.	Algoritmo para estimación de BG usando modelos ARs: Algoritmo basado en modelos auto-regresivos para el SMCG (ABMAR-SMCG)	72
3.3.1.	Conceptos básicos de modelos ARs	73
3.3.2.	Pre-procesamiento de los datos	75
3.3.3.	Entrenamiento del modelo AR: Identificación de la estructura del modelo	78
3.3.4.	ABMAR-SMCG	81
3.3.5.	Evaluación del ABMAR-SMCG usando criterios analíticos	83
3.3.6.	Evaluación del ABMAR-SMCG usando criterios clínicos	84
3.3.7.	Análisis de la capacidad de estimación de hipoglucemias e hiperglucemias del ABMAR-SMCG	86
3.3.8.	Análisis de los resultados obtenidos	90
3.4.	Resumen	93
4.	Evaluación de la precisión de los SMCGs en una Unidad de Cuidados Intensivos (UCI) según el estado séptico de los pacientes	95
4.1.	Generalidades	97
4.2.	Descripción de los pacientes y del entorno clínico del ensayo	99
4.3.	MCG en la UCI del Hospital Josep Trueta: protocolo de actuación	101
4.4.	Resultados	105
4.4.1.	Evaluación de la precisión del A-Guardian [®] RTCGMS en pacientes críticos utilizando criterios analíticos	105
4.4.2.	Evaluación de la precisión del A-Guardian [®] RTCGMS en pacientes críticos utilizando criterios clínicos	110
4.5.	Discusión	111
4.5.1.	Limitaciones del estudio	113
4.6.	Resumen	115
5.	Aplicación de Análisis de Componentes Principales (PCA) y Razonamiento Basado en Casos (CBR) para detectar medidas terapéuticamente correctas (TC) e incorrectas (TI) en SMCGs	117
5.1.	Descripción de los datos	118
5.2.	Etiquetado y pre-procesamiento de los datos	120
5.3.	Extracción de características usando PCA	122
5.3.1.	Conceptos básicos de PCA	122
5.3.2.	Construcción del modelo PCA	124
5.4.	Detección de medidas TC y TI usando CBR	128
5.4.1.	Conceptos básicos de CBR	128
5.4.2.	Definición de los descriptores	129
5.4.3.	Construcción de las BCs y de los conjuntos de validación	129
5.4.4.	Aplicación del CBR	130

5.5. Puesta a punto de la metodología PCA-CBR	132
5.6. Evaluación comparativa de la metodología PCA-CBR con SIMCA y PLS-DA . . .	136
5.7. Evaluación de la metodología PCA-CBR en dos pacientes representativos	138
5.8. Análisis de los resultados obtenidos	146
5.9. Resumen	148
6. Detección de medidas correctas e incorrectas en los SMCGs aplicando estrategias de post-procesamiento de Máquinas de Soporte Vectorial (SVMs)	149
6.1. Descripción, etiquetado y pre-procesamiento de los datos	150
6.2. Máquinas de Soporte Vectorial (SVMs)	151
6.2.1. Conceptos básicos de SVMs	151
6.2.2. SVM-1: Aplicación de una estrategia de post-procesamiento para ajustar el sesgo en función de la G_{media}	156
6.2.3. SVM-2: Aplicación de una estrategia de post-procesamiento para ajustar el sesgo en función de la <i>Sensibilidad</i>	159
6.3. Detección de medidas correctas e incorrectas en los SMCGs aplicando nuevas estrategias de post-procesamiento de SVM	160
6.4. Resultados	162
6.4.1. Evaluación comparativa de los enfoques de SVM abordados	162
6.4.2. Evaluación de la metodología propuesta basada en SVM-1	163
6.4.3. Evaluación de la metodología propuesta basada en SVM-2	165
6.4.4. Ejemplos de la metodología propuesta para detectar medidas correctas e incorrectas realizadas por los SMCGs aplicando las estrategias de post-procesamiento de SVMs	167
6.5. Análisis de los resultados obtenidos	175
6.6. Resumen	176
7. Conclusiones y Trabajo Futuro	179
7.1. Contribuciones	179
7.2. Trabajo futuro	181
Referencias	185
A. Conceptos básicos de SIMCA y PLS-DA	203

RESUMEN

En los últimos años, ha habido un progreso tecnológico sin precedentes en el desarrollo de los sistemas de monitorización continua de glucosa (SMCGs), los cuáles constituyen un campo de trabajo de gran importancia dentro de la línea de investigación tecnológica que busca mejorar la calidad de vida de las personas diabéticas. Estos dispositivos son potencialmente una herramienta eficaz para mejorar el control glucémico, tanto en los pacientes diabéticos como en los pacientes ingresados en las unidades de cuidados intensivos (UCIs). Su principal característica es estimar concentraciones de glucosa en plasma (BG) de manera continua en el tiempo, permitiendo con ello identificar las tendencias de la glucosa a lo largo del día, y facilitar la detección de episodios glucémicos críticos (hipoglucemias e hiperglucemias).

Actualmente, desde el punto de vista de su capacidad invasiva, existen múltiples dispositivos para la monitorización continua de glucosa. Sin embargo, los únicos dispositivos que se encuentran aprobados por las agencias reguladoras para uso comercial son los SMCGs mínimamente invasivos. Estos sistemas ofrecen una estimación de los niveles de BG a partir de mediciones en el fluido intersticial (ISF), un compartimento alterno al plasma, donde se inserta el sensor de glucosa subcutáneo. Por tanto, los niveles de BG se estiman de manera indirecta desde la glucosa intersticial (IG).

El SMCG, además de constituir una herramienta complementaria a las mediciones de glucosa en sangre capilar para ayudar al personal médico en la toma de decisiones terapéuticas, también es un elemento clave en el desarrollo del páncreas artificial (AP); un sistema de control automático de glucosa en lazo cerrado, en el cual se integra un SMCG, una bomba de infusión de insulina y un algoritmo de control. Sin embargo, existen algunos aspectos claves de estos dispositivos que se deben abordar para mejorar su eficiencia. Lo más crítico de todos es que sus estimaciones de glucosa resulten imprecisas y poco fiables, particularmente en el rango hipoglucémico. Esto se debe, parcialmente, a la baja calidad de los algoritmos utilizados para estimar BG, y al no disponer de capacidades de auto-monitorización, las cuáles les permitirían detectar sus propios fallos. Esta carencia de precisión y fiabilidad es un factor limitante en su aprobación como una herramienta independiente en la toma de decisiones terapéuticas, y en el desarrollo del AP. Por tanto, contribuir a dar solución a este problema es un reto de gran interés e importancia desde el punto de vista clínico.

Esta tesis doctoral se enfoca en el desarrollo y mejora de los SMCGs mínimamente invasivos, haciendo énfasis en resolver el problema de la precisión y fiabilidad de las concentraciones de BG estimadas por los SMCGs. En particular, el objetivo de esta investigación es contribuir con nuevas estrategias metodológicas que incorporen estos problemas en su formulación. En el contexto de la precisión, se proponen dos nuevos algoritmos para estimación de BG: 1) un algoritmo retrospectivo, basado en técnicas de regresión lineal, el cuál es capaz de mejorar el ajuste de las estimaciones de BG en el rango hipoglucémico; y 2) un algoritmo que considera la relación dinámica entre BG e IG, basando su principio de funcionamiento en modelos auto-regresivos, y puede implementarse en tiempo real. Con respecto a la fiabilidad, se exploran dos técnicas de aprendizaje y reconocimiento de patrones: 1) una combinación de Análisis de Componentes Principales (PCA) y Razonamiento Basado en Casos (CBR), y 2) Máquinas de Soporte Vectorial (SVMs) no convencionales adaptadas para trabajar con conjuntos de datos desbalanceados (mediante dos estrategias de post-procesamiento). Estos planteamientos utilizan la información suministrada por el propio SMCG y variables relacionadas con la condición clínica del paciente, para detectar estimaciones de BG correctas e incorrectas, sin requerir de alguna medida de refe-

rencia. Con el fin de evaluar el rendimiento de las metodologías propuestas, se han definido dos nuevos esquemas binarios de clasificación, tomando como base los criterios de precisión clínicos y analíticos aprobados por las agencias reguladoras.

El rendimiento de los enfoques metodológicos propuestos se ha validado con datos de diferentes estudios clínicos, tanto de pacientes diabéticos como de pacientes críticos.

RESUM

En els darrers anys, hi ha hagut un progrés tecnològic sense precedents en el desenvolupament dels sistemes de monitoratge continu de glucosa (SMCGs), els quals constitueixen un camp de treball de gran importància dins de la línia de recerca tecnològica que busca millorar la qualitat de vida de les persones diabètiques. Aquests dispositius són potencialment una eina eficaç per millorar el control glucèmic, tant en els malalts diabètics com en els malalts ingressats en unitats de cures intensives (UCIs). La seva principal característica és estimar concentracions de glucosa en plasma (BG) de manera contínua en el temps, permetent així identificar les tendències de la glucosa al llarg del dia, i facilitar la detecció d'episodis glucèmics crítics (hipoglucèmies i hiperglucèmies).

Actualment, des del punt de vista de la seva capacitat invasiva, existeixen múltiples dispositius per al monitoratge continu de glucosa. No obstant això, els únics dispositius que es troben aprovats per les agències reguladores per a l'ús comercial són els SMCGs mínimament invasius. Aquests sistemes ofereixen una estimació dels nivells de BG a partir de mesures al fluid intersticial (ISF), un compartiment altern al plasma, on s'insereix el sensor subcutani de glucosa. Per tant, els nivells de BG s'estimen de manera indirecta a partir de la glucosa intersticial (IG).

L'SMCG, a més de constituir una eina complementària a les mesures de glucosa en sang capil·lar per ajudar el personal mèdic en la presa de decisions terapèutiques, també és un element clau en el desenvolupament del pàncrees artificial (AP); un sistema de control automàtic de glucosa en llaç tancat, en el qual s'integra un SMCG, una bomba d'infusió d'insulina i un algorisme de control. No obstant això, existeixen alguns aspectes claus d'aquests dispositius que s'han d'abordar per millorar-ne la seva eficiència. El més crític de tots ells és que les seves estimacions de glucosa resulten imprecises i poc fiables, particularment al rang hipoglucèmic. Això es deu, parcialment, a la baixa qualitat dels algorismes utilitzats per estimar la BG, i perquè no disposa de capacitats d'auto-monitoratge, les quals els hi permetrien detectar les seves pròpies fallades. Aquesta manca de precisió i fiabilitat és un factor limitant en la seva aprovació com a eina independent en la presa de decisions terapèutiques i en el desenvolupament de l'AP. Per tant, contribuir a la solució d'aquests problemes és un repte de gran interès i importància des del punt de vista clínic.

Aquesta tesi doctoral s'enfoca en el desenvolupament i millora dels SMCGs mínimament invasius, fent èmfasi en la resolució del problema de la precisió i fiabilitat de les concentracions de BG estimats pels SMCGs. En particular, l'objectiu d'aquesta investigació es contribuir amb noves estratègies metodològiques que incorporin aquests problemes en la seva formulació. En el context de la precisió, es proposen dos nous algorismes d'estimació de BG: 1) un algorisme retrospectiu, basat en tècniques de regressió lineal, que millora la capacitat d'ajustar les estimacions de BG en el rang hipoglucèmic; i 2) un algorisme que té en compte la relació dinàmica entre BG i IG basada en models auto-regressius, i que és implementable en temps real. En relació a la fiabilitat, s'exploren dues metodologies basades en tècniques d'aprenentatge i reconeixement de patrons: 1) una combinació d'Anàlisi de Components Principals (PCA) i Raonament Basat en Casos (CBR); i 2) Màquines de Suport Vectorial (SVMs) adaptades per treballar amb conjunts de dades no balancejats (mitjançant dues estratègies de post-processament). Aquests plantejaments utilitzen la informació subministrada pel propi SMCG i variables relacionades amb la condició clínica del pacient, per detectar estimacions de BG correctes i incorrectes, sense requerir de cap mesura de referència. Per tal d'avaluar el rendiment de les metodologies proposades, s'han definit dos nous esquemes binaris de classificació, prenent com a base els criteris de

precisió clínica i analítics aprovats per les agències reguladores.

El rendiment dels mètodes proposats s'ha validat amb dades de diferents estudis clínics, tant de malalts diabètics com de malalts crítics.

ABSTRACT

In the past years, there has been an unprecedented technological progress in the development of continuous glucose monitoring systems (CGMSs), which constitute a field of work of great importance inside of the technological research line that seeks to improve the quality of life of diabetic persons. These devices are potentially an effective tool to improve glycemic control in both diabetic patients and patients admitted to intensive care units (ICUs). Its main characteristic is to estimate blood glucose (BG) concentrations continuously along time, which allows to identify the glucose trends throughout the day, and to facilitate critical glycemic episodes detection (hypo- and hyperglycemia).

Nowadays, from the viewpoint of their degree of invasiveness, there are many devices for continuous glucose monitoring. However, the only devices that have been approved by the regulatory agencies for commercial use are the minimally invasive CGMSs. These systems provide an estimation of BG levels from measurements in the interstitial fluid (ISF), an alternative compartment to plasma, where the subcutaneous glucose sensor is inserted. Therefore, BG levels are estimated indirectly from the interstitial glucose (IG).

The CGMS, besides being a complementary tool to the measurements of capillary BG to assist to medical staff in order to make clinical decisions, it's also a key element in the development of the artificial pancreas (AP); a system of automatic close-loop control of glucose, in which are integrated a CGMS, an insulin pump and a control algorithm. However, there are some key aspects of these devices to address in order to improve their efficiency. The most critical of them is that the glucose estimations are inaccurate and unreliable, particularly in the hypoglycemic range. This is partially due to the low quality of the algorithms used to estimate the BG, and their lack self-monitoring capabilities, which would allow them to detect their own faults. This lack accuracy and reliability is a limiting factor in its approval as a standalone tool, in order to make clinical decisions, and in the development of the AP. Therefore, contributions to solve this problem are a challenge of great interest and importance from a clinical point of view.

This dissertation focuses on the development and improvement of the minimally invasive CGMSs, emphasizing in the problem of the accuracy and reliability of the BG concentrations estimated by CGMSs. In particular, this research is intended to contribute with new methodological strategies that incorporate these issues in its formulation. Regarding accuracy, two new algorithms for BG estimation are proposed: 1) a retrospective algorithm, based on linear regression techniques, which is able to improve the fit of BG estimations in the hypoglycemic range; and 2) an algorithm that takes into account the dynamic relationship between BG and IG, and based on autoregressive models that can be implemented in real time. With respect to reliability, two pattern recognition and learning techniques are explored: 1) a combination of Principal Component Analysis (PCA) and Case-Based Reasoning (CBR), and 2) news Support Vector Machines (SVMs) able to work with imbalanced datasets (thanks to two post-processing strategies). These approaches combine the information provided by the monitor itself and variables related to the patient's clinical condition, to detect correct and incorrect measurements, without requiring any reference measurement. In order to evaluate the proposed methodologies, two new binary classification schemes have been defined, based on the clinical and analytical accuracy criteria approved by the regulatory agencies.

The performance of the proposed methodological approaches has been validated with data from different clinical studies, in both diabetic and critically-ill patients.

Capítulo 1

Introducción

Este capítulo introductorio presenta una visión general de la tesis, a partir de las consideraciones generales que motivaron esta investigación. Se explican brevemente los antecedentes, los principales problemas y desafíos, la metodología y los objetivos de este estudio. Por último, se presenta una descripción de la estructura y el contenido de la tesis.

1.1. Motivación

La *glucosa* es la mayor fuente de energía o el “*combustible*” más importante para el cuerpo humano y su concentración en la sangre se controla de manera precisa por la insulina, una hormona reguladora producida por las células β de los Islotes de Langerhans en el páncreas, que actúa como una llave permitiendo la difusión de la glucosa a las células a través de la membrana celular.

La diabetes mellitus (DM) es una enfermedad metabólica cuya característica primordial es la presencia de altos niveles de glucosa en sangre (hiperglucemia) (Williams and Pickup, 2004). En los pacientes diabéticos, el páncreas no es capaz de secretar cantidades suficientes de insulina (diabetes mellitus tipo 1 o DM-1), lo que conlleva a que requieran de infusión de insulina exógena para sobrevivir; o se presenta una resistencia a la insulina por parte del cuerpo, derivando en una eventual pérdida en la liberación de insulina desde el páncreas (diabetes mellitus tipo 2 o DM-2).

Existen múltiples complicaciones microvasculares y macrovasculares asociadas con la DM. Las complicaciones microvasculares incluyen la nefropatía, que conduce a insuficiencia renal, la retinopatía, que conduce a la ceguera y la neuropatía, que conduce a úlceras o amputación del pie. Dentro del grupo de complicaciones macrovasculares se encuentran el infarto cardíaco y los accidentes cerebrovasculares, los cuales tienen asociada una alta mortalidad. En la década de 1990, el Estudio sobre Control de la Diabetes y sus Complicaciones (-del inglés *Diabetes Control and Complications Trial (DCCT)*-) (DCCT, 1993) mostró que cualquier mejora en el control de la glucosa, determinado este control por el nivel de hemoglobina glicada $A1_c$ ($HbA1_c$)¹, reduce el riesgo de sufrir complicaciones crónicas microvasculares asociadas con la DM-1. Por tanto, dentro del proceso de prevención de las complicaciones mencionadas anteriormente, es esencial que los pacientes con DM-1 establezcan como objetivo de control mantener los niveles de glucosa

¹La hemoglobina glicada es una heteroproteína de la sangre cuya medición es una prueba de laboratorio muy utilizada en la diabetes para saber si el control que realiza el paciente sobre la enfermedad ha sido bueno durante las últimas seis a doce semanas. De hecho el 50,0% del resultado depende sólo de entre las cuatro y seis últimas semanas. Una $HbA1_c$ por debajo de siete, habla de un buen control glucémico, y a largo plazo, se traduce en una reducción de las complicaciones en pacientes diabéticos (MED, 2013a).

en el rango normoglucémico (70 mg/dL – 180 mg/dL).

Por otra parte, la hiperglucemia es un fenómeno muy frecuente no sólo en los pacientes diabéticos, sino también en todo tipo de pacientes críticos. Clásicamente se ha interpretado la hiperglucemia como una manifestación más de gravedad, provocada por el estímulo de hormonas de estrés y contrareguladoras, así como por la resistencia a la insulina que típicamente presentan estos enfermos críticos, aún sin historia previa de diabetes. Actualmente, se sabe que la hiperglucemia, además de representar un marcador de gravedad, genera efectos nocivos en el organismo, y por tanto, condiciona un peor pronóstico en estos pacientes y una mayor morbi-mortalidad. Los efectos deletéreos de la hiperglucemia en los pacientes críticos pueden ser explicados por el importante efecto pro-inflamatorio que posee la glucosa a nivel molecular. Aunque todavía es objeto de debate entre los investigadores cuál debe ser el rango objetivo de control glucémico, lo que no es una cuestión de debate es la necesidad de tratar la glucemia en los pacientes críticos.

Para cubrir las deficiencias de insulina en el cuerpo, los pacientes diabéticos y los pacientes críticos son tratados mediante terapia con insulina o *insulinoterapia*. La terapia con insulina es vital para los pacientes afectados por DM-1, esencial para el control de glucosa en sangre y la prevención de complicaciones en muchos pacientes con DM-2, y de gran valor para los pacientes no diabéticos en situaciones críticas. Sin embargo, la terapia intensiva de insulina requerida para lograr los objetivos de control de glucosa está limitada por un aumento en el riesgo de sufrir episodios de hipoglucemias severas (niveles de glucosa demasiado bajos), con todas sus consecuencias.

En el caso de los pacientes con DM-1, que son quienes requieren administración de insulina exógena para reemplazar su secreción fisiológica, la terapia de insulina se administra mediante múltiples inyecciones diarias o por medio de la infusión continua de insulina subcutánea (utilizando una bomba externa o el tratamiento convencional). Generalmente, antes de cada comida, para ajustar la dosis de insulina en relación a la ingesta de carbohidratos planeada, los pacientes miden su concentración de glucosa en plasma (-del inglés *Blood Glucose (BG)*-) por medio de glucómetros convencionales, también conocidos como sistemas de auto-monitorización de glucosa en plasma (-del inglés *Self-Monitoring Blood Glucose (SMBG)*-). Estos dispositivos miden los niveles de glucemia de forma puntual, mediante la extracción de una pequeña muestra sanguínea capilar (tres a cuatro mediciones por día).

En el caso de los pacientes críticos, en la mayoría de las unidades de cuidados intensivos (UCIs), la determinación y control habitual de la glucemia frecuentemente se hace usando protocolos basados en la experiencia médica, los cuáles determinan la glucemia de manera intermitente mediante punción y extracción de sangre capilar, venosa o arterial, utilizando diferentes SMBGs convencionales. Seguidamente, se utiliza el valor medido de glucosa y de acuerdo con el protocolo de control establecido por el personal sanitario de dicha UCI, se administra la insulina pertinente en función del valor de glucosa obtenido (si existe hiperglucemia y esta es o no moderada, se administra insulina rápida endovenosa en perfusión continua o insulina rápida subcutánea cada seis horas).

Sin embargo, las escasas muestras diarias que suministran los SMBGs no son suficientes para monitorizar toda la variabilidad glucémica y detectar las excursiones críticas de la glucosa (hiperglucemias postprandiales e hipoglucemias debidas a dosis extras de insulina) que se producen durante todo el día, tanto en pacientes diabéticos como en pacientes críticos. Es debido a esta importante limitación de los glucómetros convencionales, que surgen a finales del siglo XX los llamados sistemas de monitorización continua de glucosa (SMCGs) (-del inglés *Continuous*

Glucose Monitoring Systems (CGMSs)-.

En los últimos tres lustros, el progreso tecnológico sin precedentes en el desarrollo de los SMCGs ha sido decisivo para incrementar su uso como herramienta complementaria a los SMBGs y ha impulsado la construcción del llamado *páncreas artificial* (-del inglés *artificial pancreas (AP)-*), que consiste en desarrollar un sistema de control en lazo cerrado en el cual se integra un SMCG, un algoritmo de control y una bomba de infusión de insulina.

La finalidad principal de un SMCG es ofrecer una *estimación de la concentración de BG de manera continua* (entre uno a cinco minutos), proporcionando al personal sanitario y al paciente información más completa sobre el perfil glucémico. Esta característica permite identificar las tendencias de la glucosa a lo largo del día, facilitar la detección de hipoglucemias e hiperglucemias y tomar las decisiones más acertadas en el manejo de la insulino terapia. Así mismo, la portabilidad de estos dispositivos y la posibilidad de recibir una alerta cuando se han superado algunos umbrales glucémicos pre-establecidos, permiten al paciente y al personal sanitario gestionar con mayor suficiencia los eventos críticos.

Los beneficios de la tecnología de monitorización continua de glucosa (MCG) son claros. El uso de los SMCGs en los pacientes diabéticos se ha asociado con un mejor control metabólico, aunque este resultado está condicionado a su frecuencia de uso (Goldstein and Little, 2004; Rossetti et al., 2010). Por otra parte, dado que en las UCIs, el uso de SMBGs impone una carga adicional para el personal médico y de enfermería, utilizar SMCGs en estas unidades, además de proporcionar un mejor control de la glucosa, reduce la carga de trabajo del personal sanitario al utilizar un menor número de mediciones manuales.

Actualmente, entre toda la tecnología de MCG existente, los dispositivos que se encuentran disponibles para uso comercial son los SMCGs mínimamente invasivos. Estos sistemas, formados por un sensor electroquímico subcutáneo (SC), un transmisor y un monitor externo, ofrecen una estimación de la concentración de BG a partir de mediciones en el fluido intersticial (-del inglés *Interstitial Fluid (ISF)-*). El sensor electroquímico es insertado por vía subcutánea en la región abdominal. Generalmente, la salida del sensor es una señal de *corriente eléctrica* (en nanoamperios, nA) proporcional al nivel de la glucosa en el ISF o glucosa intersticial (-del inglés *Interstitial Glucose (IG)-*). El monitor externo transforma las señales de corriente generadas por el sensor en concentraciones de glucosa, empleando un algoritmo de estimación de BG, también conocido como *algoritmo de calibración*. Este algoritmo estima la glucosa considerando una regresión lineal entre la señal de corriente proveniente del sensor y la información de muestras proporcionadas por los SMBGs (también conocidas como *calibraciones*) para corregir la salida del sensor, tomadas estas muestras con la misma frecuencia con la cuál se toman las mediciones discontinuas (Bequette, 2010). Para medir la precisión de las estimaciones de BG, para cualquier SMCG, las lecturas de glucosa que proporcionan estos dispositivos se tienen que comparar con medidas de referencia, también llamadas “*gold standard*”, las cuales pueden ser muestras de glucosa en sangre venosa, capilar o arterial, dependiendo del contexto.

Sin embargo, existen algunos aspectos críticos de estos dispositivos que se deben estudiar y abordar para mejorar su eficiencia, siendo el más relevante de todos el que sus estimaciones de glucosa resultan imprecisas y poco fiables (Ginsberg, 2007), particularmente en el rango hipoglucémico, donde las tasas de falsos negativos y falsos positivos son inaceptablemente altas (Clarke et al., 2005; Wentholt et al., 2005; Kovatchev et al., 2008; Rossetti et al., 2010). Este es el motivo por el cuál, en la actualidad, ninguno de los SMCGs aprobados por la Agencia de

Drogas y Alimentos (-del inglés *Food and Drug Administration (FDA)*-) y otras agencias reguladoras ha sido aceptado como un sustituto de los tradicionales SMBGs y exclusivamente están aprobados como una herramienta complementaria a las mediciones de glucosa en sangre capilar (Laffel and Kowalski, 2007). Por tanto, con el objetivo que el SMCG sea aprobado para ser usado de forma independiente en la toma de decisiones terapéuticas, y dado que este dispositivo es un componente clave del páncreas artificial (Bondia et al., 2010), se necesita un mayor desarrollo en la tecnología de MCG para lograr la precisión y la fiabilidad requeridas.

1.2. Problemas abiertos y desafíos

Una lista de los principales problemas abiertos en los SMCGs se relaciona a continuación:

- **Aspectos fisiológicos y modelado de la cinética BG-IG.** Debido a que el metabolismo y regulación de la glucosa en el cuerpo humano se basan en la concentración plasmática y los SMCGs disponibles comercialmente se insertan en el tejido SC, se debe estudiar y modelar en detalle la relación que caracteriza la cinética entre los dos compartimentos, plasma e ISF, la cuál puede ser variable en el tiempo, así como el complejo proceso de transporte entre ambos (Sternberg et al., 1996; Rebrin et al., 1999; Monsod et al., 2002; Cengiz and Tamborlane, 2009).
- **Precisión de los algoritmos de calibración.** Un factor determinante en la precisión de los SMCGs es la calidad de los algoritmos de calibración (King et al., 2007). Los algoritmos actuales tienen un enfoque simplista, basado en técnicas de regresión lineal. Dado que la medida de BG estimada por estos algoritmos se obtiene de manera indirecta en un compartimento alternativo a la sangre, el ISF, en su desarrollo no se debe obviar la variación fisiológica introducida por la cinética de la glucosa entre el plasma y este compartimento, la cuál juega un rol clave en la calibración. Este problema ha sido parcialmente investigado en la literatura (Knobbe and Buckingham, 2005; Facchinetti et al., 2010a; Barceló-Rico et al., 2012, 2013).

Otro factor que tampoco se debe obviar es el retardo intrínseco del sensor, que introduce retardos en su respuesta a los cambios en las concentraciones de glucosa (Davey et al., 2010).

Por otra parte, sigue siendo una debilidad importante de la tecnología de MCG la calibración del sensor basándose en las muestras proporcionadas por los SMBGs, lo que condiciona que la calidad de la señal subcutánea sea fuertemente dependiente de la calidad de la calibración, la cuál depende de la variabilidad del medidor de glucosa, así como del elemento humano. Errores en la introducción de estas medidas de referencia pueden dar lugar a un sesgo importante en la señal continua de glucosa (Bequette, 2010).

- **Eliminación de ruido.** Aunque en el hardware de los SMCGs existen filtros para suavizar la señal de corriente, uno de los principales problemas a enfrentar en estos dispositivos es el ruido presente en las señales del sensor, el cuál se debe eliminar. El uso de valores de glucosa filtrados es particularmente importante en aplicaciones que buscan generar alertas y predecir niveles de glucosa. Sin embargo, aunque el problema de encontrar una estrategia óptima para realizar un adecuado filtrado en línea de los datos ha sido tratado (Knobbe

and Buckingham, 2005; Palerm et al., 2005; Kuure-Kinsey et al., 2006; Facchinetti et al., 2010b), las soluciones propuestas no han sido enteramente satisfactorias.

- **Predicción de hipoglucemias y/o hiperglucemias.** Los diferentes métodos de predicción de hipoglucemias e hiperglucemias propuestos en la literatura (Sparacino et al., 2007; Reifman et al., 2007; Gani et al., 2009, 2010; Eren-Oruklu et al., 2010) tienen fortalezas y debilidades. La importancia de estos métodos radica en que al indicar cuando un evento hipoglucémico o hiperglucémico va a ocurrir, permiten prevenir excursiones glucémicas críticas. Sin embargo, estas estrategias se deben mejorar para reducir los errores e incrementar el horizonte de predicción de la glucosa.
- **Fiabilidad de los niveles de BG estimados.** Como se indica en (Hovorka, 2008; Bequette, 2010), un desafío aún mayor es la fiabilidad de los SMCGs actuales. Es importante desarrollar sistemas de detección de fallos y averías para ser embebidos en los SMCGs que permitan a estos dispositivos tener capacidades de auto-monitorización en tiempo real (-del inglés *Real-Time (RT)*). En este caso, la detección de una secuencia de mediciones incorrectas consecutivas puede desencadenar la exigencia de introducir un nuevo punto de calibración por parte del paciente o del personal sanitario.

Adicionalmente, el uso frecuente de los SMCGs ha mostrado que ocasionalmente se presentan pérdidas de señal que no reflejan el verdadero nivel de glucosa y puede producir errores en la calibración. Las estrategias para detectar pérdidas de señal pueden incluir establecer un límite en los cambios de medida permitidos, incorporar alarmas que puedan ser activadas después de un cierto número de muestras sin una señal válida, o cuando existan evidencias que indiquen que una persona está posiblemente entrando en una situación de inseguridad clínica (Bequette, 2010).

La integración de algoritmos de detección de fallos en RT en las nuevas estrategias avanzadas de calibración, en conjunto con la incorporación de la cinética del transporte de glucosa entre compartimentos, podría abrir la puerta a satisfacer las demandas de mejor precisión y fiabilidad de estos dispositivos, requeridas por las agencias reguladoras para su aprobación como herramientas independientes en la toma de decisiones terapéuticas.

1.3. Metodología propuesta

De acuerdo con los desafíos que existen actualmente en la tecnología de MCG, el trabajo de investigación de esta tesis doctoral se enfocará en dos de las líneas de investigación abiertas mencionadas previamente: **precisión de los algoritmos de calibración** y **fiabilidad de los SMCGs**.

Respecto a los problemas de precisión en los SMCGs, estos se centran fundamentalmente en el hecho de que estos dispositivos no miden directamente la concentración de BG. Por tanto, el desarrollo de mejores algoritmos de estimación de BG basados en modelos que describan la relación entre BG e IG, especialmente en condiciones dinámicas, puede contribuir a mejorar significativamente su precisión.

En este contexto, esta investigación aborda los problemas de precisión de los algoritmos de calibración de los SMCGs presentando dos nuevos algoritmos de estimación de BG. El primero de ellos es un algoritmo retrospectivo² estático basado en las mismas técnicas de regresión lineal que utilizan los dispositivos disponibles comercialmente. Este algoritmo presenta un mejor ajuste de las estimaciones de BG, particularmente en la estimación de episodios hipoglucémicos, aunque con la limitación de no poder ser implementado en RT y de no considerar la cinética de la glucosa entre compartimentos para realizar sus estimaciones. El otro algoritmo que se plantea, se caracteriza por considerar la relación BG-IG en condiciones dinámicas basando su principio de funcionamiento en modelos auto-regresivos (-del inglés *AutoRegressive (ARs)*-). Además, tiene la ventaja de poder implementarse en RT. Los modelos ARs se han utilizado en la literatura en numerosas aplicaciones y en el estudio de la DM en particular, se han empleado como herramientas para desarrollar modelos de predicción de glucosa a corto plazo considerando varios horizontes de predicción (Sparacino et al., 2007; Reifman et al., 2007; Eren-Oruklu et al., 2009; Gani et al., 2009, 2010). Sin embargo, en ninguno de estos trabajos se direcciona el problema de desarrollar un nuevo algoritmo de calibración basado en estas técnicas. Este trabajo fue específicamente diseñado para abordar este problema.

Por otra parte, en estos momentos los SMCGs no son lo suficientemente fiables para ser utilizados en la toma de decisiones terapéuticas en los pacientes diabéticos y en los pacientes críticos. En consecuencia, la fiabilidad también es uno de los principales requerimientos para un SMCG, y con el fin de mejorarla, se deben incluir en estos dispositivos sistemas de detección de fallos y averías. El número de trabajos desarrollados con este propósito presentes actualmente en la literatura es limitado (Ward et al., 2004; Bondia et al., 2008; Shen et al., 2010; Facchinetti et al., 2011; Signal et al., 2012). Por tanto, es necesario realizar progresos significativos en este sentido que permitan dotar a los SMCGs con capacidades de auto-monitorización para detectar mediciones erróneas (Hovorka, 2008; Bequette, 2010), así como proporcionar un diagnóstico al personal médico que sea de utilidad para la toma de decisiones clínicas.

En concordancia con lo expresado anteriormente, la capacidad para detectar glucemias erróneas estimadas por un SMCG, a partir de información suministrada por el mismo dispositivo y algunas variables propias de la condición clínica del paciente, adquiere gran importancia clínica debido a que no sería necesario tener una medida de BG como “*gold standard*” para emitir un diagnóstico acerca de la fiabilidad de una medida de glucosa estimada por un SMCG. Estas estrategias de detección de errores en el SMCG posibilitarían el desarrollo del sistema de auto-monitorización que podría ser embebido en el dispositivo.

Sin embargo, al abordar el problema de la detección de fallos en los SMCGs, es muy importante tener en cuenta que el número de muestras de glucemias incorrectas que se obtienen es minoritario. En el contexto de este problema, es fundamental obtener modelos que exhiban un alto rendimiento en la detección de esta clase minoritaria, debido a que usualmente representa el objetivo de la práctica clínica (en una UCI las medidas incorrectas realizadas por el SMCG pueden corresponder a situaciones de alta inestabilidad glucémica en el paciente como episodios hiperglucémicos e hipoglucémicos), sin que ello conlleve a perjudicar la detección de la clase mayoritaria que corresponden a las glucemias estimadas correctamente. En síntesis, el objetivo deben ser modelos con una buena *sensibilidad*, pero al mismo tiempo, con una buena

²Es un algoritmo cuyo procedimiento de estimación de las concentraciones glucosa no puede ser implementado en RT. Después de registrar toda la información correspondiente a la vida útil del sensor SC, el SMCG se le retira al paciente y el algoritmo de calibración trata de ajustar la estimación de BG para minimizar el error relativo absoluto de esta estimación con respecto a BG capilar en los puntos de calibración (Rossetti et al., 2010).

especificidad.

Como paso previo para abordar este problema, en esta investigación se evalúa la precisión de un algoritmo de calibración en RT de un SMCG disponible comercialmente, el Guardian[®] REAL-Time CGMS, utilizado en una UCI polivalente, enfocándose en determinar si el estado séptico del paciente tiene alguna influencia en la precisión de este dispositivo.

Ya entrando en materia de detección de fallos en SMCGs, en este estudio se exploran dos enfoques metodológicos basados en técnicas de aprendizaje y reconocimiento de patrones. Estos enfoques utilizan la información suministrada por el propio SMCG, e incorporando algunas variables de la condición clínica del paciente, permiten detectar estimaciones de BG correctas e incorrectas, sin requerir de una medida de referencia. La clasificación se realiza de acuerdo con esquemas binarios de clasificación, definidos a partir de los criterios de precisión clínicos y analíticos aprobados por las agencias reguladoras para evaluar el rendimiento de estos dispositivos.

El primer enfoque metodológico que se plantea utiliza Análisis de Componentes Principales (-del inglés *Principal Component Analysis (PCA)*-) en combinación con Razonamiento Basado en Casos (-del inglés *Case-Based Reasoning (CBR)*-) para aprender patrones de acuerdo con un esquema binario de clasificación, definido tomando como base un criterio de precisión clínico, y determinar cuando una lectura del SMCG en RT es terapéuticamente correcta (TC) e incorrecta (TI). PCA y CBR son técnicas de monitorización y detección de situaciones anormales, frecuentemente utilizadas en la industria de procesos químicos y también aplicadas con éxito en diferentes campos (Boye et al., 2008; Mujica et al., 2011; Ruiz et al., 2011).

El segundo enfoque metodológico que se propone aplica Máquinas de Soporte Vectorial (-del inglés *Support Vector Machines (SVMs)*-) no convencionales al problema de la detección de fallos en un SMCG, empleando dos estrategias de post-procesamiento específicas para ser usadas en conjuntos de datos desbalanceados. Las SVMs no convencionales son una atractiva opción para abordar este problema por dos razones: 1) su mecanismo de aprendizaje considera un pequeño subconjunto de patrones para separar situaciones con fallos de situaciones sin fallos (Cristianini and Shawe-Taylor, 2000), y 2) utilizan diversas estrategias que buscan mejorar su rendimiento sobre conjuntos de datos desbalanceados. En este caso, el aprendizaje de patrones se realiza de acuerdo con un esquema binario de clasificación, definido tomando como base un criterio de precisión analítico.

Los enfoques metodológicos desarrollados en esta tesis pueden ser utilizados por los fabricantes de SMCGs disponibles comercialmente para mejorar la exactitud y fiabilidad de estos dispositivos e incluir en ellos capacidades de auto-monitorización.

1.4. Objetivos

El objetivo global de esta tesis es desarrollar nuevos enfoques metodológicos para mejorar la precisión y fiabilidad de los SMCGs.

A continuación, se definen los objetivos específicos de esta tesis.

1. Diseñar un algoritmo retrospectivo de estimación de BG, basado en técnicas de regresión lineal, con la capacidad de mejorar el ajuste de las estimaciones de glucosa en el rango hipoglucémico.
2. Diseñar un algoritmo de estimación de BG, que considere la relación dinámica entre BG e IG y satisfaga el requerimiento de poder implementarse en RT, con el fin de mejorar la precisión de las concentraciones de glucosa estimadas, particularmente en el rango hipoglucémico.
3. Desarrollar un enfoque metodológico mediante PCA y CBR, que a partir de información del propio SMCG en RT y de algunas variables relacionadas con la condición clínica del paciente, permita detectar si las estimaciones de BG realizadas por este dispositivo son correctas o incorrectas, de acuerdo con un esquema binario de clasificación, definido tomando como base un criterio de precisión clínico.
4. Desarrollar un enfoque metodológico basándose en SVMs no convencionales (aplicando dos estrategias de post-procesamiento) específicas para trabajar con conjuntos de datos desbalanceados, que a partir de información suministrada por el propio SMCG en RT y la dosis de insulina endovenosa, permita detectar si las estimaciones de BG realizadas por este dispositivo son correctas o incorrectas, de acuerdo con un esquema binario de clasificación, definido tomando como base un criterio de precisión analítico.

1.5. Estructura de la tesis

A continuación, se expone una descripción más detallada de todos los capítulos de la tesis:

- *Capítulo 1.* Es el capítulo introductorio. En este capítulo se exponen la motivación de la investigación, el planteamiento del problema y los objetivos de la investigación.
- *Capítulo 2.* Inicia con una visión general de la DM, sus complicaciones, las cifras actuales y previsiones futuras. A continuación, se describen las características y el estado actual de la tecnología de MCG, haciendo énfasis en los SMCGs mínimamente invasivos, específicamente en los que se han utilizado para proporcionar los datos utilizados en esta investigación, para luego abordar los diferentes problemas abiertos en este campo. En particular, se profundiza en estudiar los algoritmos de estimación de BG usados y/o desarrollados en la literatura, así como los enfoques metodológicos planteados para mejorar la fiabilidad en los SMCGs. Luego, se hace una descripción de los diferentes criterios que se emplean para evaluar su precisión y fiabilidad. Finalmente, se mencionan algunas limitaciones para avanzar en esta área de investigación.
- *Capítulo 3.* Está orientado a proponer y describir dos nuevos algoritmos de estimación de BG para abordar los problemas de precisión presentados por los SMCGs. Empieza con una descripción del proceso de experimentación clínica usado para la recolección de los datos en pacientes con DM-1. Más adelante, se introduce el algoritmo del CGMS® Gold™ y se describe el algoritmo retrospectivo para estimación de BG que se propone (*algoritmo retrospectivo propuesto (ARP-SMCG)*). Por otra parte, se describe el algoritmo de estimación de BG, el cuál considera la relación dinámica entre BG e IG y satisface el requerimiento de poder ser implementado en RT (*algoritmo basado en modelos ARs (ABMAR-SMCG)*).

Además, se presentan los resultados de la validación de los dos algoritmos propuestos, empleando criterios de precisión analíticos y clínicos, así como una comparación entre los perfiles de glucemia obtenidos con cada uno de los algoritmos planteados y los que se obtienen usando el algoritmo del CGMS® Gold™.

Los resultados de este capítulo se publicaron en:

- [Leal2009] Leal, Y., Garcia-Gabin, W., Bondia, J., Vehí, J. (2009) “Real-time glucose estimation algorithm for continuous glucose monitoring using autoregressive models.” Póster, En *9th Diabetes Technology Meeting (DTM)*. 5-7 de Noviembre, San Francisco, EE. UU.
- [Leal2010a] Leal, Y., Garcia-Gabin, W., Bondia, J., Esteve, E., Ricart, W., Fernández-Real, J.M., Vehí, J. (2010) “Real-time glucose estimation algorithm for continuous glucose monitoring using autoregressive models.” *Journal of Diabetes Science and Technology*. 4(2):391-403. PMID: 20307401 [PubMed - indexado en MEDLINE].
- [Leal2010b] Leal, Y., Garcia-Gabin, W., Bondia, J., Esteve, E., Ricart, W., Fernández-Real J.M., Vehí, J. (2010) “Enhanced algorithm for glucose estimation using the continuous glucose monitoring system.” *Medical Science Monitor*. 16(6):MT51-8. IF: 1.699 (2010) Q3 [indexado en JCR]. PMID: 20512100 [PubMed - indexado en MEDLINE].

- *Capítulo 4*. En primer lugar, se introducen algunos conceptos generales de uso común en las UCIs. A continuación, se realiza una descripción de los pacientes y del entorno clínico del estudio y se expone el protocolo del estudio para MCG en los pacientes críticos. Luego, se determina la precisión del SMCG utilizando criterios de precisión analíticos y clínicos. Más adelante, se presenta el análisis de los resultados obtenidos. Finalmente, se indican algunas limitaciones del estudio, las cuáles deberían ser consideradas en investigaciones futuras.

Los resultados de este capítulo se publicaron en:

- [Lorenzo2010] Lorenzo, C., Leal, Y.T., Bonet, A., Vehí, J., Bondia, J., Taché, A., Sirvent, J.M. (2010) “Accuracy assessment of real-time continuous glucose monitors in critically ill patients: Preliminary results.” En *23rd European Society of Intensive Care Medicine Annual Congress*. 9-13 de Octubre, Barcelona, España.
- [Leal2010c] Leal, Y., Lorenzo, C., Bondia, J., Palerm, C.C., Bonet, A., Vehí, J. (2010) “Accuracy assessment of real-time continuous glucose monitoring systems in an intensive care unit: A pilot study in medical and surgical patients.” Póster, En *10th Diabetes Technology Meeting (DTM)*. 11-13 de Noviembre, Bethesda, EE. UU.
- [Leal2011a] Leal, Y., Lorenzo, C., Bondia, J., Bonet, A., Vehí, J. (2011) “Accuracy assessment of real-time continuous glucose monitoring system in critically-ill patients.” En *Workshop on Control, Dynamics, Monitoring and Applications*. 7-9 de Febrero, Caldes de Montbui, España.
- [Leal2011b] Leal, Y., Lorenzo, C., Bondia, J., Palerm, C.C., Sirvent, J.M., Bonet, A., Vehí, J. (2011) “Numerical-accuracy assessment of a real-time continuous glucose monitoring system in medical, surgical, and trauma critically-ill patients.” Póster, En *4th International Conference on Advanced Technologies and Treatments for Diabetes (ATTD)*. 16-19 de Febrero, Londres, Reino Unido.

- [**Lorencio2011a**] Lorencio, C., **Leal, Y.**, Bonet, A., Vehi, J., Bondia, J., Taché, A., Sirvent, J.M., Palerm, C.C. (2011) “Valoración de la precisión de la monitorización continua de la glucosa intersticial en pacientes críticos.” Póster, En *XXVI Congreso Nacional de la Sociedad Española de Nutrición Parenteral y Enteral (SENPE)*. 24-27 de Mayo, Gerona, España.
- [**Lorencio2011b**] Lorencio, C., **Leal, Y.**, Bonet, A., Vehi, J., Bondia, J., Taché, A., Palerm, C.C., Sirvent, J.M. (2011) “Variabilidad de la precisión de un sistema de monitorización continua de glucosa en pacientes críticos dependiendo de su situación hemodinámica.” Póster, En *XLVI Congreso Nacional de la Sociedad Española de Medicina Intensiva, Crítica y Unidades Coronarias (SEMICYUC)*. 12-15 de Junio, Bilbao, España.
- [**Lorencio2011c**] Lorencio, C., **Leal, Y.**, Bonet, A., Vehí, J., Bondia, J., Palerm, C.C., Taché, A., Sirvent, J.M. (2011) “Accuracy assessment of subcutaneous continuous glucose monitoring in an intensive care unit according to the septic shock state.” En *24th European Society of Intensive Care Medicine Annual Congress*. 1-5 de Octubre, Berlín, Alemania.
- [**Lorencio2012**] Lorencio, C., **Leal, Y.**, Bonet, A., Bondia, J., Palerm, C.C., Tache, A., Sirvent, J.M., Vehi, J. (2012) “Real-time continuous glucose monitoring in an intensive care unit: better accuracy in patients with septic shock.” *Diabetes Technology and Therapeutics*, 14(7):568-75. IF: 2.205 (2012) Q3 [indexado en JCR]. PMID: 22512288 [PubMed - indexado en MEDLINE].
- *Capítulo 5.* Comienza con la descripción de los datos utilizados, y la explicación de la transformación del problema de detección de fallos en un SMCG en un problema de bi-clasificación, de acuerdo con un criterio de precisión clínico. Luego, se introducen los conceptos de PCA y se describe la construcción del modelo PCA. Más adelante, se presentan los conceptos básicos de CBR y se explica la implementación de la metodología propuesta para detectar medidas de glucosa TC y TI en un SMCG. Seguidamente, se presentan los resultados experimentales de la implementación del enfoque metodológico propuesto. Además, otras técnicas de clasificación, tales como Análisis Discriminante por Mínimos Cuadrados Parciales (-del inglés *Partial Least Squares Discriminant Analysis (PLS-DA)*-) y Modelado Blando Independiente de Analogías entre Clases (-del inglés *Soft Independent Modeling of Class Analogy (SIMCA)*-), se incluyen para realizar una comparación. Finalmente, se analizan los resultados obtenidos.

Los resultados de este capítulo se publicaron en:

- [**Leal2012a**] **Leal, Y.**, Ruiz, M., Lorencio, C., Bondia, J., Vehí, J. (2012) “Statistical methods and expert systems for the detection of therapeutically correct and incorrect measurements in RTCGMS.” Póster, En *5th International Conference on Advanced Technologies and Treatments for Diabetes (ATTD)*. 8-11 de Febrero, Barcelona, España.
- [**Leal2013b**] **Leal, Y.**, Ruiz, M., Lorencio, C., Bondia, J., Mujica, L., Vehí, J. (2013) “Principal component analysis in combination with case-based reasoning for detecting therapeutically correct and incorrect measurements in continuous glucose monitoring systems.” *Biomedical Signal Processing and Control. En impresión*, (DOI: 10.1016/j.bspc.2013.05.008). IF: 1.074 (2012) Q3 [indexado en JCR].

- *Capítulo 6.* En primer lugar, se explica la transformación del problema de detección de fallos en un SMCG en un problema de bi-clasificación, de acuerdo con un criterio de precisión analítico. A continuación, se introducen los conceptos de SVMs y se describen las estrategias de post-procesamiento para ajustar el umbral de la función decisión. Luego, se expone el enfoque metodológico desarrollado para detectar medidas de glucosa correctas e incorrectas en un SMCG aplicando nuevas estrategias de post-procesamiento de SVMs. Los resultados experimentales y el análisis de los mismos, también se incluyen al final de este capítulo.

Los resultados de este capítulo se publicaron en:

- [Leal2012b] Leal, Y., Gonzalez-Abril, L., Lorenzo, C., Ruiz, M., Bondia, J., Vehí, J. (2012) “Un nuevo enfoque para detectar medidas de glucosa erróneas en los Sistemas de Monitorización Continuos de Glucosa.” En *XIV Jornadas sobre Sistemas cualitativos y sus Aplicaciones en Diagnosis, Robótica e Inteligencia Ambiental (JARCA)*. 24-28 de Junio, Tarragona, España.
- [Leal2012c] Leal, Y., Gonzalez-Abril, L., Ruiz, M., Lorenzo, C., Bondia, J., Ortega, JA, Vehí, J. (2012) “Applying new support vector machines post-processing strategies for the detection of correct and incorrect measurements in CGMS.” Póster, En *12th Diabetes Technology Meeting (DTM)*. 8-10 de Noviembre, Bethesda, EE. UU.
- [Leal2013c] Leal, Y., Gonzalez-Abril, L., Lorenzo, C., Vehí, J. (2013) “Detección de medidas incorrectas en los SMCGs aplicando una nueva estrategia de post-procesamiento de SVM para ajustar el sesgo en función de la sensibilidad.” En *XV Jornadas sobre Sistemas cualitativos y sus Aplicaciones en Diagnosis, Robótica e Inteligencia Ambiental (JARCA)*. 24-27 de Junio, Murcia, España.
- [Leal2013a] Leal, Y., Gonzalez-Abril, L., Lorenzo, C., Bondia, J., Vehí, J. (2013) “Detection of correct and incorrect measurements in real-time continuous glucose monitoring systems by applying a post-processing support vector machine.” *IEEE Transactions on Biomedical Engineering*. En impresión, (DOI: 10.1109/TBME.2013.2244092). IF: 2.348 (2012) Q2 [indexado en JCR].

- *Capítulo 7.* Resume las conclusiones y aportes de esta investigación, así como el trabajo futuro.

Capítulo 2

Sistemas de Monitorización Continua de Glucosa (SMCGs)

La diabetes mellitus (DM) es una enfermedad crónica que está adquiriendo proporciones epidémicas, con alrededor de 366,2 millones de personas afectadas por esta enfermedad a nivel mundial en el 2011 (International Diabetes Federation, 2011), un número que se espera que crezca a 551,8 millones en 2030. Estas cifras probablemente infraestimadas sustentan el gran interés social y económico que existe en la actualidad en la búsqueda de un control eficiente de la glucosa.

Los SMCGs son elementos claves para realizar un control glucémico eficiente tanto en las personas con DM como en los pacientes no diabéticos en situaciones críticas. Por esta razón, uno de los objetivos de la línea de investigación de enfoque tecnológico que busca mejorar la calidad de vida de estas personas consiste en desarrollar y mejorar los dispositivos para estimar glucosa en plasma (-del inglés *Blood Glucose (BG)*-) de manera continua.

Este capítulo se enfoca en presentar el problema de la monitorización continua de glucosa (MCG) y en revisar el estado del arte de los algoritmos de estimación de BG usados y/o desarrollados en la literatura, así como de las estrategias planteadas para mejorar la fiabilidad en los SMCGs. El capítulo se organiza de la siguiente manera: la sección 2.1 introduce una visión general de la DM, sus complicaciones, las cifras actuales y las previsiones futuras de esta enfermedad. A continuación, en la sección 2.2 se describen las características y el estado actual de la tecnología de MCG, haciendo énfasis en los SMCGs mínimamente invasivos, específicamente en los que se han utilizado para proporcionar los datos utilizados en esta investigación, para luego en la sección 2.3 abordar los diferentes problemas abiertos en este campo. En particular, se profundiza en estudiar los algoritmos de estimación de BG usados y/o desarrollados en la literatura, así como los enfoques metodológicos planteados para mejorar la fiabilidad en los SMCGs. Luego, en la sección 2.4 se hace una descripción de los diferentes criterios que se emplean para evaluar la precisión y fiabilidad de estos dispositivos. Finalmente, en la sección 2.5 se mencionan algunas limitaciones para avanzar en esta área de investigación y se presenta un resumen del capítulo en la sección 2.6.

2.1. Visión de conjunto sobre la DM

Considerada como una epidemia mundial, según la definición de la Federación Internacional de Diabetes (-del inglés *International Diabetes Federation (IDF)*-), la DM es una enfermedad

crónica que aparece cuando el páncreas no produce suficiente *insulina*, o cuando el cuerpo no puede utilizar eficazmente la insulina que produce. La insulina es una hormona reguladora secretada por las células β de los Islotes de Langerhans en el páncreas (Figura 2.1), que estimula el flujo de transportadores de *glucosa*¹, permitiendo su difusión a las células a través de la membrana celular en el músculo y tejido adiposo, así como su almacenamiento en forma de glucógeno en el hígado. De esta manera, la insulina permite que las células utilicen la glucosa de la sangre para obtener energía. El fracaso en la producción de insulina, o de la insulina para actuar apropiadamente, o ambos, conducen a niveles anormalmente elevados de glucosa en la sangre. A este estado, al que se le conoce como hiperglucemia, se le asocia con daños a largo plazo para el cuerpo y fallos de varios órganos y tejidos.

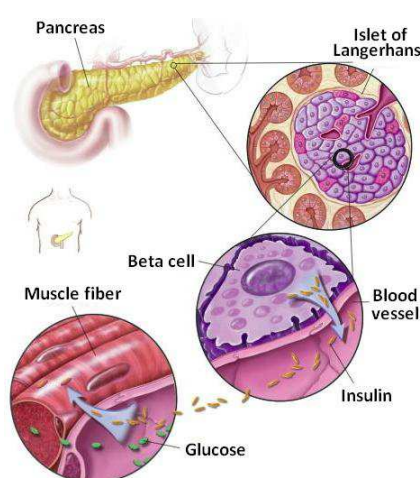


Figura 2.1: Producción de insulina en el páncreas humano.²

Para ser considerados normales, los niveles de concentración de BG en el ser humano deben permanecer entre 70 – 140 mg/dL. A este estado se le conoce como normoglucemia o euglucemia. Para una persona sana, el nivel de glucosa basal³ es de 90 mg/dL. Para un paciente diabético, el nivel basal se encuentra entre 110 – 140 mg/dL. Sin embargo, si el paciente no realiza un seguimiento de su DM y un control glucémico adecuado, estos niveles normales se incrementan de manera significativa, llegando a desarrollar diferentes complicaciones a corto y largo plazo asociadas con su enfermedad. Por lo tanto, el objetivo de control de los pacientes con DM debe ser tratar de mantener sus niveles de glucosa de manera constante en un rango entre 70 – 180 mg/dL.

Clínicamente, el control glucémico se evalúa mediante la medición de la HbA_{1c}. Esta prueba, en conjunto con la monitorización de la glucosa en casa, se utiliza como base para los ajustes en el tratamiento. Es importante resaltar, que el hecho de tener un conocimiento frecuente y preciso del comportamiento de la BG y sus variaciones, es fundamental para realizar un buen control glucémico.

¹Moléculas transportadoras de glucosa que al tener un peso molecular inferior a 100 pueden atravesar fácilmente la membrana celular. La glucosa tiene un peso molecular de 180. Las más comunes son la familia de proteínas transportadoras conocidas como GLUT.

²Imagen tomada de (National Institutes of Health, 2006).

³Nivel de glucemia después del ayuno nocturno.

2.1.1. Cifras actuales y previsiones futuras

Según los últimos datos de la IDF, se estima que el número de diabéticos en todo el mundo aumentará de 366,2 millones en 2011 hasta 551,8 millones en 2030, representando el 8,3% y el 9,9%, respectivamente, de la población adulta mundial con edades comprendidas entre 20 y 79 años (Figura 2.2) (International Diabetes Federation, 2011). El aumento mundial de estas cifras es causado por el envejecimiento y el crecimiento poblacional, la urbanización y la alta prevalencia de la obesidad y la vida sedentaria. Para hacerse una idea del gran impacto social de esta enfermedad, se puede apreciar en la Figura 2.2 que para el año 2030, el menor incremento en el porcentaje de pacientes con DM será superior al 20,0% (en Europa), una cifra que es alarmante, teniendo en cuenta además, que en las restantes regiones del planeta se superará este valor y que se ha proyectado un incremento mundial del 51,0%. En España, según cifras de la IDF para el año 2011, 2,84 millones de adultos (el 8,1% de la población nacional) sufren esta enfermedad (Whiting et al., 2011). Sin embargo, la mayor preocupación se relaciona con el hecho que estas estadísticas son probablemente una subestimación de la prevalencia futura de la pandemia de la DM.

Map: IDF Regions and global projections of the number of people with diabetes (20-79 years), 2011 and 2030

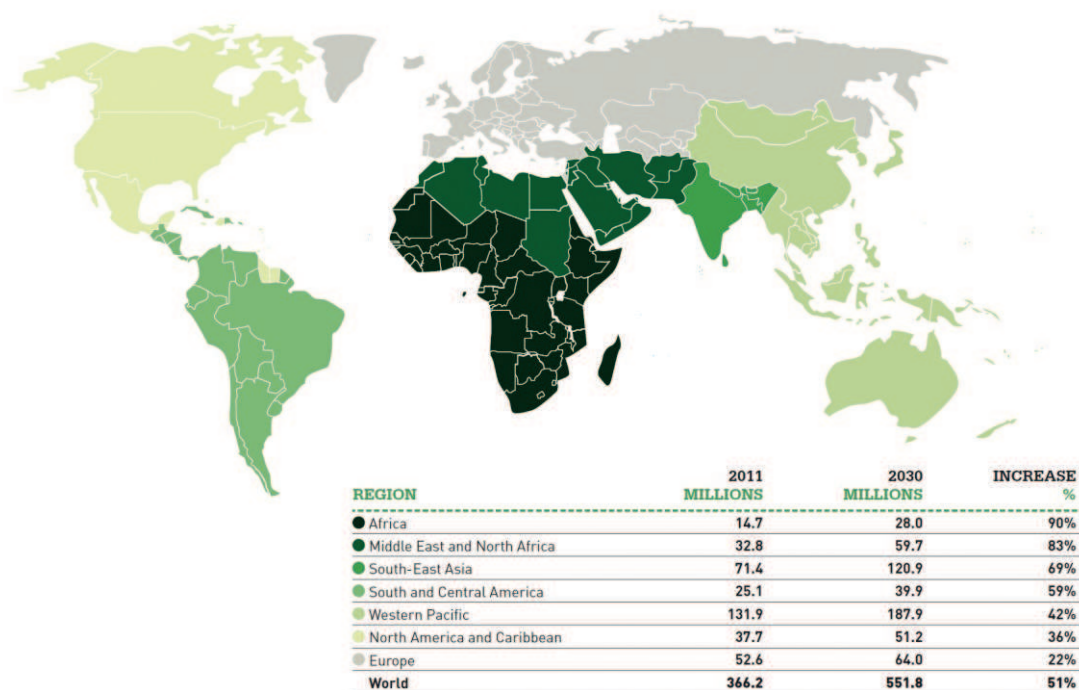


Figura 2.2: Proyecciones globales y regionales de la IDF sobre el número de personas con diabetes entre los 20-79 años, 2011 y 2030.⁵

La Figura 2.3 (International Diabetes Federation, 2011) ilustra el actual panorama mundial de la DM, indicando la prevalencia de la enfermedad y el porcentaje de DM no diagnosticada por región. En esta figura se puede observar que alrededor de la mitad de las personas con DM

⁵Imagen tomada de (International Diabetes Federation, 2011).

no son conscientes de que padecen la enfermedad.

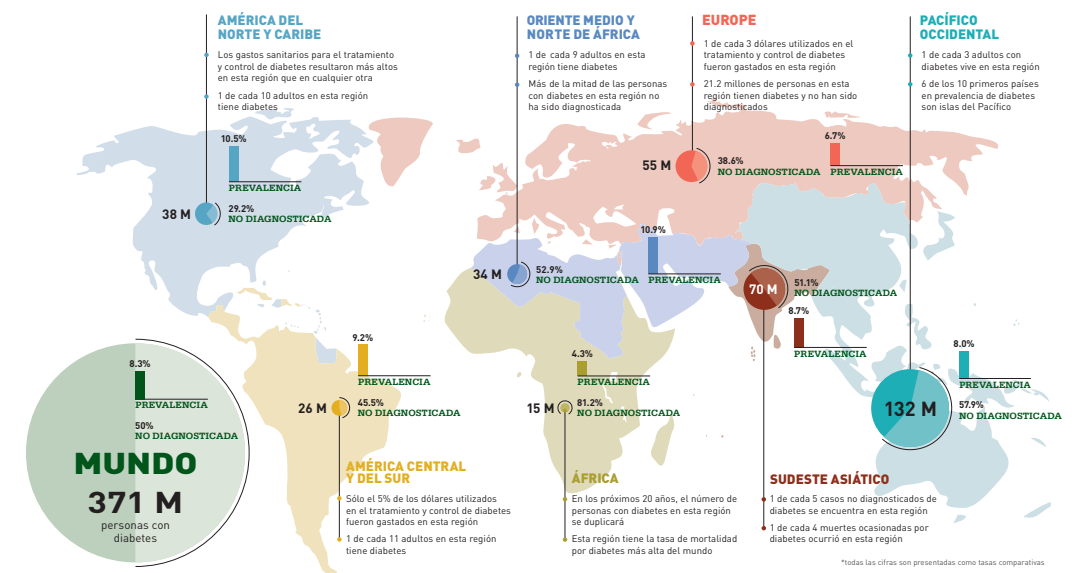


Figura 2.3: Panorama mundial de la diabetes en el 2011 por región.⁶

Los costos de la atención médica de una persona con DM duplican los costos sanitarios de una persona que no tiene la enfermedad. En Europa, el gasto sanitario por persona diabética se estima entre US \$3.000 y US \$6.500 al año, siendo estas cifras superadas en los Estados Unidos (Roglic and Unwin, 2010). Además, según un estudio de (FEN, 2008) y la IDF, la diabetes es responsable de más del 10,0% del gasto sanitario en la Unión Europea. Por las razones expuestas anteriormente, existen beneficios económicos y sociales claros en la identificación de terapias eficaces para la DM.

2.1.2. Tipos de DM

De acuerdo a una clasificación etiológica de la DM realizada por la Asociación Americana de la Diabetes (-del inglés *American Diabetes Association (ADA)*-) (American Diabetes Association, 2010), existen principalmente dos tipos de diabetes:

- Diabetes mellitus tipo 1 (DM-1). También conocida como diabetes insulino-dependiente o de inicio en la infancia. Se presenta ante la incapacidad del páncreas de segregar insulina en sangre debido a la destrucción de las células β . La sección endocrina del páncreas se concentra en los Islotes de Langerhans, un grupo de células integrada por células α , encargadas de segregar el glucagón⁷, las células β , que producen insulina y las células δ , generadoras de somatostatina⁸. La destrucción de las células β se produce por parte del sistema inmunológico del organismo, el cuál como consecuencia de la forma como reacciona frente a infecciones virales y otros procesos aún desconocidos (se cree predisposición

⁶Imagen tomada de (International Diabetes Federation, 2011).

⁷Hormona contrareguladora que permite el aumento del nivel de BG.

⁸Hormona que inhibe la secreción de insulina y glucagón

genética), las ataca y elimina.

Las personas con DM-1 producen muy poca o ninguna insulina. Sin insulina, la glucosa permanece en el torrente sanguíneo, por lo que aumentan los niveles de glucosa en sangre, especialmente después de la ingesta de alimentos (Hanas, 2004). Las personas con este tipo de diabetes necesitan inyectarse insulina exógena todos los días con el fin de controlar los niveles de glucosa en la sangre. La enfermedad puede afectar a personas de cualquier edad, pero generalmente es diagnosticada en la infancia o en la adolescencia.

Los síntomas de la DM-1 son cansancio, hambre, poliuria⁹, polidipsia¹⁰ y pérdida de peso. La hiperglucemia no tratada puede conducir a complicaciones graves a largo plazo, incluyendo enfermedades cardiovasculares, insuficiencia renal, ceguera, y accidentes cerebrovasculares.

- Diabetes mellitus tipo 2 (DM-2). También conocida como diabetes no insulino-dependiente o de inicio en la edad adulta. Se manifiesta clínicamente por resistencia a la insulina, lo que conduce a niveles elevados de glucosa en la sangre, derivando en una eventual pérdida en la liberación de insulina desde el páncreas. Inicialmente se trata con fármacos orales, para incrementar la sensibilidad insulínica o para aumentar la producción de insulina. Con el tiempo, el paciente debe ser tratado con inyecciones de insulina, debido a la disminución de las células β , de manera análoga a como se hace en la DM-1.

El diagnóstico de la DM-2 por lo general se produce después de los 40 años, pero puede ocurrir antes, sobre todo en poblaciones con alta prevalencia de diabetes. A menudo se asocia con la obesidad, que en sí misma puede causar resistencia a la insulina y conducir a niveles elevados de glucosa en la sangre. La DM-2 representa entre el 90,0 – 95,0% de los casos de diabetes.

2.1.3. Complicaciones por DM

La DM-1 y la DM-2 son enfermedades crónicas que requieren un seguimiento y un control cuidadoso. Sin un manejo adecuado, pueden llevar a niveles muy elevados de glucosa en sangre que pueden causar daños a diferentes órganos y tejidos. Las complicaciones asociadas a la DM pueden ser de dos tipos:

Complicaciones a corto plazo

- Hipoglucemia (glucosa inferior a 70 mg/dL)¹¹. Es un efecto secundario muy común ocasionado por la terapia intensiva con insulina y el no ejercer un buen control de la enfermedad. Los síntomas de la hipoglucemia incluyen: debilidad o temblores, sudoración, mareos, hambre, nerviosismo y confusión. Si la hipoglucemia no se trata, puede causar dolor de cabeza, irritabilidad, mala circulación, mala coordinación, y adormecimiento en la boca y la len-

⁹El volumen urinario excesivo, cuya denominación médica es poliuria, se refiere a la emisión anormal de grandes cantidades de orina cada día.

¹⁰Es una sensación anormal de requerir siempre el consumo de líquidos.

¹¹En numerosos estudios se fija el límite de la hipoglucemia en 80 mg/dL.

gua. Una hipoglucemia grave puede producir alteración de la conciencia, convulsiones, e incluso puede desembocar en un coma diabético o la muerte, en casos extremos.

- Hiperglucemia (glucosa superior a 180 mg/dL)¹². Es una consecuencia de la enfermedad y se caracteriza porque en el organismo se elevan las concentraciones de BG. Los síntomas de la hiperglucemia son sed excesiva y hambre, micción frecuente, náuseas, vómitos y falta de energía. La hiperglucemia prolongada puede llevar a una condición conocida como *cetoacidosis*¹³.

Complicaciones a largo plazo

- Enfermedades microvasculares¹⁴. Pueden ocasionar complicaciones oculares, renales o del sistema nervioso.
- Enfermedades macrovasculares¹⁵. Pueden ocasionar enfermedades cardíacas, accidentes cerebrovasculares o enfermedades vasculares periféricas.

Una descripción de las características de las diferentes complicaciones por DM se presenta en la Tabla 2.1.

Tejido u órgano afectado	Enfermedad	Complicaciones
Enfermedades microvasculares		
Ojos	Glaucoma y Retinopatía	Pérdida de la visión y en última instancia, la ceguera
Riñones	Nefropatía	Función renal deficiente, insuficiencia renal y necesidad de diálisis o trasplante de riñón.
Sistema Nervioso	Neuropatía	Entumecimiento y hormigueo, especialmente en los pies y las piernas, sensibilidad al tacto o debilidad muscular, dolor crónico, entumecimiento y pérdida de masa muscular, llagas y úlceras, mareos o desmayos debidos a disminución de la presión arterial después de estar de pie o sentado.
Enfermedades macrovasculares		
Corazón	Coronaria	Ataques al corazón.
Cerebro	Cerebrovascular	Afasia, mareos, pérdida de equilibrio o coordinación, repentina visión borrosa, visión doble o disminución de la visión en uno o ambos ojos, accidentes cerebrovasculares.
Arterias periféricas	Vascular periférica	Úlceras en el pie, infecciones, e incluso la pérdida de un dedo del pie, el pie o la pierna.

Tabla 2.1: Complicaciones de la DM a largo plazo.¹⁶

¹²Al igual que en el apartado anterior este límite puede variar en diferentes estudios clínicos, generalmente entre 180 a 240 mg/dL.

¹³Es un problema que ocurre en personas con diabetes y se presenta cuando el cuerpo no puede usar la glucosa como fuente de energía, debido a que no hay insulina o ésta es insuficiente. Cuando la glucosa no está disponible, se metaboliza la grasa en su lugar. A medida que las grasas se descomponen, los ácidos llamados cetonas se acumulan en la sangre y la orina. En niveles altos, las cetonas son tóxicas. Cuando esto sucede, la afección se conoce como *cetoacidosis diabética (CAD)*. Si no se trata rápidamente, una CAD puede inducir un coma diabético.

¹⁴Enfermedades en vasos sanguíneos pequeños.

¹⁵Enfermedades en vasos sanguíneos grandes.

¹⁶Tabla adaptada de (García-Jaramillo, 2011).

2.1.4. Tratamiento para la DM-1

En 1993, el Estudio sobre Control de la Diabetes y sus Complicaciones (-del inglés *Diabetes Control and Complications Trial (DCCT)*-) (DCCT, 1993) demostró que un buen control glucémico, determinado este control por el nivel de HbA_{1c}, reduce el riesgo de complicaciones crónicas microvasculares asociadas con la DM-1. Por esta razón, se ha establecido el rango entre 70 mg/dL – 180 mg/dL como el rango permanente dentro del cuál los pacientes con DM-1 deben tratar de tener sus niveles de glucosa, es decir, este es el rango objetivo a controlar, excepto cuando exista alguna contraindicación. Si el nivel de BG es menor a 70 mg/dL, el paciente está en hipoglucemia y si el nivel es superior a 180 mg/dL, el paciente está en hiperglucemia. El tratamiento requerido para lograr este objetivo de control de la glucosa implica la inyección de la insulina en el cuerpo, el ejercicio y una dieta saludable.

El tratamiento de la DM-1 consiste básicamente en compensar los bajos niveles de insulina endógena en el cuerpo humano administrando infusiones de insulina exógena (basal¹⁷ y bolus¹⁸). El páncreas libera insulina en una persona sana después de una comida, entre las comidas y durante la noche. Por lo tanto, en un paciente con DM-1, la insulina en bolus emula la secreción de insulina después de una comida, mientras que la insulina basal emula la secreción de insulina entre las comidas y durante la noche. Una combinación de insulina basal y bolus, la denominada *terapia intensiva de insulina* o *insulinoterapia*, se puede utilizar como tratamiento para lograr niveles normales de glucosa. La terapia intensiva de insulina se basa en tres puntos (Becker, 2001):

- Medida del nivel de glucosa en sangre preprandial¹⁹.
- Estimación de la cantidad de carbohidratos a ingerir en la comida.
- Estimación e inyección de la dosis apropiada de insulina exógena, calculada esta dosis en relación con la cantidad de carbohidratos a ingerir, y en concordancia con las normas prescritas por el médico.

La terapia intensiva de insulina utiliza análogos de insulina²⁰, los cuáles actúan como un suplemento facilitando la absorción de glucosa por parte de las demás células y tejidos del cuerpo humano. La frecuencia y cantidad de la dosis depende del peso de cada paciente, la actividad física, los carbohidratos consumidos, la sensibilidad a la insulina, etc.

El suministro de la insulina se hace por vía subcutánea mediante la terapia MDI (-del inglés *Multiple Daily Injections*-) que consiste en varias inyecciones al día, o por medio de la terapia CSII (-del inglés *Continuous Subcutaneous Insulin Infusion*-) que utiliza una bomba de infusión continua de insulina. Sin embargo, la principal desventaja de la terapia intensiva de insulina, en cualquiera de sus formas, es que aumenta el número de episodios hipoglucémicos (DCCT, 1993).

¹⁷Controla el nivel de glucosa en la sangre entre las comidas y durante la noche. Generalmente se da una o dos veces al día y se puede utilizar sola o en combinación con los medicamentos antidiabéticos orales o con las insulinas de acción rápida. Incluye las insulinas basales de acción intermedia (NPH) y prolongada (Glargina y Detemir).

¹⁸Se da antes de las comidas para controlar el aumento de glucosa después de comer. Incluye las insulinas en bolo de acción rápida (Lispro, Aspart, Glulisina) y de corta acción (Insulina regular)

¹⁹Concentración de glucosa en sangre antes de la ingesta de alimentos

²⁰Los análogos de insulina o insulinas “de diseño”, los cuales han sido diseñados genéticamente para asemejarse más a la insulina producida por el páncreas. Se inyectan con una jeringa, un “bolígrafo” de insulina o una bomba de insulina. Los análogos de insulina incluyen las insulinas basales de acción prolongada (Glargina y Detemir) y las insulinas en bolo de acción rápida (Lispro, Aspart y Glulisina).

2.1.5. Estado de la investigación actual

De acuerdo con (Dia, 2013; JDR, 2013), en estos momentos las dos principales líneas de investigación para mejorar la calidad de vida de los pacientes con DM-1 se centran en:

1. Las *células madre* para sustituir las células β del páncreas.
2. Una solución tecnológica. Este segundo enfoque involucra diferentes líneas de investigación (Barceló-Rico, 2012):
 - La producción de nuevos análogos de insulina, que sean más eficientes y que tengan una mejor capacidad para emular la secreción pancreática.
 - El desarrollo de nueva tecnología invasiva, mínimamente invasiva y no invasiva para MCG, así como mejoras en el desempeño de estos dispositivos (Rapp et al., 2011).
 - El desarrollo de nuevas tecnologías para administrar la insulina y mejoras en el rendimiento de las ya existentes: bombas de infusión de insulina por vía subcutánea y “bolígrafos” de insulina.
 - El desarrollo de sistemas de soporte a la toma de decisiones y de plataformas de telemedicina.
 - El desarrollo de un sustituto de la célula β del páncreas, el cuál pretende simular su función endocrina, el llamado control automático de la infusión de insulina o también denominado “*páncreas artificial*” (AP).

El trabajo que se desarrolla en esta tesis doctoral se enmarca dentro de la segunda línea de investigación, es decir, dentro de la solución tecnológica.

En relación con el AP, su finalidad es proporcionar un control glucémico totalmente automatizado, o en lazo cerrado (Kumareswaran and Hovorka, 2009). Un AP consiste de un SMCG acoplado a un dispositivo de suministro de insulina y un algoritmo de lazo cerrado que proporciona la dosis correcta de insulina en el momento adecuado sin intervención humana (Harvey et al., 2010). De acuerdo a esta definición, los elementos claves para el desarrollo de un AP son (Barceló-Rico, 2012):

- SMCGs. Estos dispositivos suministran lecturas frecuentes de los niveles de glucosa (generalmente cada cinco minutos), las cuáles permiten identificar las variaciones y tendencias de la glucosa a lo largo del día. Las lecturas de glucosa que proporcionan corresponden a estimaciones de BG medidas desde el fluido intersticial (-del inglés *Interstitial Fluid (ISF)*-).
- Bombas para la infusión de insulina. Son dispositivos externos que permiten la administración de insulina por vía subcutánea de manera continua. Si bien existen bombas de infusión de insulina implantables, hasta la fecha su uso está limitado a Francia. Estos dispositivos tienen un compartimento donde se coloca la insulina, la cuál se administra por medio de la conexión que existe entre este compartimento y una aguja que se inserta en el tejido subcutáneo (SC). Estas bombas están programadas para liberar la insulina basal durante 24 horas, así como para suministrar el bolus de insulina requerido antes de cada ingesta, previa programación por parte del paciente.
- Algoritmos de control. Este es el elemento que toma la “*decisión*” sobre la cantidad de insulina a ser administrada en cada instante de manera automática (García-Jaramillo et al., 2012).

En el desarrollo del AP, la ruta que ha mostrado el mayor potencial para abordar el problema de control en lazo cerrado es la denominada *ruta s.c-s.c.*, es decir, con monitorización continua de la glucosa subcutánea (s.c) e infusión de insulina subcutánea (s.c) (Figura 2.4) (Bondia et al., 2010). Esta ruta se caracteriza por su carácter mínimamente invasivo y la aceptación de los SMCGs y las bombas de insulina SCs.

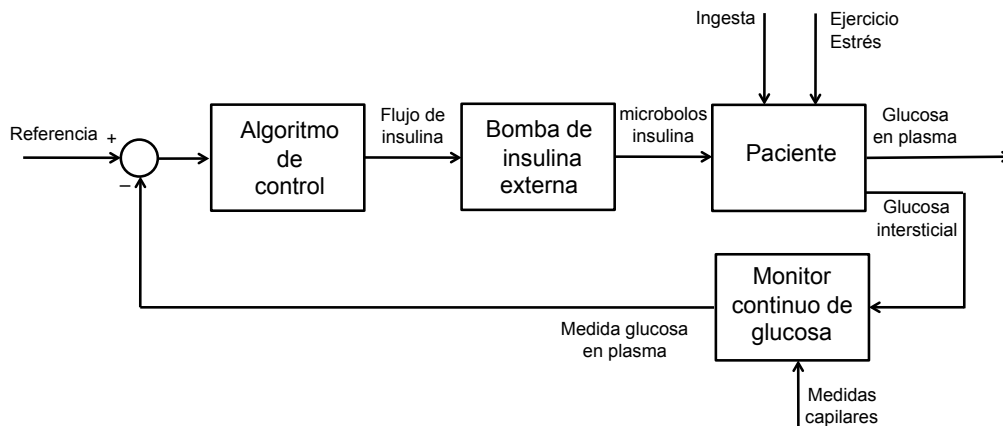


Figura 2.4: Bucle de control de glucosa en lazo cerrado empleando la ruta con monitorización continua de la glucosa subcutánea (s.c) e infusión de insulina subcutánea (s.c).²²

En los últimos años se han logrado avances significativos en el desarrollo del AP. Hoy en día, el problema radica en las exigencias para “*cerrar el lazo*”, puesto que se requiere encontrar modelos adecuados que representen la fisiología de la glucosa, SMCGs con mejores resultados (más fiables y precisos), mejor desempeño de las bombas de infusión de insulina, algoritmos de control que tengan en cuenta las desventajas del uso de SMCGs y bombas por vía subcutánea, así como del apoyo de la telemedicina.

De acuerdo con (Bondia et al., 2010), los grandes retos abiertos para el desarrollo del AP son los siguientes:

- Adecuada corrección a la lenta respuesta de los controladores debida a retardos en el lazo de control, imponiendo restricciones en la infusión de insulina con el objetivo de minimizar el riesgo de hipoglucemia. El uso de bombas externas tiene como limitación que al no ser fisiológica la infusión por vía subcutánea, se introduce un retraso adicional en la acción de control, asociada a la absorción subcutánea de insulina. Este retraso fisiológico envuelve una sobre-actuación excesiva que ocasionaría una hipoglucemia al paciente debido a la acumulación de la insulina.
- Abordar el problema de la sobre-actuación en la compensación de las ingestas, donde los errores son significativamente elevados.
- Robustez frente a la variabilidad entre los pacientes o inter-paciente (-del inglés *inter-patient*-) y frente a la gran variabilidad del comportamiento fisiológico interno del paciente o intra-paciente (-del inglés *intra-patient*-).

²²Imagen tomada de (Bondia et al., 2010).

- Hacer frente a las grandes perturbaciones que afectan al sistema, como el efecto de la composición de la comida sobre la absorción de la glucosa y el efecto del estrés y el ejercicio sobre la dinámica de la glucosa.
- Mejorar los problemas de exactitud y fiabilidad de los SMCGs con el objetivo de reducir los grandes errores que los dispositivos actuales presentan en la estimación de BG.
- Integrar estrategias que garanticen la seguridad en las bombas de infusión de insulina, así como la detección de fallos en estos dispositivos.

En cuanto a los principales actores en el desarrollo del AP, además de la Comisión Europea, ha sido muy importante el trabajo realizado por la Fundación para la Investigación de la Diabetes Juvenil (-del inglés *Juvenile Diabetes Research Foundation (JDRF)*-) y por el Instituto Nacional de Salud en los EE. UU. (-del inglés *National Institute of Health (NIH)*-). Globalmente, los avances más importantes en el campo del desarrollo del AP se han hecho a través del *Proyecto Páncreas Artificial* (Kowalski and Lum, 2009), puesto en marcha por la JDRF en 2006 para validar la eficacia de las nuevas tecnologías en MCG y fomentar el desarrollo de simuladores que permitieran “*cerrar el lazo*”, enlazando los SMCGs con bombas de insulina.

2.2. Monitorización Continua de la Glucosa (MCG)

Hasta principios del siglo XIX la forma más común de diagnosticar la DM-1 era mediante la catación de la orina del paciente o exponiéndola a la interperie para ver si esta atraía a los insectos, pues la orina de estos pacientes contiene grandes cantidades de glucosa. La primera prueba química para determinar la presencia de glucosa en la orina aparece en 1841, por tanto, hasta mediados del siglo XIX, la única forma que tenían los pacientes con DM-1 para determinar la glucemia era la prueba de glucosa en la orina. Sin embargo, a principios de la década de 1980, se logra un avance significativo en la medición de la glucosa, al empezar a utilizarse de manera directa por los pacientes los primeros glucómetros de uso domiciliario, también conocidos como SMBGs.

Los SMBGs miden los niveles de glucemia de forma puntual en el tiempo mediante la extracción de una pequeña muestra sanguínea por punción capilar, arterial o venosa, la cual se coloca en una tira reactiva o en un sensor para ser analizada. Estos dispositivos también se caracterizan por:

- Proporcionar datos discretos y altamente precisos acerca de los niveles de glucosa en sangre.
- No tener capacidad predictiva.
- Ser invasivos y dolorosos.
- Ser de fácil análisis.

Generalmente, un paciente que controla su DM-1 realiza alrededor de cuatro punciones diarias (una antes de cada comida y una antes de ir a dormir).

En la actualidad, los SMBGs constituyen un elemento indispensable para que los pacientes puedan ejercer un mejor control glucémico (Goldstein and Little, 2004; Rossetti et al., 2010). Sin embargo, las escasas determinaciones de glucosa diarias que realizan no son suficientes para seguir toda la variabilidad glucémica y detectar los episodios críticos que se producen diariamente,

por ejemplo, hiperglucemias postprandiales²³ e hipoglucemias debidas a excesos en las dosis de insulina. Esta incapacidad mostrada por los SMBGs para reflejar las tendencias globales en los niveles de glucosa de un paciente y monitorizar todas sus excursiones glucémicas ha motivado la investigación en nuevos dispositivos, los denominados SMCGs, con el objetivo de medir la concentración de BG de forma continua y no de manera intermitente.

En la última década, ha habido un progreso tecnológico sin precedentes en el desarrollo de los SMCGs. Hoy en día, este tipo de dispositivos miden la concentración de glucosa de manera frecuente (entre uno a cinco minutos), y tienen una vida útil mayor (hasta siete días). Según (Vaddiraju et al., 2010a), un SMCG típicamente consiste de:

- Un sensor de glucosa que continuamente mide niveles fisiológicos de glucosa (en el plasma sanguíneo o en el ISF).
- Una unidad de procesamiento electrónico que está en comunicación con el sensor de glucosa (cableada o inalámbrica).
- Una unidad de visualización de datos.

Mayores detalles acerca de los diferentes dispositivos que integran la tecnología de MCG se proporcionarán en la sección 2.2.1.

Se ha demostrado que los SMCGs han ayudado a los pacientes con DM-1 a realizar una mejor gestión de su enfermedad, proporcionando información detallada acerca de la variabilidad de la glucosa: dirección, magnitud, duración y frecuencia de hipoglucemias o hiperglucemias (Ellis et al., 2008; Ginsberg, 2007; Klonoff, 2005). Por otra parte, al permitir identificar las tendencias de la glucosa a lo largo del día, estos dispositivos a su vez pueden llegar a ser predictivos del nivel de glucemia en el futuro.

Adicionalmente, su desarrollo ha impulsado la investigación en el AP, con el objetivo de lograr un control glucémico lo más cercano a la normalidad en los pacientes con DM-1. Existen evidencias importantes que indican que el uso de SMCGs en estos pacientes podría mejorar su control glucémico en comparación con lo que se logra con los SMBGs (Kovatchev and Clarke, 2007; JDRF et al., 2008), dado que los SMCGs han mostrado efectos positivos en la reducción de la HbA_{1c} (Deiss et al., 2006; JDRF et al., 2008).

Sin embargo, existen algunos aspectos críticos de estos dispositivos que se deben estudiar y abordar para mejorar su eficiencia. Una descripción más detallada de estos aspectos se realizará en la sección 2.3, pero es importante mencionar que de todos ellos, el más relevante, está relacionado con la inexactitud y poca fiabilidad de sus estimaciones de glucosa. En la actualidad, ninguno de los SMCGs aprobados por la FDA y otras agencias reguladoras tiene una precisión analítica y una fiabilidad comparable a la de los SMBGs, particularmente en el rango hipoglucémico, donde las tasas de falsos negativos y falsos positivos son inaceptablemente altas (Clarke et al., 2005; Wentholt et al., 2005; Kovatchev et al., 2008; Rossetti et al., 2010). Por este motivo, los SMCGs sólo están aprobados como una herramienta para complementar las mediciones capilares realizadas mediante glucómetros, pero no para sustituirlas (Bondia et al., 2010). Hoy en día, estos dispositivos siguen siendo considerados el cuello de botella en el diseño del AP y los problemas de precisión y fiabilidad que presentan son un obstáculo relevante para que sean usados de manera independiente en la toma de decisiones terapéuticas.

