

Uso de herramientas de análisis de punto de cambio para evaluar el desempeño de sistemas de Páncreas Artificial*

María Cecilia Serafini, Emilia Fushimi, Hernan De Battista, Fabricio Garelli
Grupo de Control Aplicado (GCA), Instituto LEICI (UNLP-CONICET)
Facultad de Ingeniería, Universidad Nacional de La Plata, Argentina

1. Introducción

La Diabetes Mellitus Tipo 1 (DMT1) es una enfermedad auto-inmune caracterizada por la imposibilidad de producir insulina. Los pacientes con DMT1 presentan altos niveles de glucemia, lo cual tiene graves consecuencias para la salud si se mantiene por tiempos prolongados (por ejemplo, ceguera, amputación, etc). La terapia tradicional para la DMT1 consiste en múltiples inyecciones subcutáneas de insulina a lo largo del día y monitoreo de los niveles de glucemia utilizando tiras reactivas con glucómetro. Sin embargo, en los últimos años se han desarrollado bombas que entregan insulina a nivel subcutáneo de forma continua (CSII), las cuales permiten que la dosificación sea más precisa y apropiada. Por otro lado, también han salido al mercado los llamados monitores continuos de glucosa (CGM), los cuales permiten obtener mediciones de la glucemia cada 5 minutos tomando muestras a nivel subcutáneo. Hoy en día, existen CGM disponibles que reemplazan por completo las mediciones con glucómetro. Con el surgimiento de la terapia CSII y los CGM, aparece la posibilidad de construir un Páncreas Artificial (AP) no invasivo. Este sistema consiste en conectar la bomba de insulina con el CGM a través de un algoritmo de control que regule la glucemia de los pacientes de forma automática. Sin embargo, la complejidad y variabilidad del sistema a controlar (el cuerpo humano) hacen que éste sea aún un problema abierto.

En el área de investigación sobre AP existe un consenso entre los mayores grupos de investigación de la temática para la presentación de resultados [1]. Sin embargo, en ese mismo reporte de consenso se incentiva a la presentación de medidas y parámetros adicionales que puedan asistir en la interpretación y el análisis del funcionamiento de los sistemas evaluados. Con esa motivación, en este trabajo se investiga el uso de la herramienta Análisis de Punto de Cambio (CPA) para la evaluación de desempeño de sistemas de AP. CPA es una herramienta estadística utilizada para detectar cambios en los parámetros de la distribución de un conjunto de datos [2]. Inicialmente, los modelos de punto de cambio se utilizaban para analizar distribuciones Gaussianas, pero alrededor de los años '80 comenzaron a desarrollarse modelos basados en test de hipótesis no-paramétricos [3]. Dependiendo de los test estadísticos que se utilicen, se obtienen Modelos de Punto de Cambio (CPM) que pueden detectar cambios en parámetros de localización (como media, mediana, moda), de escala (como el rango, la desviación estándar, la dispersión), combinaciones de ambas o cambios más generales.

Para este caso particular de AP, se cuenta con datos de los primeros ensayos clínicos realizados en Latinoamérica. Durante la segunda fase de estos ensayos se monitorearon pacientes con un sistema a lazo abierto y luego se ensayó un sistema de control a lazo cerrado basado en el algoritmo de control Automatic Regulation of Glucose (ARG) (ver sección 3).

El objetivo del presente análisis es determinar si las herramientas de CPA pueden detectar en qué momento se cierra el lazo de control del sistema, y de esa manera, evaluar la significancia del cambio en el control de glucosa luego de ese cierre. En particular, se evalúan CPM basados en los tests de Mood, y Mann-Whitney (MW).

2. Modelos de punto de cambio

El problema de la detección de puntos de cambio en la distribución de un conjunto de datos fue planteado inicialmente por Page [4], utilizando Sumas Acumulativas (CUSUMs) y por otros autores considerando tests de hipótesis para detectar cambios en la media de distribuciones Gaussianas [5], entre otros abordajes. Luego, cerca de 1980, se comenzaron a considerar test de hipótesis para distribuciones no-paramétricas o desconocidas [3].

