

Caractérisation de la variabilité glycémique par analyse statistique multivariée

Fei ZHENG^{1,2}, Stéphane BONNET¹, Florence FORBES², Manon JALBERT³, Sandrine LABLANCHE³, Pierre-Yves BENHAMOU³

¹Univ. Grenoble Alpes, CEA-LETI, Grenoble, France

²Univ. Grenoble Alpes, Inria, CNRS, G-INP, Grenoble, France

³Endocrinologie Diabétologie Nutrition, CHU Grenoble-Alpes, France

fei.zheng@inria.fr, stephane.bonnet@cea.fr, florence.forbes@inria.fr,
mjalbert@chu-grenoble.fr, SLablanche@chu-grenoble.fr, PYBenhamou@chu-grenoble.fr

Résumé – La variabilité glycémique est une composante importante du contrôle glycémique chez le patient diabétique de type 1. Les systèmes de mesure glycémique interstitielle en continu (CGM) permettent de quantifier précisément les variations intra- et inter-journalières de la glycémie. On propose dans ce papier d'utiliser des techniques de modélisation et d'inférence statistique multivariées pour proposer un nouveau score qui combine statistiquement différentes métriques de la littérature. Une approche basée sur la modélisation d'une population de référence est proposée; le score est simplement la log-vraisemblance d'une observation selon le modèle obtenu. La population de référence est construite à l'aide de données issues d'un essai clinique où des patients sont greffés avec des îlots pancréatiques et retrouvent ainsi un contrôle glycémique. La méthode est comparée à l'annotation de tracés CGM journaliers par des diabétologues et montre un très bon degré de concordance avec ceux-ci.

Abstract – Glycemic variability is a key component in glycemic control in patients with type-1 diabetes. Continuous Glucose Monitoring (CGM) systems allow to quantify accurately the intra- and inter-day variations of interstitial glycemia. This paper aims at using modelization techniques taken from multivariate statistics and machine learning to propose a new score that combines in a statistical way the different metrics of glycemia variability taken from the literature. An approach based on a reference population modeling is described, the score is simply the log-likelihood of an observation according to the reference model. The reference population dataset is built from data collected in a clinical trial where patients have been transplanted with pancreatic islets and thus restore their glycemic control after islet transplantation. The method is compared to the annotations of daily CGM time waveforms by two diabetologists.

1 Introduction

Le diabète de type 1 (DT1) touche 180,000 personnes en France. Il est défini comme l'incapacité pour le pancréas de produire de l'insuline. Dans certains cas, il est proposé une greffe d'îlots de Langerhans pour restaurer cette fonction. La justification d'une greffe se base sur différents critères cliniques.

La variabilité glycémique (VG) est un des critères appréciés par le diabétologue dans son suivi de la maladie, [2]. Plus la VG est élevée et plus le risque d'épisodes hypoglycémiques est élevé. La mesure de la glycémie est aujourd'hui possible de manière continue à l'aide de capteurs électrochimiques CGM. La série temporelle ainsi acquise sur plusieurs jours est analysée par des outils mathématiques pour caractériser cette variabilité, [8]. Différentes métriques ont été proposées mais actuellement aucun réel consensus ne prévaut dans le domaine, [7].

On propose dans ce papier une approche originale pour fusionner différentes métriques et ainsi établir un score de stabilité glycémique. L'idée consiste à caractériser la distribution conjointe de ces métriques pour une population de référence. L'évaluation du score consiste à calculer la vraisemblance d'une

observation selon ce modèle dit de référence. Plus la vraisemblance sera élevée et plus la glycémie associée sera 'proche' de celle du groupe de référence. Cette méthodologie est tout à fait générale. Elle a par exemple été employée pour la localisation et la caractérisation automatique de lésions tumorales dans le cerveau [1].

Le papier décrira la construction de la base de données dite de référence, le calcul des métriques de variabilité glycémique, et la construction du modèle statistique de référence. Dans un second temps, le papier présentera quelques résultats sur des données issues d'une étude clinique et le score de VG sera comparé à des annotations médicales.