²³Concentración de glucosa en sangre después de la ingesta de alimentos

2.2.1. Tecnología actual para MCG

Desde el punto de vista del grado de inserción del sensor de glucosa (si penetra la piel o no lo hace), los SMCGs pueden clasificarse en tres categorías (Figura 2.5) (Vaddiraju et al., 2010a):

- Dispositivos invasivos: Constituidos por sensores de larga duración completamente implantables (Vaddiraju et al., 2010b; Renard, 2008).
- Dispositivos mínimamente invasivos: Su objetivo es evitar la presencia permanente de un objeto extraño (sensor) en el cuerpo. Se caracterizan por medir las concentraciones de glucosa en un fluido (plasma o ISF) (Rao et al., 1993; Mitragotri et al., 2000; Robert, 2002; Gebhart et al., 2003; Newman and A.P., 2005; Rhee et al., 2007; Park EJ, 2009). Se proporcionarán más detalles sobre estos dispositivos en la sección 2.2.2.
- Dispositivos no invasivos: Utilizan técnicas de espectroscopía para analizar diferentes fluidos o gases corporales tales como saliva, lágrimas y gases respiratorios que permitan medir la concentración de glucosa transcutánea (Tura et al., 2007; De Block et al., 2008; Ferrante do Amaral and Wolf, 2008; Oliver et al., 2009).

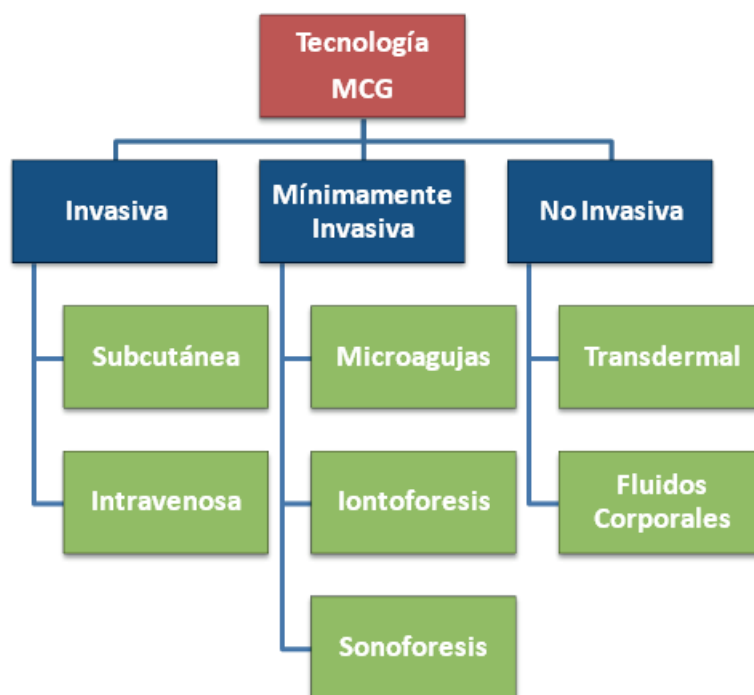


Figura 2.5: Tecnologías de MCG desde el punto de vista de su capacidad invasiva.²⁴

La Tabla 2.2 describe las características más relevantes de las diferentes tecnologías de MCG sus ventajas e inconvenientes. Una revisión más exhaustiva acerca del campo de tecnologías de MCG existentes y sus aspectos críticos se ha publicado en (Oliver et al., 2009; Vaddiraju et al., 2010a; Ramchandani and Heptulla, 2012).

²⁴Imagen adaptada de (Vaddiraju et al., 2010a).

Tecnología de MCG	Modalidad	Ventajas	Inconvenientes	Bibliografía
<i>Invasiva</i>	Subcutánea	<ul style="list-style-type: none"> ■ La herida no permanece abierta. ■ No presentan variabilidad intrapaciente. ■ Comodidad y facilidad de adaptación. ■ <u>Facilidad de implantación.</u> 	<ul style="list-style-type: none"> ■ Inexactitud en la calibración debida a la falta de correlación entre el plasma y el ISF. ■ Respuesta del organismo a la presencia de un cuerpo extraño. ■ Contaminación biológica inducida por degradación del sensor. ■ Migración del sensor y dificultad en su extracción. 	<ul style="list-style-type: none"> ■ (Vaddiraju et al., 2010b) ■ (Renard, 2008)
	Intravenosa	<ul style="list-style-type: none"> ■ La herida no permanece abierta. ■ No presentan variabilidad intrapaciente. ■ Comodidad y facilidad de adaptación. 	<ul style="list-style-type: none"> ■ Respuesta del organismo a la presencia de un cuerpo extraño. ■ Contaminación biológica inducida por degradación del sensor y daño del mismo ocasionado por la fuerza del flujo sanguíneo. ■ Migración del sensor y dificultad en su extracción, así como <u>un procedimiento de implantación tedioso.</u> 	
<i>Mínimamente invasiva</i>	Microagujas	<ul style="list-style-type: none"> ■ Determina BG de manera directa, por tanto, la precisión no está influenciada por la calibración. 	<ul style="list-style-type: none"> ■ El paciente es propenso a padecer infección, irritación o sangrado debido a la generación de múltiples heridas. ■ Requiere de mucho tiempo para recoger suficiente cantidad de <u>sangre para determinar la concentración de glucosa.</u> 	<ul style="list-style-type: none"> ■ (Robert, 2002) ■ (Newman and A.P., 2005)
	Iontoforesis	<ul style="list-style-type: none"> ■ La herida no permanece abierta. ■ La cantidad de fluido extraído es baja y, por tanto, no hay efecto del oxígeno en el sensor de glucosa enzimático. ■ Menor contaminación del electrodo. 	<ul style="list-style-type: none"> ■ Requiere de un largo tiempo de inicialización para obtener las lecturas de glucosa. ■ Inexactitud en la calibración debida a la falta de correlación entre el plasma y el ISF. ■ No se debe utilizar en condiciones de excesiva transpiración. ■ Aplicación de pequeñas corrientes pueden causar eritema en la piel. 	
	Sonoforesis	<ul style="list-style-type: none"> ■ La herida no permanece abierta. 	<ul style="list-style-type: none"> ■ No presentan variabilidad sujeto a sujeto. ■ Inexactitud en la calibración debida a la falta de correlación entre el plasma y el ISF. 	
<i>No invasiva</i>	Transdermal	<ul style="list-style-type: none"> ■ Sin dolor. ■ Comodidad y adaptabilidad al paciente. 	<ul style="list-style-type: none"> ■ Inexactitud como resultado de la pigmentación de la piel, contenido de agua en el cuerpo, e hidratación. ■ Fuerte efecto de la temperatura en la respuesta del nivel de glucosa. 	<ul style="list-style-type: none"> ■ (Tura et al., 2007) ■ (Oliver et al., 2009) ■ (De Block et al., 2008) ■ (Ferrante do Amaral and Wolf, 2008)
	Fluidos Corporales	<ul style="list-style-type: none"> ■ Sin dolor. ■ Comodidad y adaptabilidad al paciente. 	<ul style="list-style-type: none"> ■ Inexactitud como resultado de la pigmentación de la piel, contenido de agua en el cuerpo, e hidratación. 	

Tabla 2.2: Comparación de diferentes dispositivos para MCG desde el punto de vista de su capacidad invasiva.²⁵

Actualmente, entre toda la tecnología de MCG disponible, los más utilizados son los SMCGs mínimamente invasivos. Estos dispositivos se utilizan en el diseño del AP, específicamente en la ruta con mayor potencial para cerrar el lazo, la llamada *ruta s.c-s.c.* Además, puesto que son los dispositivos que se comercializan hoy en día, también son los que comúnmente se utilizan como herramienta complementaria para ayudar al personal médico en la toma de decisiones terapéuticas.

Por otra parte, los dispositivos no invasivos, si bien se caracterizan por ser una tecnología prometedora dada su capacidad para medir la glucosa de una manera indirecta a partir de ciertas señales, aún no se comercializan. La razón es que las agencias reguladoras consideran que sus estimaciones no son lo suficientemente precisas y fiables. Además, presentan un rendimiento inferior al que muestran los dispositivos mínimamente invasivos, que ya de por sí son cuestionados por sus limitaciones de precisión y fiabilidad.

2.2.2. SMCGs mínimamente invasivos

Los SMCGs mínimamente invasivos disponibles actualmente en el mercado estiman las concentraciones de glucosa mediante sensores electroquímicos insertados en el tejido SC por medio de una cánula. El principio básico es, en todos ellos, la oxidación de la glucosa en el ISF (en otras palabras, la oxidación de la glucosa intersticial (-del inglés *Interstitial Glucose (IG)*-)), en presencia de la enzima glucosa-oxidasa (-del inglés *Glucose-Oxidase (GOx)*-) (Bondia et al., 2010). Esta reacción electroquímica produce una señal de *corriente eléctrica* (en, nA) al aplicar un potencial en el electrodo, la cuál es representativa de la concentración de glucosa en el ISF (Vaddiraju et al., 2010a). Por tanto, la corriente eléctrica proporcionada por el sensor SC corresponde a la concentración de IG. Dado que el objetivo es obtener la estimación de los niveles de BG, después de un proceso de filtrado, los SMCGs incorporan algoritmos que a partir de las medidas de corriente eléctrica, alcanzan este objetivo. Estos métodos, también conocidos como algoritmos de calibración, requieren ser calibrados por medio de determinaciones de glucosa capilar al menos cuatro veces al día (Figura 2.6).

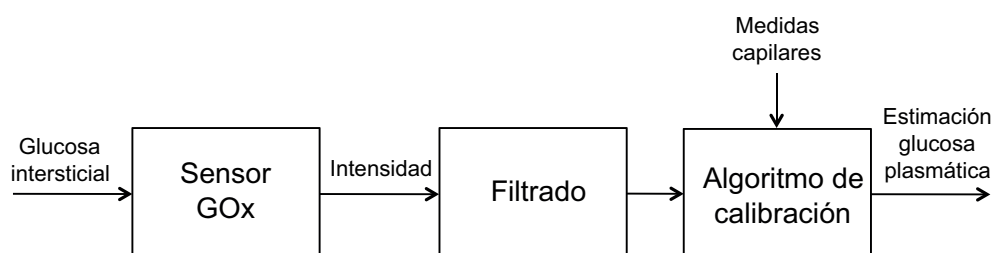


Figura 2.6: El proceso de estimación de BG en los SMCGs mínimamente invasivos.²⁶

En la literatura, los SMCGs mínimamente invasivos más utilizados incluyen cuatro sistemas que emplean sensores amperométricos SCs y un dispositivo que utiliza una técnica basada en *microdiálisis*²⁷ Una revisión casi exclusivamente enfocada en los principios químicos básicos de

²⁵Tabla adaptada de (Vaddiraju et al., 2010a).

²⁶Imagen tomada de (Bondia et al., 2010).

²⁷Esta técnica consiste en insertar una microfibras para microdiálisis en la pared abdominal, la cuál conecta una

la mayoría de los dispositivos que se mencionan a continuación se presenta en (McGarraugh, 2009).

- Medtronic CGMS[®] Gold[™]. Este dispositivo fue el primer SMCG aprobado por la FDA en 1999 para uso domiciliario. Fue introducido en el mercado en el año 2000 (Gross et al., 2000) en su modelo original llamado Medtronic MiniMed[®] CGMS. El sistema incluye un sensor que se inserta en el tejido SC durante tres días, y el hardware necesario para registrar los datos de corriente obtenidos por el sensor y las determinaciones de glucosa capilar utilizadas para la calibración del dispositivo. La información registrada se procesa y analiza retrospectivamente. Aunque no está aprobado para proporcionar estimaciones de los niveles de glucosa en RT, por tanto se le considera como un dispositivo “ciego”, el análisis de los datos de manera retrospectiva suministra información relevante para mejorar la insulino terapia. Además, se caracteriza por mostrar un buen rendimiento en la estimación de estados normoglucémicos e hiperglucémicos y una alta sensibilidad durante las primeras horas de inserción del sensor. En la actualidad, este dispositivo se comercializa en Europa y en EE.UU. en su versión mejorada con el nombre de iPro[™] 2. En América Latina se sigue comercializando como CGMS[®] Gold[™].
- Medtronic Guardian[®] REAL-Time (Mastrototaro et al., 2002). Originalmente fue aprobado por la FDA en el año 2005 con el nombre de Guardian[®] RT como una herramienta complementaria para MCG cada cinco minutos, pero no para sustituir las pruebas de BG realizadas con un SMBG. El modelo actual se empezó a comercializar desde el 2007 integrando un sensor cuya vida útil era de tres días. Actualmente, utiliza un nuevo sensor, llamado por el fabricante Enlite[®], el cuál se caracteriza por tener un tamaño más reducido y una vida útil de seis días. Para calibrar este dispositivo se exige un punto de calibración cada 12 horas y se recomienda la realización de dos a cuatro glucemias capilares al día.

Además, en el 2006 fue lanzado al mercado un producto que integraba un Guardian[®] RT con una bomba de insulina y que originalmente se denominó Paradigm[®] REAL-Time. Los modelos más recientes de este producto son el MiniMed Paradigm[®] Veo[™] y el MiniMed Paradigm[®] REVEL[™]. En conjunto con el Guardian[®] RT, la bomba de insulina puede dar más flexibilidad al paciente, al permitirle personalizar la administración de insulina para satisfacer las necesidades específicas de su terapia (MED, 2013b).

- Menarini GlucoDay[®] S (Poscia et al., 2003). Es el único dispositivo basado en el método de *microdiálisis* que ha sido comercializado a la fecha. Fue introducido en el mercado en el 2002 y aunque no está aprobado por la FDA, recibió la aprobación de la marca CE con arreglo a la directiva relativa a productos sanitarios, requisito indispensable para la comercialización de un producto sanitario en la Unión Europea. Sólo ha sido aprobado para ser utilizado por personal médico cualificado y debidamente formado, quienes insertan la fibra de microdiálisis y calibran el dispositivo. Este SMCG puede proporcionar datos de manera retrospectiva y en RT durante un período de 48 horas, periodo después del cuál el medico remueve la fibra (MEN, 2013).
- DexCom[™] STS[™]-7 (Simpson et al., 2008.). El dispositivo original denominado DexCom[™] STS[™]-3 recibió la aprobación de la FDA en 2006 para ser utilizado como una herramienta adjunta a los test de glucosa en sangre capilar, limitando su uso a proporcionar datos en

microbomba a un biosensor. Las microbomba bombea una solución de perfusión a través de la microfibrá; como el fluido fluye a través de la microfibrá bajo la piel del paciente, se recoge la glucosa a través de la membrana de diálisis y se transporta al biosensor, donde se mide la concentración de glucosa.

RT cada cinco minutos durante tres días. En 2007, pasó a ser el único dispositivo cuyo sensor ha sido aprobado para funcionamiento continuo durante siete días. La calibración se realiza mediante glucosa capilar cada 12 horas (DEX, 2013).

- Abbott FreeStyle NavigatorTM (Feldman and McGarraugh, 2007.). Este dispositivo también fue aprobado como una herramienta complementaria a los SMBGs por la FDA en 2008. La vida útil del sensor es de cinco días, pero no se reporta el valor de la glucosa durante las 10 primeras horas. La calibración utilizando un glucómetro se requiere después de 10, 12, 24 y 72 horas durante la vida útil del sensor. Este dispositivo no se comercializa en EE.UU. desde marzo de 2012, pero se mantiene su comercialización a nivel europeo y mundial (ABB, 2013).

La Tabla 2.3 sintetiza las principales características de los SMCGs mínimamente invasivos más utilizados en la literatura. Esta tabla ha sido construida a partir de la información suministrada en (Reach, 2008; McGarraugh, 2009; Oliver et al., 2009; MED, 2013b; MEN, 2013; DEX, 2013; ABB, 2013).

Producto	Fabricante	Uso	Principio	Método de análisis	Duración vida útil del sensor	Frecuencia de la calibración	Frecuencia de muestreo
CGMS [®] Gold TM	Medtronic-MiniMed	Análisis retrospectivo	Sensor amperométrico SC	Reacción enzimática GOx	72 horas	Entre tres a cuatro al día	Cinco minutos
GlucoDay [®]	A. Menarini Diagnostics	Visualización en RT y análisis retrospectivo	Fibra microdiálisis	Microdiálisis	48 horas	Una en 48 horas (en el modo en RT se recomiendan dos)	Tres minutos
Guardian [®] REAL-Time	Medtronic-MiniMed	RT	Sensor amperométrico SC	Reacción enzimática GOx	144 horas	Entre tres a cuatro al día	Cinco minutos
STS TM -7	DexCom	RT	Sensor amperométrico SC	Reacción enzimática GOx	168 horas	Una cada 12 horas	Cinco minutos
FreeStyle Navigator TM	Abbott	RT	Sensor amperométrico SC	Reacción enzimática GOx	120 horas	Período de inicialización de 10 horas. A las 12, 24 y 72 horas de uso	Un minuto

Tabla 2.3: Características de los SMCGs mínimamente invasivos más utilizados en la literatura.

Es importante clarificar que actualmente todos los SMCGs mencionados en la Tabla 2.3 están disponibles para ser comercializados en diferentes países.

Por otra parte, entre los SMCGs enunciados previamente, los dispositivos que se han utilizado para proporcionar los datos de este trabajo de investigación son el Medtronic CGMS[®] GoldTM y el Medtronic Guardian[®] REAL-Time. Una descripción más detallada de estos dos SMCGs se presenta a continuación.

El Medtronic CGMS[®] GoldTM

El CGMS[®] GoldTM (Medtronic MiniMed, Northridge, CA) (Figura 2.7) fue el primer dispositivo de monitorización de glucosa desarrollado para calcular la concentración de glucosa de manera continua. Su uso se ha evaluado en diversas poblaciones tales como niños (DirecNet, 2005; Gandrud et al., 2007), mujeres embarazadas (Kerssen et al., 2004; Buhling et al., 2004) y

pacientes con DM-2 (Zhou et al., 2008), tanto en enfermos hospitalarios como ambulatorios.

El CGMS[®] Gold[™] está integrado por cuatro componentes principales: un sensor SC de glucosa, un cable de comunicaciones, un monitor externo y una estación de enlace de comunicaciones de datos adaptada al monitor (-del inglés *Com-Station*-). El CGMS[®] Gold[™] registra los valores de glucosa cada cinco minutos en un rango de 40 mg/dL a 400 mg/dL (2,2 mmol/L a 22,2 mmol/L), proporcionando hasta 288 lecturas en un periodo de 24 horas.

El sensor es un electrodo pequeño, estéril y flexible compuesto por un soporte de platino recubierto por un sistema enzimático GOx. El electrodo está cubierto con un polímero de poliuretano biocompatible para protección del sensor y comodidad del paciente, y está unido a un conector que se adhiere a la superficie de la piel y establece una estrecha conexión con el cable. El sensor de glucosa se inserta en el tejido celular SC mediante una aguja introductora rígida. A continuación la aguja se retira y el conector se sujeta a la piel con un vendaje.

En el tejido SC entra en contacto con el ISF, el cuál contiene IG en una cantidad proporcional a BG. La enzima GOx reacciona electroquímicamente con IG, generando electrones en una cantidad proporcional a la glucosa, y por tanto, convirtiendo las concentraciones de IG en señales eléctricas en la superficie del sensor. El sensor de glucosa envía estas señales al monitor de manera continua, a través del cable. El monitor externo capta las señales eléctricas generadas por el sensor cada 10 segundos. A estas señales se les realiza un filtrado inicial, considerando las seis muestras obtenidas en un minuto, de las cuáles se descarta el valor más alto y el valor más bajo y se promedian los cuatro valores restantes para obtener una muestra cada un minuto. Seguidamente, se aplica un nuevo proceso de filtrado, ignorando el valor más bajo y el más alto de la señal muestreada cada un minuto y promediando las tres muestras restantes para obtener una señal promedio cada cinco minutos.

Las señales eléctricas (con tiempo de muestreo de cinco minutos) y las determinaciones de glucosa capilar introducidas para calibración se almacenan en la memoria del monitor. Mediante la *Com-Station*, las señales eléctricas y los valores de glucosa capilar almacenados se descargan a un ordenador personal y por medio de un algoritmo de calibración basado en técnicas de regresión lineal, se convierten en concentraciones de BG. Para realizar este proceso, este dispositivo utiliza el “MiniMed Solutions Software versión 3.0”, al que denominaremos a partir de ahora algoritmo del CGMS[®] Gold[™]. Información adicional sobre este algoritmo se proporcionará en el Capítulo 3.

La calibración del CGMS[®] Gold[™]²⁸ la realiza el paciente o el personal de enfermería mediante la determinación de glucosa por punción digital²⁹. Para conseguir unos resultados de calibración óptimos, se debe introducir el valor de la glucosa capilar tres o cuatro veces a lo largo del día. Estos valores de calibración deben estar comprendidos en el rango de 40 mg/dL a 400 mg/dL.

La precisión del CGMS[®] Gold[™] se ha evaluado en diferentes estudios clínicos (Gross et al.,

²⁸Proceso de ajuste de la manera en que el SMCG convierte las señales eléctricas procedentes del sensor de glucosa en valores de BG. El SMCG se calibra introduciendo periódicamente en su memoria las lecturas de un SMBG (MiniMed, 2003).

²⁹Método diagnóstico para medir la glucosa capilar, en el cuál se pincha la yema de un dedo, se coloca una gota de sangre en una tira reactiva, y ésta se coloca en un instrumento espectrofotométrico que indica la concentración de glucosa (MiniMed, 2003).

2000; Steil et al., 2003; Kerksen et al., 2005; Clarke et al., 2005; Hoi-Hansen et al., 2005; Fiallo-Scharer and Direcnet, 2005; Gandrud et al., 2007). Algunos estudios muestran una buena precisión entre los valores de glucosa estimados por este dispositivo y los valores obtenidos con un analizador de glucosa, considerados como medidas de referencia (Gross et al., 2000; Steil et al., 2003; Gandrud et al., 2007). Otro estudio evaluó la reproducibilidad del dispositivo en mujeres embarazadas con DM-1 y concluyó que su reproducibilidad es adecuada (Kerksen et al., 2005). En contraste, otros estudios indican que el CGMS[®] Gold[™] reporta resultados inexactos, especialmente en el rango hipoglucémico (Clarke et al., 2005; Hoi-Hansen et al., 2005), particularmente sobrestimando los bajos niveles de glucemia que se presentan durante la noche (Fiallo-Scharer and Direcnet, 2005). Sin embargo, la fiabilidad del CGMS[®] Gold[™] en la estimación de estados normoglucémicos e hiperglucémicos y la alta sensibilidad que mantiene durante las primeras horas después de la inserción del sensor, indican que este dispositivo se puede considerar como una buena herramienta de monitorización de la glucosa, particularmente durante cortos períodos de tiempo (Lesperance et al., 2007).



Figura 2.7: El CGMS[®] Gold[™]30

El Medtronic Guardian[®] REAL-Time CGMS (Guardian[®] RTCGMS)

El Guardian[®] RTCGMS (Medtronic MiniMed, Northridge, CA) (Figura 2.8) fue el primer SMCG aprobado por la FDA como un dispositivo auxiliar para monitorizar la glucosa de manera continua en RT. Al igual que el CGMS[®] Gold[™], el Guardian[®] RTCGMS consta de cuatro componentes principales: un sensor SC de glucosa, un transmisor, un monitor externo y un dispositivo de enlace de comunicaciones de datos (-del inglés *ComLink*-). El Guardian[®] RTCGMS también registra los valores de glucosa cada cinco minutos en un rango de 40 mg/dL a 400 mg/dL. Además, este dispositivo incluye alarmas para indicar si los niveles de glucemia sobrepasan los niveles predefinidos o caen por debajo de ellos, así como alarmas de tendencia y alarmas predictivas.

El sensor SC es un electrodo de características similares al sensor del CGMS[®] Gold[™]. Las diferencias están en que la transmisión de la señal del sensor al monitor se realiza por radiofrecuencia. El proceso de calibración de este dispositivo también se asimila al descrito para el CGMS[®] Gold[™]. Sin embargo, a diferencia del CGMS[®] Gold[™], este monitor indica en RT las concentraciones de BG calculadas por su algoritmo de calibración. Por lo tanto, no es necesario esperar a descargar los datos para conocer el valor de la glucosa, puesto que estos valores se

³⁰Imagen tomada de (MED, 2013b).

pueden observar en la unidad de visualización del dispositivo. Por medio de la *ComLink*[®], la información almacenada en la memoria del monitor se puede descargar a un ordenador para su posterior análisis e interpretación utilizando el software “Medtronic CareLink[™]”.



Figura 2.8: El Guardian[®] RTCGMS³¹

La precisión del Guardian[®] RTCGMS se ha evaluado en diversos estudios (Deiss et al., 2006; Kovatchev et al., 2008; Mazze et al., 2009). En términos generales, estos estudios han demostrado un incremento en el tiempo durante el cuál los pacientes se mantienen en estado normoglucémico cuando utilizan el Guardian[®] RTCGMS. Su rendimiento ha sido bueno en los rangos normoglucémico e hiperglucémico, pero ha mostrado limitaciones en el rango hipoglucémico. Además, con su uso se ha reportado una reducción de la HbA_{1c}, lo que indica un mejor control de la DM-1 (Deiss et al., 2006).

2.3. Técnicas propuestas en la literatura para abordar los problemas abiertos en la tecnología de MCG mínimamente invasiva

Con el objetivo de abordar los diferentes problemas abiertos en la tecnología de MCG mínimamente invasiva previamente mencionados en el Capítulo 1, diversas técnicas, con varios grados de sofisticación, se han propuesto en la literatura, las cuáles se relacionan a continuación. Debido a que la investigación realizada aquí se centra en dispositivos para MCG SCs mínimamente invasivos, no se relacionan las contribuciones realizadas en SMCG invasivos o no invasivos.

Es importante recordar que el interés principal del trabajo de investigación de esta tesis doctoral se enfoca sólo en dos de las líneas de investigación abiertas, las que representan la mayor debilidad de los SMCGs: **precisión de los algoritmos de calibración** y **fiabilidad de los niveles de glucosa estimados**. Por lo tanto, aunque se hará mención a las principales contribuciones en los demás desafíos existentes, se enfatiza en las dos líneas de investigación indicadas previamente.

Algunas revisiones literarias muy completas sobre estos problemas abiertos se pueden encontrar en (Cengiz and Tamborlane, 2009; Rossetti et al., 2010; Bequette, 2010; Sparacino et al., 2010).

³¹Imagen tomada de (MED, 2013b).

2.3.1. Dinámica entre BG e IG

El intercambio de glucosa del plasma al ISF (y viceversa) a través de las paredes de los capilares se produce por difusión simple a través de un gradiente de concentración de glucosa (Figura 2.9).

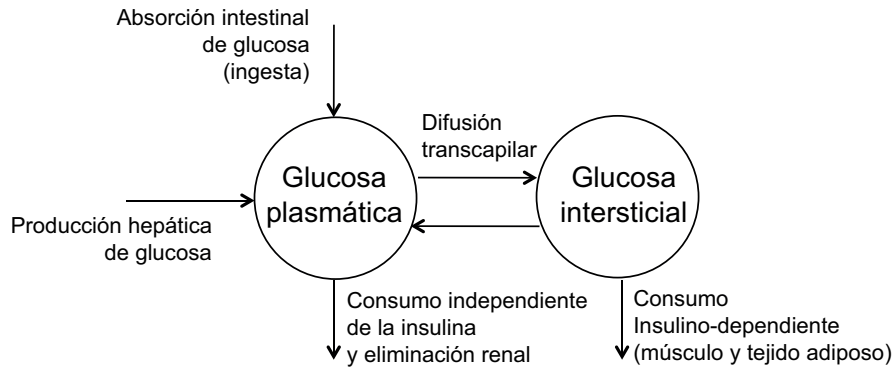


Figura 2.9: Transporte de glucosa plasma-ISF y procesos fisiológicos implicados.³²

En condiciones dinámicas, por ejemplo, después de una comida, IG y BG pueden ser particularmente diferentes, debido a la existencia de una cinética BG-IG la cuál, en la literatura, ha sido descrita por un modelo de dos compartimentos, que la aproxima como una ecuación diferencial de primer orden con la forma (Rebrin et al., 1999):

$$\frac{dIG}{dt} = -\frac{1}{\tau} IG(t) + \frac{g}{\tau} BG(t) \quad (2.1)$$

donde $BG(t)$ e $IG(t)$ representan las perturbaciones de BG e IG, respectivamente, en condiciones basales, g representa la ganancia estática del sistema BG-IG (se puede considerar que en estado estacionario, las concentraciones de glucosa en plasma y en el ISF son iguales, por tanto se considera igual a uno) y τ es una constante de tiempo (que puede variar entre los individuos).

Sin embargo, este proceso no es instantáneo y parece estar influenciado por el flujo sanguíneo y la permeabilidad capilar (Crone, 1984; Renkin, 1984). La consecuencia del intercambio transcapilar limitado de glucosa es que cualquier cambio en la concentración de BG, así como en la tasa metabólica y en la absorción de glucosa por las células, introduce un retraso fisiológico entre BG-IG (Sternberg et al., 1996; Monsod et al., 2002). La presencia de tal retraso está ampliamente difundida y reconocida (Rossetti et al., 2010). La mayoría de investigadores en este campo establecen el tiempo del retraso dentro del rango de 5 a 15 minutos (Stout et al., 2001; Kulcu et al., 2003; Boyne et al., 2003; Wentholt et al., 2007; Bailey et al., 2009; Garg et al., 2009; Kamath et al., 2009; Kovatchev et al., 2009b). Sin embargo, se ha observado mucha variabilidad, con tiempos de retardo desde 0 hasta 45 minutos e incluso adelantos (Steil et al., 2003; Sternberg et al., 1996; Monsod et al., 2002). Por esta razón, sigue siendo necesario un mejor conocimiento de la relación BG-IG, especialmente en condiciones cinéticas, lo que se traduciría en una mejora significativa de la exactitud de los SMCGs (King et al., 2007; Leal et al., 2010a).

³²Imagen tomada de (Bondia et al., 2010).

2.3.2. Calibración y retardo del sensor

Un punto importante en relación con la tecnología de MCG mínimamente invasiva es que los sensores amperométricos se colocan en el tejido SC y por lo tanto, la señal de salida del sensor es proporcional al nivel de IG, y no a la concentración de BG. Es aquí donde aparece el concepto de *calibración*. Para la mayoría de los SMCGs, la *calibración* significa transformación de una señal eléctrica generada por el sensor de glucosa (una corriente, típicamente en nA) en una señal de BG, utilizando una o más muestras de sangre capilar medidas con un SMBG convencional (en mmol/L o mg/dL) (Lodwig et al., 2003). El algoritmo que se utiliza para calibración de los SMCGs se conoce como *algoritmo de estimación de BG* o *algoritmo de calibración*. Un factor determinante en la precisión de los SMCGs es la calidad de los algoritmos de calibración (King et al., 2007). A continuación se relacionan algunas de las estrategias encontradas en la literatura para abordar este problema.

- **Técnicas lineales simples** (Bequette, 2010). La ecuación general simple usada para calcular la salida es:

$$y = m x + b \tag{2.2}$$

donde x es la variable de entrada (corriente del sensor) e y es la variable de salida (BG estimada). Los parámetros m (sensibilidad del sensor o pendiente de la ecuación lineal) y b (*Offset*) se estiman en función del número de puntos de calibración.

- **Calibración a partir de un punto.** Si se supone que b es conocido (usualmente $b = 0$), luego, usando sólo un par de datos x/y , tomados en un cierto instante de tiempo, se calcula la pendiente m :

$$m = \frac{(y_1 - b)}{x_1} \tag{2.3}$$

donde x_1 es la corriente del sensor en el instante de tiempo en que se toma el punto de calibración, e y_1 es la concentración de glucosa capilar, venosa o arterial utilizada para calibrar el dispositivo.

Para futuros instantes (k), con m y b ya calculados, el nivel de glucosa (\hat{y}_k) se calcula a partir de las muestras de corriente del sensor (x_k).

- **Calibración a partir de dos puntos.** Si se supone que b es desconocido, se emplean dos pares de datos x_1/y_1 y x_2/y_2 :

$$\begin{aligned} y_1 &= m x_1 + b \\ y_2 &= m x_2 + b \end{aligned} \tag{2.4}$$

siendo x_1 y x_2 la corriente del sensor en los instantes de tiempo en que se toman el primer y segundo punto de calibración, y siendo y_1 e y_2 las concentraciones de glucosa capilar, venosa o arterial utilizadas en cada uno de los dos puntos de calibración. Luego, m y b se estiman desde:

$$\begin{aligned} m &= \frac{(y_2 - y_1)}{(x_2 - x_1)} \\ b &= y_2 - m x_2 \end{aligned} \tag{2.5}$$

- **Calibración a partir de múltiples puntos.** Si existen múltiples puntos de calibración disponibles, se utiliza una regresión lineal para ajustar m y b a los datos. Las técnicas de regresión lineal estándar encuentran m y b que minimizan la suma de los cuadrados de los errores e_k^2 (diferencias entre medidas y_k y predicciones del modelo $\hat{y}_k = m x_k + b, k = 1, 2, \dots, N$):

$$\min_{m,b} \sum_{k=1}^N e_k^2 = \min_{m,b} \sum_{k=1}^N (y_k - (m x_k + b))^2 \quad (2.6)$$

En algunos estudios se ha evaluado la influencia de la cantidad, la exactitud, y el instante de tiempo en que se toman las muestras utilizadas para calibración, así como la tendencia de la concentración de glucosa en los momentos en que se toman estas muestras (Panteleon et al., 2003; Choleau et al., 2002a). También se han hecho cambios en la variable independiente de la regresión lineal y se concluyó que se obtienen mejores resultados en la calibración del sensor cuando la señal de corriente sin procesar o “*en bruto*” se utiliza como la variable independiente en la regresión lineal (Panteleon et al., 2003).

- **Técnicas de re-calibración** (Sparacino et al., 2010). En algunas ocasiones, los perfiles que proporcionan los SMCGs se ven afectados por un cambio del comportamiento del sensor después de su calibración inicial. Por esta razón, es deseable una *re-calibración* en RT de los datos del SMCG. Por *re-calibración*, se entiende un proceso donde la salida del sensor (en mg/dL) se procesa por medio de un algoritmo (externo al dispositivo) con el fin de mejorar su precisión.

Diferentes estudios en la literatura han tratado de hacer frente al problema de re-calibración, la mayoría de ellos retrospectivamente (Direcnet et al., 2006). En (Direcnet et al., 2006), se analiza la mejora en la precisión del SMCG retrospectivamente, modificando el número y la frecuencia de los puntos de calibración. Los resultados ponen en evidencia que la frecuencia con la cuál se toman los puntos de calibración es incluso más importante que la cantidad de ellos. En particular, la calibraciones realizadas durante los períodos de relativa estabilidad de la glucosa mejoran la precisión.

Por otra parte, (King et al., 2007) presentó un procedimiento de re-calibración pensado para una aplicación retrospectiva. Este procedimiento se basa en la siguiente ecuación:

$$y = a x + b \quad (2.7)$$

donde y representa la glucosa de referencia, x la glucosa estimada por el SMCG y a y b son los parámetros de re-calibración, los cuáles se determinan por ajuste, utilizando un par de datos de x/y recogidos simultáneamente.

Por último, en (Kuure-Kinsey et al., 2006) se presentó una estrategia de re-calibración más sofisticada utilizando un filtro de Kalman dual (-del inglés *dual-rate Kalman Filter*-), el cuál usa tanto la información de las medidas continuas del sensor como la de las medidas discontinuas del SMBG para estimar un factor de corrección para la ganancia del sensor en RT. Sin embargo, no tiene en cuenta la cinética entre BG-IG.

- **Técnicas que consideran la cinética entre BG-IG.**

Ninguno de los anteriores algoritmos de re-calibración considera explícitamente la distorsión introducida por la cinética entre BG-IG. Las estrategias que se mencionan a continuación, han sido propuestas con ese objetivo:

- Técnicas de estimaciones óptimas. La primera contribución que considera las cinéticas de la glucosa entre compartimentos corresponde a (Knobbe and Buckingham, 2005). El modelo de las cinéticas entre BG-IG fue considerado explícitamente con el fin de reconstruir los niveles de BG a partir de mediciones de un SMCG. Este trabajo utiliza un filtro de Kalman extendido (-del inglés *Extended Kalman Filter (EKF)*-) de cinco estados para estimar: 1) concentración de IG, 2) concentración de BG, 3) retardo fisiológico entre BG-IG, 4) tasa de cambio del nivel de BG y 5) factor de escala del sensor de glucosa SC. Estos estados son representados por: G_B , G_S , τ , $\frac{dG_S}{dt}$, k_S , respectivamente. El modelo del proceso usado se basa en el trabajo de (Rebrin et al., 1999):

$$\frac{dG_S}{dt} = -\frac{1}{\tau} G_S + \frac{1}{\tau} G_B \quad (2.8)$$

En este desarrollo, se asume una medida periódica del nivel de glucosa subcutánea del sensor (I_S), definida como:

$$I_S = -k_S G_S + b_S + \epsilon_S \quad (2.9)$$

donde b_S y ϵ_S representan el sesgo de la medida del sensor y las variaciones aleatorias del error del sensor, respectivamente.

El rendimiento de este modelo fue probado con datos de cuatro pacientes con DM, demostrando el potencial de esta metodología para mejorar la precisión de los SMCGs.

En (Facchinetti et al., 2010a), se desarrolló aún más la idea de (Knobbe and Buckingham, 2005) con el fin de hacer frente a la re-calibración. En este trabajo se incluye la relación dinámica entre BG-IG (ecuación (2.8)) y se propone un método de calibración bayesiano mejorado (-del inglés *enhanced Bayesian calibration method (BCM)*-) basado en un EKF de seis estados: 1) concentración de BG en el instante actual, 2) concentración de BG en el instante anterior, 3) concentración de IG, 4) ganancia del sensor de glucosa SC en el instante actual, 5) ganancia del sensor de glucosa SC en el instante anterior y 6) ganancia del sensor de glucosa SC en los dos instantes anteriores. Estos estados son representados por: $G_{B,k}$, $G_{B,k-1}$, $G_{S,k}$, α_k , α_{k-1} , α_{k-2} , respectivamente.

Una característica importante de este método es que proporciona una ecuación para definir las desviaciones de la ganancia del sensor de glucosa SC (α) en el tiempo (desde un punto de vista teórico se parece al k_S definido en (Knobbe and Buckingham, 2005)):

$$\alpha_{k+1} = 3\alpha_k - 3\alpha_{k-1} + \alpha_{k-2} + w_{\alpha,k} \quad (2.10)$$

donde $w_{\alpha,k}$ representa un ruido blanco de media cero con varianza λ_α^2 .

Las salidas del sistema son dos: la concentración de glucosa de referencia medida por el SMBG (G_B) y la concentración de glucosa estimada por el SMCG (G_M). Estas salidas se definen como:

$$G_{B,k+1} = 2G_{B,k} - G_{B,k-1} + w_{G,k} \quad (2.11)$$

donde $w_{G,k}$ representa un ruido blanco de media cero con varianza λ_G^2 .

$$G_{M,k} = (1 + \alpha_k)G_{S,k} + w_{\alpha,k} \quad (2.12)$$

Aunque este método se basa en un EKF de seis estados, es más simple que el método propuesto en (Knobbe and Buckingham, 2005), porque las variables desconocidas son sólo tres (G_B , G_S y α) y tan sólo dos varianzas (λ_α^2 y λ_G^2).

Este método puede ser utilizado en cascada a cualquier algoritmo de calibración construido para SMCGs comerciales, mejorando la salida del monitor con el fin de mejorar su precisión. Su validación se realizó con datos simulados representativos de 10 pacientes diabéticos durante siete días, teniendo en cuenta algunas consideraciones para tratar de hacer estos datos realistas (Kovatchev et al., 2009a). Los resultados mostraron una mejor precisión que la reportada con el método de (Knobbe and Buckingham, 2005). Sin embargo, la aplicación de este método a datos reales tiene limitaciones importantes (Rossetti et al., 2010). Para abordar esta dificultades, los mismos autores han propuesto recientemente algunas estrategias que incorporan procesos de desconvolución (Guerra et al., 2012).

- Técnicas basadas en modelos locales. En un trabajo reciente (Barceló-Rico et al., 2012) desarrolló un nuevo algoritmo de calibración que describe la relación entre las concentraciones de BG y la señal de corriente del sensor SC, el cuál se basa en un conjunto de modelos locales (-del inglés *Local Models (LM)*-) ponderados que tiene la ventaja de describir comportamientos locales cuando estos se presentan. Básicamente, este algoritmo se fundamenta en la integración de varios LMs, cada uno caracterizado por su independencia y validez regional, con el fin de obtener un modelo global interpretable (-del inglés *Global Model (GM)*-) definido como:

$$GM = \sum_{i=1}^c WLM_i \quad (2.13)$$

donde WLM_i es el i -ésimo LM ponderado. Cada WLM_i es el resultado del LM LM_i multiplicado por su función de validez V_i :

$$WLM_i = V_i LM_i \quad (2.14)$$

La estructura de cada LM podría ser cualquiera, pero en este algoritmo de calibración se utilizan modelos lineales para adicionar interpretabilidad al GM:

$$LM_i = \beta_{1i} x_1 + \beta_{2i} x_2 + \cdots + \beta_{Di} x_D + \beta_0 = \beta_{1Di} x_{1D} + \beta_{0i} \quad (2.15)$$

donde x_{1D} es el vector D -dimensional de entradas del sistema y $[\beta_{1Di}, \beta_{0i}]$ es el vector $D+1$ -dimensional de los coeficientes de regresión β de cada LM lineal con β_0 como el término independiente. Como se representa en la ecuación (2.14), cada LM se pondera

por algunas funciones de validez (V_i) que, en el algoritmo propuesto, se seleccionan como funciones hiper-Gaussianas. Para la implementación de este algoritmo se requiere normalizar adecuadamente las señales de entrada. En este caso se utilizaron parámetros de normalización poblacionales.

Finalmente, en un trabajo muy reciente de los mismos autores (Barceló-Rico et al., 2013), se propuso un algoritmo de calibración adaptativo que estima nuevos parámetros de normalización para el paciente en RT, incluyendo información específica de la sensibilidad paciente-sensor de glucosa en el modelo poblacional. La estimación de los parámetros se realiza a través de una minimización de la media de la diferencia absoluta relativa (-del inglés *Relative Absolute Difference (RAD)*-) en los puntos de calibración (glucosa capilar) usando una estrategia de ventana temporal deslizante. Este método se ha validado con datos de pacientes reales tomados con un SMCG retrospectivo y con datos simulados.

Otro factor que tampoco se debe obviar es el retardo intrínseco del sensor, que introduce retardos en su respuesta a los cambios en las concentraciones de glucosa (Davey et al., 2010).

2.3.3. Eliminación de ruido

Además de los errores de calibración, la señal de glucosa que proporciona el SMCG también se ve afectada por una componente de ruido aleatorio. Entre las soluciones propuestas para abordar el problema del filtrado se encuentran:

- Filtros de media móvil (-del inglés *Moving Average (MA)*-). Evidencias inferidas de la literatura de patentes indican que técnicas de pre-filtrado no lineal y filtros MA se utilizan muy a menudo para eliminar el ruido en los SMCGs (Poscia et al., 2003; Feldman and McGarraugh, 2007.; Simpson et al., 2008.; Keenan et al., 2009). Las patentes de Medtronic indican que las señales sin procesar se obtienen a intervalos de 10 segundos (Mastrototaro et al., 2002; Shin et al., 2005). Sin embargo, como se ha indicado en la sección 2.2.2 se remueven los valores más bajos y más altos y se promedian los demás valores para obtener la señal muestreada cada cinco minutos. Remover los valores más bajos y más altos tiene un efecto similar al que producen los filtros MA.
- Filtros de respuesta finita e infinita al impulso (-del inglés *Finite and Infinite Impulse Response Filters (FIR and IIR)*-). En (Pantoleon et al., 2003) se utilizó un FIR de séptimo orden para muestrear las señales a intervalos de un minuto. Una patente asignada a DexCom (Goode et al., 2005) sugiere el uso de filtros IIR para filtrado de señales sin procesar.
- Filtros basados en la teoría de estimación óptima. En varios trabajos (Knobbe and Buckingham, 2005; Palerm et al., 2005; Kuure-Kinsey et al., 2006; Chase et al., 2006; Facchinetti et al., 2010b) se pueden encontrar algunos métodos con enfoques más sofisticados que los filtros MA, los cuáles se pueden utilizar para eliminar el ruido en las señales de los SMCGs.

2.3.4. Predicción

Una aplicación de los SMCGs es la prevención de eventos hipo/hiperglucémicos. La importancia de estos métodos radica en que al indicar cuando va a suceder un evento hipoglucémico o

hiperglucémico, permiten prevenir excursiones glucémicas críticas. Entre las técnicas planteadas para hacer frente a este problema se tienen:

- Métodos de proyección. Estos métodos generan alertas cuando la tendencia actual del perfil de la concentración de glucosa sugiere que es probable que se produzca un evento hipoglucémico a corto plazo como en (Choleau et al., 2002b,c). Metodologías similares se han implementado en dispositivos comerciales con el fin de detectar oportunamente las tendencias peligrosas (Feldman and McGarraugh, 2007.).
- Modelos de entrada - salida. Estas estrategias generan alertas en base a predicciones de concentraciones de glucosa adelantadas en el tiempo, las cuáles se pueden calcular a partir de las concentraciones previas reportadas por un SMCG, utilizando modelos de series temporales. Esta posibilidad se propuso originalmente en (Bremer and Gough, 1999). Algunos trabajos usando modelos de series temporales lineales de bajo orden para predicción de la concentración de glucosa, específicos para un paciente, se han presentado en (Sparacino et al., 2007; Eren-Oruklu et al., 2009).

En los enfoques mencionados anteriormente, se utilizan modelos con estructura fija, mínima complejidad, pero con parámetros variables en el tiempo, los cuáles en cada instante de muestreo, se reajustan en base a la nueva muestra de glucosa recogida. Algunos trabajos han utilizado modelos de series temporales de mayor orden y por tanto, de mayor complejidad, específicos para un paciente como en (Reifman et al., 2007) o poblacionales como en (Gani et al., 2009, 2010). Finalmente, en (Lu et al., 2010), el mismo equipo concluyó que no se esperan mejoras en la predicción de glucosa a corto plazo si se incluye en los modelos de series temporales información sobre las comidas y las dosis de insulina.

- Técnicas de estimaciones óptimas. En (Palerm et al., 2005; Palerm and Bequette, 2007) se proponen enfoques para predicción de glucosa basados en el filtro de Kalman y en un modelo no paramétrico estocástico considerando tres diferentes horizontes de predicción.
- Otros métodos. Recientemente, modelos de redes neuronales artificiales (-del inglés *artificial neural network model (ANNM)*-) también se han aplicado para predecir concentraciones de glucosa, utilizando los datos de un SMCG (Pappada et al., 2008; Perez-Gandia et al., 2010; Zecchin et al., 2012).

2.3.5. Fiabilidad

Como se indica en (Hovorka, 2008; Bequette, 2010), un desafío aún mayor es la fiabilidad de los SMCGs actuales. En (Mazze et al., 2009), la fiabilidad medida durante el uso ambulatorio de estos dispositivos reportó 72,0% (DexComTM SEVEN) y 81,0% (Guardian[®] REAL-Time).

Para mejorar la fiabilidad, se deben integrar en los SMCGs capacidades de auto-monitorización que les permitan detectar sus propios fallos. Sin embargo, de acuerdo a nuestro conocimiento, no existen estudios a gran escala documentados en la literatura que aborden este problema. Por tanto, dado que las contribuciones realizadas hasta la fecha para detectar fallos en SMCGs son muy limitadas (Ward et al., 2004; Bondia et al., 2008; Shen et al., 2010; Facchinetti et al., 2011; Signal et al., 2012) y ante su importancia, es necesario realizar avances significativos en esta área (Bequette, 2010).

- Estrategias para detectar pérdidas de señal (conocidas en inglés como *dropouts*). En muchos de los perfiles glucémicos que suministran los SMCGs se observan *dropouts*. Un *dropout*

no es un único dato atípico, es una serie de puntos los cuáles se asemejan a un pico negativo. La característica más importante de un *dropout* es que abruptamente cambia a altas frecuencias y las concentraciones de BG suelen ser más bajas comparadas con las variaciones normales de glucosa. Además, los *dropouts* pueden incidir significativamente en la calibración de los SMCGs, dando lugar a problemas importantes en el cálculo de las concentraciones de BG en RT. Por tanto, es importante desarrollar estrategias que permitan su detección.

Para abordar específicamente este problema, la única contribución propuesta en la literatura corresponde a (Shen et al., 2010). En este trabajo se propone desarrollar un detector de *dropouts* en línea para SMCGs basándose en una transformada Wavelet discreta (-del inglés *Discrete Wavelet Transform (DWT)*-). Este método ha sido validado utilizando datos de cerdos no diabéticos y datos clínicos de pacientes.

- Estrategias para detección de medidas erróneas. Este tipo de estrategias posibilitarían el desarrollo de sistemas de auto-monitorización que podría ser embebidos en los SMCGs. A continuación se mencionan las contribuciones que se han realizado en este campo:
 - Una técnica consiste en utilizar estrategias basadas en medidas redundantes para detectar errores. En (Ward et al., 2004), se utilizó un arreglo de cuatro sensores de glucosa amperométricos implantables de manera subcutánea. Luego, se desarrolló un algoritmo basado en la mediana de las señales, para detectar cuando las señales de un sensor son inconsistentes con las señales de los otros sensores.

Este algoritmo se validó en ratas con y sin DM-1, las cuáles fueron sometidas a perturbaciones de sus niveles glucémicos administrando insulina o glucosa.

- Otra contribución se basa en la aplicación de métodos de clasificación desde el campo del aprendizaje estadístico. En (Bondia et al., 2008) se utilizaron SVMs estándar para detectar medidas de glucosa incorrectas realizadas por el CGMS[®] Gold[™], a partir de la información suministrada por el propio dispositivo y de acuerdo con un criterio de precisión clínico conocido como Clarke EGA (más detalles sobre este criterio en la sección 2.4.1). Este método se validó utilizando validación cruzada en pacientes con DM-1.

Sin embargo, en este trabajo no se tuvo en cuenta que en el problema de la detección de fallos en un SMCG existen muchos ejemplos de muestras correctas y pocos de muestras erróneas. Cuando se utilizan SVMs convencionales en conjuntos de datos significativamente desbalanceados, como en este contexto, estos modelos favorecen a la clase mayoritaria, exhibiendo un buen rendimiento en dicha clase, pero en perjuicio de la clase minoritaria. Esta es una limitación importante, puesto que la correcta detección de medidas erróneas realizadas por los SMCGs constituye el principal objetivo de este tipo de estrategias, especialmente si las medidas erróneas corresponden a situaciones en las cuáles el paciente muestra alta inestabilidad glucémica (eventos hipoglucémicos e hiperglucémicos). Otra limitación es que este método empleó un SMCG de análisis retrospectivo, por tanto es deseable un enfoque basado en un SMCG de análisis en RT.

- Por otra parte, una estrategia propuesta en (Facchinetti et al., 2011) se fundamentó en técnicas de estimaciones óptimas. En esta contribución se presenta un método para detectar en RT posibles fallos del sistema SMCG-bomba de CSII durante el periodo nocturno, utilizando simultáneamente datos de un SMCG y de una bomba de CSII. Este método se ha desarrollado en un contexto de espacio de estados, y se ha implementado a través de un filtro de Kalman. El método se probó en datos simulados y en datos de un paciente real con DM-1. Los resultados muestran que el método tiene posibilidades para generar correctamente alertas de fallos del sistema SMCG-bomba de CSII y anima a realizar una investigación más profunda en su desarrollo.
- Finalmente, una contribución muy reciente propuesta en (Signal et al., 2012), se sustenta en un modelo estocástico, el cuál utiliza un método de densidad basado en funciones núcleo (-del inglés *Kernel*-), y en datos históricos proporcionados por un SMCG retrospectivo, el CGMS[®] Gold[™]. El objetivo de este trabajo fue desarrollar una herramienta de ayuda a los médicos para identificar comportamientos inusuales en un SMCG (no excursiones glucémicas infrecuentes) y para destacar los datos suministrados por este dispositivo que potencialmente necesiten una interpretación más cuidadosa. Este método se validó en recién nacidos ingresados en una UCI, con resultados prometedores.

2.4. Criterios para evaluar la precisión y la fiabilidad de los SMCGs

2.4.1. Criterios para evaluar la precisión de los SMCGs

El rendimiento de los monitores de glucosa, tanto de los SMBGs como de los SMCGs, se puede clasificar basándose en la precisión analítica o en la precisión clínica (Wentholt et al., 2008; Clarke and Kovatchev, 2009; Krouwer and Cembrowski, 2010). La precisión analítica representa un enfoque cuantitativo para describir que tan cercano es el resultado de un método de medición que está siendo evaluado en comparación con una medida obtenida por un método de referencia (sangre capilar, arterial o venosa). La precisión clínica representa un enfoque cualitativo que describe como es el resultado clínico si se toman decisiones acerca del tratamiento a seguir basándose en el resultado de un método de medición que está siendo evaluado. En ambos casos, el método de medición que se evalúa corresponde a la medida proporcionada por un SMBG o un SMCG (Boren and Clarke, 2010).

Criterios analíticos para evaluar la precisión de los SMCGs

Generalmente, los requerimientos de las agencias reguladoras (por ejemplo la FDA) para demostrar la precisión analítica de los SMCGs han sido idénticos a los prescritos para los SMBGs. La precisión analítica se mide por medio de diferentes métricas estadísticas utilizando pares de datos obtenidos de las lecturas proporcionadas por un SMCG emparejadas con las determinaciones suministradas por un método de referencia (sangre capilar, arterial o venosa). Los criterios de uso común para describir la precisión analítica de los SMCGs se enuncian a continuación:

- Los coeficientes de correlación. Se han utilizado los coeficientes de correlación de Pearson y de Spearman. Sin embargo estos coeficientes han sido ampliamente criticados y conside-

rados inapropiados para evaluar la precisión de los SMCGs, debido a que están altamente influenciados por el rango y la distribución de los valores medidos (Wentholt et al., 2008).

- La media (DE) (-del inglés *Mean (Standard Deviation (SD))*-) o la mediana (IQR) (-del inglés *Median (Interquartile Range (IQR))*-) de la diferencia absoluta relativa (-del inglés *Relative Absolute Difference (RAD)*-) también conocidas como *Mean RAD (SD)* o *Median RAD (IQR)* (Clarke and Kovatchev, 2007; Wentholt et al., 2008; Clarke and Kovatchev, 2009). *RAD* es una métrica porcentual que indica cuanto se desvía una estimación de glucosa realizada por un SMCG del método de referencia. *RAD* se calcula como el valor absoluto de la diferencia entre el nivel de glucosa obtenido por un método de referencia y el estimado por un SMCG, expresado como un porcentaje del valor de referencia (ecuación (2.16)).

En general, cuando se trabaja con datos que se obtienen de experimentos médicos, estos presentan valores anómalos que hacen que la media no sea representativa. Por esta razón, debe utilizarse la mediana, al ser más representativa que la media en este tipo de datos. En los SMCGs, de acuerdo a recomendaciones de (Wentholt et al., 2008), se sugiere utilizar el valor de *Median RAD (IQR)*. El valor de *Median RAD (IQR)* resulta inferior al valor de *Mean RAD (SD)*. Sin embargo, es importante mencionar que en el rango hipoglucémico (definido según el criterio médico como valores de glucosa de referencia ≤ 70 mg/dL o en otros casos ≤ 80 mg/dL) es de uso común calcular la media (DE) o la mediana (IQR) de la diferencia absoluta (-del inglés *Absolute Difference (AD)*-) (ecuación (2.17)), en lugar de calcular *RAD* (Mastrototaro et al., 2008).

$$RAD (\%) = \left| \frac{\text{Glucosa de referencia} - \text{Glucosa estimada SMCG}}{\text{Glucosa de referencia}} \right| 100 \quad (2.16)$$

$$AD = |\text{Glucosa de referencia} - \text{Glucosa estimada SMCG}| \quad (2.17)$$

- Los criterios definidos en la norma ISO 15197, también conocidos como criterios ISO (-del inglés *International Organization for Standardization criteria (ISO criteria)*-) (ISO, 2003) (Figura 2.10). De acuerdo con estos criterios:
 - Para que una medida de un SMCG pueda ser considerada correcta:
 - Si la medida de BG considerada como referencia es ≤ 75 mg/dL, la medida de glucosa del SMCG sólo puede variar entre ± 15 mg/dL de la medida de referencia.
 - Si la medida de BG considerada como referencia es > 75 mg/dL, la medida de glucosa del SMCG sólo puede variar entre $\pm 20,0\%$ de la medida de referencia.
 - En caso tal que la medida del SMCG no cumpla estos criterios, se considera incorrecta.
- Los criterios de concordancia dentro de $\pm 20,0\%$ y $\pm 30,0\%$ de la medida de referencia. Como su nombre lo indica, estos criterios definen como correcta una medida de un SMCG si esta sólo puede variar entre $\pm 20,0\%$ o $\pm 30,0\%$ de la medida de referencia, en el caso de los criterios de concordancia dentro de $\pm 20,0\%$, o dentro de $\pm 30,0\%$, respectivamente. En caso tal que la medida del SMCG no cumpla estos criterios, se considera incorrecta (Mastrototaro et al., 2008).

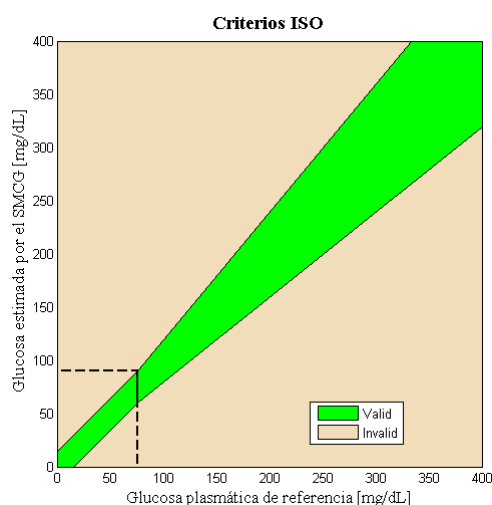


Figura 2.10: Gráfico que ilustra los criterios ISO.

- El método de Bland-Altman (Bland and Altman, 1986). Este método evalúa la concordancia entre dos métodos de medida, en este caso entre la medida de un SMCG y la de un método de referencia. Consiste en graficar las diferencias entre cada par de valores de los dos métodos de medida (eje y) contra la media de cada par de valores de ambos métodos (eje x). Si las medidas del SMCG son exactas y no existe error, los datos se distribuirán sobre la recta correspondiente a la diferencia cero entre medidas. En el gráfico se traza una línea horizontal para representar la media de las diferencias (sesgo) entre las muestras de BG y las lecturas del SMCG. Un aspecto muy importante de este método es que proporciona unos límites, conocidos como límites de concordancia, a partir del cálculo del intervalo de confianza para la diferencia de las dos mediciones. En el gráfico también se representan por medio de líneas horizontales los límites de concordancia, los cuáles se especifican como la media de las diferencias $\pm 1,96$ desviaciones estándar alrededor de la media de esas diferencias. Como es bien sabido, el intervalo de 1,96 desviaciones estándar alrededor de la media de las diferencias incluye el 95,0 % de las diferencias observadas. El valor 1,96 se corresponde con el percentil 95,0 % del modelo normal tipificado.

En este trabajo de investigación se proponen diferentes enfoques metodológicos con el objetivo de mejorar la precisión de los SMCGs. Para evaluar el rendimiento de las metodologías propuestas se utilizarán los siguientes criterios de precisión analítica: *Median RAD (IQR)*, los criterios ISO, los criterios de concordancia dentro de $\pm 20,0\%$ y $\pm 30,0\%$ de la medida de referencia y el método de Bland-Altman.

Criterios clínicos para evaluar la precisión de los SMCGs

La precisión clínica se mide comparando pares de datos, obtenidos de la misma forma como se describió en los criterios analíticos. Estos pares de datos se trazan por medio de puntos que se representan en una rejilla, conocida como rejilla de errores (Klonoff, 2012). En una rejilla de errores, cada punto dibujado (representando en dos dimensiones ambos valores, SMCG y referencia) se puede trazar dentro de una zona de riesgo. Este enfoque permite definir los conjuntos de datos con base en el porcentaje de puntos que caen dentro de cada zona o categoría clínica. Por tanto, se asigna a cada punto una zona de riesgo en función de si existe o no un efecto en

la acción clínica y si es así, como va a afectar el resultado clínico.

Dos tipos de rejillas de errores, diseñadas inicialmente para evaluar los SMBGs, se utilizan de manera frecuente para describir la precisión clínica de los SMCGs (Clarke et al., 1987; Parkes et al., 2000). Adicionalmente, otras rejillas de errores han sido propuestas específicamente para evaluar los SMCGs, tales como (Garg et al., 1999; Kovatchev et al., 2004). El método propuesto por (Kovatchev et al., 2004) combina la evaluación de la precisión puntual y la de la tasa y dirección de cambio de la glucosa. En cuanto a la propuesta de (Garg et al., 1999), esta podría producir una noción exagerada de la verdadera precisión de un SMCG (Kollman et al., 2005). Otras rejillas de errores para cuantificar el rendimiento clínico de los monitores de glucosa se han propuesto. Una de ellas permite ajustar la dosis de insulina utilizando pares de datos, trazados en zonas de riesgo, definidas a partir de como la severidad del error podría afectar la decisión de cuanta insulina administrar a un paciente hospitalario (Ellmerer et al., 2006).

De todas las rejillas de errores mencionadas anteriormente, las de uso común para describir la precisión clínica de los SMCGs, las cuáles son aceptadas por las agencias reguladoras se detallan a continuación:

- Rejilla de análisis de errores (-del inglés *Error Grid Analysis (EGA)*-) de Clarke, también conocida como Clarke EGA (Clarke et al., 1987). Esta rejilla divide el área de la gráfica en cinco zonas (A, B, C, D y E). El tamaño y la forma de cada zona dependen de la relevancia clínica de los puntos que allí se sitúen, siendo consideradas las zonas A y B aceptables desde el punto de vista clínico y las zonas C, D y E potencialmente peligrosas (Figura 2.11).

Los puntos que se localizan en la zona A son aquellas medidas del SMCG que difieren de la medida de referencia por no más de 20,0 %. Los puntos en la zona B son medidas del SMCG que difieren de la medida de referencia por más de 20,0 %, pero no dan lugar a alteraciones en el tratamiento, por lo que son considerados errores benignos. Los puntos en la zona C son medidas del SMCG que pueden ocasionar sobre-correcciones de valores de glucosa aceptables, lo que daría lugar a correcciones innecesarias que podrían ocasionar hipos o hiperglucemias. Los puntos en la zona D se consideran fallos peligrosos que se deberían detectar, los cuáles darían lugar a errores en el tratamiento, y finalmente, los puntos en la zona E se traducen en decisiones de tratamiento opuestas. Cuando una cantidad superior al 95,0 % de los puntos se trazan en las zonas A y B, generalmente un sistema es considerado preciso (Clarke, 2005).

- Rejilla de análisis de errores consensuada (-del inglés *Consensus Error Grid Analysis (Consensus EGA)*-) (Parkes et al., 2000). Fue desarrollada a partir de una encuesta realizada a 100 médicos quienes tratan pacientes diabéticos. Esta rejilla mantiene el formato de cinco zonas de riesgo de Clarke EGA pero esta métrica por consenso utilizó un panel experto para trazar los límites exactos entre las zonas (Figura 2.12). Consensus EGA fue diseñada para eliminar varias discontinuidades que se presentan en Clarke EGA, donde un cambio infinitesimal en los niveles de glucosa sanguínea causa que la zona de riesgo se incremente en dos o incluso tres niveles de riesgo (Klonoff, 2012).
- Rejilla de análisis de errores para glucosa continua (-del inglés *Continuous Glucose Error Grid Analysis (CG-EGA)*-) (Kovatchev et al., 2004). Se diferencia de la tradicional Clarke EGA en el ajuste dinámico de las zonas de la rejilla de errores, en función de la tasa y dirección de cambio momentánea de la medida de referencia. Este ajuste dinámico está diseñado para acondicionarse a un posible retardo fisiológico entre la referencia y las lecturas

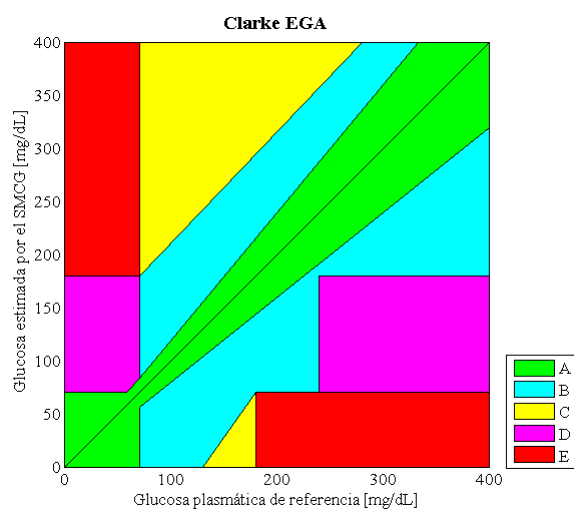


Figura 2.11: Rejilla Clarke EGA.

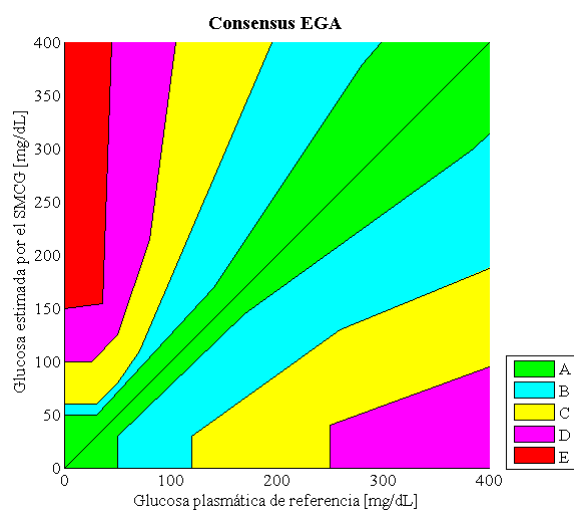


Figura 2.12: Rejilla Consensus EGA.

del SMCG.

CG-EGA incluye dos componentes que interactúan: la rejilla de análisis de errores puntuales (-del inglés *Point-Error Grid Analysis* o *P-EGA*-), la cuál evalúa la precisión de las lecturas puntuales de un SMCG respecto a una medida de referencia, y la rejilla de análisis de errores en la tasa de cambio (-del inglés *Rate-Error Grid Analysis* o *R-EGA*-), la cuál evalúa la capacidad de un SMCG para capturar la tasa y dirección de cambio de las fluctuaciones de la glucosa de referencia (Figura 2.13).

P-EGA (Figura 2.13(a)) es similar a la rejilla original de Clarke, excepto que permite un desplazamiento de los límites superior e inferior de las zonas A, B y D, en proporción a la tasa y dirección de cambio de la glucosa y del retardo fisiológico entre IG y BG. El retardo fisiológico asumido por P-EGA es un retardo constante igual a siete minutos.

R-EGA (Figura 2.13(b)) computa para cada uno de los pares de datos la tasa y dirección de cambio, y categoriza los errores en la dirección y la tasa de cambio en cinco zonas de riesgo (zonas A - E), las cuáles son clínicamente similares a las definidas en Clarke EGA.

Finalmente, CG-EGA computa la precisión clínica de un SMCG combinando R-EGA y P-EGA en tres regiones clínicamente relevantes: hipoglucémica definida como $BG \leq 70$ mg/dL, normoglucémica como $70 \text{ mg/dL} < BG \leq 180 \text{ mg/dL}$ e hiperglucémica como $BG > 180$ mg/dL.

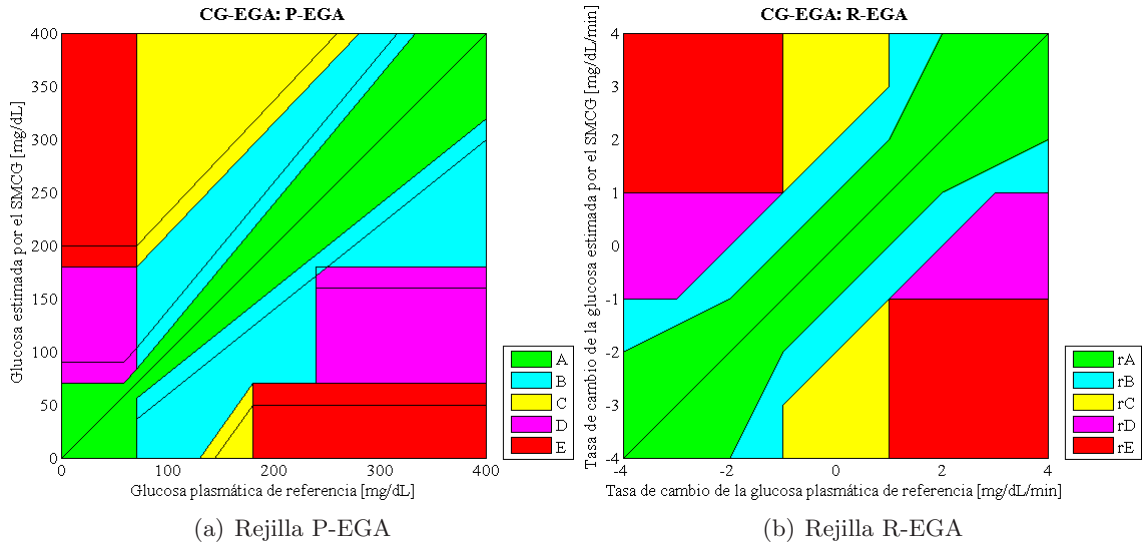


Figura 2.13: Rejilla CG-EGA.

Para evaluar el rendimiento de los enfoques metodológicos propuestos en esta investigación se utilizarán los criterios de precisión clínica Clarke EGA y Consensus EGA. CG-EGA no se aplicará debido a que requiere de numerosas medidas de referencia, las cuáles deben tener un tiempo de muestreo muy frecuente, una característica que no poseen los datos usados como medidas de referencia en esta investigación, debido a limitaciones de tipo hospitalario.

2.4.2. Métricas para evaluar la fiabilidad de los SMCGs

Con el objetivo de evaluar el rendimiento de las diferentes metodologías propuestas para detectar si las concentraciones de glucosa estimadas por un SMCG son correctas o incorrectas, en el aprendizaje supervisado se consideran frecuentemente algunas métricas con base en la matriz de confusión (Tabla 2.4). A continuación se definen estas métricas, algunas de las cuáles serán consideradas en este trabajo de investigación.

La *proporción de aciertos* de un clasificador $f(x)$, $f : \mathcal{X} \rightarrow \mathcal{Y}$, en un conjunto de test $\mathcal{D} \subset \mathcal{X} \times \mathcal{Y}$ con $\mathcal{Y} = \{+1, -1\}$, siendo $\mathcal{Y} = \{-1\}$ la clase negativa o clase correcta y $\mathcal{Y} = \{+1\}$ la clase positiva o clase incorrecta, se define como:

$$\text{Proporción de aciertos } (f, \mathcal{D}) = \frac{t_{pos} + t_{neg}}{N_{pos} + N_{neg}} \quad (2.18)$$

		Clase real	
		<i>Positiva</i>	<i>Negativa</i>
Clase predicha	<i>Positiva</i>	Verdaderos Positivos (t_{pos})	Falsos Positivos (f_{pos})
	<i>Negativa</i>	Falsos Negativos (f_{neg})	Verdaderos Negativos (t_{neg})
Número de casos		Número de Positivos (N_{pos})	Número de Negativos (N_{neg})

Tabla 2.4: Matriz de confusión para un problema de clasificación binario.

Es decir, la *proporción de aciertos* es una métrica que computa la proporción de casos que son correctamente clasificados por un modelo (ecuación (2.18)), donde, t_{pos} es el número de casos de la clase positiva correctamente clasificados y t_{neg} es el número de casos de la clase negativa correctamente clasificados. N_{pos} y N_{neg} corresponden al total de casos de la clase positiva y la clase negativa, respectivamente. En esta métrica, las dos clases (positiva y negativa) tienen la misma prioridad para el propósito de clasificación. Por lo tanto, esta métrica es inapropiada cuando difieren bastante las probabilidades a priori de las clases, debido a que no tiene en cuenta los costos en las clasificaciones erróneas y es muy sensible al sesgo entre las clases (He and Garcia, 2009).

Para evitar esta situación, en problemas de aprendizaje no balanceado se consideran otras métricas de evaluación que permitan realizar evaluaciones más completas de este tipo de problemas y medir el rendimiento sobre cada una de las clases de manera independiente. Estas métricas son la *precisión*, la *especificidad*, la *sensibilidad*, el *Valor-F* y la G_{media} .

La *precisión* o valor predictivo positivo, definida en la ecuación (2.19), es una medida de la exactitud que computa, de todos los casos etiquetados como positivos por el clasificador, cuantos son etiquetados correctamente (He and Garcia, 2009).

$$Precisión(f, \mathcal{D}) = \frac{t_{pos}}{t_{pos} + f_{pos}} \quad (2.19)$$

Además, a partir de la matriz de confusión, también se definen la *especificidad* (o tasa de verdaderos negativos) y la *sensibilidad* (o tasa de verdaderos positivos). La *especificidad*, definida en la ecuación (2.20), es una medida de exactitud negativa que determina la proporción de casos negativos correctamente clasificados. En contraposición, la *sensibilidad*, definida en la ecuación (2.21), es una medida de la exactitud positiva que calcula la proporción de casos positivos correctamente clasificados.

$$Especificidad(f, \mathcal{D}) = \frac{t_{neg}}{N_{neg}} \quad (2.20)$$

$$Sensibilidad(f, \mathcal{D}) = \frac{t_{pos}}{N_{pos}} \quad (2.21)$$

Por otra parte, el *Valor-F* combina la *precisión* y la *sensibilidad* como una medida de la eficacia de la clasificación (ecuación (2.22)). Esta métrica se puede interpretar como una media ponderada de la *precisión* y la *sensibilidad*, determinada por el coeficiente β establecido por el usuario. β es un coeficiente para ajustar la importancia relativa de la *precisión* versus la *sensibilidad* (generalmente, $\beta = 1$) (He and Garcia, 2009).

$$Valor-F(f, \mathcal{D}) = \frac{(1 + \beta)^2 \cdot Sensibilidad(f, \mathcal{D}) \cdot Precisión(f, \mathcal{D})}{\beta^2 \cdot Sensibilidad(f, \mathcal{D}) + Precisión(f, \mathcal{D})} \quad (2.22)$$

Asimismo, una métrica que es más sensible a un bajo número de casos positivos y que combina una evaluación del rendimiento en términos de la *especificidad* y la *sensibilidad* es la G_{media} (ecuación (2.23)). Esta métrica se define como la media geométrica entre la *especificidad* y la *sensibilidad* y en ella las dos clases, la clase positiva y la clase negativa, tienen la misma relevancia para propósitos de clasificación:

$$G_{media}(f, \mathcal{D}) = \sqrt{\frac{t_{pos} \cdot t_{neg}}{N_{pos} \cdot N_{neg}}} = \sqrt{Sensibilidad(f, \mathcal{D}) \cdot Especificidad(f, \mathcal{D})} \quad (2.23)$$

Finalmente, el análisis de la curva característica operativa del receptor (-del inglés *Receiver Operating Characteristic* (ROC)-) es una representación gráfica que muestra las relaciones entre la *sensibilidad* (eje y) y $1 - especificidad$ (eje x). Por tanto, cualquier punto en el espacio de una curva ROC representa el rendimiento global de un clasificador en una distribución dada. Este gráfico es útil porque proporciona una representación visual de comparación entre diferentes clasificadores (He and Garcia, 2009). Además, esta técnica permite evaluar el rendimiento de diferentes clasificadores calculando el área bajo la curva ROC (-del inglés *Area Under the Curve* (AUC)), la cuál se puede determinar a partir de la ecuación (2.24). El AUC se define para todo el conjunto de pares sensibilidad-especificidad en un clasificador. Esta métrica puede tomar valores entre 0 y 1, donde a mayor valor, mejor rendimiento de clasificación.

$$AUC(f, \mathcal{D}) = \frac{Sensibilidad(f, \mathcal{D}) + Especificidad(f, \mathcal{D})}{2} \quad (2.24)$$

En este trabajo de investigación se proponen diferentes enfoques metodológicos con el objetivo de mejorar la fiabilidad de los SMCGs. Para evaluar el rendimiento de las metodologías propuestas se utilizarán la *proporción de aciertos*, la *especificidad*, la *sensibilidad* y la G_{media} .

2.5. Limitaciones para avanzar en esta área de investigación

Es muy importante mencionar las principales limitaciones que existen en esta área de investigación. La principal de todas está relacionada con la gran variabilidad inter-pacientes e intra-paciente. La causa de esta variabilidad se encuentra en que el cuerpo humano es un sistema complejo. Por tanto, la variabilidad inter-paciente se produce porque el metabolismo y los estados fisiológicos de cada persona son heterogéneos, en tanto que la variabilidad intra-paciente se genera porque las condiciones metabólicas pueden variar en la misma persona en diferentes horas del día, dependiendo de diversos factores.

Pues bien, esta variabilidad condiciona el desarrollo de algoritmos de estimación de BG y la implementación de estrategias para detectar fallos en los SMCGs que sean representativos de una población. Para lograr este carácter representativo se requiere utilizar en su desarrollo numerosos datos obtenidos de múltiples pacientes, con características heterogéneas. Este requerimiento es difícil de satisfacer, entre otras cosas por las razones que se exponen a continuación, las cuáles han ralentizado y limitado, en general, el desarrollo de este campo de investigación, y en particular, el trabajo realizado en esta tesis doctoral:

- *Obstáculos en el desarrollo de los ensayos clínicos.* En primer lugar, dado que se trata de ensayos con seres humanos, se deben satisfacer ciertos requerimientos de tipo ético. Por tanto, se debe no sólo diseñar el ensayo, sino también elaborar un protocolo clínico de

actuación, el cuál debe ser sometido a la aprobación de un Comité de Ética, con el fin de garantizar que se respetarán los principios éticos establecidos para experimentar con seres humanos. Además, es necesario proporcionar una adecuada formación específica al personal medico y de enfermería en el uso eficaz de los dispositivos involucrados en el ensayo, para evitar que errores debidos al “*factor humano*” interfieran en el correcta ejecución del mismo y en los productos que se generen a partir de la información allí recolectada.

- *El alto costo que implica la implementación de esta tecnología y la realización de ensayos clínicos.* Los SMCGs no tienen un bajo precio y aunque se obtengan los fondos para comprar estos dispositivos, se requiere adquirir constantemente sensores SCs que también tienen un costo. Por tanto, si no existen fondos suficientes para adquirir estos productos y llevar a cabo estos ensayos, los algoritmos se desarrollan y validan a partir de datos *in-silico*, los cuáles no reportan los mismos resultados ni permiten llegar a las mismas conclusiones que se obtienen con datos reales.
- *La no disponibilidad de datos públicos.* Aunque existe una base de datos pública denominada DirecNet (Dir, 2013), los datos allí almacenados difícilmente cumplen con las características requeridas para el desarrollo de algoritmos de calibración y detección de medidas erróneas, como por ejemplo la señal de corriente sin procesar que mide el sensor SC o frecuentes medidas de glucosa de referencia.
- *La escasa colaboración entre los grupos propietarios de datos.* Es una consecuencia de las razones mencionadas anteriormente. Debido a las dificultades y los altos costos que implican obtener datos clínicos, los diferentes grupos de investigación que poseen esta información se muestran renuentes a compartirlas con otros grupos, lo que se traduce en que cada grupo, generalmente, sólo posee pocos datos de sus propios ensayos clínicos, frenando con ello la posibilidad de llegar a resultados concluyentes.

Además, otra limitación importante de esta área de investigación, está relacionada con el hecho que los SMCGs disponibles en el mercado no ofrecen información detallada acerca de las técnicas utilizadas para calcular las concentraciones de BG.

2.6. Resumen

En este capítulo, se ha presentado una visión general de la DM, sus complicaciones y las cifras relacionadas con la enfermedad. Los SMCGs permiten el desarrollo de nuevas estrategias para el tratamiento de la DM y el control glucémico en los pacientes críticos. Por tanto, este capítulo se ha enfocado en estudiar el estado de la técnica de MCG y en él se han discutido los principales avances propuestos en la literatura y algunas cuestiones abiertas. Se han considerado cinco problemas específicos: análisis y modelado de la cinética entre BG e IG, calibración para mejorar la precisión de las estimaciones continuas de glucosa, filtrado para eliminar el ruido de las señales, predicción de concentraciones de glucosa y estrategias para mejorar la fiabilidad en estos dispositivos. Dar solución a estos problemas es crucial para mejorar el desempeño de los SMCGs, lo que permitirá su integración de manera confiable en el AP, así como su aprobación como un sustituto de los SMBGs en la toma de decisiones terapéuticas.

Capítulo 3

Algoritmos de estimación de glucosa en plasma para SMCGs

En los Capítulos 1 y 2, se indicó que el problema de falta de precisión que aún presentan los *sistemas de monitorización continua de glucosa (SMCGs)* (-del inglés *Continuous Glucose Monitoring Systems (CGMSs)*-) se debe a diversos factores. Un factor determinante en la precisión de los SMCGs es la calidad de los *algoritmos de estimación de glucosa en plasma* (-del inglés *Blood Glucose (BG)*-), también conocidos como *algoritmos de calibración*. Estos algoritmos presentan varias deficiencias especialmente en la estimación de hipoglucemias, presentando problemas en la detección de las mismas con una alta tasa de falsos positivos y de falsos negativos (Clarke et al., 2005; Wentholt et al., 2005; Kovatchev et al., 2008; Rossetti et al., 2010; Bondia et al., 2010).

En ese contexto, este capítulo está orientado a proponer y describir dos nuevos algoritmos de estimación de BG para abordar los problemas de precisión presentados por los SMCGs: 1) un algoritmo retrospectivo basado en técnicas de regresión lineal, al cuál denominaremos *algoritmo retrospectivo propuesto para el SMCG (ARP-SMCG)*; y 2) el otro algoritmo que se propone considera la relación entre IG y BG en condiciones dinámicas con base en modelos autorregresivos (-del inglés *AutoRegressive (ARs)*-), al cuál denominaremos *algoritmo basado en modelos ARs para el SMCG (ABMAR-SMCG)*. En la investigación desarrollada en este capítulo, se ha utilizado como dispositivo de monitorización continua de glucosa (MCG) el CGMS[®] Gold[™]. Estudios clínicos previos concluyeron que el algoritmo de estimación de BG de este dispositivo, al que hemos denominado A-Gold[™], presenta varias deficiencias, las cuáles se observan especialmente en la estimación de hipoglucemias (Buckingham, 2003; Weinzimer et al., 2003; Direcnet et al., 2006).

El capítulo se organiza como sigue: la sección 3.1 describe el proceso de experimentación clínica usado para la recolección de los datos. En la sección 3.2, se introduce el A-Gold[™] y a continuación, se describe el algoritmo retrospectivo que se propone. Seguidamente, se presentan los resultados de la validación del ARP-SMCG, empleando criterios de precisión analíticos y clínicos. Además, se realiza una comparación entre los perfiles de glucemia obtenidos usando el A-Gold[™] y el ARP-SMCG. Por otra parte, en la sección 3.3, se describe el ABMAR-SMCG. En esta sección, se explica el entrenamiento de los modelos ARs y se describe el algoritmo de estimación de BG que se propone con base en estos modelos. De la misma forma que en la sección 3.2, aquí se presentan los resultados de la validación de este algoritmo, empleando criterios de precisión analíticos y clínicos. A continuación, se realiza una comparación entre los perfiles de glucemia obtenidos usando el A-Gold[™] y el ABMAR-SMCG. Finalmente, se presenta un

resumen del capítulo en la sección 3.4.

Un artículo que describe del ARP-SMCG, los resultados y las conclusiones de este trabajo se ha publicado en (Leal et al., 2010b). Un artículo que describe el ABMAR-SMCG, los resultados y las conclusiones obtenidas se ha publicado en (Leal et al., 2010a). Además, algunos resultados parciales del ABMAR-SMCG, se presentaron en el Noveno Congreso Internacional sobre Tecnologías para la Diabetes (Leal et al., 2009).

3.1. Estudio observacional y recolección de datos

3.1.1. Descripción de los pacientes y del entorno clínico del ensayo

En el desarrollo de los algoritmos de estimación de BG presentados en este capítulo se utilizan datos clínicos obtenidos en un estudio observacional. La población a estudio se define como enfermos con DM-1 de la Unidad de Diabetes, Endocrinología y Nutrición del Hospital Universitario de Girona Doctor Josep Trueta (en adelante Hospital Josep Trueta). Estos pacientes se seleccionaron para el estudio de acuerdo con los criterios para diabéticos tipo 1 definidos por la Asociación Americana de Diabetes (ADA). El periodo de reclutamiento fue de cuatro meses.

Los pacientes seleccionados debían cumplir todos los criterios de inclusión que se definen a continuación:

- Tener una edad entre 18 y 40 años.
- Presentar DM-1 mayor a un año de evolución y que reúnan los siguientes criterios:
 - Péptido C basal¹ menor a 0,5 ng/ml (realizado como mínimo durante los últimos seis meses).
 - Debut en cetoacidosis o anticuerpos antiGAD o antiIA2 positivos².
 - Índice de masa corporal (IMC) entre 18,5 y 27,5 kg/m².
- Presentar un control metabólico no óptimo, definido como HbA_{1c} mayor a 6,5%.
- Tener cumplimentada la hoja de consentimiento informado.