En la actualidad, las técnicas de CPA son muy utilizadas en el campo del control estadístico de procesos industriales. Sin embargo, su utilidad está siendo extendida a estudios diversos como análisis de variabilidad climática y resultados de estudios clínicos. Actualmente existen múltiples herramientas para realizar CPA, muchas de las cuales están incluidas en paquetes de funciones para programas de análisis estadístico como **R**.

Para este trabajo se utilizaron algunos de los tests contenidos en el paquete **cpm** diseñado por Ross [6], en el cual se considera lo siguiente: Se asumen variables aleatorias independientes X_1, X_2, \dots

* publicado en AADECA 2018 – 26º Congreso Argentino de Control Automático

y se nota x_i a una realización particular de X_i observada en un tiempo t . Se supone que se observan n puntos. Para cualquier $k < n$, la hipótesis de que exista un punto de cambio puede expresarse como

$$H_0 : X_i \sim F_0(x; \theta_0), \quad i = 1, \dots, n$$

$$H_1 : X_i \sim \begin{cases} F_0(x; \theta_0) & i = 1, \dots, k \\ F_1(x; \theta_1) & i = k + 1, k + 2, \dots, n \end{cases}, \quad k \leq n$$

donde θ representa los parámetros desconocidos de la distribución. Si la hipótesis nula H_0 es aceptada, significa que la distribución es la misma para las n muestras, con parámetros θ_0 . Si la hipótesis nula es rechazada, significa que F tiene una distribución F_0 con parámetros θ_0 hasta el punto k , y otra distinta (F_1 con parámetros θ_1) en adelante. Para evaluar este cambio se puede utilizar un test de hipótesis de dos muestras, con la elección de un estadístico elegido apropiadamente dependiendo de qué datos se tengan sobre la distribución y qué tipo de cambio se quiera detectar. En principio, los test de hipótesis de dos muestras solamente responden a la pregunta "¿hay un cambio en la distribución?", y por eso es necesario adaptarlos para que, además de esa respuesta, se pueda obtener el punto donde se produce ese cambio, si lo hay.

Luego de la elección de un test específico puede computarse el valor de su estadístico, al que denotaremos $\tilde{D}_{k,n}$, y compararlo con un umbral adecuado en el posible punto de cambio k (la elección de este umbral es discutida más adelante). Como no se conoce a priori donde se encuentra el punto de cambio, el estadístico es calculado iterativamente en todos los valores $1 < k < n$ y se toma como punto de cambio el valor de k donde se produce el máximo. En el paquete **cpm**, el punto de cambio se supone en el valor máximo del estadístico, que es de la forma:

$$D_n = \max_{k=2, \dots, n-1} D_{k,n} = \max_{k=2, \dots, n-1} \left| \frac{\tilde{D}_{k,n} - \mu_{\tilde{D}_{k,n}}}{\sigma_{\tilde{D}_{k,n}}} \right|$$

donde el estadístico del test elegido es estandarizado para tener media $\mu = 0$ y varianza $\sigma = 1$. Para esta estandarización, es necesario conocer estos parámetros de $\tilde{D}_{k,n}$, lo cual en general no es trivial y según la prueba de hipótesis elegida, el método para aproximar esos valores es distinto. Para este análisis se utilizan los tests de Mood y MW, adaptados a CPA por Ross [7], y Hawkins [8] respectivamente, este último basándose en el trabajo preliminar de Pettitt [3].