2 Analyse multivariée

2.1 Description des données

Les données utilisées dans cette étude sont issues de l'essai clinique TRIMECO. Cette étude visait, en essai contrôlé randomisé, à comparer l'efficacité métabolique de la transplantation d'îlots pancréatiques pour le traitement du diabète de type 1

[6]. 50 patients ont ainsi été sélectionnés entre 2010 et 2013 sur 15 centres français. Pour la plupart des patients, on dispose de données CGM en pré-greffe (M0), 6 mois (M6) et 12 mois (M12) après la première injection d'îlots. Le tableau 1 indique le nombre de patients disponibles à chaque étape ainsi que le nombre de jours total que cela représente. Comme on peut le voir, un nombre de patients inférieur au nombre total est disponible pour chaque cohorte du fait de tracés CGM peu exploitables ou manquants.

Les données sont analysées de manière rétrospective. Pour chaque période analysée, on dispose des données CGM sur plusieurs jours. Le médecin a également établi pour chaque période un score fonctionnel dit de β -score, [9]. Ce score est un score composite, de 0 à 8, qui prend en compte les niveaux de glycémie à jeun, l'hémoglobine glyquée HbA_{1c} , la dépendance à l'insuline ou l'utilisation d'insuline ou d'agents oraux hypoglycémisants et enfin le niveau de C-peptide.

Cohorte	M0	M6	M12
Nombre de sujets	32	35	31
Nombre de jours	217	205	175

La population de référence choisie correspond aux patients dont le beta-score est supérieur ou égal à 6, c-a-d avec un très bon contrôle. Ce choix correspond au cas de patients stables dans le contrôle de leur diabète. Ce choix est justifié dans la mesure où, pour une population non diabétique, on aurait eu un comparateur biaisé car la glycémie d'un sujet sain est une des variables biologiques la plus constante grâce à l'évolution humaine.

Pour chacun des patients sélectionnés, on dispose de plusieurs jours d'enregistrement CGM et on néglige la corrélation entre les observations d'un même sujet. Deux médecins différents ont rétrospectivement classés chaque jour comme "stable", "instable" ou "indéterminé" au vu du tracé CGM correspondant.

2.2 Calcul des métriques

La figure 1 montre un exemple de tracé CGM pour un même sujet mais à deux jours distincts. La variabilité glycémique est naturelle avec par exemple des variations postprandiales. L'incapacité pour le patient à contrôler sa glycémie en lien avec des perturbations endogènes ou exogènes rend ce patient instable et candidat potentiel pour une greffe d'îlots. Les métriques proposées dans la littérature visent à rendre compte de cette variabilité en quantifiant les variations observées dans le signal CGM à l'aide d'une valeur numérique. Chaque jour peut être ainsi résumé à un ensemble de métriques de variabilité glycémique. Pour ce faire, comme la glycémie est mesurée avec une période d'échantillonnage de 5 minutes, on dispose quotidiennement de 144 observations de la variable aléatoire G . On note ΔG la série temporelle obtenue par différentiation de G . On note respectivement $m_G = \mathbb{E}(G)$ et $s_G = \sqrt{\mathbb{V}(G)}$ la moyenne et la déviation standard de cette série temporelle.

Nous considérons 7 métriques issues de la littérature

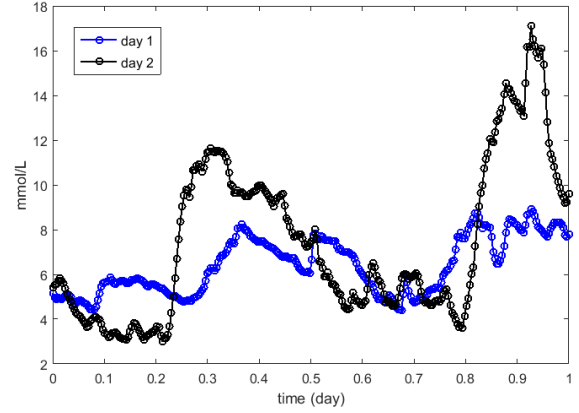


FIGURE 1 – Illustration de la variabilité glycémique

- Coefficient de variation (CV) : $x_1 = s_G/m_G$
- J-index, $x_2 = 0.324(m_G + s_G)^2$
- M-value, $x_3 = \mathbb{E}\{10 \log_{10}(G/IGV)\} + \text{dyn}(G) \cdot 18/20$, où $\text{dyn}(G)$ désigne la dynamique journalière de glycémie et $IGV = 6.67$ mmol/L.
- ADRR (Average Daily Risk Ratio), $x_4 = \mathbb{E}\{r(G)\}$ où la glycémie est convertie à chaque instant en une valeur de risque $r(\cdot)$, cf. [5].
- CONGA, $x_5 = \sqrt{\mathbb{V}\{\Delta G\}}$
- GVP, $x_6 = \mathbb{E}\{\sqrt{\Delta G^2 + \Delta t^2}\} / \mathbb{E}\{\Delta t\} - 1$.
- avgMAGE (Mean Amplitude of Glycaemic Excursion), cf. [11].