Los criterios de exclusión que se consideraron fueron:

- Gestación o lactancia.
- Mujeres en edad fértil que no sigan un método anticonceptivo adecuado.
- Tratamiento previo con fármacos que puedan cambiar el perfil glucémico.

¹Es un examen de sangre que mide la cantidad de péptido C, un subproducto que se crea cuando se produce la hormona insulina. Los valores normales deben estar entre 0,5 a 2,0 ng/mL (nanogramos por mililitro). Los valores del péptido C se basan en el nivel de glucosa en la sangre. El péptido C es una señal de que el cuerpo está produciendo insulina. Los valores bajos (o la ausencia de péptido C de insulina) indican que el páncreas no está produciendo o está produciendo poca insulina (MED, 2013a).

²Es un examen que se hace para determinar si el cuerpo ha producido anticuerpos contra la insulina. Este examen se puede realizar si una persona tiene o está en riesgo de padecer DM-1. Normalmente, no hay anticuerpos contra la insulina en la sangre (MED, 2013a).

- Historia de abuso de drogas o de alcohol.
- Enfermedad que pueda afectar la ingesta-absorción de alimentos.
- Insuficiencia renal, definida por una creatinina $\geq 2,0$ mg/dL.
- Estado mental del paciente que le impida comprender la naturaleza y/o el objetivo del estudio.
- Evidencia de actitud no cooperadora o imposibilidad de acudir a las visitas de seguimiento.

De acuerdo con la definición de la población de estudio anteriormente descrita y los criterios de inclusión y exclusión que se definieron previamente, se incluyen en total 21 pacientes (valor de HbA_{1c} (%), 8,0 (7,6 – 9,9)). Las características clínicas del grupo de pacientes se describen en la Tabla 3.1.

Variable	Valor
Número (%)	21 (100)
Edad (años)	27,0 (23,0 – 30,0)
Sexo (Femenino/%)	9/42,9
IMC (kg/m ²)	24,0 (22,0 – 26,0)
Presión sanguínea sistólica (mmHg)	125,0 (116,0 – 131,0)
Presión sanguínea diastólica (mmHg)	80,0 (62,0 – 71,0)
Colesterol total (mg/dL)	189,0 (151,0 – 213,0)
Triglicéridos (mg/dL)	87,0 (69,0 – 108,0)
HDL-C (mg/dL)	69,0 (54,0 – 72,2)
LDL-C (mg/dL)	101,0 (73,0 – 121,0)
Duración de la diabetes (años)	13,0 (10,0 – 17,8)
HbA _{1c} (%)	8,0 (7,6 – 9,9)

Los datos son expresados utilizando la mediana (IQR).

Los números entre paréntesis indican el rango intercuartílico (IQR) (Q1 - Q3, respectivamente).

Tabla 3.1: Características basales y condiciones primarias de los pacientes con DM-1 del Hospital Josep Trueta utilizados en el desarrollo del ARP-SMCG.

3.1.2. MCG en la Unidad de Endocrinología del Hospital Josep Trueta: protocolo de actuación MODAS

En el ensayo clínico que se describe en esta sección, se determinó la glucemia mediante el análisis de la concentración de glucosa en sangre venosa, a la que denominaremos VBG (medida en mg/dL), determinada en un laboratorio central usando el Glucose Analyzer II (Beckman Instruments, Brea, CA). Además, a cada enfermo incluido en el estudio se le determinó de manera continua (cada cinco minutos) la glucosa mediante el CGMS[®] Gold[™] (MiniMed referencia MMT-7102), a la que denominaremos G_{Gold} (medida en mg/dL). El sensor de glucosa del CGMS[®] Gold[™] se insertó en la región subcutánea abdominal de cada paciente, y se utilizó continuamente por un periodo de 72 horas (tres días, de los cuáles uno estuvo ingresado en el hospital y dos días en su domicilio). Después de un periodo de inicialización de dos horas, se introdujo la primera calibración. El CGMS[®] Gold[™] se calibró por medio de concentraciones de glucosa en capilar, a la que denominaremos $G_{Capillary}$ (medida en mg/dL), usando para su determinación un glucómetro convencional. Estas calibraciones se realizaron de acuerdo con las instrucciones del fabricante (MiniMed, 2003).

El protocolo de actuación que se siguió en el estudio clínico se denominó MODAS-0114. Los aspectos más relevantes de este protocolo se describen a continuación:

- Adquisición de datos hospitalarios en el Hospital Josep Trueta.
 - Determinación de los signos vitales y verificación de que el paciente estuviera en ayunas, así como el tipo y la dosis de insulina y el lugar de administración de la misma.
 - Canalización de una vía endovenosa en una vena antecubital, usando para ello una llave de tres pasos conectada a un suero fisiológico de mantenimiento (ritmo 500 ml cada 12 horas), para la extracción de muestras de sangre venosa.
 - Recolección de la primera muestra de sangre venosa a las ocho horas y 30 minutos. En esta muestra se determinaba un hemograma³, HbA_{1c} y concentraciones de glucosa, urea, creatinina, insulina y glucagón.
 - Inserción del sensor subcutáneo (SC) e inicialización del CGMS[®] Gold[™], para medir y registrar la glucemia cada cinco minutos. No se definió una hora exacta para el inicio de la monitorización continua.
 - Administración de las dosis habituales de insulina del paciente a las nueve horas, las 13 horas y las 19 horas, previamente a las ingestas de alimentos correspondientes al desayuno, la comida y la cena, respectivamente. La dosis usual de insulina, incluyendo el tipo, la cantidad y el lugar de administración, se mantuvo durante el estudio. Las tres ingestas estaban estandarizadas para el paciente con cuantificación de los macronutrientes. Además, la ingesta de comida se mantuvo con comidas normales, aunque la preparación de la comida fue supervisada por expertos en nutrición.
 - Extracciones de sangre venosa para la determinación de glucosa cada 15 minutos hasta transcurridas dos horas después de cada ingesta de alimentos, tras lo cual la frecuencia de extracción era de 30 minutos.
 - Última extracción de sangre a las 21 horas, tras lo cuál se retiraba el catéter endovenoso y el paciente abandonaba el hospital.

- Adquisición de datos domiciliarios en el domicilio del paciente.
 - Continuación de la MCG hasta completar 72 horas desde el instante en que se inserta el sensor SC de glucosa.
 - Seguimiento de una dieta cuantitativa y de hábitos rutinarios normales por parte del paciente.
 - Mantenimiento y control del CGMS[®] Gold[™] por parte del paciente, según las instrucciones dadas previamente por el personal médico del Hospital Josep Trueta.
 - Cálculo de la dosis de insulina mediante el procedimiento convencional.
 - Registro adecuado y oportuno de la dosis de insulina y del tiempo de inyección, así como del tiempo y composición de las ingestas en el diario del paciente.

- Descarga de los datos del CGMS[®] Gold[™] a un ordenador después de completadas 72 horas de monitorización continua, utilizando el software MiniMed Solutions Software versión 3.0C.

³Hematías, hemoglobina, hematocrito, leucocitos, recuento diferencial y plaquetas.

Previamente al inicio del estudio, se capacitó a los pacientes en el manejo del CGMS® Gold™ y se les formó en el cálculo del porcentaje de macronutrientes de la dieta. El protocolo clínico del estudio fue aprobado por el Comité de Ética del Hospital Josep Trueta. Todos los pacientes que participaron en el estudio fueron informados minuciosamente del mismo y decidieron participar de forma libre y voluntaria en él, firmando el correspondiente consentimiento informado.

3.2. Algoritmo para estimación de BG usando técnicas de regresión lineal: Algoritmo retrospectivo propuesto para el SMCG (ARP-SMCG)

3.2.1. Descripción del A-Gold™

El A-Gold™ estima la concentración de BG a partir de la señal eléctrica generada de la reacción electroquímica entre la glucosa intersticial (-del inglés *Interstitial Glucose (IG)*-) y GOx, descrita en la sección 2.2.2 (MiniMed, 2003; Mastrototaro et al., 2002). Este es un algoritmo retrospectivo, cuyo principio de funcionamiento se basa en una técnica de regresión lineal modificada, para estimar BG a partir de la corriente sin procesar o medida en “bruto” en el fluido intersticial (-del inglés *Interstitial Fluid (ISF)*-). Es un algoritmo retrospectivo porque después de recoger todos los datos correspondientes a 72 horas de funcionamiento del sensor, el algoritmo trata de ajustar la estimación de BG, minimizando el error absoluto relativo de esta medida con respecto a la glucosa capilar en los puntos de calibración. Después de la inserción del sensor SC y una vez que el proceso de estabilización esté completo, el monitor mide la señal de corriente eléctrica continua, denominada por el fabricante *ISIG* (medida en nA), generada por el sensor de glucosa a una tasa de muestreo de 10 segundos. A esta señal *ISIG* se le realiza un filtrado inicial, considerando las seis muestras obtenidas en un minuto, de las cuáles se descarta el valor más alto y el valor más bajo y se promedian los cuatro valores restantes para obtener una muestra cada un minuto. Seguidamente, se aplica un nuevo proceso de filtrado, ignorando el valor más bajo y el más alto de la señal *ISIG* muestreada cada minuto y promediando las tres muestras restantes para obtener un valor promedio cada cinco minutos, el cual se almacena en la memoria del monitor del CGMS® Gold™. Cada valor de *ISIG* almacenado en la memoria se considera válido a menos que ocurra un evento de cancelación, y la señal se avanza en el tiempo en dos períodos de muestreo (10 minutos) para tener en cuenta el retardo fisiológico entre IG y BG, es decir, se considera la corriente adelantada 10 minutos para estimar la concentración de BG. Esta señal de corriente adelantada 10 minutos es llamada por el fabricante *Valid ISIG* (medida en nA).

Utilizando una técnica de regresión lineal modificada, este algoritmo determina los factores de calibración que transforman la lectura de *Valid ISIG* en una concentración de glucosa (en un rango de 40 mg/dl a 400 mg/dl). La ecuación de la regresión lineal ha sido denominada por el fabricante *Razón de la Sensibilidad de la Regresión Lineal Modificada* (-del inglés *Modified Linear Regression Sensitivity Ratio (MLRSR)*-) (ecuación (3.1)).

$$MLRSR = \frac{\sum_{i=1}^N [X_i - Offset] [Y_i]}{\sum_{i=1}^N [X_i - Offset]^2} \quad (3.1)$$

donde X_i , Y_i y N hacen referencia a los pares de datos para calibración (determinaciones de glucosa en capilar utilizadas para calibrar el CGMS® Gold™ y las medidas de *Valid ISIG*

asociadas a ellas), siendo X_i el i -ésimo valor de *Valid ISIG*, Y_i el i -ésimo valor de glucosa en capilar (considerado como la determinación de referencia), y N el número total de pares de datos.

En este algoritmo retrospectivo, dado un valor de *Valid ISIG*, *MLRSR* se calcula usando los pares de datos para calibración, en una ventana de tiempo de 24 horas (12 horas antes y 12 horas después del instante de tiempo en que se obtiene en nuevo valor de *Valid ISIG*), considerando una relación lineal entre ambas medidas y usando cinco diferentes valores de *Offset* asignados empíricamente por el fabricante. El valor de *Offset* se selecciona empíricamente de un rango de -6 a $+6$, con un paso de cambio de 3 ($-6, -3, 0, 3, 6$). La pendiente aplicada corresponde al valor de *MLRSR* que minimice la media de *AD* o *Mean AD* (ecuación (2.17)) de los pares de datos de calibración dentro de la ventana de tiempo. La pendiente se suaviza utilizando una ventana temporal de dos horas (-1 a $+1$ hora), para evitar cambios bruscos de pendiente, denominada por el fabricante *Slope*. Finalmente, el A-GoldTM estima la concentración de glucosa utilizando *Valid ISIG*, *Slope* y *Offset*. La implementación de este algoritmo se realizó basándose en la información proporcionada por la guía de usuario del CGMS[®] GoldTM (MiniMed, 2003) y por una antigua versión de su patente (Mastrototaro et al., 2002).

La información proporcionada por el CGMS[®] GoldTM en el proceso de descarga de datos es la siguiente:

- *Time*. Instante de tiempo en que se almacena el valor promedio de corriente eléctrica en la memoria del monitor del CGMS[®] GoldTM. El intervalo de tiempo de muestreo es de cinco minutos.
- *ISIG*. Valor promedio de corriente eléctrica generada por el sensor de glucosa y almacenada en la memoria del monitor continuo de glucosa, medida en nA.
- *Valid ISIG*: Valor de *ISIG* adelantado 10 minutos y seleccionado para la conversión a valores de glucosa en mg/dL. Se refiere a los valores de corriente para los cuales la calibración ha sido considerada adecuada y también es medida en nA.
- *VCTR*. Es el voltaje en el electrodo del sensor de glucosa, medido en V.
- *Meter BG*. Valor de glucosa en capilar usado para calibrar el monitor. A esta variable la hemos denominado $G_{Capillary}$.
- *MLRSR*. Valor de *MLRSR* que minimiza el error relativo absoluto de la estimación de BG con respecto a los pares de datos de calibración encontrados dentro de la ventana de tiempo.
- *Slope*. Pendiente suavizada utilizando valores de *MLRSR* en una ventana temporal de dos horas (-1 a $+1$ hora), para evitar cambios bruscos de pendiente. Se aplica en la estimación de BG proporcionada por el A-GoldTM.
- *Offset*. Valor de *Offset* (seleccionado del conjunto de cinco posibles valores) aplicado en la estimación de BG proporcionada por el A-GoldTM.
- *Glucose Level*. Concentración de BG estimada por el A-GoldTM (medida en mg/dL). A esta variable la hemos llamado G_{Gold} .

Para mejorar la precisión del CGMS[®] GoldTM, se han propuesto en la literatura algunos algoritmos alternativos para la estimación de la glucosa (Panteleon et al., 2003; Lesperance et al.,

2007). El método propuesto por Panteleon et al. propone realizar cambios en la variable que se considera como variable independiente de la regresión lineal (Panteleon et al., 2003). Este método ha permitido concluir que la precisión del CGMS[®] Gold[™] se ve afectada en función de cuál variable de la calibración se considera como variable independiente de la regresión lineal: *Valid ISIG* o glucosa en capilar. Si se utiliza *Valid ISIG* como variable independiente, se reduce el error entre la glucosa estimada por el algoritmo de calibración y la determinación de referencia BG, pero se presentan sobrestimaciones de los eventos hipoglucémicos. Si se utiliza la glucosa en capilar como variable independiente, el error se incrementa. En otro método propuesto por Lesperance et al. se plantea un algoritmo que utiliza transformaciones logarítmicas y técnicas de regresión múltiples para cuantificar un factor de dependencia temporal (Lesperance et al., 2007).

3.2.2. ARP-SMCG

Considerando que la glucosa estimada por el A-Gold[™] se calcula a partir de *Valid ISIG* y $G_{Capillary}$, se utilizan estas mismas variables y el principio de funcionamiento básico de este algoritmo (descrito en la sección 3.2.1) para proponer un nuevo algoritmo retrospectivo que estima concentraciones de BG. Este nuevo algoritmo retrospectivo, al que hemos denominado ARP-SMCG, también se basa en una técnica de regresión lineal modificada.

De la información proporcionada por el CGMS[®] Gold[™] en el proceso de descarga de datos, el ARP-SMCG sólo utiliza los datos de *Valid ISIG*, $G_{Capillary}$ y *Time*. A diferencia del A-Gold[™], esta metodología considera que *Offset* puede tomar 25 diferentes valores, en un rango de -6 a $+6$, con un paso de cambio de $0, 5$ ($-6, -5, 5, -5, -4, 5, \dots +6$).

A continuación se describe el proceso que realiza el ARP-SMCG para estimar la concentración de glucosa cada cinco minutos:

1. Tomando como base la información proporcionada por el CGMS[®] Gold[™], para cada una de las muestras de *Valid ISIG*, el ARP-SMCG aplica una ventana temporal deslizante de 24 horas (-12 a $+12$ horas), partiendo del instante de tiempo en el que se obtuvo la muestra de *Valid ISIG*, y busca dentro de ella los pares de datos $G_{Capillary}/Valid ISIG$. Luego, utilizando los pares de datos encontrados, a partir de la ecuación (3.1), se calcula una matriz $MLRSR (Offset)$ de $m \times 1$, donde m se calcula en función de cada uno de los 25 valores que puede tomar (*Offset*). $MLRSR (Offset)$ puede escribirse como (ecuación (3.2)),

$$MLRSR (Offset) = \frac{\sum_{i=1}^{Z_1} [Valid ISIG_i - Offset] [G_{Capillary} i]}{\sum_{i=1}^{Z_1} [Valid ISIG_i - Offset]^2} \quad (3.2)$$

donde Z_1 representa el número de pares de datos $G_{Capillary}/Valid ISIG$ que se encuentran dentro de la ventana temporal deslizante. Por lo tanto, cada una de las pendientes que integran el conjunto $MLRSR (Offset)$ se asocia con un (*Offset*) diferente.

2. Una vez finalizado el cálculo de $MLRSR (Offset)$ para cada una de las muestras de *Valid ISIG*, se busca evitar cambios bruscos de pendiente cuando se realiza la estimación de BG. Por lo tanto, se aplica una ventana temporal deslizante de dos horas (-1 a $+1$ h), partiendo del instante de tiempo en que se obtuvo el valor *Valid ISIG* que está siendo evaluado, para calcular la pendiente media, a la que denominaremos *Slope (Offset)*, definida como el promedio

de los valores de $MLRSR (Offset)$ que contiene la ventana (ecuación (3.3)). Debido a que $Slope (Offset)$ también se calcula en función de $Offset$, es una matriz de $m \times 1$, donde cada uno de los m valores se asocia con un $Offset$ diferente. Teniendo en cuenta que las muestras de *Valid ISIG* se tomaron cada cinco minutos, la ventana temporal deslizante de dos horas comprende 25 valores de $MLRSR (Offset)$.

$$Slope (Offset) = \frac{\sum_{i=1}^{25} MLRSR (Offset)_i}{25} \quad (3.3)$$

3. Para cada una de las muestras de *Valid ISIG* se calcula una matriz de $m \times 1$ concentraciones de glucosa, a la que denominaremos $Glucose (Offset)$ (medida en mg/dL), donde cada uno de los m valores de $Glucose (Offset)$ se asocia a un valor de $Offset$ diferente (ecuación (3.4)).

$$Glucose (Offset) = Slope (Offset) (Valid ISIG - Offset) \quad (3.4)$$

4. Luego, para cada una de las muestras de *Valid ISIG*, se selecciona el valor de $Offset$ que permite minimizar $Mean RAD$. Para ello, se define a partir de la (ecuación (2.16)) una matriz de $m \times 1$, a la que denominaremos $Mean RAD (Offset)$, donde cada uno de los m valores se calcula en función de cada uno de los diferentes valores de $Offset$ (ecuación (3.5)). En el cálculo de cada uno de los valores de $Mean RAD (Offset)$, se aplica una ventana temporal deslizante de longitud variable que garantice la existencia de por lo menos tres pares de datos $Glucose (Offset)/G_{Capillary}$ dentro de ella. La longitud inicial de la ventana es de por lo menos 24 horas (-12 a $+12$ horas). Si en la longitud inicial no se cumple el requisito anterior respecto al número de pares de datos, la ventana va ampliando su longitud de búsqueda en dos (2) horas más (-1 a $+1$ horas) hasta alcanzar una longitud que permita encontrar por lo menos tres calibraciones dentro de ella. N ($N \geq 3$) representa el número de pares de datos $Glucose (Offset)/G_{Capillary}$ dentro de la ventana de longitud variable.

$$Mean RAD (Offset) = \frac{\sum_{i=1}^N \left| \frac{G_{Capillary i} - Glucose (Offset)_i}{G_{Capillary i}} \right|}{N} 100 \quad (3.5)$$

5. Se determina el valor de $Offset$ que permite minimizar $Mean RAD (Offset)$, al que denominaremos $Offset_{min}$, y al mínimo valor de $Mean RAD (Offset)$ lo denominaremos $Mean RAD_{min}$.

6. A continuación, se asigna el correspondiente $Offset_{min}$ a cada una de las muestras de *Valid ISIG*. Sin embargo, el $Offset_{min}$ no se puede considerar como el $Offset$ definitivo (al que denominaremos $Offset_F$) para calcular la concentración de glucosa que reporta el ARP-SMCG, debido a que se podrían presentar cambios bruscos entre los valores de $Offset_{min}$ asignados a cada muestra. Para evitar drásticas diferencias en el $Offset$ entre muestras adyacentes, se imponen ciertas restricciones en la asignación de $Offset_F$. Para ello, se define $\Delta Offset$ (ecuación (3.6)) como la diferencia entre $Offset_{min}$ en el instante actual ($Offset_{min} (S)$) y el $Offset_{min}$ en el instante anterior ($Offset_{min} (S-1)$).

$$\Delta Offset = Offset_{min} (S) - Offset_{min} (S-1) \quad (3.6)$$

7. Partiendo del valor obtenido en la ecuación (3.6), la asignación de $Offset_F$ para cada una de las muestras de *Valid ISIG* se hace teniendo en cuenta:

- Si $|\Delta Offset| > 0,5$, la asignación de $Offset_F (S)$ está sujeta a los siguientes criterios (ecuación (3.7)):

$$Offset_F(S) = \begin{cases} Offset_{min}(S-1) + 0,5 & \text{Si } \Delta Offset > 0 \\ Offset_{min}(S-1) - 0,5 & \text{Si } \Delta Offset < 0 \end{cases} \quad (3.7)$$

- Si $|\Delta Offset| \leq 0,5$, la asignación de $Offset_F (S)$ sería (ecuación (3.8)):

$$Offset_F(S) = Offset_{min}(S) \quad (3.8)$$

8. Luego, se calcula nuevamente $MLRSR$ para cada una de las muestras de *Valid ISIG* en función de $Offset_F$, de manera análoga a como se hizo en la ecuación (3.2). La nueva $MLRSR$ la denominaremos $MLRSR_F$ y se define como (ecuación (3.9)).

$$MLRSR (Offset)_F = \frac{\sum_{i=1}^{Z_1} [Valid ISIG_i - Offset_F] [G_{Capillary} i]}{\sum_{i=1}^{Z_1} [Valid ISIG_i - Offset_F]^2} \quad (3.9)$$

9. Nuevamente se utiliza una ventana de tiempo de dos horas (-1 a $+1$ horas) para suavizar $MLRSR_F$. La nueva pendiente media, a la que denominaremos $Slope_F$ se define como:

$$Slope_F = \frac{\sum_{i=1}^{25} MLRSR_{Fi}}{25} \quad (3.10)$$

10. Finalmente, utilizando cada una de las muestras de *Valid ISIG* y sus correspondientes valores de $Slope_F$ y $Offset_F$, se calcula la concentración de glucosa que reporta el ARP-SMCG cada cinco (5) minutos (ecuación (3.11)), a la que denominaremos $G_{Proposed}$ (medida en mg/dL).

$$G_{Proposed} = Slope_F (Valid ISIG - Offset_F) \quad (3.11)$$

El rango del *Offset* entre -6 y 6 se determina considerando el mismo rango de *Offset* que utiliza el A-GoldTM, pero reduciendo el paso de cambio. Este paso de cambio se seleccionó de manera empírica realizando numerosas simulaciones considerando diferentes pasos de cambio en el *Offset*. Utilizando un paso de cambio de 1 no se logró reducir el error. Reduciendo el paso de cambio a 0,1 se lograron similares resultados a los obtenidos con un paso de cambio de 0,5. Sin embargo, esta consideración implica un tiempo computacional más alto sin una significativa reducción en el error. Por tanto, se consideran 25 diferentes valores de *Offset*, correspondientes a un rango de -6 a $+6$, con un paso de cambio de 0,5 ($-6, -5,5, -5, -4,5, \dots +6$).

3.2.3. Evaluación de la precisión del ARP-SMCG usando criterios analíticos

La precisión analítica se mide por medio de diferentes métricas estadísticas (descritas en la sección 2.4.1) utilizando los pares de datos obtenidos al emparejar las lecturas proporcionadas por un algoritmo de calibración con las suministradas por un método de referencia (sangre capilar, arterial o venosa). En este caso, se evalúa la precisión analítica considerando como método

de referencia la determinación de glucosa capilar ($G_{Capillary}$). La glucosa capilar es considerada por los pacientes con DM-1 como el “*gold standard*” de las pruebas de control glucémico que se realizan en el cuidado diario de su diabetes. Estas muestras de $G_{Capillary}$ utilizadas como referencia para la evaluación son las que se emplean como puntos de calibración del A-GoldTM y del ARP-SMCG, debido a que no se disponía de muestras adicionales de glucosa capilar para realizar la evaluación durante las 72 horas de MCG, de forma similar a como se hizo en (Gandrud et al., 2007). Esta decisión también se soporta en que durante el uso del CGMS[®] GoldTM en un periodo de 72 horas para obtener información que permita ajustar la terapia del paciente diabético, el A-GoldTM utiliza como medidas de referencia para calcular el valor del RAD, las determinaciones de glucemia capilar que se introducen para calibrar el dispositivo. Durante las 72 horas de duración del estudio clínico, de los 21 pacientes que se monitorizaron se obtuvieron 358 determinaciones de $G_{Capillary}$.

Es importante mencionar que el ensayo clínico descrito en la sección 3.1.1 no fue diseñado específicamente para esta investigación. Por lo tanto, no se hizo un análisis *a priori* de la potencia del estudio. Sin embargo, si se hizo un análisis *a posteriori* de dicha potencia, para determinar si el número de determinaciones de $G_{Capillary}$ era suficiente para evaluar la precisión del CGMS[®] GoldTM utilizando criterios analíticos. Para un tamaño muestral de 358 muestras, se obtiene una potencia de 90,0% (se considera un nivel de significancia del 5,0%).

En este estudio, las muestras de referencia de $G_{Capillary}$ se comparan con las correspondientes estimaciones del A-GoldTM y del ARP-SMCG, con el fin de comprobar como es la capacidad en el ajuste de las estimaciones de BG de ambos algoritmos. En particular, se compara cuál de los dos algoritmos presenta un mejor rendimiento cuando trata de ajustar la estimación de BG (G_{Gold} y $G_{Proposed}$) para minimizar el error relativo absoluto de esta estimación con respecto a BG capilar ($G_{Capillary}$) en los puntos de calibración. Solamente las lecturas de G_{Gold} y $G_{Proposed}$ emparejadas con las muestras de $G_{Capillary}$ se emplean en el cálculo de los criterios analíticos, debido a que las medidas de $G_{Capillary}$ no estaban disponibles para el resto de las muestras de G_{Gold} y $G_{Proposed}$. La comparación se realiza sincronizando en el tiempo las determinaciones de $G_{Capillary}$ con las lecturas de G_{Gold} y $G_{Proposed}$, dado que estas muestras se obtuvieron en los mismos instantes de tiempo.

La evaluación de la precisión de las lecturas del A-GoldTM y del ARP-SMCG utilizando criterios analíticos se realiza considerando tanto la población global, como los rangos glucémicos de los pacientes con DM-1 (Tabla 3.2). En este caso, los rangos glucémicos se han definido siguiendo las recomendaciones del personal médico del Hospital Josep Trueta cuando se considera como método de referencia la determinación de glucosa capilar. Estos rangos glucémicos se corresponden con los definidos en (Bode et al., 2004; Mastrototaro et al., 2008), donde también se considera la determinación de glucosa capilar como el “*gold standard*”, aunque el SMCG utilizado no es el mismo. Los rangos glucémicos se han definido como: hipoglucémico ($G_{Capillary} \leq 80$ mg/dL), normoglucémico ($G_{Capillary}$ entre 81 mg/dL – 120 mg/dL), hiperglucémico moderado ($G_{Capillary}$ entre 121 mg/dL – 240 mg/dL) e hiperglucémico severo ($G_{Capillary} > 240$ mg/dL).

En primer lugar, se determina la concordancia entre los pares de datos comparados evaluando las diferencias entre $G_{Capillary}/G_{Gold}$ y $G_{Capillary}/G_{Proposed}$. Previamente a la aplicación de estos criterios de evaluación, se realiza un análisis de normalidad a los pares de datos, para determinar si estos siguen o no una distribución normal, empleando para ello el Test de Shapiro-Wilk⁴.

⁴En estadística, el Test de Shapiro-Wilk se usa para contrastar la normalidad de un conjunto de datos. Se

Rango de la glucemia (mg/dL)	<i>Median RAD (IQR) (%)</i>		<i>Criterios ISO (%)</i>		<i>Test de suma de rangos de Wilcoxon</i>
	A-Gold TM	ARP-SMCG	A-Gold TM	ARP-SMCG	<i>p-valor</i>
Población global	8,1 (3,2 – 15,1)	6,1 (2,1 – 12,9)	86,3	92,5	<0,05
≤80	14,6 (7,5 – 25,0)	5,2 (2,0 – 18,0)	80,3	94,7	<0,05
81 – 120	9,6 (4,9 – 18,8)	7,8 (2,9 – 18,0)	79,7	76,6	>0,05
121 – 240	6,7 (2,8 – 11,4)	5,7 (2,1 – 11,4)	89,2	90,4	>0,05
>240	6,0 (2,0 – 10,7)	6,0 (2,0 – 11,2)	93,4	93,4	>0,05

^aPara el caso de $VBG \leq 80$ mg/dL, se calcula *Median AD* (mg/dL).

Tabla 3.2: Resultados de evaluar la precisión analítica del A-GoldTM y del ARP-SMCG, usando los pares de datos $G_{Capillary}/G_{Gold}$ y $G_{Capillary}/G_{Proposed}$, considerando tanto la población global como el estado glucémico.

Debido a que los pares de datos no siguen una distribución normal, el estadístico a utilizar es la mediana (IQR) del *RAD*, también conocida como *Median RAD (IQR)* (Tabla 3.2). El Test de suma de rangos de Wilcoxon⁵ se utiliza para comparar el valor de *Median RAD (IQR)* del A-GoldTM y el ARP-SMCG, y determinar si existen diferencias estadísticamente significativas entre ellos. Al comparar los resultados de ambos algoritmos, se observa un contraste estadísticamente significativo (*Median RAD (IQR)* = 8,1 (3,2 – 15,1) % versus 6,1 (2,1 – 12,9) %, *p-valor* < 0,05, respectivamente).

Los resultados de la evaluación considerando el estado glucémico del enfermo con DM-1 (76 muestras de $G_{Capillary} \leq 80$ mg/dL, 64 muestras de $G_{Capillary}$ entre 81 mg/dL – 120 mg/dL, 157 muestras de $G_{Capillary}$ entre 121 mg/dL – 240 mg/dL y 61 muestras de $G_{Capillary} > 240$ mg/dL) también se presentan en la Tabla 3.2. Estos resultados indican diferencias significativas en el rango hipoglucémico entre ambos algoritmos (*Median RAD (IQR)* = 14,6 (7,5 – 25,0) % versus 5,2 (2,0 – 18,0) %, *p-valor* < 0,05, respectivamente). Estos resultados también muestran que los dos algoritmos son igualmente precisos cuando los pacientes se encuentran en estado normoglucémico, hiperglucémico moderado e hiperglucémico severo (*p-valor* > 0,05). Para realizar el contraste de hipótesis se aplica el Test de suma de rangos de Wilcoxon para comparar el valor de *Median RAD (IQR)*.

Otro criterio de precisión analítico que se considera son los criterios ISO, previamente descritos en la sección 2.4.1. La Tabla 3.2 muestra el porcentaje de pares de datos que satisfacen estos criterios, tanto en el conjunto global de pares de datos, como en cada uno de los rangos glucémicos evaluados. En el conjunto global de pares de datos, en comparación con el A-GoldTM, con el ARP-SMCG se obtiene un incremento en el porcentaje de los datos que satisfacen los criterios ISO (Criterios ISO = 86,3 % versus 92,5 %). Los resultados obtenidos de la aplicación de estos

plantea como hipótesis nula que una muestra x_1, \dots, x_n proviene de una población normalmente distribuida. Se considera uno de los test más potentes para el contraste de normalidad, sobre todo para muestras pequeñas ($n < 30$) (Shapiro, 1965).

⁵En estadística, este test también llamado la prueba de suma de rangos Wilcoxon, el test de Mann-Whitney-Wilcoxon, la prueba U de Mann-Whitney, o la prueba de Wilcoxon-Mann-Whitney, es una técnica no paramétrica aplicada a dos muestras independientes. Se usa para comprobar la heterogeneidad de dos muestras mediante el cálculo del estadístico U, comparando la mediana de las muestras relacionadas y determinando si existen diferencias entre ellas. Es, de hecho, la versión no paramétrica de la habitual prueba t de Student. Fue propuesta inicialmente en 1945 por Frank Wilcoxon para muestras de igual tamaño y extendida a muestras de tamaño arbitrario por Henry B. Mann y D. R. Whitney en 1947 (Wilcoxon, 1945).

criterios permiten corroborar que el ARP-SMCG presenta mejor capacidad en el ajuste de las estimaciones de BG en el rango hipoglucémico, en comparación con el rendimiento presentado por el A-GoldTM (Tabla 3.2). Según los criterios ISO, tanto en los rangos normoglucémico, hiperglucémico moderado e hiperglucémico severo, el rendimiento de los dos algoritmos es similar.

Finalmente, en la Tabla 3.3 también se presentan los criterios de precisión analíticos de acuerdo con la tasa absoluta de cambio del valor de la glucosa (202 muestras $<0,5$ mg/dL/min, 89 muestras entre $0,5$ mg/dL/min – $1,0$ mg/dL/min, 50 muestras entre $1,0$ mg/dL/min – $1,5$ mg/dL/min y 17 muestras $\geq 1,5$ mg/dL/min). Al comparar el valor de *Median RAD (IQR)*, se observa que cuando la tasa absoluta de cambio del valor de la glucosa durante la calibración es más baja ($<0,5$ mg/dL/min), el ARP-SMCG es significativamente mejor que el algoritmo del CGMS[®] GoldTM. Los criterios ISO por su parte reportaron diferencias relevantes entre los dos algoritmos cuando la tasa absoluta de cambio es más alta ($\geq 1,5$ mg/dL/min). En los demás rangos evaluados, no se aprecian diferencias significativas entre ambos algoritmos. A manera de referencia, en un estudio de DirecNet (Direcnet et al., 2006) los valores de *Median RAD (IQR)* y de los criterios ISO fueron $11,0\%$ y $73,0\%$, respectivamente, cuando la tasa absoluta de cambio fue $\leq 0,5$ mg/dL/min; $13,0\%$ y $69,0\%$, respectivamente, para una tasa absoluta de cambio entre $0,5 - 1,0$ mg/dL/min; $14,0\%$ y $72,0\%$, respectivamente, para una tasa absoluta de cambio entre $1,0 - 1,5$ mg/dL/min; y $17,0\%$ y $62,0\%$, respectivamente, cuando la tasa absoluta de cambio fue $>1,5$ mg/dL/min.

Tasa de cambio de la glucemia (mg/dL/min)	<i>Median RAD (IQR) (%)</i>		<i>Criterios ISO (%)</i>		<i>Test de suma de rangos de Wilcoxon</i>
	A-Gold TM	ARP-SMCG	A-Gold TM	ARP-SMCG	<i>p-valor</i>
$<0,5$	8,4 (3,5 – 16,1)	5,6 (2,2 – 13,2)	85,7	86,6	$<0,05$
$0,5 - 1,0$	7,4 (3,4 – 13,7)	5,9 (2,0 – 11,8)	89,2	91,0	$>0,05$
$1,0 - 1,5$	8,9 (4,0 – 14,6)	6,7 (1,9 – 16,3)	88,1	88,0	$>0,05$
$\geq 1,5$	8,5 (4,9 – 23,9)	8,7 (4,7 – 15,7)	69,2	76,5	$>0,05$

Tabla 3.3: Resultados de evaluar la precisión analítica del A-GoldTM y del ARP-SMCG, usando los pares de datos $G_{Capillary}/G_{Gold}$ y $G_{Capillary}/G_{Proposed}$, de acuerdo con la tasa de cambio de la glucemia.

3.2.4. Evaluación de la precisión del ARP-SMCG usando criterios clínicos

En la sección 3.1.2 se indicó que en el ensayo clínico se determinó BG venosa en un laboratorio. Al igual que la precisión analítica, la precisión clínica se mide comparando pares de datos. Sin embargo, en la aplicación de estos criterios clínicos, se evalúa la precisión considerando como método de referencia *VBG*.

En la evaluación de la precisión de los algoritmos utilizando criterios clínicos se utilizan las determinaciones de *VBG* recogidas durante el primer día del estudio clínico. Por lo tanto, en esta evaluación se utilizan pares de datos VBG/G_{Gold} y $VBG/G_{Proposed}$. Solamente las lecturas de *VBG* se emplean en el cálculo de los criterios clínicos, por lo cuál estos criterios sólo se calculan para el primer día del estudio, debido a que estas determinaciones no estaban disponibles para los restantes dos días. De los 21 pacientes que se monitorizaron se obtuvieron 644 determinaciones de *VBG*. Estas determinaciones se tomaron cada 15 minutos durante las dos

horas sucesivas a cada ingesta y cada 30 minutos en las horas siguientes.

Para obtener los pares de datos, se deben sincronizar las determinaciones de VBG con las muestras simultáneas de G_{Gold} y $G_{Proposed}$. Sin embargo, las muestras G_{Gold} y $G_{Proposed}$ tienen tiempos de muestreo diferentes a las determinaciones de VBG . Por lo tanto, se aplica un método de interpolación para muestrear las determinaciones de VBG cada cinco minutos a los mismos instantes de tiempo en que se tomaron las muestras de G_{Gold} y $G_{Proposed}$. Aplicar este método de interpolación permite también incrementar el número de muestras de VBG . Esta decisión también se soportó en un estudio previo donde se trabajó con el mismo conjunto de datos clínicos utilizado en el desarrollo del ARP-SMCG, en el cuál se interpolaron las determinaciones de BG consideradas como el “gold standard” (Bondia et al., 2008).

Se prueban cuatro métodos de interpolación: interpolación por el vecino más cercano, interpolación lineal, interpolación cúbica e interpolación por medio de “splines” clásicos, es decir, “splines” de orden tres que buscan suavizar la función objetivo subyacente hasta el segundo orden. Tanto la interpolación por el vecino más cercano como la interpolación lineal retornan algunos valores que no son posibles de calcular (-del inglés *Not and Number (NaN)*-), por tanto se descartan. Los resultados de la interpolación por medio de splines indican pequeñas diferencias con los valores de VBG . Finalmente, se decidió considerar el método de interpolación cúbica, un método rápido, eficiente y estable de la función de interpolación, el cuál presentaba las menores diferencias con las determinaciones de VBG . Además, la interpolación cúbica había sido utilizada en (Bondia et al., 2008). En la Figura 3.1 se puede observar, a manera de ejemplo, los resultados de aplicar diferentes métodos de interpolación para muestrear las determinaciones de VBG cada cinco minutos a los mismos instantes de tiempo en que se tomaron las muestras de G_{Gold} . En consecuencia, después de aplicar el método de interpolación cúbica a las determinaciones de VBG se obtienen 2450 pares de datos VBG/G_{Gold} y $VBG/G_{Proposed}$ para usarse en la evaluación.

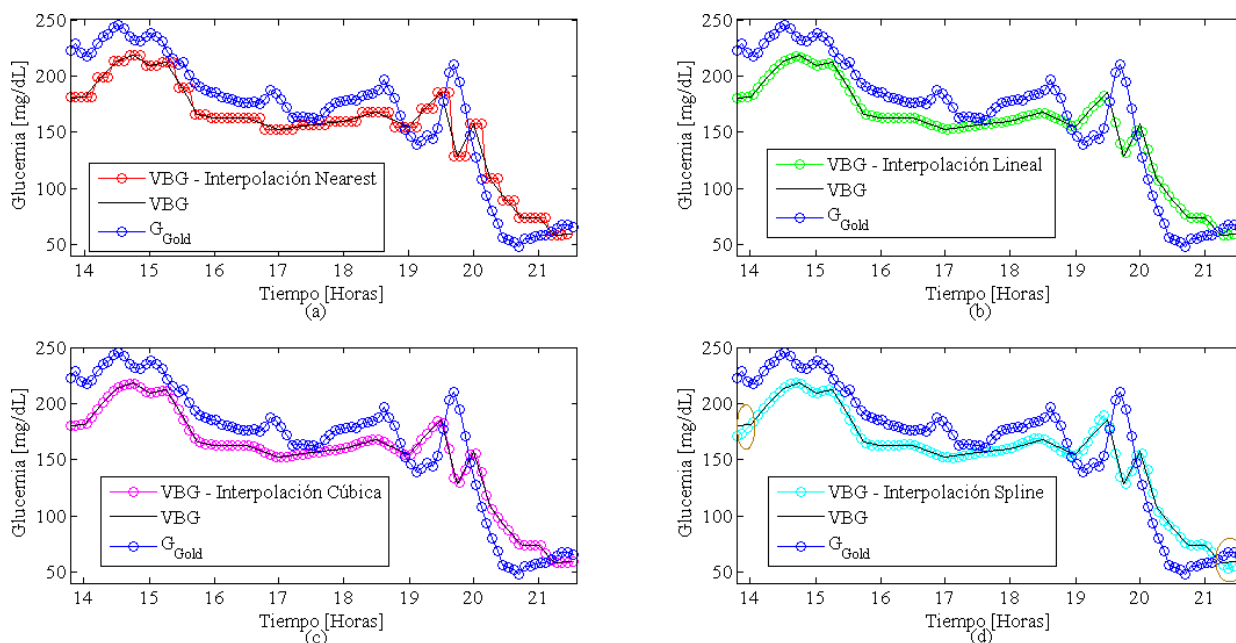


Figura 3.1: Ejemplo de aplicar cuatro diferentes métodos de interpolación para muestrear las determinaciones de VBG cada cinco minutos.

Las rejillas de errores de Clarke y Consensus, previamente explicadas en la sección 2.4.1, se utilizan para determinar la relevancia clínica de las diferencias existentes entre las determinaciones de VBG y las estimaciones de ambos algoritmos. El porcentaje de puntos que caen dentro de cada zona o categoría clínica para cada uno de los dos algoritmos, tanto en Clarke EGA como en Consensus EGA, se presenta en la Tabla 3.4. Esta evaluación se realiza considerando como población global los pacientes que participaron en el estudio. Los resultados obtenidos en las dos rejillas de errores evidencian un mayor número de lecturas precisas con el ARP-SMCG. Analizando los resultados de Clarke EGA, con el ARP-SMCG se incrementa la distribución de datos en la zona A y se reduce en las zonas B, D y E, con respecto al A-GoldTM. Analizando los resultados de Consensus EGA, con el ARP-SMCG se aumenta la concentración de datos en la zona A y se disminuye en las zonas B y C.

Tipo	Clarke EGA					Consensus EGA				
	A(%)	B(%)	C(%)	D(%)	E(%)	A(%)	B(%)	C(%)	D(%)	E(%)
A-Gold TM	76,0	17,6	0,0	6,1	0,3	78,8	17,7	3,6	0,0	0,0
ARP-SMCG	79,3	14,5	0,0	6,2	0,0	82,7	14,8	2,5	0,0	0,0

Tabla 3.4: Resultados de evaluar la precisión clínica del A-GoldTM y del ARP-SMCG, usando los pares de datos VBG/G_{Gold} y $VBG/G_{Proposed}$ obtenidos de todos los pacientes con DM-1.

La Figura 3.2 ilustra la distribución de los puntos que representan los pares de datos VBG/G_{Gold} y $VBG/G_{Proposed}$ dentro de cada zona clínica de Clarke EGA y de Consensus EGA, tanto del A-GoldTM (Figuras 3.2(a) y 3.2(c), respectivamente), como del ARP-SMCG (Figuras 3.2(b) y 3.2(d), respectivamente).

3.2.5. Análisis de la capacidad de estimación de hipoglucemias e hiperglucemias del ARP-SMCG

Con el objetivo de comparar la capacidad del ARP-SMCG y del A-GoldTM para estimar eventos hipoglucémicos e hiperglucémicos correctamente, se obtienen los perfiles glucémicos de los 21 pacientes con DM-1. En las Figuras 3.3, 3.4 y 3.5 se ilustran como ejemplos los perfiles glucémicos de tres pacientes. Estos pacientes se seleccionan porque en ellos las determinaciones consideradas como referencia o “gold standard” (VBG y $G_{Capillary}$) indican eventos hipoglucémicos e hiperglucémicos importantes, los cuáles permiten analizar la capacidad de estimación de hipoglucemias e hiperglucemias del ARP-SMCG.

En estas gráficas se representan los perfiles glucémicos proporcionados por los dos algoritmos a comparar, así como las determinaciones consideradas como referencia (VBG y $G_{Capillary}$). La línea magenta representa las determinaciones de VBG , la línea verde las muestras de $G_{Proposed}$, la línea negra corresponde a G_{Gold} y los rombos rojos representan las determinaciones de $G_{Capillary}$. El umbral para definir una hipoglucemia (VBG o $G_{Capillary} \leq 80$ mg/dL) se representa por una línea cian punteada y el umbral para definir una hiperglucemia severa (VBG o $G_{Capillary} > 240$ mg/dL) se representa por una línea amarilla punteada. Adicionalmente, para cada uno de estos pacientes representativos, también se indican los resultados de la evaluación considerando *Median RAD (IQR)*, donde la medida de referencia es $G_{Capillary}$, y Clarke EGA y Consensus EGA, donde la medida de referencia es VBG .

En el perfil que se ilustra en la Figura 3.3, se muestra el comportamiento del ARP-SMCG durante la estancia hospitalaria del paciente y se evalúa la capacidad de este algoritmo para

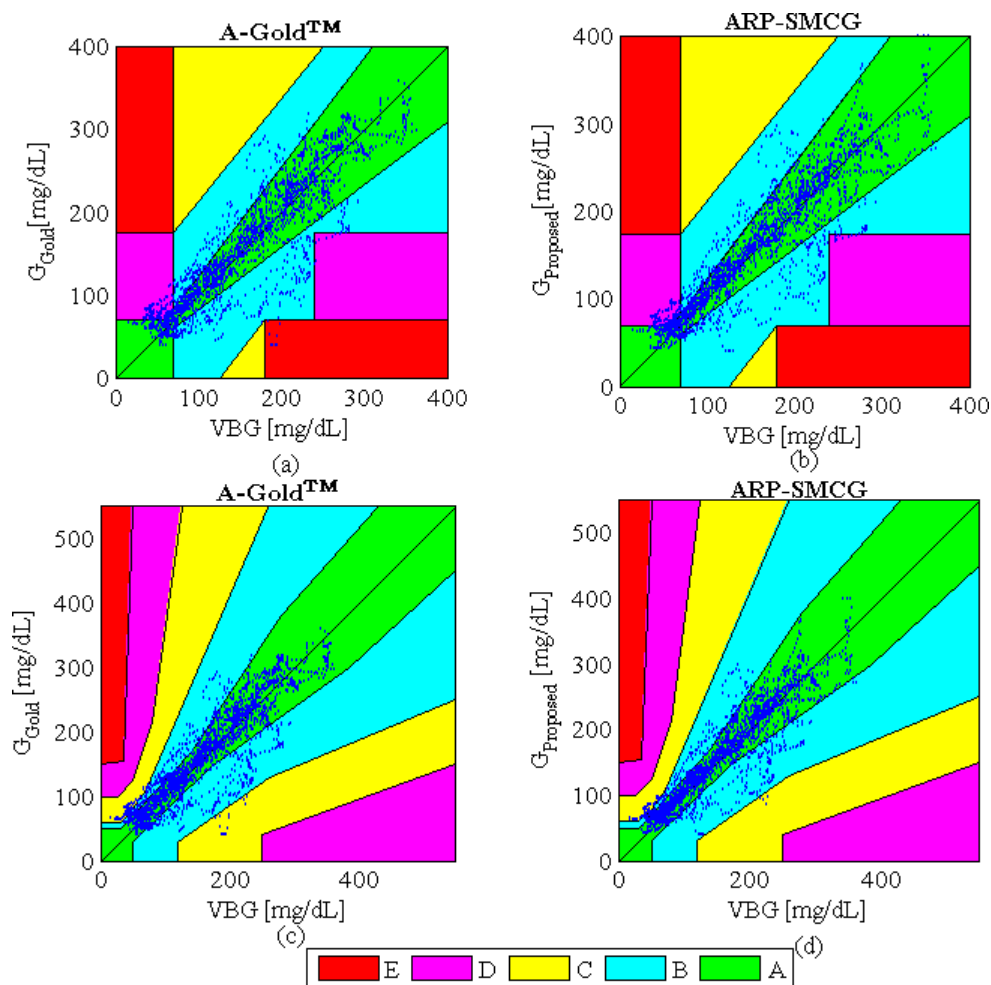


Figura 3.2: Rejillas de Clarke EGA y Consensus EGA para evaluar el A-Gold™ y el ARP-SMCG, usando los pares de datos VBG/G_{Gold} y $VBG/G_{Proposed}$ obtenidos de todos los pacientes con DM-1.

estimar eventos hipoglucémicos. El episodio hipoglucémico indicado en VBG , el cuál es omitido por el A-Gold™, es estimado correctamente por el ARP-SMCG. Considerando $G_{Capillary}$ como *gold standard*, el valor de *Median RAD (IQR)* del ARP-SMCG es 6,2 (5,3)%, mientras el del A-Gold™ es 7,1 (6,3)%. Analizando los resultados obtenidos de Clarke EGA y Consensus EGA, se evidencia una mayor precisión del ARP-SMCG. Clarke EGA clasifica el 96,0% de los puntos que representan los pares de datos $VBG/G_{Proposed}$ en la zona A versus el 72,0% de VBG/G_{Gold} . Consensus EGA clasifica el 99,2% de los puntos que representan los pares de datos $VBG/G_{Proposed}$ en la zona A versus el 73,6% de VBG/G_{Gold} .

Por otra parte, en el perfil que se ilustra en la Figura 3.4, también se muestra el comportamiento del ARP-SMCG mientras el paciente permanece recluido en el hospital, y se evalúa la capacidad de dicho algoritmo para estimar eventos hiperglucémicos. Para este paciente, el A-Gold™ muestra falsos eventos hipoglucémicos, estimando glucemias menores a 50 mg/dL, cuando las determinaciones de VBG realmente indican eventos hiperglucémicos. Aunque el ARP-SMCG no reporta este estado hiperglucémico, si proporciona glucemias más cercanas a VBG . El

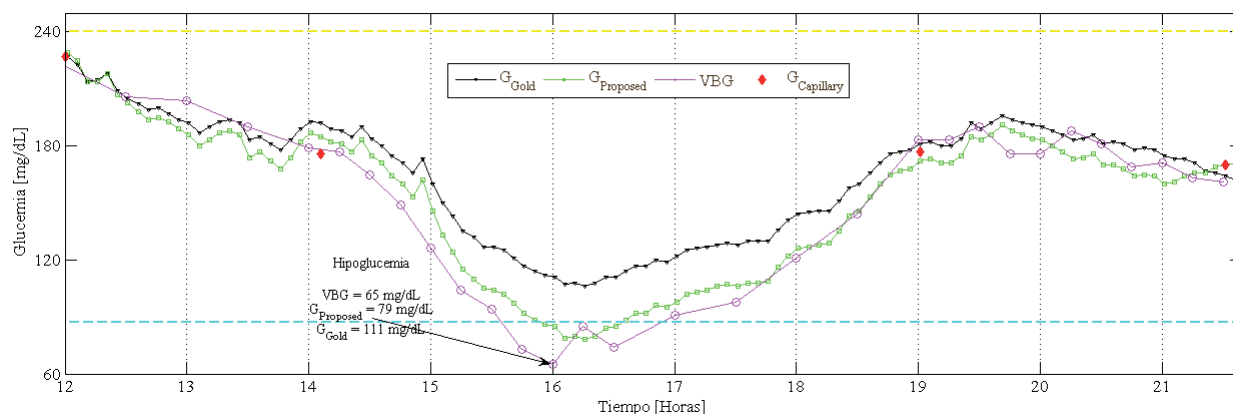


Figura 3.3: Ejemplos de perfiles glucémicos hospitalarios estimados por el A-GoldTM y el ARP-SMCG: Estimación de episodios hipoglucémicos.

valor de *Median RAD (IQR)* del ARP-SMCG es 13,4 (16,9) %, mientras el del A-GoldTM es 16,7 (17,7) %. La evaluación utilizando criterios clínicos indica que Clarke EGA clasifica el 86,7 % de los puntos que representan los pares de datos $VBG/G_{Proposed}$ en la zona A, en comparación con el 70,4 % de VBG/G_{Gold} . Además, con el ARP-SMCG no se clasifica ningún dato en las zonas D o E, a diferencia del A-GoldTM, con el cual 1,5 % de los puntos están en zona D y 5,9 % en zona E. Consensus EGA clasifica el 71,1 % de VBG/G_{Gold} en zona A, en comparación con el 80,7 % de $VBG/G_{Proposed}$.

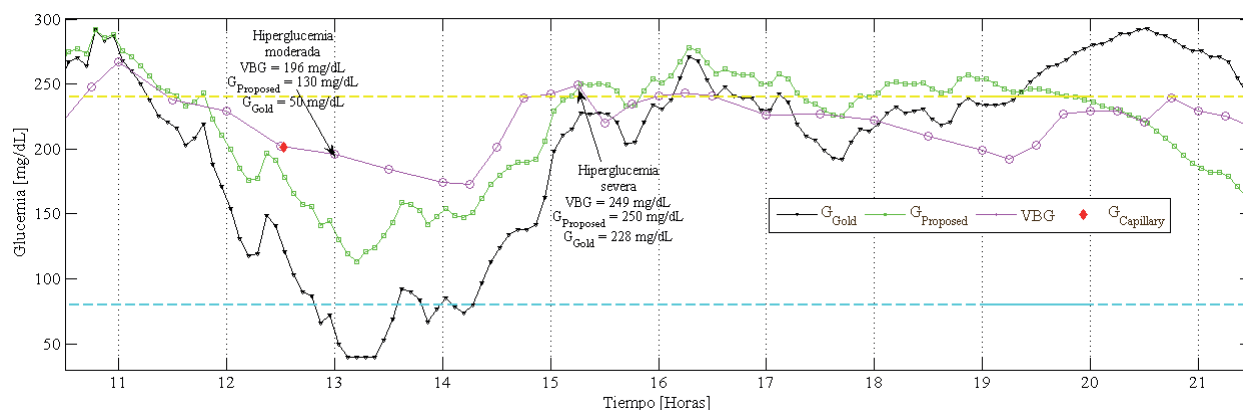


Figura 3.4: Ejemplos de perfiles glucémicos hospitalarios estimados por el A-GoldTM y el ARP-SMCG: Estimación de episodios hiperglucémicos.

Finalmente, en el perfil que se ilustra en la Figura 3.5, se muestra el comportamiento del ARP-SMCG solamente para los datos domiciliarios, después que el paciente ha abandonado el hospital. En este caso se considera como medida de referencia $G_{Capillary}$, debido a que no se dispone de las determinaciones de VBG . Al comparar $G_{Proposed}$ con $G_{Capillary}$, se observa que el ARP-SMCG estima correctamente los eventos hipoglucémicos. Por su parte, el A-GoldTM indica que se trata de eventos dentro del rango normoglucémico. En esta figura se puede observar que el ARP-SMCG estima glucemias más cercanas a $G_{Capillary}$ que el A-GoldTM, particularmente

en el rango hipoglucémico. El valor de *Median RAD (IQR)* del ARP-SMCG es 12,4 (11,1) %, mientras el del A-GoldTM es 17,0 (14,5) %. En este caso, no se presentan los resultados de los criterios clínicos, debido a que se trata de las muestras domiciliarias y no se dispone de las determinaciones de *VBG*.

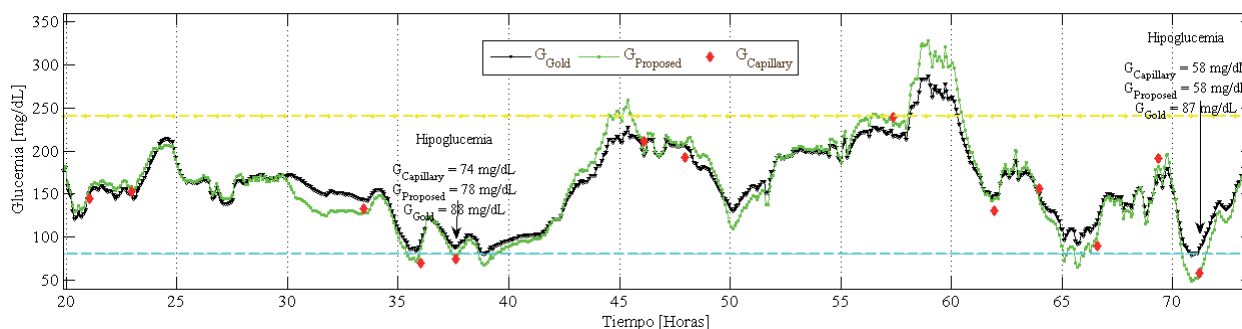


Figura 3.5: Ejemplos de perfiles glucémicos domiciliarios estimados por el A-GoldTM y el ARP-SMCG.

Es importante mencionar, de acuerdo a los resultados obtenidos del análisis que se realiza a los eventos hipoglucémicos capilares que se presentaron en los 21 pacientes durante los tres días del estudio clínico ($G_{Capillary} \leq 80$ mg/dL), que el ARP-SMCG es más sensible y mejora el ajuste de las estimaciones de glucosa en el rango hipoglucémico en comparación con el A-GoldTM. El porcentaje de episodios hipoglucémicos capilares estimados correctamente por el ARP-SMCG es 94,7 %, mientras que el del A-GoldTM es 86,8 %. Por lo tanto, el porcentaje de episodios hipoglucémicos omitidos por el ARP-SMCG es menor que en el A-GoldTM (5,3 % versus 13,2 %, respectivamente).

3.2.6. Análisis de los resultados obtenidos

El A-GoldTM ha mostrado problemas de precisión en la determinación de BG, particularmente en la estimación de hipoglucemias. Para mejorar el ajuste en la estimación de BG, se propone el ARP-SMCG. Al igual que el A-GoldTM, el ARP-SMCG utiliza las determinaciones de glucosa capilar y las lecturas de corriente válida medidas por el sensor para estimar BG. Sin embargo, en el cálculo de las concentraciones de BG, el ARP-SMCG aumenta el número de valores considerados en el *Offset*, en comparación con los valores que considera el A-GoldTM en este mismo factor. Además, el ARP-SMCG garantiza que el *Offset* seleccionado para cada muestra de *Valid ISIG* sea un *Offset* variable, el cuál permite calcular la *MLRSR* que genere el mínimo error para cada muestra. Este algoritmo se considera como retrospectivo, debido a que para calcular la concentración de BG en un instante determinado, utiliza datos previos y posteriores al instante de tiempo en que se toma la muestra *Valid ISIG*, aplicando ventanas temporales deslizantes de longitud variable. Además, en este algoritmo, se suaviza el valor del *Offset* en el tiempo, utilizando ciertos criterios de ajuste para evitar cambios bruscos en la variación del *Offset* entre muestras adyacentes.

Realizando un análisis de la precisión desde el punto de vista analítico, si se comparan los resultados del ARP-SMCG con los resultados del A-GoldTM en las mismas condiciones de eva-

luación, el ARP-SMCG muestra un error global más reducido en la población total de pacientes con DM-1. En comparación con el A-GoldTM, este algoritmo mejora los valores de *Median RAD (IQR)* ($Median\ RAD\ (IQR) = 8,1\ (3,2-15,1)\%$ versus $6,1\ (2,1-12,9)\%$, $p\text{-valor} < 0,05$, respectivamente) y de los criterios ISO (Criterios ISO = $86,3\%$ versus $92,5\%$, respectivamente). Estos resultados globales se deben a la alta diferencia observada en el rango hipoglucémico, puesto que en los otros rangos glucémicos los resultados fueron similares en ambos algoritmos y no se observaron diferencias significativas entre ellos.

Por otra parte, a manera de referencia, se pueden contrastar los resultados indicados por el ARP-SMCG con los obtenidos en otros estudios donde se evalúa la precisión del A-GoldTM (Tabla 3.5):

Fuente	<i>Median RAD (%)</i>	<i>Criterios ISO (%)</i>	Glucosa de Referencia
(DirecNet, 2005)	12,0 (6,0 – 23,0)	72,0	Capilar (Glucómetro One Touch [®] Ultra [®])
(Direcnet et al., 2006) ^a	14,0	66,0	Capilar (Glucómetro One Touch [®] Ultra [®])
	15,0	67,0	
	13,0	71,0	
	13,0	72,0	
(Gandrud et al., 2007)	8,0	85,0	Capilar (Glucómetro One Touch [®] Ultra [®])

^aCuando realizan 3, 4, 5, y 7 calibraciones diarias, respectivamente.

Tabla 3.5: Resultados de algunos criterios de precisión analítica del A-GoldTM reportados en la literatura.

En contraste con el A-GoldTM, el ARP-SMCG tiene una mejor capacidad en el ajuste de la estimación de BG para minimizar el valor del RAD de esta estimación con respecto a BG capilar en los puntos de calibración. Por ello, el ARP-SMCG estima concentraciones de BG más próximas a las determinaciones de glucosa capilar que el A-GoldTM, en particular en el estado hipoglucémico. Esta característica permite estimar episodios hipoglucémicos en numerosos casos en los cuáles el A-GoldTM los omite. La evidencia que soporta esta conclusión se obtiene del análisis de los resultados descritos en la Tabla 3.2.

Finalmente, los resultados obtenidos realizando un análisis de la precisión desde el punto de vista clínico muestran que esta mejora en el ajuste de la estimación de BG por parte del ARP-SMCG, permiten obtener una mejor correspondencia de la glucemia estimada con la glucemia plasmática venosa, como se observa en la localización de los puntos que representan los pares de datos $VBG/G_{Proposed}$ tanto en Clarke EGA como en Consensus EGA.

3.3. Algoritmo para estimación de BG usando modelos ARs: Algoritmo basado en modelos auto-regresivos para el SMCG (ABMAR-SMCG)

La glucosa se difunde del plasma al ISF (y viceversa) a través de la pared capilar mediante difusión simple según el gradiente de concentración de glucosa (Figura 2.9). Este proceso introduce un retardo fisiológico entre ambas señales (Bondia et al., 2010). En el capítulo anterior, se indicó que este retardo es uno de los aspectos que influyen en los problemas de precisión

de los SMCGs (King et al., 2007). En términos generales, el retardo fisiológico entre IG y BG se considera de entre 5-15 minutos, aunque se ha observado mucha variabilidad (Bondia et al., 2010). Incluso, algún estudio ha reportado retrasos de hasta 40 minutos y otros estudios de hipoglucemia inducida mediante infusión de insulina han reportado adelantos, en los cuáles la caída de la IG precede a la caída de BG (Sternberg et al., 1996; Monsod et al., 2002).

Dado que considerar la relación entre IG y BG, especialmente en condiciones dinámicas, podría mejorar significativamente la precisión de los SMCGs, en esta investigación se busca desarrollar un algoritmo de estimación de BG que tenga en cuenta esta relación dinámica, en lugar de los tradicionales algoritmos de estimación de glucosa basados en técnicas de regresión lineal. Para ello se propone entrenar un modelo poblacional AR que modele la relación dinámica que existe entre IG y BG e incorporarlo a un algoritmo de estimación de glucosa. Las determinaciones de *VBG* y las muestras de corriente medidas por el sensor SC se utilizan para entrenar un modelo poblacional AR. En el entrenamiento del modelo poblacional AR se sigue una metodología basada en los planteamientos de Ljung (Ljung, 1999), sobre las etapas de identificación de un proceso. Este modelo AR se incorpora a un algoritmo para estimar las concentraciones de BG a partir de las muestras de corriente. Las glucemias calculadas por el modelo AR se corrigen cada vez que se introduce un nuevo punto de calibración, mediante las determinaciones de $G_{Capillary}$.

3.3.1. Conceptos básicos de modelos ARs

En el tratamiento de la diabetes, los métodos paramétricos clásicos, también conocidos como modelos ARs, se han utilizado como herramientas para predecir los niveles de glucosa durante cortos intervalos de tiempo, a partir de los datos de un SMCG. Varios autores han utilizado estos modelos para desarrollar modelos de predicción de glucosa específicos para cada individuo (Finan et al., 2007; Sparacino et al., 2007; Reifman et al., 2007; Eren-Oruklu et al., 2009) o poblacionales (Gani et al., 2009, 2010). Por ejemplo, Sparacino et al. utilizó un modelo AR de primer orden, el cuál continuamente adapta los coeficientes del modelo para predecir concentraciones de glucosa con un horizonte de predicción de hasta 30 minutos (Sparacino et al., 2007). Sin embargo, este modelo se debe adaptar continuamente para cada individuo. En un trabajo paralelo, un modelo AR de orden 10 con parámetros fijos se desarrolló en (Reifman et al., 2007). Sin embargo, este modelo mostró retardos relativamente grandes que reducen sus beneficios clínicos. En otro trabajo más reciente, los mismos autores utilizan datos de un paciente diabético para entrenar un modelo poblacional AR de orden 30 y posteriormente aplicarlo para predecir las concentraciones de glucosa de otros pacientes, incluso de aquellos con un tipo de diabetes diferente a la del paciente usado en el entrenamiento (Gani et al., 2009, 2010).

Los modelos ARs, permiten estimar los parámetros de los diferentes polinomios que integran la estructura de un modelo determinado, minimizando el error de predicción. Su representación se realiza mediante funciones de transferencia. De acuerdo al criterio a minimizar entre el modelo y el proceso real, la estimación de los parámetros se realiza usando técnicas en el dominio frecuencial (Kollár, 1993) o en el dominio temporal (Ljung, 1987).

La forma general de representar un modelo discreto es según (Ljung, 1987):

$$y(t) = G(q^{-1})u(t) + H(q^{-1})e(t) \tag{3.12}$$

donde $y(t)$ corresponde a la salida del modelo, $G(q^{-1})$ es un operador que realiza una representación de la dinámica del sistema (componente determinista o causal), $u(t)$ indica la entrada del sistema, $H(q^{-1})$ es un operador que representa el modelo del ruido (componente estocástica) y $e(t)$ es un término que incluye los errores de modelización y se asocia a una fuente de ruido blanco, formada por variables aleatorias independientes uniformemente distribuidas con media 0 y varianza λ . q^{-1} es el operador retardo, por lo tanto $[q^{-1} y(t) = y(t-1)]$. Por conveniencia, se puede usar una forma generalizada para representar toda la familia de modelos AR mediante la ecuación (3.13), la cuál se esquematiza en la Figura 3.6:

$$A(q^{-1}) y(t) = q^{-n_k} \frac{B(q^{-1})}{F(q^{-1})} u(t) + \frac{C(q^{-1})}{D(q^{-1})} e(t) \quad (3.13)$$

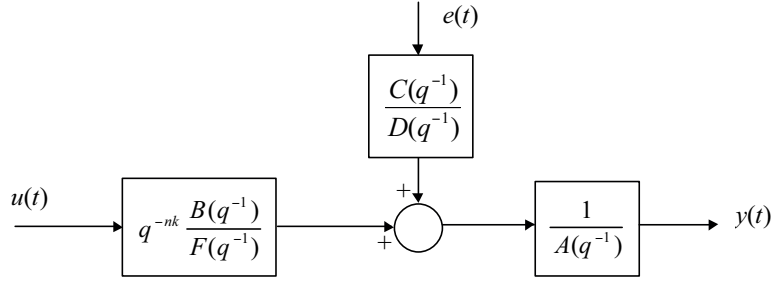


Figura 3.6: Esquema de bloques general de un modelo AR.⁶

La ecuación (3.13) puede dar lugar a 32 diferentes estructuras de modelos AR, dependiendo de cuáles de los cinco polinomios AR ($A(q^{-1})$, $B(q^{-1})$, $C(q^{-1})$, $D(q^{-1})$ y $F(q^{-1})$) se utilicen. Desde la ecuación (3.13) se definen los polinomios $A(q^{-1})$ (ecuación (3.14)), $B(q^{-1})$ (ecuación (3.15)), $C(q^{-1})$ (ecuación (3.16)), $D(q^{-1})$ (ecuación (3.17)) y $F(q^{-1})$ (ecuación (3.18)) de grados n_a , n_b , n_c , n_d y n_f , respectivamente, y n_k se define como el retardo puro del proceso.

$$A(q^{-1}) = 1 + a_1 q^{-1} + \dots + a_{n_a} q^{-n_a} \quad (3.14)$$

$$B(q^{-1}) = b_0 + b_1 q^{-1} + \dots + b_{n_b-1} q^{-(n_b-1)} \quad (3.15)$$

$$C(q^{-1}) = 1 + c_1 q^{-1} + \dots + c_{n_c} q^{-n_c} \quad (3.16)$$

$$D(q^{-1}) = 1 + d_1 q^{-1} + \dots + d_{n_d} q^{-n_d} \quad (3.17)$$

$$F(q^{-1}) = 1 + f_1 q^{-1} + \dots + f_{n_f} q^{-n_f} \quad (3.18)$$

En las ecuaciones (3.14), (3.15), (3.16), (3.17) y (3.18), los coeficientes a_i , b_i , c_i , d_i y f_i son desconocidos y se identifican usando regresión por mínimos cuadrados.

Las estructuras de modelos AR que se describen a continuación pertenecen a las estructuras más comúnmente usadas en la práctica (Ljung, 1999).

⁶Imagen tomada de (Ljung, 1999)

- Modelos autorregresivos con entrada exógena (-del inglés *AutoRegressive models with eXogeneous input (ARX models)*-): En esta estructura se considera que las componentes determinista y estocástica tienen el mismo denominador. Luego $F(q^{-1})$, $C(q^{-1})$ y $D(q^{-1})$ son iguales a uno (ecuación (3.19)).

$$y(t) = q^{-nk} \frac{B(q^{-1})}{A(q^{-1})} u(t) + \frac{1}{A(q^{-1})} e(t) \quad (3.19)$$

- Modelos autorregresivos de media móvil con entrada exógena (-del inglés *AutoRegressive Moving Average models with eXogeneous input (ARMAX models)*-): En esta estructura se extiende el modelo ARX agregando ceros particulares a la componente estocástica (ecuación (3.20)).

$$y(t) = q^{-nk} \frac{B(q^{-1})}{A(q^{-1})} u(t) + \frac{C(q^{-1})}{A(q^{-1})} e(t) \quad (3.20)$$

- Modelos del error en la salida (-del inglés *Output Error models (OE models)*-): En esta estructura se separa la dinámica de la entrada de la dinámica del error y se asume que la dinámica del error es uno (ecuación (3.21)):

$$y(t) = q^{-nk} \frac{B(q^{-1})}{F(q^{-1})} u(t) + e(t) \quad (3.21)$$

- Modelos Box-Jenkins (-del inglés *Box-Jenkins models (BJ models)*-): En esta estructura se separa la dinámica de la entrada de la dinámica del error, por tanto, las componentes determinista y estocástica no tienen parámetros comunes (ecuación (3.22)).

$$y(t) = q^{-nk} \frac{B(q^{-1})}{F(q^{-1})} u(t) + \frac{C(q^{-1})}{D(q^{-1})} e(t) \quad (3.22)$$

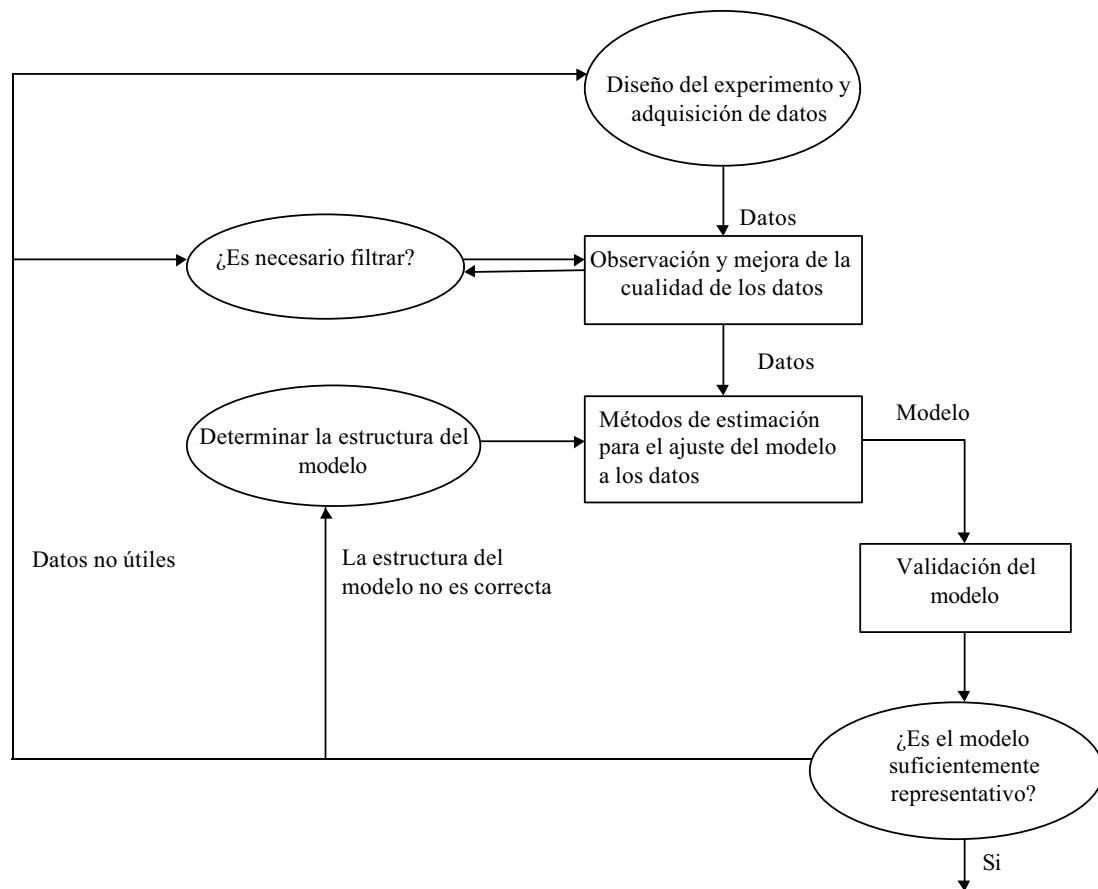
La Figura 3.7 describe las etapas a seguir para la identificación de un modelo paramétrico. Las decisiones indicadas en los rectángulos son principalmente responsabilidad del ordenador y las decisiones contenidas en los óvalos responsabilidad del usuario.

3.3.2. Pre-procesamiento de los datos

El objetivo es entrenar un modelo poblacional AR que modele la relación dinámica que existe entre IG y BG. Este modelo se ajusta utilizando como entrada la información suministrada por el sensor SC, obtenida a partir de la reacción electroquímica entre GOx e IG en el ISF, y como salida las concentraciones de referencia de VBG .

En la sección 3.2.1 se describió la información proporcionada por el CGMS[®] Gold[™] en el proceso de descarga de datos. De esta información, solamente *ISIG*, *Valid ISIG* y *VCTR* corresponden a información proporcionada por el sensor SC al monitor, al reaccionar la enzima GOx que lo recubre con la concentración de glucosa presente en el ISF. Las demás variables corresponden a información proporcionada por el algoritmo de calibración del CGMS[®] Gold[™]. Dado que se pretende desarrollar un nuevo algoritmo de estimación de BG, se descarta la información que corresponde al A-Gold[™]. Por tanto, se requiere determinar de la información obtenida desde

⁷Imagen tomada de (Ljung, 1999)

Figura 3.7: Ciclo de identificación de un modelo AR.⁷

el ISF, cuáles de estas variables son idóneas para ser consideradas como entradas del modelo AR.

Recordemos que *ISIG* y *Valid ISIG* contienen la misma información. *ISIG* contiene la corriente eléctrica que mide el sensor SC y *Valid ISIG* contiene la misma corriente adelantada 10 minutos para compensar el retardo fisiológico que existe entre IG y BG. Debido a que un modelo AR puede modelar la dinámica entre IG y BG, no es necesario considerar este retraso fisiológico y se puede trabajar directamente con *ISIG*.

Por lo tanto, descartada *Valid ISIG*, sólo queda por considerar *ISIG* y *VCTR*. Atendiendo al criterio de seleccionar como entrada la variable que presente la mejor correlación lineal con la concentración de BG de referencia, que en este caso es *VBG*, se calcula el coeficiente de correlación lineal de Pearson entre *VBG* e *ISIG* y entre *VBG* y *VCTR*. El mayor coeficiente de correlación lineal de Pearson se obtuvo entre *VBG* e *ISIG* ($r = 0.87$), en contraste con el valor obtenido entre *VBG* y *VCTR* ($r = 0.38$). Partiendo de este resultado se concluye que existe una mayor correlación lineal entre *VBG* e *ISIG*. Por tanto, se asume que la mejor excitación se puede obtener empleando *ISIG* como entrada del modelo AR.

En consecuencia, sea un instante de tiempo t , donde t representa el tiempo de muestreo del CGMS[®] Gold[™], se definen $ISIG(t)$ como la variable independiente o de entrada (-del inglés *input*-) y $VBG(t)$ como la variable dependiente o de salida (-del inglés *output*-) del proceso en

función de t . Por lo tanto, en el entrenamiento del modelo poblacional AR se utilizan pares de datos $ISIG(t)/VBG(t)$. Sin embargo, debido a que las determinaciones de VBG sólo se hicieron durante el primer día del estudio clínico, el modelo poblacional AR se entrena usando solamente la información proporcionada por el CGMS[®] Gold[™] durante el primer día de monitorización continua. Para obtener los pares de datos $ISIG(t)/VBG(t)$, se deben sincronizar las muestras de $ISIG(t)$ con las determinaciones simultáneas de $VBG(t)$. En este caso, se aplicó una estrategia similar a la que se utilizó en (Garg et al., 2009) para emparejar los datos. Dado que las muestras de $ISIG$ se obtuvieron en diferentes instantes de tiempo que las determinaciones de VBG , esta estrategia consiste en seleccionar el valor de $ISIG$ más cercano en el tiempo a la medida de VBG , dentro de un intervalo de tiempo de $\pm 2,5$ minutos.

En esta investigación, se utiliza un conjunto de datos de 18 pacientes (valor de $HbA1_c$ (%), 8,0 (7,3 – 9,9)) con DM-1 obtenidos del grupo de 21 enfermos incluidos en el estudio clínico descrito en la sección 3.1.1. Se excluyeron los datos de tres pacientes, debido a que disponían de muy pocas muestras de VBG que pudieran emparejarse con las muestras de $ISIG$ proporcionadas por el sensor. Para estos tres pacientes, se insertó el sensor SC en las horas de la tarde, cuando ya había transcurrido más de la mitad del periodo de estancia hospitalaria. De los 18 pacientes se obtienen 549 determinaciones de VBG . Las características clínicas de los 18 pacientes utilizados se presentan en la Tabla 3.6.

Variable	Valor
Número (%)	18 (100)
Edad (años)	28,5 (24,0 – 30,8)
Sexo (Femenino/%)	8/44,4
IMC (kg/m^2)	24,0 (22,3 – 26,0)
Presión sanguínea sistólica (mmHg)	124,5 (111,5 – 131,8)
Presión sanguínea diastólica (mmHg)	68,5 (61,3 – 78,5)
Colesterol total (mg/dL)	193,0 (156,8 – 212,8)
Triglicéridos (mg/dL)	86,0 (63,5 – 107,5)
HDL-C (mg/dL)	67,2 (52,0 – 72,2)
LDL-C (mg/dL)	101,5 (75,5 – 127,0)
Duración de la diabetes (años)	13,0 (10,0 – 19,3)
$HbA1_c$ (%)	8,0 (7,3 – 9,9)

Los datos son expresados utilizando la mediana (IQR).

Los números entre paréntesis indican el rango intercuartílico (IQR) (Q1 - Q3, respectivamente).

Tabla 3.6: Características basales y condiciones primarias de los pacientes con DM-1 del Hospital Josep Trueta utilizados en el desarrollo del ABMAR-SMCG.

Sin embargo, es importante recordar que las determinaciones de $VBG(t)$ se tomaron cada 15 y cada 30 minutos, a diferencia de las muestras de $ISIG(t)$, de las cuáles se tiene información cada cinco minutos. Además, las determinaciones de VBG sólo se recogieron durante el primer día del estudio clínico, por lo que su tamaño muestral es reducido, lo cuál constituye una limitación en el entrenamiento de un modelo poblacional AR. Por lo tanto, para aumentar la proporción de muestras $VBG(t)$ que se utilizaron en el entrenamiento del modelo poblacional AR, se aplicó un método de interpolación cúbica similar al que se describió en la sección 3.2.4 muestreando las determinaciones de VBG cada 5 minutos. De esta forma se obtiene un conjunto de datos que contiene 2206 pares de datos $ISIG(t)/VBG(t)$.

3.3.3. Entrenamiento del modelo AR: Identificación de la estructura del modelo

Para identificar la estructura del modelo poblacional, los datos de los pacientes se dividieron en dos conjuntos: un conjunto de entrenamiento - del inglés *training set* - (conformado por 16 pacientes) y un conjunto de prueba - del inglés *testing set* - (conformado por los dos pacientes restantes). Se aplica una estrategia de validación cruzada dejando dos fuera (conocida en inglés como *leave-two-out cross-validation*), la cuál implica separar los datos de forma que para cada iteración tengamos dos pacientes para el conjunto de prueba y el resto conformando el conjunto de entrenamiento; por ejemplo, en la primera iteración, el conjunto de entrenamiento se constituye por los datos de los pacientes No. 3, hasta el No. 18; por lo tanto, los pacientes No. 1 y No. 2 constituyen el conjunto de prueba. En la siguiente iteración, el paciente No. 2 pasa a formar parte del conjunto de entrenamiento, y el paciente No. 3 se excluye del conjunto de entrenamiento anterior para pasar al conjunto de prueba; y así sucesivamente hasta que se completan todas las posibles combinaciones entre el paciente No. 1 y los demás pacientes (17 iteraciones en total). Luego, se continua evaluando todas las combinaciones entre el paciente No. 2 y los demás pacientes en el conjunto de prueba, excepto la combinación pacientes No. 1 y No. 2, la cuál ya se probó, por lo tanto, en este caso se tienen 16 iteraciones en total. Este proceso iterativo se realiza hasta que se completan todas las combinaciones posibles en el conjunto de prueba (153 iteraciones en total).

En primer lugar, se realiza el entrenamiento de los diferentes modelos ARs utilizando los datos que constituyen el conjunto de entrenamiento. En el entrenamiento de estos modelos se considera como entrada $ISIG(t)$ y como salida $VBG(t)$, donde t representa el tiempo de muestreo del CGMS[®] Gold[™]. Las variables $ISIG(t)$ y $VBG(t)$ se normalizan dividiendo cada una de sus muestras entre 400 mg/dL y 100 nA, respectivamente. A estas variables normalizadas las hemos denominado $ISIG_N(t)$ (ecuación (3.23)) y $VBG_N(t)$ (ecuación (3.24)), respectivamente. La constante para normalizar $ISIG(t)$ se seleccionó a partir del máximo valor que puede tomar esta variable de acuerdo con las instrucciones del fabricante del sensor SC del CGMS[®] Gold[™]. La constante para normalizar $VBG(t)$ se seleccionó a partir del máximo valor que puede tomar esta variable de acuerdo con el A-Gold[™]. Luego, se obtienen los parámetros de los modelos ARs utilizando el conjunto de entrenamiento previamente normalizado.

$$ISIG_N(t) = \frac{ISIG(t)}{100 \text{ nA}} \quad (3.23)$$

$$VBG_N(t) = \frac{VBG(t)}{400 \text{ mg/dL}} \quad (3.24)$$

En el entrenamiento del modelo poblacional se utilizan cuatro estructuras de modelos ARs, las cuáles se han descrito en la sección 3.3.1, con el objetivo de encontrar la estructura que modele de la manera más precisa la concentración de BG, a partir de la corriente del sensor SC. Se entrenaron múltiples modelos ARX, ARMAX, OE y Box Jenkins considerando como grados de los polinomios ARs: $n_a = n_b = n_c = n_d = n_f = [1, 2 \dots 8]$, y como retardo puro del proceso: $n_k = 0$.

En segundo lugar, se realiza la validación de los diferentes modelos ARs obtenidos, utilizando los datos de los dos pacientes que no se incluyeron en el entrenamiento. Para ello, se normalizan los valores de $ISIG(t)$ del conjunto de prueba, dividiendo cada una de las muestras entre 100 nA, para obtener $ISIG_N(t)$. Luego, los parámetros de los modelos ARs obtenidos se utilizan para calcular las concentraciones de glucosa a partir de los valores de $ISIG_N(t)$. Partiendo de la

ecuación (3.13), la ecuación que describe estos modelos se puede escribir de manera generalizada como:

$$A(q^{-1}) BG_E(t) = q^{-n_k} \frac{B(q^{-1})}{F(q^{-1})} ISIG_N(t) + \frac{C(q^{-1})}{D(q^{-1})} e(t) \quad (3.25)$$

Los valores de $BG_E(t)$ corresponden a valores normalizados, por lo tanto, se multiplican por 400 mg/dL. Para cuantificar la precisión de las estimaciones de los modelos ARs se utiliza el criterio del mejor ajuste (conocido en inglés como *Best Fit*). *Best Fit* es una métrica que indica el porcentaje de la salida, es decir, el porcentaje de $VBG(t)$ que el modelo reproduce, y se define en la ecuación (3.26):

$$Best\ Fit = \left(1 - \frac{|VBG(t) - BG_E(t)|}{|VBG(t) - \overline{VBG(t)}|} \right) 100\% \quad (3.26)$$

donde $VBG(t)$ corresponde a las determinaciones de $VBG(t)$ de cada uno de los pacientes que se utilizan en la validación, $BG_E(t)$ indica los valores de glucosa estimadas por el respectivo modelo AR y $\overline{VBG(t)}$ es la media de $VBG(t)$ de cada uno de los pacientes que se emplean para validar. Para un ajuste perfecto, *Best Fit* debe ser igual a 100%. Si por el contrario, la glucosa estimada por el modelo ($BG_E(t)$) y la glucemia real ($VBG(t)$) no se corresponden, *Best Fit* será muy bajo.