2.1. Test de Mood

Este CPM detecta cambios en la escala de la secuencia de datos y utiliza un test estadístico basado en la observación de dos secuencias de datos S y T , con $n_S = k$ y $n_T = n - k$ el número de muestras de S y T , que en total suman n muestras. Este test evalúa que tanto se aleja cada x_k de su posición esperada en el conjunto ordenado de muestras a través del estadístico

$$\tilde{D} = \sum_{x_k} (R(x_k) - (n+1)/2)^2$$

donde x_k es la muestra en punto k y $R(x_k)$ denota la posición que ocupa en el conjunto ordenado (*ranking*). Luego, para este caso la media y varianza pueden obtenerse como

$$\mu_{\tilde{D}} = n_S(n^2 - 1)/12 \quad \sigma_{\tilde{D}}^2 = n_S n_T (n+1)(n^2 - 4)/180$$

y luego, como ya se había descrito para el caso general, la forma estandarizada del estadístico sería

$$D = \left| (\tilde{D} - \mu_{\tilde{D}}) / \sigma_{\tilde{D}} \right|$$

y el x_k en el cual se maximiza ese estadístico se considera el lugar donde se encuentra el punto de cambio. Para mayor detalle en el desarrollo matemático referirse a [7].

2.2. Test de Mann-Whitney

En 1979 Pettitt propuso un estadístico basado en la prueba U de Mann-Whitney [3], que detecta cambios en los parámetros de localización de la distribución. En base a ello se han desarrollado varios criterios para determinar la existencia de puntos de cambio. En particular, en el paquete **cpm**, se utiliza el desarrollo de Hawkins y Deng [8]: Sea

$$\text{sgn}(X_i - X_j) = \begin{cases} 1 & \text{si } X_i > X_j \\ 0 & \text{si } X_i = X_j \\ -1 & \text{si } X_i < X_j \end{cases}$$

y $\tilde{D}_{k,n}$, una función antisimétrica de $\text{sgn}(X_i - X_j)$

$$\tilde{D}_{k,n} = \sum_{i=1}^k \sum_{j=k+1}^n \text{sgn}(X_i - X_j)$$

$\tilde{D}_{k,n}$ define el estadístico de Mann-Whitney para verificar si las dos muestras X_1, \dots, X_k y X_{k+1}, \dots, X_n tienen la misma distribución o no. Si consideramos $R(X_i)$, el ranking de X_i , puede demostrarse que

$$\tilde{D}_{k,n} = 2 \sum_{i=1}^k R(X_i) - k(n+1)$$

y esto permite calcular una esperanza y una varianza para el estadístico:

$$E(\tilde{D}_{k,n}) = 0 \quad \text{Var}(\tilde{D}_{k,n}) = \frac{k(n-k)(n+1)}{3}$$

Schechtman y Wolfe [9] propusieron estandarizar el estadístico para eliminar la dependencia de la varianza definiendo

$$D_{k,n} = \frac{\tilde{D}_{k,n}}{\sqrt{k(n-k)(n+1)/3}} \quad (1)$$

y considerando la normalidad asintótica del test de MW, puede considerarse $D_{k,n} \sim N(0,1)$ cuando k y $n-k$ tienden a infinito.

Luego, por la definición dada para el caso general, el estadístico final y el punto estimado de cambio quedan definidos como:

$$D_{max,n} = \max_{1 \leq k \leq n-1} |D_{k,n}| \quad \tilde{\tau}_D = \operatorname{argmax}_{1 \leq k \leq n-1} |D_{k,n}| \quad (2)$$

2.3. Determinación de umbral

Hawkins y Zamba [10] recomiendan calcular los umbrales de manera tal que la probabilidad de una detección falsa se mantenga constante. En el caso del paquete **cpm** se utiliza la simulación Monte Carlo para computar los umbrales y se los tabula para distintos valores de significancia [7]. En este trabajo se utilizó, para todos los test el mismo valor de significancia: $\alpha = 0,05$, donde α es la probabilidad de un error de tipo 1 en el resultado del test de hipótesis (rechazo erróneo de la hipótesis nula, es decir, detección errónea de un punto de cambio) y el tiempo promedio entre falsos positivos está definido por $ARL_0 = 1/\alpha$.

