A Markov chain probability model of glucose tolerance in post gestational diabetes follow up study

Angela Grassi[a,b], Laura Gaetano[c], Giovanni Pacini[a], Alexandra Kautzky-Willer[d], Andrea Tura[a]

[a] *Institute of Biomedical Engineering, Italian National Research Council, Padova, Italy*
[b] *Department of Information Engineering, University of Padova, Padova, Italy*
[c] *Department of Mechanics, Politecnico di Torino, Torino, Italy*
[d] *Department of Internal Medicine 3, Medical University of Vienna, Vienna, Austria*

Abstract

Women with gestational diabetes mellitus (GDM) are at increased risk of developing type 2 diabetes (T2DM). However, the degree of risk and the timing of progression from normal to a pre-diabetic or diabetic state have not been clearly quantified. In this study we analyzed data from a longitudinal study on a group of women with a history of GDM, that were inserted in an oral glucose tolerance test (OGTT) annual screening program and followed up for 5 years after partum. A three state Markov chain model was proposed to represent the dynamics of changes between metabolic states. We used the data to empirically estimate the one-year transition parameters of the model and make predictions about the possibility that women with normal glucose tolerance or impaired glucose metabolism just after partum will develop overt T2DM in three or five years. Results show that subjects with an impaired glucose metabolism few months after partum will hardly (10%) be in the same state after three years. Women with normal glucose tolerance after partum will have a high probability (0.80) to be in the same state three years after.

Keywords:

Gestational Diabetes Mellitus, Markov chains, Follow-up studies, Oral Glucose Tolerance Test.

Introduction

Gestational diabetes mellitus (GDM) is the specific type of diabetes that may develop during pregnancy [1]. GDM is a quite common condition with a prevalence ranging from 1 to 14% of pregnancies, depending on the population under study. GDM represents nearly 90% of all pregnancies complicated by diabetes. The incidence of gestational diabetes is rising throughout the world [2,3], and the same holds for type 2 diabetes [4]. Women with previous gestational diabetes, after delivery (pGDM), often normalize their glucose levels, but they are at increased risk of developing type 2 diabetes [5,6], especially if they have other risk factors (i.e. obesity, hypertension, family history of type 2 diabetes). In this context, it is important to undertake follow up studies that can provide

some insights about the actual risk and timing of possible onset of type 2 diabetes. Data analyzed in this paper refers to a 5 year follow up study of glucose tolerance in pGDM women. The analysis of the data was carried out using a three state Markov chain model to study the transition between different glucose tolerance conditions and making predictions about the onset of type 2 diabetes in the population after three years from the delivery. Markov chains have often been used to model disease's natural history, see [7] for an example of application in the context of diabetes.

Subjects

A group of 116 women participated in the study. They were selectively recruited from the outpatient department of the University Clinic of Vienna. Written informed consent was obtained from all subjects and the protocol was approved by the local ethics committee. For all women, GDM was found in their first pregnancy; repeated pregnancies were excluded.

Table 1 – Characteristics of the studied pGDM women at the basal state (mean ± SE).

	NGT	**IGM**	**T2DM**
Number of subjects	87	21	8
Age (years)	32.91 ± 0.51	34.66 ± 0.93	33.88 ± 2.25
BMI (kg m^{-2})	26.57 ± 0.55	30.97 ± 1.00	31.35 ± 2.15
Fasting plasma glucose (pmol L^{-1})	4.82 ± 0.05	5.89 ± 0.18	6.91 ± 0.35
Plasma glucose at 120 min (pmol L^{-1})	5.88 ± 0.11	9.77 ± 0.36	11.85 ± 0.63

GDM was diagnosed in all the participants according to the criteria of the 4th Workshop Conference of Gestational Diabetes [5], through a 75-g oral glucose tolerance test (OGTT). During pregnancy the GDM women were not subjected to any treatment; they received only dietary and physical exercise advertisements. The study started with a basal OGTT screen-

ing taken 4 to 6 months after partum and then the women were followed, with an annual OGTT screening program, up to 5 years later on. The characteristics of the pGDM women immediately after partum are described in Table 1. Subjects' conditions (or metabolic states) were classified, according to the American Diabetes Association (ADA) 1997 criteria, as: normal glucose tolerance (NGT); impaired glucose metabolism (IGM), including subjects with impaired glucose tolerance (IGT) and/or impaired fasting glucose (IFG); and type 2 diabetes (T2DM). A complete depiction of the data is given by Figure 1, where the distribution of the metabolic states at each year is reported.
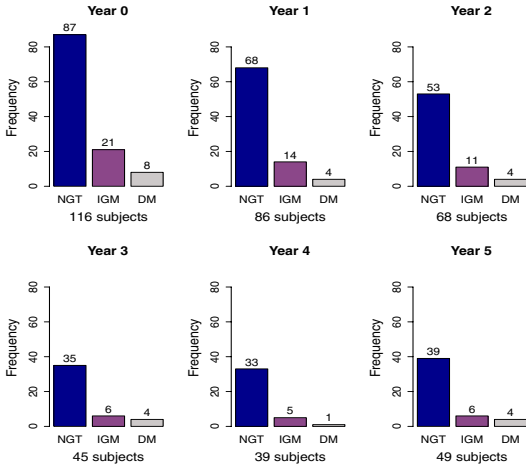

Figure 1- Frequency distribution of the three states (NGT, IGM, T2DM) at each year.