La metodología explicada anteriormente se repite para las diferentes particiones de datos, es decir, combinaciones de 16 pacientes en el conjunto de entrenamiento y dos pacientes en el conjunto de prueba.

En la Tabla 3.7 se presentan los resultados de la mediana de la métrica *Best Fit* obtenidos de los experimentos realizados con las diferentes estructuras de modelos ARs abordadas, aplicando una estrategia de validación cruzada dejando dos fuera.

Estructura AR	n_a	n_b	n_c	n_d	n_f	<i>Best Fit</i> (%)
Modelos ARs considerando polinomios de grado 1						
ARX	1	1	-	-	-	16,5
ARMAX	1	1	1	-	-	30,8
OE	-	1	-	-	1	35,5
BJ	-	1	1	1	1	45,4
Modelos ARs considerando polinomios de grado 2						
ARX	2	2	-	-	-	19,8
ARMAX	2	2	2	-	-	35,0
OE	-	2	-	-	2	39,2
BJ	-	2	2	2	2	50,5
Modelos ARs considerando polinomios de grado 3						
ARX	3	3	-	-	-	58,0
ARMAX	3	3	3	-	-	77,0
OE	-	3	-	-	3	74,7
BJ	-	3	3	3	3	83,1
Modelos ARs considerando polinomios de grado 4						
ARX	4	4	-	-	-	60,3
ARMAX	4	4	4	-	-	78,5
OE	-	4	-	-	4	76,0
BJ	-	4	4	4	4	84,9

Tabla 3.7: Resultados de la mediana de la métrica *Best Fit* obtenidos aplicando una estrategia de validación cruzada dejando dos fuera.

A partir de estos resultados se puede llegar a las siguientes conclusiones:

- Los resultados de la métrica *Best Fit* indican que de todas las estructuras evaluadas, la estructura AR de más bajo orden que describe mejor la dinámica del sistema es una estructura Box-Jenkins (BJ) con polinomios de grado tres (*Best Fit* de 83,1%) con una tasa de muestreo de cinco minutos ($n_b = n_c = n_d = n_f = 3$). En una estructura BJ (ecuación (3.22)), una señal de salida $y(t)$, en este caso representada como $BG_E(t)$, se describe como una combinación lineal de señales observadas previamente, donde la componente determinista, denotada por $B(q^{-1})$ y $F(q^{-1})$, y la componente estocástica, denotada por $C(q^{-1})$ y $D(q^{-1})$, no comparten coeficientes.
- Las estructuras polinomiales de orden inferior a tercer orden no tienen la capacidad de modelar la relación dinámica que existe entre IG y BG. Al considerar polinomios de grado uno y dos, los modelos ARX indicaban un mal modelado entre $BG_E(t)$ y $VBG(t)$, con valores de *Best Fit* inferiores a 20,0%, presentándose la mayor diferencia en los modelos de primer orden. Con las estructuras ARMAX y OE se reduce el error y *Best Fit* presenta valores entre 30,8% y 39,2%. En el caso de la estructura BJ se incrementa el valor de la métrica *Best Fit* a 45,4% y 50,5%, respectivamente.
- Los modelos más apropiados son los que consideran polinomios de grado tres. Al considerar estructuras ARs (ARX, ARMAX, OE y BJ) de tercer orden, estas estructuras indican una mayor precisión y buen ajuste entre $BG_E(t)$ y $VBG(t)$. Estos modelos son estables y al cuantificar la precisión de la estimación, se obtienen valores de *Best Fit* entre 58,0% y 83,1%, proporcionando las estructuras ARMAX y BJ los ajustes más altos, en comparación con ARX y OE.
- Aunque los modelos ARs de grado cuatro reproducían en forma más precisa $VBG(t)$ (*Best Fit* entre 60,3% y 84,9%), su problema consistía en que varios de ellos eran inestables. Además, las diferencias no son relevantes con respecto a las estructuras de tercer orden y se incrementa la complejidad del modelo.
- Se hicieron algunas pruebas con modelos de grado ocho e incluso con estructuras de orden superior, buscando optimizar la precisión, pero en la mayor parte de los modelos obtenidos se observaron problemas de inestabilidad.

Una vez identificada la estructura del modelo poblacional como una estructura BJ con polinomios de grado tres ($n_b = n_c = n_d = n_f = 3$), se repite el entrenamiento aplicando una estrategia de validación cruzada dejando dos fuera. De todas las 153 posibles particiones probadas, se obtienen los coeficientes de 153 modelos BJ con polinomios de grado tres. Para cuantificar la precisión de las estimaciones de los modelos BJ se utiliza la métrica *Best Fit*, tal como se hizo previamente (ecuación (3.26)). En la Tabla 3.8 se presentan los resultados de la mediana de dicha métrica en cada uno de los pacientes con DM-1 utilizados en esta investigación, obtenidos al aplicar una estrategia de validación cruzada dejando dos fuera.

Los coeficientes del modelo poblacional AR propuesto se computan como los coeficientes promedios de los modelos BJ. En la estimación de los parámetros del modelo poblacional AR, se descartaron cinco de estos modelos BJ porque eran inestables, por tanto, sólo se usaron los restantes 148 modelos BJ. El modelo poblacional AR está definido por la ecuación (3.27), donde $G_{AR}(t)$ representa la concentración de VBG estimada por el modelo poblacional, $ISIG_N(t)$ es el valor de la corriente suministrada por el sensor SC de glucosa previamente normalizada, $e(t)$ representa una fuente de ruido blanco gaussiano con media 0 y una pequeña varianza arbitraria

$\lambda = 3,37 \cdot 10^{-11}$ (Gani et al., 2009) y $B(q^{-1})$, $C(q^{-1})$, $D(q^{-1})$ y $F(q^{-1})$ denotan polinomios ARs de tercer orden, definidos por las ecuaciones (3.28), (3.29), (3.30) y (3.31), respectivamente. Los coeficientes b_i , c_i , d_i y f_i se presentan en la Tabla 3.9. Los valores de $G_{AR}(t)$ corresponden a valores normalizados, por lo tanto, se multiplican por 400 mg/dL.

$$G_{AR}(t) = q^{-n_k} \frac{B(q^{-1})}{F(q^{-1})} ISIG_N(t) + \frac{C(q^{-1})}{D(q^{-1})} e(t) \quad (3.27)$$

$$B(q^{-1}) = b_0 + b_1q^{-1} + b_2q^{-2} \quad (3.28)$$

$$C(q^{-1}) = 1 + c_1q^{-1} + c_2q^{-2} + c_3q^{-3} \quad (3.29)$$

$$D(q^{-1}) = 1 + d_1q^{-1} + d_2q^{-2} + d_3q^{-3} \quad (3.30)$$

$$F(q^{-1}) = 1 + f_1q^{-1} + f_2q^{-2} + f_3q^{-3} \quad (3.31)$$

Paciente No.	Best Fit (%)
1	65,9
2	75,9
3	51,9
4	45,4
5	74,8
6	47,4
7	66,3
8	81,0
9	84,3
10	71,7
11	64,5
12	44,9
13	90,4
14	83,2
15	83,3
16	80,6
17	74,5
18	71,5
Mediana (IQR)	73,1 (64,9 – 80,9)

Tabla 3.8: Resultados de la mediana de la métrica *Best Fit* en cada uno de los pacientes con DM-1, obtenidos con una estructura BJ con polinomios de tercer orden ($n_b = n_c = n_d = n_f = 3$) aplicando una estrategia de validación cruzada dejando dos fuera.

3.3.4. ABMAR-SMCG

El modelo poblacional AR definido por la ecuación (3.27) se utiliza para desarrollar un algoritmo que calcule el valor de la BG desde la corriente medida por el sensor SC. Por tanto, la entrada del algoritmo de estimación de BG es el valor de $ISIG(t)$ medido por el sensor SC del CGMS[®] Gold[™]. Este valor de $ISIG(t)$ se normaliza previamente para obtener $ISIG_N(t)$ (ecuación 3.23). Las mismas determinaciones de $G_{Capillary}$ usadas para calibrar el A-Gold[™] se utilizan como puntos de calibración en este algoritmo. Para ser introducidas al ABMAR-SMCG,

Polinomios ARs del modelo	Parámetros de los polinomios		
	b_0	b_1	b_2
B	1,465452	-1,786699	0,334811
F	f_1	f_2	f_3
	-1,204988	0,188096	0,024471
C	c_1	c_2	c_3
	-0,137914	-0,010064	0,008919
D	d_1	d_2	d_3
	-1,605645	0,700240	-0,083482

Tabla 3.9: Parámetros del modelo poblacional BJ de tercer orden.

estas determinaciones de $G_{Capillary}$ se normalizan dividiendo entre 400 mg/dL (ecuación (3.32)). Estas determinaciones normalizadas se definen como $G_C(t)$.

$$G_C(t) = \frac{G_{Capillary}(t)}{400 \text{ mg/dL}} \quad (3.32)$$

A continuación se describe el proceso que realiza el ABMAR-SMCG para estimar la concentración de BG cada cinco minutos:

1. En el instante inicial ($t = 0$), se introduce el valor de $G_{Capillary}$ que se utiliza como primer punto de calibración. Luego se calculan $G_C(t)$ (ecuación (3.32)) e $ISIG_N(t)$ (ecuación 3.23), y se asume $G_{AR}(t) = G_C(t)$.

2. Para la siguiente muestra ($t+1$), $G_{AR}(t+1)$ se calcula como la salida del modelo poblacional AR definido por la ecuación (3.27). Debido a que este modelo corresponde a una estructura AR de tercer orden, para calcular $G_{AR}(t+1)$ se requiere los valores de las tres muestras previas de $G_{AR}(t+1)$, es decir, $G_{AR}(t)$, $G_{AR}(t-1)$ y $G_{AR}(t-2)$, así como el valor de $ISIG_N(t+1)$ y los valores de sus dos muestras previas, es decir, $ISIG_N(t)$ e $ISIG_N(t-1)$. El valor de $G_{AR}(t)$ es conocido, pero como no se conoce el valor de la glucemia antes de introducir el primer punto de calibración, se asume $G_{AR}(t-1) = G_{AR}(t-2) = G_C(t)$. Los valores de $ISIG_N(t+1)$ e $ISIG_N(t)$ son conocidos. $ISIG_N(t-1)$ también es un valor conocido, pero es un valor de la corriente antes de introducir el primer punto de calibración, y no se considera un valor fiable debido a que fue medido durante el proceso de inicialización del sensor SC, cuando este aún no se había estabilizado. Por esta razón, se asume $ISIG_N(t-1) = ISIG_N(t)$.

3. Para la siguiente muestra ($t+2$), $G_{AR}(t+2)$ nuevamente se calcula como la salida del modelo poblacional AR definido por la ecuación (3.27). En este caso, para calcular $G_{AR}(t+2)$ se requiere de $G_{AR}(t+1)$, $G_{AR}(t)$ y $G_{AR}(t-1)$, así como de $ISIG_N(t+2)$, $ISIG_N(t+1)$ e $ISIG_N(t)$. Aquí sólo se asume $G_{AR}(t-1) = G_C(t)$. Los demás valores son conocidos.

4. Para las siguientes muestras tomadas hasta que se retira el sensor SC, todos los valores son conocidos y los valores de $G_{AR}(t)$ se calculan como la salida del modelo poblacional AR definido por la ecuación (3.27).

5. Para las siguientes muestras, cuando se introduce un nuevo punto de calibración, el ABMAR-SMCG reemplaza el valor de $G_{AR}(t)$ calculado por el modelo AR por el valor de $G_C(t)$. Adicionalmente, los dos instantes previos a la introducción del punto de calibración $G_{AR}(t-1)$ y $G_{AR}(t-2)$ también se reemplazan por este nuevo valor de $G_C(t)$. Esta consideración se uti-

liza para evitar drásticas diferencias entre el punto de calibración y las estimación del modelo poblacional AR.

6. Los valores de $G_{AR}(t)$ corresponden a valores normalizados, por lo tanto, se retornan a su rango original multiplicando por 400 mg/dL.

3.3.5. Evaluación del ABMAR-SMCG usando criterios analíticos

En la sección 3.3.2 se explicó que durante la recolección de muestras de glucosa plasmática en el hospital, de los 18 pacientes que se utilizan en esta investigación se obtuvieron 549 determinaciones de VBG , las cuáles se utilizan en la evaluación del ABMAR-SMCG usando criterios analíticos. Las determinaciones de $G_{Capillary}$, las cuáles si se recogieron durante los tres días de monitorización, no se pueden considerar como medidas de referencia en la evaluación, porque cada vez que se introduce un punto de calibración en el ABMAR-SMCG, este sustituye el valor de la glucosa estimada por el modelo AR por el valor de $G_{Capillary}$. Por lo tanto, si se considera $G_{Capillary}$ como la medida de referencia, no se estaría realizando una evaluación objetiva, porque el valor de la muestra de G_{AR} sería el mismo que el valor de la medida de referencia y los errores serían nulos. Por las razones expuestas anteriormente, sólo es posible realizar el cálculo de los criterios de precisión analíticos para el primer día del estudio clínico, debido a que no hay muestras disponibles que puedan ser consideradas como medidas de referencia para los restantes dos días.

En esta evaluación, las muestras de VBG se comparan con las correspondientes lecturas del A-GoldTM y del ABMAR-SMCG, con el fin de comprobar el rendimiento de ambos algoritmos. Solamente las glucemias estimadas por ambos algoritmos (G_{Gold} y G_{AR}) emparejadas con VBG se emplean en el cálculo de los criterios analíticos, debido a que las medidas de VBG no estaban disponibles para el resto de las muestras. Para sincronizar las determinaciones de G_{Gold} y G_{AR} con las determinaciones de VBG , se aplica una estrategia similar a la que se utilizó en (Garg et al., 2009) para emparejar los datos. Dado que las muestras de G_{Gold} y G_{AR} se obtuvieron en diferentes instantes de tiempo que las determinaciones de VBG , esta estrategia consiste en seleccionar los valores de G_{Gold} y G_{AR} más cercanos en el tiempo a la medida de VBG , dentro de un intervalo de tiempo de $\pm 2,5$ minutos. Las muestras de G_{Gold} y G_{AR} si se obtuvieron de manera simultánea.

La evaluación de la precisión de las estimaciones de ambos algoritmos utilizando algunos de los criterios analíticos descritos en la sección 2.4.1, se presenta en la Tabla 3.10 considerando tanto la población global, como los rangos glucémicos de los pacientes con DM-1. En este caso, los rangos glucémicos también se han definido siguiendo los criterios del personal médico del Hospital Josep Trueta, pero atendiendo a sus recomendaciones cuando se considera como método de referencia la determinación de BG venosa. Estos rangos se corresponden con los definidos en (Weinstein et al., 2007), donde también se considera la determinación de BG venosa como el “gold standard”. Los rangos glucémicos se han definido como: hipoglucémico ($VBG \leq 70$ mg/dL), normoglucémico leve (VBG entre 71 mg/dL – 140 mg/dL), normoglucémico moderado (VBG entre 141 mg/dL – 180 mg/dL), hiperglucémico moderado (VBG entre 181 mg/dL – 240 mg/dL) e hiperglucémico severo ($VBG > 240$ mg/dL). A diferencia de como se definieron los rangos glucémicos en la sección 3.2.3, en este caso la definición de los rangos hipoglucémicos e hiperglucémicos es menos estricta.

En primer lugar, se determina la concordancia entre los pares de datos que se comparan, evaluando la diferencia entre VBG y G_{Gold} y entre VBG y G_{AR} . Previamente se aplica el Test de Shapiro-Wilk (Shapiro, 1965) para determinar si los pares de datos siguen o no una distribución normal. Debido a que los pares de datos no siguen una distribución normal, se calcula el valor de *Median RAD (IQR)* (Tabla 3.10). El Test de suma de rangos de Wilcoxon (Wilcoxon, 1945) se utiliza para comparar el valor de *Median RAD (IQR)* entre el A-GoldTM y el ABMAR-SMCG y determinar si existen diferencias estadísticamente significativas entre ellos. Al comparar los resultados de ambos algoritmos, se observa un contraste estadísticamente significativo (*Median RAD (IQR)* = 11,3 (5,6 – 23,3) % versus 6,7 (2,9 – 13,8) %, p -valor < 0,05, respectivamente).

Los resultados de la evaluación considerando el estado glucémico del enfermo con DM-1 (83 muestras de $VBG \leq 70$ mg/dL, 197 muestras de VBG entre 71 mg/dL – 140 mg/dL, 84 muestras de VBG entre 141 mg/dL – 180 mg/dL, 104 muestras de VBG entre 181 mg/dL – 240 mg/dL y 81 muestras de $VBG > 240$ mg/dL) también se presentan en la Tabla 3.10. Estos resultados indican diferencias significativas en todos los rangos glucémicos, excepto en el rango hiperglucémico severo. Es importante mencionar que las diferencias más relevantes se presentan en el rango hipoglucémico (*Median RAD (IQR)* = 15,0 (10,0 – 26,0) % versus 6,0 (3,0 – 12,0) %, p -valor < 0,05, respectivamente). Estos resultados también muestran que los dos algoritmos son igualmente precisos cuando los pacientes se encuentran en estado hiperglucémico severo (p -valor > 0,05). Aquí también se aplica un Test de suma de rangos de Wilcoxon el cuál permite comparar el valor de *Median RAD (IQR)* entre los dos algoritmos.

Otro criterio de precisión analítico que se considera son los criterios ISO. La Tabla 3.10 muestra el porcentaje de pares de datos que satisfacen estos criterios. En el conjunto global de datos, con el ABMAR-SMCG se obtiene un incremento en el porcentaje de los datos que satisfacen estos criterios con respecto al reportado para el A-GoldTM (Criterios ISO = 88,7% versus 73,0%). Los resultados obtenidos corroboran que el ABMAR-SMCG presenta mejor precisión analítica en el rango hipoglucémico, en comparación con el rendimiento presentado por el A-GoldTM (Tabla 3.10). En los demás rangos glucémicos, estos resultados también son favorables para el ABMAR-SMCG, tan sólo en el rango hiperglucémico severo la precisión de los dos algoritmos es similar, en los demás rangos la precisión es superior.

Finalmente, en la Tabla 3.10 también se muestran los resultados de calcular los criterios de concordancia⁸, los cuáles se describieron en la sección 2.4.1. Estos resultados también reportan diferencias relevantes entre los dos algoritmos y en ellos se observa que el ABMAR-SMCG es más preciso que el A-GoldTM, excepto en el rango hiperglucémico severo, en el cuál no se aprecian diferencias significativas entre los dos algoritmos.

3.3.6. Evaluación del ABMAR-SMCG usando criterios clínicos

En la evaluación de la precisión de los algoritmos utilizando criterios clínicos se utilizan como medidas de referencia las mismas determinaciones de VBG usadas en la evaluación desde el punto de vista analítico (sección 3.3.5). Por lo tanto, también se utilizan pares de datos VBG/G_{Gold} y VBG/G_{AR} y los criterios de precisión clínicos sólo se calculan para el primer día del estudio.

⁸Para $VBG \leq 70$ mg/dL, se calculan los criterios de concordancia dentro de $\pm 20,0$ mg/dL y no dentro del $\pm 20,0$ % y del $\pm 30,0$ % de la medida de referencia, porque en el rango hipoglucémico no se puede considerar como buena una medida de glucemia con una diferencia tan amplia de la medida de referencia.

Rango de la glucemia (mg/dL)	<i>Median RAD (IQR) (%)</i>		<i>Criterios ISO (%)</i>		<i>Medidas dentro de ±20% de la referencia (%)</i>		<i>Medidas dentro de ±30% de la referencia (%)</i>	
	A-Gold™	ABMAR-SMCG	A-Gold™	ABMAR-SMCG	A-Gold™	ABMAR-SMCG	A-Gold™	ABMAR-SMCG
≤70 ^a	15,0 (9,0 – 26,0)	6,0 (3,5 – 13,5)	55,4	86,7	66,3	95,2	–	–
71 – 140	11,1 (6,2 – 23,4)	6,9 (2,7 – 15,9)	69,0	83,2	69,0	83,2	82,7	95,4
141 – 180	9,0 (4,7 – 14,8)	5,5 (2,8 – 9,0)	82,1	90,5	82,1	90,5	88,1	96,4
181 – 240	11,1 (4,3 – 20,5)	5,3 (2,4 – 11,9)	74,0	94,2	74,0	94,2	83,7	99,0
>240	8,7 (4,9 – 13,4)	6,6 (2,9 – 10,3)	90,1	95,1	90,1	95,1	93,8	100,0

^aPara $VBG \leq 70$ mg/dL, se calcula *Median AD* (mg/dL).

Tabla 3.10: Resultados de evaluar la precisión analítica del A-Gold™ y del ABMAR-SMCG, usando pares de datos VBG/G_{Gold} y VBG/G_{AR} obtenidos de todos los pacientes con DM-1 de acuerdo con el estado glucémico.

Rango de la glucemia (mg/dL)	A (%)		B (%)		C (%)		D (%)		E (%)	
	A-Gold™	ABMAR-SMCG	A-Gold™	ABMAR-SMCG	A-Gold™	ABMAR-SMCG	A-Gold™	ABMAR-SMCG	A-Gold™	ABMAR-SMCG
<i>Clarke EGA</i>										
≤70	57,8	89,2	1,2	1,2	0,0	0,0	41,0	9,6	0,0	0,0
71 – 140	73,1	84,8	26,9	15,2	0,0	0,0	0,0	0,0	0,0	0,0
141 – 180	88,1	94,0	11,9	6,0	0,0	0,0	0,0	0,0	0,0	0,0
181 – 240	79,8	96,2	18,3	3,8	0,0	0,0	0,0	0,0	1,9	0,0
>240	90,1	97,5	8,6	2,5	0,0	0,0	1,3	0,0	0,0	0,0
<i>Consensus EGA</i>										
≤70	68,6	94,6	14,5	2,4	16,9	3,0	0,0	0,0	0,0	0,0
71 – 140	76,5	87,8	22,8	11,7	0,7	0,5	0,0	0,0	0,0	0,0
141 – 180	83,3	91,7	16,7	8,3	0,0	0,0	0,0	0,0	0,0	0,0
181 – 240	79,8	95,2	17,3	4,8	2,9	0,0	0,0	0,0	0,0	0,0
>240	90,1	100,0	9,9	0,0	0,0	0,0	0,0	0,0	0,0	0,0

Tabla 3.11: Resultados de evaluar la precisión clínica del A-Gold™ y del ABMAR-SMCG, usando los pares de datos VBG/G_{Gold} y VBG/G_{AR} obtenidos de todos los pacientes con DM-1 de acuerdo con el estado glucémico.

Nuevamente se utilizan las rejillas de errores de Clarke EGA y Consensus EGA, para determinar la relevancia clínica de las diferencias existentes entre las muestras de VBG y las de G_{Gold} y G_{AR} (Tabla 3.11).

Esta evaluación se realiza considerando tanto la población global de pacientes con DM-1, como el estado glucémico del paciente. Los resultados obtenidos en las dos rejillas de errores evidencian que los criterios de precisión clínicos son más altos con el ABMAR-SMCG. Analizando los resultados de Clarke EGA, con el ABMAR-SMCG se incrementa la distribución de puntos en la zona A y se reduce en las zonas B, D y E. De la misma forma, analizando los resultados de Consensus EGA, con el ABMAR-SMCG se aumenta la concentración de puntos en la zona A y se disminuye en las zonas B y C.

La evaluación desde el punto de vista clínico en los diferentes rangos glucémicos también confirma que el ABMAR-SMCG reporta valores glucémicos más precisos que el A-GoldTM. Particularmente en el rango hipoglucémico, la precisión mejora notablemente con el algoritmo que se propone (Tabla 3.11). En los demás rangos glucémicos también se observan mejoras importantes.

En la Figura 3.8 se ilustra la distribución de los puntos que representan los pares de datos VBG/G_{AR} y VBG/G_{Gold} dentro de cada zona clínica de Clarke EGA y de Consensus EGA, tanto del ABMAR-SMCG (Figuras 3.8(a) y 3.8(c), respectivamente), como del A-GoldTM (Figuras 3.8(b) y 3.8(d), respectivamente).

3.3.7. Análisis de la capacidad de estimación de hipoglucemias e hiperglucemias del ABMAR-SMCG

Con el objetivo de comparar la capacidad del ABMAR-SMCG y del A-GoldTM para estimar eventos hipoglucémicos e hiperglucémicos correctamente, se obtienen los perfiles glucémicos de los 18 pacientes con DM-1. En las Figuras 3.9, 3.10(a), 3.10(b), 3.11(a) y 3.11(b) se ilustran como ejemplos los perfiles glucémicos de cinco pacientes. En estas gráficas se representan los perfiles glucémicos proporcionados por los dos algoritmos a comparar, así como las determinaciones de VBG y los puntos de calibración. La línea roja continua representa las determinaciones de VBG , la línea verde punteada representa las muestras de G_{AR} , la línea azul punteada corresponde a las muestras de G_{Gold} y los rombos magenta representan las determinaciones de $G_{Capillary}$. El umbral para definir una hipoglucemia ($VBG \leq 70$ mg/dL) se representa por una línea cian punteada y el umbral para definir una hiperglucemia severa ($VBG > 240$ mg/dL) se representa por una línea amarilla punteada. Adicionalmente, para cada uno de estos pacientes también se indican los resultados de la evaluación aplicando *Median RAD (IQR)*, Clarke EGA y Consensus EGA, considerando VBG como la medida de referencia. En estos perfiles sólo se muestra el comportamiento de los algoritmos durante la estancia hospitalaria de los pacientes, debido a que no se dispone de determinaciones de referencia para el periodo de monitorización domiciliaria.

En el perfil que se ilustra en la Figura 3.9, se evalúa la capacidad del ABMAR-SMCG para estimar eventos hipoglucémicos. Recordemos que en este caso el umbral para definir un evento hipoglucémico es $VBG \leq 70$ mg/dL. Los episodios hipoglucémicos, los cuales son omitidos por el A-GoldTM, son estimados correctamente por el ABMAR-SMCG. De hecho, si se compara este perfil glucémico con el perfil ilustrado en la Figura 3.3, se observa por parte del ABMAR-SMCG una mejor capacidad para estimar concentraciones de glucosa más próximas a VBG . Considerando VBG como el *gold standard*, el valor de *Median RAD (IQR)* del ABMAR-SMCG es 5,2

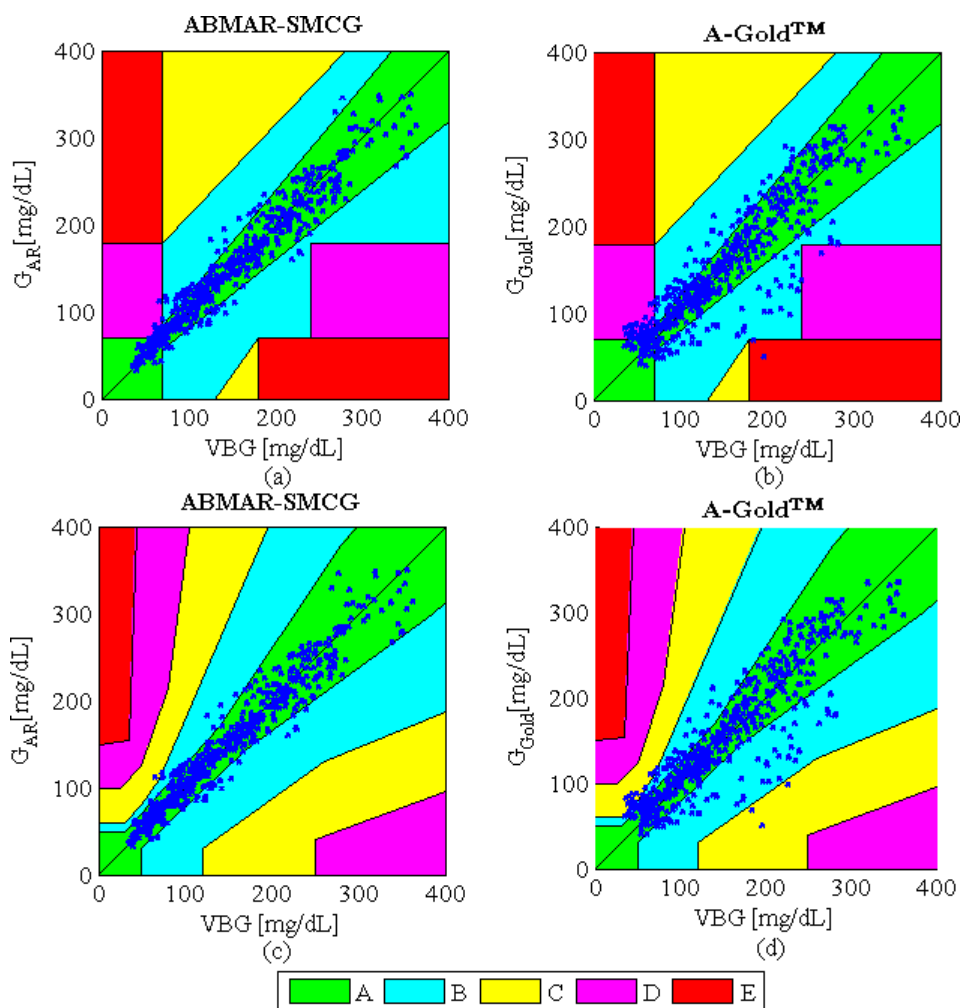


Figura 3.8: Rejillas de Clarke EGA y Consensus EGA para evaluar el ABMAR-SMCG y el A-Gold™, usando los pares de datos VBG/G_{AR} y VBG/G_{Gold} obtenidos de todos los pacientes con DM-1.

(3,9) %, mientras el del A-Gold™ es 7,9 (24,0) %. Analizando los resultados obtenidos de Clarke EGA y Consensus EGA, se evidencia una mayor precisión del ABMAR-SMCG. Clarke EGA clasifica el 100,0 % de los puntos que representan VBG/G_{AR} en la zona A versus el 70,0 % de VBG/G_{Gold} . Sin embargo, lo más importante es que no se reporta ningún punto en zona D con el ABMAR-SMCG, a diferencia del A-Gold™ que reporta 3,3 %. Consensus EGA clasifica el 100,0 % de los puntos que representan VBG/G_{AR} en la zona A versus el 73,3 % de VBG/G_{Gold} . En este caso las reducciones más relevantes se dieron en las zonas B y C.

Adicionalmente, en las Figuras 3.10(a) y 3.10(b), también se muestran otros episodios hipoglucémicos estimados de manera correcta por el ABMAR-SMCG. Para el paciente de la Figura 3.10(a), el valor de *Median RAD (IQR)* del ABMAR-SMCG es 4,1 (6,7) %, mientras el del A-Gold™ es 11,5 (12,3) %. Clarke EGA clasifica el 84,6 % de los pares de datos VBG/G_{AR} en la zona A versus el 80,8 % de VBG/G_{Gold} . Consensus EGA clasifica el 96,2 % de los puntos VBG/G_{AR} en la zona A versus el 92,3 % de VBG/G_{Gold} . Por otra parte, para el paciente de la Figura 3.10(b), el valor de *Median RAD (IQR)* del ABMAR-SMCG es 5,0 (5,0) %, mientras el

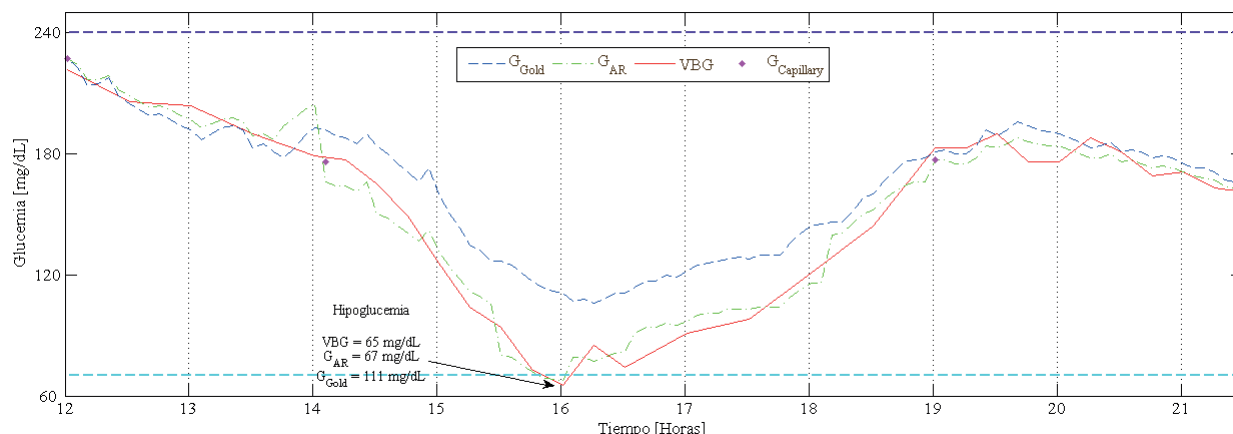


Figura 3.9: Ejemplos de perfiles glucémicos hospitalarios estimados por el A-Gold™ y el ABMAR-SMCG: Estimación de episodios hipoglucémicos.

del A-Gold™ es 4,2 (9,5) %. En este caso, si bien con el ABMAR-SMCG se estiman correctamente episodios hipoglucémicos omitidos, y los dos algoritmos reportan valores de *Median RAD (IQR)* muy bajos, desde el punto de vista analítico es más preciso el A-Gold™. Analizando los resultados obtenidos desde el punto de vista clínico, se aprecia un mejor rendimiento en la estimación de hipoglucemias por parte del ABMAR-SMCG. Clarke EGA clasifica el 97,0 % de los puntos VBG/G_{AR} en la zona A versus el 90,9 % de VBG/G_{Gold} . En este caso, lo más relevante es que no existen puntos en la zona D (el A-Gold™ reporta 9,1 %). Consensus EGA no muestra diferencias entre los dos algoritmos, reportando para ambos 100,0 % de los puntos en la zona A. Esta diferencia entre Clarke EGA y Consensus EGA se debe a la fuerte restricción que impone Clarke EGA en el rango hipoglucémico, la cuál posibilita pasar directamente de zona D a zona A.

Por otra parte, en el perfil que se ilustra en la Figura 3.11(a), se evalúa la capacidad del ABMAR-SMCG para estimar eventos hiperglucémicos. Para este paciente, el A-Gold™ muestra falsos eventos hipoglucémicos, estimando glucemias menores a 50 mg/dL, cuando las determinaciones de VBG realmente indican eventos hiperglucémicos. Aunque el ABMAR-SMCG no reporta este estado hiperglucémico, si proporciona glucemias más cercanas a VBG . El valor de *Median RAD (IQR)* del ABMAR-SMCG es 4,2 (7,9) %, mientras el del A-Gold™ es 18,5 (22,6) %. Clarke EGA clasifica el 94,1 % de los puntos VBG/G_{AR} en la zona A, en comparación con el 73,5 % de VBG/G_{Gold} . Además, con el ABMAR-SMCG no se clasifica ningún punto en la zona E, a diferencia del A-Gold™, con el cuál 5,9 % de los puntos están en dicha zona. Consensus EGA reporta idénticos resultados a los de Clarke EGA para ambos algoritmos en zona A. Aquí la reducción significativa se dio en la zona C, porque allí no se reportó ningún punto VBG/G_{AR} , en comparación con el 5,9 % de VBG/G_{Gold} .

Finalmente, en el perfil que se ilustra en la Figura 3.11(b), al comparar G_{AR} con G_{Gold} , se observa que el ABMAR-SMCG estima glucemias más cercanas a VBG . El valor de *Median RAD (IQR)* para el ABMAR-SMCG es 11,9 (21,2) %, mientras el del A-Gold™ es 14,1 (21,4) %. En este caso, los resultados de Clarke EGA reportan el mismo porcentaje de datos en zona A para ambos algoritmos, es decir, 66,7%; mientras los resultados de Consensus EGA reportan 72,7 % en zona A del ABMAR-SMCG y 66,7 % en zona A del A-Gold™.

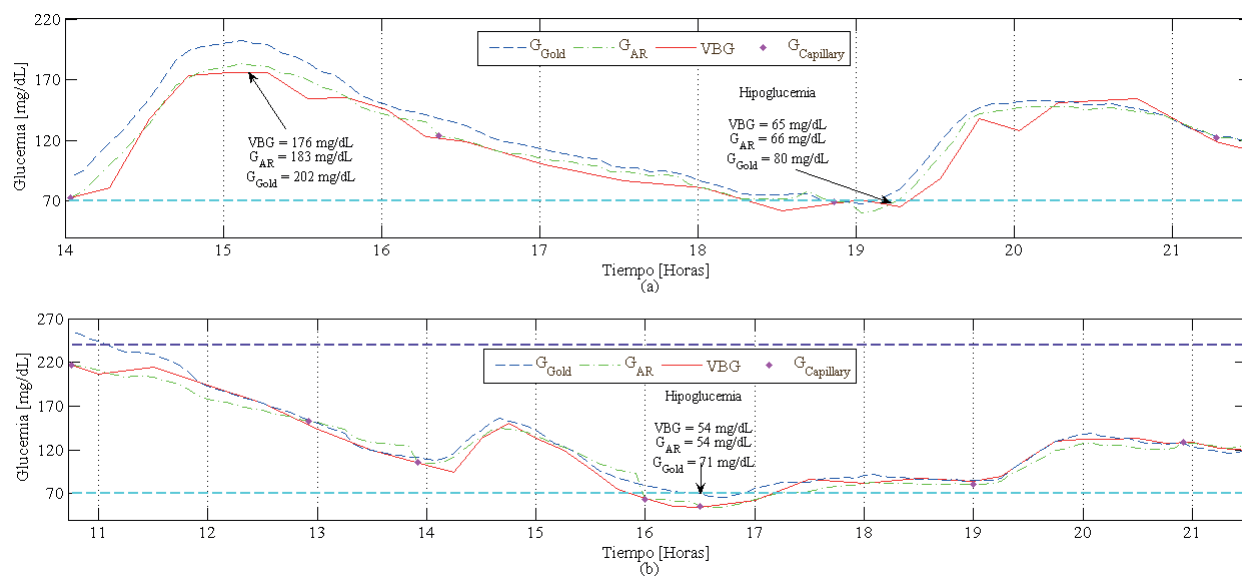


Figura 3.10: Otros ejemplos de perfiles glucémicos hospitalarios estimados por el A-GoldTM y el ABMAR-SMCG: Estimación de episodios hipoglucémicos.

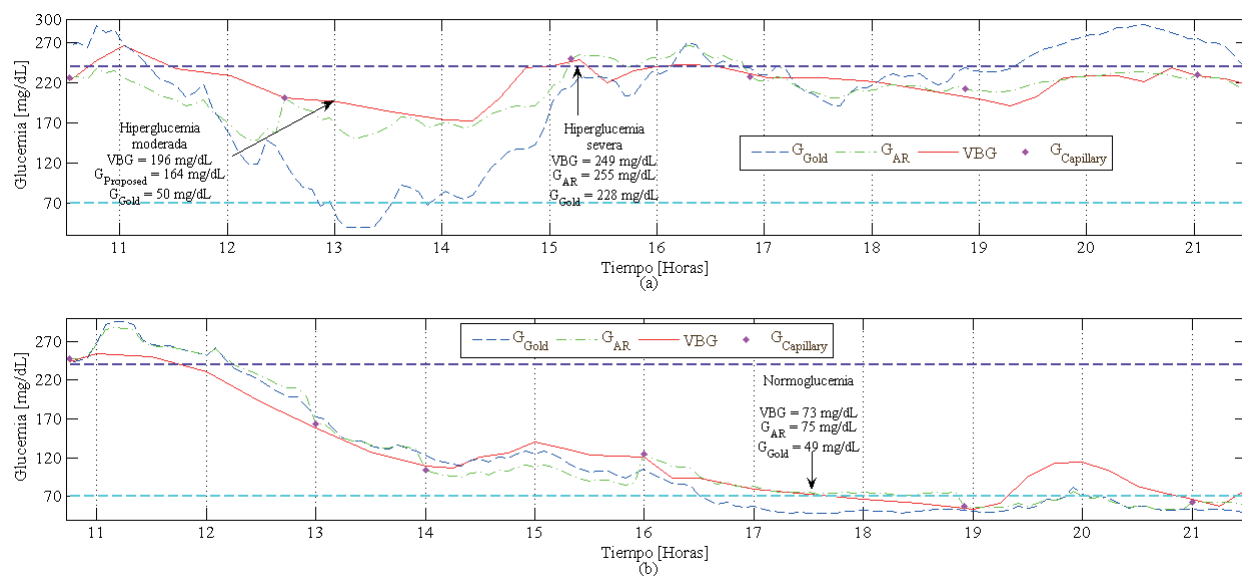


Figura 3.11: Ejemplos de perfiles glucémicos hospitalarios estimados por el A-GoldTM y el ABMAR-SMCG: Estimación de episodios hiperglucémicos y normoglucémicos.

De acuerdo a los resultados obtenidos del análisis que se realiza a los eventos hipoglucémicos que se presentaron en los 18 pacientes durante el primer día del estudio clínico (83 muestras de $VBG \leq 70$ mg/dL), es importante mencionar que el ABMAR-SMCG es más sensible y mejora las estimaciones de glucosa en este rango en comparación con el A-GoldTM. El porcentaje de episodios hipoglucémicos estimados correctamente por el ABMAR-SMCG es 78,3%, mientras que el del A-GoldTM es 45,8%. Por lo tanto, el porcentaje de episodios hipoglucémicos omi-

tidos por el ABMAR-SMCG es menor (21,7% versus 54,2%, respectivamente). A manera de referencia, en la literatura se ha reportado 67,0% y 88,0% de episodios hipoglucémicos ($VBG \leq 70$ mg/dL) estimados correctamente, para los algoritmos del Guardian[®] REAL-Time y del DexCom[™] STS[™], respectivamente (Bode et al., 2004; Garg and Jovanovic, 2006).

3.3.8. Análisis de los resultados obtenidos

Para reducir los errores que presenta el A-Gold[™] en la determinación de BG, particularmente en la estimación de hipoglucemias, se propone el ABMAR-SMCG. Este algoritmo también utiliza las determinaciones de glucosa capilar y las lecturas de corriente suministradas por el sensor para calcular las concentraciones de glucosa. Sin embargo, no se trata de un algoritmo retrospectivo basado en técnicas de regresión lineal. El algoritmo que se propone considera la relación entre IG y BG en condiciones dinámicas. La metodología descrita aquí muestra que es factible usar estructuras ARs para entrenar un modelo poblacional, que desde la información proporcionada por el sensor SC insertado en el ISF, modele la relación dinámica que existe entre IG y BG. Este modelo constituye el núcleo del algoritmo de estimación de BG que se propone. El modelo poblacional AR se obtiene con una estructura de tercer orden con parámetros fijos, la cuál muestra buena capacidad de generalización. El algoritmo para estimación de glucosa basado en modelos ARs tiene la ventaja de poder implementarse en RT, debido a que para calcular la concentración de BG en un instante determinado ($G_{AR}(t)$), el modelo AR sólo utiliza la información de $ISIG$ y de G_{AR} de dos y tres muestras previas al instante de tiempo actual, respectivamente.

La precisión de las estimaciones de glucosa del modelo poblacional AR no es completamente satisfactoria. Es necesario incorporar puntos de calibración para corregir las estimaciones de BG del modelo poblacional, tal como se hace en los algoritmos de calibración que se usan actualmente en los SMCGs. En este algoritmo, cuando se introduce un nuevo punto de calibración, no sólo se ajusta G_{AR} al mismo valor de la calibración, además, las determinaciones de G_{AR} en los dos instantes anteriores se hacen iguales a la glucemia utilizada como calibración, a fin de corregir las estimaciones del modelo. De hecho, si no se considera esta corrección en la estimación de G_{AR} , las estimaciones de glucosa después de introducir la calibración pueden no ser precisas y presentar drásticas diferencias con respecto al *gold standard*.

Un par de ejemplos de esta situación se ilustran en las Figuras 3.12(a) y 3.12(b). En estos casos, sólo se considera la corrección en el valor de G_{AR} cuando se introduce la calibración, en los dos instantes anteriores se mantuvo la medida estimada por el modelo AR. En este caso, la línea verde punteada representa el perfil obtenido de implementar el ABMAR-SMCG tal como se ha propuesto en la sección 3.3.4 y la línea negra punteada representa el perfil del mismo algoritmo sin considerar la corrección en el valor de G_{AR} . En el paciente que se representa en la Figura 3.12(a), el valor de *Median RAD (IQR)* se incrementa desde 11,9 (21,2)% cuando se considera la corrección en los instantes previos hasta 14,2 (23,4)% cuando no se tiene en cuenta. Para el caso del paciente que se muestra en la Figura 3.12(b), el valor de *Median RAD (IQR)* se incrementa desde 5,0 (5,0)% cuando se considera la corrección hasta 8,0 (6,5)% cuando no se tiene en cuenta. Como resultado general, considerando la población global de 18 pacientes con DM-1, el valor de *Median RAD (IQR)* se incrementa (*Median RAD (IQR)* = 6,7 (10,9)% cuando se considera la corrección en los instantes previos versus 15,3 (11,6)% cuando no se tiene en cuenta, p -valor < 0,05).

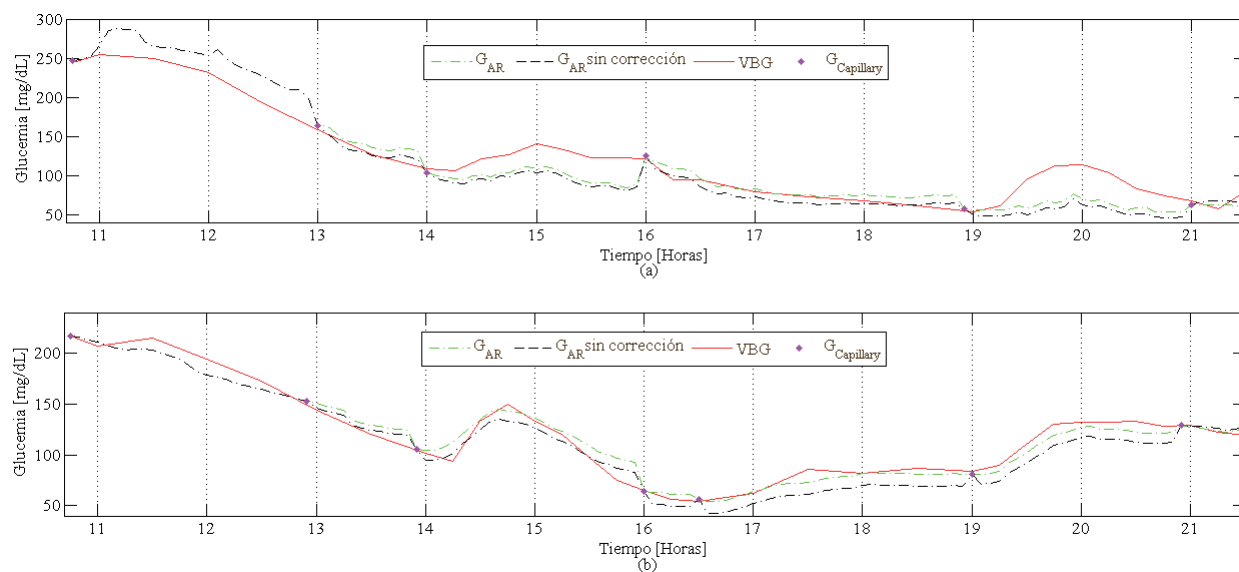


Figura 3.12: Ejemplos de perfiles glucémicos hospitalarios estimados por el ABMAR-SMCG con y sin corregir las dos estimaciones de G_{AR} anteriores a la introducción de la calibración.

Los perfiles que se muestran en las Figuras 3.10(a), 3.10(b), 3.11(a) y 3.11(b) parecen sugerir la necesidad de introducir calibraciones de manera frecuente para obtener un buen rendimiento del ABMAR-SMCG. En las Figuras 3.13(a), 3.13(b), 3.14(a) y 3.14(b) se representan los mismos ejemplos mostrados en las Figuras 3.10(a), 3.10(b), 3.11(a) y 3.11(b), pero introduciendo solamente puntos de calibración antes de cada ingesta, para evaluar si el buen rendimiento del ABMAR-SMCG está condicionado a introducir calibraciones frecuentemente. En este caso, la línea roja continua representa las determinaciones de VBG , la línea verde punteada representa las muestras de G_{AR} obtenidas cuando sólo se consideran las calibraciones antes de cada ingesta, la línea azul punteada corresponde a las muestras de G_{Gold} , los rombos magenta representan las determinaciones de $G_{Capillary}$ que usa el ABMAR-SMCG, y los cuadrados negros las determinaciones de $G_{Capillary}$ que usa el A-GoldTM. En todos estos casos, el ABMAR-SMCG continua mostrando un buen rendimiento y reporta glucemias más cercanas a las determinaciones de VBG , a pesar de utilizar un menor número de calibraciones que el A-GoldTM. Los resultados muestran que el valor de *Median RAD (IQR)* considerando la población global de 18 pacientes con DM-1 se incrementa cuando se reduce el número de calibraciones (*Median RAD (IQR)* = 6,7 (10,9) % versus 8,1 (11,1) %), sin que este incremento indique cambios relevantes en el rendimiento del ABMAR-SMCG. Sin embargo, en investigaciones futuras es necesario determinar la frecuencia con que se debe realizar el proceso de calibración para esta nueva técnica.

Por otra parte, desde el punto de vista de los criterios de precisión analíticos, es importante destacar que el ABMAR-SMCG logra reducir los valores de *Median RAD (IQR)* con respecto al A-GoldTM (*Median RAD (IQR)* = 6,7 (2,9 – 13,8) % versus 11,3 (5,6 – 23,3) %, p -valor < 0,05, respectivamente). Sin embargo, es más relevante que se presente un valor de *Median RAD (IQR)* más bajo en el rango hipoglucémico (*Median RAD (IQR)* = 6,0 (3,0 – 12,0) % versus 15,0 (10,0 – 26,0) %, p -valor < 0,05, respectivamente). A manera de referencia, estos resultados se pueden contrastar con los obtenidos en otros estudios donde se evalúa la precisión del A-GoldTM (Tabla 3.5). Sin embargo, recordemos que en estos estudios las medidas de referencia

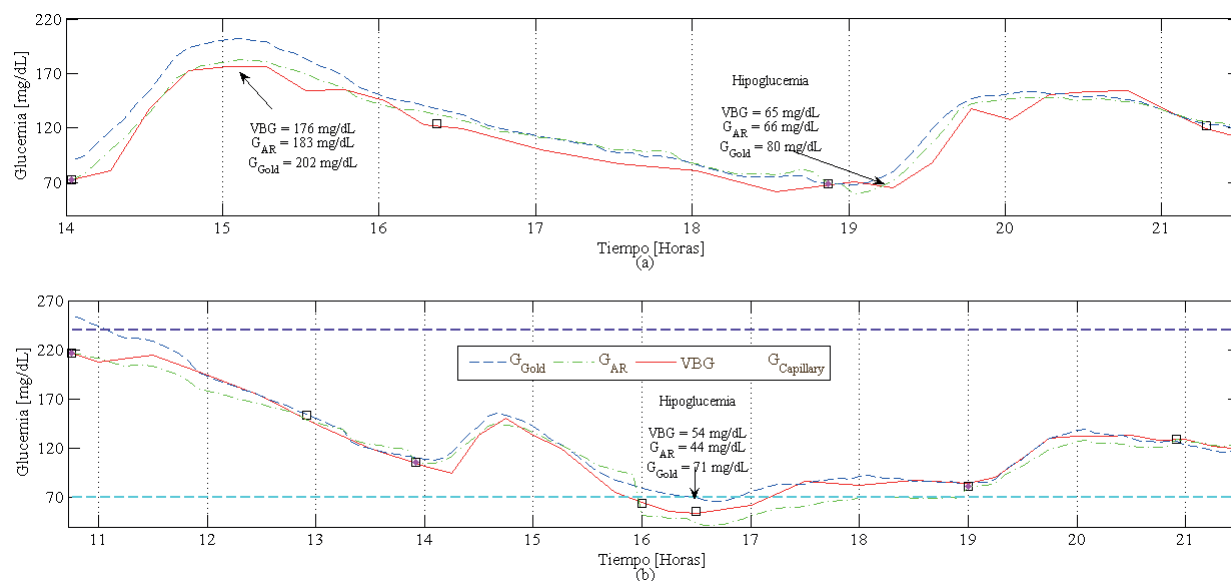


Figura 3.13: Ejemplos de perfiles glucémicos hospitalarios estimados por el A-Gold™ y el ABMAR-SMCG introduciendo solamente puntos de calibración antes de cada ingesta: Estimación de episodios hipoglucémicos.

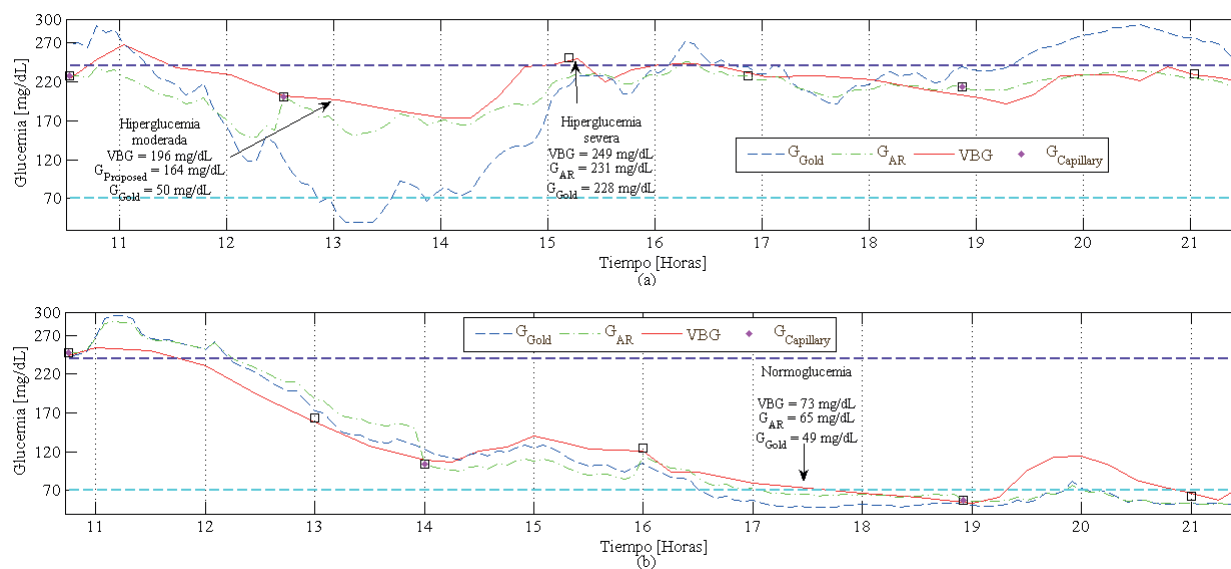


Figura 3.14: Ejemplos de perfiles glucémicos hospitalarios estimados por el A-Gold™ y el ABMAR-SMCG introduciendo solamente puntos de calibración antes de cada ingesta: Estimación de episodios hiperglucémicos y normoglucémicos.

son las determinaciones de glucosa capilar. En la Tabla 3.12 se presentan datos tomados de la literatura reportados por otros estudios de precisión de SMCGs empleando como referencia las determinaciones de VBG. Estos resultados corresponden al A-Gold™, así como a los algoritmos de estimación de glucosa en RT de otros dispositivos disponibles actualmente para uso comercial.

SMCG	Fuente	Median RAD (%)	Criterios ISO (%)	Glucosa de Referencia
Medtronic CGMS [®] Gold TM	(DirecNet, 2003)	18,0	56,0	Venosa (YSI ^a)
Abbott FreeStyle Navigator TM	(Weinstein et al., 2007)	9,3	–	Venosa (YSI ^a)
DexCom TM SEVEN [®]	(Zisser et al., 2009)	13,2	–	Venosa (YSI ^a)
DexCom TM SEVEN [®] Plus	(Bailey et al., 2009)	13,0	–	Venosa (YSI ^a)
Medtronic Guardian [®] REAL-Time	(Mazze et al., 2009)	16,9	–	Venosa (YSI ^a)
DexCom TM STS TM		14,2	–	
Rango hipoglucémico.				
Abbott FreeStyle Navigator TM	(Weinstein et al., 2007) ^b	15,4 (16,1 – 15,1)	–	Venosa (YSI ^a)
Medtronic Guardian [®] REAL-Time	(Kovatchev et al., 2008) ^c	13,8	76,5	Venosa (YSI ^a)
Abbott FreeStyle Navigator TM		7,4	79,4	
DexCom TM STS TM		22,5	52,9	

^aYellow Springs Instrument, Yellow Springs, OH.

^bPara $VBG \leq 70$ mg/dL, se calcula *Median AD* (mg/dL).

^cLa hipoglucemia fue inducida considerando como nivel objetivo $VBG = 45$ mg/dL.

Tabla 3.12: Resultados de algunos criterios de precisión analítica para algoritmos de calibración en RT reportados en la literatura.

Finalmente, desde el punto de vista de los criterios de precisión clínicos, los resultados obtenidos confirman que el ABMAR-SMCG mejora la precisión en la estimación de la glucosa.

3.4. Resumen

En este capítulo, se han estudiado dos nuevos algoritmos para estimación de BG: 1) el ARP-SMCG, un algoritmo retrospectivo basado en técnicas de regresión lineal similares a las que utilizan los algoritmos de estimación de BG de los SMCGs disponibles para uso comercial; 2) el ABMAR-SMCG, un algoritmo que considera la relación entre IG y BG en condiciones dinámicas basándose en un modelo poblacional AR (ABMAR-SMCG). En los dos algoritmos, la estimación de las concentraciones de BG se realizó a partir de las muestras de corriente que suministra el sensor SC y de las determinaciones de glucosa capilar empleadas como puntos de calibración. Sin embargo, el ABMAR-SMCG tiene la ventaja de poder implementarse en RT, debido a que para calcular la concentración de BG en un instante determinado, el modelo AR sólo utiliza la información de la corriente y la glucosa de dos y tres muestras previas al instante de tiempo actual, respectivamente.

El análisis de las estimaciones de BG indicó que ambos algoritmos presentaron mejoras significativas en la precisión, especialmente en el rango hipoglucémico. Los resultados de las evaluaciones para determinar la bondad de su rendimiento desde el punto de vista analítico y clínico, reflejan una buena concordancia entre la respuesta de los algoritmos propuestos y las determinaciones de VBG consideradas como referencia. Particularmente, el ABMAR-SMCG reportó concentraciones de BG con un alto grado de precisión. De igual forma, fue evidente su capacidad para estimar correctamente episodios hipoglucémicos e hiperglucémicos omitidos por el algoritmo de calibración del dispositivo de uso comercial.

Capítulo 4

Evaluación de la precisión de los SMCGs en una Unidad de Cuidados Intensivos (UCI) según el estado séptico de los pacientes

En el capítulo anterior se abordó el problema de precisión en los SMCGs, presentando dos nuevos algoritmos de estimación de BG. Sin embargo, en estos momentos los SMCGs no sólo son imprecisos, también son poco fiables para ser utilizados en la toma de decisiones terapéuticas en los pacientes diabéticos y en los pacientes críticos. Como paso previo para abordar el problema de fiabilidad en los SMCGs, en esta investigación se evalúa el rendimiento de un algoritmo de calibración en RT de un SMCG disponible comercialmente, utilizado en una UCI polivalente.

La hiperglucemia es un fenómeno muy frecuente no sólo en los pacientes diabéticos, sino también en todo tipo de pacientes críticos. De manera clásica se ha interpretado la hiperglucemia como una manifestación más de gravedad, provocada por el estímulo de hormonas contrarreguladoras y de estrés, así como por la resistencia a la insulina que comúnmente presentan estos enfermos críticos, aún sin historia previa de diabetes. Actualmente, se sabe que la hiperglucemia además de representar un marcador de gravedad, está fuertemente relacionada con un peor pronóstico en estos pacientes (McAlister et al., 2005; Capes et al., 2000, 2001; Yendamuri et al., 2003). Los efectos deletéreos de la hiperglucemia en los pacientes críticos pueden ser explicados por el importante efecto pro-inflamatorio que posee la glucosa a nivel molecular, siendo capaz de estimular diferentes citoquinas¹ e interleucinas² pro-inflamatorias, así como estimular la liberación de ácidos grasos libres, entre otros (McCowen et al., 2001; Collier et al., 2008; Montori et al., 2002; Vanhorebeek and Langouche, 2009; Esposito et al., 2002). Por todas las razones anteriormente expuestas, el manejo de la glucemia en los pacientes ingresados en las unidades de cuidados intensivos (UCIs) ha sido objeto de múltiples estudios.

¹Conjunto de proteínas que regulan interacciones de las células del sistema inmune. Su acción fundamental es en la regulación del mecanismo de la inflamación. Hay citocinas pro-inflamatorias y otras anti-inflamatorias. Su función inmunorreguladora es clave en la respuesta inmune, en la inflamación y en la hematopoyesis de distintos tipos celulares. Según la célula que las produzca se denominan linfocinas (linfocito), monocinas (monocitos, precursores de los macrófagos), adipocinas (células adiposas o adipocitos) o interleucinas (células hematopoyéticas).

²Conjunto de citoquinas que son sintetizadas principalmente por los leucocitos, aunque en algún caso también pueden intervenir células endoteliales o de la médula ósea. Su principal función es regular los eventos que atañen a las funciones de estas poblaciones de células del sistema inmunitario, como la activación, la secreción de anticuerpos, entre otras.

Diferentes estudios se han enfocado en definir el rango objetivo dentro del cual se debe mantener la glucosa sanguínea en los pacientes de UCI, con el fin de mejorar su pronóstico. Algunos estudios han afirmado que el control estricto de la glucemia, es decir, mantener niveles de la misma entre 80 mg/dL y 110 mg/dL el máximo de horas posibles al día, disminuye significativamente la morbilidad y mortalidad de los pacientes en estado crítico (Van den Berghe et al., 2006, 2001; Krinsley, 2004; Chase et al., 2010). Sin embargo, han sido muchas las objeciones y limitaciones de dichos estudios. Una de las más relevantes hace referencia a que un rango objetivo muy estricto también se asocia con un aumento en la producción de eventos hipoglucémicos. En consecuencia, otros estudios (Preiser et al., 2009; NICE-SUGAR et al., 2009; Brunkhorst et al., 2008) han indicado que un rango objetivo menos estricto reduce el riesgo de que estos pacientes presenten eventos hipoglucémicos sin aumento de la mortalidad.

Aunque todavía es objeto de debate entre los investigadores cuál debe ser el rango objetivo de control glucémico, lo que no es una cuestión de debate es la necesidad de tratar la glucemia en los pacientes críticos. Actualmente en las UCIs, la determinación de la glucosa sanguínea se realiza de manera intermitente utilizando diferentes glucómetros convencionales, lo que impone una carga adicional para el personal médico y de enfermería. Por esta razón, utilizar SMCGs en estas unidades, además de proporcionar un mejor control de la glucosa, porque permitiría al personal de la UCI identificar en tiempo real (-del inglés *Real-Time (RT)*) las fluctuaciones de la glucosa sanguínea y anticiparse a episodios de hiperglucemia e hipoglucemia, reduce la carga de trabajo del personal sanitario al utilizar un menor número de mediciones manuales.

Diferentes estudios previos han evaluado la precisión de los SMCGs en pacientes críticos (Goldberg et al., 2004; Vriesendorp et al., 2005; Corstjens et al., 2006; Jacobs et al., 2010; De Block et al., 2006; Piper et al., 2006; Logtenberg et al., 2009; Holzinger et al., 2009; Rabiee et al., 2009; Bridges et al., 2010; Brunner et al., 2011). Algunos de estos estudios han concluido que la glucosa estimada por el SMCG tiene una correcta correlación con la glucemia plasmática (Goldberg et al., 2004; Piper et al., 2006; Holzinger et al., 2009; Bridges et al., 2010; Brunner et al., 2011). Otros estudios han concluido que si bien los SMCGs puede ser clínicamente útiles en la UCI, en la actualidad estos no son suficientemente exactos y fiables para ser utilizados en la toma de decisiones terapéuticas (Vriesendorp et al., 2005; Corstjens et al., 2006; Jacobs et al., 2010; De Block et al., 2006; Logtenberg et al., 2009). En contraste, un estudio indicó que aunque los SMCGs proporcionan resultados precisos, la alta tasa de subestimación de hipoglucemias que muestran los hacen poco fiables para ser utilizados en la UCI (Rabiee et al., 2009). Estos estudios utilizaron diferentes tipos de SMCGs y diversos tipos de pacientes.

El presente capítulo está dirigido a describir el protocolo clínico y los resultados obtenidos de un estudio observacional prospectivo realizado en la UCI del Hospital Josep Trueta. El estudio clínico abordado en este capítulo tiene dos objetivos fundamentales. El primer objetivo es evaluar la precisión de las lecturas de glucosa estimadas por el algoritmo en RT del Guardian[®] REAL-Time CGM System (Guardian[®] RTCGMS) (Medtronic, Northridge, CA), cuando se utiliza en una UCI polivalente ³. El segundo objetivo es determinar si el estado séptico del paciente crítico, teniendo en cuenta los diferentes grados de sepsis definidos en (Bone et al., 1992), tiene alguna influencia en la precisión de este SMCG. En particular, la alteración de la permeabilidad capilar, una característica de los pacientes en shock séptico es una condición que puede influir en la relación entre la glucosa plasmática y la glucosa intersticial.

³UCI que admite pacientes con patología de diversa etiología (médica, quirúrgica, traumática)

Este capítulo se organiza de la siguiente forma: la sección 4.1 introduce algunos conceptos generales de uso común en las UCIs. A continuación, la sección 4.2 realiza una descripción de los pacientes y del entorno clínico del estudio. El protocolo del estudio clínico para monitorizar la glucemia en los pacientes críticos se expone en la sección 4.3. Los resultados se reportan en la sección 4.4. La precisión del SMCG se determina utilizando algunos criterios de precisión analíticos y clínicos aceptados actualmente por las agencias reguladoras para evaluar el rendimiento de estos dispositivos. Los criterios de evaluación se describieron previamente en la sección 2.4.1. La sección 4.5 presenta el análisis de los resultados obtenidos. Seguidamente, en la misma sección se indican algunas limitaciones del estudio, que deberían ser consideradas en investigaciones futuras. Finalmente, se presenta un resumen del capítulo en la sección 4.6.

El protocolo de actuación, los resultados y las conclusiones de este estudio se presentaron en forma parcial en las ediciones número 23 y 24 del Congreso Anual de la Sociedad Europea de Medicina Intensiva (Lorencio et al., 2010, 2011), en el Décimo Congreso Internacional sobre Tecnologías para la Diabetes (Leal et al., 2010c) y en la Cuarta Conferencia Internacional sobre Tecnologías Avanzadas y Tratamientos para la Diabetes (Leal et al., 2011). Este estudio en su totalidad se ha publicado en (Lorencio et al., 2012).

4.1. Generalidades

Es importante definir previamente algunos términos de uso común en las UCIs, con el objetivo de comprender mejor el contenido de este capítulo, los resultados obtenidos y las conclusiones extraídas.

- APACHE II (-del inglés *Acute Physiology and Chronic Health Evaluation*-). El APACHE II (Knaus et al., 1985) es una escala de gravedad que se utiliza de forma rutinaria en la UCI. Indica la probabilidad de mortalidad que presenta un determinado paciente en función de la edad, de sus patologías crónicas, y en función de la gravedad de la enfermedad actual valorando las diferentes disfunciones orgánicas que causa. El rango de la escala APACHE es desde 0 hasta 71. Mediante el valor obtenido se puede establecer el riesgo de mortalidad del paciente. A mayor puntuación, mayor gravedad y, por tanto, mayor probabilidad de mortalidad del paciente. Un valor de APACHE superior a 12 se considera como una gravedad importante en el paciente.
- Evaluación de Fallo Orgánico Secuencial (SOFA) (-del inglés *Sepsis-related Organ Failure Assessment*-). El SOFA (Vincent et al., 1996) es una escala de gravedad que evalúa la función de seis órganos diferentes en el cuerpo humano, con puntuaciones para cada uno de ellos de cero a cuatro. Se define como disfunción de órgano si se puntúa de uno a dos, y como fallo de órgano si se obtiene puntuación de tres ó cuatro. Se trata de un sistema sencillo para valorar la gravedad y el riesgo de mortalidad de un paciente crítico, especialmente en situación de sepsis. El rango de puntuación oscila de 0 a 24.
- Shock hemodinámico. Para que los diferentes órganos de nuestro organismo funcionen correctamente, necesitan recibir sangre con la suficiente presión y de forma continua, de manera que puedan recibir correctamente los nutrientes (oxígeno, glucosa, etc.) que precisan. Por diferentes causas, la presión sanguínea sistémica puede disminuir y por tanto, no llegar a perfundir correctamente todos los órganos del cuerpo. Este fenómeno es conocido como shock hemodinámico. Cuando el shock hemodinámico se mantiene en el tiempo

y produce una disfunción de diferentes órganos por hipoperfusión, se puede producir el denominado síndrome de fracaso multiorgánico (disfunción de dos o más órganos).

- Soporte inotrópico (también conocido como soporte vasoactivo, terapia vasopresora o vasoconstrictora). Los inotropos, también llamados fármacos vasoactivos o vasopresores, hacen referencia a un grupo de fármacos que actúan para intentar aumentar la presión sanguínea en situación de shock hemodinámico e intentar evitar o frenar la aparición del síndrome de fracaso multiorgánico. Son inotropos la dopamina, la dobutamina y la noradrenalina⁴.
- Perfusión tisular. Es la presión con la cuál la sangre llega a los diferentes tejidos y órganos. Diferentes mecanismos pueden producir una mala perfusión tisular, entre ellos el shock hemodinámico. Debido a la mala perfusión tisular, se presenta un desarreglo en el funcionamiento de las células intersticiales.
- Edema. Se define como acumulación de líquido en el espacio intersticial.
- Nutrición artificial. Un paciente crítico frecuentemente no puede ingerir alimentos por vía oral. De manera que, cuando el paciente no se puede nutrir de forma normal, se recurre a la nutrición artificial, nutriendo al paciente por métodos no habituales, tales como:
 - Nutrición artificial parenteral: hace referencia a realizar el aporte nutricional que el paciente necesita por vía endovenosa, mediante un catéter.
 - Nutrición artificial enteral: hace referencia a realizar el aporte nutricional que el paciente requiere utilizando la vía digestiva, habitualmente con una sonda nasogástrica. Es mucho más fisiológica que la nutrición parenteral y siempre es la primera elección en el paciente crítico. Sólo se utiliza la nutrición parenteral cuando la nutrición enteral está totalmente contraindicada (cirugía gástrica, cirugía intestinal, obstrucción intestinal por tumoración, etc.).
 - Nutrición artificial mixta: hace referencia a realizar el aporte nutricional por las dos vías, la enteral y la parenteral, de manera simultánea. Habitualmente se da en aquellos pacientes que reciben nutrición enteral, pero que por cualquier motivo no toleran el aporte adecuado totalmente por esta vía y necesitan recibir el resto por vía parenteral.
- Clasificación de los pacientes según el estado séptico. De acuerdo con las definiciones del Colegio Americano de Neumología y del Comité de Consenso de la Conferencia de la Sociedad Americana de Medicina Intensiva (Bone et al., 1992), los pacientes críticos se pueden clasificar según su estado séptico en tres grupos: pacientes no sépticos, con sepsis y en shock séptico.
 - Pacientes no sépticos. La sepsis es una infección grave generalizada, por lo tanto, se consideran pacientes no sépticos los enfermos críticos que no presentan esta condición.
 - Pacientes con sepsis. Son aquellos enfermos críticos que presentan una infección grave generalizada o sepsis, pero en los cuáles esta situación no provoca una vasodilatación de todas las arterias del organismo ni una disminución de la presión sanguínea.

⁴Aminas simpaticomiméticas (aminas que emulan la acción del sistema nervioso simpático) que actúan en los diferentes receptores adrenérgicos y cuyas propiedades mimetizan las de la hormona adrenalina. Estas drogas provocan aumento de la contractilidad miocárdica, de la frecuencia cardíaca o de la vasoconstricción del sistema arterial (Braunwald, 2002).

- Pacientes en shock séptico. Se definen como aquellos enfermos críticos en los cuáles su condición séptica produce un shock hemodinámico. Generalmente presentan mala perfusión tisular después de una correcta resucitación con terapia vasoconstrictora.
- Clasificación de los pacientes según la etiología de la condición crítica del paciente. Esta clasificación depende de la causa variable que ocasiona la enfermedad del paciente, bien por una complicación de una enfermedad médica, post-operatoria o debido a un gran traumatismo.
 - Pacientes médicos. Son enfermos graves por una complicación de una enfermedad médica, por ejemplo una neumonía, una meningitis, etc. que precisan de cuidados en una UCI.
 - Pacientes post-quirúrgicos. Son enfermos que han sufrido una intervención quirúrgica urgente de extrema gravedad (perforación gástrica por una úlcera, perforación de colon por un cáncer, etc.) y precisan de cuidados en una UCI para tratar el cuadro de gravedad posterior a la cirugía. También se clasifican dentro de este grupo los pacientes que son sometidos a una cirugía electiva (programada), pero cuyo estado clínico posteriormente se complica (por sangrado, infección, etc.) y que precisan posteriormente cuidados en UCI para ser estabilizados.
 - Pacientes traumáticos. Son enfermos que han sufrido un traumatismo y que por este motivo se encuentran en una situación crítica. Habitualmente son pacientes que han sufrido accidentes de tránsito o laborales, agresión física con heridas de arma blanca o de fuego, etc. En este grupo también se encuentran aquellos pacientes con grandes quemaduras o que han sido intoxicados.
- Alteración de la permeabilidad capilar. Los capilares son los vasos sanguíneos más delgados que existen en el organismo del ser humano. Cuando aumenta la actividad orgánica, las células endoteliales que forman la pared de los capilares, dejan unos espacios entre ellas para realizar el intercambio de nutrientes con las células que irrigan. Existe un equilibrio perfecto entre los nutrientes que dejan pasar las células endoteliales⁵ a las células que irrigan para nutrir las y los detritos⁶ que se recogen de ellas. En situación de shock séptico, se produce una vasodilatación extrema, es decir, aumenta el tamaño de los poros entre las células endoteliales, rompiéndose el equilibrio en el intercambio de sustancias. Esta disfunción de las células endoteliales ocasiona que pierdan el control del paso de sustancias hacia el intersticio. Entre las sustancias que pasan de manera descontrolada al intersticio se encuentra la glucosa en plasma. A este fenómeno se le conoce como alteración de la permeabilidad capilar.

4.2. Descripción de los pacientes y del entorno clínico del ensayo

El estudio clínico que se presenta en este capítulo corresponde a un estudio observacional prospectivo cuyo objetivo es evaluar la precisión de los SMCGs en una UCI mixta y determinar si el estado séptico del paciente tiene alguna influencia en la precisión de estos dispositivos. La población a estudio se define como aquellos enfermos ingresados en la UCI del Hospital Josep

⁵Una célula endotelial es un tipo de célula aplanada que recubre el interior de los vasos sanguíneos y sobre todo de los capilares, formando parte de su pared. El endotelio es un tejido que recubre el interior de todos los vasos sanguíneos, incluido el corazón, donde se llama endocardio.

⁶En biología los detritos son residuos, generalmente sólidos permanentes, que provienen de la descomposición de fuentes orgánicas (vegetales y animales).

Trueta, recibiendo nutrición artificial (parenteral, enteral o mixta) de forma continua, a dosis ya protocolizadas y como único aporte nutricional. Para que un enfermo crítico fuera incluido en el estudio, debía alcanzar el criterio preestablecido de gravedad, definido como una puntuación superior a 12 en la escala APACHE II.

El estudio se llevó a cabo en la UCI del Hospital Josep Trueta. Esta UCI es una unidad polivalente que consta de 18 camas e ingresa anualmente una media de 700 pacientes, entre los cuáles se incluyen enfermos médicos, post-quirúrgicos y traumáticos.

Los pacientes seleccionados debían cumplir todos los criterios de inclusión que se definen a continuación:

- Pacientes de ambos sexos, ingresados en forma prospectiva en la UCI.
- Tener una edad superior a 18 años.
- Presentar una valoración mayor a 12 en la escala de gravedad APACHE II.
- Presentar hiperglucemia y precisar de insulinoterapia endovenosa en perfusión continua (-del inglés *Continuous Intravenous Insulin Infusion (CIII)*-).
- Estar recibiendo nutrición artificial (parenteral, enteral o mixta) como única forma de soporte metabólico nutricional.
- Tener colocado un catéter arterial debido a su situación clínico-patológica.
- Tener cumplimentada la hoja de consentimiento informado, bien sea firmada por ellos mismos o por su tutor legal o representante.

Los criterios de exclusión que se consideraron fueron:

- Presentar una valoración inferior a 12 en la escala de gravedad APACHE II.
- No precisar de CIII.
- Haber recibido tratamiento con corticoides⁷ de forma crónica en el mes previo al ingreso en la UCI.
- Estar incluido el paciente en otro estudio clínico.
- Mujeres embarazadas.
- Rechazo por parte del enfermo o familiares a participar en el estudio.

De acuerdo con la definición de la población de estudio anteriormente descrita y los criterios de inclusión y exclusión que se definieron previamente, se incluyen en total 41 casos (valor de APACHE II, 18,0 (14,3 – 21,8); valor de SOFA, 8,5 (6,0 – 10,0)). Estos pacientes presentaron hiperglucemia y recibieron CIII. Además, se les colocó un catéter arterial al ingreso para control hemodinámico o para monitorización gasométrica, según características del paciente. Los

⁷Los corticoides son una variedad de hormonas del grupo de los esteroides (producida por la corteza de las glándulas suprarrenales) y sus derivados. Los corticoides están implicados en una variedad de mecanismos fisiológicos, incluyendo aquellos que regulan la inflamación, el sistema inmunitario, el metabolismo de hidratos de carbono, el catabolismo de proteínas, los niveles electrolíticos en plasma y, por último, los que caracterizan la respuesta frente al estrés.

pacientes admitidos corresponden a 14 pacientes médicos, 21 pacientes post-quirúrgicos y 6 pacientes traumáticos. Las características clínicas de los 41 pacientes incluidos se presentan en la Tabla 4.1. El seguimiento clínico del paciente se realizó hasta que este fue dado de alta de la UCI y posteriormente hasta el alta hospitalaria, o en caso de mortalidad del paciente, hasta su fallecimiento.

Variable	Valor
Número (%)	41 (100)
Tipo de paciente (médico/post-quirúrgico/traumático)	14/21/6 (34,2/51,2/14,6)
Edad (años)	63,0 (57,0 – 71,0)
Sexo (femenino/%)	17/41,5
IMC (kg/m ²)	30,0 (28,0 – 35,0)
Conocimiento previo de diabetes (si/%)	19/46,3
Estancia en UCI (días)	20,5 (9,8 – 32,0)
Estancia en el hospital (días)	34,0 (26,3 – 52,0)
Mortalidad en UCI (si/%)	10/24,4
Mortalidad en el hospital (si/%)	13/31,7
SOFA ^a	8,5 (6,0 – 10,0)
APACHE II ^a	18,0 (14,3 – 21,8)
Sepsis ^a (si/%)	31/75,6
Sedación y ventilación mecánica (si/%)	38/92,7
HDFVVC ^b (si/%)	3/7,3
Nutrición (enteral/parenteral)	25/16

Los datos son expresados utilizando la mediana (IQR).

Los números entre paréntesis indican el rango intercuartílico (IQR) (Q1-Q3, respectivamente).

^aEstos valores corresponden a la condición inicial del paciente.

^bHemodiafiltración Venovenosa Continua, se trata de una técnica de terapia de sustitución renal en casos de insuficiencia renal aguda o de insuficiencia renal crónica reagudizada.

Tabla 4.1: Características basales y condiciones primarias de los pacientes de la UCI del Hospital Josep Trueta incluidos en el estudio clínico.

4.3. MCG en la UCI del Hospital Josep Trueta: protocolo de actuación

La regulación de hiperglucemia en UCIs frecuentemente se hace usando protocolos basados en la experiencia médica, los cuáles utilizan la información de los niveles de glucosa capilar para ajustar la dosis de insulina que reciben los enfermos críticos. El valor medido de glucosa determina la cantidad de insulina que se suministra al paciente. La determinación y control habitual de la glucemia en enfermos críticos en la UCI del Hospital Josep Trueta, así como en la mayoría de las UCIs, se realiza determinando la glucemia capilar mediante punción y extracción de sangre del pulpejo digital y medición de la glucosa por reacciones de oxidación enzimática en un glucómetro convencional (Glucocard Gmeter[®], Menarini). Seguidamente y de acuerdo con el protocolo de control de la glucemia en el enfermo crítico, establecido por el personal sanitario de dicha UCI, se administra la insulina pertinente en función del valor de glucosa capilar obtenido. Dependiendo de los valores obtenidos y en función de si la hiperglucemia es o no moderada, se administra insulina rápida endovenosa en perfusión continua o insulina rápida subcutánea cada seis horas.

Aunque la glucemia capilar se encuentra validada como medición para glucemia en enfermos críticos de UCI, en el ensayo clínico que se describe en este capítulo, el “*gold standard*” para la

determinación de la glucemia es la glucosa en plasma arterial (-del inglés *Arterial Blood Glucose (ABG)*-, medida en mg/dL). En este estudio, se siguió el siguiente protocolo de actuación: en primer lugar, a cada enfermo incluido en el estudio, se procedió a determinársele la glucemia plasmática inicial, recogiendo una muestra sanguínea a través de un catéter arterial. Esta muestra se procesó en un analizador de glucosa plasmática, el sistema HemoCue[®] 201 DM (HemoCue AB, Ängelholm, Sweden), el cuál permitió determinar la concentración de *ABG*. Este dispositivo, utilizando una pequeña muestra de sangre entera, calcula por fotometría el valor de la glucosa plasmática mediante reacciones enzimáticas de oxidación, utilizando la enzima glucosa deshidrogenasa⁸. Para garantizar la calidad de las muestras de *ABG* se realizaron pruebas de control de calidad de las determinaciones realizadas por el HemoCue[®] 201 DM utilizando los líquidos de control recomendados en las instrucciones del fabricante.

Seguidamente, se tomaron otras muestras de *ABG* siguiendo el protocolo de control de la glucemia establecido en la UCI. Estos valores de *ABG* fueron utilizados por las enfermeras para administrar la insulino-terapia, en caso de esta ser necesaria. El protocolo de infusión de insulina de la UCI del Hospital Josep Trueta se basa en el protocolo de infusión de insulina descrito por Goldberg (Goldberg et al., 2004).

Según el protocolo de control glucémico, el objetivo es mantener la glucemia plasmática entre 120 mg/dL y 160 mg/dL. A continuación se describen los aspectos más relevantes del protocolo de control de la glucemia establecido en la UCI del Hospital Josep Trueta:

- Si el valor de *ABG* es >150 mg/dL, se inicia administración de insulina endovenosa en perfusión continua (medida en UI/h). Se utiliza insulina de acción rápida (Actrapid, Novo Nordisk, Bagsvaerd, Denmark) diluida en 50 mL de 0,9 % NaCl.
- Una vez iniciada la CIII, se procede a la determinación de *ABG* a la media hora del inicio de la terapia insulínica.
- Si *ABG* es <150 mg/dL, se considera buena respuesta de la hiperglucemia a la insulino-terapia y se procede a la determinación de *ABG* cada hora.
- Si en dos determinaciones horarias consecutivas de *ABG* esta resulta <140 mg/dL, se procede a la determinación de *ABG* cada dos horas, debido a que se considera que existe un buen control de la hiperglucemia. Durante estas dos horas, la perfusión de insulina se mantiene a la misma dosis.
- Si en dos determinaciones consecutivas de *ABG* a intervalos de dos horas esta resulta <140 mg/dL, se interpreta este resultado como un muy buen control de la hiperglucemia y se procede a la determinación de *ABG* cada tres horas. Durante estas tres horas, la perfusión de insulina se mantiene a la misma dosis.
- Si en una de las determinaciones de *ABG* se obtiene un valor de glucemia >150 mg/dL, se restaura la toma de muestras a intervalos de una hora.
- Si en una de las determinaciones de *ABG* se obtiene un valor de glucemia >200 mg/dL, se procede a la determinación de *ABG* cada media hora hasta que los valores sean <140 mg/dL.

⁸Enzima catalizador que interviene en la ruta de las pentosa fosfato para la obtención de energía y de diversas sustancias precursoras en el organismo (Horton et al., 2008).

- Si ABG es <110 mg/dL se suspende la perfusión de insulina durante una hora y se determina ABG cada media hora durante el paro de la insulino terapia. Se reinicia la insulino terapia endovenosa si ABG vuelve a presentar valores >120 mg/dL.
- Si durante el paro de la insulino terapia se registran valores de $ABG <60$ mg/dL se procede a la administración de suero endovenoso glucosado al 33 %.
- Si la ingesta nutricional del paciente se detiene por alguna razón, el control de la glucemia del paciente se realiza con más frecuencia, incluso durante eventos normoglucémicos.

Las determinaciones de ABG utilizadas por el personal de la UCI para administrar la insulino terapia y las dosis de insulina administradas se registran en la base de datos interna del Hospital Josep Trueta. Las dosis de insulina están sincronizadas con las determinaciones de ABG .