3. Descripción del ensayo clínico y tratamientos de DMT1

La terapia tradicional para la DMT1 consiste en un bolo prandial para compensar las comidas, el cual es calculado e infundido manualmente por los pacientes. Este tratamiento está basado en parámetros individualizados del paciente que se analizan periódicamente por el médico para lograr un control glucémico correcto. Sin embargo, este control manual o de lazo abierto posee varias desventajas (error en el conteo de carbohidratos, por ejemplo), por lo que interesa desarrollar estrategias de lazo cerrado o puramente automáticas para abordar el tratamiento de la DMT1 [11].

Recientemente se realizaron en el Hospital Italiano de Buenos Aires (HIBA) los primeros ensayos clínicos de AP en Latinoamérica. Allí se puso a prueba el algoritmo ARG, desarrollado en conjunto por la Universidad Nacional de La Plata (UNLP), el Instituto Tecnológico de Buenos Aires (ITBA) y la Universidad Nacional de Quilmes (UNQ) [12]. El ensayo contaba con 5 participantes con DMT1. El protocolo aprobado por la ANMAT estipulaba 54hs de monitoreo con un monitor continuo de glucosa (CGM) utilizando la terapia tradicional de lazo abierto, seguido de 36hs de control a lazo cerrado utilizando el algoritmo ARG [13].

La figura 1 muestra un diagrama en bloques del algoritmo ARG. Éste consiste en un controlador principal Switched Linear Quadratic Gaussian (SLQG) al que se le suma el suministro de insulina basal de lazo abierto del paciente. El SLQG, conmuta entre un controlador agresivo, que se encarga de reemplazar al bolo de lazo abierto para compensar las comidas, y uno conservador para mantener al paciente en el rango deseado de glucemia ([70 – 180] mg/dl) cuando no come. La dosis de insulina que indica el SLQG es multiplicada por una señal γ calculada en el bloque Safety Auxiliary Feedback Element (SAFE), que tiene como función modular la infusión de insulina si esta llegara a implicar un riesgo para el paciente. Para una explicación más detallada del funcionamiento del algoritmo ARG referirse a [12, 13].

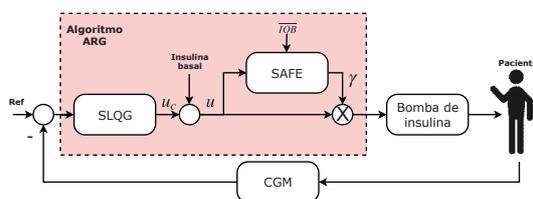
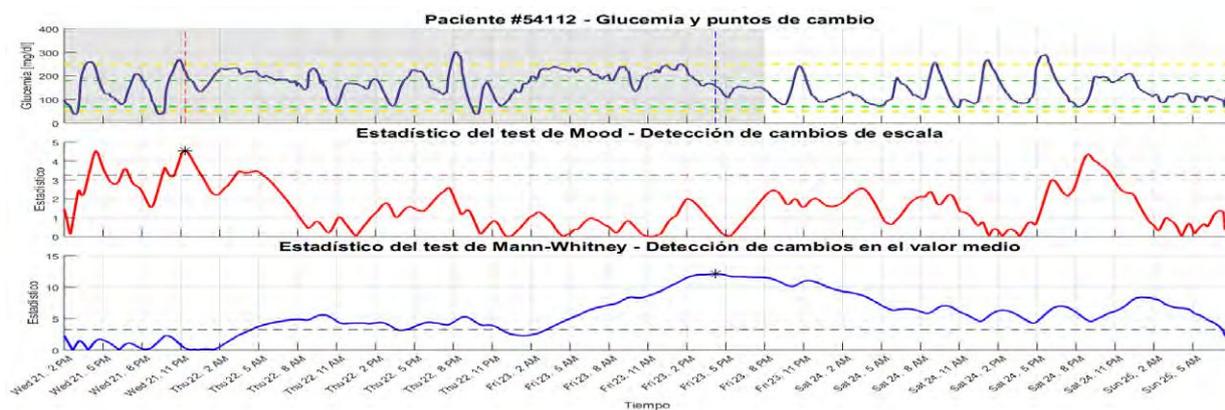