There were two main causes for ending the program: the natural deadline (5 years after partum), or the onset of type 2 diabetes. The data expected from the protocol are not complete, because some participants skipped one or more visits and then reentered the program subsequently.

Methods

Oral Glucose Tolerance Test

After an overnight fast, all women underwent a standard 75-g OGTT. Venous blood samples were collected immediately before glucose ingestion (fasting sample, $t = 0$) and at 10, 20, 30, 60, 90, 120, 150 and 180 min afterwards for glucose, insulin and C-peptide measurements.

Data analysis

Markov chains

A Markov chain (MC) is a simple stochastic process often used to model uncertain phenomena evolving in time. In a stochastic process changes of state (value of the process) are governed by probabilistic laws. The laws describing present and future values of the process in terms of its past state history are called transition laws. A Markov chain, specifically, is characterized by transition laws depending only the most recent past state and not on the whole past history. More formally, consider a stochastic process $\{X_t, t = 0, 1, 2, \ldots\}$ that takes values on a finite set S, called state space. Let the state space size be N, with $S = \{s_1, s_2, \ldots, s_N\}$. If $X_t = s_i$, then the process is said to be in state s_i at time t. Suppose further that the Markov property holds:

$$\Pr(X_{t+1} = s_j \mid X_t = s_i, X_{t-1} = s_{t-1}, \ldots, X_0 = s_0) =$$
$$= \Pr(X_{t+1} = s_j \mid X_t = s_i) \quad \forall s_0, \ldots, s_{t-1}, s_i, s_j; \; \forall t. \quad (1)$$

Such a process is known as a Markov chain, and the above property reads that for a MC the conditional distribution of any future state given the past states and the present state is independent on the past states and depends only on the present state. If $X_t = s_i$ and $X_{t+1} = s_j$, then we can write

$$\Pr(X_{t+1} = s_j \mid X_t = s_i) = p_{ij}^{t+1} \qquad \forall t. \qquad (2)$$

The value p_{ij}^{t+1} represents the probability that the process, when in state s_i at time t will make a transition to state s_j at time $t + 1$. Clearly we have that $p_{ij}^{t+1} \geq 0 \quad \forall i, j$; and $\sum_{j=1}^{N} p_{ij}^{t+1} = 1 \quad \forall i$. We can arrange the transition probabilities into a $N \times N$ matrix, $P^{t+1} = \left\| p_{ij}^{t+1} \right\|$, obtaining the one-step transition matrix of the MC from time t to $t + 1$. As the transition probability between any two states of the chain does depend of t, the chain is called time-inhomogeneous Markov chain.

Once the one-step transition probabilities are known it is possible to compute the conditional probability that the chain at time $t+T$ will be in state s_j given that at time t was in state s_i. As an example, we report the formula for the case $T=3$:

$$\Pr(X_{t+3} = s_j \mid X_t = s_i) = \sum_{k=1}^{N} \sum_{l=1}^{N} p_{ik}^{t+1} p_{kl}^{t+2} p_{lj}^{t+3}, \qquad (3)$$
$$\forall s_i, s_j; \; \forall t.$$

A state of the chain, say s_i, is called an adsorbing state if $p_{ii} = 1$, that is once the process is entered in that state it will never leave it. A state is called transient, if, starting the process from that state, there is a positive probability that the process will never reenter that state. For further details on the basic concepts introduced in this paragraph, we refer the reader to

any book including an introduction on discrete time Markov chain, i.e. [7].