A cada enfermo incluido en el estudio, se le determinó de manera continua (cada cinco minutos) la glucosa mediante el sistema subcutáneo (SC) Guardian[®] RTCGMS (MiniMed referencia CSS72), a la que denominaremos G_{RTCGMS} (medida en mg/dL). Las características de este dispositivo han sido descritas en la sección 2.2.2. La monitorización continua de glucosa (MCG) en los pacientes se realizó durante 72 horas. El sensor de glucosa del Guardian[®] RTCGMS se colocó a nivel del tejido SC del muslo superior del paciente, previa limpieza de la zona con alcohol. Se decidió ubicar el sensor en esta posición debido a la dificultad que generaba su colocación en otros lugares. No fue posible ubicar el sensor en la zona pectoral, en la región abdominal o en la zona glútea, porque los enfermos críticos se encuentran generalmente sujetos a radiografías de tórax, laparatomías⁹ y la mayoría de ellos se encuentran en posición supina¹⁰. El sensor se insertó a través de la piel mediante un dispositivo de inserción denominado Sen-serter[®]. Una vez adherido el sensor, se retiró la aguja introductora, dejando bajo la piel el catéter de lectura. Posteriormente se protegió la zona para evitar la salida accidental del catéter. Después de la inserción del sensor y previa comprobación de que no se ha presentado ninguna hemorragia, se conectó el transmisor MiniLink[™] REAL-Time al sensor. Una vez realizada esta conexión, se utilizó el monitor para iniciar la comunicación con el sensor por radiofrecuencia y se inicializó el sensor. Después de un periodo de inicialización de dos horas, se introdujo el primer punto de calibración. La calibración del Guardian[®] RTCGMS se realizó utilizando determinaciones adicionales de ABG , diferentes a las realizadas para aplicar el protocolo de control glucémico.

Luego de realizarse la calibración del dispositivo, se iniciaron las mediciones continuas de G_{RTCGMS} . Al personal de enfermería se le entregó un cuadernillo facilitado por la casa comercial Medtronic MiniMed[®] (Northridge, California), en el cuál se explicaba como introducir los valores de ABG , con el fin de calibrar las medidas del SMCG de forma correcta. Por tanto, las calibraciones se realizaron de acuerdo a las instrucciones del fabricante del Guardian[®] RTCGMS (de tres a cuatro por día) (MiniMed, 2007). El cuadernillo también explicaba los mecanismos de protección del SMCG y las diferentes alarmas que puede mostrar el dispositivo, así como la forma de actuar ante ellas. Las lecturas de glucosa estimadas por el Guardian[®] RTCGMS no se utilizaron para modificar la insulino terapia endovenosa de los pacientes.

⁹Sección quirúrgica a nivel abdominal con el propósito de abrir, examinar y explorar la cavidad intra-abdominal (Brunicardi, 2010).

¹⁰Posición anatómica que se caracteriza por presentar posición corporal recostada sobre la espalda, cuello en posición neutra, miembros superiores pegados al tronco y palmas de las manos hacia abajo, extremidades inferiores extendidas, pies en flexión neutra y punta de los dedos hacia arriba (Brunicardi, 2010).

Un ejemplo representativo de la monitorización continua de la glucosa en un paciente de la UCI del Hospital Josep Trueta se muestra en la Figura 4.1. Las líneas sólidas azules indican las estimaciones de glucosa realizadas por el Guardian[®] RTCGMS (G_{RTCGMS}), los círculos rojos indican las muestras de *ABG* y los triángulos verdes apuntando hacia abajo representan las muestras de *ABG* empleadas para calibrar el Guardian[®] RTCGMS.

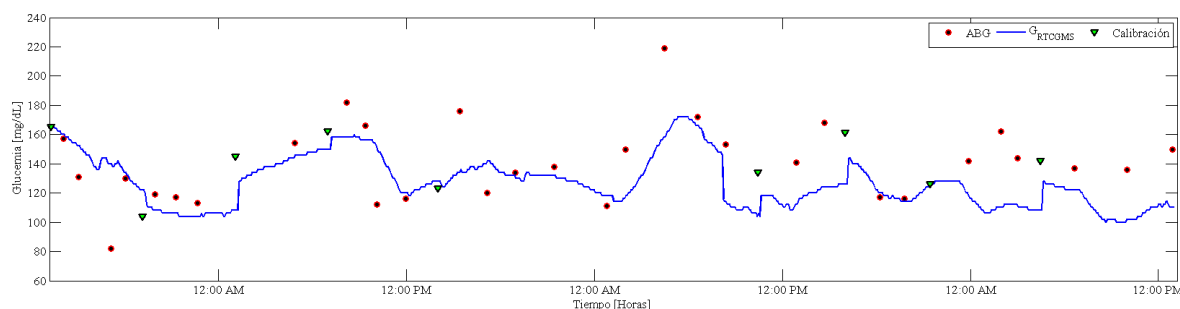


Figura 4.1: Ejemplo del perfil glucémico estimado por el A-Guardian[®] RTCGMS en un paciente crítico.

Después del tercer día, los datos del Guardian[®] RTCGMS y los datos de *ABG* se descargaron a un ordenador utilizando el software respectivo, es decir, el software Medtronic CareLink[®] Pro versión 2.0B, para el Guardian[®] RTCGMS y el software HemoCue[®] 201 DM versión 3.1, para el HemoCue[®] 201 DM.

Así mismo, para cada uno de los pacientes que conforman la población a estudio, se recogieron los criterios diagnósticos y pronósticos de sepsis. El estado séptico de los pacientes se registró cada 24 horas mientras el paciente se encontraba bajo monitorización continua. El grado de estado séptico de los pacientes se clasificó en tres grupos: pacientes en shock séptico, con sepsis y no sépticos. El equipo médico dio el diagnóstico acerca del estado séptico del paciente con base en las definiciones del Colegio Estadounidense de Médicos del Tórax y la Sociedad Norteamericana de Cuidados Críticos (Bone et al., 1992). El equipo médico emitió un diagnóstico de presunción considerando el historial médico del paciente y los resultados de pruebas de laboratorio tales como leucocitosis¹¹, proteína C reactiva (PCR)¹² y procalcitonina (PCT)¹³. La causa del shock hemodinámico fue siempre el shock séptico, ya que se descartaron como causa del shock hemodinámico, el shock cardiogénico¹⁴, el shock hemorrágico¹⁵ y el shock anafiláctico¹⁶. Tampoco se registró el uso de glucocorticoides ni el balance de fluidos.

De igual forma, se recogieron otros aspectos de la situación clínica del paciente, tales como la temperatura corporal axilar y el aporte nutricional. Se registró la temperatura corporal

¹¹Conteo alto de glóbulos blancos

¹²La proteína C reactiva es producida por el hígado y su nivel se eleva cuando hay inflamación en todo el cuerpo. La proteína C-reativa es un examen general para verificar si hay una inflamación en todo el cuerpo. No es un examen específico; es decir, puede revelar que el paciente tiene una inflamación en alguna parte del cuerpo, pero no puede señalar la localización exacta (MED, 2013a).

¹³Es un polipéptido idéntico a la prohormona de la calcitonina que se eleva rápidamente en los pacientes con infección bacteriana grave en ausencia de aumento de la hormona calcitonina. En humanos sanos los valores séricos de PCT son casi indetectables, ascendiendo muy ligeramente en infecciones víricas (Casado and Blanco, 2001).

¹⁴Shock hemodinámico producido por insuficiencia cardíaca, de cualquier etiología.

¹⁵Shock hemodinámico producido por pérdida aguda de sangre, de cualquier etiología.

¹⁶Shock hemodinámico producido por reacción alérgica aguda.

axilar (medida en °C) del paciente utilizando un termómetro electrónico (Thermoval Basic, Hartmann[®], Germany). Al mismo tiempo en que se realizó la primera determinación de *ABG* para administrar la insulino terapia, se registró una muestra de temperatura. Seguidamente, se continuó registrando la temperatura a intervalos de 30 minutos. Además, cada día se valoró el aporte nutricional del paciente, así como la vía de administración de la nutrición, ya sea en forma parenteral, enteral o mixta. También se realizó el cálculo total de kilocalorías y gramos de glucosa recibidos en el día por el paciente.

El protocolo clínico del estudio fue aprobado por el Comité de Ética del Hospital Josep Trueta. Todos los pacientes que participaron en el estudio fueron informados minuciosamente del mismo y decidieron participar de forma voluntaria y libre en él, firmando el correspondiente consentimiento informado, en caso de estar conscientes y capacitados. En el caso de estado de inconsciencia por parte del paciente, se consultó con su tutor legal o representante y fue este quien firmó el consentimiento informado.

4.4. Resultados

Para abordar los dos objetivos planteados al inicio del estudio clínico, se evalúa la precisión de las muestras proporcionadas por el algoritmo de estimación de BG del Guardian[®] RTCGMS (al que denominaremos A-Guardian[®] RTCGMS) en la población mixta de pacientes de la UCI polivalente del Hospital Josep Trueta, y adicionalmente, se determina si el estado séptico del paciente (pacientes en shock séptico, con sepsis y no sépticos) tiene alguna influencia en la precisión de este dispositivo. Por esta razón, se aplican algunos de los criterios de precisión analíticos y clínicos (descritos en la sección 2.4.1) tanto a la población mixta como a las tres cohortes de pacientes definidas según el grado de estado séptico.

4.4.1. Evaluación de la precisión del A-Guardian[®] RTCGMS en pacientes críticos utilizando criterios analíticos

De acuerdo con los criterios enunciados en la sección 2.4.1, la precisión analítica se mide por medio de diferentes métricas estadísticas utilizando pares de datos correspondientes a las medidas proporcionadas por un algoritmo de estimación de BG y a las determinaciones de glucosa consideradas como medidas de referencia. En este caso, se evalúa la precisión analítica de las muestras proporcionadas por el A-Guardian[®] RTCGMS considerando como método de referencia las determinaciones de *ABG*, consideradas por el equipo médico de la UCI como el *gold standard* del protocolo de control glucémico que se aplica a los pacientes críticos incluidos en este estudio. La evaluación de la precisión de las lecturas del Guardian[®] RTCGMS utilizando criterios analíticos se realiza considerando tanto la población mixta de pacientes críticos, como la clasificación realizada según el estado séptico del paciente.

En este estudio clínico se hizo un análisis *a priori* de la potencia del estudio para determinar si el número de determinaciones de *ABG* era suficiente para evaluar la precisión del Guardian[®] RTCGMS utilizando criterios analíticos. Para obtener como mínimo una potencia de 80,0 %, se requerían 250 muestras en cada una de las cohortes de pacientes de acuerdo a su condición séptica. Durante las 72 horas de duración del estudio clínico, de los 41 pacientes que conforman la población mixta de pacientes críticos del estudio se obtuvieron 956 determinaciones de *ABG* (se excluyen las muestras de *ABG* usadas para la calibración del Guardian[®] RTCGMS). Según el

estado séptico de los pacientes, estas determinaciones corresponden a 326 muestras de pacientes en shock séptico, 360 muestras de pacientes con sepsis y 270 muestras de pacientes no sépticos. En la Tabla 4.2 se presentan los rangos glucémicos de ABG tanto de la población mixta como de cada una de las cohortes de pacientes, de acuerdo a su estado séptico. La prueba de Kruskal-Wallis¹⁷ se utiliza para determinar si existen diferencias significativas entre los rangos glucémicos de las determinaciones de ABG de los tres grupos (el umbral de significancia estadística que se considera es del 5,0%). De acuerdo con los resultados de la prueba de Kruskal-Wallis, los rangos de las muestras de ABG de los tres grupos son similares ($p\text{-valor} > 0,05$), sin diferencias significativas entre ellos.

En esta evaluación, las muestras de ABG se comparan con las correspondientes lecturas del A-Guardian[®] RTCGMS. En la Tabla 4.2 también se presentan los rangos glucémicos de G_{RTCGMS} tanto de la población mixta como de cada una de las cohortes de pacientes, de acuerdo a su estado séptico. Solamente las muestras de G_{RTCGMS} emparejadas con las muestras de referencia de ABG se emplean en el cálculo de los criterios de precisión analíticos, debido a que las medidas de ABG no estaban disponibles para el resto de las muestras de G_{RTCGMS} . Para sincronizar en el tiempo las determinaciones de G_{RTCGMS} con las determinaciones de ABG , se aplica una estrategia similar a la que se utilizó en (Garg et al., 2009) para emparejar los datos. Dado que las muestras de G_{RTCGMS} se obtuvieron en diferentes instantes de tiempo que las determinaciones de ABG , esta estrategia consiste en seleccionar los valores de G_{RTCGMS} más cercanos en el tiempo a la medida de ABG , dentro de un intervalo de tiempo de $\pm 2,5$ minutos. De los 41 pacientes que se monitorizaron se obtuvieron 956 pares de datos (ABG/G_{RTCGMS}), los cuáles corresponden a 326 pares de datos de pacientes en shock séptico, 360 pares de datos de pacientes con sepsis y 270 pares de datos de pacientes no sépticos.

La calibración del Guardian[®] RTCGMS se hizo de tres a cuatro veces al día, usando muestras de ABG y siguiendo el procedimiento explicado por el fabricante. Si en cada una de las cohortes de pacientes, se utilizan determinaciones de ABG en diferentes rangos glucémicos para calibrar el SMCG, estas muestras podrían potencialmente influir en el proceso de determinación de la glucemia, estimando valores de glucosa muy diferentes en cada grupo de pacientes, lo que implicaría que cada cohorte de pacientes presentaría porcentajes de precisión muy diversos.

Para evaluar la influencia de la calibración en el funcionamiento del Guardian[®] RTCGMS, se calculan los rangos glucémicos de los puntos de calibración para cada una de las cohortes de pacientes. 493 determinaciones de ABG se utilizaron como puntos de calibración para el Guardian[®] RTCGMS, es decir, en promedio 12 muestras por paciente. El rango de las muestras de ABG usadas para calibración en cada una de las cohortes de pacientes (Tabla 4.2) es similar ($p\text{-valor} > 0,05$), sin diferencias significativas entre ellas (de acuerdo con una prueba de Kruskal-Wallis). Las muestras de ABG usadas para la calibración del Guardian[®] RTCGMS se encuentran dentro del rango de control glucémico considerado idóneo, de acuerdo con el protocolo clínico aprobado.

El primer criterio de precisión analítico que se aplica para evaluar la concordancia entre los pares de datos es el método de Bland-Altman. El método de Bland-Altman grafica las diferen-

¹⁷En estadística, la prueba de Kruskal-Wallis (de William Kruskal y W. Allen Wallis) es un método no paramétrico para probar si un grupo de datos proviene de la misma población. Intuitivamente, es idéntico al ANOVA con los datos reemplazados por categorías. Es una extensión de la prueba U de Mann-Whitney para tres o más grupos. Ya que es una prueba no paramétrica, la prueba de Kruskal-Wallis no asume normalidad en los datos, en oposición al tradicional ANOVA. Sí asume, bajo la hipótesis nula, que los datos vienen de la misma distribución (Kruskal and Wallis, 1952).

Tipos de pacientes	G_{RTCGMS} (mg/dL)	ABG (mg/dL)	ABG para calibración (mg/dL)	No. de pares de datos
Población mixta	128 (108 – 154)	136 (120 – 156)	130 (116 – 147)	956
Pacientes en shock séptico	130 (110 – 158)	136 (118 – 154)	132 (117 – 147)	326
Pacientes con sepsis	124 (103 – 152)	135 (120 – 155)	128 (113 – 144)	360
Pacientes no sépticos	130 (108 – 154)	138 (121 – 158)	138 (114 – 156)	270

Los datos son expresados utilizando la mediana (IQR).

Los números entre paréntesis indican el rango intercuartílico (IQR) (Q1-Q3, respectivamente).

Tabla 4.2: Rangos glucémicos de la población mixta y de cada una de las cohortes de pacientes de acuerdo con su estado séptico.

cias entre cada pareja de valores de ABG y G_{RTCGMS} (eje y) contra la media de cada pareja de valores de ambos métodos (eje x). En la Figura 4.2 se puede observar el gráfico de Bland-Altman para todos los pares de datos obtenidos de la población mixta de pacientes del estudio, así como para los pares de datos correspondientes a cada una de las cohortes de pacientes críticos según su estado séptico. En el eje de las abscisas se grafica la media de ABG y G_{RTCGMS} . En el eje de las ordenadas se grafica la diferencia entre ABG y G_{RTCGMS} . La línea roja representa la media de las diferencias (sesgo) entre ABG y G_{RTCGMS} . Las líneas verdes representan los límites de concordancia, los cuáles se especifican como el sesgo $\pm 1,96$ desviaciones estándar de la media de las diferencias entre ambos métodos.

En la Figura 4.2(a), la cuál representa todos los pares de datos obtenidos de los 41 pacientes incluidos en el estudio, la media de las diferencias (sesgo) entre ABG y G_{RTCGMS} es $6,4 \pm 30,3$ mg/dL, con límites de concordancia de $+65,8$ mg/dL y $-53,1$ mg/dL. En el caso de la Figura 4.2(b), se representa los datos correspondientes a los pacientes en shock séptico, el sesgo entre los dos métodos de medición es $1,7 \pm 25,3$ mg/dL, con límites de concordancia de $+51,3$ mg/dL y $-47,8$ mg/dL. Los datos correspondientes a los pacientes con sepsis se ilustran en la Figura 4.2(c). En este caso, el sesgo es $8,8 \pm 34,2$ mg/dL, y los límites de concordancia son $+75,8$ mg/dL y $-58,3$ mg/dL. Finalmente, los datos correspondientes a los pacientes no sépticos se muestran en la Figura 4.2(d). El sesgo entre los dos métodos de medición es $8,8 \pm 29,8$ mg/dL con los límites de concordancia de $+67,2$ mg/dL y $-49,7$ mg/dL. Al analizar el gráfico de Bland-Altman, este muestra una buena precisión del Guardian[®] RTCGMS en relación con los valores de ABG en los pacientes en shock séptico.

Por otra parte, la concordancia entre los pares de datos comparados también se determina evaluando la diferencia entre las lecturas del Guardian[®] RTCGMS y las muestras del HemoCue[®] 201 DM. Previamente a la aplicación de estos criterios de evaluación, se realiza un análisis de normalidad a los pares de datos, para determinar si estos siguen o no una distribución normal, empleando para ello el Test de Shapiro-Wilk. Debido a que los pares de datos ABG/G_{RTCGMS} no siguen una distribución normal, se calcula *Median RAD (IQR)* (Tabla 4.3). El Test de suma de rangos de Wilcoxon se utiliza para comparar el valor de *Median RAD (IQR)* entre los grupos de pacientes críticos y determinar si existen diferencias estadísticamente significativas entre ellos. De acuerdo a los resultados de esta prueba, los datos en los pacientes en shock séptico son significativamente más precisos que en los pacientes con sepsis (p -valor = $0,03$). Al comparar datos en los pacientes en shock séptico con los datos de los pacientes no sépticos, se detectan diferencias significativas en el *Median RAD (IQR)* entre las dos cohortes de pacientes (p -valor = $0,01$). Por otra parte, no se reportan diferencias estadísticamente significativas entre la cohorte de pacientes con sepsis y la cohorte de pacientes no sépticos (p -valor > $0,05$).

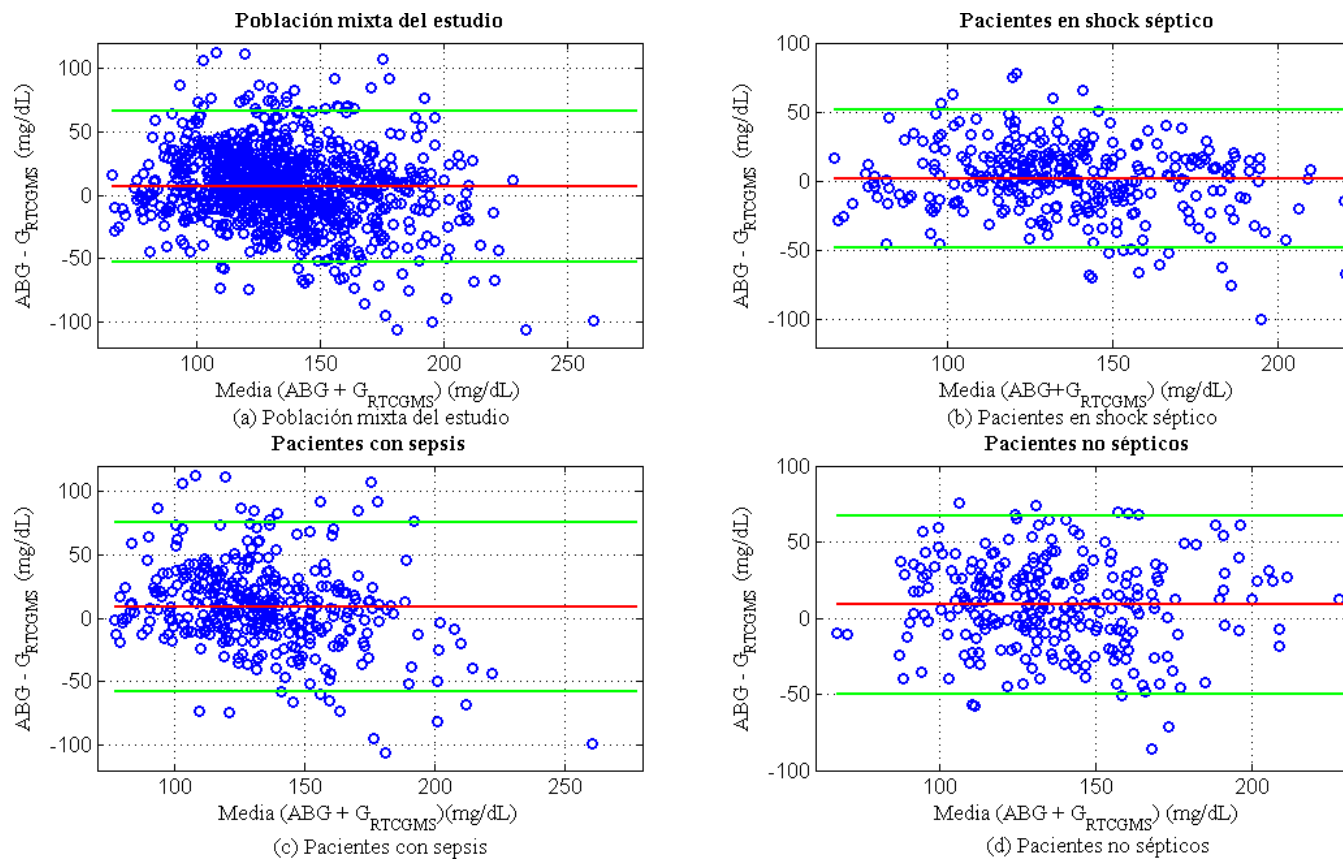


Figura 4.2: Gráficos de Bland-Altman para evaluar el A-Guardian[®] RTCGMS, utilizando los pares de datos $ABG/G_{RTC}GMS$ obtenidos de todos los pacientes de UCI incluidos en el estudio. Los gráficos representan la población mixta y cada una de las cohortes de pacientes de UCI de acuerdo con su estado séptico.

Otro criterio de precisión analítico que se considera son los criterios ISO. La Tabla 4.3 muestra el porcentaje de pares de datos que cumplieron estos criterios, tanto en el conjunto de datos mixto, como en cada uno de los grupos de pacientes según su estado séptico. Los resultados obtenidos de la aplicación de estos criterios permiten corroborar que el Guardian[®] RTCGMS presenta mejor precisión analítica en la cohorte de pacientes en shock séptico en comparación con el rendimiento presentado por el dispositivo en las otras dos cohortes de pacientes. Según los criterios ISO, tanto en la cohorte de pacientes con sepsis como en la cohorte de pacientes no sépticos, la precisión del Guardian[®] RTCGMS es similar.

Finalmente, se realiza un análisis de varianza para determinar si existen diferencias significativas entre el número de determinaciones de ABG realizadas por hora, en cada uno de los tres grupos de pacientes. El número de muestras de referencia tomadas por hora no es estadísticamente diferente entre los tres grupos de pacientes (p -valor $> 0,05$). En la Tabla 4.3 se pueden observar el número de muestras de ABG por hora, tanto para la población mixta de 41 pacientes, como para los grupos de pacientes según su estado séptico.

Tipos de pacientes	Median RAD (IQR) (%)	Criterios ISO (%)	No. Datos	No. de muestras por hora
Población mixta de pacientes	13,5 (6,0 – 24,1)	68,1	956	-
Pacientes en shock séptico	11,2 (5,4 – 20,5)	74,5	326	0,43 (0,36 – 0,60)
Pacientes con sepsis	14,3 (6,7 – 26,9)	65,6	360	0,40 (0,30 – 0,55)
Pacientes no sépticos	16,3 (6,3 – 25,6)	63,7	270	0,50 (0,35 – 0,60)

Los datos son expresados utilizando la mediana (IQR).

Los números entre paréntesis indican el rango intercuartílico (IQR) (Q1-Q3, respectivamente).

Tabla 4.3: Resultados de evaluar la precisión analítica del A-Guardian[®] RTCGMS, usando los pares de datos ABG/G_{RTCGMS} , de acuerdo con su estado séptico.

En adición a la evaluación para determinar si el estado séptico de los pacientes tiene alguna influencia en la precisión del Guardian[®] RTCGMS, también se determina la concordancia entre las muestras de ABG y las estimaciones de G_{RTCGMS} , considerando otros factores, tales como, el estado glucémico del paciente (hipoglucémico, normoglucémico e hiperglucémico), o la causa de su condición crítica (médica, post-quirúrgica o traumática).

Los resultados de la evaluación considerando el estado glucémico del enfermo crítico (27 muestras de $ABG < 80$ mg/dL, 727 muestras de ABG entre 80 mg/dL – 160 mg/dL y 202 muestras de $ABG > 160$ mg/dL) muestran que el Guardian[®] RTCGMS es igualmente preciso cuando los pacientes se encuentran en estado normoglucémico e hiperglucémico (p -valor $> 0,05$). Estos resultados se obtienen aplicando un Test de suma de rangos de Wilcoxon. El Guardian[®] RTCGMS es menos exacto en el intervalo hipoglucémico, pero los datos de esta cohorte de pacientes no se comparan debido a su bajo tamaño muestral (Tabla 4.4).

Finalmente, en la Tabla 4.4 también se presentan los criterios de precisión analíticos de acuerdo con la causa de la condición crítica del paciente (370 pares de datos corresponden a pacientes médicos, 433 corresponden a pacientes post-quirúrgicos y 153 corresponden a pacientes traumáticos). Conforme a los resultados esperados, se reportan resultados similares en los criterios de precisión analíticos en los pacientes médicos y en los pacientes post-quirúrgicos. Al comparar el valor de *Median RAD (IQR)* entre la cohorte médica y la cohorte post-quirúrgica, aplicando un Test de suma de rangos de Wilcoxon, no se encuentran diferencias significativas en la precisión entre estas dos cohortes de pacientes (p -valor $> 0,05$). Los pacientes traumáticos

Tipos de pacientes	Median RAD (IQR) (%)	Criterios ISO (%)	No. Datos
Según el estado glucémico del paciente			
Hipoglucémicos (<80 mg/dL)	15,0 (4,2 – 27,3)	51,9	27
Normoglucémicos (80 mg/dL – 160 mg/dL)	13,9 (6,0 – 23,7)	68,8	727
Hiperoglucémicos (>160 mg/dL)	12,0 (6,0 – 24,1)	67,8	202
Según la causa de la condición crítica del paciente			
Médicos	14,5 (6,6 – 25,0)	67,0	370
Post-quirúrgicos	11,7 (5,4 – 22,4)	71,1	433
Traumáticos	15,5 (7,0 – 28,3)	62,1	153

Los datos son expresados utilizando la mediana (IQR).

Los números entre paréntesis indican el rango intercuartílico (IQR) (Q1-Q3, respectivamente).

Tabla 4.4: Resultados de evaluar la precisión analítica del A-Guardian[®] RTCGMS, usando los pares de datos ABG/G_{RTCGMS} , de acuerdo con otros factores clínicos.

no se comparan por el insuficiente tamaño de esta cohorte.

4.4.2. Evaluación de la precisión del A-Guardian[®] RTCGMS en pacientes críticos utilizando criterios clínicos

En la evaluación de la precisión del A-Guardian[®] RTCGMS utilizando criterios clínicos se utilizan como medidas de referencia las mismas determinaciones de ABG usadas en la evaluación desde el punto de vista analítico (sección 4.4.1). Por lo tanto, también se utilizan pares de datos ABG/G_{RTCGMS} .

Las rejillas de errores de Clarke y Consensus, descritas en la sección 2.4.1, se utilizan para determinar la relevancia clínica de las diferencias existentes entre las estimaciones del Guardian[®] RTCGMS y las determinaciones del HemoCue[®] 201 DM (Tabla 4.5). Es importante mencionar que algunos investigadores consideran que estas rejillas no son relevantes para evaluar la precisión de los SMCGs en la UCI, debido a se diseñaron para aplicarse en pacientes diabéticos y no en pacientes críticos. En los pacientes críticos, el rango de control objetivo, la naturaleza del paciente, y la variabilidad nutricional son completamente diferentes de los que se establecen en los pacientes diabéticos. Sin embargo, dado que estas rejillas han sido empleadas en diferentes estudios de MCG en UCI para evaluar la precisión clínica de los SMCGs (Goldberg et al., 2004; Vriesendorp et al., 2005; Corstjens et al., 2006; De Block et al., 2006; Piper et al., 2006; Logtenberg et al., 2009; Rabiee et al., 2009; Bridges et al., 2010; Jacobs et al., 2010), se ha decidido utilizarlas como herramienta de evaluación complementaria. No obstante, en opinión de algunos autores (De Block et al., 2006; Logtenberg et al., 2009), desde el punto de vista de la seguridad de los pacientes de UCI, la lecturas del SMCG distribuidas en la zona B de Clarke EGA no son aceptables como lecturas precisas, puesto que estos puntos indican una desviación >20,0% entre las lecturas del SMCG y los valores de referencia. Por lo tanto, sólo los puntos del SMCG en la zona A se consideran clínicamente aceptables en el caso de pacientes críticos.

Esta evaluación se realiza considerando tanto la población mixta de pacientes críticos que participaron en el estudio, como el estado séptico de estos pacientes. Al realizar la evaluación de la población mixta, los resultados muestran que el 72,0% de los puntos que representan los pares de datos ABG/G_{RTCGMS} se distribuyen en la zona A en Clarke EGA, mientras que el 65,7% de estos puntos se distribuyen en la zona A en Consensus EGA. La Figura 4.3(a) ilustra

la distribución de estos puntos dentro de cada zona clínica de Clarke EGA.

De acuerdo con el estado séptico del paciente y desde el punto de vista clínico, los resultados de aplicar las rejillas de errores de Clarke y Consensus también muestran que la precisión del Guardian[®] RTCGMS es mayor en los pacientes en shock séptico en comparación con la precisión obtenida en los pacientes con sepsis y no sépticos (Tabla 4.5). Por lo tanto, en la cohorte de pacientes en shock séptico, se observa en las dos rejillas de errores una mayor cantidad de puntos distribuidos en la zona A y una menor cantidad de ellos en la zona B. Es importante destacar que en Clarke EGA se observa que la mayor concentración de puntos en zona D corresponde a pacientes en shock séptico. Esta situación se presenta porque es precisamente esta cohorte de pacientes la que contiene el mayor número de episodios hipoglucémicos (Figura 4.3(b)). Como bien se sabe, la rejilla de Clarke EGA presenta una discontinuidad en el paso de zona A a zona D, cuando se trata de eventos hipoglucémicos. Al comparar las cohortes de pacientes con sepsis y no sépticos, la distribución de puntos en todas las zonas son similares en ambas cohortes, tanto en Clarke EGA como en Consensus EGA (Tabla 4.5 y Figuras 4.3(c) y 4.3(d)).

Tipo	No. Datos	Clarke EGA					Consensus EGA				
		A (%)	B (%)	C (%)	D (%)	E (%)	A (%)	B (%)	C (%)	D (%)	E (%)
Población mixta del estudio	956	72,0	27,3	0,1	0,6	0,0	65,7	33,4	0,9	0,0	0,0
Pacientes en shock séptico	326	79,2	19,6	0,0	1,2	0,0	74,5	25,2	0,3	0,0	0,0
Pacientes con sepsis	360	68,9	30,6	0,3	0,3	0,0	62,2	36,4	1,4	0,0	0,0
Pacientes no sépticos	270	67,4	32,2	0,0	0,4	0,0	59,6	39,3	1,1	0,0	0,0

Tabla 4.5: Resultados de evaluar la precisión clínica del A-Guardian[®] RTCGMS, usando los pares de datos ABG/G_{RTCGMS} obtenidos de la población mixta de pacientes de UCI que participaron en el estudio, y de cada una de las cohortes de pacientes de acuerdo con su estado séptico.

4.5. Discusión

El principal beneficio de utilizar los SMCGs en la UCI es que estos dispositivos pueden proporcionar información detallada y continua acerca de los niveles de glucosa en los pacientes críticos, lo que permitiría realizar un mejor control glucémico. Sin embargo, para que el control glucémico sea bueno, los SMCGs deben ser suficientemente precisos y fiables para ser utilizados en la toma de decisiones terapéuticas. Además, entre sus principales beneficios está evitar las múltiples extracciones sanguíneas en el paciente crítico. En este aspecto, se podrían evitar las frecuentes, incómodas y dolorosas punciones en el pulpejo digital o en el lóbulo de la oreja del paciente, asociadas con la determinación de la glucemia capilar. Otro beneficio es una importante reducción en la carga de trabajo del personal de enfermería debido a que la determinación intermitente de la glucemia, que es la que se realiza actualmente en las UCIs, requiere de la extracción de muestras sanguíneas y estimación de la glucemia mediante glucómetros convencionales varias veces al día, e incluso cada hora, cuando el control glucémico es más difícil. Además, a favor de la seguridad clínica para el personal sanitario, el uso de los SMCGs reduce de forma significativa el número de extracciones sanguíneas que se deben hacer a los pacientes y, por tanto, disminuye radicalmente el riesgo de punción accidental con el riesgo de transmisión de

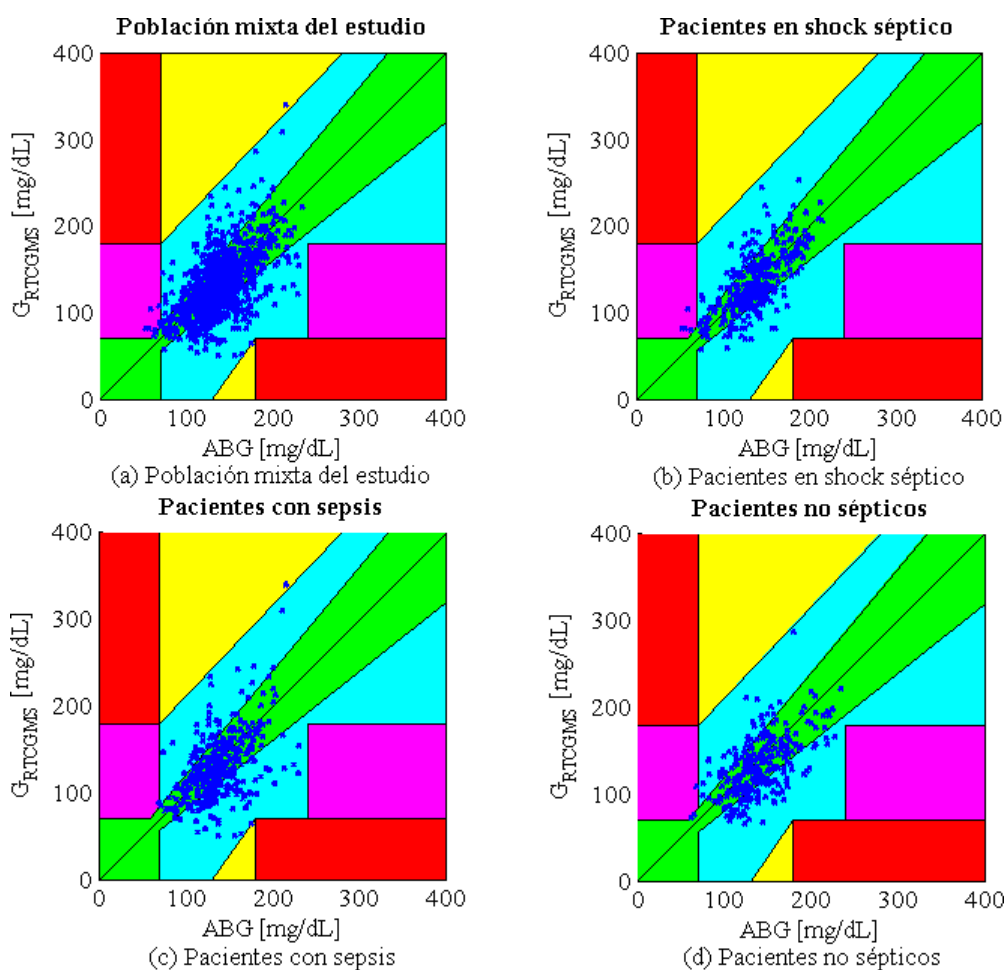


Figura 4.3: Rejillas de Clarke EGA para evaluar el A-Guardian[®] RTCGMS, utilizando los pares de datos ABG/G_{RTCGMS} obtenidos de todos los pacientes de UCI incluidos en el estudio. Los gráficos representan la población mixta y cada una de las cohortes de pacientes de UCI de acuerdo con su estado séptico.

enfermedades que esto conlleva.

Los resultados obtenidos en este estudio observacional aplicando los criterios de precisión analíticos y clínicos son similares a los resultados publicados previamente sobre MCG en pacientes críticos (Vriesendorp et al., 2005; Corstjens et al., 2006; Jacobs et al., 2010; De Block et al., 2006; Piper et al., 2006; Logtenberg et al., 2009). Además, la precisión de los SMCGs en la UCI fue similar a la observada en pacientes con DM-1 (DirecNet, 2008; Mazze et al., 2009; Kovatchev et al., 2008). En concordancia con las conclusiones publicadas en (Vriesendorp et al., 2005; Corstjens et al., 2006; Jacobs et al., 2010; De Block et al., 2006; Logtenberg et al., 2009), los resultados obtenidos permiten concluir que aunque el SMCG podría ser clínicamente útil en la UCI, en la actualidad no es suficientemente preciso y fiable para ser utilizado en la toma de decisiones terapéuticas.

Sin embargo, el principal aporte de esta investigación consiste en determinar si el estado séptico de los pacientes, teniendo en cuenta los diferentes grados de sepsis definidos en (Bone

et al., 1992), influye de alguna manera en la precisión del algoritmo de estimación de BG en RT de un SMCG, en este caso, el A-Guardian[®] RTCGMS. Evaluar la precisión del SMCG desde este enfoque es muy importante porque las directrices internacionales para el tratamiento del shock séptico y la sepsis (dos de los grados de sepsis definidos en (Bone et al., 1992)) recomiendan en este tipo de pacientes la práctica del control glucémico (Dellinger et al., 2008). Los pacientes en shock séptico y los pacientes con sepsis son los más frecuentes en la UCI. En estos pacientes críticos, los valores de la glucemia plasmática se encuentran generalmente en el rango hiperglucémico y el control glucémico en ellos es a menudo más difícil que en los pacientes no sépticos (McCowen et al., 2001; Collier et al., 2008; Montori et al., 2002; Vanhorebeek and Langouche, 2009; Esposito et al., 2002). Por otra parte, la evaluación realizada desde este punto de vista no sólo permite determinar si la condición séptica de los pacientes influye de alguna manera en la precisión de un SMCG, también permite identificar en que tipo de pacientes críticos se puede utilizar un SMCG con mayor fiabilidad. Además, si un mejor rendimiento de un SMCG en una determinada cohorte de pacientes se produce como consecuencia de las características de su condición fisiopatológica, identificar esta condición fisiopatológica podría ser de ayuda en el desarrollo de futuras herramientas para mejorar la precisión y fiabilidad de estos dispositivos.

Los resultados reportados en este estudio indican que la condición séptica de los pacientes si influye en la precisión de un SMCG, y que la precisión del Guardian[®] RTCGMS es significativamente mejor cuando se estima la glucosa en los pacientes en shock séptico. Este resultado puede ser causado por la alteración de la permeabilidad capilar, una característica de la condición fisiopatológica observada en los pacientes en shock séptico. Al alterarse la permeabilidad capilar se presenta un incremento de la difusión de la glucosa intravascular al espacio intersticial. Por consiguiente, el retardo fisiológico para alcanzar el equilibrio entre ABG e IG puede ser más pequeño que en los pacientes que no presentan esta condición fisiopatológica, y, por tanto, el sensor SC del Guardian[®] RTCGMS mide mejor, aumentando la precisión del algoritmo de estimación de BG. La importancia de este resultado está en que a los pacientes en shock séptico es importante realizarles un buen control glucémico, debido a que en ellos es mayor el riesgo de presentar una disfunción de diferentes órganos por hipoperfusión, con la posibilidad de producirse un síndrome de fracaso multiorgánico.

Además, los resultados obtenidos también demuestran que el tipo de paciente crítico, de acuerdo con la causa de la condición de ingreso en UCI del paciente (es decir, médico, postquirúrgico o traumático) no tiene una influencia significativa en la precisión de los SMCGs.

Otro aspecto importante de mencionar es que en la inserción de los sensores SCs no se presentaron complicaciones locales (por ejemplo, infecciones locales, inflamaciones o sangrados).

Este estudio sólo estuvo centrado en un SMCG específico, uno basado en un sensor SC para medir la concentración de glucosa en el líquido intersticial. Por lo tanto, no es posible extrapolar los resultados obtenidos aquí a otras tecnologías de sensores de glucosa, especialmente a los métodos intravenosos.

4.5.1. Limitaciones del estudio

Es importante mencionar que este estudio tiene algunas limitaciones técnicas. En primer lugar, aunque se lograron diferencias estadísticamente significativas en las muestras recolectadas, se requieren estudios futuros con un tamaño muestral más grande. Estos análisis deben enfocarse

en la precisión del SMCG de acuerdo con el estado séptico del paciente.

Por otra parte, los episodios hipoglucémicos fueron poco frecuentes. Por esta razón, no fue posible extraer conclusiones acerca del rendimiento del A-Guardian[®] RTCGMS en el rango hipoglucémico. En la UCI del Hospital Josep Trueta, los eventos hipoglucémicos normalmente son muy poco frecuentes, debido a que se utiliza un protocolo de control glucémico aprobado y validado por el comité científico del hospital, el cuál aunque no está diseñado para realizar un control estricto de la glucemia, si lo está para reducir significativamente el número de eventos hipoglucémicos.

En este estudio se observa una mejor precisión del A-Guardian[®] RTCGMS en pacientes en shock séptico. Recordemos que el shock séptico se presenta cuando la condición séptica de un paciente produce un shock hemodinámico. Para intentar aumentar la presión sanguínea en esta situación e intentar evitar o frenar la aparición del síndrome de fracaso multiorgánico se utiliza la terapia vasopresora, también conocida como soporte inotrópico. En este estudio, el soporte inotrópico se registró de manera cualitativa, indicando si se había usado noradrenalina, dopamina o dobutamina. Sin embargo, una limitación importante es que no se registraron las dosis de los fármacos inotrópicos necesarios en cada caso, ni el tiempo durante el cuál los pacientes requirieron estos fármacos. Previamente, otros autores han reportado una mayor precisión en los SMCGs cuando se utilizan en pacientes pediátricos con soporte inotrópico (Piper et al., 2006) y en pacientes pediátricos con medidas de soporte (tales como ventilación mecánica o terapia vasopresora) (Bridges et al., 2010), aunque sus resultados no fueron estadísticamente significativos. En contraste, otros estudios recientes no reportaron ninguna diferencia en la precisión del SMCG en pacientes con soporte inotrópico (Holzinger et al., 2009; Brunner et al., 2011). Sin embargo, en (Brunner et al., 2011) la dosis promedio de noradrenalina administrada era muy baja ($0,04 \mu\text{g}/\text{Kg}/\text{min}$) y no se explica el motivo por el cuál se administró este fármaco. En estudios futuros, sería interesante evaluar si la precisión del Guardian[®] RTCGMS varía en función de las dosis de los fármacos inotrópicos.

Además, tampoco se consideraron otros factores que podrían influir en la precisión del Guardian[®] RTCGMS en pacientes críticamente enfermos, tales como el grado de edema o el nivel de la perfusión tisular, determinado este último mediante la medición de la concentración de lactato arterial¹⁸. Otros factores que podrían ser de interés, tales como el uso de glucocorticoides o el balance de fluidos, tampoco se registraron.

Aunque los resultados del Guardian[®] RTCGMS no se comparan con los valores de glucosa plasmática obtenidos en pruebas de laboratorio o con un analizador de gases sanguíneos, si se comparan con un dispositivo de uso regular en UCIs. Las muestras de *ABG* se determinaron con el HemoCue[®] 201 DM configurado en sangre entera. La selección del HemoCue[®] 201 DM se hizo con base en que se requerían mediciones más frecuentes de glucosa plasmática, en concordancia con el protocolo aprobado. El HemoCue[®] 201 DM es adecuado para este propósito, debido a sus mínimos requerimientos de volumen de sangre entera, un sencillo procedimiento de calibración, respuesta rápida, bajo costo y durabilidad en el ambiente clínico. Además, el HemoCue[®] 201 DM proporciona resultados fiables sobre un amplio rango de valores de hematocrito (Wiener,

¹⁸El ácido láctico o lactato se produce principalmente en las células musculares y en los glóbulos rojos. Dicho ácido se forma cuando el cuerpo descompone carbohidratos para utilizarlos como energía durante momentos de niveles bajos de oxígeno. El nivel de oxígeno en el cuerpo podría bajar si la persona tiene una infección o una enfermedad. Los resultados anormales indican que los tejidos corporales no están obteniendo suficiente oxígeno (MED, 2013a).

1993). Sin embargo, existen dudas acerca de su precisión (Rice et al., 2010; Watkinson et al., 2012). Mientras que estudios anteriores han encontrado una buena correlación entre el sistema HemoCue[®] 201 DM y las muestras de glucemia plasmática obtenidas mediante pruebas de laboratorio (Stork et al., 2005; Torjman et al., 2001), otro estudio no recomienda su uso debido a que considera que es significativamente afectado por los niveles de hematocrito, cuando se compara con las glucemias obtenidas mediante pruebas de laboratorio (Watkinson et al., 2012). Sin embargo, este dispositivo también ha sido utilizado como referencia en otros estudios importantes para analizar el control glucémico (Van den Berghe et al., 2006; NICE-SUGAR et al., 2009; Brunkhorst et al., 2008). Por lo tanto, para evitar estos problemas en estudios futuros, es vital comparar con mediciones obtenidas mediante pruebas de laboratorio.

Otra limitación de este estudio es que no se evaluó la precisión del Guardian[®] RTCGMS en función de la localización del sensor, debido a la complejidad fisiológica de los pacientes en estado crítico. La implantación del sensor en otros sitios anatómicos como el abdomen es difícil, particularmente en los pacientes post-quirúrgicos, debido a su condición clínica. Su ubicación en la zona pectoral queda descartada debido a que no es recomendable exponer el transmisor del sensor a los rayos X (MiniMed, 2007). Además, los pacientes críticamente enfermos están usualmente en posición supina, y ellos permanecen sedados durante mucho tiempo, por lo que es imposible localizar el Guardian[®] RTCGMS en la zona glútea.

Por último, un problema importante de este estudio ha sido el largo intervalo de aprendizaje requerido por el personal médico y de enfermería de la UCI, para adquirir los conocimientos necesarios en el correcto uso del Guardian[®] RTCGMS. Para usar de manera rutinaria los SMCGs en las UCIs, es importante proporcionar una adecuada formación específica al personal médico y de enfermería en el uso eficaz de estos dispositivos. Se debe formar a todo el personal implicado, tanto en la inserción del sensor SC de glucosa, como en el uso y calibración del monitor, e interpretación de sus resultados.

4.6. Resumen

En este capítulo, se ha descrito el protocolo clínico y los resultados obtenidos de un estudio observacional prospectivo realizado en la UCI del Hospital Josep Trueta. Este estudio evaluó la precisión del algoritmo de estimación de BG del Guardian[®] RTCGMS cuando se utiliza en una UCI polivalente y se enfocó en determinar si el estado séptico del paciente tiene alguna influencia en la precisión de este dispositivo.

El análisis de las estimaciones de glucosa del Guardian[®] RTCGMS indicó que su precisión en este estudio es similar a la reportada en otros estudios previos realizados en pacientes críticos, es decir, que aunque el SMCG podría ser clínicamente útil en la UCI, en la actualidad no es suficientemente preciso y fiable para ser utilizado en la toma de decisiones terapéuticas. Por otra parte, estos resultados también mostraron que el estado séptico del paciente influyó en la precisión del Guardian[®] RTCGMS. La precisión fue significativamente mejor en los pacientes en shock séptico en comparación con las otras cohortes de pacientes. Este resultado puede ser causado por la alteración de la permeabilidad capilar, una característica de la condición fisiopatológica de los pacientes en shock séptico. La importancia de este resultado está en que a los pacientes en shock séptico es vital realizarles un buen control glucémico, debido a que en ellos es mayor el riesgo de presentar una disfunción de diferentes órganos por hipoperfusión, con la

posibilidad de producirse un síndrome de fracaso multiorgánico.

Sin embargo, para confirmar los beneficios que potencialmente tiene la MCG en comparación con la monitorización convencional en los pacientes críticos, se hace necesario realizar investigaciones futuras con un tamaño muestral más grande y considerando otros factores fisiopatológicos, tales como las dosis de los fármacos inotrópicos, el grado de edema o el nivel de la perfusión tisular, los cuáles podrían influir en la precisión de los SMCGs, abordando esta investigación desde la influencia de la condición séptica de los pacientes.

Capítulo 5

Aplicación de Análisis de Componentes Principales (PCA) y Razonamiento Basado en Casos (CBR) para detectar medidas terapéuticamente correctas (TC) e incorrectas (TI) en SMCGs

En el capítulo anterior se mencionaron los beneficios de utilizar sistemas de monitorización continua de glucosa (SMCGs) (-del inglés *Continuous Glucose Monitoring Systems (CGMSs)*-) en la unidad de cuidados intensivos (UCI). Aunque los SMCGs podrían ser clínicamente útiles en la UCI, actualmente no son lo suficientemente precisos y fiables para ser utilizados en la toma de decisiones terapéuticas. Sus carencias de precisión y fiabilidad han sido importantes factores que limitan la administración de la insulina de manera automática, y en su uso clínico. Por esta razón, ninguno de los SMCGs disponibles comercialmente han sido aprobados por las agencias reguladoras como sustitutos del método tradicional de monitorización de glucosa en sangre de manera intermitente. Por consiguiente, mejorar la exactitud y fiabilidad de estos dispositivos es esencial.

En el Capítulo 3 se han propuesto y descrito dos nuevos algoritmos de estimación de glucosa en plasma (-del inglés *Blood Glucose (BG)*-) para abordar los problemas de precisión presentados por los SMCGs. Sin embargo, la fiabilidad también es uno de los principales requerimientos para un SMCG, y con el fin de mejorarla, se deben incluir en estos dispositivos sistemas de detección de fallos y averías. El número de trabajos desarrollados con este propósito presentes actualmente en la literatura es limitado (Ward et al., 2004; Bondia et al., 2008; Shen et al., 2010; Signal et al., 2012). Por lo tanto, es necesario realizar progresos significativos en este sentido (Bequette, 2010).

El objetivo principal de este capítulo es desarrollar una metodología basada en datos para detectar las estimaciones de BG realizadas por el Guardian[®] RTCGMS, determinando cuáles son terapéuticamente correctas (TC) e incorrectas (TI), y de esta manera proporcionar un diagnóstico al personal de la UCI que les sea de utilidad para la toma de decisiones clínicas. La metodología que aquí se propone utiliza Análisis de Componentes Principales (-del inglés *Prin-*

Principal Component Analysis (PCA)-) y el Razonamiento Basado en Casos (-del inglés *Case-Based Reasoning (CBR)*-) para determinar cuando una lectura del Guardian[®] RTCGMS es TC o TI. PCA se utiliza para extraer características: el estadístico Q (conocido en inglés como Q -statistic) y las nuevas observaciones (conocidas en inglés como *scores*). Luego, el Q -statistic y los *scores* obtenidos a partir del modelo PCA, junto con el estado séptico del paciente, se proponen como descriptores para constituir las estructuras de casos del CBR. CBR se aplica para reconocer patrones y clasificar los datos. En el CBR se han considerado dos enfoques diferentes de vecinos más cercanos (conocidos en inglés como *z-Nearest Neighbours (z-NN)*) para recuperar los casos pasados que permiten clasificar los nuevos casos: uno basado en un criterio de distancia en dos pasos (utilizando exclusivamente las salidas del modelo PCA), al cuál llamaremos *recuperación en dos pasos* (-del inglés *Two-steps retrieval (2SR)*-) y el otro basado en un criterio de distancia en tres pasos (empleando las salidas del modelo PCA y la similitud del estado séptico entre los casos), al cuál llamaremos *recuperación en tres pasos* (-del inglés *Three-steps retrieval (3SR)*-).

El capítulo se organiza como sigue: la sección 5.1 describe el proceso de experimentación clínica usado para capturar los datos. Después, la sección 5.2 explica la transformación del problema de detección de fallos en un SMCG en un problema de bi-clasificación. En la sección 5.3, se introducen los conceptos de PCA y a continuación, se describe la construcción del modelo PCA. Por otra parte, en la sección 5.4, se presentan los conceptos básicos de CBR y seguidamente, se explica la implementación de la metodología CBR para detectar medidas de glucosa TC y TI en un SMCG. Seguidamente, en la sección 5.5, se reporta la puesta a punto de la metodología PCA-CBR, empleando las métricas para evaluar la fiabilidad de los SMCGs descritas en 2.4.2. A continuación, en la sección 5.6, se realiza una evaluación comparativa entre la metodología PCA-CBR, Análisis Discriminante por Mínimos Cuadrados Parciales (-del inglés *Partial Least Squares Discriminant Analysis (PLS-DA)*-) y Modelado Blando Independiente de Analogías entre Clases (-del inglés *Soft Independent Modeling of Class Analogy (SIMCA)*-). Luego, en la sección 5.7 se presentan los resultados de la evaluación de la metodología PCA-CBR en dos pacientes críticos. En la sección 5.8 se analizan los resultados obtenidos. Finalmente, se presenta un resumen del capítulo en la sección 5.9.

Una descripción de este trabajo se presentó en forma parcial en la Quinta Conferencia Internacional sobre Tecnologías Avanzadas y Tratamientos para la Diabetes (Leal et al., 2012b). Un artículo con la metodología completa ha sido aceptado para publicación (Leal et al., 2013b).

5.1. Descripción de los datos

En esta investigación se utiliza un conjunto de datos obtenidos de 23 pacientes en estado crítico admitidos en la UCI mixta del Hospital Josep Trueta (valor de APACHE II¹, 20.0 (15.0 – 21.5); valor de SOFA², 8.0 (6.0 – 10.0)). Estos pacientes forman parte del grupo de 41 enfermos críticos incluidos en el estudio clínico descrito en la sección 4.2. Los resultados globales del estudio se presentaron en (Lorenzo et al., 2012). Se utilizan datos de sólo 23 pacientes debido a que no se disponía de la corriente eléctrica generada por el sensor para el grupo completo de 41 enfermos críticos. Las características clínicas de los 23 pacientes utilizados se presentan en la Tabla 5.1. Este conjunto de datos corresponde a pacientes quienes al ingreso en la UCI presentaron hiperglucemia y requerían insulino terapia endovenosa en perfusión continua (-del

¹Definido en 4.1

²Definido en 4.1

inglés *Continuous Intravenous Insulin Infusion (CIII)*-). Este conjunto de datos está constituido por información proporcionada por el Guardian[®] RTCGMS e información de algunas variables acerca de la condición clínica del paciente.

Variable	Valor
Número (%)	23 (100)
Tipo de paciente (médico/post-quirúrgico/traumático)	7/11/5 (30,5/47,8/21,7)
Edad (años)	62,0 (54,5 – 74,0)
Sexo (femenino/%)	11/47,8
IMC (kg/m ²)	29,0 (28,0 – 32,5)
Conocimiento previo de diabetes (si/%)	13/56,5
Estancia en UCI (días)	21,0 (14,5 – 33,0)
Estancia en el hospital (días)	34,0 (28,0 – 52,0)
Mortalidad en UCI (si/%)	7/30,4
Mortalidad en el hospital (si/%)	9/39,1
SOFA ^a	8,0 (6,0 – 10,0)
APACHE II ^a	20,0 (15,0 – 21,5)
Sepsis ^a (si/%)	18/78,3
Ventilación mecánica (si/%)	23/100,0
Nutrición (enteral/parenteral)	15/8

Los datos son expresados utilizando la mediana (IQR).

Los números entre paréntesis indican el rango intercuartílico (IQR) (Q1-Q3, respectivamente).

^aEstos valores corresponden a la condición inicial del paciente.

Tabla 5.1: Características basales y condiciones primarias de los pacientes de UCI utilizados en el desarrollo de la metodología PCA-CBR.

Las determinaciones de *ABG* se consideran como el método de referencia para etiquetar el conjunto de datos en medidas TC y TI. El proceso de etiquetado de los datos se explica en la sección 5.2. Para mayor información, en la sección 4.3 está descrito el proceso de obtención de *ABG*.

El Guardian[®] RTCGMS proporciona las siguientes medidas: la corriente eléctrica generada por el sensor (*ISIG*, medida en nA) y la estimación de la cantidad de glucosa en plasma (*G_{RTCGMS}*, medida en mg/dL). El protocolo clínico que describe el proceso de monitorización de los pacientes utilizando el Guardian[®] RTCGMS está explicado en la sección 4.3.

Además, se ha considerado información de algunas variables acerca de la condición clínica del paciente: la dosis de insulina endovenosa en perfusión continua (a la que denominaremos *Insulin*, medida en UI/h), la temperatura corporal axilar (a la que denominaremos *Temperature*, medida en °C) y el estado séptico de los pacientes (a la que denominaremos *Sepsis*). Para mayor información, en el protocolo clínico descrito en la sección 4.3 también está explicado el proceso para obtener estas variables. La variable *Sepsis* se considera como una variable categórica que podría tomar los siguientes valores: *Sepsis* = 0 en pacientes no sépticos, *Sepsis* = 1 en pacientes con sepsis y *Sepsis* = 2 en pacientes en shock séptico.

De los 23 pacientes se obtienen 537 determinaciones de *ABG* (se excluyen las medidas de *ABG* usadas para la calibración del Guardian[®] RTCGMS). Estas determinaciones de *ABG* se sincronizan con *ISIG*, *G_{RTCGMS}*, *Insulin*, *Temperature* y *Sepsis*. En el caso de los datos del Guardian[®] RTCGMS (*ISIG* y *G_{RTCGMS}*), dado que estos se obtuvieron en diferentes instantes de tiempo que las determinaciones de *ABG*, para emparejar los datos se seleccionan los valores de *ISIG* y *G_{RTCGMS}* más cercanos en el tiempo a la medida de *ABG*, dentro de un intervalo de tiempo de $\pm 2,5$ minutos como en (Garg et al., 2009). Respecto a *Insulin*, esta variable

se registró al mismo tiempo que *ABG*, por tanto, las muestras de *Insulin* están sincronizadas con las determinaciones de *ABG*. En relación con *Temperature*, puesto que al tiempo en que se realizó la primera determinación de *ABG* en un paciente para administrar la insulino-terapia, se registró una muestra de *Temperature*, y esta variable se continuó registrando cada 30 minutos, cada vez que se obtenía una determinación de *ABG* existía un registro de *Temperature*. Por otro lado, en relación con *Sepsis*, este estado sólo se registra cada 24 horas, por lo tanto, cada una de las determinaciones de *ABG* se asocia con el respectivo estado séptico previamente diagnosticado por el equipo médico del paciente.

Para el conjunto de datos de 23 pacientes, la calibración del Guardian[®] RTCGMS se hizo de la misma forma como se explicó en la sección 4.4.1 para el conjunto de 41 pacientes, utilizando en este caso 241 lecturas de *ABG* como puntos de calibración. Las determinaciones de *ABG* usadas para la calibración de este conjunto de datos también se encuentran dentro del rango considerado correcto de acuerdo con el protocolo clínico aprobado. En la Tabla 5.2 se presentan los rangos glucémicos de *G_{RTCGMS}*, *ABG* y de los puntos de calibración.

Tipos de pacientes	<i>G_{RTCGMS}</i> (mg/dL)	<i>ABG</i> (mg/dL)	<i>ABG</i> para calibración (mg/dL)	No. de pares de datos
Población mixta	128 (106 – 156)	138 (121 – 160)	129 (116 – 147)	537
Pacientes en shock séptico	129 (104 – 163)	137 (113 – 157)	130 (117 – 145)	226
Pacientes con sepsis	126 (106 – 152)	138 (125 – 161)	129 (117 – 148)	208
Pacientes no sépticos	128 (106 – 153)	138 (121 – 166)	129 (107 – 147)	103

Los datos son expresados utilizando la mediana (IQR).

Los números entre paréntesis indican el rango intercuartílico (IQR) (Q1-Q3, respectivamente).

Tabla 5.2: Rangos glucémicos de la población mixta y de cada una de las cohortes de pacientes utilizados.

5.2. Etiquetado y pre-procesamiento de los datos

Por lo que se ha explicado anteriormente, el conjunto de datos contiene 537 observaciones recogidas de 23 pacientes. Sin embargo, en el trabajo que se presenta en este capítulo, se utilizan datos de sólo 22 pacientes debido a que no se disponía de la variable *Temperature* para uno de ellos, específicamente un paciente traumático. Por lo tanto, el conjunto de datos con el que se desarrolla este capítulo contiene 501 observaciones y cada una de ellas está compuesta por cinco variables: *ISIG*, *G_{RTCGMS}*, *Insulin*, *Temperature* y *Sepsis*. Adicionalmente, se han obtenido 501 determinaciones de *ABG* que se utilizan como medidas de referencia para etiquetar los datos, basándose en un esquema binario de clasificación definido a partir de un criterio de precisión clínico.

En primer lugar, se etiqueta el conjunto de datos en medidas TC y TI usando la rejilla de errores de Clarke (Clarke EGA se describió en la sección 2.4.1). Para comparar los pares de datos *ABG/G_{RTCGMS}*, estos se trazan por medio de puntos en las diferentes zonas de riesgo de Clarke EGA. Sin embargo, en opinión de algunos autores (De Block et al., 2006; Logtenberg et al., 2009), desde el punto de vista de los estándares de seguridad de la UCI, la estimaciones de glucosa del SMCG distribuidas en la zona B de Clarke EGA no son aceptables como precisas, puesto que estos puntos indican una desviación >20,0% entre las lecturas del SMCG y los valores de referencia. Por lo tanto, sólo los puntos del SMCG en la zona A se consideran clínicamente aceptables en el caso de pacientes críticos. Así, se define el siguiente esquema de clasificación

para etiquetar el conjunto de datos, basado en la distribución de los puntos en las diferentes zonas de riesgo (Figura 5.1):

- Clase terapéuticamente correcta (Clase TC): Los puntos que se trazan en la zona A. Del conjunto entero de datos 72,7% (364 observaciones) son medidas TC.
- Clase terapéuticamente incorrecta (Clase TI): Los puntos que se trazan en las zona B, C, D y E. Del conjunto entero de datos 27,3% (137 observaciones) son medidas TI.

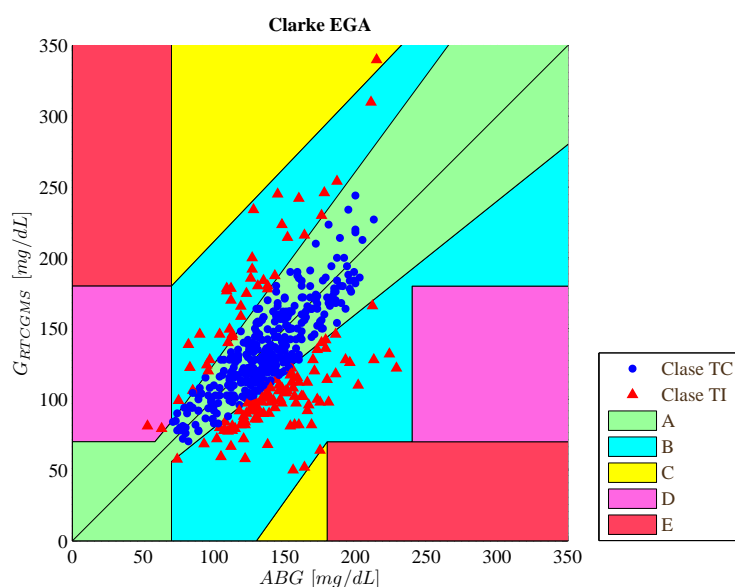


Figura 5.1: Distribución de los datos de acuerdo con Clarke EGA

Por consiguiente, el conjunto de datos contiene 501 observaciones etiquetadas en dos clases (364/137) y cada observación está compuesta por cinco variables: *ISIG*, *GRTCGMS*, *Insulin*, *Temperature* y *Sepsis* (Figura 5.2)

Información sobre la muestra				Referencia	Variables del proceso					Etiquetado de los datos	
Paciente No.	Caso No.	Fecha	Hora	ABG	ISIG	<i>GRTCGMS</i>	Insulin	Temperature	Sepsis	Clase	Zona Clarke EGA
3	65	09/08/2008	21:36:35	189	16,37	214	6	37,7	2	TC (-1)	A
3	66	09/08/2008	22:38:44	159	14,74	190	5	37,8	2	TC (-1)	A
3	67	10/08/2008	0:16:21	138	13,21	170	2	38,3	2	TC (-1)	A
3	68	10/08/2008	1:25:18	215	27,21	340	2	38,3	2	TI (1)	C
3	69	10/08/2008	2:08:06	211	24,73	310	2	38,3	2	TI (1)	B
3	70	10/08/2008	3:27:37	212	15,06	166	2	38,1	2	TI (1)	B

Figura 5.2: Ejemplo de una porción del conjunto de datos utilizado para construir la metodología PCA-CBR

Para construir la metodología PCA combinada con CBR que permita detectar fallos en los SMCGs, se utilizan los datos de 20 pacientes (432 observaciones, de las cuáles 321 son TC y 111 son TI). Los restantes dos pacientes (representativos de las cohortes de pacientes post-quirúrgicos y médicos) no se emplean en el desarrollo de dicha metodología. Estos pacientes se utilizan

para evaluar el rendimiento de uso generalizado de la metodología PCA-CBR. Más detalles sobre estos dos pacientes se darán en la sección 5.7.

5.3. Extracción de características usando PCA

5.3.1. Conceptos básicos de PCA

PCA es una técnica utilizada para encontrar combinaciones lineales de variables o factores que describe las principales tendencias en un conjunto de datos confuso (Jolliffe, 2002). El objetivo de PCA es encontrar el sub-espacio que describe la mayor varianza del espacio original de las variables. Las variables originales, comúnmente correlacionadas, se transforman linealmente en un menor número de variables no correlacionadas: las llamadas componentes principales (-del inglés *Principal Components (PCs)*-) (Camacho et al., 2010).

Sea X una matriz de datos $n \times m$ con n filas (observaciones) y m columnas (variables). El objetivo es transformar el conjunto de datos X en otro conjunto de datos de menor dimensión.

$$X = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,m} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n,1} & x_{n,2} & \cdots & x_{n,m} \end{bmatrix} \quad (5.1)$$

En primer lugar, las m variables se autoescalán para tener media igual a cero (restando la media de cada columna) y varianza unitaria (se divide cada columna con su desviación estándar) (Nomikos and MacGregor, 1994). El proceso de autoescalado de los datos se realiza con el objetivo de obtener variables con rangos similares y que tengan el mismo peso durante su análisis.

Luego, para datos autoescalados, se calcula la matriz de covarianza S con la siguiente expresión (ecuación (5.2)):

$$S = \frac{1}{n-1} X^T X \quad (5.2)$$

A continuación, se definen los diferentes sub-espacios mediante la descomposición de la matriz de covarianzas S en sus vectores y valores propios (ecuación (5.3)):

$$S P = P \Lambda \quad (5.3)$$

donde los vectores propios de S son las columnas de la matriz ortogonal $P = (\vec{p}_1 | \vec{p}_2 | \cdots | \vec{p}_j | \cdots | \vec{p}_m)$. La matriz P se denota como las PCs del conjunto de datos, también llamada matriz de *loadings*. Mientras que los valores propios de S se encuentran en la diagonal de la matriz $\Lambda = \text{diag}(\lambda_1, \lambda_2, \cdots, \lambda_m)$ (los términos fuera de la diagonal son cero) e indican la varianza de cada PC. Los vectores propios se ordenan de forma descendente (de mayor a menor) en función de los valores propios, ya que valores altos en los segundos indican una mayor variabilidad explicada en relación a los datos autoescalados. El hecho de sólo retener las k primeras PCs para el modelo (k , con $k \ll m$) permite explicar la mayor fuente de varianza de los datos originales con un número menor de variables.

La matriz de datos originales X se transforma por medio de la matriz P . A las proyecciones de X en P se les denomina *scores* y se almacenan en la matriz T . La matriz T se denota como la matriz proyectada o transformada al espacio de las PCs, también llamada matriz de *scores* (ecuación (5.4)).

$$T = XP \quad (5.4)$$

En el caso de la dimensión completa, esta proyección es invertible (desde $PP^T = I$) y los datos originales se pueden recuperar como $X = TP^T$. Ahora con la T dada no es posible totalmente recuperar X , pero T se puede proyectar de nuevo en el espacio de observaciones original m -dimensional y obtener otra matriz de datos de la siguiente manera:

$$\hat{X} = T_k P_k^T = \sum_{i=1}^k \vec{t}_i \vec{p}_i^T = \vec{t}_1 \vec{p}_1^T + \vec{t}_2 \vec{p}_2^T + \cdots + \vec{t}_k \vec{p}_k^T \quad (5.5)$$

donde \hat{X} es la proyección de la matriz de datos X en las k PCs seleccionadas, $T_k = [\vec{t}_1, \vec{t}_2, \cdots, \vec{t}_k]$ es la matriz de *scores* $n \times k$ que contiene las proyecciones de las observaciones de X en el subespacio de las PCs y $P_k = [\vec{p}_1, \vec{p}_2, \cdots, \vec{p}_k]$ es la matriz de *loadings* $m \times k$ que contiene las combinaciones lineales de las variables representadas en cada una de las PCs.

El modelo PCA completo se puede expresar matemáticamente como sigue:

$$X = \hat{X} + E = \sum_{i=1}^k \vec{t}_i \vec{p}_i^T + E = \vec{t}_1 \vec{p}_1^T + \vec{t}_2 \vec{p}_2^T + \cdots + \vec{t}_k \vec{p}_k^T + E \quad (5.6)$$

donde E es la matriz de residuos $n \times m$, la cuál resume la varianza no capturada por las k componentes seleccionadas en el nuevo espacio. En resumen, el conjunto de datos puede expresarse como una combinación lineal de k nuevas variables, asumiendo una matriz de error o residual E . Las matrices de *scores* y de *loadings* se computan para reflejar relaciones relevantes entre observaciones y variables.

Por otra parte, la matriz T y la matriz E se pueden utilizar para detectar un comportamiento anómalo en un sistema. Con este objetivo se utilizan dos estadísticos asociados con PCA: el estadístico *D-statistic* (también denotado como estadístico T^2 de Hotelling o en inglés *Hotelling's T^2 statistic*), y el estadístico *Q-statistic* (también conocido como error de predicción cuadrático o en inglés *Squared Prediction Error (SPE)*).

El *D-statistic* se pueden representar de la siguiente manera (MacGregor, 2003):

$$D - statistic(x) = \sum_{j=1}^k \frac{t_{jx}^2}{\lambda_j} \quad (5.7)$$

donde t_{jx} es la proyección de una observación \vec{x} en la j -ésima PC (ecuación (5.7)). *D-statistic* se computa para cada observación adicionando el cuadrado de las k componentes t_{jx}^2 ponderadas por sus respectivos valores propios λ_j . Esto resulta en una medida de distancia (distancia de Mahalanobis) de cada observación al centro del modelo.

Q-statistic representa los cambios en los eventos que no son explicados por el modelo PCA (ecuación (5.8)):

$$Q - statistic(x) = \sum_{j=1}^k e_{jx}^2 \quad (5.8)$$

donde $Q - statistic$ se obtiene para cada observación \vec{x} a partir de la suma de cuadrados de cada fila de E (e_{jx}). Durante un comportamiento normal en un sistema, $Q - statistic$ es muy pequeño, y por tanto, cualquier cambio menor en las características del sistema podría ser observable (Mujica et al., 2011).

$D - statistic$ y $Q - statistic$ asociados con un nuevo dato, se pueden calcular y comparar con los estadísticos correspondientes a los datos de entrenamiento. Si los estadísticos para el nuevo dato violan un nivel de confianza razonable que describe los datos de entrenamiento, por ejemplo un nivel de confianza del 95,0%, entonces el nuevo dato se clasifica como anormal. Los límites de los estadísticos $D - statistic$ y $Q - statistic$ correspondientes a los datos de entrenamiento se representan en la Figura 5.3. El límite de $D - statistic$ define una elipse en el plano dentro del cuál los datos normalmente se proyectan (representado en color rojo en la Figura 5.3). El $Q - statistic$ define la distancia fuera del plano (representado en color azul en la Figura 5.3).

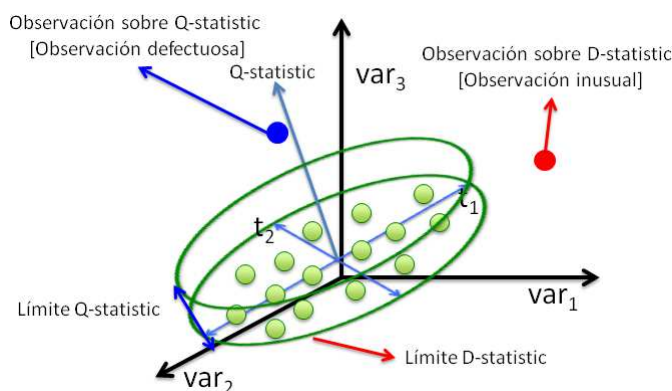


Figura 5.3: Representación gráfica de PCA

5.3.2. Construcción del modelo PCA

Para construir el modelo PCA se utiliza sólo la información de cuatro variables originales: *ISIG*, *GRTCGMS*, *Insulin* y *Temperature*. La variable *Sepsis* no se considera en la construcción del modelo PCA debido a que se trata de una variable categórica (que podría tomar los valores de 0, 1 y 2) mientras que X es una matriz de variables de proceso. Sin embargo, se realizaron varios experimentos previos donde se le considera como una variable de entrada y sus resultados indican que esta variable no contribuye a mejorar el modelo. Contrariamente, la variable *Sepsis* adiciona ruido al modelo PCA. Por tanto, se construye el modelo PCA de la siguiente manera:

- Se considera el conjunto de datos de 20 pacientes (432 observaciones) como una matriz X de $n \times m$, donde n representa el número de observaciones (432 observaciones) y m representa las 4 variables medidas en cada paciente (*ISIG*, *GRTCGMS*, *Insulin* y *Temperature*). Este conjunto de datos se compone de 321 observaciones etiquetadas como TC y 111 etiquetadas como TI. Los dos pacientes restantes se reservan para la evaluación de la metodología.

- Se realiza la identificación de los datos atípicos utilizando el criterio del diagrama de caja (Velleman and Hoaglin, 1981). Sólo se excluyen los datos atípicos con un nivel de severidad extremo, es decir, aquellos casos que están por encima del tercer cuartil más tres veces el rango intercuartílico ($Q3 + 3IQR$). 13 de las 432 observaciones originales se excluyen por ser valores extremos. Por lo tanto, el conjunto de datos para construir el modelo PCA se compone de 419 observaciones.
- Los datos a su vez se dividen en dos subconjuntos: un subconjunto de entrenamiento para generar el modelo PCA (consiste de 308 observaciones etiquetadas como TC) y un subconjunto de validación para proyectar sobre el modelo generado (formado por 111 observaciones etiquetadas como TI).
- Los datos del subconjunto de entrenamiento se autoescalan para tener media cero y varianza unitaria, lo cuál garantiza que todas las variables tengan la misma importancia.
- Utilizando los datos del subconjunto de entrenamiento previamente autoescalados se calculan los vectores propios y los valores propios de S .
- El criterio del porcentaje de varianza explicada por el modelo se utiliza para determinar el número específico de PCs a retener. El número de PCs se selecciona con el fin de explicar por lo menos 80,0% de la varianza de los datos originales.
- Usando los datos del subconjunto de entrenamiento previamente autoescalados se construye el modelo PCA considerando el número de PCs que se han seleccionado. De acuerdo con el criterio mencionado anteriormente, se mantienen las tres primeras PCs en el modelo PCA, las cuáles capturan el 91,2% de la varianza total acumulada. Los datos originales del subconjunto de entrenamiento se transforman por medio de la matriz P (proyección), obteniéndose los índices que caracterizan cada una de estas observaciones. Estos índices son los *scores*, el *D-statistic* y el *Q-statistic*. La Figura 5.4 muestra los valores propios del modelo PCA construido y la Tabla 5.3 indica la varianza capturada por cada una de las PCs del modelo y el porcentaje de varianza total acumulada.

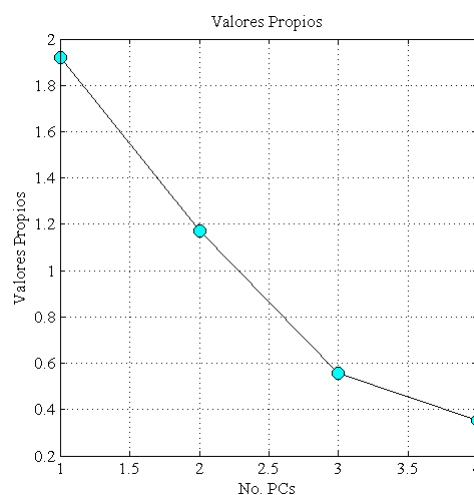


Figura 5.4: Valores propios

- Los datos del subconjunto de validación se autoescalan empleando para ello los valores de medias y desviaciones estándar calculados a partir de los datos del subconjunto de

Porcentaje de varianza capturada por el modelo PCA			
No. PC	Valores propios de S	Varianza explicada por cada PC (%)	Varianza acumulada (%)
1	1,920	48,00	48,00
2	1,170	29,24	77,24
3	0,558	13,95	91,19
4	0,352	8,81	100,00

Tabla 5.3: Varianza capturada por el modelo PCA.

entrenamiento.

- Los datos del subconjunto de validación previamente autoescalados se proyectan sobre el modelo PCA. Los *scores*, el *D-statistic* y el *Q-statistic* también se calculan para el subconjunto de validación.

En la Tabla 5.4 se puede observar el porcentaje de explicación de cada una de las variables originales para cada una de las PCs. La Figura 5.5 representa los *scores* (Figura 5.5(a)), el *D-statistic* (Figura 5.5(b)) y el *Q-statistic* (Figura 5.5(c)) tanto para el subconjunto de entrenamiento como para el subconjunto validación.

Variable original	Porcentaje de explicación		
	PC 1	PC 2	PC 3
<i>ISIG</i>	22,5	34,8	5,3
<i>GCGMS</i>	39,5	0,8	11,1
<i>Insulin</i>	30,4	4,8	64,0
<i>Temperature</i>	7,7	59,6	19,6

Tabla 5.4: Porcentaje de explicación de cada variable para cada PC.

Inicialmente se pensó aplicar PCA como una técnica de control estadístico multivariante, utilizándola para distinguir entre medidas TC y medidas TI realizadas por un SMCG. Se realizaron pruebas proyectando nuevos datos no utilizados en la construcción del modelo PCA con el fin de determinar si eran “similares” a los datos de entrenamiento, es decir, si estos nuevos datos eran descritos de manera correcta por el modelo PCA (recordemos que el modelo sólo se construye con medidas TC). Se calcularon los dos estadísticos (*D-statistic* y *Q-statistic*) para los nuevos datos y se compararon con los correspondiente límites de confianza para los datos de entrenamiento, en un procedimiento secuencial en el que primero se analizó si las nuevas observaciones se comportaban de acuerdo al modelo (ni el *D-statistic* ni el *Q-statistic* superaban sus respectivos límites) o no (cualquiera de los dos índices rebasaba su límite estadístico). Sin embargo, PCA no fue capaz de distinguir entre medidas TC y TI realizadas por un SMCG, debido a que los dos estadísticos eran similares para ambas tipologías. En la Figura 5.6 se puede observar la superposición de los estadísticos en las dos clases. Por esta razón, se decidió combinar PCA y CBR (como clasificador no lineal). En primer lugar, PCA se utiliza para extraer características que alimentan el sistema CBR. En segundo lugar, CBR se usa como una herramienta de decisión para distinguir entre ambas tipologías.

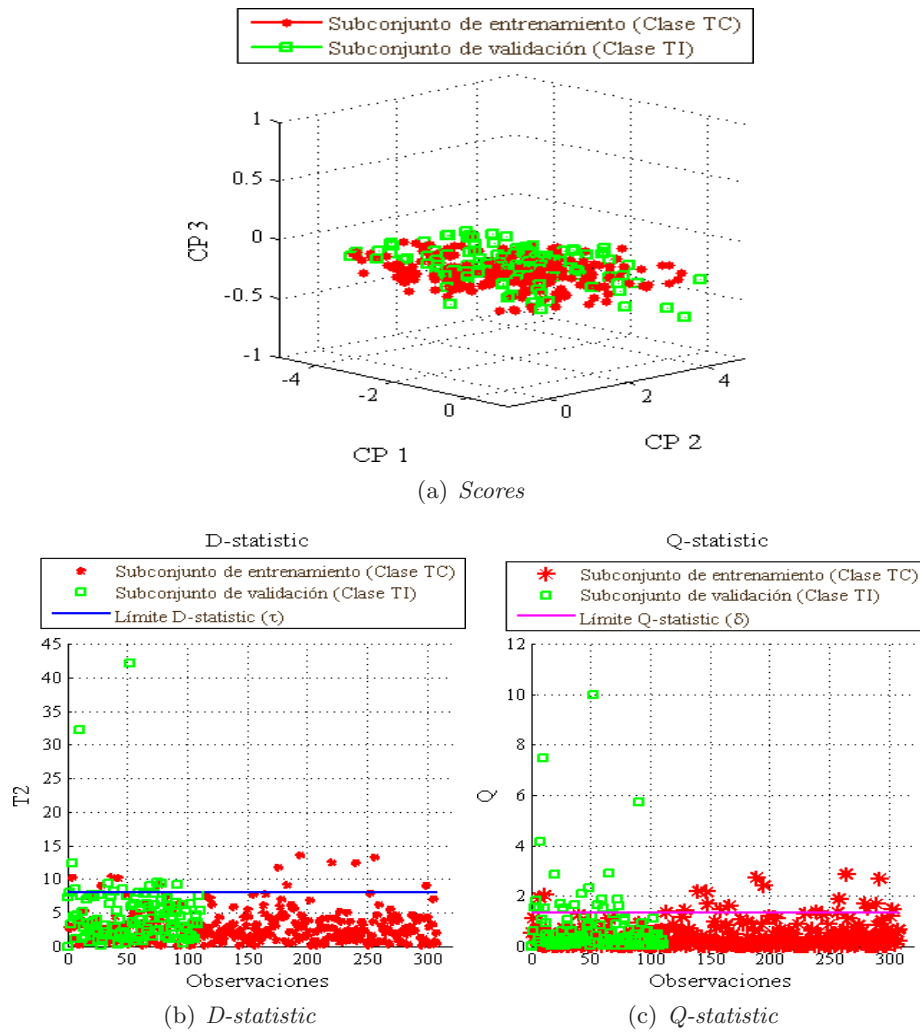
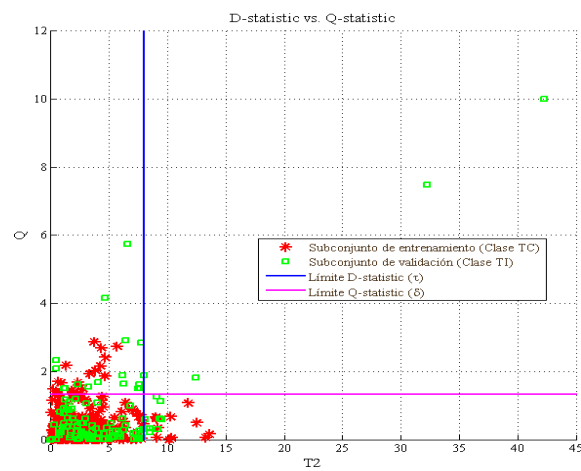


Figura 5.5: Scores, *D*-statistic y *Q*-statistic para los subconjuntos de entrenamiento y validación.



5.4. Detección de medidas TC y TI usando CBR

5.4.1. Conceptos básicos de CBR

CBR es una técnica de inteligencia artificial, la cuál se aplica en este capítulo para reconocer patrones y clasificar los datos extraídos de un modelo PCA. La característica principal del CBR es su capacidad para resolver problemas mediante el uso del conocimiento adquirido a partir de casos pasados, con la posibilidad de aprender para resolver nuevos casos. En el caso de la detección de fallos, la solución del problema significa que el sistema CBR propone un caso similar para identificar un fallo.

La base del CBR es la definición de caso. Un caso es la mínima representación de una experiencia pasada y su solución (Leake, 1996). Cuando existen varios casos disponibles, estos se pueden agrupar en una Base de Casos (BC). Una importante ventaja del CBR es que la BC se construye sólo una vez, y su mantenimiento y actualización se lleva a cabo a través de la capacidad de aprendizaje que tiene esta herramienta.

En el CBR los procesos implicados pueden ser representados por un ciclo esquemático de cuatro etapas básicas que son conocidas como las cuatro R's (Figura 5.7) (Aamodt and Plaza, 1994):

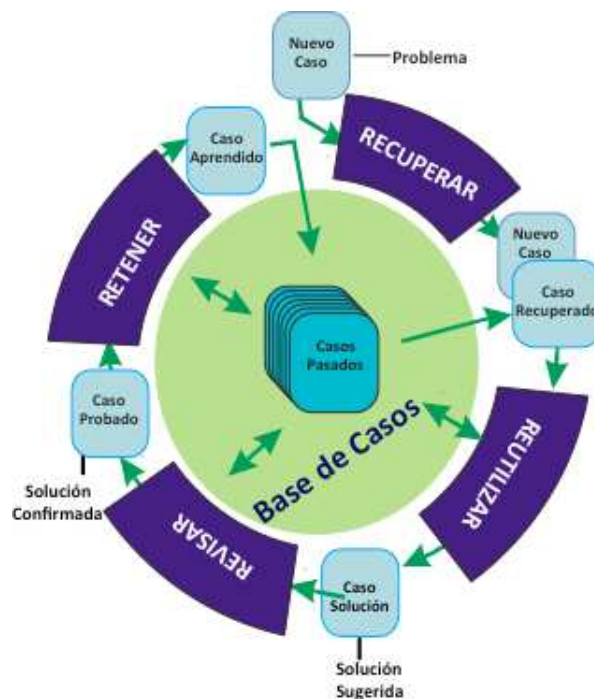


Figura 5.7: Ciclo de las 4-Rs

- Recuperar: Dado un problema objetivo, se recuperan los casos previos más similares para resolverlo. Es necesario definir un criterio y el número de casos (z) para recuperar.
- Reutilizar: Una vez que se han recuperado los z -casos más cercanos, estos se utilizan para proponer una posible solución. De acuerdo con un índice de similitud, el diagnóstico se da usando una técnica de votación o de adaptación de casos.