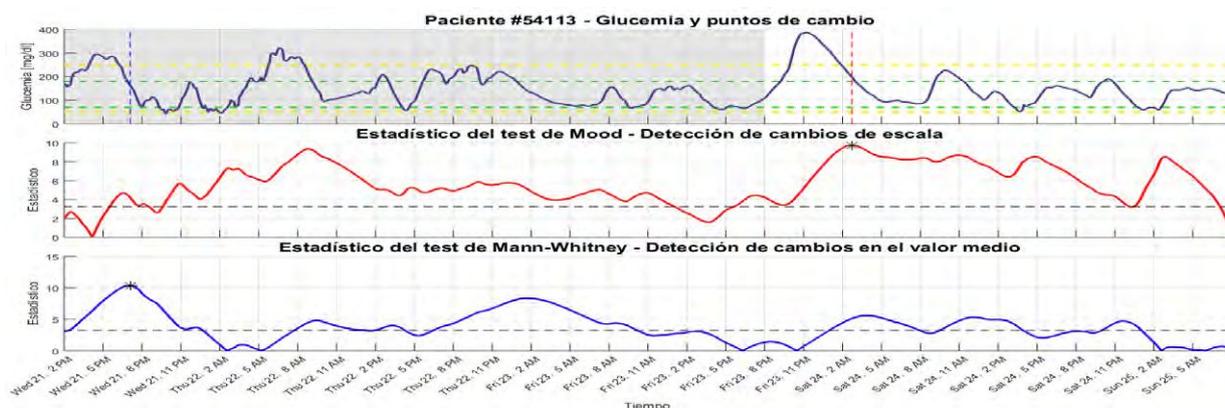
Figura 1: Diagrama en bloques del algoritmo ARG

4. Resultados del análisis de datos

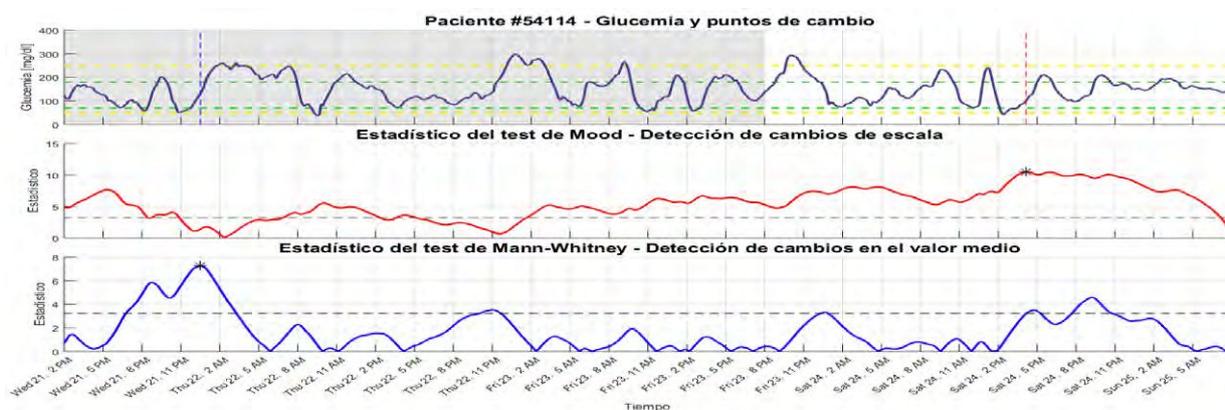
En la figura 2 se muestran los resultados de la aplicación de las herramientas descritas en la sección 2 en los datos de los 5 pacientes del ensayo clínico. Considerando que el cierre del lazo se realizó cerca de las 20 hs del día viernes 23 (demarcado en las figuras por el cambio de color de fondo en la gráfica de glucemia), sería esperable detectar algún cambio en la distribución después de ese punto. Un factor a tomar en cuenta que puede dificultar la evaluación de las herramientas es que, durante el ensayo clínico, el cierre del lazo se hizo justo antes de una comida, por lo cual la mayoría de los pacientes presenta un pico muy pronunciado en el valor glucémico justo después del cierre del lazo, y eso puede dificultar la tarea de las herramientas de CPA. A continuación, se hace un análisis con mayor detalle de los resultados individuales:



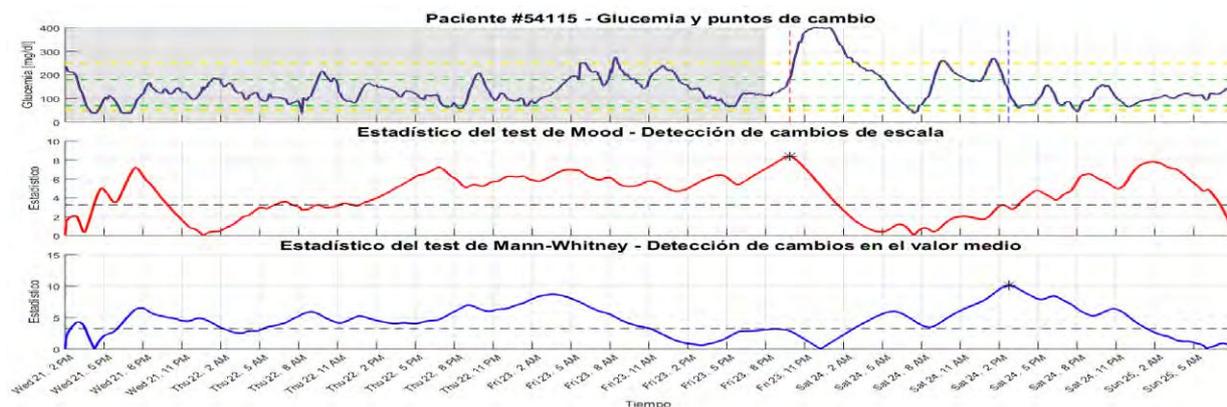
(a)



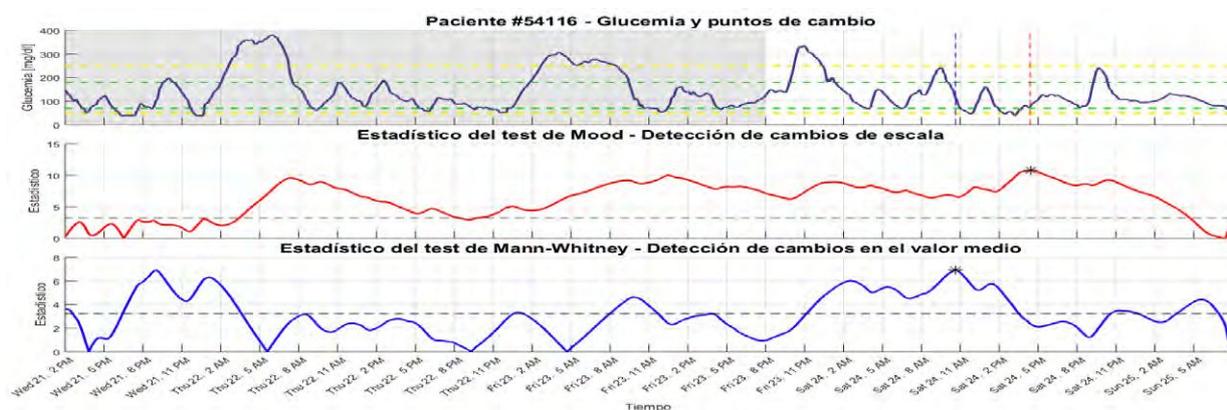
(b)