Les métriques sont normalisées par un z-score global sur l'ensemble de la base de données.

$$x_k \leftarrow \frac{x_k - \mathbb{E}(X_k)}{\sqrt{\mathbb{V}(X_k)}} \quad (1)$$

Le pouvoir discriminant de chaque caractéristique est évalué à l'aide du critère de Fisher. Il s'agit ici de calculer de manière indépendante pour chaque métrique

$$FS = \frac{(m_1 - m_2)^2}{\sigma_1^2 + \sigma_2^2} \quad (2)$$

avec (m_c, σ_c^2) les moyennes et variances de la métrique pour les observations de la classe c .

2.3 Modélisation de la distribution jointe

Dans la population de référence choisie, un individu correspond à un jour pour un patient de score $\beta > 6$. Nous disposons ainsi de $N = 597$ individus. Pour chacun de ces N jours nous pouvons calculer les 7 métriques évoquées ci-dessus. Certaines de ces métriques étant connues pour être corrélées, nous commençons par faire une analyse en composantes principales (ACP) de la matrice $N \times 7$ des données. L'analyse est ensuite poursuivie dans l'espace réduit ainsi obtenu, sur une matrice notée \mathbf{X} de dimensions $N \times M$ avec $M < 7$. La distribution de probabilité de cette population de N individus est alors modélisée par un modèle de mélange de lois à K composantes.

Les métriques retenues n'exhibent pas en général une distribution gaussienne ni même elliptique. Pour capter des formes plus complexes, nous proposons de considérer des lois plus générales introduites dans [4] sous le nom de loi de Student Multichéelles (\mathcal{MS}). Ces lois sont construites à partir de mélanges infinis de gaussiennes et en conservent la tractabilité tout en autorisant des contours non elliptiques. La densité de probabilité d'une telle loi est notée $\mathcal{MS}(\mathbf{x}|\phi)$ et est donnée par :

$$\mathcal{MS}(\mathbf{x}|\phi) = \prod_{m=1}^M \frac{\Gamma[(\nu_m + 1)/2]}{\Gamma[\nu_m/2] (\lambda_m \nu_m \pi)^{1/2}} \left(1 + \frac{[\mathbf{u}_m^T (\mathbf{x} - \boldsymbol{\mu})]^2}{\lambda_m \nu_m} \right)^{-(\nu_m + 1)/2} \quad (3)$$

où $\phi = \{\boldsymbol{\mu}, \mathbf{U}, \boldsymbol{\Lambda}, \boldsymbol{\nu}\}$ est l'ensemble des paramètres : $\boldsymbol{\mu}$ est un vecteur de localisation (moyenne), \mathbf{U} est une matrice orthogonale dont les colonnes sont notées \mathbf{u}_m , $\boldsymbol{\Lambda}$ est une matrice diagonale dont les éléments sont notés λ_m , et $\boldsymbol{\nu}$ est un vecteur de degrés de liberté notés ν_m . Cette loi peut être vue comme une généralisation de la loi de Student multivariée standard. Voir [4] pour plus de détails. Pour la modélisation de nos données, nous considérons donc un mélange de ces lois :

$$p(\mathbf{x}|\boldsymbol{\psi}) = \sum_{k=1}^K \pi_k \mathcal{MS}(\mathbf{x}|\phi_k) \quad (4)$$

avec $\boldsymbol{\psi} = \{\pi_k, \phi_k\}_{k=1}^K$. L'estimation de $\boldsymbol{\psi}$ se réalise à l'aide d'un algorithme Expectation-Maximization (EM) [3], qui cherche à maximiser de manière itérative la vraisemblance $p(\mathbf{X}|\boldsymbol{\psi})$ de la série d'observations indépendantes $\mathbf{X} = \{\mathbf{x}_1, \dots, \mathbf{x}_N\}$. Les détails de l'algorithme EM sont donnés dans [4]. Le nombre de composantes K du mélange étant à fixer a priori, nous utilisons le critère BIC [10] (*Bayesian Information Criterion*) pour sélectionner la valeur de K qui correspond le mieux aux données observées.