A Markov model for the transitions between metabolic states

Metabolic (or glucose tolerance) states can be defined for each subject - using ADA 1997 criteria - as NGT, IGM, and T2DM. We consider this simplified three state classification, even if the intra-class variability may be large in the IGM class (including both IFG and IGT subjects). Furthermore, we hypothesize that the dynamics of the glucose tolerance derived from the OGTT depends on the past screenings only through the previous year dynamics. Under this hypothesis and recalling that the subject state classification depends on the fasting and on the 2-h plasma glucose level during the OGTT, the future and the past subject's metabolic states are conditionally independent given the state in which the subject is at present, so the Markov property holds. Thus, we model the transitions between subject's glucose tolerance states as a one order time-inhomogeneous Markov chain with categorical state space $S = \{NGT, IGM, T2DM\}$. Let X_t be the metabolic state at year t, if $X_t = s_j$ and $X_{t-1} = s_i$, then

$$\Pr(X_t = s_j \mid X_{t-1} = s_i) = p_{ij}^t \quad \forall t.$$

From the screening protocol, we know that women entering the T2DM state are no longer followed up because they might be treated with anti-diabetic drugs. This information allows us to define a Markov chain model with two transient states (NGT and IGM) and one absorbing state (T2DM). Thus the transition matrix, from year $t-1$ to t, has the following structure:

$$P^t = \begin{Vmatrix} p_{11}^t & p_{12}^t & p_{13}^t \\ p_{21}^t & p_{22}^t & p_{23}^t \\ 0 & 0 & 1 \end{Vmatrix},$$

where state 1 correspond to NGT, state 2 to IGM, and the absorbing state 3 to T2DM. In Figure 2 the graphical representation of the transition probabilities between any couple of states of the chain is reported.

Alternative Markov chain models with longer memory of the chain or with higher number of states might also be considered, but such models would not be suitable in the case of sample size of 116 (that is our case) because they would lead to many transition parameter estimates being zero for lack of empirical data.
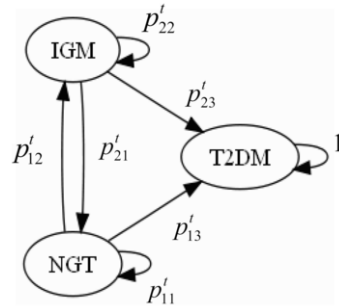

Figure 2 - One-step transitions between the states of the chain; labels on arrows are the transition probabilities.

Estimation of transition probabilities

Let denote by n_{ij}^t the number of women that were in state i at year $t-1$ and are in state j at year t after partum. We can empirically estimate the probability of a woman being in state j at year t after partum given that she was in state i at year $t-1$, $\hat{p}_{ij}^t$, using the formula

$$\hat{p}_{ij}^t = \frac{n_{ij}^t}{\sum_j n_{ij}^t}, \tag{4}$$

that corresponds also to the maximum likelihood estimator. This estimator is showed to be consistent but biased, with bias tending to zero as the sample size increases [9]. For the reader interested in further statistical properties of the transition probability estimator of a MC the paper just cited is a very good reference.

Thus, Equation (4) states that the estimated transition probability from any given state i to state j is equal to the proportion of women that started in state i and ended in state j with respect to all women that started in state i. As we are considering inhomogeneous Markov chains, these probabilities are different for different times.

Results

One year transition matrices

By using the procedure described in the data analysis section we have estimated the one year transition matrices for each possible transition (0 to 1, 1 to 2, 2 to 3, 3 to 4, 4 to 5). Due to the incompleteness of the longitudinal data (some screenings were skipped by the women or they had developed T2DM, thus they were excluded from further analysis) for the last two transition we had only few cases available to estimate some of the transition probabilities. Therefore we decide to stop our analysis at the third year.

As regards the estimated transition matrices for the transitions from 0 to 1, 1 to 2 and 2 to 3, we found the following results:

$$\hat{P}^1 = \begin{Vmatrix} 0.89 & 0.10 & 0.01 \\ 0.33 & 0.47 & 0.20 \\ 0 & 0 & 1 \end{Vmatrix},$$

$$\hat{P}^2 = \begin{Vmatrix} 0.88 & 0.12 & 0.00 \\ 0.36 & 0.45 & 0.18 \\ 0 & 0 & 1 \end{Vmatrix},$$

$$\hat{P}^3 = \begin{Vmatrix} 0.90 & 0.10 & 0.00 \\ 0.50 & 0.10 & 0.40 \\ 0 & 0 & 1 \end{Vmatrix}.$$

As can be appreciated, the transition from year 0 to 1 and from year 1 to 2 lead to approximately the same estimated transition matrix. In other words, we can claim that the differences in the first two transition matrices are negligible from a clinical point of view. In the transition from year 2 to 3 we notice a strong change in the behavior of IGM subjects. The percentage of women passing from IGM to T2DM is doubled with respect to the transitions occurred in the first 2 years after partum.