- Revisar: Una vez que se propone la solución al nuevo caso presentado, se revisa si la solución propuesta es adecuada, ya sea utilizando un modelo o por un experto.
- Retener: De acuerdo con la solución propuesta, se decide la conveniencia de mantener el nuevo caso para futuras clasificaciones. Si la solución se considera correcta y es lo suficientemente precisa, no es necesario retener el nuevo caso. Por otra parte, si la solución se considera errónea o tiene poca precisión, se actualiza la BC almacenando el nuevo caso para futuras decisiones.

5.4.2. Definición de los descriptores

La BC debe contener los descriptores y las soluciones o etiquetas asociados a cada caso. En la formulación del problema para distinguir entre glucemias TC y TI estimadas por un SMCG en una UCI, se propone una estructura de caso básico (ecuación (5.9)). Cada uno de estos casos (los cuáles contienen la descripción del problema y su solución) se almacenan en la BC. Los descriptores que componen la BC son los resultados derivados del modelo PCA. *Q-statistic* y *scores* los hemos denominado descriptores cuantitativos, y la variable categórica *Sepsis* la hemos denominado descriptor cualitativo (Figura 5.8):

$$\text{Caso}_x = \{Q_x, t_{1x}, t_{2x}, \dots, t_{kx}, \text{Sepsis}, l\} \quad (5.9)$$

donde Q_x es el *Q-statistic*, $t_{1x}, t_{2x}, \dots, t_{kx}$ son los k *scores* retenidos, y *Sepsis* es el valor asociado al estado séptico del paciente. Finalmente, l indica la etiqueta de cada caso, es decir si una medida es TC o TI. El *D-statistic* no se utiliza explícitamente para constituir una estructura de caso para evitar redundancia de información. Más detalles sobre esta cuestión se proporcionan en la sección 5.4.4.

Caso No.	Descriptores cuantitativos				Descriptor cualitativo
	Descriptor 1	Descriptor 2	...	Descriptor K	Descriptor K+1
Caso 1	Q-statistic	score1	...	scorek	Sepsis
Caso 2					
...					
Caso n					

Figura 5.8: Descriptores que constituyen una estructura de caso

5.4.3. Construcción de las BCs y de los conjuntos de validación

De acuerdo a la metodología CBR, una vez que se seleccionan como descriptores las variables cuantitativas y categóricas, se construyen las BCs iniciales. Teniendo en cuenta que el número de datos en cada clase está desbalanceado (308 TC/111 TI), para garantizar en la construcción de las diferentes BCs iniciales un equilibrio entre las clases, estas se construyen usando el mismo número de casos TC y TI, de manera similar a como se hizo en (Ruiz, 2008). Esta consideración se hace para que en el proceso de aprendizaje el sistema CBR no esté sesgado hacia la clase mayoritaria. La porción de datos restante se utiliza para construir los conjuntos para validar la metodología.

Teniendo en cuenta esta distribución de los casos en BCs y conjuntos de validación, se construyen aleatoriamente diez BCs diferentes con sus respectivos conjuntos de validación. Cada

una de estas BCs, a las que hemos denominado BCs iniciales está conformada por el 70% de los casos etiquetados como TI (77 casos) y el 25% de los casos etiquetados como TC (77 casos), todos ellos seleccionados aleatoriamente. Como se ha mencionado anteriormente, se utilizan porcentajes diferentes para cada clase con el fin de garantizar que todas las clases tengan el mismo número de casos en la BC. Por lo tanto, el resto de los casos, es decir el 30% de los casos etiquetados como TI (34 casos) y el 75% de los casos etiquetados como TC (231 casos) se utilizan para construir los respectivos conjuntos de validación. En consecuencia, cada BC inicial contiene 154 casos, y cada conjunto de validación contiene 265 casos (Tabla 5.5 y Figura 5.9).

No. BC	No. de datos BC inicial	No. conjunto de validación	No. de datos conjunto de validación
1	154	1	265
2	154	2	265
3	154	3	265
4	154	4	265
5	154	5	265
6	154	6	265
7	154	7	265
8	154	8	265
9	154	9	265
10	154	10	265

Tabla 5.5: BCs y sus correspondientes conjuntos de validación.

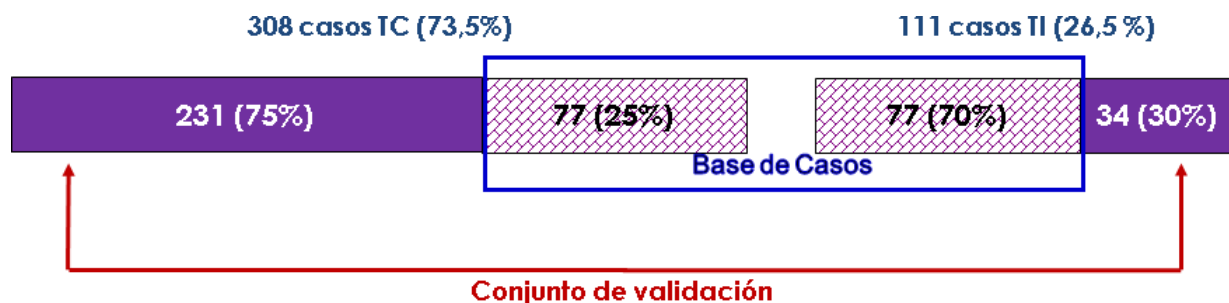


Figura 5.9: Distribución de los casos en BC y conjunto de validación

5.4.4. Aplicación del CBR

Una vez se han construido las BCs iniciales y sus respectivos conjuntos de validación, los casos que conforman los conjuntos de validación se deben comparar con las BCs iniciales. Con el objetivo de encontrar los mejores descriptores para recuperar los casos similares desde las diferentes BCs, se hicieron varias simulaciones combinando las diferentes salidas del modelo PCA, similar a como se hizo en (Ruiz, 2008). De todas estas pruebas, los enfoques denominados recuperación en dos pasos (-del inglés *Two-steps retrieval (2SR)*-) y recuperación en tres pasos (-del inglés *Three-steps retrieval (3SR)*-) ofrecen los mejores resultados entre las otras combinaciones posibles. *2SR* se basa en un criterio de distancia en dos pasos, el cuál utiliza los descriptores *Q-statistic* (en un primer nivel) y los *scores* (en un segundo nivel), y *3SR* se basa en un criterio de distancia en dos pasos similar al mencionado anteriormente y adicionando un criterio de similitud, el estado séptico. En ambos enfoques (*2SR* y *3SR*), cuando se recuperan

los casos similares, estos se seleccionan y reutilizan para diagnosticar la clase para el nuevo caso a validar. Si la solución para este nuevo caso a validar no es satisfactoria, esta se revisa y se retiene dentro de la BC para futuros diagnósticos, conformando una BC actualizada a la que hemos denominado BC final. Mayores detalles sobre *2SR* y *3SR* se explican a continuación:

- *2SR*: Los z casos más cercanos se recuperan basándose sólo en un criterio de distancia en dos pasos. En el primer paso se utiliza sólo el descriptor *Q-statistic* y se calcula la distancia de un nuevo caso a cada uno de los casos que conforman la BC respectiva, por medio de la ecuación (5.10). Esta distancia a la que se ha denominado *Q-distance* (d_Q) es la medida de la distancia absoluta entre el descriptor *Q-statistic* del nuevo caso ($Q_{x\ new}$) y el descriptor *Q-statistic* del i -ésimo caso almacenado en la BC respectiva ($Q_{x\ iCB}$). Luego, los primeros $z_1 = 30$ casos más cercanos, es decir, los 30 casos que tienen la menor *Q-distance*, se almacenan para usarse en el siguiente paso. Inmediatamente después, se aplica el segundo paso, donde se utilizan los *scores* y se computa la distancia entre el nuevo caso y cada uno de los z_1 casos recuperados, por medio de la ecuación (5.11). Esta distancia a la que se ha denominado *score-distance* (d_t) es la medida de la distancia euclidiana ponderada que se calcula entre cada descriptor *score* del nuevo caso ($t_{1x\ new}, t_{2x\ new}, \dots, t_{kx\ new}$) y cada descriptor *score* del i -ésimo z_1 caso recuperado de la BC ($t_{1x\ iCB}, t_{2x\ iCB}, \dots, t_{kx\ iCB}$). De este segundo nivel de cómputo de la distancia, sólo se recuperan los primeros $z_2 = 15$ casos más cercanos, es decir, los que tienen la menor *score-distance*. Finalmente, se obtiene la clase del nuevo caso por votación a partir de los z_2 casos recuperados, esto significa que la clase con la mayor población entre los 15 casos recuperados se selecciona como la clase del nuevo caso.

$$d_Q = |Q_{x\ new} - Q_{x\ iCB}| \quad (5.10)$$

$$d_t = \sqrt{\sum_{j=1}^k \frac{(t_{jx\ new} - t_{jx\ iCB})^2}{\lambda_j}} \quad (5.11)$$

Respecto a la selección de $z_1 = 30$ y $z_2 = 15$, se hicieron múltiples experimentos y esta combinación del número de vecinos es la que ofrece los mejores resultados entre las otras combinaciones posibles, porque presenta los valores más altos de las métricas para evaluar el rendimiento del clasificador.

Un punto crítico en la metodología CBR es asignar los pesos a las variables, porque generalmente esta decisión está sujeta al conocimiento del experto. Este problema se resuelve asignando a cada descriptor el valor propio calculado por PCA asociado a cada *score* como peso (λ_j), con lo cuál se garantiza que todas las variables puedan tener la misma importancia. Estos valores propios representan la varianza para cada *score*.

- *3SR*: Los z casos más cercanos se recuperan en tres pasos: utilizando el mismo criterio de distancia en dos pasos y el criterio de similitud entre los casos. Tal como se desarrolló en *2SR*, en los dos primeros pasos de este enfoque se consideran d_Q y d_t para recuperar los primeros z_1 y z_2 casos más cercanos, respectivamente. En el tercer paso, se aplica un criterio de similitud entre los casos a partir de la similitud de la variable *Sepsis*, empleándose este criterio sólo en los z_2 casos recuperados. Así, se identifican y recuperan los z_3 casos que toman el mismo valor en la variable *Sepsis* que el nuevo caso. Este criterio de similitud se

propone teniendo en cuenta las conclusiones del estudio clínico descrito en el Capítulo 4, donde los resultados reportados mostraron que el estado séptico del paciente puede influir en la precisión de las estimaciones de un SMCG cuando se utiliza en pacientes de UCI. Una vez se han recuperado los z_3 casos, se determina la solución del nuevo caso por votación. Si ninguno de los z_2 casos recuperados toma el mismo valor en la variable *Sepsis* que el nuevo caso, por tanto $z_3 = 0$ y se obtiene la clase del nuevo caso por votación a partir de los z_2 casos recuperados, como en *2SR*.

Es importante mencionar que en ambos procesos de recuperación (*2SR* y *3SR*) se implementa completamente el ciclo de las 4-Rs y se actualiza cada una de las BCs obteniendo las BCs finales. Dependiendo de cuál sea el procedimiento de recuperación que se utilice cuando se aplica CBR, a la metodología se le ha denominado *PCA-CBR-2SR* o *PCA-CBR-3SR*.

PCA-CBR-2SR y *PCA-CBR-3SR* se basan en el *Q-statistic* y los *scores*, por tanto, ellos contienen información complementaria (el *Q-statistic* mira en el subespacio residual, mientras los *scores* buscan dentro del subespacio de proyección). Además, aunque el *D-statistic* no se utiliza explícitamente como entrada del sistema CBR, d_t es una diferencia *score a score* entre observaciones, y por tanto, tiene en cuenta la misma información que los valores del *D-statistic*. La combinación del *D-statistic* y los *scores* fue descartada porque proporciona redundancia de información y no mejoraba los resultados.

5.5. Puesta a punto de la metodología PCA-CBR

En primer lugar, se utilizan los 10 conjuntos de validación descritos en la Tabla 5.5 con el objetivo de evaluar la metodología PCA-CBR. Para ello, cada caso del conjunto de validación se considera como un nuevo caso, y se compara con todos los casos que integran la BC inicial asociada al conjunto de validación (también descritas en la Tabla 5.5) para recuperar los casos similares, considerando los dos enfoques (*PCA-CBR-2SR* y *PCA-CBR-3SR*) descritos en la sección 5.4.4. A partir del re-uso de los casos recuperados basándose en los z vecinos más cercanos se propone una clase para el caso a validar (TC o TI). Seguidamente, se revisa la clase predicha comparándola con la clase real. Si la clase que se propone no coincide con la clase real, el caso a validar se retiene dentro de la BC para futuros diagnósticos y pasa a conformar la BC final.

La evaluación de la revisión se realiza utilizando las siguientes métricas basadas en la matriz de confusión definidas en la sección 2.4.2: la *proporción de aciertos*, la *especificidad*, la *sensibilidad* y la G_{media} . Los resultados promedios obtenidos de todos los experimentos se presentan en la Tabla 5.6.

A partir de esta tabla se puede llegar a las siguientes conclusiones:

- La principal conclusión es que la mediana (Q1 - Q3) de los resultados de PCA-CBR-2SR es mejor que la de PCA-CBR-3SR en *Proporción de aciertos* = 67,2% (64,0% - 68,5%) versus 59,2% (56,0% - 62,8%) y *Especificidad* = 68,6% (64,2% - 69,2%) versus 58,2% (54,1% - 61,6%), pero no lo es en la *Sensibilidad* = 63,2% (61,8% - 66,9%) versus 70,6% (60,3% - 75,7%). La causa de las diferencias en los resultados de PCA-CBR-2SR y PCA-CBR-3SR se debe al re-uso de los casos recuperados, debido a que el método para determinar la clase de un nuevo caso es diferente entre ambos enfoques. En PCA-CBR-3SR se tiene en cuenta

<i>PCA-CBR-2SR</i>							
No. BC	No. de datos BC inicial	No. de datos BC final	No. de datos conjunto de validación	Proporción de aciertos (%)	Especificidad (%)	Sensibilidad (%)	G_{media} (%)
1	154	215	265	77,0	76,2	82,4	79,2
2	154	251	265	63,4	63,6	61,8	62,7
3	154	239	265	67,9	68,8	61,8	65,2
4	154	252	265	63,0	62,3	67,6	64,9
5	154	264	265	58,5	58,0	61,8	59,9
6	154	240	265	67,5	68,8	58,8	63,6
7	154	237	265	68,7	69,3	64,7	66,9
8	154	245	265	65,7	65,8	64,7	65,3
9	154	242	265	66,8	68,4	55,9	61,8
10	154	208	265	79,6	79,2	82,4	80,8
Media±DS				67,8±6,3	68,1±6,3	66,2±9,1	67,0±7,1
Mediana (IQR)				67,2 (4,5)	68,6 (5,0)	63,2 (5,1)	65,1 (3,6)
<i>PCA-CBR-3SR</i>							
No. BC	No. de datos BC inicial	No. de datos BC final	No. de datos conjunto de validación	Proporción de aciertos (%)	Especificidad (%)	Sensibilidad (%)	G_{media} (%)
1	154	228	265	72,1	70,6	82,4	76,2
2	154	257	265	61,1	60,6	64,7	62,6
3	154	271	265	55,8	55,4	58,8	57,1
4	154	269	265	56,6	53,7	76,5	64,1
5	154	282	265	51,7	50,6	58,8	54,6
6	154	251	265	63,4	61,9	73,5	67,5
7	154	260	265	60,0	58,0	73,5	65,3
8	154	264	265	58,5	58,4	58,8	58,6
9	154	272	265	55,5	53,7	67,6	60,3
10	154	219	265	75,5	73,6	88,2	80,6
Media±DS				61,0±7,5	59,7±7,4	70,3±10,3	64,7±8,3
Mediana (IQR)				59,2 (6,8)	58,2 (7,5)	70,6 (15,4)	63,3 (7,9)

Tabla 5.6: Resultados de las métricas *proporción de aciertos*, *especificidad*, *sensibilidad* y G_{media} obtenidos al evaluar la metodología PCA-CBR-2SR y PCA-CBR-3SR.

el estado séptico de los pacientes para indicar si el caso es TC o TI, una variable que no se considera en PCA-CBR-2SR.

- En los conjuntos de validación, se presentan mejores resultados en la *Sensibilidad* y se incrementa también la variabilidad cuando se aplica PCA-CBR-3SR. Sin embargo, estos resultados podrían ser un indicador de una mejor capacidad para detectar casos TI. La consideración de la variable *Sepsis* como criterio de similitud para determinar la clase predicha ha producido este efecto, puesto que condiciona la forma como se propone una clase para el caso a validar. Por tanto se infiere que esta variable puede influir de alguna manera en una mejor identificación de los casos TI.
- En la Tabla 5.6 también se puede observar que el número de casos en las BCs iniciales tanto en PCA-CBR-2SR como en PCA-CBR-3SR es 154, tal como se indicó en la Tabla 5.5. Sin embargo, el número de casos en las BCs finales para los dos enfoques es mayor que en las BCs iniciales. Esto es debido al aprendizaje de las BCs, ya que los casos para los cuáles la clase predicha era errónea, identificados durante la etapa de revisión, son retenidos por el CBR. Además, se puede apreciar que el número de casos retenidos dentro de cada una de las BCs es superior en PCA-CBR-3SR, indicando que los errores en la clasificación son más altos en PCA-CBR-3SR que en PCA-CBR-2SR.
- Basándose en la métrica G_{media} para seleccionar cuál es la mejor BC, se observa que la BC No. 1 y la BC No. 10 obtienen mejores resultados para ambos enfoques. Siendo estas BCs

las que producen un mejor rendimiento en la clasificación, también son las que retienen un menor número de casos en el aprendizaje. De manera análoga, la peor BC tanto para PCA-CBR-2SR como para PCA-CBR-3SR es la BC No. 5 y por ende, es la que reporta el mayor número de casos aprendidos.

Una importante observación, durante la puesta a punto de la metodología también se puede identificar que existen varios casos cuya clase ha sido predicha de manera equivocada, es decir, frecuentemente se han clasificado mal, independientemente de la BC que se utilice para recuperar los z vecinos más cercanos. Por tanto, antes de actualizar la BC con ellos, es importante decidir si es idóneo mantenerlos o removerlos de las BCs (y de los conjuntos de validación). Muchos de estos casos que frecuentemente se clasifican mal se derivan de errores humanos en el registro de las variables de entrada, o de errores humanos en el uso efectivo de los dispositivos para medir dichas variables (por ejemplo en el uso del Guardian[®] RTCGMS y del HemoCue[®] 201 DM). Por lo tanto, se determina que si al usar el mismo enfoque (PCA-CBR-2SR y PCA-CBR-3SR), un caso se clasifica erróneamente en por lo menos seis de las diez BCs, este se remueve. En consecuencia, cuando se aplica PCA-CBR-2SR, 49 casos se excluyen tanto de las BCs como de los conjuntos de validación. Del mismo modo, cuando se aplica PCA-CBR-3SR, 63 casos se excluyen tanto de las BCs como de los conjuntos de validación. Una vez que se han eliminado los casos poco fiables, se aplica nuevamente todo el procedimiento descrito anteriormente. Los nuevos resultados promedios obtenidos de todos los experimentos se presentan en la Tabla 5.7.

<i>PCA-CBR-2SR</i>							
No. BC	No. de datos BC inicial	No. de datos BC final	No. de datos conjunto de validación	Proporción de aciertos (%)	Especificidad (%)	Sensibilidad (%)	G_{media} (%)
1	148	186	222	79,7	78,8	84,8	81,8
2	143	221	227	65,6	65,3	67,6	66,4
3	144	205	226	73,0	75,1	60,6	67,5
4	151	215	219	70,8	71,0	69,7	70,3
5	147	216	223	69,1	69,8	64,7	67,2
6	143	197	227	76,2	77,2	70,6	73,8
7	145	209	225	71,6	71,9	69,7	70,8
8	147	203	223	74,9	76,3	66,7	71,3
9	148	203	222	75,2	78,7	55,9	66,4
10	144	180	226	83,2	82,4	87,9	85,1
Media±DS				73,9±5,1	74,7±5,1	69,8±9,8	72,1±6,5
Mediana (IQR)				73,9 (5,0)	75,7 (7,1)	68,7 (5,2)	70,6 (5,9)
<i>PCA-CBR-3SR</i>							
No. BC	No. de datos BC inicial	No. de datos BC final	No. de datos conjunto de validación	Proporción de aciertos (%)	Especificidad (%)	Sensibilidad (%)	G_{media} (%)
1	144	184	212	76,9	75,4	84,8	80,0
2	136	206	220	68,2	66,7	76,5	71,4
3	142	209	214	68,7	69,6	63,6	66,6
4	149	208	207	71,5	70,1	78,8	74,3
5	143	220	213	63,8	64,8	58,8	61,7
6	139	205	217	69,6	68,9	73,5	71,2
7	146	202	210	73,3	72,9	75,8	74,3
8	144	206	212	70,8	71,5	66,7	69,0
9	145	207	211	70,6	71,8	64,7	68,1
10	136	174	220	86,4	85,0	93,9	89,4
Media±DS				72,0±6,1	71,7±5,6	73,7±10,7	72,6±7,7
Mediana (IQR)				70,7 (4,0)	70,8 (3,6)	74,6 (13,0)	71,3 (6,0)

Tabla 5.7: Resultados de las métricas *proporción de aciertos*, *especificidad*, *sensibilidad* y G_{media} obtenidos al evaluar la metodología PCA-CBR-2SR y PCA-CBR-3SR después de remover los datos que no son fiables.

De esta tabla también se pueden extraer algunas conclusiones:

- PCA-CBR-2SR reporta resultados ligeramente mayores que PCA-CBR-3SR en la detección de casos TC (*Especificidad* = 75,7 % (71,2 % - 78,3 %) versus 70,8 % (69,0 % - 72,6 %)). En contraste, PCA-CBR-3SR indica resultados ligeramente mayores en la detección de casos TI (*Sensibilidad* = 74,6 % (65,2 % - 78,2 %) versus 68,7 % (65,2 % - 70,4 %)). En general, la metodología PCA-CBR alcanza un rendimiento balanceado aceptable entre *especificidad* y *sensibilidad*.
- Respecto a determinar cuál es la mejor BC (considerando como criterio de decisión la G_{media}), los mejores resultados en ambos enfoques se siguen obteniendo con la BC No. 1 y la BC No. 10. En relación con la peor BC, para PCA-CBR-2SR es la BC No. 2 y para PCA-CBR-3SR es la BC No. 5. El número de casos retenidos en estas BCs es inversamente proporcional al rendimiento dado por cada una de ellas.
- Se puede observar que el número de casos aprendidos para actualización de cada una de las BCs se ha reducido en ambos enfoques, aunque continua siendo superior en PCA-CBR-3SR.

La Figura 5.10 resume la metodología PCA-CBR para detectar glucemias TC y TI estimadas por un SMCG a partir de información suministrada por el mismo dispositivo y algunas variables propias de la condición clínica del paciente.

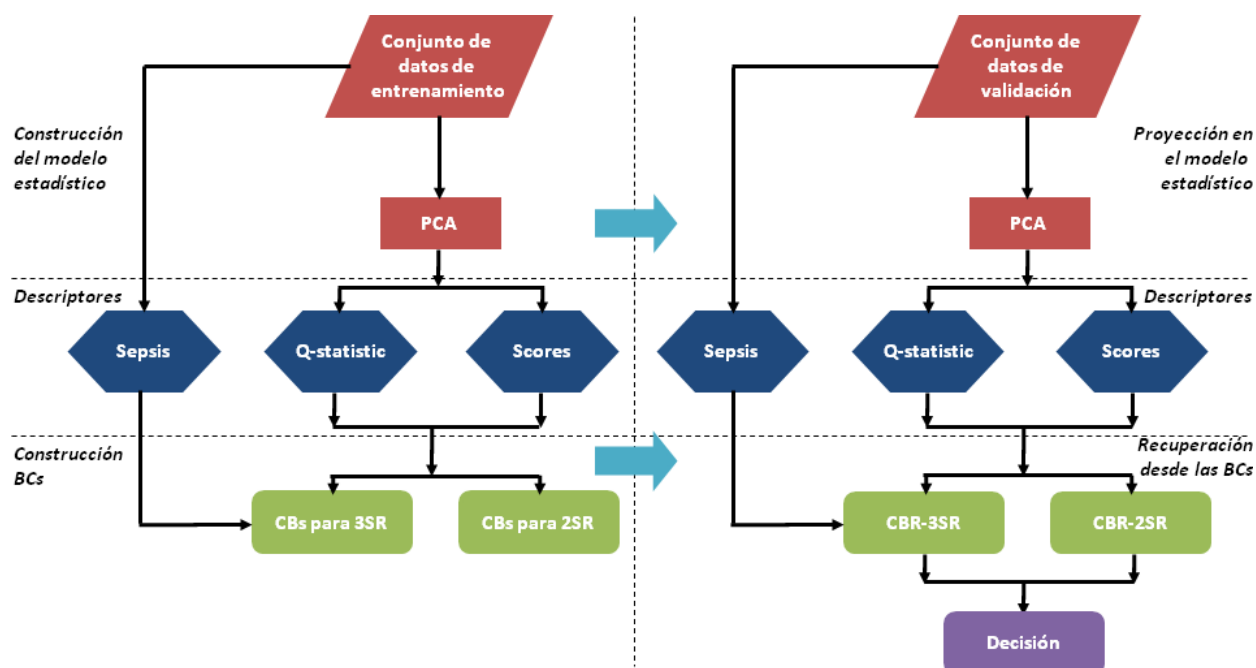


Figura 5.10: Diagrama metodología PCA-CBR

5.6. Evaluación comparativa de la metodología PCA-CBR con SIMCA y PLS-DA

Dado que se está proponiendo una nueva metodología, su rendimiento se puede validar comparándola con dos de los más comunes métodos multivariantes utilizados para clasificación: SIMCA (Wold et al., 1984) y PLS-DA (Sjöstrom et al., 1986). Algunos conceptos básicos acerca de estos dos métodos multivariantes se proporcionan en el Apéndice A.

Se realizan nuevos experimentos considerando SIMCA y PLS-DA. La comparación entre la metodología PCA-CBR con SIMCA y PLS-DA utiliza los mismos casos utilizados en cada una de las BCs iniciales y los respectivos conjuntos de validación (presentados en la Tabla 5.7). Sin embargo, estos conjuntos de entrenamiento y de validación no se construyen utilizando el Q -*statistic* y los *scores*. SIMCA y PLS-DA se aplican utilizando los datos antes de ser proyectados en el modelo PCA, es decir, considerando las variables originales (autoescaladas) como variables de entrada para construir el modelo. Dos enfoques diferentes se usan en las dos técnicas para construir los conjuntos de entrenamiento y sus correspondientes conjuntos de validación:

- Enfoque 1. Utilizando *ISIG*, *G_{RTCGMS}*, *Insulin* y *Temperature* como variables de entrada.
- Enfoque 2. Usando *ISIG*, *G_{RTCGMS}*, *Insulin*, *Temperature* y *Sepsis* como variables de entrada.

Por tanto, para efectos de comparar con PCA-CBR-2SR se emplea el enfoque 1 y se construyen 10 conjuntos de entrenamiento y 10 conjuntos de validación (utilizando en cada uno de ellos los mismos casos empleados en la Tabla 5.7 para PCA-CBR-2SR). De manera análoga, para efectos de comparar con PCA-CBR-3SR se emplea el enfoque 2 y también se construyen 10 conjuntos de entrenamiento y 10 conjuntos de validación (utilizando en cada uno de ellos los mismos casos empleados en la Tabla 5.7 para PCA-CBR-3SR). Con cada uno de los conjuntos de entrenamiento se construye un modelo aplicando SIMCA y otro aplicando PLS-DA. Luego, se proyectan los datos del conjunto de validación en el modelo construido para evaluar el rendimiento del mismo en términos de *proporción de aciertos*, *especificidad*, *sensibilidad* y G_{media} . Las Tablas 5.8 y 5.9 muestran los resultados obtenidos de esta evaluación con SIMCA y PLS-DA respectivamente, para los enfoques 1 y 2.

De este análisis se puede concluir:

- SIMCA reporta resultados similares tanto para el enfoque 1 como para el enfoque 2, lo cuál indica que considerar la variable *Sepsis* como una variable de entrada no contribuye a mejorar el rendimiento del modelo. PLS-DA también reporta resultados similares en ambos enfoques, por tanto, esta conclusión también es válida para este método.
- En SIMCA las diferencias entre los valores de *sensibilidad* y *especificidad* son amplias, tanto en el enfoque 1. como en el 2. Específicamente, en este problema SIMCA exhibe un rendimiento muy bueno en la detección de medidas TI (*Sensibilidad* = 85,3% (80,1% - 93,4%) y 86,8% (83,1% - 91,2%) para los enfoques 1 y 2, respectivamente), pero muy pobre en la detección de medidas TC (*Especificidad* = 30,3% (26,4% - 36,3%) y 29,4% (24,6% - 37,2%) para los enfoques 1 y 2, respectivamente). Por tanto, aunque SIMCA garantiza una máxima tasa de detección de estimaciones TI, la máxima *especificidad* que se logra es demasiado baja para ser aceptada clínicamente. Dado que la clase TI es la clase minoritaria, los resultados de la *proporción de aciertos* y G_{media} también son muy bajos.

5.6. Evaluación comparativa de la metodología PCA-CBR con SIMCA y PLS-DA

			SIMCA considerando como variables de entrada al modelo el enfoque 1			
No. conjunto de entrenamiento	No. de datos conjunto de entrenamiento	No. de datos conjunto de validación	Proporción de aciertos (%)	Especificidad (%)	Sensibilidad (%)	G_{media} (%)
1	148	222	23,8	13,9	91,2	35,5
2	143	227	30,2	22,9	79,4	42,7
3	144	226	45,3	39,4	85,3	58,0
4	151	219	37,7	29,4	94,1	52,6
5	147	223	34,0	29,0	67,6	44,3
6	143	227	41,9	35,9	82,4	54,4
7	145	225	38,1	31,2	85,3	51,6
8	147	223	41,5	36,4	76,5	52,7
9	148	222	34,7	25,5	97,1	49,8
10	144	226	49,1	42,4	94,1	63,2
Media±DS			37,6±7,4	30,6±8,5	85,3±9,2	50,5±7,9
Mediana (IQR)			37,9 (7,6)	30,3 (9,8)	85,3 (13,2)	52,1 (8,3)
			SIMCA considerando como variables de entrada al modelo el enfoque 2			
No. conjunto de entrenamiento	No. de datos conjunto de entrenamiento	No. de datos conjunto de validación	Proporción de aciertos (%)	Especificidad (%)	Sensibilidad (%)	G_{media} (%)
1	144	212	32,1	22,9	94,1	46,5
2	136	220	41,9	37,2	73,5	52,3
3	142	214	44,5	37,2	94,1	59,2
4	149	207	32,8	24,2	91,2	47,0
5	143	213	32,1	25,5	76,5	44,2
6	139	217	43,4	37,2	85,3	56,4
7	146	210	32,5	24,2	88,2	46,2
8	144	212	38,5	31,6	85,3	51,9
9	145	211	35,5	27,3	91,2	49,9
10	136	220	46,0	40,7	82,4	57,9
Media±DS			37,9±5,6	30,8±6,8	86,2±7,1	51,1±5,3
Mediana (IQR)			37,0 (10,5)	29,4 (12,7)	86,8 (8,1)	50,9 (8,7)

Tabla 5.8: Resultados de aplicar SIMCA.

- En PLS-DA se observan diferencias importantes entre los valores de *sensibilidad* y *especificidad* en ambos enfoques. En particular, los resultados favorecen la *especificidad* en detrimento de la *sensibilidad*, aunque los valores que se obtienen no son altos (*Especificidad* = 68,6% (65,5% - 76,3%) y 71,0% (67,6% - 73,2%) para los enfoques 1 y 2, respectivamente).
- SIMCA y PLS-DA son clasificadores lineales. De acuerdo con los resultados obtenidos con estos dos métodos y lo observado cuando se aplicó PCA como una técnica de control estadístico multivariante, los métodos lineales tienen problemas en fijar la “frontera” para la detección de medidas entre TC y TI en los SMCGs. La combinación de PCA y CBR tiene un mejor comportamiento porque utiliza un sistema no-lineal para determinar la clase de las nuevas observaciones, mientras que los dos métodos lineales (SIMCA y PLS-DA) presentan valores por debajo de los resultados obtenidos con esta metodología.

En la Figura 5.11 se ilustran y comparan los resultados de la media (DS) obtenidos usando PCA-CBR-2SR y los resultados de SIMCA y PLS-DA (obtenidos considerando el enfoque 1).

En la Figura 5.12 se ilustran y comparan los resultados de la media (DS) obtenidos utilizando PCA-CBR-3SR y los resultados de SIMCA y PLS-DA (obtenidos considerando el enfoque 2).

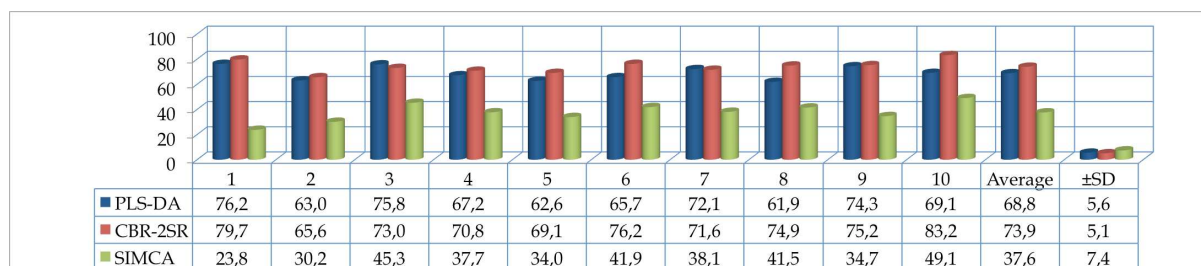
			PLS-DA considerando como variables de entrada al modelo el enfoque 1			
No. conjunto de entrenamiento	No. de datos conjunto de entrenamiento	No. de datos conjunto de validación	Proporción de aciertos (%)	Especificidad (%)	Sensibilidad (%)	G_{media} (%)
1	148	222	76,2	81,0	44,1	59,8
2	143	227	63,0	63,6	58,8	61,2
3	144	226	75,8	77,5	64,7	70,8
4	151	219	67,2	67,1	67,6	67,4
5	147	223	62,6	63,2	58,8	61,0
6	143	227	65,7	68,4	47,1	56,7
7	145	225	72,1	72,7	67,6	70,1
8	147	223	61,9	64,9	41,2	51,7
9	148	222	74,3	79,7	38,2	55,2
10	144	226	69,1	68,8	70,6	69,7
Media±DS			68,8±5,6	70,7±6,6	55,9±12,2	62,9±6,8
Mediana (IQR)			68,1 (10,1)	68,6 (10,8)	58,8 (22,1)	61,1 (11,6)
			PLS-DA considerando como variables de entrada al modelo el enfoque 2			
No. conjunto de entrenamiento	No. de datos conjunto de entrenamiento	No. de datos conjunto de validación	Proporción de aciertos (%)	Especificidad (%)	Sensibilidad (%)	G_{media} (%)
1	144	212	75,8	81,4	38,2	55,8
2	136	220	70,6	72,7	55,9	63,8
3	142	214	74,0	75,8	61,8	68,4
4	149	207	67,5	67,5	67,6	67,6
5	143	213	67,5	69,3	55,9	62,2
6	139	217	71,3	73,2	58,8	65,6
7	146	210	72,1	73,2	64,7	68,8
8	144	212	64,5	67,5	44,1	54,6
9	145	211	62,3	63,6	52,9	58,0
10	136	220	67,5	68,0	64,7	66,3
Media±DS			69,3±4,2	71,2±5,1	56,5±9,4	63,4±5,3
Mediana (IQR)			69,1 (4,3)	71,0 (5,5)	57,4 (10,3)	64,7 (8,2)

Tabla 5.9: Resultados de aplicar PLS-DA.

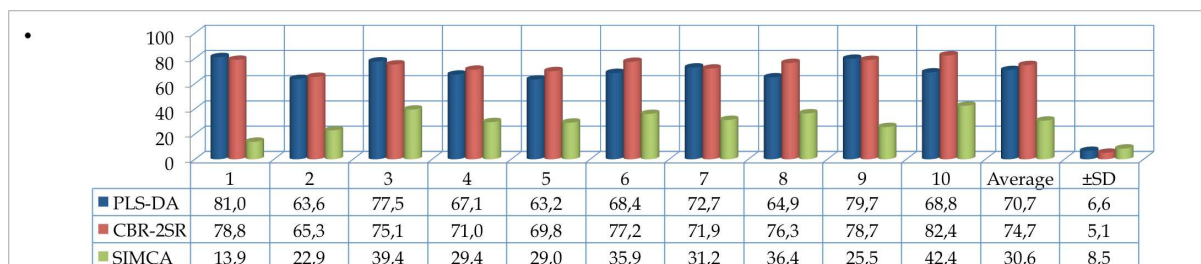
De acuerdo a estos resultados, la metodología PCA-CBR logra un rendimiento más balanceado entre *sensibilidad* y *especificidad* que el que se logra con SIMCA o PLS-DA. Por un lado, no existen diferencias importantes entre la metodología PCA-CBR y PLS-DA en cuanto a *proporción de aciertos* y *especificidad*; sin embargo, el uso de PCA-CBR reporta mejores resultados en términos de *sensibilidad* y G_{media} . Por otra parte, la metodología PCA-CBR presenta importantes diferencias con respecto a SIMCA. Más específicamente, SIMCA favorece mucho la detección correcta de medidas TI, por tanto la *sensibilidad* que se obtiene con este método es alta, pero a expensas de reducir notablemente la *especificidad*.

5.7. Evaluación de la metodología PCA-CBR en dos pacientes representativos

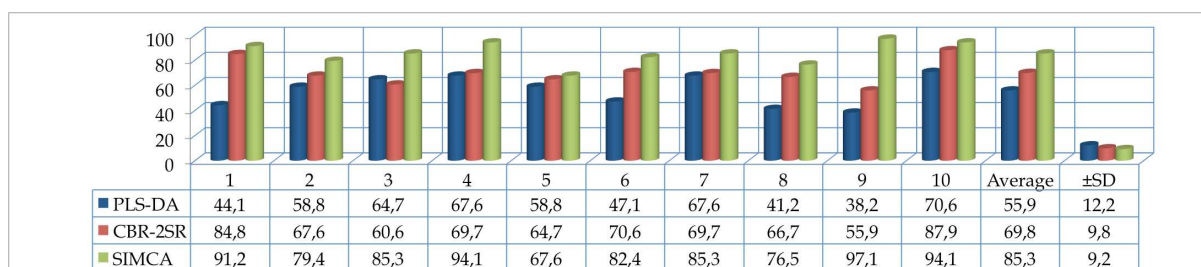
Dado que la metodología PCA-CBR logra un rendimiento más balanceado entre *sensibilidad* y *especificidad* en comparación con SIMCA y PLS-DA, se evalúa desde el punto de vista del uso generalizado, su capacidad para detectar medidas TC y TI hechas por el Guardian[®] RTCGMS. Para ello, se realiza el análisis de los datos de dos pacientes críticamente enfermos que se han reservado en la sección 5.2, excluidos con este propósito. Estos pacientes corresponden



(a) Proporción de aciertos



(b) Especificidad



(c) Sensibilidad

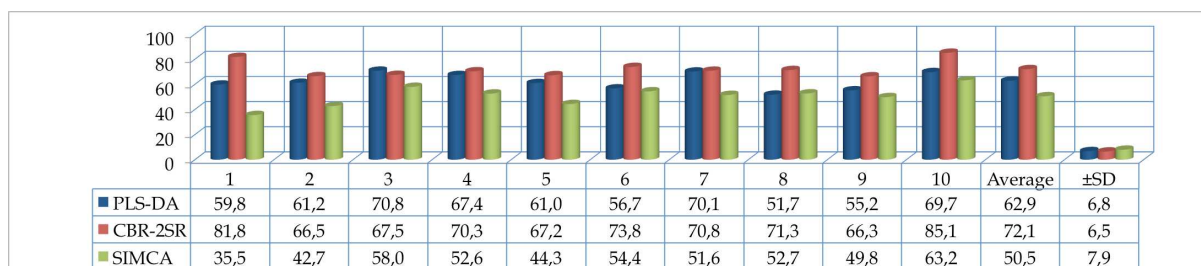
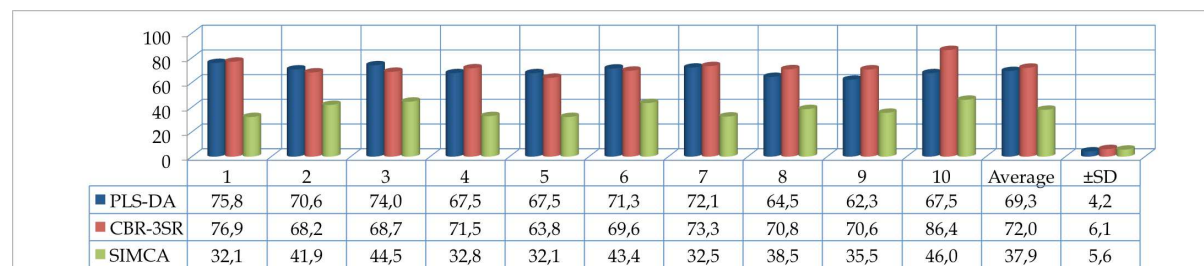
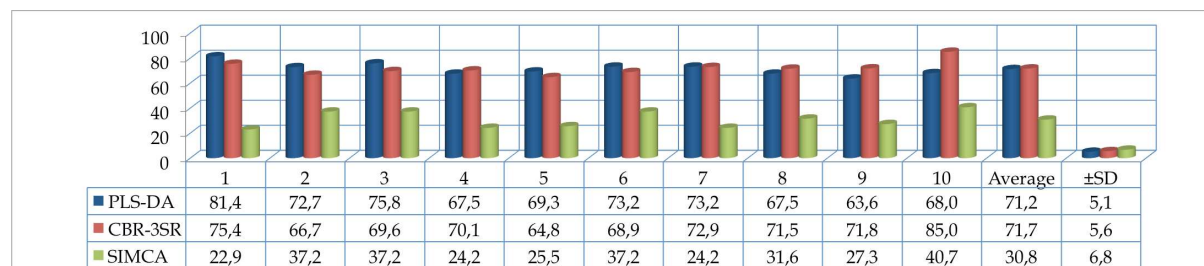

 (d) G_{media}

Figura 5.11: Resultados comparativos de PCA-CBR-2SR (obtenidos considerando como variables de entrada: Q -statistic y scores) y resultados de SIMCA y PLS-DA (obtenidos considerando como variables de entrada a los modelos el enfoque 1). Los resultados mostrados en las diferentes tablas son porcentajes (%).

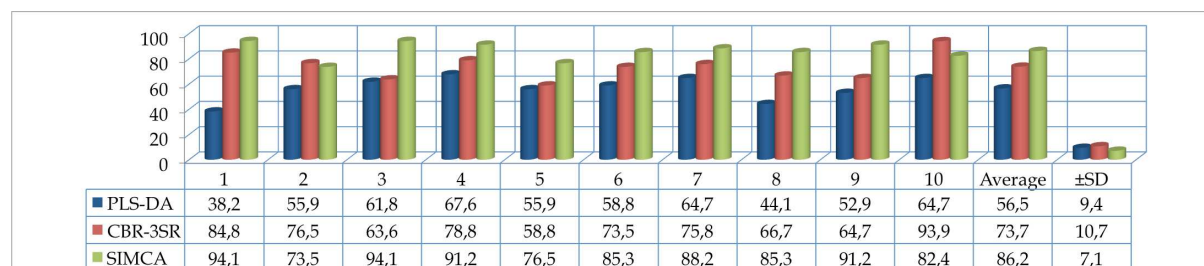
a un paciente médico (39 observaciones, de las cuáles 25 son TC (64,1%) y 14 son TI (35,9%)) y un paciente post-quirúrgico (30 observaciones, de las cuáles 18 son TC (60,0%) y 12 son TI (40,0%)). Los dos pertenecen a la cohorte de pacientes en shock séptico, en la cuál según resultados reportados en el Capítulo 4, la precisión del Guardian[®] RTCGMS fue mejor en comparación con las otras cohortes de pacientes. Estos dos pacientes fueron seleccionados porque en ellos el número de observaciones TC y TI no es tan desequilibrado, por tanto se puede realizar una



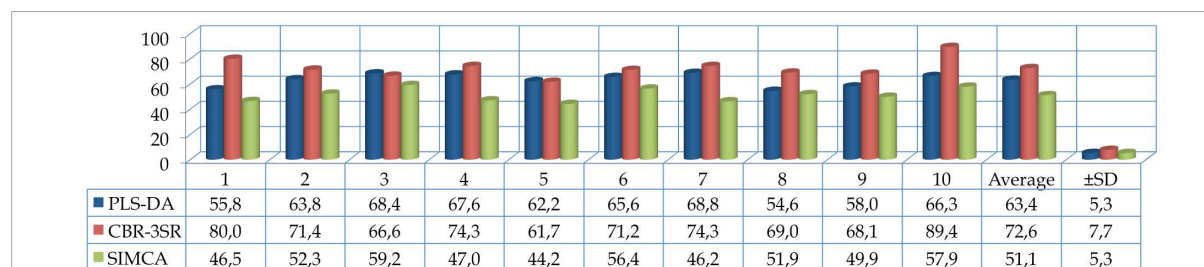
(a) Proporción de aciertos



(b) Especificidad



(c) Sensibilidad



(d) G_{media}

Figura 5.12: Resultados comparativos de PCA-CBR-3SR (obtenidos considerando como variables de entrada: Q -statistic, scores y Sepsis) y resultados de SIMCA y PLS-DA (obtenidos considerando como variables de entrada a los modelos el enfoque 2). Los resultados mostrados en las diferentes tablas son porcentajes (%).

evaluación más equitativa de la metodología desarrollada.

Cada una de las observaciones de los dos pacientes se considera como un nuevo dato a validar, por tanto, se aplica el modelo PCA construido en la sección 5.3.2 con el objetivo de obtener los descriptores, los cuáles caracterizarán los dos nuevos pacientes. Luego, cada uno de los casos, de los dos nuevos pacientes a validar, se compara con todos los casos que conforman cada una

de las 10 BCs finales de la Tabla 5.7, para recuperar los z vecinos más cercanos y proponer una clase para el nuevo caso, considerando tanto PCA-CBR-2SR como PCA-CBR-3SR.

Para todas las estimaciones de G_{RTCGMS} obtenidas cada cinco minutos para los dos pacientes se propone una clase TC o TI. Sin embargo, solamente las estimaciones de G_{RTCGMS} emparejadas con las determinaciones de ABG se emplean en la revisión (comparar la clase predicha con la clase real), y por ende, en el cálculo de las métricas que permiten evaluar el rendimiento de la metodología PCA-CBR, debido a que las determinaciones de ABG no están disponibles para todo el conjunto de datos. La Tabla 5.10 muestra los resultados obtenidos para cada uno de estos dos pacientes críticos, en términos de *proporción de aciertos*, *especificidad*, *sensibilidad* y G_{media} , considerando los dos enfoques (PCA-CBR-2SR y PCA-CBR-3SR).

No. BC	PCA-CBR-2SR				PCA-CBR-3SR			
	Proporción de aciertos (%)	Especificidad (%)	Sensibilidad (%)	G_{media} (%)	Proporción de aciertos (%)	Especificidad (%)	Sensibilidad (%)	G_{media} (%)
Paciente médico								
1	66,7	76,0	50,0	61,6	66,7	60,0	78,6	68,7
2	61,5	68,0	50,0	58,3	64,1	52,0	85,7	66,8
3	64,1	80,0	35,7	53,5	61,5	56,0	71,4	63,2
4	71,8	84,0	50,0	64,8	64,1	56,0	78,6	66,3
5	69,2	80,0	50,0	63,2	61,5	60,0	64,3	62,1
6	66,7	76,0	50,0	61,6	48,7	56,0	35,7	44,7
7	66,7	80,0	42,9	58,6	64,1	60,0	71,4	65,5
8	56,4	64,0	42,9	52,4	61,5	56,0	71,4	63,2
9	59,0	64,0	50,0	56,6	66,7	56,0	85,7	69,3
10	69,2	88,0	35,7	56,1	59,0	56,0	64,3	60,0
Media±DS	65,1±4,9	76,0±8,2	45,7±6,0	58,7±4,1	61,8±5,2	56,8±2,5	70,7±14,5	63,0±7,0
Mediana (IQR)	66,7 (6,4)	78,0 (10,0)	50,0 (7,1)	58,4 (5,5)	62,8 (2,6)	56,0 (3,0)	71,4 (12,5)	64,4 (4,3)
Paciente post-quirúrgico								
1	83,3	94,4	66,7	79,3	56,7	72,2	33,3	49,1
2	70,0	83,3	50,0	64,5	56,7	61,1	50,0	55,3
3	70,0	77,8	58,3	67,4	56,7	77,8	25,0	44,1
4	66,7	77,8	50,0	62,4	53,3	66,7	33,3	47,1
5	66,7	88,9	33,3	54,4	56,7	66,7	41,7	52,7
6	66,7	83,3	41,7	58,9	60,0	77,8	33,3	50,9
7	80,0	94,4	58,3	74,2	53,3	72,2	25,0	42,5
8	70,0	88,9	41,7	60,9	46,7	61,1	25,0	39,1
9	63,3	77,8	41,7	56,9	63,3	66,7	58,3	62,4
10	73,3	88,9	50,0	66,7	53,3	66,7	33,3	47,1
Media±DS	71,0±6,3	85,6±6,5	49,2±10,0	64,6±7,7	55,7±4,5	68,9±6,	35,8±11,1	49,0±6,7
Mediana (IQR)	70,0 (5,8)	86,1 (9,7)	50,0 (14,6)	63,5 (7,8)	56,7 (3,3)	66,7 (5,6)	33,3 (12,5)	48,1 (7,4)

Tabla 5.10: Resultados de evaluar la metodología PCA-CBR en dos pacientes representativos.

Analizando cada escenario de casos reportado en la Tabla 5.10 se observa que:

- En el caso del paciente médico, PCA-CBR-2SR reporta resultados muy superiores a PCA-CBR-3SR en *Especificidad* = 78,0 % (70,0 % - 80,0 %) versus 56,0 % (56,0 % - 59,0 %). Sin embargo, es el PCA-CBR-3SR el que reporta resultados mucho más altos en *Sensibilidad* = 71,4 % (66,1 % - 78,6 %) versus 50,0 % (42,9 % - 50,0 %). De acuerdo a G_{media} , el mejor resultado para PCA-CBR-2SR se obtiene con la BC No. 4 (64,8 %) y para PCA-CBR-3SR con la BC No. 9 (69,3 %). El peor resultado para PCA-CBR-2SR se obtiene con la BC No. 8 (52,4 %) y para PCA-CBR-3SR con la BC No. 6 (44,7 %). Por tanto, PCA-CBR-3SR con la BC No. 9 se elige como referencia para la evaluación.
- En relación con el paciente post-quirúrgico, PCA-CBR-2SR supera notablemente los resultados de PCA-CBR-3SR en todas las métricas: *Proporción de aciertos* = 70,0 % (66,7 % - 72,5 %) versus 56,7 % (53,3 % - 56,7 %), *Sensibilidad* = 50,0 % (41,7 % - 56,3 %) versus

33,3% (27,1% - 39,6%), *Especificidad* = 86,1% (79,2% - 88,9%) versus 66,7% (66,7% - 72,2%) y G_{media} = 63,5% (59,4% - 67,2%) versus 48,1% (44,9% - 52,3%). En este caso, de acuerdo a G_{media} , el mejor resultado para PCA-CBR-2SR se obtiene con la BC No. 1 (79,3%) y para PCA-CBR-3SR con la BC No. 9 (62,4%). Por otra parte, el peor resultado para PCA-CBR-2SR se obtiene con la BC No. 5 (54,4%) y para PCA-CBR-3SR con la BC No. 8 (39,1%). Por tanto, PCA-CBR-2SR con la BC No. 1 se elige como referencia para la evaluación.

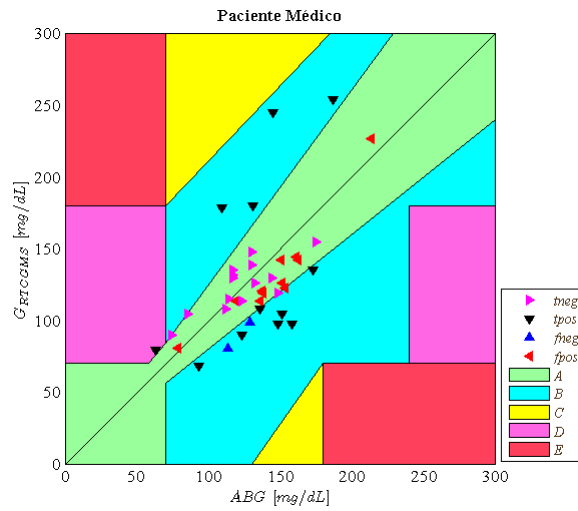
- Considerando los resultados obtenidos en estos dos pacientes, estos indican desde el punto de vista del uso generalizado de la metodología desarrollada, que no se logra un rendimiento más balanceado entre *sensibilidad* y *especificidad*. Sin embargo, es necesaria su validación con más pacientes para confirmar o invalidar esta conclusión.

A continuación se representan gráficamente los mejores resultados de acuerdo con G_{media} reportados en la Tabla 5.10, contrastando estos resultados derivados de aplicar la metodología PCA-CBR con lo indicado por Clarke EGA (Figura 5.13). Cada figura ilustra t_{neg} , t_{pos} , f_{neg} y f_{pos} . t_{neg} en este caso son medidas de G_{RTCGMS} que según el esquema de clasificación definido en la sección 5.2 son TC y la metodología PCA-CBR las clasifica correctamente, y t_{pos} son medidas de G_{RTCGMS} que son TI y la metodología PCA-CBR las clasifica correctamente. Por otra parte, f_{neg} son medidas de G_{RTCGMS} que se han etiquetado originalmente como TI, pero la metodología PCA-CBR las clasifica como TC. Finalmente, f_{pos} son medidas de G_{RTCGMS} que son TC y la metodología PCA-CBR las clasifica erróneamente como TI. En la Figura 5.13, el triángulo magenta apuntando hacia la derecha indica un t_{neg} , el triángulo negro apuntando hacia abajo indica un t_{pos} , el triángulo azul apuntando hacia arriba indica un f_{neg} , y finalmente, el triángulo rojo apuntando hacia la izquierda indica un f_{pos} .

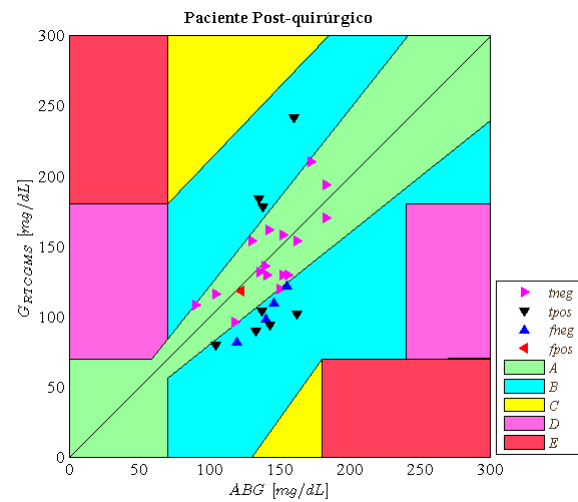
La clasificación de los pares de datos ABG/G_{RTCGMS} del paciente médico aplicando la metodología PCA-CBR-3SR (métricas reportadas en la Tabla 5.10 usando la BC No. 9) se ilustra en la Figura 5.13(a). Esta figura muestra que este paciente presenta una buena *sensibilidad*. De 14 medidas TI hechas por el Guardian[®] RTCGMS, dos se clasifican erróneamente (es decir, dos son f_{neg}); por tanto, se detecta el 85,7% de las medidas TI. Sin embargo, este paciente presenta una baja *especificidad* de 56,0%. 11 de 25 medidas TC (44,0%) se clasifican erróneamente y se reportan como medidas TI (es decir, 11 son f_{pos}).

De igual importancia, en la Figura 5.13(b) se muestra el caso para el paciente post-quirúrgico aplicando la metodología PCA-CBR-2SR (métricas reportadas en la Tabla 5.10 usando la BC No. 1). Se puede observar que este paciente presenta una baja *sensibilidad*. De 12 medidas TI hechas por el Guardian[®] RTCGMS, cuatro se clasifican erróneamente (es decir, cuatro son f_{neg}); por tanto, se detecta el 66,7% de las medidas TI. Sin embargo, la Figura 5.13(b) ilustra una alta *especificidad* de 94,4%. Una de 18 (5,6%) medidas correctas se clasifica erróneamente y se reporta como medida TI (es decir, 1 es f_{pos}).

En las Figuras 5.14 y 5.15 se ilustran los perfiles glucémicos para los dos pacientes. Los perfiles mostrados son ejemplos de los resultados obtenidos aplicando la metodología PCA-CBR. El cuadrado verde y el círculo rojo corresponden a medidas de G_{RTCGMS} realizadas por el Guardian[®] RTCGMS que son clasificadas como TC y TI, respectivamente. El diamante magenta corresponde a una medida de ABG que es utilizada como medida de referencia. En las medidas de G_{RTCGMS} el triángulo cian apuntando hacia la derecha indica un verdadero negativo (t_{neg}), el triángulo amarillo apuntando hacia abajo indica un verdadero positivo (t_{pos}), el



(a) Resultados en un paciente médico aplicando la metodología PCA-CBR-3SR usando la BC No. 9.



(b) Resultados en un paciente post-quirúrgico aplicando la metodología PCA-CBR-2SR usando la BC No. 1.

Figura 5.13: Clasificación de los pares de datos ABG/G_{RTCGMS} de dos pacientes representativos al comparar la metodología PCA-CBR con la clasificación de acuerdo a Clarke EGA.

triángulo negro apuntando hacia la izquierda indica un falso positivo (f_{pos}) y el triángulo azul apuntando hacia arriba indica un falso negativo (f_{neg}).

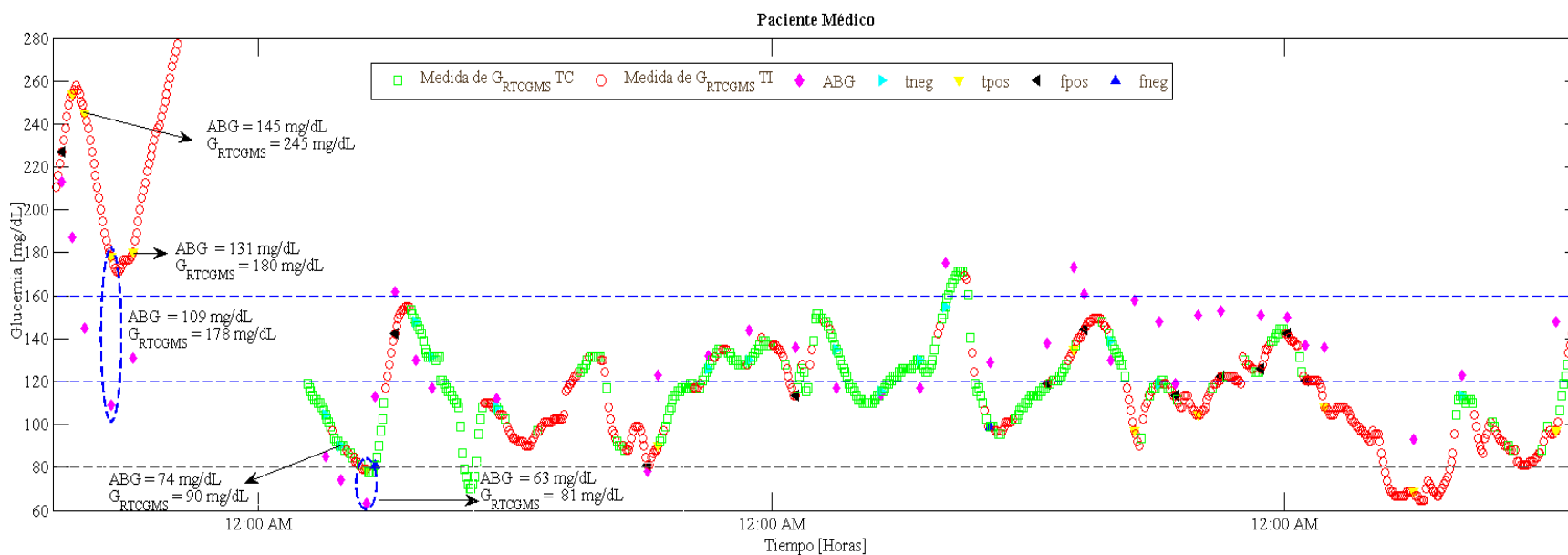


Figura 5.14: Perfil de un paciente médico obtenido al aplicar la metodología PCA-CBR-3SR usando la BC No. 9.

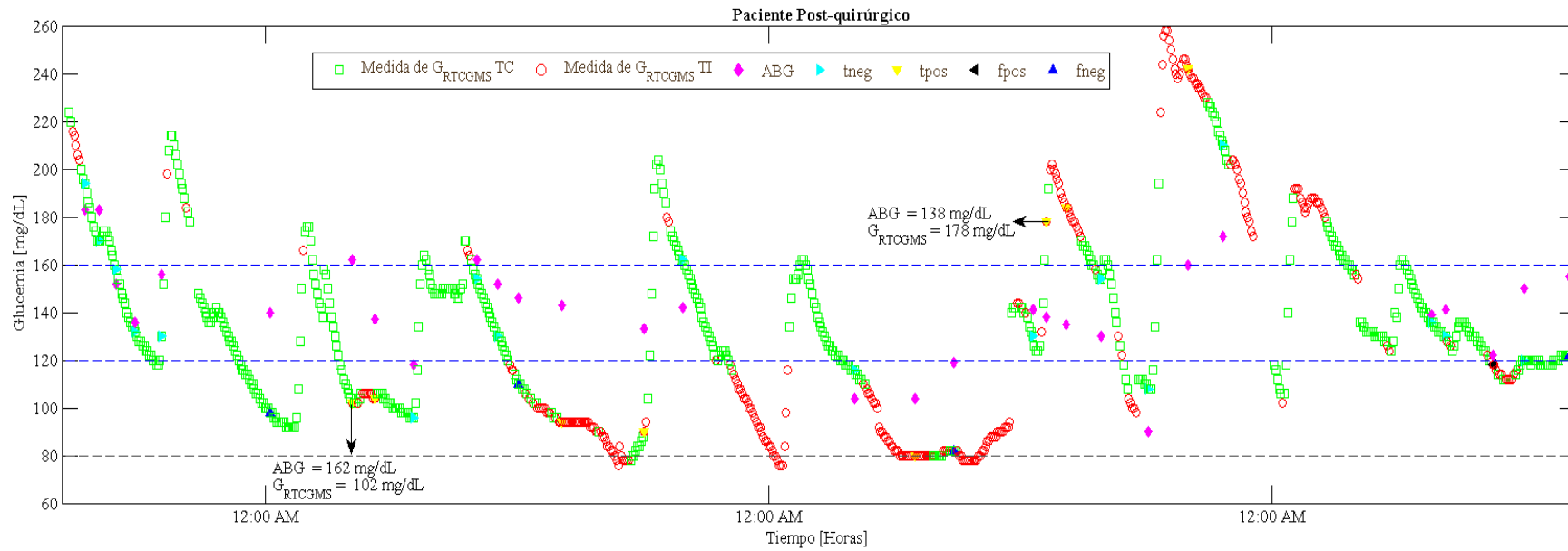


Figura 5.15: Perfil de un paciente post-quirúrgico obtenido al aplicar la metodología PCA-CBR-2SR usando la BC No. 1.

En el paciente médico (Figura 5.14), la metodología PCA-CBR-3SR detecta como medidas TI tres falsos eventos hiperglucémicos estimados por el Guardian[®] RTCGMS (G_{RTCGMS} superior a 160 mg/dL cuando ABG es inferior a 160 mg/dL). De acuerdo a la opinión del médico, es de gran importancia la correcta detección del segundo falso episodio hiperglucémico (resaltado con un círculo azul discontinuo). Para este caso, G_{RTCGMS} es 178 mg/dL mientras ABG es 109 mg/dL. Si el médico considera como fiable la medida indicada por el monitor, se inicia una dosis de CIII para mantener la concentración de ABG entre 120 y 160 mg/dL. Esta decisión podría luego causar un severo evento hipoglucémico en el paciente. Adicionalmente, la metodología PCA-CBR-3SR detecta correctamente un relevante evento hipoglucémico (ABG inferior a 80 mg/dL). En este evento (resaltado con un círculo azul discontinuo), mientras la medida de ABG indica 63 mg/dL, el Guardian[®] RTCGMS estima 81 mg/dL. Considerando la opinión del médico especialista, la correcta detección de este episodio hipoglucémico permite detectar una medida de glucosa erróneamente estimada por el Guardian[®] RTCGMS. Es importante mencionar que aunque se reportaron falsos positivos en este paciente, el desempeño de la metodología PCA-CBR no compromete la seguridad del paciente. El mayor perjuicio de los falsos positivos es que llevan al personal de la UCI a tomar determinaciones de ABG con más frecuencia para garantizar la seguridad del paciente, aumentando el número de punciones.

En el paciente post-quirúrgico (Figura 5.15) se puede observar que la metodología PCA-CBR permite detectar correctamente algunas observaciones TI estimadas por el Guardian[®] RTCGMS (por ejemplo, un episodio hiperglucémico y un episodio euglucémico). Sin embargo, su baja sensibilidad implica la presencia de un importante número de falsos negativos, los cuáles llevan al personal de la UCI a modificar la dosis de insulina. En este caso particular, los falsos negativos no tendrían un impacto clínico significativo en lo que respecta a la administración de la terapia insulínica al paciente, pero estos errores de clasificación podrían causar alteraciones en el tratamiento que aunque no comprometen la vida del paciente, son inadecuadas.

5.8. Análisis de los resultados obtenidos

El objetivo principal en este capítulo es determinar si la metodología PCA-CBR es capaz de detectar estimaciones de glucosa TC y TI realizadas por un SMCG. PCA calcula un nuevo espacio vectorial basándose en el cálculo de los vectores propios y de los valores propios de las variables originales ($ISIG$, G_{RTCGMS} , $Insulin$ y $Temperature$) destinadas a maximizar la varianza y minimizar la correlación entre ellas. Por lo tanto, estas nuevas variables “virtuales” (el Q -statistic y los $scores$ obtenidos a partir del modelo PCA) no pueden ser consideradas como variables clínicas o fisiológicas, ya que son combinaciones lineales de las variables originales.

Inicialmente, se aplicó PCA como una técnica de monitorización estadística multivariante para determinar si se podría usar para distinguir entre estimaciones TC y TI. Para este propósito se utilizó el D -statistic y el Q -statistic en un procedimiento secuencial que analizó si la nueva observación se comportaba de acuerdo al modelo (ni el D -statistic ni el Q -statistic superaban sus respectivos límites) o no (cualquiera de los dos índices rebasaba su límite estadístico). Sin embargo, este procedimiento no fue capaz de distinguir entre ambas tipologías.

Por la razón anteriormente expuesta, se decidió combinar PCA y CBR (como clasificador no lineal). En primer lugar, PCA se utiliza para extraer características que alimentan al sistema CBR. Más específicamente, el Q -statistic y los $scores$ junto con el estado séptico del paciente se

proponen como descriptores para la herramienta CBR. En segundo lugar, CBR se utiliza como herramienta de decisión que reconoce patrones y clasifica las estimaciones de BG realizadas por un SMCG. Las siguientes conclusiones se extraen a partir de los resultados obtenidos:

La principal ventaja de la metodología PCA-CBR es que permite el uso de la información proporcionada por el propio SMCG, e incorporar variables sobre el estado clínico del paciente para detectar cuando una glucemia proporcionada por este dispositivo es TC o TI. Esta metodología da mejores resultados para abordar el problema de detección de fallos en las estimaciones de BG de los SMCGs que el uso de PLS-DA o SIMCA.

Por otra parte, una característica de la metodología PCA-CBR es que permite la identificación, análisis y remoción de casos que de manera frecuente se clasifican erróneamente, para la detección de glucemias TC y TI. PCA por si misma no fue lo suficientemente robusta para detectar errores de entrada como los que a menudo ocurren en una UCI. Sin embargo, la metodología PCA-CBR si es robusta para identificar y analizar casos que pueden ser poco fiables. Más específicamente, en la etapa de puesta a punto de la metodología se pueden identificar frecuentemente casos mal clasificados, los cuáles se llevan a la etapa de retención para decidir la idoneidad de mantenerlos o de expulsarlos de la BC (y del conjunto de validación). La existencia de datos que continuamente se clasificaban mal, debidos a errores humanos en la entrada de las variables originales o en el uso efectivo de los dispositivos para medirlas, podría conducir a un diagnóstico erróneo. Por lo tanto, cuando se excluyen estos datos, la metodología PCA-CBR presenta una mejor clasificación global, en contraste con la que presenta cuando se utiliza las BCs originales sin limpieza.

Se observan mejores métricas utilizando PCA-CBR-2SR que usando PCA-CBR-3SR. Sin embargo, cuando se considera PCA-CBR-3SR mejora la *sensibilidad* aunque en detrimento de la *especificidad*. Los resultados indican que PCA-CBR-3SR no sólo tiende a identificar de manera correcta un alto porcentaje de estimaciones de glucosa TI, sino que también produce muchas falsas alarmas. Generalmente la presencia de falsos positivos no tiene consecuencias clínicas para los pacientes, debido a que no dan lugar a la modificación de la terapia intravenosa de insulina por parte del personal de la UCI. Sin embargo, en este estudio los falsos positivos conducen al personal de la UCI a tomar determinaciones de *ABG* con más frecuencia para garantizar la seguridad del paciente. Por otra parte, la presencia de falsos negativos, los cuáles llevan al personal de la UCI a modificar la dosis de insulina, si tendrían un impacto clínico significativo en lo que respecta a la administración de la terapia insulínica al paciente, debido a que estos errores de clasificación provocarían tratamientos inadecuados.

Es importante mencionar algunas limitaciones de la metodología PCA-CBR. Una limitación está en su capacidad para generalizar desde un conjunto de entrenamiento a otros conjuntos de validación. La generalización desde un conjunto de entrenamiento a otros conjuntos de validación depende entre otros factores de la naturaleza de los datos de validación. Esto es, si los conjuntos de entrenamiento y validación tienen las mismas características (en este estudio los datos que constituyen las BCs se recogieron en una UCI polivalente), la probabilidad de generalizar desde los datos de entrenamiento a un nuevo conjunto de datos de validación debe ser buena. Sin embargo, si el conjunto de validación se basa en una cohorte de pacientes diferente, entonces la capacidad de generalización puede ser pobre. En este caso, se recomienda utilizar esta cohorte de pacientes con nuevas características para construir nuevos conjuntos de entrenamiento. Sin embargo, es importante señalar que aunque existe una limitación en la capacidad de generalizar desde un conjunto de entrenamiento específico a un nuevo conjunto de validación, la generalidad

de la metodología PCA-CBR hace posible su aplicación a diferentes cohortes de pacientes, con el único requerimiento de definir nuevos conjuntos de entrenamiento para representar el comportamiento de la nueva cohorte.

Por otro lado, si el diagnóstico sobre el estado séptico del paciente se hace de manera subjetiva (es decir, basándose en la intuición del médico o en micro-cultivo), este hecho podría adicionar variabilidad a los resultados y confundir el sistema CBR, lo cual podría afectar su rendimiento. Por lo tanto, este tema constituye otra limitación de la metodología PCA-CBR. Sin embargo, es importante aclarar que esta limitación sólo afectaría el rendimiento de la metodología PCA-CBR-3SR, porque recupera los casos basándose en d_Q y d_t y luego propone la clase para un nuevo caso utilizando únicamente los casos recuperados con un estado séptico similar al del nuevo caso.

Finalmente, se destacan dos cuestiones importantes. El estudio se enfoca sólo en un SMCG específico, que se basa en un sensor subcutáneo (SC) que mide la concentración de glucosa en el ISF. Por lo tanto, son necesarios nuevos estudios para determinar si es posible extrapolar los resultados a otras tecnologías de sensores de glucosa, especialmente a los métodos intravenosos. Por otra parte, existen preocupaciones acerca de la exactitud del Hemocue (Watkinson et al., 2012), tal como se indicó en el Capítulo 4. Para evitar estos problemas, en estudios futuros es preciso utilizar como valores de referencia las glucemias obtenidas en pruebas de laboratorio.

5.9. Resumen

En este capítulo, se ha presentado una metodología basada en datos para la detección de estimaciones de BG TC o TI realizadas por el Guardian[®] RTCGMS, la cual proporciona un diagnóstico al personal de la UCI siendo de gran utilidad en la toma de decisiones clínicas. La metodología desarrollada utilizó PCA y CBR para determinar si una estimación de BG era TC o TI. PCA se utilizó para extraer características que fueron usadas como los descriptores de la herramienta CBR (*Q-statistic* y *scores*). CBR se aplicó para reconocer patrones y clasificar los datos. Además, se consideraron dos enfoques de recuperación de los z vecinos más cercanos: 2SR, basado en un criterio de distancia que utilizó sólo el *Q-statistic* y los *scores*, y 3SR, basado en un criterio de distancia similar al anterior junto con el criterio de similitud basado en el estado séptico de los pacientes.

Los resultados experimentales mostraron que la metodología PCA-CBR, puede ser una potencial herramienta para desarrollar un sistema de detección de fallos y averías que permita auto-monitorizar los SMCGs, utilizando la información proporcionada por el propio dispositivo, e incorporando algunas variables sobre el estado clínico del paciente. La metodología aquí descrita mostró un desempeño más adecuado que el logrado por SIMCA y PLS-DA, dos clasificadores lineales de uso común. Sin embargo, una de sus principales limitaciones está en su capacidad para generalizar desde un conjunto de entrenamiento a otros conjuntos de validación.

Capítulo 6

Detección de medidas correctas e incorrectas en los SMCGs aplicando estrategias de post-procesamiento de Máquinas de Soporte Vectorial (SVMs)

El presente capítulo aborda un enfoque diferente al explicado en el Capítulo 5 para detectar medidas de glucosa correctas e incorrectas realizadas por un SMCG, puesto que aquí se aplican estrategias específicas para mejorar el rendimiento de un clasificador sobre conjuntos de datos desbalanceados. Recordemos que en el capítulo anterior, en la evaluación del rendimiento de uso generalizado de la metodología propuesta en dos pacientes representativos, la metodología que se propuso favorecía la detección de la clase TC y exhibía un menor rendimiento en la detección de la clase TI. Para evitar este desbalance, en este capítulo se utilizan Máquinas de Soporte Vectorial (-del inglés *Support Vector Machines (SVMs)*-) no convencionales para detectar medidas de glucosa correctas e incorrectas realizadas por un SMCG, empleando estrategias específicas para mejorar su rendimiento en conjuntos de datos desbalanceados. Las SVMs son una atractiva opción para abordar este problema debido a que su mecanismo de aprendizaje considera un pequeño subconjunto de patrones para separar situaciones con fallos de situaciones sin fallos (Cristianini and Shawe-Taylor, 2000).

En un trabajo previo, las SVMs convencionales se usaron para detectar medidas de glucosa TC y TI realizadas por el CGMS Gold[®] de acuerdo a criterios clínicos (Bondia et al., 2008). Sin embargo, al abordar el problema de la detección de fallos en un SMCG mediante el uso de SVMs, se debe tener en cuenta que existen mayoritariamente glucemias correctas, pero escasas glucemias incorrectas, es decir, el número de estimaciones de glucosa incorrectas es limitado. En el contexto de este problema, es muy importante obtener modelos que exhiban un alto rendimiento en la detección de esta clase minoritaria, debido a que esta usualmente representa el objetivo de la práctica clínica (en una UCI las medidas incorrectas realizadas por el SMCG pueden corresponder a situaciones de alta inestabilidad glucémica en el paciente como episodios hiperglucémicos e hipoglucémicos), sin que ello conlleve a perjudicar la detección de las glucemias correctas. En resumen, el objetivo deben ser modelos con una buena *sensibilidad*, pero al mismo tiempo, con una buena *especificidad*. Sin embargo, la SVM convencional no satisface este objetivo, puesto que favorece la clase mayoritaria cuando clasifica conjuntos de datos desbalan-

ceados, una premisa no considerada en (Bondia et al., 2008).

En este capítulo se introducen y aplican dos estrategias de post-procesamiento de SVMs para clasificar conjuntos de datos desbalanceados, al problema de la detección de fallos en un SMCG. El capítulo se organiza como sigue: la sección 6.1 explica la transformación del problema de detección de fallos en un SMCG en un problema de bi-clasificación. La sección 6.2 introduce los conceptos de SVM y a continuación, en la misma sección, se describen las estrategias de post-procesamiento para ajustar el umbral de la función decisión de la SVM en función de la G_{media} y de la *sensibilidad*. La metodología propuesta para detectar medidas de glucosa correctas e incorrectas en un SMCG aplicando nuevas estrategias de post-procesamiento de SVMs se expone en la sección 6.3. Los resultados experimentales de la implementación de la metodología propuesta se reportan en la sección 6.4. La sección 6.5 presenta el análisis de los resultados obtenidos. Finalmente, se presenta un resumen del capítulo en la sección 6.6.

La metodología propuesta, los resultados y las conclusiones de este trabajo se presentaron en forma parcial en el Doceavo Congreso Internacional sobre Tecnologías para la Diabetes (Leal et al., 2012a). Un artículo basado en este trabajo ha sido aceptado para publicación (Leal et al., 2013a).

6.1. Descripción, etiquetado y pre-procesamiento de los datos

Teniendo como base el trabajo previo realizado por (Bondia et al., 2008), se transforma el problema de la detección de fallos en los SMCGs en un problema de clasificación binario. En esta investigación, se utilizó el mismo conjunto de datos obtenidos desde 23 pacientes críticamente enfermos admitidos en la unidad de cuidados intensivos (UCI) del Hospital Josep Trueta. Recordemos que en la sección 5.1 se describió este conjunto de datos. Sin embargo, en este trabajo sólo se consideró la información proporcionada por el Guardian[®] RTCGMS (*ISIG* y *G_{RTCGMS}*) y cómo variable propia de la condición clínica del paciente, la dosis de *Insulin*. Las razones por las cuáles no se consideran las variables *Sepsis* y *Temperature* se mencionan en la sección 6.6. Por lo tanto, el conjunto de datos contiene 537 observaciones y cada una de ellas está compuesta por tres variables: *ISIG*, *G_{RTCGMS}* e *Insulin*. Además, se han obtenido 537 determinaciones de *ABG* que se utilizan como medidas de referencia para etiquetar los datos, basándose en un esquema de clasificación definido a partir de un criterio de precisión analítico.

En primer lugar, se etiqueta el conjunto de datos en medidas correctas e incorrectas usando los criterios ISO. Los criterios ISO se describieron en la sección 2.4.1. De acuerdo con estos criterios:

- Para que una medida de un SMCG pueda ser considerada correcta:
 - a) Si *ABG* es ≤ 75 mg/dL, *G_{RTCGMS}* sólo puede variar entre ± 15 mg/dL de *ABG*.
 - b) Si *ABG* es > 75 mg/dL, *G_{RTCGMS}* sólo puede variar entre $\pm 20,0\%$ de *ABG*.
- En caso tal que la medida del SMCG no cumpla estos criterios, se considera incorrecta.

Por tanto, se define el siguiente esquema de clasificación, basado en la distribución de los datos de acuerdo con los criterios ISO, para etiquetar el conjunto de datos (Figura 6.1):

- Clase correcta (Clase -1): Las medidas que satisfacen los criterios ISO se consideran correctas. Del conjunto entero de datos, 69,3% (372 observaciones) fueron medidas correctas.
- Clase incorrecta (Clase 1): Las medidas que no satisfacen los criterios ISO se consideran incorrectas. Del conjunto entero de datos, 30,7% (165 observaciones) fueron medidas incorrectas.

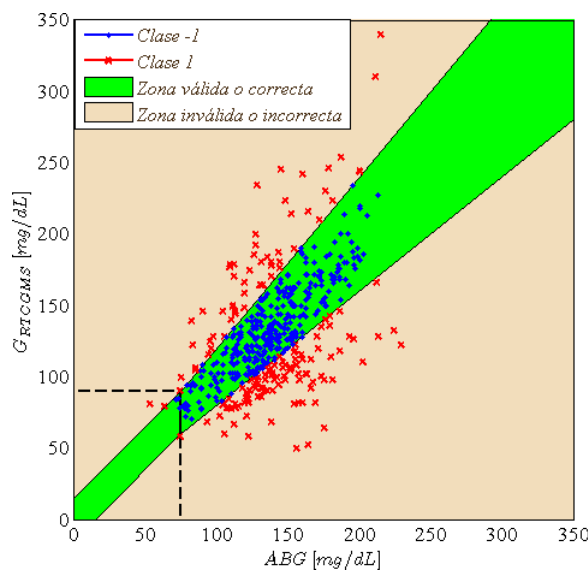


Figura 6.1: Distribución de los datos de acuerdo con los criterios ISO

En consecuencia, el conjunto de datos contiene 537 observaciones etiquetadas en dos clases desbalanceadas (372/165) y cada observación está compuesta por tres variables: *ISIG*, *G_RTCGMS* e *Insulin*.

6.2. Máquinas de Soporte Vectorial (SVMs)

6.2.1. Conceptos básicos de SVMs

Las máquinas de soporte vectorial son modelos de aprendizaje que implementan el principio inductivo de minimización del riesgo estructural para poder generalizar a partir de un número limitado de patrones de aprendizaje (Vapnik, 1998). Están fundamentadas en la teoría del aprendizaje estadístico y han sido aplicadas con éxito en problemas de clasificación y regresión en diferentes dominios (Vapnik, 1999; Cristianini and Shawe-Taylor, 2000). El espacio de hipótesis de las SVMs son hiperplanos (superficies de decisión lineales) y su principal objetivo es buscar durante el entrenamiento el hiperplano de separación que permite obtener el margen máximo de separación entre las clases (Núñez et al., 2012).

Para resolver problemas de clasificación de conjuntos de datos no separables linealmente, se proyecta el espacio de entradas \mathcal{X} en un espacio con producto interno (escalar) de dimensión superior, al que se denomina espacio de características \mathcal{F} , utilizando para ello una función no lineal ϕ . En este espacio de características \mathcal{F} se puede realizar una separación lineal de los datos

transformados, determinando los parámetros del hiperplano de separación óptimo. Este hiperplano corresponderá en el espacio original \mathcal{X} a una función de decisión no lineal, cuya forma estará determinada por la función no lineal ϕ .

Sea $\mathcal{Z} = \{(x_1, y_1), \dots, (x_N, y_N)\}$ un conjunto de datos de entrenamiento, con $x_i \in \mathcal{X}$ como el espacio de entradas, $y_i \in \mathcal{Y} = \{+1, -1\}$ como el espacio de salidas y $z_i = (x_i, y_i)$. Sea $\phi : \mathcal{X} \rightarrow \mathcal{F}$, $\mathbf{x} = \phi(x)$, un mapeo de características y \mathcal{F} es un espacio dotado con un producto escalar denotado por $\langle \cdot, \cdot \rangle$. Se busca un clasificador lineal binario $f : \mathcal{X} \rightarrow \mathbb{R}$, definido como $f(x) = \langle \mathbf{x}, \mathbf{w} \rangle - b$, donde $\mathbf{w} \in \mathcal{F}$ es el vector perpendicular al hiperplano, el cuál define su orientación, y el parámetro $b \in \mathbb{R}$ determina su posición. Las salidas se obtienen mediante la función decisión $h(x) = \text{signo}(f(x))$. El hiperplano óptimo se determina mediante la formulación del problema primal estándar de las SVMs, el cuál conduce al siguiente problema de optimización (Gonzalez-Abril et al., 2006):

$$\begin{aligned} \min_{\mathbf{w} \in \mathcal{F}, b \in \mathbb{R}} \quad & \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^N \xi_i \\ \text{s.a.} \quad & \begin{cases} y_i (\langle \mathbf{x}_i, \mathbf{w} \rangle - b) + \xi_i \geq 1, & \forall i \\ \xi_i \geq 0, & \forall i \end{cases} \end{aligned} \quad (6.1)$$

donde el parámetro $C > 0$ es un parámetro de regularización y las variables de holgura o de relajación ξ_i miden el error sobre los casos que violan la restricción $y_i (\langle \mathbf{x}_i, \mathbf{w} \rangle - b) \geq 1$, es decir, miden el número de errores que se cometen en el conjunto de entrenamiento. El parámetro C en la ecuación (6.1), es un parámetro definido por el usuario, el cuál determina el balance entre maximizar el margen, definido como $\frac{1}{\|\mathbf{w}\|}$ y minimizar el error. El parámetro C penaliza los errores de tal manera que cuanto más grande es, se está asignando un peso a los errores más alto y por lo tanto, la SVM se enfoca más en minimizar los errores; mientras más pequeño es el valor de C , el objetivo principal de la SVM será maximizar el margen.