Une fois que le modèle est appris, la log-vraisemblance d'une observation au modèle $\log p(\mathbf{x}|\hat{\boldsymbol{\psi}})$ est utilisée pour comparer une observation au groupe de référence. Plus la log-vraisemblance sera faible, plus cette observation a des chances d'être éloignée de la population de référence et donc plus la variabilité glycémique sera jugée comme instable.

Une difficulté est de décider à partir de quand la log-vraisemblance est considérée comme faible. Pour cela, on peut déterminer un ou plusieurs seuils visant à séparer les scores de log-vraisemblances en 2 ou plusieurs groupes. La méthode la plus traditionnelle consiste à fixer un taux de faux positifs, en général à quelques pourcents, et à calculer le seuil qui satisfait ce taux d'erreur. Ceci revient à calculer un quantile de la loi du log-score vu comme une variable aléatoire. Ce log-score a une expression explicite en fonction de la loi de référence mais sa loi n'est en général pas connue et ses quantiles encore moins. Une solution consiste alors à calculer ces quantiles de manière empirique à l'aide de réalisations simulées du log-score. Pour des taux de faux positifs acceptables (petits) cette démarche

demande souvent de faire appel à la théorie des valeurs extrêmes comme illustré dans [1]. D'autres approches sont également possibles pour déterminer un seuil comme par exemple l'analyse par modèle de mélange des valeurs de log-scores obtenues sur un mélange de sujets stables et instables. L'objectif est alors de faire un *clustering* de ces données et d'utiliser la frontière trouvée entre deux groupes comme seuil. Dans cette étude, nous avons envisagé plusieurs solutions mais pas encore tranché sur une proposition définitive.

3 Illustration

3.1 Analyse des métriques

Les métriques de variabilité glycémique rappelées dans la section précédente sont étudiées dans un premier temps. On montre l'évolution du CV pour les différentes cohortes de M0 à M12. Comme attendu, la greffe d'ilots permet de réduire la variabilité glycémique au sens du coefficient de variation. Le seuil de 36% proposé par [8] pour séparer les diabètes stables et labiles est figuré par une ligne verticale sur la figure 2.

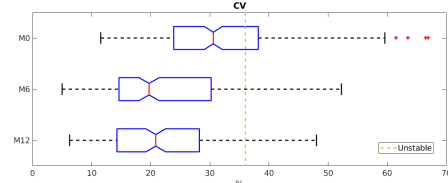


FIGURE 2 – Boxplot du coefficient de variation

On peut calculer le score de Fisher pour chaque métrique en se basant sur les annotations des médecins. Le tableau 2 présente l'analyse de Fisher des métriques journalières. Le tableau contient le vecteur de poids utilisé pour construire la première composante de l'ACP. On peut noter la bonne correspondance entre les deux analyses. La métrique ADRR semble être intéressante dans la discrimination entre les deux populations.

TABLE 2 – Relation entre caractéristiques et classification analysée par différents scores.

Métrique	CV	J-index	M value	ADRR	CONGA	MAGEavg	GVP
PCI coeff (PCA)	0.86	0.85	0.89	0.95	0.92	0.93	0.84
Fisher score	0.57	0.66	0.48	0.85	0.79	0.81	0.45

On montre à la figure 3 la distribution en 2 dimensions des caractéristiques après ACP et on superpose les contours des niveaux de confiance associés aux deux composantes du mélange de distributions \mathcal{MS} ajusté aux données de référence. Chaque point correspond à une variabilité journalière dans le nouvel espace. Pour une meilleure visualisation, on attribue une couleur à chaque point selon sa classe, classe stable (bleu), instable (rouge) ou non-déterminée (gris). Notre choix de population de référence est donc justifié à posteriori par le label des médecins. En effet, notre population de référence est définie à partir du β -score et non pas à partir des annotations des médecins.

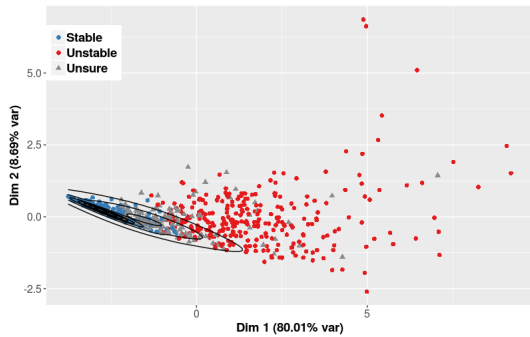


FIGURE 3 – Scatter plot après analyse en composantes principales

Le critère BIC conduit à choisir une ACP à 2 composantes, bien adaptée pour une tâche de classification ultérieure. De même, la notion de distance au modèle de référence semble pertinente pour séparer des glycémies journalières régulières ou labiles.