(c)



(d)



(e)

Figura 2: Resultados individuales de los 5 pacientes en los ensayos clínicos con los estadísticos resultantes para cada test. En el gráfico de glucemia: La línea violeta indica los valores de glucemia medidos con el CGM, las líneas verdes punteadas marcan el rango deseado, las líneas amarillas punteadas marcan el rango aceptable, la línea vertical roja marca el punto de cambio con el test de Mood y la línea vertical azul marca el punto de cambio utilizando el test de MW. El cambio en el color de fondo indica lazo abierto (fondo gris) y lazo cerrado (fondo blanco). En el gráfico del estadístico de Mood: la línea roja es el valor que toma el estadístico, la línea punteada negra es el umbral de significancia ($\alpha = 0,05$). En el gráfico del estadístico de MW: la línea roja es el valor que toma el estadístico, la línea punteada negra es el umbral de significancia ($\alpha = 0,05$). Los asteriscos marcan los puntos máximos del valor de estadístico.

Paciente 54112: La figura 2a muestra la excursión de glucemia del paciente 54112 desde el inicio de la evaluación del lazo abierto, hasta el fin del lazo cerrado. Puede observarse que en el caso de este paciente, el cambio en la escala (test de Mood) se detecta al inicio del ensayo, en el período de lazo abierto. Por definición, esto indicaría que no hubo otro cambio más significativo que este en la excursión de la glucemia de este paciente entre lazo abierto y lazo cerrado. Este resultado es esperable ya que el paciente 54112 presenta diabetes lábil, por lo que resulta complejo mejorar la variabilidad en la glucemia. Sin embargo, si se analizan los valores del estadístico correspondiente al test de Mood se puede reconocer la capacidad de detección de al menos un punto significativo (picos por encima del umbral) luego del cierre del lazo, con un valor de estadístico muy similar al punto detectado al principio. Por otra parte, si se tiene en cuenta el test de MW, se puede observar que el punto de cambio más significativo se produce unas horas antes del cierre del lazo. Considerando el estadístico de MW al momento del cierre del lazo, el cambio sigue siendo muy significativo (alto valor del estadístico). Esto quiere decir que, si bien no se logró reducir significativamente la excursión de este paciente, sí se pudo cambiar su valor medio utilizando la terapia de lazo cerrado.

Pacientes 54113 y 54115: La figura 2b y 2d muestran resultados similares para el test de Mood y una detección de punto de cambio, con este test, posterior al cierre del lazo. En ambos casos, el pico de glucemia posterior al cierre del lazo es pronunciado pero aún así se detecta una reducción significativa en la variabilidad glucémica. Esta detección indica que la reducción esperada en la variabilidad es detectable a través de este test. El test de MW no detecta un cambio significativo en el valor medio de la glucosa

posterior al cierre del lazo para el paciente 54113, pero sí en el caso del paciente 54115, lo cual indicaría un desempeño aún mejor, considerando que hay reducción tanto de valor medio como de excursión.

Pacientes 54114 y 54116: En las figuras 2c y 2e se puede observar que el test de Mood detecta, para cada uno de los pacientes, un cambio significativo en la variabilidad glucémica. Este cambio, si bien es más de 12hs después del cierre del lazo, indicaría que el algoritmo logra modificar la dispersión de los valores de glucosa. El test de MW para el paciente 54114 no detecta cambios significativos luego del cierre del lazo pero para el paciente 54116 indica que existe, como en el caso del paciente 54115, reducción del valor medio, además de la excursión glucémica.