En su forma dual, el problema de optimización expresado en la ecuación (6.1) se puede expresar como:

$$\begin{aligned} \max \quad & \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i,k=1}^N \alpha_i y_i \alpha_k y_k \langle \mathbf{x}_i, \mathbf{x}_k \rangle \\ \text{s.a.} \quad & \begin{cases} 0 \leq \alpha_i \leq C & \forall i, \sum_{i=1}^N \alpha_i y_i = 0 \end{cases} \end{aligned} \quad (6.2)$$

que conduce a la solución $\mathbf{w}_0 = \sum_i \alpha_i y_i \mathbf{x}_i$, donde α_i son los multiplicadores de Lagrange para la formulación dual, sujeto a $0 \leq \alpha_i \leq C, \forall i$ y $\sum_i \alpha_i y_i = 0$. El parámetro b se calcula a posteriori (Gonzalez-Abril et al., 2008) y se denota como b_0 (sesgo estándar).

Entonces, desde el punto de vista del problema de optimización, dado un conjunto de vectores de entrenamiento \mathcal{Z} , no necesariamente separable, se realiza una transformación ϕ de los vectores de entrada $\mathcal{X} = \{x_1, \dots, x_N\}$, tal que convierta el conjunto transformado, $\mathcal{Z}_\phi = \{(\phi(x_1), y_1), \dots, (\phi(x_N), y_N)\}$, en un espacio adecuado. Esta transformación se realiza dentro de un espacio de características \mathcal{F} de dimensión muy superior, siguiendo una transformación no lineal elegida a priori. Dentro de este nuevo espacio, se busca el hiperplano separador óptimo (función de decisión) como un desarrollo lineal de funciones. Por lo tanto, planteando el problema de optimización de la SVM a los vectores transformados \mathcal{Z}_ϕ , se obtiene un problema de optimización que proporciona, según la clase de funciones dada, una solución lineal en el

espacio de características, pero no necesariamente lineal en el espacio de entradas.

Sin embargo, en la elección de la función solución, se debe considerar el problema de la generalización. Esta consideración implica que la función solución no puede ser cualesquiera. No obstante, esta elección no plantea ningún inconveniente ya que la solución al problema de optimización se expresa en términos de las denominadas funciones núcleo (-del inglés *Kernel*-), denotadas por K , sin tener en ningún momento que conocer la transformación $\phi(x)$.

Una función Kernel $K : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$, es una función real de dos variables, cuya principal característica es que debe expresarse a través de un producto escalar de una transformación de los vectores de entrada de \mathcal{X} en un espacio de características de dimensión superior, \mathcal{F} . No obstante, no todas las funciones de tipo $K : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ corresponden a funciones núcleo. Para que una función $K : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ sea considerada una función Kernel, esta debe satisfacer el teorema de Mercer. El teorema de Mercer indica si una función se corresponde a un producto escalar en \mathcal{F} . El teorema de Mercer (ecuación (6.3)) dado en (Courant and Hilbert, 1953) indica lo siguiente: existirá una aplicación ϕ y un desarrollo $K(x, y) = \sum_{i=1}^{\dim \mathcal{F}} \langle \phi(x)_i, \phi(y)_i \rangle$, si y sólo si, para cualquier función real $g(x)$ tal que $\int g^2(x) dx < \infty$, se cumple:

$$\int K(x, y) g(x) g(y) dx dy \geq 0 \quad (6.3)$$

Por lo tanto, una función Kernel se define como $K(x, y) = \langle x, y \rangle = \langle \phi(x), \phi(y) \rangle$. Un estudio detallado de las funciones núcleos y de la condición de Mercer que debe satisfacer una función de dos variables para ser una función núcleo se presenta en (Gonzalez-Abril, 2002).

Finalmente, el clasificador en términos de la función Kernel K puede ser escrito como:

$$f(x) = \sum_{i=1}^N \alpha_i y_i K(x_i, x) - b_0 = \langle x, w_0 \rangle - b_0 \quad (6.4)$$

donde $w_0 = \sum_i \alpha_i y_i \phi(x_i)$ y $x = \phi(x)$.

La función decisión en términos de la función Kernel K se puede formular como:

$$h(x) = \text{signo} \left(\sum_{i=1}^N \alpha_i y_i K(x_i, x) - b_0 \right) \quad (6.5)$$

Entre todos los vectores del conjunto de entrenamiento, sólo aquellos ubicados en el margen de decisión, tienen asociado un peso α_i mayor que cero en la ecuación (6.4). A estos vectores se les denomina vectores soporte (-del inglés *Support Vector*-). La ecuación (6.4) es una medida de la distancia de un caso x al hiperplano, es decir, el “score” de una entrada x . La ecuación (6.5) determina la etiqueta, que puede ser clase positiva o negativa.

La teoría explicada anteriormente fue desarrollada sobre la base de un problema de clasificación binario separable, donde el criterio de optimización es el ancho del margen entre los casos etiquetados como clase positiva y los casos etiquetados como clase negativa (Vapnik, 1999). Esta teoría proporciona un buen rendimiento de uso generalizado, como se muestra en numerosas aplicaciones (Cristianini and Shawe-Taylor, 2000), aún en aquellas que contienen conjuntos de datos moderadamente desbalanceados. En estos escenarios moderadamente desbalanceados, a diferencia de otras máquinas de aprendizaje, los resultados empíricos de diferentes aplicaciones

han mostrado que no es necesario realizar ninguna modificación durante el proceso de aprendizaje de las SVMs para mejorar su rendimiento (Akbari et al., 2004; Imam et al., 2006). Este fenómeno se puede explicar porque las SVMs sólo utilizan los vectores soportes para construir los modelos de clasificación, por lo tanto, en la construcción de estos modelos no se consideran los casos alejados del límite de decisión.

Sin embargo, al igual que otras máquinas de aprendizaje que construyen modelos de clasificación, las SVMs tienden a minimizar el error global en el conjunto de datos. Se ha observado que en escenarios donde clasifican conjuntos de datos significativamente desbalanceados entre las clases, el hiperplano de separación aprendido por la SVM está muy cercano a la clase minoritaria (He and Ghodsi, 2010; Wu and Chang, 2005), lo que resulta en un rendimiento de uso generalizado que favorece la clase mayoritaria. Este fenómeno, se puede apreciar más claramente si se expresa la *proporción de aciertos* (ecuación (2.18)) como la media aritmética ponderada de la *sensibilidad* (ecuación (2.21)) y la *especificidad* (ecuación (2.20)):

$$\text{Proporción de aciertos } (f, \mathcal{D}) = \frac{N_{pos} \text{ Sensibilidad } (f, \mathcal{D}) + N_{neg} \text{ Especificidad } (f, \mathcal{D})}{N_{pos} + N_{neg}} \quad (6.6)$$

En la ecuación (6.6), los pesos asociados con la *sensibilidad* y la *especificidad* corresponden al número de casos positivos y negativos, respectivamente. En escenarios donde las clases son significativamente desbalanceadas, $N_{pos} \ll N_{neg}$, y de la ecuación (6.6) se puede determinar que la *proporción de aciertos* y la *especificidad* son muy similares, es decir, que la *sensibilidad* escasamente altera el valor de la métrica *proporción de aciertos*. Para abordar estos escenarios desbalanceados, se han propuesto diversas técnicas de pre-procesamiento, de entrenamiento y de post-procesamiento, que buscan mejorar el rendimiento de las SVMs. Un resumen de estas estrategias se presenta a continuación.

Estrategias propuestas para mejorar el rendimiento de las SVMs sobre conjuntos de datos no balanceados

Se han propuesto diversas estrategias para mejorar el rendimiento de la SVM sobre conjuntos de datos no balanceados (Núñez et al., 2012). De acuerdo al momento en que pueden ser aplicadas durante el proceso de aprendizaje, estas estrategias se pueden clasificar en:

- Estrategias de pre-procesamiento. Se basan principalmente en utilizar técnicas de muestreo para re-balancear el conjunto de datos, con el fin de construir un nuevo conjunto de aprendizaje donde la clase minoritaria esté mejor representada. Dentro de este tipo de estrategias se han propuesto:
 - Estrategias de sobre-muestreo (-del inglés *over-sampling*-), como el algoritmo SMO-TE, el cuál se basa en la técnica de k-vecinos más cercanos para aumentar la proporción de los datos de la clase minoritaria en el conjunto de datos (Chawla et al., 2002; Vilariño et al., 2005). Estas estrategias requieren la entonación de nuevos parámetros como la tasa de muestreo.
 - Estrategias de sub-muestreo (-del inglés *under-sampling*-), como la aplicación de algoritmos de agrupación para disminuir la proporción de los datos de la clase mayoritaria en el conjunto de datos (Li et al., 2008b; Yu et al., 2006). Estas técnicas requieren la entonación de nuevos parámetros como la tasa de muestreo.

- Estrategias que buscan aumentar la representatividad de la clase minoritaria sólo en los límites de decisión entre las dos clases, obteniendo la información de los límites de decisión mediante reglas heurísticas basadas en la técnica de k-vecinos (Castro et al., 2009). Estas estrategias requieren la entonación de nuevos parámetros como el número de k-vecinos.
 - Estrategias que utilizan métodos de muestreo en combinación con *ensembles*. Los *ensembles* se basan en la construcción de un conjunto de clasificadores, por ejemplo varias SVMs, cuyas decisiones son combinadas para producir un resultado final (Kang and Cho, 2006; Yang et al., 2006; Waske et al., 2009; Yang et al., 2011). Estas técnicas pueden ser computacionalmente costosas.
- Estrategias de entrenamiento. Se incluyen aquí todas aquellas estrategias que modifican el problema de optimización estándar para el entrenamiento de las SVMs, con el fin de incorporar la información relacionada con las proporciones de las clases en el conjunto de datos. Dentro de este tipo de propuestas se encuentran:
- Estrategias basadas en el aprendizaje sensitivo al costo. Consisten en incorporar en el problema de aprendizaje, la información relacionada al costo asociado a las predicciones incorrectas, pero asignando un mayor peso a los errores sobre la clase minoritaria (Cohen et al., 2006; Veropoulos et al., 1999). Esto se puede lograr modificando el parámetro C o las variables de relajación ξ_i (He and Ghodsi, 2010; Yang et al., 2008; Rukshan and Vasile, 2010). Estas técnicas pueden producir modelos sobreajustados.
 - Estrategias combinando el aprendizaje sensitivo al costo con otras técnicas, como el algoritmo SMOTE (Akbari et al., 2004), los *ensembles* (Wang and Japkowicz, 2008) o el algoritmo KBA, este último un algoritmo iterativo que propone modificar la matriz Kernel de acuerdo al desbalance observado en la distribución de los datos (Wu and Chang, 2005). Estas técnicas pueden ser computacionalmente costosas.
- Estrategias de post-procesamiento. Estos trabajos de investigación se enfocan en la modificación del vector w de la función decisión o en la determinación de un nuevo sesgo o umbral, con el objetivo de ajustar el límite de decisión aprendido por la SVM, de tal forma que suministre un buen margen de separación para la clase minoritaria. Algunas de las técnicas propuestas han sido:
- Estrategias para modificar el vector w de la función decisión, como el método z-SVM, el cuál determina el valor de un nuevo parámetro al que han denominado z , que pondera la contribución de los vectores soportes de la clase minoritaria en w (Imam et al., 2006).
 - Estrategias para modificar el umbral o sesgo de la función decisión, las cuáles sugieren determinar un nuevo umbral para la función decisión de la SVM, considerando la proporción de las clases en el conjunto de datos, para mejorar su desempeño sobre la clase minoritaria, como en (Haydemar et al., 2011). Este tipo de estrategias no afectan directamente el entrenamiento de las SVMs.
 - Estrategias para sustituir el esquema de clasificación estándar, el cuál se basa en el signo de la función decisión, por otro que permita otorgar una interpretación diferente a las salidas de las SVMs en escenarios no balanceados, como por ejemplo en (Li et al., 2008a), donde se utiliza una función de decisión difusa cuyos parámetros son estimados a partir de la distribución observada en el conjunto de datos.

En el contexto de este trabajo, en el cuál se utiliza un conjunto de datos significativamente desbalanceado, se utilizan las estrategias de post-procesamiento para modificar el umbral o sesgo de la función decisión que se han propuesto en (Haydemar et al., 2011), con el objetivo de mejorar el rendimiento de las SVMs sobre conjuntos de datos no balanceados. En particular, se han considerado dos estrategias de post-procesamiento para ajustar el sesgo, la primera, la cuál denominaremos SVM-1, en función de la G_{media} (Haydemar et al., 2011), diseñada para conseguir un mejor rendimiento en ambas clases; y la segunda, la cuál denominaremos SVM-2, en función de la *sensibilidad*, diseñada para aplicarse a problemas de clasificación binarios donde una clase se considera de mayor importancia que la otra. Algunos criterios que se consideraron para seleccionar estas técnicas fueron que no suponen la entonación de nuevos parámetros, que no requieren modificar el problema de optimización estándar para el entrenamiento de las SVMs y que tampoco requieren pasos adicionales de re-entrenamiento.

6.2.2. SVM-1: Aplicación de una estrategia de post-procesamiento para ajustar el sesgo en función de la G_{media}

Cuando se aplica una SVM convencional para resolver un problema de clasificación binaria, la *proporción de aciertos* es la métrica que usualmente se maximiza. Sin embargo, existen problemas donde se obtiene esta solución con una exactitud desbalanceada, es decir, donde una clase alcanza una exactitud mucho más alta que la otra clase. Esta situación es común cuando el aprendizaje se realiza a partir de un conjunto de datos que contiene observaciones etiquetadas en clases significativamente desbalanceadas, tal como el discutido en el contexto de este trabajo. Por otra parte, recientes investigaciones han demostrado empíricamente que el hiperplano aprendido por una SVM estándar en la presencia de un conjunto de datos desbalanceado tiene aproximadamente la misma orientación que el hiperplano ideal (He and Ghodsi, 2010; Yang et al., 2006; Wu and Chang, 2005). La reducida capacidad de generalización en la clase minoritaria está realmente asociada al sesgo b , debido a que los casos positivos se encuentran mucho más lejos de este límite ideal, es decir, la SVM convencional aprende un límite que está muy cercano a la clase minoritaria.

Por lo tanto, dado que la *proporción de aciertos* es inapropiada cuando difieren bastante las probabilidades a priori de las clases, se pueden considerar otras estrategias para mejorar el rendimiento de las SVMs, en las cuáles el objetivo sea maximizar otras métricas de evaluación basadas en la matriz de confusión. En ese sentido, siguiendo las sugerencias presentadas en algunos trabajos sobre SVM (Sun et al., 2009), que proponen investigar en estrategias dirigidas a determinar un nuevo umbral para la función decisión de la SVM, las cuáles se sustentan en la distribución de las clases en el conjunto de datos, se decidió aplicar una estrategia de post-procesamiento para ajustar el sesgo en función de la G_{media} (ecuación (2.23)). Esta estrategia de post-procesamiento, la cuál denominaremos SVM-1, considera que el sesgo b de la ecuación (6.1) es un parámetro que se puede sintonizar (Haydemar et al., 2011), es decir, se basa en el cálculo de un nuevo sesgo b considerando la proporción de las clases en el conjunto de datos. La función aprendida por la SVM estándar se ajusta mediante el sesgo b para mejorar su desempeño sobre la clase minoritaria, medido en términos de la G_{media} . La solución que se aplica no supone la entonación de nuevos parámetros. Tampoco requiere modificar el problema de optimización estándar para entrenar la SVM, ni pasos adicionales, ni re-entrenamiento. Esta estrategia se basa en los desarrollos presentados en (Gonzalez-Abril et al., 2008).

Por consiguiente, una vez que se obtiene el vector solución $w_0 = \sum_i \alpha_i y_i x_i$ del problema

descrito en la ecuación (6.2), se considera el conjunto de clasificadores:

$$\mathcal{F}(w_0) = \{f_b : \mathcal{X} \rightarrow \mathbb{R}, f_b(\cdot) = \langle \phi(\cdot), w_0 \rangle - b, b \in \mathbb{R}\} \quad (6.7)$$

Se define un mapa $\Theta_b : \mathcal{X} \rightarrow \{-1, +1\}$, usando el clasificador $f_b(x) \in \mathcal{F}(w_0)$, donde $f_b(x) = \langle x, w_0 \rangle - b$, de tal manera que, dado un vector de entradas x , se asigne una etiqueta de la siguiente forma:

$$\Theta_b(x) = \text{signo}(f_b(x)) = \begin{cases} +1 & \text{si } \langle x, w_0 \rangle \geq b \\ -1 & \text{si } \langle x, w_0 \rangle < b \end{cases} \quad (6.8)$$

Se define *Sensibilidad* (b) = *Sensibilidad* (f_b, \mathcal{Z}) desde la ecuación (2.21). Si se considera $\mathcal{Z}_{pos} = \{(x_1, +1), \dots, (x_{N_{pos}}, +1)\} = \{(p_1, +1), \dots, (p_{N_{pos}}, +1)\}$ como el conjunto de patrones de entrenamiento con etiquetas positivas, con $p_i = x_{\sigma^*(i)}$, donde σ^* es una permutación de N_{pos} elementos de tal manera que:

$$\langle p_1, w_0 \rangle \leq \dots \leq \langle p_i, w_0 \rangle \leq \dots \leq \langle p_{N_{pos}}, w_0 \rangle \quad (6.9)$$

Se consideran los valores β y β^* , siendo β y β^* los valores mínimo y máximo del hiperplano sin sesgo aplicado al conjunto \mathcal{Z}_{pos} :

$$\beta = \min_{z_i \in \mathcal{Z}_{pos}} \langle x_i, w_0 \rangle = \langle p_1, w_0 \rangle \quad (6.10)$$

$$\beta^* = \max_{z_i \in \mathcal{Z}_{pos}} \langle x_i, w_0 \rangle = \langle p_{N_{pos}}, w_0 \rangle \quad (6.11)$$

Por otra parte, se define *Especificidad* (b) = *Especificidad* (f_b, \mathcal{Z}) desde la ecuación (2.20). Si se considera $\mathcal{Z}_{neg} = \{(x_1, -1), \dots, (x_{N_{neg}}, -1)\} = \{(q_1, -1), \dots, (q_{N_{neg}}, -1)\}$ como el conjunto de patrones de entrenamiento con etiquetas negativas, con $q_j = x_{\sigma'(j)}$, donde σ' es una permutación de N_{neg} elementos de tal manera que:

$$\langle q_1, w_0 \rangle \leq \dots \leq \langle q_j, w_0 \rangle \leq \dots \leq \langle q_{N_{neg}}, w_0 \rangle \quad (6.12)$$

Se consideran los valores α^* y α , siendo α^* y α los valores mínimo y máximo del hiperplano sin sesgo aplicado al conjunto \mathcal{Z}_{neg} :

$$\alpha^* = \min_{z_i \in \mathcal{Z}_{neg}} \langle x_i, w_0 \rangle = \langle q_1, w_0 \rangle \quad (6.13)$$

$$\alpha = \max_{z_i \in \mathcal{Z}_{neg}} \langle x_i, w_0 \rangle = \langle q_{N_{neg}}, w_0 \rangle \quad (6.14)$$

Un ejemplo de las propiedades de la *Sensibilidad* (b) y la *Especificidad* (b) se muestra en la Figura 6.2. Los valores α^* , α , β y β^* son elementos claves en el desarrollo de las SVMs, debido a que los valores óptimos en las métricas habitualmente usadas para evaluar su rendimiento, las cuáles se basan en la matriz de confusión, se encuentran en los intervalos determinados a partir de estos elementos. En esta gráfica se puede apreciar que $\alpha^* \leq \alpha$ y $\beta \leq \beta^*$.

Finalmente, se define $G_{media}(b) = G_{media}(f_b, \mathcal{Z})$ desde la ecuación (2.23). En la ecuación (6.9), los sesgos $b_i = \langle p_i, w_0 \rangle$, $i = 1, \dots, N_{pos}$ están ordenados en orden creciente. Además, se puede demostrar que:

²Imagen tomada de (Haydemar et al., 2011).

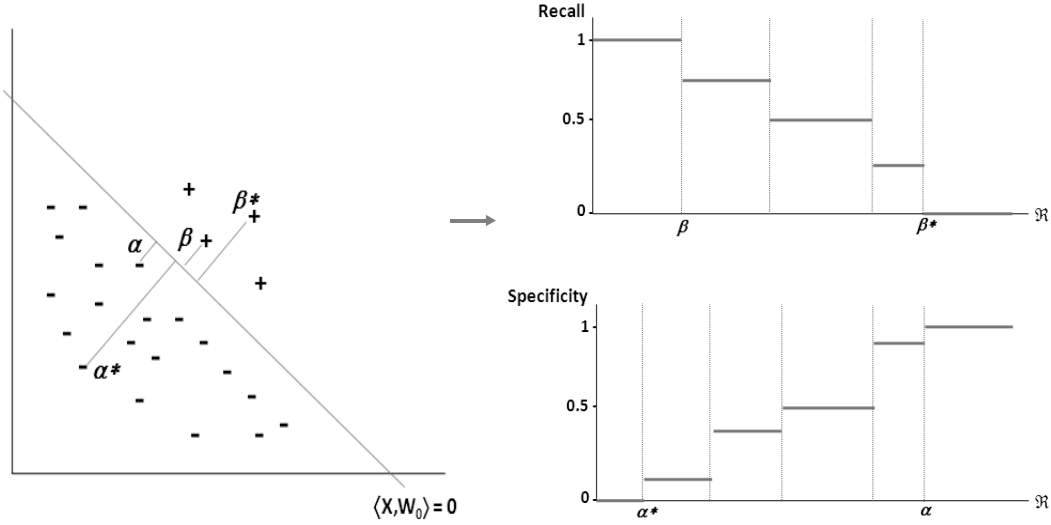


Figura 6.2: Representación gráfica de un ejemplo de las propiedades de las métricas *sensibilidad* y *especificidad*².

$$\max_{b \in \mathbb{R}} G_{media}(b) = \max_{i=1, \dots, N_{pos}} G_{media}(b_i) \quad (6.15)$$

En consecuencia, se define el clasificador $\Theta_{b_G}(x)$ desde la ecuación (6.8) como:

$$\Theta_{b_G}(x) = \text{signo}(f_{b_G}(x)) = \begin{cases} +1 & \text{si } \langle x, w_0 \rangle \geq b_G \\ -1 & \text{si } \langle x, w_0 \rangle < b_G \end{cases} \quad (6.16)$$

donde el sesgo b_G se define considerando los dos casos posibles:

- Si $\alpha < \beta$, \mathcal{Z} es un conjunto de entrenamiento linealmente separable. En este caso, el sesgo $b_G = \frac{\alpha + \beta}{2}$, es decir, el sesgo b_G es igual al sesgo estándar b_0 .
- Si $\alpha \geq \beta$, \mathcal{Z} es un conjunto de entrenamiento no linealmente separable. En este caso, el sesgo $b_G = \arg \max_{i=1, \dots, N_{pos}} G_{media}(b_i)$.

De acuerdo a los resultados teóricos enunciados anteriormente y considerando el rendimiento del clasificador con base en la métrica G_{media} , cuando el conjunto de datos es linealmente separable en el espacio kernel, el rendimiento de la SVM-1 y la SVM estándar es el mismo. Por otra parte, cuando el conjunto de datos no es linealmente separable en el espacio kernel, el rendimiento de la SVM-1 sería más alto que el de la SVM estándar, lo cuál significa que el nuevo clasificador alcanza un rendimiento más balanceado entre la tasa de verdaderos positivos y la tasa de verdaderos negativos.

6.2.3. SVM-2: Aplicación de una estrategia de post-procesamiento para ajustar el sesgo en función de la *Sensibilidad*

Como se ha mencionado anteriormente, la estrategia de post-procesamiento presentada en esta sección, está especialmente diseñada para aplicarse a problemas de clasificación binarios donde una clase (la clase positiva) se considera de mayor importancia que la otra clase (la clase negativa). En escenarios desbalanceados, en los cuáles una clase es considerada de mayor importancia que la otra clase, maximizar la *proporción de aciertos* es inapropiado. Por lo tanto, en este tipo de situaciones, tiene mayor prioridad maximizar la *sensibilidad* que mejorar la *proporción de aciertos* del clasificador global.

En el contexto de este problema, es muy importante obtener modelos que exhiban un alto rendimiento en la clase prioritaria, que en este contexto clínico corresponde a la clase asignada a las medidas incorrectas realizadas por el Guardian[®] RTCGMS, pero sin que esta mejora en el rendimiento vaya en detrimento de la clase asignada a las medidas correctas. En una UCI, es especialmente importante detectar las glucemias estimadas de manera errónea por el Guardian[®] RTCGMS, debido a que estas glucemias inadecuadas pueden llevar al personal sanitario a tomar decisiones equivocadas en el tratamiento que se suministra a los pacientes, especialmente si estas medidas incorrectas corresponden a situaciones de alta inestabilidad glucémica en el paciente (hiperglucemias e hipoglucemias).

De manera análoga a la estrategia de post-procesamiento presentada en la sección 6.2.2, esta estrategia de post-procesamiento, la cuál denominaremos SVM-2, también se basa en la consideración de que el sesgo b es un parámetro que se puede sintonizar. Por lo tanto, la función aprendida por la SVM estándar también se ajusta mediante el sesgo b . Sin embargo, a diferencia de la SVM-1, la estrategia de post-procesamiento descrita aquí se basa en ajustar el sesgo b , de tal manera que dado un valor fijo para la *sensibilidad* (ecuación (2.21)), se maximiza la *especificidad* (ecuación (2.20)). En otras palabras, el parámetro b se ajusta para mejorar el rendimiento de uso generalizado sobre la clase mayoritaria, mientras se mantiene un umbral para garantizar el rendimiento de uso generalizado sobre la clase minoritaria. Al igual que en la estrategia de post-procesamiento descrita en la sección 6.2.2, la solución que se aplica no modifica el problema de optimización original para entrenar la SVM y tampoco introduce nuevos hiper-parámetros.

El punto clave del enfoque propuesto aquí, es que una vez se obtiene el vector solución $w_0 = \sum_i \alpha_i y_i x_i$ del problema descrito en la ecuación (6.2), el sesgo b se puede considerar como un parámetro a ser ajustado. Con el nuevo sesgo, el hiperplano aprendido por la SVM es desplazado para maximizar la *especificidad*, manteniendo la *sensibilidad* a un nivel específico.

Por lo tanto, se considera el conjunto de clasificadores $\mathcal{F}(w_0)$ descrito en la ecuación (6.7). Se define un mapa $\Theta_b : \mathcal{X} \rightarrow \{-1, +1\}$, usando el clasificador $f_b(x) \in \mathcal{F}(w_0)$, donde $f_b(x) = \langle x, w_0 \rangle - b$, de tal manera que, dado un vector de entradas x , se asigne una etiqueta, en forma análoga a como se describe en la ecuación (6.8).

Dentro del conjunto de clasificadores $\mathcal{F}(w_0)$, se busca encontrar un balance o tradeoff entre la *Sensibilidad* (b), una función decremental de b y la *Especificidad* (b), una función incremental de b , para maximizar el rendimiento del clasificador. Para ello, se consideran las definiciones dadas en las ecuaciones (6.9), (6.10) y (6.11), para el caso de la *Sensibilidad* (b); y en las ecuaciones (6.12), (6.13) y (6.14), para el caso de la *Especificidad* (b).

Un camino para abordar este problema dado el conjunto de clasificadores $\mathcal{F}(w_0)$, es fijar un umbral $0 \leq r \leq 1$, de tal manera que la *Sensibilidad* (b) $\geq r$, a fin de garantizar una mínima tasa de verdaderos positivos en la clase positiva y maximizar la *Especificidad* (b). Por consiguiente, se considera el siguiente problema:

$$\max_{b \in \mathbb{R}} \text{Especificidad } (b)$$

$$\text{t.q. } \text{Sensibilidad } (b) \geq r, 0 \leq r \leq 1, f_b \in \mathcal{F}(w_0)$$

Por lo tanto, dado un valor de r , de tal manera que $0 \leq r \leq 1$, se considera el clasificador $\Theta_{b_r}(x)$ desde la ecuación (6.8) como:

$$\Theta_{b_r}(x) = \text{signo}(f_{b_r}(x)) = \begin{cases} +1 & \text{si } \langle x, w_0 \rangle \geq b_r \\ -1 & \text{si } \langle x, w_0 \rangle < b_r \end{cases} \quad (6.17)$$

donde el sesgo b_r se define considerando los dos casos posibles:

- Si $\alpha < \beta$, \mathcal{Z} es un conjunto de entrenamiento linealmente separable. En este caso, el sesgo $b_r = b_r^* = (1 - r)\beta + r\alpha$, de tal manera que 100r % del espacio entre α y β se da para los casos positivos. Por lo tanto, si $0 \leq r \leq 1$, está cerca de 1, más espacio se le da a la clase positiva (clase prioritaria) que a la clase negativa.
- Si $\alpha \geq \beta$, \mathcal{Z} es un conjunto de entrenamiento no linealmente separable. En este caso, el sesgo $b_r = b_{ir}$, donde $ir = \max\{i, i \leq 1 + N_{pos}(1 - r)\}$.

6.3. Detección de medidas correctas e incorrectas en los SMCGs aplicando nuevas estrategias de post-procesamiento de SVM

Para evaluar el rendimiento de las estrategias de post-procesamiento SVM-1 y SMV-2, los datos de los pacientes se dividieron en dos conjuntos: un conjunto de entrenamiento (-del inglés *training set*-, conformado por 20 pacientes) y un conjunto de prueba (-del inglés *testing set*-, conformado por los tres pacientes restantes).

En primer lugar, se realiza el entrenamiento de los diferentes clasificadores SVM-1 y SVM-2. En el entrenamiento de estos clasificadores se considera como entrada el vector de características $X = (ISIG, G_{RTCGMS}, Insulin)$, que se obtiene del conjunto de entrenamiento. El vector de características X se autoescala para tener media cero y varianza unitaria, con el fin de prevenir problemas con datos atípicos y garantizar que todas las variables tengan la misma importancia. Luego, se sintonizan los parámetros de los clasificadores utilizando el conjunto de entrenamiento previamente autoescalado.

En segundo lugar, los parámetros de los clasificadores obtenidos se utilizan para clasificar las estimaciones de glucosa de los restantes tres pacientes, los cuáles no se incluyeron en el entrenamiento. Previamente, se autoescala apropiadamente este conjunto de prueba o datos de *testing* usando los valores de medias y desviaciones estándar que se obtuvieron de los datos de *training*.

La metodología explicada anteriormente se repite para diferentes particiones de datos, es decir, combinaciones de 20 pacientes en el conjunto de *training* y 3 pacientes en el conjunto de

testing (ocho particiones en total). En cada partición, se sintonizan y validan diferentes clasificadores SVMs-1 y SVMs-2 usando funciones Kernel polinomial (a los cuáles denominaremos SVMs-1-Poly y SVMs-2-Poly) y funciones Kernel de base radial Gaussiana (RBF) (a los cuáles denominaremos SVMs-1-RBF y SVMs-2-RBF). La función Kernel polinomial se define como $K(x_i, x) = (x_i^T x + 1)^d$, donde d representa el grado del polinomio. La función Kernel RBF se define como $K(x_i, x) = e^{(-0,5 \frac{\|x_i - x\|^2}{\sigma^2})}$, donde σ representa el ancho de la función de base radial Gaussiana. El criterio empleado en el entrenamiento de las SVMs-1 y las SVMs-2 es la validación cruzada de cinco capas. Este procedimiento se repite 25 veces con el fin de garantizar un buen comportamiento estadístico.

En el caso de la estrategia de post-procesamiento SVM-1, las SVMs-1-Poly se sintonizan considerando 60 posibles combinaciones entre el valor de C (parámetro de penalización o regularización) y el valor de d (grado del polinomio), buscando estas combinaciones en una cuadrícula de búsqueda de dos dimensiones, donde $C = [2^{-4}, 2^{-3} \dots, 2^6, 2^7]$ y $d = [1, 2 \dots, 5]$. Por lo tanto, se resuelven en total 7500 problemas.

De igual forma, las SVMs-1-RBF se sintonizan considerando 132 posibles combinaciones entre el valor de C y el valor de σ (ancho de la función RBF) explorando estas combinaciones en una cuadrícula de búsqueda de dos dimensiones, donde $C = [2^{-4}, 2^{-3} \dots, 2^6, 2^7]$ y $\sigma = [2^{-3} \dots, 2^6, 2^7]$. Por consiguiente, se resuelven en total 16500 problemas. Todos los modelos obtenidos desde cada combinación de parámetros se validan usando su correspondiente conjunto de validación.

En el caso de la estrategia de post-procesamiento SVM-2, las SVMs-2-Poly se sintonizan considerando 60 posibles combinaciones entre el valor de C y el valor de d , buscando estas combinaciones en una cuadrícula de búsqueda de dos dimensiones, donde $C = [2^{-4}, 2^{-3} \dots, 2^6, 2^7]$ y $d = [1, 2 \dots, 5]$. A fin de garantizar una máxima tasa de verdaderos positivos en la clase incorrecta (*Sensibilidad* ($b \geq r$)), se fijó el umbral r considerando los siguientes siete posibles valores $r = [0,7, 0,75, 0,8, 0,85, 0,9, 0,95, 0,99]$. Por lo tanto, para cada uno de los valores del umbral r se resuelven en total 7500 problemas.

Asimismo, las SVMs-2-RBF se sintonizan considerando 132 posibles combinaciones entre el valor de C y el valor de σ , explorando estas combinaciones en una cuadrícula de búsqueda de dos dimensiones, donde $C = [2^{-4}, 2^{-3} \dots, 2^6, 2^7]$ y $\sigma = [2^{-3} \dots, 2^6, 2^7]$. El umbral r se fijó considerando los siete posibles valores mencionados anteriormente. Por consiguiente, para cada uno de los valores del umbral r se resuelven en total 16500 problemas. Todos los modelos obtenidos desde cada combinación de parámetros se validan usando su correspondiente conjunto de validación.

Finalmente, en cada partición, se seleccionan las combinaciones de parámetros de sintonización de los clasificadores que proporcionan el mejor valor de G_{media} , en el caso de las SVMs-1 (tanto en las SVMs-1-Poly como en las SVMs-1-RBF); y las combinaciones de parámetros que proporcionan el mejor valor de *sensibilidad* (tanto en las SVMs-2-Poly como en las SVMs-2-RBF). A continuación, se evalúa cada uno de los tres pacientes excluidos para validar el rendimiento general de las metodologías propuestas SVM-1 y SVM-2. Para validar cualquier clasificador propuesto en una aplicación en tiempo real (-del inglés *Real-Time (RT)*), los datos del paciente a evaluar se autoescalán apropiadamente utilizando los valores de medias y desviaciones estándar obtenidos de los datos con los que se construye el clasificador. Esto significa que en el contexto de datos en RT, no se considera la media y la desviación estándar del vector de características X de cada uno de los pacientes de validación para normalizar sus propios datos.

Un artículo basado en la aplicación de la estrategia de post-procesamiento SVM-1 para detectar medidas correctas e incorrectas en los SMCGs ha sido aceptado para publicación en (Leal et al., 2013a). De igual forma, una contribución sobre una comparativa entre la SVM convencional, la estrategia de post-procesamiento SVM-1 y la estrategia de post-procesamiento SVM-2 se ha presentado en (Leal et al., 2012a).

6.4. Resultados

6.4.1. Evaluación comparativa de los enfoques de SVM abordados

En primer lugar, se presentan los resultados de una evaluación comparativa entre los diferentes enfoques de SVM abordados. Diferentes clasificadores SVMs estándar, SVMs-1 y SVMs-2 usando funciones Kernel polinomial y RBF se sintonizan y validan empleando validación cruzada de cinco capas. Los resultados promedios de todos los experimentos se muestran en la Tabla 6.1.

	<i>Proporción de aciertos</i> promedio (%)	<i>Especificidad</i> promedio (%)	<i>Sensibilidad</i> promedio (%)	G_{media} promedio (%)	
SVMs usando funciones Kernel polinomial					
SVM-Poly	71,3	89,5	27,4	49,5	
SVM-1-Poly	68,9	70,3	66,0	68,1	
SVM-2-Poly	$r = 0,70$	68,5	68,3	68,5	68,4
	$r = 0,75$	43,1	30,8	73,5	47,6
	$r = 0,80$	62,4	56,1	77,8	66,1
	$r = 0,85$	55,3	44,4	82,7	60,6
	$r = 0,90$	47,2	31,5	86,6	52,2
	$r = 0,95$	39,4	18,6	91,7	41,3
	$r = 0,99$	32,6	6,7	98,0	25,6
SVMs usando funciones Kernel RBF					
SVM-RBF	69,9	76,9	53,5	64,1	
SVM-1-RBF	69,7	69,7	69,5	69,6	
SVM-2-RBF	$r = 0,70$	67,5	68,2	66,2	67,2
	$r = 0,75$	64,9	61,2	73,9	67,3
	$r = 0,80$	59,3	51,1	79,6	63,8
	$r = 0,85$	54,7	43,0	83,4	59,9
	$r = 0,90$	46,5	30,2	87,3	51,3
	$r = 0,95$	40,8	20,2	93,1	43,4
	$r = 0,99$	32,5	6,4	98,1	25,1

Tabla 6.1: Resultados promedios de las métricas *proporción de aciertos*, *especificidad*, *sensibilidad* y G_{media} para la validación cruzada de cinco capas.

A partir de esta tabla se puede llegar a las siguientes conclusiones:

- En escenarios sobre conjuntos de datos desbalanceados, se observa que el desempeño de las SVMs se mejora utilizando nuevos sesgos que se determinan teniendo en cuenta la distribución de las clases. Además, es importante destacar como una ventaja de utilizar este tipo de estrategias de post-procesamiento, el hecho de que no se modifica el problema de optimización para hallar la SVM y no se requiere la entonación de nuevos hiper-parámetros.
- Aunque la SVM-1 presenta un valor más bajo en la métrica *proporción de aciertos* que la SVM estándar, la SVM-1 muestra una mejor G_{media} que la SVM estándar. Como se

mencionó previamente, esto es debido a que la SVM-1 se basa en ajustar el sesgo b en función de la $G_{media}(b)$, la cuál considera que las dos clases positiva y negativa son igualmente relevantes para propósitos de clasificación, y por lo tanto, se logra un rendimiento más balanceado entre *sensibilidad* y *especificidad*.

- La SVM-2 (para todos los valores de r considerados) presenta el valor más bajo en la métrica *proporción de aciertos* de los tres enfoques de SVM abordados. Sin embargo, esta estrategia mejora considerablemente los resultados a nivel de *sensibilidad*, con respecto a la SVM convencional. Su principal inconveniente radica en que presenta los valores más bajos de *especificidad*, lo que indica que la mejora en la *sensibilidad* se logra a expensas de un significativo detrimento de la *especificidad*. Por tanto, el número de falsos positivos se incrementa de manera considerable, ocasionando un aumento significativo en el número de determinaciones de ABG que se toman al paciente, lo que va en contra de uno de los principales beneficios de utilizar los SMCGs en la UCI, reducir el número de punciones en el paciente. Para maximizar la *especificidad* y mantener una buena *sensibilidad*, el valor más óptimo para fijar el umbral r es 0,8, con el cuál se garantiza que la *especificidad* sea superior al 50,0% y una buena *sensibilidad*.
- El hecho de considerar funciones Kernel polinomial y RBF, no representa grandes diferencias en los resultados, a excepción de los valores reportados en las métricas *sensibilidad* y G_{media} en la SVM convencional, donde se observan valores mucho más altos cuando se usan funciones Kernel RBF.

6.4.2. Evaluación de la metodología propuesta basada en SVM-1

En cada partición, se valida cada uno de los tres pacientes excluidos, usando la combinación de parámetros que permite obtener la mejor G_{media} en las SVMs-1. Todas las lecturas de G_{RTCGMS} obtenidas cada 5 minutos en cada uno de los pacientes incluidos en el estudio se clasifican como medidas correctas e incorrectas. Sin embargo, al igual que en la sección solamente los pares de datos ABG/G_{RTCGMS} se emplean en el cálculo de las métricas que permiten evaluar el rendimiento de la metodología propuesta. La Tabla 6.2 muestra los resultados para los 23 pacientes en términos de *proporción de aciertos*, *especificidad*, *sensibilidad* y G_{media} , considerando diferentes funciones Kernel (SVMs-1-Poly y SVMs-1-RBF) y dependiendo del estado séptico de los pacientes. Además, en secciones posteriores se presentan algunos ejemplos representativos de la aplicación de esta metodología.

Ciertas conclusiones se pueden extraer de la Tabla 6.2:

- La principal conclusión que se puede extraer es que la SVM-1 proporciona un rendimiento balanceado y aceptable desde el punto de vista clínico para los pacientes en shock séptico y para los pacientes con sepsis, pero el rendimiento es deficiente para los pacientes no sépticos. Los valores de G_{media} reportados en cada una de las tres cohortes de pacientes confirman esta conclusión.

Tipo de paciente	No. Datos Pareados	Proporción de aciertos (%)		Especificidad (%)		Sensibilidad (%)		G_{media} (%)	
		SVM-1 Poly	SVM-1 RBF	SVM-1 Poly	SVM-1 RBF	SVM-1 Poly	SVM-1 RBF	SVM-1 Poly	SVM-1 RBF
Pacientes en shock séptico									
Post-quirúrgico	16	81,3	87,5	86,7	93,3	0,0	0,0	0,0	0,0
Médico	36	66,7	69,4	60,0	64,0	81,8	81,8	70,1	72,4
Médico	23	78,3	87,0	83,3	88,9	60,0	80,0	70,7	84,3
Post-quirúrgico	30	70,0	83,3	52,9	82,4	92,3	84,6	69,9	83,5
Post-quirúrgico	16	81,3	75,0	83,3	83,3	75,0	50,0	79,0	64,5
Post-quirúrgico	17	76,5	76,5	90,9	90,9	50,0	50,0	67,4	67,4
Post-quirúrgico	20	80,0	85,0	85,7	92,9	66,7	66,7	75,6	78,7
Post-quirúrgico(*)	11	90,9	90,9	90,9	90,9	-	-	-	-
Médico	39	71,8	79,5	69,6	78,3	75,0	81,3	72,2	79,8
Traumático	18	72,2	77,8	60,0	80,0	87,5	75,0	72,5	77,5
Media±DS		74,8±7,0	80,1±6,7	74,4±14,3	82,7±9,0	75,7±27,9	74,3±27,2	75,0±24,3	78,4±26,2
Mediana (IQR)		77,4 (9,1)	81,4 (9,7)	83,3 (24,1)	86,1 (10,3)	75,0 (21,8)	75,0 (31,3)	70,7 (2,6)	77,5 (12,4)
Pacientes con sepsis									
Post-quirúrgico	32	81,3	84,4	87,0	91,3	66,7	66,7	76,2	78,0
Post-quirúrgico	28	60,7	75,0	40,0	70,0	72,2	77,8	53,7	73,8
Post-quirúrgico	17	88,2	100,0	85,7	100,0	100,0	100,0	92,6	100,0
Médico	18	66,7	77,8	60,0	80,0	75,0	75,0	67,1	77,5
Médico	28	57,1	71,4	50,0	72,7	83,3	66,7	64,5	69,6
Post-quirúrgico	25	76,0	88,0	76,2	90,5	75,0	75,0	75,6	82,4
Traumático	12	91,7	100,0	90,9	100,0	100,0	100,0	95,3	100,0
Traumático	20	75,0	85,0	71,4	85,7	83,3	83,3	77,1	84,5
Médico	28	64,3	64,3	33,3	33,3	100,0	100,0	57,7	57,7
Media±DS		71,6±12,1	80,8±12,2	67,1±21,3	80,7±20,6	80,9±13,1	80,9±13,9	73,7±14,3	80,8±13,6
Mediana (IQR)		75,0 (17,0)	84,4 (13,0)	71,4 (35,7)	85,7 (18,6)	83,3 (25,0)	77,8 (25,0)	75,6 (12,6)	78,0 (10,7)
Pacientes sin sepsis									
Traumático	23	73,9	78,3	78,9	78,9	50,0	75,0	62,8	76,9
Médico	34	58,8	61,8	60,0	64,0	55,6	55,6	57,8	59,7
Post-quirúrgico	10	60,0	60,0	66,7	66,7	0,0	0,0	0,0	0,0
Traumático	36	58,3	44,4	69,6	39,1	38,5	53,8	51,8	45,9
Media±DS		62,1±7,5	59,2±13,8	68,4±7,8	60,5±16,7	44,5±25,0	55,6±32,2	55,2±29,1	58,0±33,0
Mediana (IQR)		59,4 (4,8)	60,9 (9,8)	68,2 (6,9)	65,4 (12,0)	44,3 (22,5)	54,7 (20,0)	54,8 (20,2)	52,8 (29,5)

(*) Significa que en este paciente $N_{pos} = 0$ y por lo tanto la *Sensibilidad* y la G_{media} no pueden ser evaluadas.

Tabla 6.2: Resultados de la evaluación de la metodología basada en SVM-1 de acuerdo con el estado séptico de los pacientes.

- Respecto a la cohorte de pacientes en shock séptico se puede apreciar un comportamiento homogéneo y valores de G_{media} moderadamente altos, excepto en el caso de un paciente post-quirúrgico. El valor de la mediana (Q1 – Q3) de la métrica G_{media} es 70,7 % (69,9 % – 72,5 %) para la SVM-1 Poly y 77,5 % (67,4 % – 79,8 %) para la SVM-1 RBF. El paciente que reporta una *sensibilidad* nula, corresponde a un paciente post-quirúrgico con $N_{pos} = 1$ y $N_{neg} = 15$. Además, los resultados son más homogéneos cuando se utilizan funciones Kernel polinomiales.
- En general, en el caso del grupo de pacientes con sepsis, también se observa un comportamiento homogéneo y valores de G_{media} moderadamente altos. El valor de la mediana (Q1 – Q3) de la métrica G_{media} es 75,6 % (64,5 % – 77,2 %) para la SVM-1 Poly y 78,0 % (73,8 % – 84,5 %) para la SVM-1 RBF. Sin embargo, el rango intercuartílico es mucho más amplio en esta cohorte de pacientes, lo que indica que existe mayor variabilidad en estos datos. Se aprecian resultados similares tanto en el uso de funciones Kernel polinomiales como funciones Kernel RBF.
- Finalmente, en el grupo de pacientes no sépticos, se aprecia un comportamiento homogéneo, pero negativo en cuanto a los valores de G_{media} se refiere, los cuáles son bastante bajos. El valor de la mediana (Q1 – Q3) de la métrica G_{media} es 54,7 % (38,8 % – 59,0 %) para la SVM-1 Poly y 52,8 % (34,4 % – 64,0 %) para la SVM-1 RBF. Analizando el rango intercuartílico, se observa una variabilidad mucho mayor en los resultados, presentándose esta disparidad de manera más frecuente cuando se utilizan funciones Kernel RBF. También existe en este grupo un paciente que reporta una *sensibilidad* nula. Este caso corresponde también a un paciente post-quirúrgico con $N_{pos} = 1$ y $N_{neg} = 9$. Es importante resaltar que el tamaño muestral de esta cohorte de pacientes es mucho más bajo, en comparación con las otras dos. Esto se debe a que este tipo de pacientes son los menos frecuentes en las UCIs.

6.4.3. Evaluación de la metodología propuesta basada en SVM-2

En cada partición, se valida cada uno de los tres pacientes excluidos, usando la combinación de parámetros que permite obtener la mejor *sensibilidad* en las SVMs-2 y fijando el umbral $r = 0,8$. La Tabla 6.3 muestra los resultados para los 23 pacientes en términos de *proporción de aciertos*, *especificidad*, *sensibilidad* y G_{media} , considerando diferentes funciones Kernel (SVMs-2-Poly y SVMs-2-RBF) y de acuerdo con el estado séptico de los pacientes. Algunos ejemplos representativos se presentan en secciones posteriores.

Analizando los resultados presentados en la Tabla 6.3, se puede llegar a las siguientes conclusiones:

- La principal conclusión es que a diferencia de la SVM-1, la SVM-2 no proporciona un rendimiento adecuado desde el punto de vista clínico para ninguna de las tres cohortes de pacientes, de acuerdo a los valores de *especificidad* reportados. Sin embargo, es importante destacar que se logra el objetivo de mantener un umbral para garantizar un buen rendimiento sobre la clase minoritaria, lo que se traduce en que los valores de *sensibilidad* reportados en los pacientes en shock séptico y en los pacientes con sepsis son muy altos. En el caso de los pacientes no sépticos, los valores de *sensibilidad* reportados corresponden a valores aceptables clínicamente.

Tipo de paciente	No. Datos Pareados	Proporción de aciertos (%)		Especificidad (%)		Sensibilidad (%)		G_{media} (%)	
		SVM-2 Poly	SVM-2 RBF	SVM-2 Poly	SVM-2 RBF	SVM-2 Poly	SVM-2 RBF	SVM-2 Poly	SVM-2 RBF
Pacientes en shock séptico									
Post-quirúrgico	16	62,5	12,5	66,7	13,3	0,0	0,0	0,0	0,0
Médico	36	52,8	25,0	32,0	0,0	100,0	81,8	56,6	0,0
Médico	23	47,8	34,8	33,3	16,7	100,0	100,0	57,7	40,8
Post-quirúrgico	30	60,0	50,0	29,4	11,8	100,0	100,0	54,2	34,3
Post-quirúrgico	16	56,3	50,0	41,7	33,3	100,0	100,0	64,5	57,7
Post-quirúrgico	17	70,6	47,1	63,6	45,5	83,3	50,0	72,8	47,7
Post-quirúrgico	20	55,0	40,0	42,9	21,4	83,3	83,3	59,8	42,3
Post-quirúrgico(*)	11	72,7	72,7	72,7	72,7	-	-	-	-
Médico	39	61,5	59,0	43,5	34,8	87,5	93,8	61,7	57,1
Traumático	18	55,6	44,4	30,0	0,0	87,5	100,0	51,2	0,0
Media±DS		58,4±7,7	42,9±16,9	43,6±16,2	22,4±22,4	91,4±31,8	88,6±33,7	63,1±20,9	44,6±24,5
Mediana (IQR)		58,1 (7,1)	45,8 (13,9)	42,3 (26,3)	19,0 (22,3)	87,5 (16,7)	93,8 (18,2)	57,7 (7,4)	40,8 (47,7)
Pacientes con sepsis									
Post-quirúrgico	32	56,3	59,4	43,5	47,8	88,9	88,9	62,2	65,2
Post-quirúrgico	28	60,7	67,9	20,0	10,0	83,3	100,0	40,8	31,6
Post-quirúrgico	17	58,8	76,5	50,0	71,4	100,0	100,0	70,7	84,5
Médico	18	61,1	61,1	30,0	30,0	100,0	100,0	54,8	54,8
Médico	28	39,3	28,6	22,7	9,1	100,0	100,0	47,7	30,2
Post-quirúrgico	25	40,0	36,0	28,6	23,8	100,0	100,0	53,5	48,8
Traumático	12	58,3	58,3	54,5	54,5	100,0	100,0	73,9	73,9
Traumático	20	40,0	35,0	21,4	14,3	83,3	83,3	42,3	34,5
Médico	28	60,7	60,7	26,7	26,7	100,0	100,0	51,6	51,6
Media±DS		52,4±9,9	52,9±16,5	32,9±12,9	31,4±21,6	92,6±7,6	97,1±6,3	55,2±11,6	55,2±19,1
Mediana (IQR)		58,3 (20,7)	59,4 (25,1)	28,6 (20,8)	26,7 (33,5)	100,0 (11,1)	100,0 (0,0)	53,5 (14,5)	51,6 (30,7)
Pacientes sin sepsis									
Traumático	23	69,6	52,2	68,4	47,4	75,0	75,0	71,6	59,6
Médico	34	47,1	35,3	36,0	20,0	77,8	77,8	52,9	39,4
Post-quirúrgico	10	50,0	20,0	44,4	11,1	100,0	100,0	66,7	33,3
Traumático	36	52,8	52,8	60,9	39,1	38,5	76,9	48,4	54,9
Media±DS		54,4±10,1	43,7±15,6	52,6±14,8	31,6±16,7	59,3±25,5	77,8±11,8	55,8±11,0	49,6±12,4
Mediana (IQR)		51,4 (7,7)	43,7 (20,9)	52,7 (20,4)	29,6 (23,4)	76,4 (17,5)	77,4 (6,9)	59,8 (16,1)	47,2 (18,1)

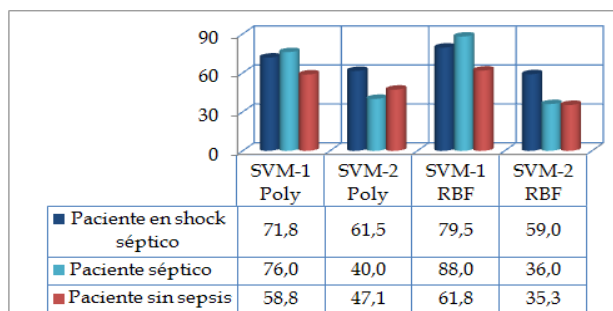
(*) Significa que en este paciente $N_{pos} = 0$ y por lo tanto la *Sensibilidad* y la G_{media} no pueden ser evaluadas.

Tabla 6.3: Resultados de la evaluación de la metodología basada en SVM-2 de acuerdo con el estado séptico de los pacientes.

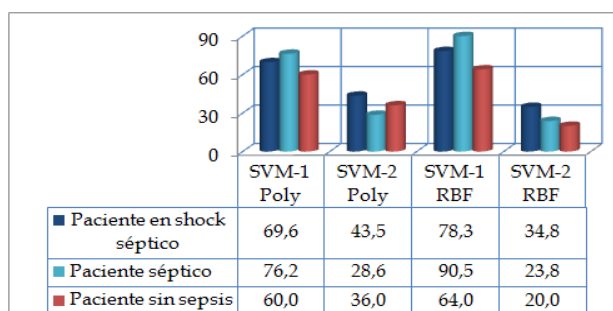
- En el grupo de pacientes en shock séptico, si bien se presentan valores de *sensibilidad* muy altos (excepto en el caso de un paciente post-quirúrgico), el comportamiento de los pacientes de este grupo respecto a esta métrica no es del todo homogéneo y es evidente la variabilidad en los resultados. El valor de la mediana (Q1 – Q3) de la métrica *sensibilidad* es 87,5 % (83,3 % – 100,0 %) para la SVM-2 Poly y 93,8 % (81,8 % – 100,0 %) para la SVM-2 RBF. Este comportamiento es el mismo independientemente del tipo de función Kernel que se utilice. Sin embargo, es importante mencionar que los valores de la *especificidad* mejoran de manera considerable cuando se usan funciones Kernel polinomiales, si bien estos resultados no corresponden a un buen rendimiento del clasificador. El valor de la mediana (Q1 – Q3) de la métrica *especificidad* es 42,3 % (32,3 % – 58,6 %) para la SVM-2 Poly y 19,0 % (12,2 % – 34,4 %) para la SVM-2 RBF, lo que muestra una amplia variabilidad en esta métrica.
- Respecto a los pacientes con sepsis, los valores de *sensibilidad* reportados también son muy altos, pero en contraste, la *especificidad* es muy baja. El valor de la mediana (Q1 – Q3) de la métrica *sensibilidad* es 100,0 % (88,9 % – 100,0 %) para la SVM-2 Poly y 100,0 % (100,0 % – 100,0 %) para la SVM-2 RBF. En cuanto a la *especificidad*, el valor de la mediana (Q1 – Q3) es 28,6 % (22,7 % – 43,5 %) para la SVM-2 Poly y 26,7 % (14,3 % – 47,8 %) para la SVM-2 RBF. Sólo se reportan resultados muy homogéneos en los valores de *sensibilidad* al utilizar funciones Kernel RBF, pero en los valores de *especificidad* la variabilidad es muy amplia para ambos tipos de funciones Kernel.
- Finalmente, en la cohorte de pacientes no sépticos, los valores de *sensibilidad* reportados también son moderadamente altos, presentándose una menor disparidad cuando se emplean funciones Kernel RBF. El valor de la mediana (Q1 – Q3) de la *sensibilidad* es 76,4 % (65,9 % – 83,3 %) para la SVM-2 Poly y 77,4 % (76,4 % – 83,3 %) para la SVM-2 RBF. En cuanto a la *especificidad*, el valor de la mediana (Q1 – Q3) es 52,7 % (42,3 % – 62,8 %) para la SVM-2 Poly y 29,6 % (17,8 % – 41,2 %) para la SVM-2, reflejando estos resultados un comportamiento similar a lo que sucede en los pacientes en shock séptico.

6.4.4. Ejemplos de la metodología propuesta para detectar medidas correctas e incorrectas realizadas por los SMCs aplicando las estrategias de post-procesamiento de SVMs

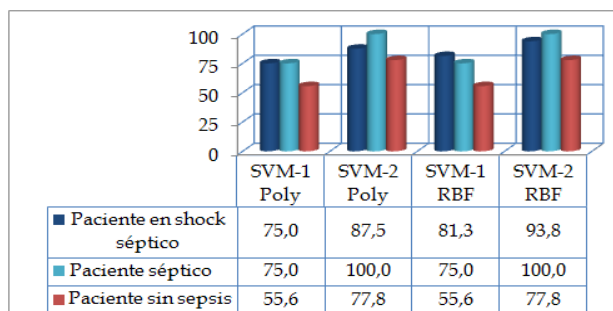
Con el fin de ilustrar la capacidad de la metodología propuesta para detectar medidas correctas e incorrectas, se realiza el análisis de los pares de datos ABG/G_{RTCGMS} de los 23 pacientes críticamente enfermos. Los resultados de la clasificación de tres pacientes representativos de acuerdo con el estado séptico del paciente utilizando la metodología propuesta con base en la SVM-1 se muestran en la Figura 6.3. En la misma Figura 6.3 se muestran los resultados para estos tres pacientes representativos utilizando la metodología propuesta con base en la SVM-2. Estos pacientes se seleccionaron porque tienen similares valores de G_{media} y *sensibilidad* a los valores medianos de G_{media} y *sensibilidad*, reportados en las Tablas 6.2 y 6.3, respectivamente, para cada una de las cohortes de pacientes. Estos pacientes corresponden a un paciente en shock séptico (39 estimaciones de glucosa, de las cuáles 23 son correctas (59,0 %) y 16 son incorrectas (41,0 %)), un paciente séptico (25 estimaciones de glucosa, de las cuáles 21 son correctas (84,0 %) y 4 son incorrectas (16,0 %)) y un paciente no séptico (34 estimaciones de glucosa, de las cuáles 25 son correctas (73,5 %) y 9 son incorrectas (26,5 %)). La Figura 6.3 muestra los resultados para los tres pacientes representativos en términos de la *proporción de aciertos* (Figura 6.3(a)), la *especificidad* (Figura 6.3(b)), la *sensibilidad* (Figura 6.3(c)) y la G_{media} (Figura 6.3(d)).



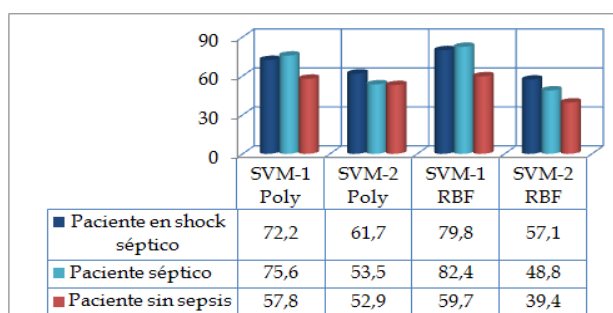
(a) Proporción de aciertos



(b) Especificidad



(c) Sensibilidad



(d) G_{media}

Figura 6.3: Resultados de la clasificación de tres pacientes representativos de cada cohorte de pacientes de acuerdo con el estado séptico, utilizando las metodologías propuestas con base en las estrategias de post-procesamiento SVM-1 y SVM-2. Los resultados mostrados en las diferentes tablas son porcentajes (%).

Del análisis de la Figura 6.3 se observa que la estrategia de post-procesamiento SVM-1 per-

mite obtener una *proporción de aciertos* balanceada y aceptable tanto en el paciente en shock séptico como en el paciente con sepsis (Figura 6.3(a)). Por lo tanto, estos pacientes muestran un desempeño adecuado de la metodología propuesta con base en la SVM-1, tanto en la detección de las medidas correctas (Figura 6.3(b)), como en la detección de las medidas incorrectas (Figura 6.3(c)). Para estos pacientes, la estrategia de post-procesamiento SVM-1 alcanza un rendimiento balanceado aceptable entre la tasa de verdaderos positivos y la tasa de verdaderos negativos (Figura 6.3(d)). Aunque el paciente no séptico también reporta valores balanceados en sus métricas, estos resultados son bajos para ser aceptables desde el punto de vista clínico.

En la Figura 6.3 también se observa que la estrategia de post-procesamiento SVM-2 da mayor importancia a la detección de las medidas incorrectas (Figura 6.3(c)) en detrimento de mejorar la *proporción de aciertos* global (Figura 6.3(a)). Esta mejora significativa de la *sensibilidad* se presenta en los tres pacientes y los valores reportados son aceptables clínicamente, aunque continúan presentándose los resultados más bajos en el paciente no séptico. Si bien los resultados reportados en los valores de *sensibilidad* confirman que la estrategia de post-procesamiento SVM-2 garantiza una máxima tasa de detección de medidas incorrectas, la máxima *especificidad* (Figura 6.3(b)) que se logra es demasiado baja para ser aceptada clínicamente.

Desde los resultados mencionados anteriormente en las Tablas 6.2 y 6.3, se concluye que la metodología basada en la estrategia de post-procesamiento SVM-1 fue la que presentó los mejores resultados para abordar el problema de detección de fallos en los SMCGs. Por esta razón, se presentan algunos de los resultados reportados de la clasificación de las estimaciones de G_{RTCGMS} en los tres pacientes representativos, contrastando los resultados obtenidos al aplicar la metodología basada en SVM-1 con lo indicado por los criterios ISO (Figura 6.4). En el caso de las estimaciones de G_{RTCGMS} clasificadas como correctas e incorrectas por la metodología basada en la estrategia de post-procesamiento SVM-1, cada figura ilustra t_{neg} , t_{pos} , f_{neg} y f_{pos} . t_{neg} en este caso son medidas de G_{RTCGMS} que según el esquema de clasificación definido en 6.1 son correctas (clase -1) y la metodología propuesta las clasifica como tal, y t_{pos} son medidas de G_{RTCGMS} que son incorrectas (clase 1) y la metodología propuesta también las clasifica como incorrectas. Por otra parte, f_{neg} en este caso son medidas de G_{RTCGMS} que se han etiquetado originalmente como incorrectas, pero la metodología propuesta las clasifica como correctas. Finalmente, f_{pos} son medidas de G_{RTCGMS} que son correctas y la metodología propuesta las clasifica erróneamente como incorrectas. En la Figura 6.4, un triángulo rojo apuntando hacia la derecha indica un t_{neg} , un triángulo magenta apuntando hacia abajo indica un t_{pos} , un triángulo azul apuntando hacia arriba indica un f_{neg} y un triángulo negro apuntando hacia la izquierda indica un f_{pos} .

El caso específico de un paciente en shock séptico usando una SVM-1-RBF se ilustra en la Figura 6.4(a). Esta figura muestra que este paciente presenta una buena *sensibilidad*. De 16 medidas incorrectas hechas por el Guardian[®] RTCGMS, tres se clasifican erróneamente; por tanto, se detecta el 81,3% de las medidas incorrectas. En forma similar, este paciente presenta una moderada *especificidad*. Cinco de 23 medidas correctas (21,7%) se clasifican erróneamente y se reportan como medidas incorrectas.

De igual importancia, en la Figura 6.4(b) se muestra el caso específico de un paciente con sepsis usando una SVM-1-RBF. Se puede observar que este paciente presenta una moderada *sensibilidad*. De cuatro medidas incorrectas hechas por el Guardian[®] RTCGMS, una se clasifica erróneamente; por tanto, se detecta el 75,0% de las medidas incorrectas. Sin embargo, la Figura 6.4(b) ilustra una alta *especificidad*. Dos de 21 (9,5%) medidas correctas se clasifican errónea-

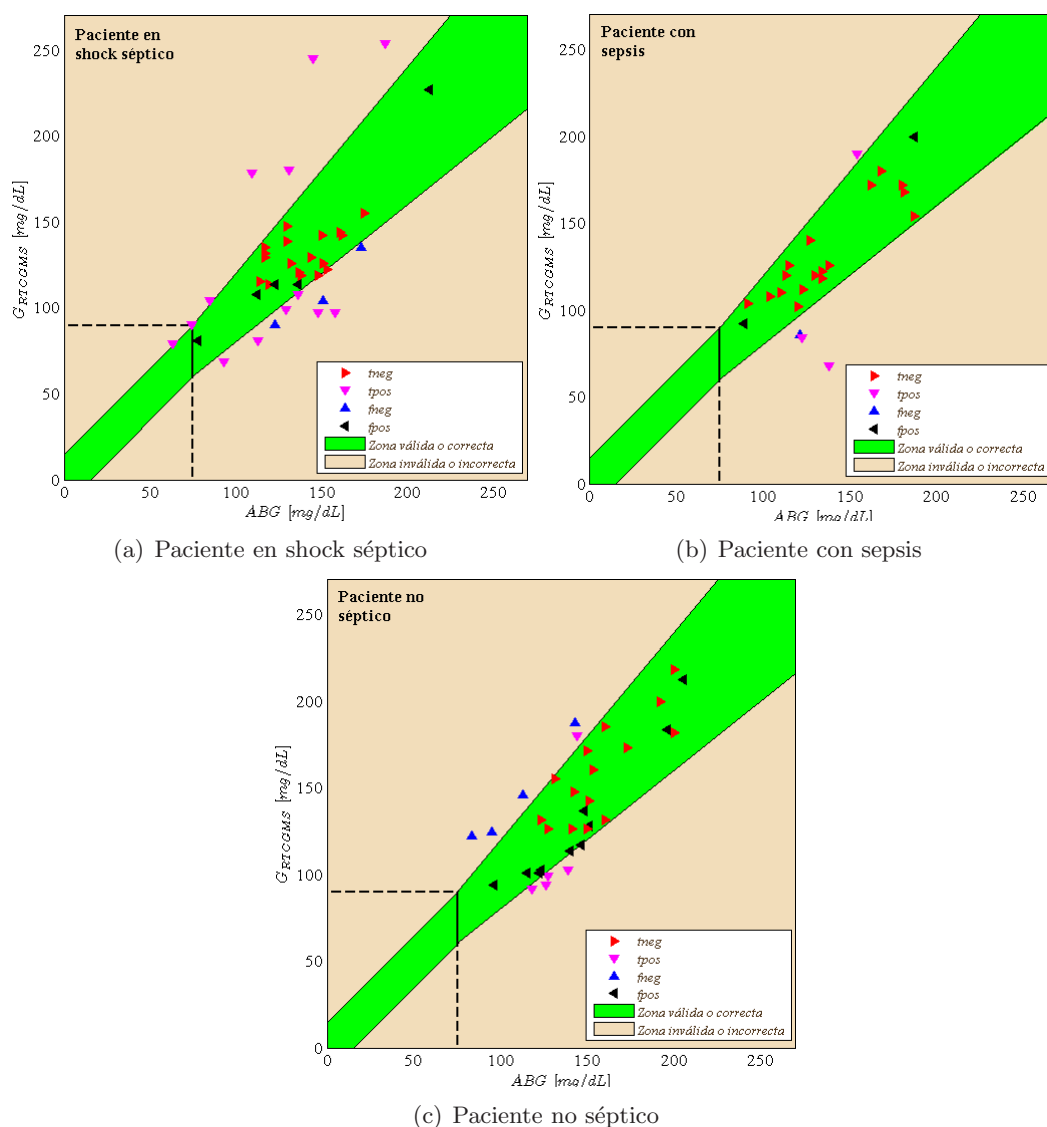


Figura 6.4: Clasificación de los pares de datos ABG/G_{RTCMS} de tres pacientes representativos contrastando los resultados obtenidos al aplicar la metodología SVM-1 con los criterios ISO.

mente y se reportan como medidas incorrectas.

De igual forma que en los casos anteriores, para el caso específico de un paciente sin sepsis usando una SVM-1-Poly, la Figura 6.4(c) muestra que este paciente presenta similares valores de *sensibilidad* y *especificidad*. De nueve medidas incorrectas realizadas por el Guardian® RTCMS, cuatro se clasifican erróneamente; por tanto, se detecta el 55,6% de las medidas incorrectas. Por otra parte, diez de 25 (40,0%) medidas correctas se clasifican erróneamente y se reportan como medidas incorrectas.

Además, en las Figuras 6.5, 6.6, 6.7 se ilustran los perfiles de estos tres pacientes representativos de cada cohorte de pacientes. Los perfiles mostrados son ejemplos de los resultados obtenidos usando la metodología propuesta basándose en la estrategia de post-procesamiento SVM-1. Un cuadrado verde y un círculo magenta corresponden a medidas de G_{RTCMS} rea-

lizadas por el Guardian[®] RTCGMS que son clasificadas como medidas correctas e incorrectas por el clasificador, respectivamente. Un diamante rojo corresponde a una medida de ABG que es utilizada como medida de referencia. Respecto a las medidas de G_{RTCGMS} , un triángulo cian apuntando hacia la derecha indica un t_{neg} , un triángulo amarillo apuntando hacia abajo indica un t_{pos} , un triángulo negro apuntando hacia la izquierda indica un f_{pos} y un triángulo azul apuntando hacia arriba indica un f_{neg} . Un asterisco negro corresponde a un punto de calibración.

Para el paciente en shock séptico (Figura 6.5), la metodología propuesta detecta tres falsos eventos hiperglucémicos estimados por el Guardian[®] RTCGMS (G_{RTCGMS} superior a 160 mg/dL cuando ABG es inferior a 160 mg/dL). De acuerdo a la opinión del médico, es de gran importancia la correcta detección del segundo falso episodio hiperglucémico (resaltado con un círculo negro discontinuo). Para este caso, G_{RTCGMS} es 178 mg/dL mientras ABG es 109 mg/dL. Si el médico considera como fiable la medida indicada por el monitor, se inicia una dosis de CIII para mantener la concentración de ABG entre 120 y 160 mg/dL. Esta decisión podría luego causar un severo evento hipoglucémico en el paciente. Adicionalmente, la metodología propuesta basada en SVM-1 detecta correctamente dos relevantes eventos hipoglucémicos (ABG inferior a 80 mg/dL). En estos eventos, mientras las medidas de ABG indican eventos hipoglucémicos, el Guardian[®] RTCGMS estima medidas de glucosa en el rango euglucémico (G_{RTCGMS} entre 80 y 160 mg/dL). Considerando la opinión del especialista médico, la correcta detección del segundo episodio hipoglucémico (resaltado con un círculo azul discontinuo), permite detectar una medida de glucosa erróneamente estimada por el Guardian[®] RTCGMS. Los falsos negativos que se presentan pueden ser considerados como errores de clasificación de menor importancia, porque el médico sólo comienza CIII si el paciente tiene valores de glucemia superiores a 150 mg/dL.

De hecho, si se contrasta este perfil glucémico con el perfil ilustrado en la Figura 5.14, se observa un rendimiento más balanceado entre *sensibilidad* y *especificidad* por parte de la metodología basada en la estrategia de post-procesamiento SVM-1 en comparación con el rendimiento obtenido aplicando la metodología basada en combinar PCA con CBR. Sin embargo, no se debe olvidar que los esquemas de clasificación en ambos enfoques son diferentes. En el capítulo anterior se tomó como base un criterio de precisión clínico (Clarke EGA) para definir el esquema de clasificación, mientras en este trabajo se toma como base un criterio de precisión analítico (los Criterios ISO).

Para el paciente con sepsis (Figura 6.6), el rendimiento de la metodología propuesta puede ser considerado fiable, porque se muestra su capacidad para detectar estados euglucémicos falsamente estimados por el Guardian[®] RTCGMS como hipoglucecias (por ejemplo, G_{RTCGMS} es 68 mg/dL mientras ABG es 138 mg/dL). Sólo se reporta un falso negativo, pero este resultado no altera las conclusiones generales obtenidas del perfil.

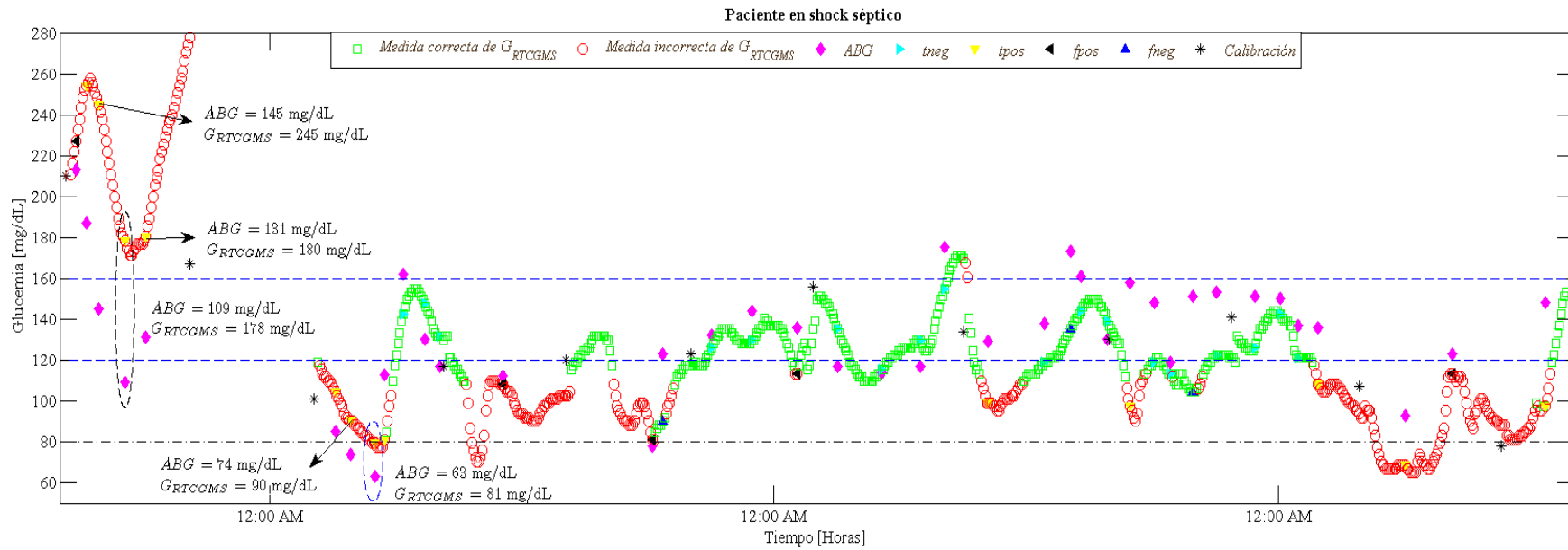


Figura 6.5: Perfil de un paciente representativo en shock séptico obtenido al aplicar la metodología basada en SVM-1.

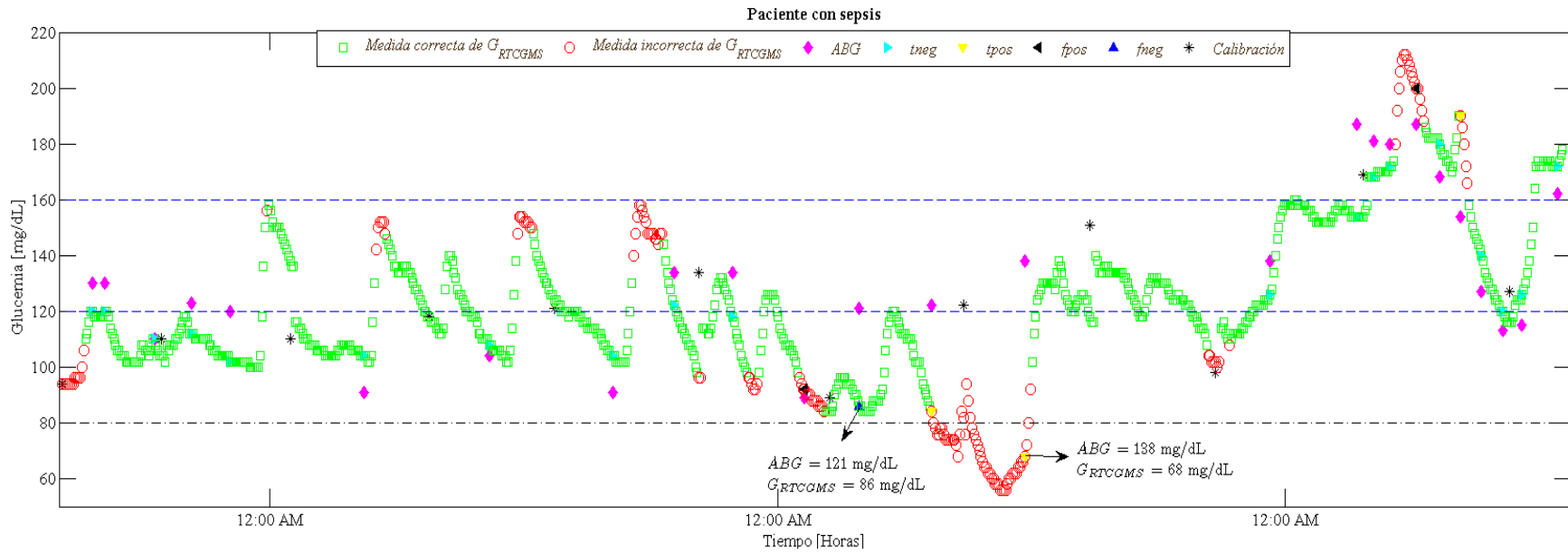


Figura 6.6: Perfil de un paciente representativo con sepsis obtenido al aplicar la metodología basada en SVM-1.

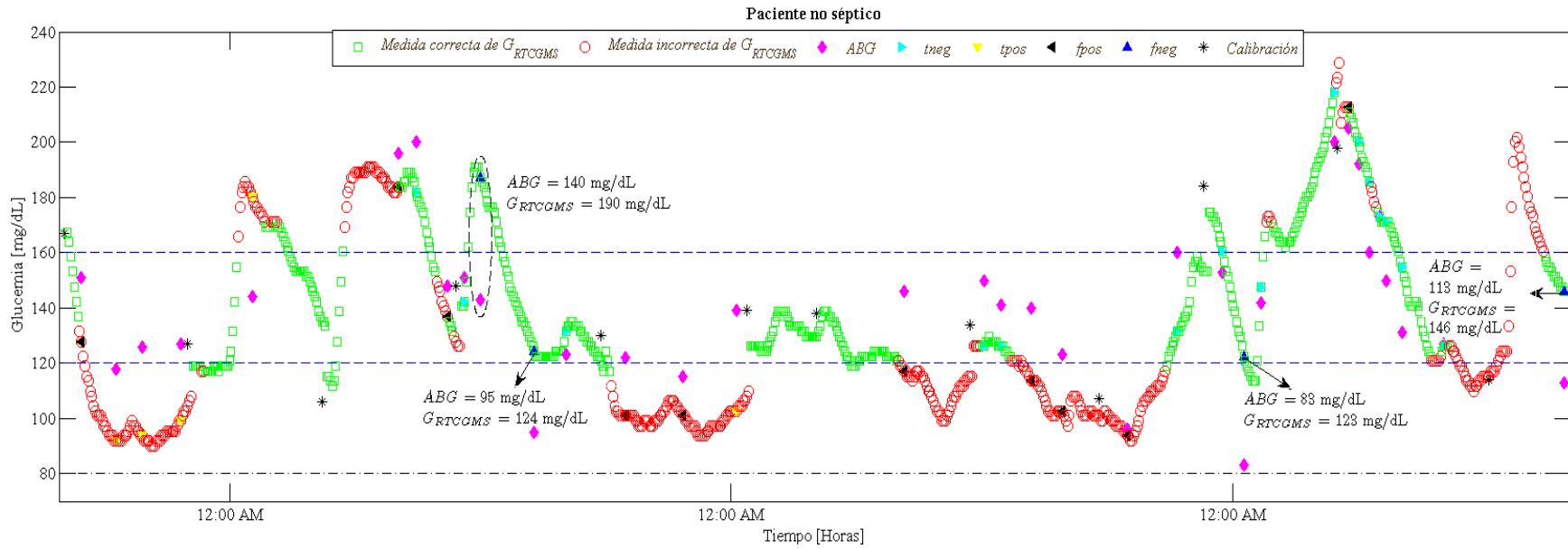


Figura 6.7: Perfil de un paciente representativo no séptico obtenido al aplicar la metodología basada en SVM-1.

Finalmente, para el paciente sin sepsis (Figura 6.7), la metodología propuesta puede ser considerada como poco fiable, debido a la baja *sensibilidad* reportada, lo que implica un número significativo de falsos negativos. Aunque los falsos negativos reportados no tienen un impacto clínico significativo en lo que respecta a la administración de la terapia de insulina al paciente, pueden ocasionar un tratamiento inadecuado. Por ejemplo, la correcta detección del falso estado hiperglucémico estimado por el Guardian[®] RTCGMS (resaltado con un círculo negro discontinuo) puede ser considerado como un error importante. Para esta sobrestimación de la glucemia producida por el Guardian[®] RTCGMS, el cuál reportó 190 mg/dL, mientras *ABG* indicó 140 mg/dL, se podría iniciar una dosis de CIII. Este tratamiento, aunque no compromete la vida del paciente, no es el tratamiento adecuado.

Aunque al aplicar la metodología con base en la estrategia de post-procesamiento SVM-1 se presentan falsos positivos en todos los pacientes, su número es reducido en los pacientes en shock séptico y en los pacientes con sepsis. Por tanto, estos errores de clasificación no alteran significativamente el rendimiento de la metodología propuesta. Recordemos que la presencia de falsos positivos no tiene consecuencias clínicas para los pacientes, debido a que no dan lugar a que el personal de la UCI modifique la terapia de insulina intravenosa. Los falsos positivos conducen a que el personal de la UCI tome determinaciones de *ABG* con mayor frecuencia con el fin de garantizar la seguridad del paciente. Si bien esta no es la situación ideal, dado que con el uso del SMCG se busca reducir el número de punciones en el paciente, tampoco se incurre en riesgos clínicos.

Por otra parte, la presencia de falsos negativos es un indicador de pérdida de *sensibilidad*. Si el personal de la UCI se fía del diagnóstico dado por el clasificador, estaría considerando estas medidas como correctas, cuando realmente se trata de falsos negativos. Este diagnóstico erróneo podría llevarlos a modificar la dosis de CIII, lo que podría provocar situaciones de inestabilidad glucémica en el paciente. En otras palabras, los falsos negativos podrían tener un significativo impacto clínico en cuanto a la administración de la terapia de insulina al paciente, debido a que estos errores de clasificación podrían ocasionar un tratamiento inadecuado.

6.5. Análisis de los resultados obtenidos

Los SMCGs no son suficientemente exactos y fiables en la actualidad para ser utilizados en la toma de decisiones terapéuticas. Sin embargo, en este capítulo se demostró que es factible el uso de SVMs no convencionales para mejorar la fiabilidad en los SMCGs empleando la información proporcionada por el propio monitor, e incorporando variables acerca de la condición clínica de los pacientes.

Las siguientes conclusiones se pueden extraer de toda la experimentación empírica realizada:

- La experimentación en pacientes críticamente enfermos muestra que la aplicación de una estrategia de post-procesamiento para ajustar el sesgo b en función de la G_{media} (SVM-1) puede ser una buena alternativa para la detección de las medidas correctas e incorrectas realizadas por los SMCGs, debido a que permite obtener un rendimiento más equilibrado entre *sensibilidad* y *especificidad*, es decir, entre falsos positivos y falsos negativos.
- Los resultados obtenidos son prometedores para pacientes en shock séptico o con sepsis, para los cuáles la metodología propuesta SVM-1 puede considerarse como fiable, pero no

puede ser considerada como buena para los pacientes sin sepsis. Esta mayor fiabilidad en pacientes en shock séptico, podría explicarse por la alteración de la permeabilidad capilar, una característica fisiopatológica de este tipo de pacientes. Esta alteración de la permeabilidad capilar se traduciría en un aumento de la difusión de la glucosa intravascular al espacio intersticial. En consecuencia, para esta cohorte de pacientes, la información proporcionada por el SMCG (*ISIG* y *G_{RTCGMS}*) como entrada para el sistema de detección de fallos propuesto sería más fiable.

- Los resultados indican que la estrategia de post-procesamiento SVM-2 presenta altos valores de *sensibilidad*, lo que implica que tiende a clasificar adecuadamente un alto porcentaje de medidas incorrectas. Sin embargo, su principal inconveniente es que genera un elevado número de falsas alarmas, lo que afecta significativamente los valores de *especificidad*. Generalmente, la presencia de falsos positivos no tiene consecuencias clínicas para los pacientes, porque su presencia no permite que el equipo médico de la UCI modifique la CIII. Sin embargo, estas clasificaciones erróneas permiten que el equipo médico de la UCI tome medidas de *ABG* con más frecuencia, con el fin de garantizar la seguridad del paciente, incrementando el número de punciones. Aunque no se incurre en riesgos clínicos, esta no es la situación ideal, debido a que uno de los principales objetivos de utilizar el SMCG en la UCI es reducir el número de punciones en el paciente.
- Finalmente, y al igual que en el capítulo anterior, es importante resaltar que la investigación realizada solamente se enfocó en un SMCG específico, el dispositivo mínimamente invasivo Guardian[®] RTCGMS. Por lo tanto, se requieren nuevos estudios para determinar si es posible extrapolar las conclusiones obtenidas a otras tecnologías de sensores de glucemia, especialmente a los sensores basados en métodos intravenosos.

Es importante mencionar que en el desarrollo de la metodología que se propone en este capítulo, se pensó inicialmente en considerar las mismas variables originales utilizadas en el Capítulo 5, es decir, *ISIG*, *G_{RTCGMS}*, *Insulin*, *Temperature* y *Sepsis*. Sin embargo, resultados de experimentos preliminares considerando las variables *Temperature* y *Sepsis* para constituir el vector de características *X*, permitieron llegar a las mismas conclusiones que las obtenidas cuando no se consideran estas variables. Dado que el aporte de estas variables no contribuye a mejorar el rendimiento de las SVMs y con el objetivo de simplificar al mínimo su implementación, se descartó el uso de *Temperature* y *Sepsis*. Sin embargo, la variable *Sepsis* se utilizó para agrupar los diferentes pacientes de acuerdo a su estado séptico y analizar en cada una de estas cohortes de pacientes como es el comportamiento de la metodología que aquí se propone.