3.2 Analyse du score de log-vraisemblance

Toutefois, on observe que le modèle de référence contient une fraction de glycémies jugées instables par le médecin. Pour mieux visualiser ce point, on donne à la figure 4 l’histogramme des scores de log-vraisemblance pour les observations de la classe stable (bleu), instable (rouge) ou non-déterminée (gris) respectivement. Une proposition de seuil est indiquée par une barre verticale. On vérifie que les tracés journaliers jugés comme indéterminés par les médecins se retrouvent répartis entre les deux populations. De plus, les jours instables couvrent une partie importante de valeurs de score, démontrant la variabilité biologique des différentes situations observées.

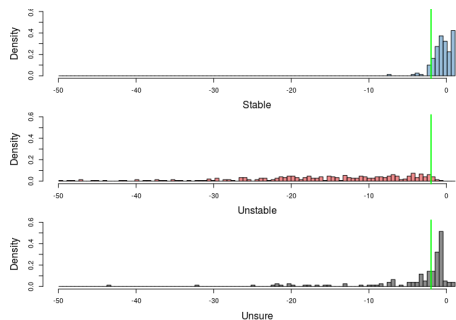


FIGURE 4 – Histogrammes des scores de log-vraisemblance pour les différentes classes de jours, stables, instables et indéterminés de bas en haut.

4 Conclusion

Ce papier a décrit une méthodologie générale pour scorer la variabilité glycémique. La méthodologie est tout à fait générale et peut s’adapter à d’autres métriques ou d’autres problématiques. Les techniques de réduction de dimensionnalité et

de modélisation statistique plus générale que l’approximation Gaussienne sont employées. Le score proposé à la communauté médicale est ainsi défini comme la log-vraisemblance d’une observation selon un modèle de référence. La méthode est suffisamment flexible pour incorporer de nouvelles observations et mettre à jour à intervalles réguliers le modèle de référence. La classification est résumée à la partition des valeurs de log-vraisemblance. Comme évoqué précédemment, plusieurs possibilités existent pour fixer un seuil. Nous allons étudier des approches non-supervisées en se basant sur la forme de l’histogramme de la log-vraisemblance des observations. Une alternative sera d’utiliser un algorithme d’apprentissage supervisé (SVM, régression logistique,...) de manière à optimiser un critère de classification en lien avec les annotations des médecins.

Références

- [1] A. Arnaud, F. Forbes et al. Fully automatic lesion localization and characterization : Application to brain tumors using multiparametric quantitative MRI data. *IEEE Trans on Medical Imaging*, 37(7) :1678–1689, 2018.
- [2] A. Ceriello, L. Monnier, D. R. Owens. Glycaemic variability in diabetes : clinical and therapeutic implications. *Lancet Diabetes Endocrinol*, 2018.
- [3] A. P. Dempster, N. M. Laird et D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, Vol. 39, No. 1. (1977), pp. 1-38.
- [4] F. Forbes et D. Wraith. A new family of multivariate heavy-tailed distributions with variable marginal amounts of tailweight : application to robust clustering. *Statistics and Computing*, 24(6) :971–984, 2014.
- [5] B. Kovatchev, E. Otto et al. Evaluation of a new measure of blood glucose variability in diabetes. *Diabetes Care*, 29(11) :2433-2438, 2006.
- [6] L. Lablanche, MC. Vantighem et al. Assessing islet transplantation compared to insulin therapy in type 1 diabetes : a randomised parallel study. *Lancet Diabetes Endocrinol*, 6(7) :527-537, 2018.
- [7] C. Marling, N. Struble et al. A consensus perceived glycemic variability metric. *Journal of Diabetes Science and Technology*, 7(4) :871–879, 2013.
- [8] L. Monnier, C. Colette, D.-R. Owens. La variabilité glycémique. *Médecine des maladies Métaboliques*, 12 :11 p, 2018.
- [9] E. Ryan, B. Paty et al. β -score. An assessment of β -cell function after islet transplantation. *Diabetes Care*, 28(12) :343–347, 2005.
- [10] G. Schwartz. Estimating the dimension of a model. *Ann. Statist.*, 6(2) :461–464, 1978.
- [11] F. Service, G. D. Molnar, et al. Mean amplitude of glycaemic excursions, a measure of diabetic instability. 19 :644–55, 10 1970.