5. Conclusiones y trabajo futuro

En líneas generales, a pesar de que ninguno de los tests dio resultados por sí solo para todos los pacientes, se puede ver que en el caso donde el test de Mood no detectó un cambio significativo luego del cierre del lazo, el test de MW sí lo hizo. Por otro lado, si se observan los valores del estadístico de Mood, se puede notar que aún en el caso donde no se detecta un cambio posterior al cierre del lazo, existen valores pico luego de este punto que indican un cambio en la variabilidad.

Una posible implementación de esta información como medida de desempeño sería considerar no solo los puntos máximos del estadístico sino el hecho de que existan puntos cercanos al máximo luego del cierre del lazo. Dado que el test de Mood detecta cambios en la escala, lo cual en un sistema de AP implica un cambio en la excursión de la glucosa, poder detectar un cambio significativo en ella es una herramienta muy útil para evaluar desempeño. Como líneas de trabajo futuro, sería provechoso estudiar otros tests que permitan evaluar la magnitud y el tipo del variación (reducción, aumento) que se da en el punto de cambio, de manera tal de poder cuantificarlo además de detectarlo. Otra posibilidad podría ser desarrollar CPM propios que se adapten a las necesidades particulares del campo de AP y evaluar las herramientas en una mayor cantidad de datos.

Referencias

- [1] D. M. Maahs, B. A. Buckingham *et al.*, "Outcome Measures for Artificial Pancreas Clinical Trials: A Consensus Report: Table 1," *Diabetes Care*, vol. 39, no. 7, pp. 1175–1179, jun 2016. 1
- [2] D. M. Hawkins, P. Qiu *et al.*, "The Change-point Model for Statistical Process Control," *Journal of Quality Technology*, vol. 35, no. 4, pp. 355–366, oct 2003. 1
- [3] A. N. Pettitt, "A Non-Parametric Approach to the Change-Point Problem," *Applied Statistics*, vol. 28, no. 2, p. 126, 1979. 1, 2
- [4] E. S. Page, "Continuous Inspection Schemes," *Biometrika*, vol. 41, no. 1-2, pp. 100–115, 1954. 1
- [5] A. Sen and M. S. Srivastava, "On Tests for Detecting Change in Mean," *The Annals of Statistics*, vol. 3, no. 1, pp. 98–108, 1975. 1
- [6] G. J. Ross, "Parametric and nonparametric sequential change detection in R: The cpm package," *Journal of Statistical Software*, vol. 66, no. 3, pp. 1–20, 2015. 1
- [7] G. J. Ross, D. K. Tasoulis *et al.*, "Nonparametric monitoring of data streams for changes in location and scale," *Technometrics*, vol. 53, no. 4, pp. 379–389, nov 2011. 2, 3
- [8] D. M. Hawkins and Q. Deng, "A Nonparametric Change-Point Control Chart," *Journal of Quality Technology*, vol. 42, no. 2, pp. 165–173, apr 2010. 2
- [9] E. Schechtman and D. A. Wolfe, "Distribution-Free Tests for the Change-point Problem," *American Journal of Mathematical and Management Sciences*, vol. 8, no. 1-2, pp. 93–119, jan 1988. 3
- [10] D. M. Hawkins and K. D. Zamba, "A Change-Point Model for a Shift in Variance," *Journal of Quality Technology*, vol. 37, no. 1, pp. 21–31, jan 2005. 3
- [11] A. Haidar, "The artificial pancreas: How closed-loop control is revolutionizing diabetes," *IEEE Control Systems*, vol. 36, no. 5, pp. 28–47, Oct 2016. 3
- [12] P. Colmegna, F. Garelli *et al.*, "Automatic regulatory control in type 1 diabetes without carbohydrate counting," *Control Engineering Practice*, pp. 22–32, 2018. 3
- [13] R. Sánchez-Peña, P. Colmegna *et al.*, "Artificial Pancreas: Clinical Study in Latin America Without Premeal Insulin Boluses," *Journal of Diabetes Science and Technology*, 2018. 3