6.6. Resumen

En este capítulo, se ha estudiado el uso de SVMs para detectar medidas de glucosa correctas e incorrectas realizadas por el Guardian[®] RTCGMS. Al abordar este problema, se debe tener en cuenta que son escasas las glucemias erróneas, las cuáles constituyen la clase minoritaria y, por tanto, el conjunto de datos a clasificar está desbalanceado. En conjuntos de datos desequilibrados, la aplicación de SVMs convencionales favorece la clase mayoritaria y exhibe un bajo rendimiento en la clase minoritaria. Por esta razón, en este trabajo se aplicaron dos estrategias de post-procesamiento de SVMs para clasificar conjuntos de datos desbalanceados, empleando la información proporcionada por el propio SMCG, e incorporando variables acerca de la condición clínica de los pacientes. En el primer enfoque, llamado SVM-1, se ajustó el umbral de la función

decisión de la SVM en función de mejorar la G_{media} entre *sensibilidad* y *especificidad*. En el segundo enfoque, llamado SVM-2, la prioridad fue mejorar la *sensibilidad* ajustando el umbral de la función decisión de la SVM.

Los resultados experimentales de la metodología propuesta para detectar medidas de glucosa correctas e incorrectas en un SMCG implementando estas estrategias de post-procesamiento indicaron que la SVM-1 podía ser una buena alternativa para mejorar la fiabilidad en los SMCGs, debido a que permitió obtener un rendimiento más equilibrado entre *sensibilidad* y *especificidad*. Respecto a la estrategia de post-procesamiento SVM-2, aunque presentó altos valores de *sensibilidad*, su principal inconveniente fue que generó un elevado número de falsas alarmas, lo que afectó significativamente los valores de *especificidad*.

Capítulo 7

Conclusiones y Trabajo Futuro

Esta tesis se concluye señalando los aportes presentados en los capítulos anteriores. Además, se discuten algunas líneas abiertas para desarrollar posibles trabajos futuros sobre la investigación aquí presentada.

7.1. Contribuciones

De acuerdo con los desafíos que existen actualmente en la tecnología de monitorización continua de glucosa (MCG), el trabajo de investigación de esta tesis doctoral se enfocó en dos de las líneas de investigación abiertas mencionadas en el Capítulo 1: precisión de los algoritmos de calibración y fiabilidad de los sistemas de monitorización continua de glucosa (SMCGs) (-del inglés *Continuous Glucose Monitoring Systems (CGMSs)*-). En consecuencia, la principal contribución de esta tesis fue el desarrollo de nuevos enfoques metodológicos para mejorar la precisión y fiabilidad de los SMCGs. Específicamente, las contribuciones de esta tesis son:

1. Se diseñó un algoritmo retrospectivo de estimación de glucosa en plasma (-del inglés *Blood Glucose (BG)*-), cuyo principio de funcionamiento son las técnicas de regresión lineal. Para calcular la concentración de BG en un instante determinado, el algoritmo retrospectivo propuesto para el SMCG (ARP-SMCG) utiliza datos previos y posteriores al instante de tiempo en que se toma la muestra de *Valid ISIG*, aplicando ventanas temporales deslizantes de longitud variable. Al igual que el algoritmo del CGMS[®] Gold[™] (A-Gold[™]), este algoritmo utiliza las calibraciones y la señal de corriente *Valid ISIG* del sensor para estimar BG. Sin embargo, en la estimación de BG, el ARP-SMCG utiliza un mayor número de valores en el *Offset*, si se compara con los valores que considera el A-Gold[™] en este mismo factor. Por otra parte, el ARP-SMCG garantiza que el *Offset* seleccionado sea un *Offset* variable, lo que permite minimizar el error. Además, se caracteriza por utilizar criterios de ajuste para evitar cambios bruscos en la variación del *Offset* entre muestras adyacentes.

La principal ventaja del ARP-SMCG es que tiene mejor capacidad en el ajuste de las estimaciones de BG que el A-Gold[™], lo que permite mejorar la precisión, minimizando la diferencia absoluta relativa (-del inglés *Relative Absolute Difference (RAD)*-) con respecto al *gold standard*, particularmente en el rango hipoglucémico, como se demuestra en los resultados que se obtuvieron al evaluar su rendimiento con datos clínicos de pacientes con diabetes mellitus tipo 1 (DM-1). Sin embargo, su principal limitación es que al requerir de datos previos y posteriores al instante de tiempo en que se toma la muestra del sensor para obtener la estimación de BG, no

puede ser implementado en tiempo real (-del inglés *Real-Time (RT)*). Por otra parte, también puede ser considerado un algoritmo “estático”, porque no considera la dinámica que existe entre el plasma y el fluido intersticial (-del inglés *Interstitial Fluid (ISF)*-).

2. Se diseñó un algoritmo de estimación de BG para funcionamiento en RT, cuyo principio de funcionamiento son los modelos autorregresivos (-del inglés *AutoRegressives (ARs)*-). En particular, se obtuvo un modelo AR Box-Jenkins (-del inglés *Box-Jenkins (BJ)*-) de tercer orden, el cuál modela la relación dinámica entre BG y la glucosa intersticial (-del inglés *Interstitial Glucose (IG)*-), desde la señal de corriente del sensor insertado en el ISF y las muestras de BG consideradas como referencia. Este modelo se incorpora a un algoritmo de estimación de BG, el algoritmo basado en modelos ARs para el SMCG (ABMAR-SMCG), el cuál utiliza las calibraciones y la señal de corriente *ISIG* del sensor para estimar BG. Las principales ventajas del ABMAR-SMCG son que considera las dinámicas de la glucosa entre el plasma y el ISF, y que satisface el requerimiento de poder implementarse en RT. El ABMAR-SMCG puede implementarse en RT, debido a que para calcular la concentración de BG en un instante determinado, el modelo AR sólo utiliza la información de la corriente y la glucosa de dos y tres muestras previas al instante de tiempo actual, respectivamente. En este algoritmo, cuando se introduce un nuevo punto de calibración por parte del paciente, no sólo se ajusta la estimación actual a la misma glucemia de la calibración, además, las determinaciones de glucosa en los dos instantes anteriores se hacen iguales a esta glucemia, a fin de corregir las estimaciones del modelo. De hecho, su limitación se encuentra en que requiere incorporar puntos de calibración para corregir los niveles de BG estimados por el modelo poblacional, tal como se hace en los algoritmos de calibración que se utilizan actualmente en los SMCGs. Si no se considera esta corrección, los valores de glucosa después de introducir la calibración pueden no ser precisos y presentar drásticas diferencias con respecto al *gold standard*.

Es importante destacar que el ABMAR-SMCG logra minimizar el valor del *RAD* de sus estimaciones de BG con respecto al *gold standard*, en todos los rangos glucémicos, si se compara con el error reportado por el A-GoldTM. En particular, es más relevante que minimiza sustancialmente el valor del *RAD* en el rango hipoglucémico. Los resultados de evaluar el rendimiento del ABMAR-SMCG mediante diferentes criterios de precisión clínicos y analíticos, confirman este resultado. De igual forma, fue evidente la capacidad del ABMAR-SMCG para estimar correctamente episodios hipoglucémicos e hiperglucémicos omitidos por el A-GoldTM, lo cuál es extremadamente importante en el contexto de la DM-1.

3. Se realizó un estudio para evaluar la precisión de un SMCG en RT en una unidad de cuidados intensivos (UCI) polivalente. La precisión fue significativamente mejor en los pacientes en shock séptico en comparación con las otras cohortes de pacientes. Este resultado puede ser causado por la alteración de la permeabilidad capilar, una característica de la condición fisiopatológica de los pacientes en shock séptico, que se traduce en un aumento de la difusión de la glucosa entre el plasma y el ISF. Realizar un buen control glucémico es vital en este tipo de pacientes, puesto que en ellos es mayor el riesgo de presentar una disfunción de diferentes órganos por hipoperfusión, con la posibilidad de producirse un síndrome de fracaso multiorgánico. Sin embargo, para confirmar los beneficios potenciales de la MCG en los pacientes en shock séptico, se requiere de investigaciones con un tamaño muestral más grande, y en las cuáles se consideren otros factores fisiopatológicos que podrían influir en la precisión de los SMCGs.

4. Se diseñó una metodología mediante Análisis de Componentes Principales (-del inglés *Principal Component Analysis (PCA)*-) en combinación con Razonamiento Basado en Casos (-del

inglés *Case-Based Reasoning (CBR)*-), que a partir de información del propio SMCG en RT y de algunas variables relacionadas con la condición clínica del paciente, detecta si las estimaciones de BG realizadas por este dispositivo son terapéuticamente correctas (TC) e incorrectas (TI) (un nuevo esquema binario de clasificación, definido a partir de un criterio de precisión clínico), sin requerir de una medida de *gold standard*. Los resultados experimentales mostraron que este enfoque metodológico, puede ser una potencial herramienta para desarrollar un sistema de detección de fallos que permita auto-monitorizar los SMCGs. La metodología PCA-CBR mostró un desempeño más adecuado que el logrado por clasificadores lineales de uso común. Sin embargo, una de sus principales limitaciones está en su capacidad para generalizar desde un conjunto de entrenamiento a otros conjuntos de validación, no logrando en todas las bases de casos (BCs) un rendimiento balanceado y aceptable desde el punto de vista clínico para los pacientes evaluados.

5. Se diseñó una metodología basándose en Máquinas de Soporte Vectorial (-del inglés *Support Vector Machines (SVMs)*-) no convencionales capaces de trabajar con conjuntos de datos desequilibrados (gracias a dos estrategias de post-procesamiento para ajustar el sesgo en función de la G_{media} (SVM-1) y la *sensibilidad* (SVM-2)), que a partir de información suministrada por el propio SMCG en RT (corriente y nivel de glucosa) y la dosis de insulina endovenosa, detectan si las estimaciones de BG realizadas por este dispositivo son correctas o incorrectas (un nuevo esquema binario de clasificación, definido tomando como base un criterio de precisión analítico), sin requerir de una medida de *gold standard*. La SVM-1 puede ser una buena alternativa para desarrollar un sistema de detección de fallos en los SMCGs, debido a que reportó un rendimiento más equilibrado entre sensibilidad y especificidad. Este resultado es muy importante, puesto que un sistema de este tipo permitiría dotar a los SMCGs con capacidades de auto-monitorización para detectar mediciones erróneas, así como proporcionar un diagnóstico al personal médico que sea de utilidad para la toma de decisiones clínicas.

El rendimiento de la SVM-1 puede considerarse bueno para las cohortes de pacientes en shock séptico o con sepsis, no así para los pacientes no sépticos. Esta mayor fiabilidad, en la cohorte de pacientes en shock séptico, estaría influenciada por la alteración de la permeabilidad capilar, la cuál posibilitaría que la corriente que mide el sensor subcutáneo (SC) y en consecuencia, la concentración de BG que proporciona el SMCG, fueran más fiables. Por otra parte, aunque la SVM-2 facilita la detección de medidas incorrectas y por tanto, presenta altos valores de sensibilidad, su principal inconveniente es que genera un elevado número de falsas alarmas, lo que afecta significativamente la especificidad, y no la hace una opción adecuada para la detección de errores en la MCG.

7.2. Trabajo futuro

En general, las conclusiones obtenidas en este trabajo son indicadores de cómo puede ser el comportamiento de los enfoques metodológicos propuestos. Esto es debido a los pequeños tamaños muestrales de los conjuntos de datos con los que se validan estas metodologías. Por ello, se requiere validar las metodologías aquí descritas utilizando conjuntos de datos con un tamaño poblacional mayor y que además, reúnan las siguientes características:

- Buena calidad de los datos.
- Tener suficientes y frecuentes determinaciones de BG, obtenidas mediante pruebas de laboratorio o con un analizador de glucosa de laboratorio, por ejemplo el YSI (Yellow

Springs Instrument, Yellow Springs, OH).

- Incluir suficientes muestras de eventos hipoglucémicos.
- Disponer de la señal de corriente eléctrica del sensor SC.

Además de las contribuciones presentadas, se debe extender esta investigación abordando las siguientes líneas de trabajo abiertas:

- Con respecto al ARP-SMCG:
 - En el conjunto de datos clínicos obtenidos de enfermos con DM-1 de la Unidad de Endocrinología del Hospital Josep Trueta:
 - Evaluar su precisión usando criterios analíticos empleando como método de referencia la concentración de glucosa en sangre venosa (*VBG*) sin interpolar, considerando tanto la población global como el estado glucémico y la tasa de cambio de la glucemia.
 - Evaluar su precisión usando criterios clínicos utilizando como método de referencia la concentración de *VBG* sin interpolar, considerando tanto la población global como el estado glucémico y la tasa de cambio de la glucemia. Comparar con los resultados reportados en la Tabla 3.4.
- Con respecto a los dos nuevos algoritmos de estimación de BG:
 - En conjuntos de datos con mayor tamaño muestral, representativos de pacientes con DM-1 y de pacientes críticos.
 - Validar el ARP-SMCG. En particular, analizar su precisión en la estimación de episodios hipoglucémicos.
 - Validar el ABMAR-SMCG. En particular, analizar su precisión en la estimación de episodios hipoglucémicos.
 - En los pacientes críticos, evaluar la precisión del ABMAR-SMCG, enfocándose en determinar si el estado séptico del paciente tiene alguna influencia en la precisión de este algoritmo. Se debe repetir el entrenamiento del modelo poblacional AR para que sea representativo del comportamiento de la nueva cohorte de pacientes.
- En relación con los enfoques metodológicos propuestos para detección de fallos en los SMCGs:
 - En un conjunto de datos con un tamaño muestral mayor, representativo de pacientes de una UCI polivalente.
 - Definir nuevas BCs con un mayor tamaño muestral y validar el rendimiento de la metodología PCA-CBR en numerosos pacientes críticos no utilizados para la puesta a punto de la metodología. Para asegurar la independencia entre la calibración y la validación, para cada BC, el auto-escalado y posterior modelado PCA debe realizarse sólo con las observaciones que la constituyen.
 - Validar el rendimiento de la estrategia de post-procesamiento SVM-1, abordando esta investigación desde el punto de vista de la condición séptica de los pacientes. En particular, centrarse en validar el potencial rendimiento mostrado por esta metodología en la detección de glucemias erróneas en pacientes en shock séptico.

- Examinar el comportamiento de las dos metodologías en la detección de los episodios hipoglucémicos, debido a la importancia de la detección correcta y oportuna de estos eventos. En este trabajo los episodios hipoglucémicos fueron muy infrecuentes y no es posible extraer conclusiones acerca del rendimiento de las metodologías propuestas en este rango.
- Analizar el rendimiento de las dos metodologías, en función de otros factores fisiopatológicos que podrían influir en la precisión y fiabilidad de los SMCGs en pacientes críticos, tales como las dosis de los fármacos inotrópicos, el grado de edema o el nivel de la perfusión tisular, determinado este último midiendo la concentración de lactato arterial. Otros factores que podrían ser de interés, son el uso de glucocorticoides o el balance de fluidos.
- En un conjunto de datos con un tamaño muestral mayor, representativo de pacientes con DM-1.
 - Adaptar la metodología PCA-CBR a las características de la nueva cohorte de pacientes, definiendo nuevos conjuntos de entrenamiento (nuevas BCs) para representar el comportamiento de la nueva cohorte y validar su rendimiento. Para asegurar la independencia entre la calibración y la validación, para cada BC, el auto-escalado y posterior modelado PCA debe realizarse sólo con las observaciones que la constituyen.
 - Adaptar la SVM-1 a las características de la nueva cohorte de pacientes, repetir el entrenamiento y validar su rendimiento.
 - Enfocarse en el comportamiento de las dos metodologías en la detección de eventos hipoglucémicos.
- En cuanto a la integración de las técnicas para mejorar la precisión y fiabilidad de los SMCGs, se debería desarrollar una aplicación de software que integre las estrategias de detección de fallos aquí propuestas en el ABMAR-SMCG o en cualquier algoritmo propuesto en la literatura. En este caso, la detección de una secuencia de mediciones incorrectas consecutivas puede desencadenar la exigencia de introducir un nuevo punto de calibración por parte del paciente o del personal sanitario.
- Finalmente, una futura línea de trabajo en este campo podría ser abordar el problema de la detección de fallos en los SMCGs como un problema de multi-clasificación, que no sólo indique si una concentración de glucosa es correcta o incorrecta, sino que también permita la inclusión de una tercera clase para clasificar las mediciones dudosas. Esta propuesta se debe a un análisis muy preliminar recientemente realizado por la autora de esta tesis en los resultados obtenidos con la SVM-1. En este análisis se observó que algunas mediciones clasificadas como falsos positivos o falsos negativos se concentran muy cerca de la frontera entre la zonas válida e inválida, definidas por los criterios ISO. Por lo tanto, se plantea la hipótesis de que las concentraciones de glucosa que caen justo en estos umbrales confunden al clasificador, y generan errores de clasificación. Basándose en esta observación, los estudios futuros podrían abordar el problema de la detección de fallos en los SMCGs como un problema de multi-clasificación. De hecho, dentro de la clase incorrecta también se podrían definir sub-clases, de acuerdo al nivel de riesgo clínico que la medida incorrecta represente para el paciente.

Referencias

- The policy puzzle: is europe making progress?, 2008. URL <http://www.fend.org/>. 22
- Abbot, 2013. URL <https://www.abbottdiabetescare.com>. 34
- Dexcom, 2013. URL <http://www.dexcom.com/>. 34
- Diabetes net web, 2013. URL <http://www.diabetesnet.com>. 26
- Diabetes research in children network (direcnet), 2013. URL <http://direcnet.jaeb.org>. 54
- Juvenile diabetes research foundation, 2013. URL <http://jdrf.org>. 26
- Medline, 2013a. URL <http://www.nlm.nih.gov/medlineplus>. 7, 56, 104, 114
- Medtronic diabetes, 2013b. URL <http://www.medtronicdiabetes.com/products>. 33, 34, 36, 37
- Menarini, 2013. URL <http://www.menarini.com>. 33, 34
- A. Aamodt and E. Plaza. Case-based reasoning: foundational issues, methodological variations, and system approaches. *AI Communications*, 7:39–59, 1994. 128
- R. Akbani, S. Kwek, and N. Japkowicz. Applying support vector machines to imbalanced datasets. In *Proc. 15th Eur. Conf. Mach. Learn.*, pages 39–50, 2004. 154, 155
- American Diabetes Association. Diagnosis and classification of diabetes mellitus. *Diabetes Care*, 33 (suppl 1):S62–S69, 2010. 22
- T. Bailey, H. Zisser, and A. Chang. New features and performance of a next-generation seven-day continuous glucose monitoring system with short lag time. *Diabetes Technol Ther*, 11: 749–755, 2009. 38, 93
- F. Barceló-Rico. *Multimodel Approaches for Plasma Glucose Estimation in Continuous Glucose Monitoring - Development of New Calibration Algorithms*. PhD thesis, Universitat Politècnica de Valencia, 2012. 26
- F. Barceló-Rico, J. Bondia, J. L. Díez, and P. Rossetti. A multiple local models approach to accuracy improvement in continuous glucose monitoring. *Diabetes Technol Ther*, 14:74–82, 2012. 10, 42
- F. Barceló-Rico, J. L. Díez, P. Rossetti, J. Vehi, and J. Bondia. Adaptive calibration algorithm for plasma glucose estimation in continuous glucose monitoring. *IEEE Journal of Biomedical and Health Informatics*, In press., 2013. 10, 43

- K. Becker. *Principles and Practices of Endocrinology and Metabolism*. Lippincott Williams and Wilkins Publishers, 2001. 25
- W. B. Bequette. Continuous glucose monitoring: real-time algorithms for calibration, filtering, and alarms. *J Diabetes Sci Technol*, 4:404–418, 2010. 9, 10, 11, 12, 37, 39, 44, 117
- J. M. Bland and D. G. Altman. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*, 1:307–310, 1986. 48
- B. Bode, K. Gross, N. Rikalo, S. Schwartz, T. Wahl, C. Page, T. Gross, and J. Mastrototaro. Alarms based on real-time sensor glucose values alert patients to hypo- and hyperglycemia: the guardian continuous monitoring system. *Diabetes Technol Ther*, 6:105–113, 2004. 64, 90
- J. Bondia, C. Tarin, W. Garcia-Gabin, E. Esteve, J. M. Fernández-Real, W. Ricart, and J. Vehí. Using support vector machines to detect therapeutically incorrect measurements by the mimimed cgms. *J Diabetes Sci Technol*, 2:622–629, 2008. 12, 44, 45, 67, 117, 149, 150
- J. Bondia, J. Vehí, CC. Palerm, and P. Herrero. El páncreas artificial: control automático de infusión de insulina en diabetes mellitus tipo 1. *Revista Iberoamericana de Automática e Informática Industrial*, 7(2):5–20, 2010. 10, 27, 29, 32, 38, 55, 72, 73
- R. C. Bone, R. A. Balk, F. B. Cerra, R. P. Dellinger, A. M. Fein, W. A. Knaus, R. M. Schein, and W. J. Sibbald. Definitions for sepsis and organ failure and guidelines for the use of innovative therapies in sepsis. the accp/sccm consensus conference committee. american college of chest physicians/society of critical care medicine. *Chest*, 101:1644–1655, 1992. 96, 98, 104, 112, 113
- SA. Boren and WL. Clarke. Analytical and clinical performance of blood glucose monitors. *J Diabetes Sci Technol*, 4:84–97, 2010. 46
- A.T. Boye, U.Q. Kristiansen, M. Billinger, O.F. Do Nascimento, and D. Farina. Identification of movement-related cortical potentials with optimized spatial filtering and principal component analysis. *Biomedical Signal Processing and Control*, 3:300–304, 2008. 13
- M.S. Boyne, D.M. Silver, J. Kaplan, and C.D. Saudek. Timing of changes in interstitial and venous blood glucose measured with a continuous subcutaneous glucose sensor. *Diabetes*, 52:2790–2794, 2003. 38
- E. Braunwald. *Harrison. Principios de medicina interna*. McGraw-Hill, 2002. 98
- T. Bremer and D.A. Gough. Is blood glucose predictable from previous values? a solicitation for data. *Diabetes*, 48(3):445–51, 1999. 44
- B. C. Bridges, C. M. Preissig, K. O. Maher, and M. R. Rigby. Continuous glucose monitors prove highly accurate in critically ill children. *Crit Care*, 14:R176, 2010. 96, 110, 114
- F. Brunicardi. *Schwartz Principios de Cirugía*. McGraw-Hill, 2010. 103
- F. M. Brunkhorst, C. Engel, F. Bloos, A. Meier-Hellmann, M. Ragaller, N. Weiler, O. Moe-rer, M. Gruendling, M. Oppert, S. Grond, D. Olthoff, U. Jaschinski, S. John, R. Rossaint, T. Welte, M. Schaefer, P. Kern, E. Kuhnt, M. Kiehntopf, C. Hartog, C. Natanson, M. Loeffler, K. Reinhart, and German Competence Network Sepsis (SepNet). Intensive insuline therapy and pentastarch resuscitation in severe sepsis. *N Engl J Med*, 358:125–139, 2008. 96, 115

- R. Brunner, R. Kitzberger, W. Miehsler, H. Herkner, C. Madl, and U. Holzinger. Accuracy and reliability of a subcutaneous continuous glucose-monitoring system in critically ill patients. *Crit Care Med*, 39:659–664, 2011. 96, 114
- B. Buckingham. Dawn of real-time continuous glucose sensing. *Diabetes Technol Ther*, 5, 2003. 55
- K. J. Buhling, B. Kurzidim, C. Wolf, K. Wohlfarth, M. Mahmoudi, C. Wascher, G. Siebert, and J. W. Dudenhausen. Introductory experience with the continuous glucose monitoring system (cgms; medtronic minimed) in detecting hyperglycemia by comparing the self-monitoring of blood glucose (smbg) in non-pregnant women and in pregnant women with impaired glucose tolerance and gestational diabetes. *Exp Clin Endocrinol Diabetes*, 112, 2004. 34
- J. Camacho, J. Picó, and A. Ferrer. Data understanding with pca: Structural and variance information plots. *Chemometrics and Intelligent Laboratory Systems*, 100:48–56, 2010. 122
- S. E. Capes, D. Hunt, K. Malmberg, and H. C. Gerstein. Stress hyperglycaemia and increased risk of death after myocardial infarction in patients with and without diabetes: a systematic overview. *Lancet*, 355:773–778, 2000. 95
- S. E. Capes, D. Hunt, K. Malmberg, P. Pathak, and H. C. Gerstein. Stress hyperglycemia and prognosis of stroke in nondiabetic and diabetic patients: a systematic overview. *Stroke*, 32: 2426–2432, 2001. 95
- J. Casado and A. Blanco. Procalcitonina: un nuevo marcador de infección bacteriana. *An Esp Pediatr*, 54:69–73, 2001. 104
- C. L. Castro, M. A. Carvalho, and A. P. Braga. An improved algorithm for SVMs classification of imbalanced data sets. In *Proc. 11th Int. Conf. Eng. Appl. Neural Netw.*, pages 108–118, London, UK, August 27–29 2009. Berlin: Springer. ISBN 978-3-642-03968-3. 155
- E. Cengiz and W. V. Tamborlane. A tale of two compartments: interstitial versus blood glucose monitoring. *Diabetes Technol Ther*, 11: Suppl 1:S11–S16, 2009. 10, 37
- J. G. Chase, C. G. Pretty, L. Pfeifer, G. M. Shaw, J. C. Preiser, A. J. Le Compte, J. Lin, D. Hewett, K. T. Moorhead, and T. Desaive. Organ failure and tight glycemic control in the sprint study. *Crit Care*, 14:R154, 2010. 96
- J.G. Chase, C.E. Hann, M. Jackson, J. Lin, T. Lotz, X.W. Wong, and G.M. Shaw. Integral-based filtering of continuous glucose sensor measurements for glycaemic control in critical care. *Comput. Methods Programs Biomed*, 82:238–247, 2006. 43
- N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer. SMOTE: Synthetic minority over-sampling technique. *J. Artif. Intell. Res.*, 16:321–357, 2002. 154
- C. Choleau, P. Dokladal, J.C. Klein, W.K. Ward, G.S. Wilson, and G. Reach. Prevention of hypoglycemia using risk assessment with a continuous glucose monitoring system. *Diabetes*, 51:3263–73, 2002a. 40
- C. Choleau, J.C. Klein, G. Reach, B. Aussedat, V. Demaria-Pesce, G.S. Wilson, R. Gifford, and W.K. Ward. Calibration of a subcutaneous amperometric glucose sensor. part 1. effect of measurement uncertainties on the determination of sensor sensitivity and background current. *Biosens Bioelectron.*, 17(8):641–6, 2002b. 44

- C. Choleau, J.C. Klein, G. Reach, B. Aussedat, V. Demaria-Pesce, G.S. Wilson, R. Gifford, and W.K. Ward. Calibration of a subcutaneous amperometric glucose sensor implanted for 7 days in diabetic patients. part 2. superiority of the one-point calibration method. *Biosens Bioelectron*, 17(8):647–54, 2002c. 44
- W. Clarke and B. Kovatchev. Continuous glucose sensors: Continuing questions about clinical accuracy. *J Diabetes Sci Technol*, 1:669–675, 2007. 47
- W. Clarke and B. Kovatchev. Statistical tools to analyze continuous glucose monitor data. *Diabetes Technol Ther*, 11 (Suppl 1):S45–S54, 2009. 46, 47
- W L. Clarke. The original clarke error grid analysis (ega). *Diabetes Technol Ther*, 7:776–779, 2005. 49
- W. L. Clarke, D. Cox, L. A. Gonder-Frederick, W. Carter, and S. L. Pohl. Evaluating clinical accuracy of systems for self-monitoring of blood glucose. *Diabetes Care*, 10:622–628, 1987. 49
- W. L. Clarke, S. Anderson, L. Farhy, M. Breton, L. Gonder-Frederick, D. Cox, and B. Kovatchev. Evaluating the clinical accuracy of two continuous glucose sensors using continuous glucose-error grid analysis. *Diabetes Care*, 28:2412–2417, 2005. 9, 29, 36, 55
- G. Cohen, M. Hilario, H. Sax, S. Hugonnet, and A. Geissbuhler. Learning from imbalanced data in surveillance of nosocomial infection. *Artif. Intell. Med.*, 37:7–18, 2006. ISSN 0933-3657. 155
- B. Collier, L. A. Dossett, A. K. May, and J. J. Diaz. Glucose control and the inflammatory response. *Nutr Clin Pract*, 23:3–15, 2008. 95, 113
- A. M. Corstjens, J. J. Ligtenberg, I. C. van der Horst, R. Spanjersberg, J. S. Lind, J. E. Tulleken, J. H. Meertens, and J. G. Zijlstra. Accuracy and feasibility of point-of-care and continuous blood glucose analysis in critically ill icu patients. *Crit Care Forum*, 10:R135, 2006. 96, 110, 112
- R Courant and D. Hilbert. *Methods of mathematical physics*. Interscience, 1953. 153
- N. Cristianini and J. Shawe-Taylor. *An introduction to Support Vector Machines and other kernel-based learning methods*. Cambridge University, 2000. ISBN 0 521 78019 5. 13, 149, 151, 153
- C. Crone. Capillary permeability to small solutes. In *In Handbook of Physiology: The Cardiovascular system, Microcirculation*, volume IV. American Physiology Society, Bethesda, MD, USA, 1984. 38
- R.J. Davey, C. Low, T.W. Jones, and P.A. Fournier. Contribution of an intrinsic lag of continuous glucose monitoring systems to differences in measured and actual glucose concentrations changing at variable rates in vitro. *J Diabetes Sci Technol*, 4(6):1393–1399, 2010. 10, 43
- DCCT. The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. *The New England Journal of Medicine*, 329:977–986, 1993. 7, 25
- C. De Block, B. Manuel-Y-Keenoy, L. Van Gaal, and P. Rogiers. Intensive insulin therapy in the intensive care unit: assessment by continuous glucose monitoring. *Diabetes Care*, 29:1750–1756, 2006. 96, 110, 112, 120

- C. De Block, J. Vertommen, B. Manuel-y Keenoy, and L. Van Gaal. Minimally-invasive and non-invasive continuous glucose monitoring systems: indications, advantages, limitations and clinical aspects. *Curr Diabetes Rev.*, 29:159–168, 2008. 30, 31
- D. Deiss, J. Bolinder, J. Riveline, T. Battelino, E. Bosi, N. Tubiana-Rufi, D. Kerr, and M. Phillip. Improved glycemic control in poorly controlled patients with type 1 diabetes using real-time continuous glucose monitoring. *Diabetes Care*, 29:2730–2732, 2006. 29, 37
- R. P. Dellinger, M. M. Levy, J. M. Carlet, J. Bion, M. M. Parker, R. Jaeschke, K. Reinhart, D. C. Angus, C. Brun-Buisson, R. Beale, T. Calandra, J. F. Dhainaut, H. Gerlach, M. Harvey, J. J. Marini, J. Marshall, M. Ranieri, G. Ramsay, J. Sevransky, B. T. Thompson, S. Townsend, J. S. Vender, J. L. Zimmerman, J. L. Vincent, International Surviving Sepsis Campaign Guidelines Committee, American Association of Critical-Care Nurses, American College of Chest Physicians, American College of Emergency Physicians, Canadian Critical Care Society, European Society of Clinical Microbiology, Infectious Diseases, European Society of Intensive Care Medicine, European Respiratory Society, International Sepsis Forum, Japanese Association for Acute Medicine, Japanese Society of Intensive Care Medicine, Society of Critical Care Medicine, Society of Hospital Medicine, Surgical Infection Society, World Federation of Societies of Intensive, and Critical Care Medicine. International guidelines for management of severe sepsis and septic shock. *Crit Care Med*, 36:296–327, 2008. 113
- DirecNet. The accuracy of the cgms in children with type 1 diabetes: Results of the diabetes research in children network (direcnet) accuracy study. *Diabetes Technol Ther*, 5:781–789, 2003. 93
- DirecNet. Accuracy of the modified continuous glucose monitoring system (cgms) sensor in an outpatient setting: Results from a diabetes research in children network (direcnet) study. *Diabetes Technol Ther*, 7:109–114, 2005. 34, 72
- DirecNet. The accuracy of the guardian rt continuous glucose monitor in children with type 1 diabetes. *Diabetes Technol Ther*, 10:266–272, 2008. 112
- Direcnet, B.A. Buckingham, C. Kollman, R. Beck, A. Kalajian, R. Fiallo-Scharer, M. J. Tansey, L. A. Fox, D. M. Wilson, S. A. Weinzimer, K. J. Ruedy, and W. V. Tamborlane. Evaluation of factors affecting cgms calibration. *Diabetes Technol Ther*, 8:318–325, 2006. 40, 55, 66, 72
- S. L. Ellis, T. Bookout, S. K. Garg, and K. E. Izuora. Use of continuous glucose monitoring to improve diabetes mellitus management. *Endocrinol Metab Clin North Am*, 36(Suppl 2): 47–68, 2008. 29
- M. Ellmerer, M. Haluzik, J. Blaha, J. Kremen, S. Svacina, W. Toller, J. Mader, L. Schaupp, J. Plank, and T. Pieber. Clinical evaluation of alternative-site glucose measurements in patients after major cardiac surgery. *Diabetes Care*, 29:1275–1281, 2006. 49
- M. Eren-Oruklu, A. Cinar, L. Quinn, and D. Smith. Estimation of future glucose concentrations with subject-specific recursive linear models. *Diabetes Technol Ther*, 11:243–253, 2009. 12, 44, 73
- M. Eren-Oruklu, A. Cinar, and L. Quinn. Hypoglycemia prediction with subject-specific recursive time-series models. *J Diabetes Sci Technol.*, 4(1):25–33, 2010. 11
- K. Esposito, F. Nappo, R. Marfella, G. Giugliano, F. Giugliano, M. Ciotola, L. Quagliari, A. Ceriello, and D. Giugliano. Inflammatory cytokine concentrations are acutely increased by hyperglycemia in humans: role of oxidative stress. *Circulation*, 106:2067–2072, 2002. 95, 113

- A. Facchinetti, G. Sparacino, and C. Cobelli. Enhanced accuracy of continuous glucose monitoring by online extended kalman filtering. *Diabetes Technol Ther*, 4:391–403, 2010a. 10, 41
- A. Facchinetti, G. Sparacino, and C. Cobelli. An online self-tuneable method to denoise cgm sensor data. *IEEE Trans Biomed Eng*, 57(3):634–641, 2010b. 11, 43
- A. Facchinetti, S. Del Favero, G. Sparacino, and C. Cobelli. Detecting failures of the glucose sensor-insulin pump system: improved overnight safety monitoring for type-1 diabetes. In *Proc. Engineering in Medicine and Biology Society, EMBC.*, pages 4947–4950, 2011. ISBN 978-0-7695-4109-9. 12, 44, 46
- B.J. Feldman and G.V. McGarraugh. Method of calibrating an analyte measurement device, and associated methods, devices and systems. U.S. Patent 7,299,082, November 20 2007. 34, 43, 44
- C.E. Ferrante do Amaral and B. Wolf. Current development in non-invasive glucose monitoring. *Med Eng Phys*, 30:541–549, 2008. 30, 31
- R. Fiallo-Scharer and Direcnet. Eight-point glucose testing versus the continuous glucose monitoring system in evaluation of glycemic control in type 1 diabetes. *J Clin Endocrinol Metab*, 90(6):3387, 2005. 36
- D. A. Finan, H. Zisser, L. Jovanovic, W. C. Bevier, and D. E. Seborg. Practical issues in the identification of empirical models from simulated type 1 diabetes data. *Diabetes Technol Ther*, 9:438–450, 2007. 73
- LM. Gandrud, D. Xing, C. Kollman, J. M. Block, B. Kunselman, D. M. Wilson, and B. A. Buckingham. The medtronic minimed gold continuous glucose monitoring system: An effective means to discover hypo- and hyperglycemia in children under 7 years of age. *Diabetes Technol Ther*, 9:307–316, 2007. 34, 36, 64, 72
- A. Gani, A. V. Gribok, S. Rajaraman, W. K. Ward, and J. Reifman. Predicting subcutaneous glucose concentration in humans: data-driven glucose modeling. *IEEE Trans Biomed Eng.*, 56:246–254, 2009. 11, 12, 44, 73, 81
- A. Gani, A. V. Gribok, S. Rajaraman, W. K. Ward, and J. Reifman. Universal glucose models for predicting subcutaneous glucose concentration in humans. *IEEE Trans Inf Technol Biomed.*, 14:157–265, 2010. 11, 12, 44, 73
- M. García-Jaramillo, R. Calm, J. Bondia, C. Tarín, and J. Vehí. Insulin dosage optimization based on prediction of postprandial glucose excursions under uncertain parameters and food intake. *Comput Methods Programs Biomed*, 105:61–69, 2012. 26
- M.A. Garcia-Jaramillo. *Prediction of postprandial blood glucose under intra-patient variability and uncertainty and its use in the design of insulin dosing strategies for type 1 diabetic patients.* PhD thesis, University of Girona, 2011. 24
- S. Garg and L. Jovanovic. Relationship of fasting and hourly blood glucose levels to hba1c values: safety, accuracy, and improvements in glucose profiles obtained using a 7-day continuous glucose sensor. *Diabetes Care*, 29:2644–2649, 2006. 90

- S. K. Garg, J. Smith, C. Beatson, B. Lopez-Baca, M. Voelmle, and P. A. Gottlieb. Comparison of accuracy and safety of the seven and the navigator continuous glucose monitoring systems. *Diabetes Technol Ther*, 11:65–72, 2009. 38, 77, 83, 106, 119
- SK. Garg, RO. Potts, NR. Ackerman, SJ. Fermi, JA. Tamada, and HP. Chase. Correlation of fingerstick blood glucose measurements with glucoWatch biographer glucose results in young subjects with type 1 diabetes. *Diabetes Care*, 22:1708–1714, 1999. 49
- S. Gebhart, M. Faupel, R. Fowler, C. Kapsner, D. Lincoln, V. McGee, J. Pasqua, L. Steed, M. Wangsness, F. Xu, and M. Vanstoy. Glucose sensing in transdermal body fluid collected under continuous vacuum pressure via micropores in the stratum corneum. *Diabetes Technol Ther*, 5:159–166, 2003. 30, 31
- P. Gemperline. *Practical Guide to Chemometrics*. CRC Press, second edition edition, 2006. 203
- B. H. Ginsberg. The current environment of cgm technologies. *J Diabetes Sci Technol*, 1:117–121, 2007. 9, 29
- P. A. Goldberg, M. D. Siegel, R. R. Russell, R. S. Sherwin, J. I. Halickman, D. A. Cooper, J. D. Dziura, and S. E. Inzucchi. Experience with the continuous glucose monitoring system in a medical intensive care unit. *Diabetes Technol Ther*, 6:339–347, 2004. 96, 102, 110
- D. E. Goldstein and Lorenz R. A. Malone J. I. Nathan D. Peterson C. M. & Sacks D. B. Little, R. R. Tests of glycemia in diabetes. *Diabetes Care*, 27:1761–1773, 2004. 9, 28
- L. Gonzalez-Abril. *Análisis Discriminante utilizando Máquinas Núcleos de Vectores Soporte. Función Núcleo Similitud*. PhD thesis, Universidad de Sevilla, 2002. 153
- L. Gonzalez-Abril, C. Angulo, F. Velasco, and A. Català. Dual unification of bi-class support vector machine formulations. *Pattern Recognition*, 39:1325–1332, 2006. 152
- L. Gonzalez-Abril, C. Angulo, F. Velasco, and J. A. Ortega. A note on the bias in SVMs for multiclassification. *IEEE Trans. Neural Netw.*, 19:723–725, 2008. 152, 156
- P.V. Jr. Goode, J.H. Brauker, and A.U. Kamath. System and methods for processing analyte sensor data. United States patent US 6,931,327 B2, Aug 16 2005. 43
- T. M. Gross, B. W. Bode, D. Einhorn, D. M. Kayne, J. H. Reed, N. H. White, and J. J. Mastrototaro. Performance evaluation of the minimed continuous glucose monitoring system during patient home use. *Diabetes Technol Ther*, 2:49–56, 2000. 33, 35, 36
- S. Guerra, A. Facchinetti, G. Sparacino, G. De Nicolao, and C. Cobelli. Enhancing the accuracy of subcutaneous glucose sensors: A real-time deconvolution-based approach. *IEEE Trans Biomed Eng*, 59(6):1658–1669, 2012. 42
- R. Hanas. Type 1 diabetes in children, adolescents and young adults. *Class publishing*, 2004. 23
- R. A. Harvey, Y. Wang, B. G. M. W. Percival, W. Bevier, D. A. Finan, H. Zisser, D. E. Seborg, L. Jovanovic, F. J. Doyle III, R.A. Harvey, and E. Dassau. Quest for the artificial pancreas: combining technology with treatment. *IEEE Engineering in Medicine and Biology Magazine*, 29:53–62, 2010. 26
- N. Haydemar, L. Gonzalez-Abril, and C. Angulo. A post-processing strategy for SVM learning from unbalanced data. In *Proc. 19th European Symposium on Artificial Neural Networks, ESANN*, pages 195–200, 2011. 155, 156, 157

- H. He and E. A. Garcia. Learning from imbalanced data. *IEEE Trans. Know. and Data Eng.*, 21:1263–1284, 2009. ISSN 1041-4347. 52, 53
- H. He and A. Ghodsi. Rare class classification by support vector machine. In *Proc. 20th Int. Conf. on Pattern Recogn.*, pages 548–551, Washington, DC, USA, 2010. IEEE Computer Society. ISBN 978-0-7695-4109-9. 154, 155, 156
- T. Hoi-Hansen, U. Pedersen-Bjergaard, and B. Thorsteinsson. Reproducibility and reliability of hypoglycaemic episodes recorded with continuous glucose monitoring system (cgms) in daily life. *Diabet Med.*, 22:858–862, 2005. 36
- U. Holzinger, J. Warszawska, R. Kitzberger, H. Herkner, P. G. Metnitz, and C. Madl. Impact of shock requiring norepinephrine on the accuracy and reliability of subcutaneous continuous glucose monitoring. *Intensive Care Med.*, 35:1383–1389, 2009. 96, 114
- H. R. Horton, J. D. Rawn, K. G. Scrimgeour, L. A. Moran, and M. D. Perry. *Principios de bioquímica*. Pearson. Prentice Hall, 2008. 102
- R. Hovorka. The future of continuous glucose monitoring: closed loop. *Current Diabetes Reviews*, 4(3):269–279, 2008. 11, 12, 44
- T. Imam, K. T. Ming, and J. Kamruzzaman. z-SVM: An SVM for improved classification of imbalanced data. In *Australian Conf. Artif. Intell.*, pages 264–273, 2006. 154, 155
- International Diabetes Federation. *Idf diabetes atlas 5th edition.*, 2011. 19, 21, 22
- ISO. International organization for standardization. *in vitro diagnostic test systems requirements for blood-glucose monitoring systems for self-testing in managing diabetes mellitus. iso 15197.*, 2003. 47
- B. Jacobs, K. Phan, L. Bertheau, G. Dogbey, F. Schwartz, and J. Shubrook. Continuous glucose monitoring system in a rural intensive care unit: a pilot study evaluating accuracy and acceptance. *J Diabetes Sci Technol*, 4:636–644, 2010. 96, 110, 112
- JDRF, W.V. Tamborlane, R.W. Beck, B.W. Bode, B. Buckingham, H.P. Chase, R. Clemons, R. Fiallo-Scharer, L.A. Fox, L.K. Gilliam, I.B. Hirsch, E.S. Huang, C. Kollman, A.J. Kowalski, L. Laffel, J.M. Lawrence, J. Lee, N. Mauras, M. O’Grady, K.J. Ruedy, M. Tansey, E. Tsalikian, S. Weinzimer, D. M. Wilson, H. Wolpert, T. Wysocki, and D. Xing. Continuous glucose monitoring and intensive treatment of type 1 diabetes. *N Engl J Med*, 359:1464–1476, 2008. 29
- I.T. Jolliffe. *Principal Component Analysis*. Springer series in statistics, 2002. 122
- A. Kamath, A. Mahalingam, and J. Brauker. Analysis of time lags and other sources of error of the dexcom seven continuous glucose monitor. *Diabetes Technol Ther*, 11:689–695, 2009. 38
- P. Kang and S. Cho. EUS SVMs: Ensemble of Under-Sampled SVMs for data imbalance problems. volume 4232 of *Lect. Not. Comp. Sc.*, pages 837–846. Springer Berlin Heidelberg, 2006. ISBN 978-3-540-46479-2. 155
- D.B. Keenan, J.J. Mastrototaro, G. Voskanyan, and G.M. Steil. Delays in minimally invasive continuous glucose monitoring devices: a review of current technology. *J Diabetes Sci Technol.*, 3(5):1207–1214, 2009. 43

- A. Kerssen, H. W. De Valk, and G. H. Visser. The continuous glucose monitoring system during pregnancy of women with type 1 diabetes mellitus: Accuracy assessment. *Diabetes Technol Ther*, 6:645–651, 2004. 34
- A. Kerssen, H. W. De Valk, and G. H. Visser. Validation of the continuous glucose monitoring system (cgms) by the use of two cgms simultaneously in pregnant women with type 1 diabetes mellitus. *Diabetes Technol Ther*, 7:699–706, 2005. 36
- C. King, S. M. Anderson, M. Breton, W. L. Clarke, and B. P. Kovatchev. Modeling of calibration effectiveness and blood-to-interstitial glucose dynamics as potential confounders of the accuracy of continuous glucose sensors during hyperinsulinemic clamp. *J Diabetes Sci Technol*, 1:317–322, 2007. 10, 38, 39, 40, 73
- D. C. Klonoff. Continuous glucose monitoring roadmap for 21st century diabetes therapy. *Diabetes Care*, 28:1231–1239, 2005. 29
- DC. Klonoff. The need for clinical accuracy guidelines for blood glucose monitors. *J Diabetes Sci Technol*, 6:1–4, 2012. 48, 49
- W. A. Knaus, E. A. Draper, D. P. Wagner, and J. L. Zimmerman. Apache ii: a severity of disease classification system. *Crit Care Med*, 13:818–829, 1985. 97
- E.J. Knobbe and B. Buckingham. The extended kalman filter for continuous glucose monitoring. *Diabetes Technol Ther*, 7:15–27, 2005. 10, 41, 42, 43
- I. Kollár. On frequency domain identification of linear system. *IEEE Trans. on Instrumentation and Measurement*, 42:2–6, 1993. 73
- C. Kollman, D. M. Wilson, T. Wysocki, W. V. Tamborlane, R. W. Beck, and Diabetes Research in Children Network Study Group. Limitations of statistical measures of error in assessing the accuracy of continuous glucose sensors. *Diabetes Technol Ther*, 7:665–672, 2005. 49
- B. P. Kovatchev and W. L. Clarke. Continuous glucose monitoring (cgm) reduces risks for hypo- and hyperglycemia and glucose variability in diabetes. *Diabetes*, 56:(Suppl. 1):A23, 2007. 29
- B. P. Kovatchev, L. A. Gonder-Frederick, D. J. Cox, and W. L. Clarke. Evaluating the accuracy of continuous glucose - monitoring sensors: Continuous glucose-error grid analysis illustrated by the sense freestyle navigator data. *Diabetes Care*, 27:1922–8, 2004. 49
- B. P. Kovatchev, S. Anderson, L. Heinemann, and W. Clarke. Comparison of the numerical and clinical accuracy of four continuous glucose monitors. *Diabetes Care*, 31:1160 – 1164, 2008. 9, 29, 37, 55, 93, 112
- B.P. Kovatchev, M. Breton, C.D. Man, and C. Cobelli. In silico preclinical trials: a proof of concept in close-loop control of type 1 diabetes. *J Diabetes Sci Technol*, 3(1):44–55, 2009a. 42
- B.P. Kovatchev, D. Shields, and M. Breton. Graphical and numerical evaluation of continuous glucose sensing time lag. *Diabetes Technol Ther*, 11:139–143, 2009b. 38
- A. Kowalski and J. W. Lum. Juvenile diabetes research foundation artificial pancreas consortium update. *Journal of Diabetes Science and Technology*, 3:1224–1226, 2009. 28
- J.S. Kronsley. Effect of an intensive glucose management protocol on the mortality of critically ill adult patients. *Mayo Clin Proc*, 79:992–1000, 2004. 96

- JS. Krouwer and GS. Cembrowski. A review of standards and statistics used to describe blood glucose monitor performance. *J Diabetes Sci Technol*, 4:75–83, 2010. 46
- W. H. Kruskal and A. W Wallis. Use of ranks in one-criterion analysis of variance. *Journal of the American Statistical Association*, 47(260):583–621, 1952. 106
- E. Kulcu, J.A. Tamada, G. Reach, R.O. Potts, and M.J. Lesho. Physiological differences between interstitial glucose and blood glucose measured in human subjects. *Diabetes Care*, 26(8):2405–2409, 2003. 38
- Evans M. L. Kumareswaran, K. and R. Hovorka. Artificial pancreas: an emerging approach to treat type 1 diabetes. *Expert Review of Medical Devices*, 6:401–410, 2009. 26
- M. Kuure-Kinsey, C.C. Palerm, and Bequette B.W. A dual-rate kalman filter for continuous glucose monitoring. In *Conf Proc IEEE Eng Med Biol Soc*, volume 1, pages 63–66, 2006. 11, 40, 43
- M. Laffel, L. and A.J. Kowalski. State of the art: continuous glucose monitoring. *American Academy of Pediatric Section on Endocrinology Newsletter*, 16, 2007. 10
- D.B. Leake. *Case-based reasoning: experiences, lessons and future direction*. Press, 1996. 128
- Y. Leal, W. Garcia-Gabin, J. Bondia, and J. Vehí. Real-time glucose estimation algorithm for continuous glucose monitoring using autoregressive models. In *9th Diabetes Technology Meeting (DTM)*, volume 4, page A84, 2009. 56
- Y. Leal, W. Garcia-Gabin, J. Bondia, E. Esteve, W. Ricart, J. M. Fernández-Real, and J. Vehí. Real-time glucose estimation algorithm for continuous glucose monitoring using autoregressive models. *J Diabetes Sci Technol*, 4:391–403, 2010a. 38, 56
- Y. Leal, W. Garcia-Gabin, J. Bondia, E. Esteve, W. Ricart, J. M. Fernández-Real, and J. Vehí. Enhanced algorithm for glucose estimation using the continuous glucose monitoring system. *Med Sci Monit*, 16:MT51–MT58, 2010b. 56
- Y. Leal, C. Lorenzo, J. Bondia, CC. Palerm, A. Bonet, and J. Vehí. Accuracy assessment of real-time continuous glucose monitoring systems in an intensive care unit: A pilot study in medical and surgical patients. In *10th Diabetes Technology Meeting (DTM)*, volume 5, page A79, 2010c. 97
- Y. Leal, C. Lorenzo, J. Bondia, C. C. Palerm, J. M. Sirvent, A. Bonet, and J. Vehí. Numerical-accuracy assessment of a real-time continuous glucose monitoring system in medical, surgical, and trauma critically-ill patients. *Diabetes Technol Ther*, 13:243, 2011. 97
- Y. Leal, L. Gonzalez-Abril, M. Ruiz, C. Lorenzo, J. Bondia, J. A. Ortega, and J. Vehí. Applying new support vector machines post-processing strategies for the detection of correct and incorrect measurements in cgms. In *Proc. 12th DTM Diabetes Technology Meeting (DTM)*, 2012a. 150, 162
- Y. Leal, M. Ruiz, C. Lorenzo, J. Bondia, and J. Vehí. Statistical methods and expert systems for the detection of therapeutically correct and incorrect measurements in rtcgms. In *5th International Conference on Advanced Technologies & Treatments for Diabetes (ATTD)*, 2012b. 118

- Y. Leal, L. Gonzalez-Abril, C. Lorencio, J. Bondia, and J. Vehí. Detection of correct and incorrect measurements in real-time continuous glucose monitoring systems by applying a post-processing support vector machine. *IEEE Trans Biomed Eng*, In press, DOI: 10.1109/TB-ME.2013.2244092, 2013a. 150, 162
- Y. Leal, M. Ruiz, C. Lorencio, J. Bondia, L.E. Mujica, and J. Vehí. Principal component analysis in combination with case-based reasoning for detecting therapeutically correct and incorrect measurements in continuous glucose monitoring systems. *Biomedical Signal Processing and Control*, In press, DOI:10.1016/j.bspc.2013.05.008, 2013b. 118
- L. M. Lesperance, A. Spektor, and K. J. Mcleod. Calibration of the continuous glucose monitoring system for transient glucose monitoring. *Diabetes Technol Ther*, 9:183–190, 2007. 36, 60, 61
- B. Li, J. Hu, and K. Hirasawa. An improved support vector machine with soft decision-making boundary. In *Proc. 26th IASTED Int. Conf. Artif. Intell. and Appl.*, pages 40–45, Anaheim, CA, USA, 2008a. ISBN 978-0-88986-710-9. 155
- P. Li, P-L. Qiao, and Y-Ch. Liu. A hybrid re-sampling method for SVM learning from imbalanced data sets. In *Proc. 5th Int. Conf. Fuzzy Syst. Know. Disc.*, volume 2, pages 65–69, Washington, DC, USA, 2008b. IEEE Computer Society. ISBN 978-0-7695-3305-6. 154
- L. Ljung. *System identification: theory for the user.*, volume 1. Englewood Cliffs (NJ): Prentice Hall, New York, 1987. 73
- L. Ljung. *System identification: theory for the user.*, volume 2. Englewood Cliffs (NJ): Prentice Hall, New York, 1999. 73, 74, 75
- V. Lodwig, L. Heinemann, and Glucose Monitoring Study Group. Continuous glucose monitoring with glucose sensors: calibration and assessment criteria. *Diabetes Technol Ther*, 5:572–586, 2003. 39
- S. J. Logtenberg, N. Kleefstra, F. T. Snellen, K. H. Groenier, R. J. Slingerland, A. P. Nierich, and H. J. Bilo. Pre- and postoperative accuracy and safety of a real-time continuous glucose monitoring system in cardiac surgical patients: a randomized pilot study. *Diabetes Technol Ther*, 11:31–37, 2009. 96, 110, 112, 120
- C. Lorencio, Y. T. Leal, A. Bonet, J. Vehí, J. Bondia, A. Taché, and J-M. Sirvent. Accuracy assessment of real-time continuous glucose monitors in critically ill patients: Preliminary results. *Intensive Care Med*, 36(Suppl 2):S217, 2010. 97
- C. Lorencio, Y. Leal, A. Bonet, J. Vehí, J. Bondia, CC. Palerm, A. Taché, and JM. Sirvent. Accuracy assessment of subcutaneous continuous glucose monitoring in an intensive care unit according to the septic shock state. In *24th European Society of Intensive Care Medicine Annual Congress*, volume 37(Suppl 1), page S227, 2011. 97
- C. Lorencio, Y. Leal, A. Bonet, J. Bondia, C. C. Palerm, A. Taché, J-M. Sirvent, and J. Vehí. Real-time continuous glucose monitoring in an intensive care unit: Better accuracy in patients with septic shock. *Diabetes Technol Ther*, 14:568–575, 2012. 97, 118
- Y. Lu, A. V. Gribok, W. K. Ward, and J. Reifman. The importance of different frequency bands in predicting subcutaneous glucose concentration in type 1 diabetic patients. *IEEE Trans Biomed Eng*, 57(8):1839 – 1846, 2010. 44

- J. R. Lucio. *Aplicación de Métodos Quimiométricos para la Caracterización y Control de Calidad de Plantas Medicinales*. PhD thesis, Universitat Autònoma de Barcelona, 2012. 203
- J.F. MacGregor. Multivariate statistical approaches to fault detection and isolation. In *SAFE-PROCESS*, page 17, 2003. 123
- J.J. Mastrototaro, T.M. Gross, and J.J. Shin. Glucose monitor calibration methods. United States patent US 6,424,847, Jul 23 2002. 33, 43, 59, 60
- J.J. Mastrototaro, J. Shin, A. Marcus, and G. Sultur. The accuracy and efficacy of real-time continuous glucose monitoring sensor in patients with type 1 diabetes. 10:385–390, 05 2008. 47, 64
- R. S. Mazze, E. Strock, S. Borgman, D. Wesley, P. Stout, and J. Racchini. Evaluating the accuracy, reliability, and clinical applicability of continuous glucose monitoring (cgm): is cgm ready for real time? *Diabetes Technol Ther*, 11:11–18, 2009. 37, 44, 93, 112
- F. A. McAlister, S. R. Majumdar, S. Blitz, B. H. Rowe, J. Romney, and T. J. Marrie. The relation between hyperglycemia and outcomes in 2,471 patients admitted to the hospital with community-acquired pneumonia. *Diabetes Care*, 28:810–815, 2005. 95
- K. C. McCowen, A. Malhotra, and B. R. Bistrrian. Stress-induced hyperglycemia. *Crit Care Clin*, 17:107–124, 2001. 95, 113
- G. McGarraugh. The chemistry of commercial continuous glucose monitors. *Diabetes Technol Ther*, 11(Suppl 1):S17–S24, 2009. 33, 34
- Medtronic MiniMed. *The Medtronic MiniMed CGMS® System Solutions Software: User Guide MMT-7310*, 3.0c edition, 2003. 35, 57, 59, 60
- Medtronic MiniMed. *Medtronic MiniMed Guardian® REAL-Time Continuous Glucose Monitoring System: User Guide.*, 3.0c edition, 2007. 103, 115
- S. Mitragotri, M. Coleman, J. Kost, and R. Langer. Analysis of ultrasonically extracted interstitial fluid as a predictor of blood glucose levels. *J Appl Physiol.*, 89:961–6., 2000. 30, 31
- T.P. Monsod, D.E. Flanagan, F. Rife, R. Saenz, S Caprio, R.S. Sherwin, and W.V. Tamborlane. Do sensor glucose levels accurately predict plasma glucose concentrations during hypoglycemia and hyperinsulinemia? *Diabetes Care*, 25:889–893, 2002. 10, 38, 73
- V. M. Montori, B. R. Bistrrian, and M. M. McMahan. Hyperglycemia in acutely ill patients. *JAMA*, 288:2167–2169, 2002. 95, 113
- L.E. Mujica, J. Rodellar, A. Fernández, and A. Guemes. Q-statistic and t2-statistic pca-based measures for damage assessment in structures. *Structural Health Monitoring*, 10:539–553, 2011. 13, 124
- National Institutes of Health. *Regenerative Medicine, chapter Are Stem Cells the Next Frontier for Diabetes Treatment?* 2006. 20
- J.D. Newman and Turner A.P. Home blood glucose biosensors: a commercial perspective. *Biosens Bioelectron*, 20:2435–2453, 2005. 30, 31

- Y Ni, Y Lai, S Brandes, and S. Kokot. Multi-wavelength hplc fingerprints from complex substances: An exploratory chemometrics study of the cassia seed example. *Anal. Chim. Acta*, 647:149–158, 2009. 203
- NICE-SUGAR, S. Finfer, D. R. Chittock, S. Y. Su, D. Blair, D. Foster, V. Dhingra, R. Bellomo, D. Cook, P. Dodek, W. R. Henderson, P. C. Hébert, S. Heritier, D. K. Heyland, C. McArthur, E. McDonald, I. Mitchell, J. A. Myburgh, R. Norton, J. Potter, B. G. Robinson, and J.J. Ronco. Intensive versus conventional glucose control in critically ill patients. *N Engl J Med*, 360:1283–1297, 2009. 96, 115
- P. Nomikos and J.F. MacGregor. Monitoring batch processes using multiway principal component analysis. *AIChE*, 40:1361–1375, 1994. 122
- H. Núñez, L. Gonzalez-Abril, and C. Angulo. Máquinas de soporte vectorial sobre conjuntos de datos no balanceados: Propuesta de un nuevo sesgo. Technical Report RT 2012-06, Universidad Central de Venezuela, 2012. 151, 154
- N.S. Oliver, C. Toumazou, A.E. Cass, and D.G. Johnston. Glucose sensors: a review of current and emerging technology. *Diabet Med.*, 26:197–210, 2009. 30, 31, 34
- C.C. Palerm and B.W. Bequette. Hypoglycemia detection and prediction using continuous glucose monitoring—a study on hypoglycemic clamp data. *J Diabetes Sci Technol.*, 1(5):624–9, 2007. 44
- C.C. Palerm, J.P. Willis, J. Desemone, and B.W. Bequette. Hypoglycemia prediction and detection using optimal estimation. *Diabetes Technol Ther*, 7(1):3–14, 2005. 11, 43, 44
- A. E. Panteleon, R. Kerstin, and G. M. Steil. The role of the independent variable to glucose sensor calibration. *Diabetes Technol Ther*, 5:401–10, 2003. 40, 43, 60, 61
- S.M. Pappada, B.D. Camero, and P.M. Rosman. Development of a neural network for prediction of glucose concentration in type 1 diabetes patients. *J Diabetes Sci Technol*, 2(5):792–801, 2008. 44
- Beebe J Chan S Smith NB. Park EJ, Werner J. Noninvasive ultrasonic glucose sensing with large pigs (approximately 200 pounds) using a lightweight cymbal transducer array and biosensors. *J Diabetes Sci Technol.*, 3:517–23., 2009. 30, 31
- J. L. Parkes, S. L. Slatin, S. Pardo, and B. H. Ginsberg. A new consensus error grid to evaluate the clinical significance of inaccuracies in the measurement of blood glucose. *Diabetes Care*, 23:1143–8, 2000. 49
- C. Perez-Gandia, A. Facchinetti, G. Sparacino, C. Cobelli, E.J. Gómez, M. Rigla, A. de Leiva, and M.E. Hernando. Artificial neural network algorithm for on-line glucose prediction from continuous glucose monitoring. *Diabetes Technol Ther*, 12:81–88, 2010. 44
- H. G. Piper, J. L. Alexander, A. Shukla, F. Pigula, J. M. Costello, P. C. Laussen, T. Jaksic, and M. S. Agus. Real-time continuous glucose monitoring in pediatric patients during and after cardiac surgery. *Pediatrics*, 118:1176–1184, 2006. 96, 110, 112, 114
- A. Poscia, M. Mascini, D. Moscone, M. Luzzana, G. Caramenti, P. Cremonesi, F. Valgimigli, C. Bongiovanni, and M. Varalli. A microdialysis technique for continuous subcutaneous glucose monitoring in diabetic patients (part 1). *Biosens Bioelectron.*, 18:891–8., 2003. 33, 43

- J. C. Preiser, P. Devos, S. Ruiz-Santana, C. Mélot, D. Annane, J. Groeneveld, G. Iapichino, X. Leverve, G. Nitenberg, P. Singer, J. Wernerman, M. Joannidis, A. Stecher, and R. Chioléro. A prospective randomised multi-centre controlled trial on tight glucose control by intensive insulin therapy in adult intensive care units: the glucontrol study. *Intensive Care Med*, 35: 1738–1748, 2009. 96
- A. Rabiee, R. N. Andreasik, R. Abu-Hamdah, B. S. Galiatsatos, Z. Khouri, B. R. Gibson, D. K. Andersen, and D. Elahi. Numerical and clinical accuracy of a continuous glucose monitoring system during intravenous insulin therapy in the surgical and burn intensive care units. *J Diabetes Sci Technol*, 3:951–959, 2009. 96, 110
- N. Ramchandani and R. Heptulla. New technologies for diabetes: a review of the present and the future. *International Journal of Pediatric Endocrinology*, 28:1–10, 2012. 30
- G. Rao, P. Glikfeld, and R.H. Guy. Reverse iontophoresis: development of a noninvasive approach for glucose monitoring. *Pharm Res.*, 10:1751–5., 1993. 30, 31
- P. Rapp, M. Mesch, H. Giessen, and C. Tarín. Regression methods for ophthalmic glucose sensing using metamaterials. *Journal of Electrical and Computer Engineering*, 11: doi:10.1155/2011/953064, 2011. 26
- G. Reach. Continuous glucose monitoring and diabetes health outcomes: a critical appraisal. *Diabetes Technol Ther*, 10:69–80, 2008. 34
- K. Rebrin, G. M. Steil, W. P. van Antwerp, and J. J. Mastrototaro. Subcutaneous glucose predicts plasma glucose independent of insulin: implications for continuous monitoring. *Am J Physiol*, 277:E561–E571, 1999. 10, 38, 41
- J. Reifman, S. Rajaraman, A. V. Gribok, and W. K. Ward. Predictive monitoring for improved management of glucose levels. *J Diabetes Sci Technol*, 1:478–86, 2007. 11, 12, 44, 73
- E. Renard. Implantable continuous glucose sensors. *Curr Diabetes Rev.*, 4:169–74., 2008. 30, 31
- E. Renkin. Capillary permeability to small solutes. In *In Handbook of Physiology: The Cardiovascular system, Microcirculation*, volume IV. American Physiology Society, Bethesda, MD, USA., 1984. 38
- S.Y. Rhee, S. Chon, G. Koh, J.R. Paeng, S. Oh, J.T. Woo, S.W. Kim, J.W. Kim, and Y.S. Kim. Clinical experience of an iontophoresis based glucose measuring system. *J Korean Med Sci.*, 22:70–3., 2007. 30, 31
- M. J. Rice, A. D. Pitkin, and D. B. Coursin. Review article: glucose measurement in the operating room: more complicated than it seems. *Anesth Analg*, 110:1056–1065, 2010. 115
- J.J. Robert. Continuous monitoring of blood glucose. *Horm Res*, 57:81–4., 2002. 30, 31
- G. Roglic and N. Unwin. Mortality attributable to diabetes: Estimates for the year 2010. *Diabetes Research and Clinical Practice*, 87(3): 15–19, 2010. ISSN 0168-8227. doi: 10.1016/j.diabres.2011.10.029. URL <http://www.sciencedirect.com/science/article/pii/S0168822711005912>. 22
- P. Rossetti, J. Bondia, J. Vehi, and C.G. Fanelli. Estimating plasma glucose from interstitial glucose: The issue of calibration algorithms in commercial continuous glucose monitoring devices. *Sensors*, 10:10936–10952, 2010. 9, 12, 28, 29, 37, 38, 42, 55

- M. Ruiz. *Multivariate statistical process control and case-based reasoning for situation assessment of sequencing batch reactors*. PhD thesis, University of Girona, 2008. 129, 130
- M. Ruiz, G. Sin, X. Berjaga, J. Colprim, S. Puig, and J. Colomer. Multivariate principal component analysis and case-based reasoning for monitoring, fault detection and diagnosis in a wwtp. *Water Science and Technology*, 64.8:1661–1667, 2011. 13
- B. Rukshan and P. Vasile. FSVM-CIL: fuzzy support vector machines for class imbalance learning. *IEEE Trans. Fuzzy Syst.*, 18:558–571, 2010. ISSN 1063-6706. 155
- S. S Shapiro. An analysis of variance test for normality (complete samples). 1965. 65, 84
- Q. Shen, S. J. Qin, and K. J. Doniger. Online dropout detection in subcutaneously implanted continuous glucose monitoring. In *Proc. American Control Conference (ACC)*, pages 4373–4378, 2010. ISBN 978-0-7695-4109-9. 12, 44, 45, 117
- J.J. Shin, K.R. Holtzclaw, N.D. Dangui, S Jr. Kanderian, J.J. Mastrototaro, and P.I. Hong. Real time self-adjusting calibration algorithm. United States patent US 6,895,263 B2., May 17 2005. 43
- M. Signal, A. Le Compte, D.L. Harris, P.J. Weston, J.E. Harding, and J.G. Chase. Using stochastic modelling to identify unusual continuous glucose monitor measurements and behaviour, in newborn infants. *BioMedical Engineering OnLine*, 11:doi:10.1186/1475-925X-11-45, 2012. 12, 44, 46, 117
- P. Simpson, M. Brister, M. Wightlin, and J. Pryor. Dual electrode system for a continuous analyte sensor. Publication Number WO/2008/042918, April 10 2008. 33, 43
- M. Sjöström, S. Wold, and B. Soderstrom. Pls discriminant plots, proceedings of parc in practice. Elsevier Science Publishers B.V., North-Holland., 1986. 136, 203
- G. Sparacino, F. Zanderigo, S. Corazza, A. Maran, A. Facchinetti, and C. Cobelli. Glucose concentration can be predicted ahead in time from continuous glucose monitoring sensor time-series. *IEEE Trans Biomed Eng*, 54:931–37, 2007. 11, 12, 44, 73
- G. Sparacino, A. Facchinetti, and C. Cobelli. Smart continuous glucose monitoring sensors: On-line signal processing issues. *Sensors*, 10:6751–6772, 2010. 37, 40
- G. M. Steil, K. Rebrin, J. Mastrototaro, B. Bernaba, and M. F. Saad. Determination of plasma glucose during rapid glucose excursions with a subcutaneous glucose sensor. *Diabetes Technol Ther*, 5, 2003. 36, 38
- F. Sternberg, C. Meyerhoff, H. Mennel, F. J. and Mayer, F. Bischof, and E. F. Pfeiffer. Does fall in tissue glucose precede fall in blood glucose? *Diabetologia*, 39:609–612, 1996. 10, 38, 73
- A. D. Stork, H. Kemperman, D. W. Erkelens, and T. F. Veneman. Comparison of the accuracy of the hemocue glucose analyzer with the yellow springs instrument glucose oxidase analyzer, particularly in hypoglycemia. *Eur. J. Endocrinol.*, 153:275–281, 2005. 115
- P.J. Stout, N. Peled, B.J. Erickson, M.E. Hilgers, J.R. Racchini, and T.B. Hoegh. Comparison of glucose levels in dermal interstitial fluid and finger capillary blood. *Diabetes Technol Ther*, 3:81–90, 2001. 38

- A. Sun, E-P. Lim, and Y. Liu. On strategies for imbalanced text classification using svm: A comparative study. *Decision Support Systems*, 48:191–201, December 2009. ISSN 0167-9236. 156
- M. C. Torjman, L. Jahn, J. I. Joseph, and K. Crothall. Accuracy of the hemocue portable glucose analyzer in a large nonhomo- 47. geneous population. *Diabetes Technol Ther*, 3:591–600, 2001. 115
- A. Tura, A. Maran, and G. Pacini. Non-invasive glucose monitoring: assessment of technologies and devices according to quantitative criteria. *Diabetes Res Clin Pract.*, 77:16–40., 2007. 30, 31
- S. Vaddiraju, D.J. Burgess, I. Tomazos, F.C. Jain, and F. Papadimitrakopoulos. Technologies for continuous glucose monitoring: Current problems and future promises. *J Diabetes Sci Technol*, 4(6):1540–1562, 2010a. 29, 30, 32
- S. Vaddiraju, I. Tomazos, D.J. Burgess, F.C. Jain, and F. Papadimitrakopoulos. Emerging synergy between nanotechnology and implantable biosensors: a review. *Biosens Bioelectron.*, 25(7):1553–65., 2010b. 30, 31
- G. Van den Berghe, P. Wouters, F. Weekers, C. Verwaest, F. Bruyninckx, M. Schetz, D. Vlasselaers, P. Ferdinande, P. Lauwers, and R. Bouillon. Intensive insulin therapy in the critically-ill patients. *N Engl J Med*, 345:1359–1367, 2001. 96
- G. Van den Berghe, A. Wilmer, G. Hermans, W. Meersseman, P. J. Wouters, I. Milants, E. Van Wijngaerden, H. Bobbaers, and R. Bouillon. Intensive insulin therapy in the medical icu. *N Engl J Med*, 354:449–461, 2006. 96, 115
- I. Vanhorebeek and L. Langouche. Molecular mechanisms behind clinical benefits of intensive insulin therapy during critical illness: glucose versus insulin. *Best Pract Res Clin Anaesthesiol*, 23:449–459., 2009. 95, 113
- V. Vapnik. *Statistical Learning Theory*. John Wiley & Sons, New York, 1998. 151
- V. Vapnik. *The Nature of Statistical Learning Theory (Information Science and Statistics)*. Springer, November 1999. ISBN 0387987800. 151, 153
- P.F. Velleman and D.C. Hoaglin. *Applications, basics, and computing of exploratory data analysis*. Boston: Duxbury Press, 1981. 125
- K. Veropoulos, C. Campbell, and N. Cristianini. Controlling the sensitivity of support vector machines. In *Proc. 16th Int. Joint Conf. Artif. Intell*, pages 55–60, 1999. 155
- F. Vilariño, P. Spyridonos, J. Vitrià, and P. Radeva. Experiments with SVM and stratified sampling with an imbalanced problem: Detection of intestinal contractions. In *Proc. Third Int. Conf. Adv. Patt. Recogn.*, volume 2, pages 783–791, August 2005. 154
- J. L. Vincent, R. Moreno, J. Takala, S. Willatts, A. De Mendonça, H. Bruining, C. K. Reinhart, and L. G. Suter, P. M. Thijs. The sofa (sepsis-related organ failure assessment) score to describe organ dysfunction/failure. *Intensive Care Med*, 22:707–710, 1996. 97
- T. M. Vriesendorp, J. H. DeVries, F. Holleman, M. Dzoljic, and J. B. Hoekstra. The use of two continuous glucose sensors during and after surgery. *Diabetes Technol Ther*, 7:315–322, 2005. 96, 110, 112

- B. X. Wang and N. Japkowicz. Boosting support vector machines for imbalanced data sets. volume 4994 of *Lect. Not. Comp. Sc.*, pages 38–47. Springer, 2008. ISBN 978-3-540-68122-9. 155
- W. K. Ward, H. M. Casey, M. J. Quinn, I. F. Federiuk, and M. D. Wood. A fully implantable subcutaneous glucose sensor array: Enhanced accuracy from multiple sensing units and a median-based algorithm. *Diabetes Technol Ther*, 5:943–952, 2004. 12, 44, 45, 117
- B. Waske, J. A. Benediktsson, and J. R. Sveinsson. Classifying remote sensing data with support vector machines and imbalanced training data. In *Proc. 8th Int. Works. Multiple Class. Syst.*, pages 375–384, Berlin, Heidelberg, 2009. Springer-Verlag. ISBN 978-3-642-02325-5. 155
- P. J. Watkinson, V. S. Barber, E. Amira, T. James, R. Taylor, and J. D. Young. The effects of precision, haematocrit, ph and oxygen tension on point-of-care glucose measurement in critically ill patients: a prospective study. *Ann Clin Biochem*, 49(Pt 2):144–151, 2012. 115, 148
- R.L. Weinstein, S. L. Schwartz, R. L. Brazg, J. R. Bugler, T. A. Peyser, and G. V. McGarraugh. Accuracy of the 5-day freestyle navigator continuous glucose monitoring system comparison with frequent laboratory reference measurements. *Diabetes Care*, 30:1125–30, 2007. 83, 93
- S. A. Weinzimer, M. C. De Lucia, E. A. Boland, A. Steffen, and W. V. Tamborlane. Analysis of continuous glucose monitoring data from non-diabetic and diabetic children: a tale of two algorithms. *Diabetes Technol Ther*, 5, 2003. 55
- I. M. Wentholt, M. A. Vollebregt, A. A. Hart, J. B. Hoekstra, and J. H. DeVries. Comparison of a needle-type and a microdialysis continuous glucose monitor in type 1 diabetic patients. *Diabetes Care*, 28:2871–2876, 2005. 9, 29, 55
- I. M. Wentholt, A. A. Hart, J. B. Hoekstra, and J. H. Devries. How to assess and compare the accuracy of continuous glucose monitors? *Diabetes Technol Ther*, 10:57–68, 2008. 46, 47
- I.M. Wentholt, A.A. Hart, J.B. Hoekstra, and J.H. Devries. Relationship between interstitial and blood glucose in type 1 diabetes patients: Delay and the push-pull phenomenon revisited. *Diabetes Technol Ther.*, 9:169–175., 2007. 38
- D.R Whiting, L. Guariguata, C. Weil, and J. Shaw. {IDF} diabetes atlas: Global estimates of the prevalence of diabetes for 2011 and 2030. *Diabetes Research and Clinical Practice*, 94(3):311 – 321, 2011. ISSN 0168-8227. doi: 10.1016/j.diabres.2011.10.029. URL <http://www.sciencedirect.com/science/article/pii/S0168822711005912>. 21
- K. Wiener. An assessment of the effect of haematocrit on the hemocue blood glucose analyser. *Ann Clin Biochem*, 30:90–93, 1993. 114
- F. Wilcoxon. Individual comparisons by ranking methods. *Biometrics*, 1:80–83, 1945. 65, 84
- G. Williams and J.C. Pickup. *Handbook of Diabetes*. Blackwell Publishing, 2004. 7
- B. M. Wise, N. B. Gallagher, R. Bro, J. M. Shaver, W. Windig, and R. S. Koch. *Chemometrics Tutorial for PLS Toolbox and Solo*. Eigenvector Research Incorporated, 2006. 203
- S Wold, C Albano, W.J. Dunn, U Edlund, K Esbensen, P Geladi, S. Hellberg, E. Johansson, W. Lindberg, and M. Sjöström. Multivariate data analysis in chemistry. In D. Reidel Publishing Company, editor, *B.R. Kowalski (ed.) Chemometrics: Mathematics and Statistics in Chemistry*, 1984. 136, 203

- G. Wu and E. Y. Chang. KBA: Kernel Boundary Alignment Considering Imbalanced Data Distribution. *IEEE Trans. Know. Data Eng.*, 17:786–795, 2005. ISSN 1041-4347. 154, 155, 156
- Ch. Y. Yang, J. Wang, Jr. S. Yang, and G. D. Yu. Imbalanced svm learning with margin compensation. In *Proc. 5th Int. Symp. Neural Netw.*, pages 636–644, Berlin, Heidelberg, 2008. Springer-Verlag. ISBN 978-3-540-87731-8. 155
- L. Yang, A. Aijun, and H. Xiangji. Boosting prediction accuracy on imbalanced datasets with SVM ensembles. In *Proc. 10th Pacific-Asia Conf. Know. Disc. Data Mining*, pages 107–118, 2006. 155, 156
- P. Yang, Z. Zhang, B. B. Zhou, and A. Y. Zomaya. Sample subset optimization for classifying imbalanced biological data. In *Proc. 15th Pacific-Asia Conf. Adv. Know. Disc. Data Min.*, volume 2, pages 333–344, Berlin, Heidelberg, 2011. Springer-Verlag. ISBN 978-3-642-20846-1. 155
- S. Yendamuri, G. J. Fulda, and G.H. Tinkoff. Admission hyperglycemia as a prognostic indicator in trauma. *J Trauma*, 55:33–38, 2003. 95
- T. Yu, J. Debenham, T. Jan, and S. Simoff. Combine vector quantization and support vector machine for imbalanced datasets. In Max Bramer, editor, *Artif. Intell. Theory and Prac.*, volume 217 of *IFIP International Federation for Information Processing*, chapter 9, pages 81–88. Springer US, 2006. ISBN 978-0-387-34654-0. 154
- C. Zecchin, A. Facchinetti, G. Sparacino, G. De Nicolao, and C. Cobelli. Neural network incorporating meal information improves accuracy of short-time prediction of glucose concentration. *IEEE Trans Biomed Eng*, 59(6):1550 – 1560, 2012. 44
- J. Zhou, W. Jia, Y. Bao, X. Ma, W. Lu, H. Li, C. Hu, and K. Xiang. Glycemic variability and its responses to intensive insulin treatment in newly diagnosed type 2 diabetes. *Med Sci Monit*, 14:CR552–8, 2008. 35
- H.C. Zisser, T.S. Bailey, S. Schwartz, R.E. Ratner, and J. Wise. Accuracy of the seven continuous glucose monitoring system: comparison with frequently sampled venous glucose measurements. *J Diabetes Sci Technol*, 3:1146–54, 2009. 93

Apéndice A

Conceptos básicos de SIMCA y PLS-DA

En este apéndice se presentan algunos conceptos básicos sobre dos métodos multivariantes de uso común para clasificación, los cuales se emplearon en el Capítulo 5 para realizar una comparativa con la metodología PCA-CBR. Estos métodos son: El Modelado Blando Independiente de Analogías entre Clases (-del inglés *Soft Independent Modeling of Class Analogy (SIMCA)*-) y el Análisis Discriminante por Mínimos Cuadrados Parciales (-del inglés *Partial Least Squares Discriminant Analysis (PLS-DA)*-). Mayor información sobre SIMCA y PLS-DA se puede encontrar en (Wold et al., 1984) y (Sjöström et al., 1986), respectivamente.

SIMCA es un método de clasificación supervisado, el cual consiste de una colección de modelos de Análisis de Componentes Principales (-del inglés *Principal Component Analysis (PCA)*-), de forma que existe un modelo por cada clase en el conjunto de datos. Cada clase se modela usando un PCA de manera independiente, por tanto las diferentes clases pueden tener el mismo número o un número diferente de PCs (Gemperline, 2006). Cada sub-modelo PCA contiene todas las partes usuales de un modelo PCA. Para saber si una muestra se ajusta a la clase, la información que se calcula es dividida en dos, una parte explicada por el modelo de la clase y otra que permanece en los residuales. Si los residuales de la muestra son significativamente mayores que los de la clase, la muestra es rechazada (Ni et al., 2009).

Una forma de representar los resultados en la clasificación de las muestras desconocidas, es determinar el espacio definido por una clase dada (estadístico T^2) y su espacio residual (estadístico Q), con la finalidad de evaluar si los valores de cada muestra predicha están mejor descritos por el espacio T^2 ; de ser así, las muestras son clasificadas como miembros de la clase (Lucio, 2012). La clase más cercana a una muestra i se define como la clase del sub-modelo que resulta en una mínima distancia de la muestra i al sub-modelo j , d_{ij} (Wise et al., 2006):

$$d_{ij} = \sqrt{(Q_r)^2 + (T_r^2)^2} \quad (\text{A.1})$$

donde T_r^2 y Q_r se definen como:

$$T_r^2 = \frac{T^2}{T_{0,95}^2} \quad (\text{A.2})$$

$$Q_r = \frac{Q}{Q_{0,95}} \quad (\text{A.3})$$

siendo $T_{0,95}^2$ y $Q_{0,95}$ los límites de los estadísticos T^2 y Q (nivel de confianza del 95 %) para el modelo en cuestión. Esta medida de distancia da el mismo peso a la distancia en el espacio del modelo (T^2) y en el espacio residual (Q). El uso de T_r^2 y Q_r permite la comparación directa de los valores de muestras específicas en los diferentes sub-modelos SIMCA. Sin el uso de estos estadísticos reducidos, los valores de T^2 y Q de cada modelo estarían en escalas numéricas muy diferentes y no serían fácilmente comparables.

El principal inconveniente de SIMCA es que los sub-modelos PCA se computan con el objetivo de capturar la variación dentro de cada clase. No se hace ningún intento de identificar en el espacio de los datos las direcciones que discriminen las clases directamente.

PLS-DA es un método de clasificación que si aborda este tema directamente. Este método de análisis supervisado que se basa en Regresión por Mínimos Cuadrados Parciales (-del inglés *Partial Least Squares* (PLS)-) combinado con Análisis Discriminante (-del inglés *Discriminant Analysis* (DA)-). PLS tiene como objetivo principal identificar un subconjunto reducido de variables latentes r que expliquen como los cambios en la matriz de datos originales X (que para PLS será la matriz predictora), afectan la matriz predicha Y . Por consiguiente, se puede utilizar Y con el fin de obtener un modelo PLS, para distinguir entre varias categorías o clases de las observaciones. Puesto que, los métodos de regresión generalmente están asociados con las variables continuas, PLS se combina con DA para restringir la regresión a un conjunto finito de clases en Y . La combinación de ambas técnicas proporciona una clara separación entre los grupos de observaciones, mezclando las variables latentes de tal manera que se alcance la separación máxima entre las clases, y basándose en que las variables observadas llevan la información de separación de las clases.

Similar a PCA, la matriz de observaciones X es una matriz de $n \times m$ con n filas (observaciones o muestras) y m columnas (variables). La matriz Y está formada por un conjunto de variables categóricas, también conocidas como variables “dummy”, existiendo tantas variables en Y como clases en la clasificación. Dada que cada una de las variables que conforman la matriz Y representa cada una de las clases, en cada una de ellas un 1 indica que una muestra pertenece a una determinada clase y un 0 indica que no es así. Sin embargo, cuando se trata de un problema de clasificación binario suele utilizarse una única variable, en la cuál un 1 indica que una muestra pertenece a una clase y un 0 que pertenece a la otra clase. El objetivo es encontrar la relación entre la matriz X con la matriz Y , y para ello se computa la matriz de *scores* $T = XW$ para una apropiada matriz de pesos W , y luego se considera el modelo de regresión lineal de la ecuación (A.4):

$$Y = TQ + E \quad (\text{A.4})$$

donde Q es la matriz de *loadings* y E es la matriz del error resultante de la regresión. Una vez que se computan los *loadings* y siendo B la matriz de coeficientes de la regresión, con $B = WQ$, la ecuación (A.4) es equivalente a:

$$Y = XB + E \quad (\text{A.5})$$

El algoritmo estándar para computar los factores de la regresión PLS es el algoritmo NIPALS (-del inglés *Non-Linear Iterative Partial Least Squares*-). Finalmente, la estimación de la variable predicha (\hat{y}_{new}) para una nueva observación (x_{new}) se obtiene de acuerdo a la ecuación A.6. En un conjunto con sólo dos clases, dado que \hat{y}_{new} no predice exactamente ya sea un 1 (si la nueva

observación está en una de las clases) o un 0 (si no lo está), se establece un límite, por ejemplo 0,5, sobre el cuál la nueva observación se estima como un 1 y bajo el cuál se estima como un 0.

$$\hat{y}_{new} = x_{new}B \tag{A.6}$$



Imagen de portada: MiniMed Gurardian® REAL-Time CGM System. Manufactured by Medtronic Diabetes.
Imagen reproducida con permiso de Medtronic, Inc., EE.UU.