Predictions

The estimated transition matrices allow us to make predictions about the possibility that women with previous GDM will develop T2DM in three years after partum. Inserting into Equation (3) the estimated transition probabilities above, the ensuing conditional distributions of the metabolic state at year 3, given that the basal state was either NGT or IGM, are

$$\Pr(X_3 \mid X_0 = NGT) = (0.8, 0.1, 0.1),$$

$$\Pr(X_3 \mid X_0 = IGM) = (0.5, 0.1, 0.4).$$

Thus, a woman with pGDM who is NGT after partum, after three years from the delivery will remain NGT with probability 0.8, will become IGM with probability 0.1 and will develop T2DM with probability 0.1. The IGM condition appears as a temporary condition: in fact, in three years we expect that only the 10% of IGM subjects will remain in the same condition, while the remaining 90% will be split between the return to a condition of normal glucose tolerance (50%) and the development of the pathological condition of type 2 diabetes (40%).

Discussion

In this study we presented a Markov chain approach for population predictions about the transition between metabolic states of women with a history of gestational diabetes. The parameters of our model were estimated using data from a follow up study. We found that a woman in NGT condition immediately after partum has high probability (0.8) to be in the same condition three years later and will move to the condition of impaired glucose metabolism or become diabetic with the same

probability: 0.1. The IGM condition appears to be a splitting condition from the normal and the pathological states. This may be due to the fact that in the IGM condition reversibility to the NGT condition is still possible, provided that the subject undergoes a change in the lifestyle, especially in terms of dietary habits and physical fitness. The quantification of the suitable energy intake and the possible degree of exercise are beyond the aims of this report; therefore, no specific suggestion is provided here. On the other hand, if no action is taken, it is known that the worsening of the metabolic condition with possible development of diabetes is likely to happen. Thus, in both cases, the IGM condition does not hold for long periods. Our results appear in fact consistent with this explanation: the 50% of IGM reversing to NGT may be the subjects that did take some actions for their health, whereas the 40% progressing to T2DM may be those not taking any significant action. Another interesting result was that, based on the estimated transition matrices, relevant changes in the probability of transition were noticed only between the second and the third year after partum. This observation indicates that in post-gestational diabetic women screening programs shorter than three years after the delivery may not be sufficient, since possible changes (worsening) in the metabolic condition are unlikely to happen before a period of at least three years after partum.

Conclusion

In this study we used a three state Markov chain model to analyze the dynamics of changes in the metabolic condition of women with a history of gestational diabetes. We found that possible transitions of the metabolic condition might not occur until three years after partum. We also found that the condition of impaired glucose metabolism is not likely to be maintained for a long time. The clinical implication of this finding is that a pGDM woman with IGM (easily detectable with a simple metabolic test) must be carefully followed in terms of lifestyle and perhaps appropriate pharmacological agents to yield in the following years a status of NGT. It will be necessary however the assessment of other parameters (such as for instance, insulin resistance, pancreatic secretion, lipid profile) to precisely identify the target for the intervention. With this regard, future works will be devoted to extend these preliminary results including considerations based on other measured variables (like plasma insulin and C-peptide), as well as on metabolic parameters describing fundamental processes such as insulin sensitivity and beta cell function.

Acknowledgments

References

[1] Bellamy L, Casas JP, Hingorani AD, and Williams D. Type 2 diabetes mellitus after gestational diabetes: a systematic review and meta-analysis. Lancet 2009; 373:1773-9.

[2] Kim C, Newton KM, and Knopp RH. Gestational diabetes and the incidence of type 2 diabetes: a systematic review. Diabetes Care 2002; 25:1862-1868.

[3] Hunt KJ, and Schuller KL. The increasing prevalence of diabetes in pregnancy. Obstetrics and Gynecology Clinics 2007; 34: 173-199.

[4] Haines L, Wan KC, Lynn R, Barrett TG, and Shield JPH. Rising incidence of type 2 diabetes in children in the U.K. Diabetes Care 2007; 30:1097-1101.

[5] Damm P, Kühl C, Bertelsen A, and Mølsted-Pedersen L. Predictive factors for the development of diabetes in women with previous gestational diabetes mellitus. Am J Obstet Gynecol. 1992; 167:607-16.

[6] American Diabetes Association. Gestational Diabetes Mellitus. Diabetes Care 2002; 25(2 Suppl): 94-96.

[7] Kuo, HS, Chang, HJ, Chou, P, Teng L, and Chen TH. A Markov chain model to assess the efficacy of screening for non-insulin dependent diabetes mellitus (NIDDM). International Journal of Epidemiology 1999; 28: 233-240.

[8] Norris JR. Markov Chains, Cambridge Series in Statistical and Probabilistic Mathematics: Cambridge University Press, 1998.

[9] Anderson TW, and Goodman L. Statistical inference about Markov chains. Annals of Mathematical Statistics 1957; 28: 89-110.

Address for correspondence

Angela Grassi
Metabolic Unit ISIB-CNR
Corso Stati Uniti, 4
I-35127 Padova (Italy)
E-mail: angela.grassi@isib.cnr.it