

Single Nucleotide Polymorphisms Associated With Fasting Blood Glucose Trajectory And Type 2 Diabetes Incidence: A Joint Modelling Approach

Mickaël Canouil¹, Philippe Froguel^{1,2}, Ghislain Rocheleau¹

¹Lille University, CNRS, Lille Pasteur Institute, UMR 8199 - EGID, Lille, France

²Department of Genomics of Common Disease, Imperial College London, London, United Kingdom

Contact informations:

CNRS UMR8199 - Institut de Biologie de Lille

1 Rue du Professeur Calmette

BP 245

F-59019 LILLE CEDEX

<http://mickael.canouil.fr>

+33(0)3-20-87-11-33

mickael.canouil@cnrs.fr

Introduction

To optimise use of existing phenotypic data, we propose a joint model [Tsiatis and Davidian, 2004] approach aimed at identifying genetic markers simultaneously associated to temporal trajectories of a trait and an event outcome. Standard formulation of the joint model involves two components: a longitudinal component and a time-to-event component. We illustrate the application of the joint model approach in genetic epidemiology by exploiting the strong link between temporal variation of blood glucose levels (FG) and onset of type 2 diabetes (T2D). Using genotypes assayed with the Metabochip DNA arrays (Illumina) from 4,500 subjects recruited in the French cohort D.E.S.I.R. (Data from an Epidemiological Study on the Insulin Resistance Syndrome), we analysed 124,095 Single Nucleotide Polymorphisms (SNPs) and reexamined previous GWAS findings for some confirmed glycaemia and T2D loci.

Methods: The Joint Modelling Approach

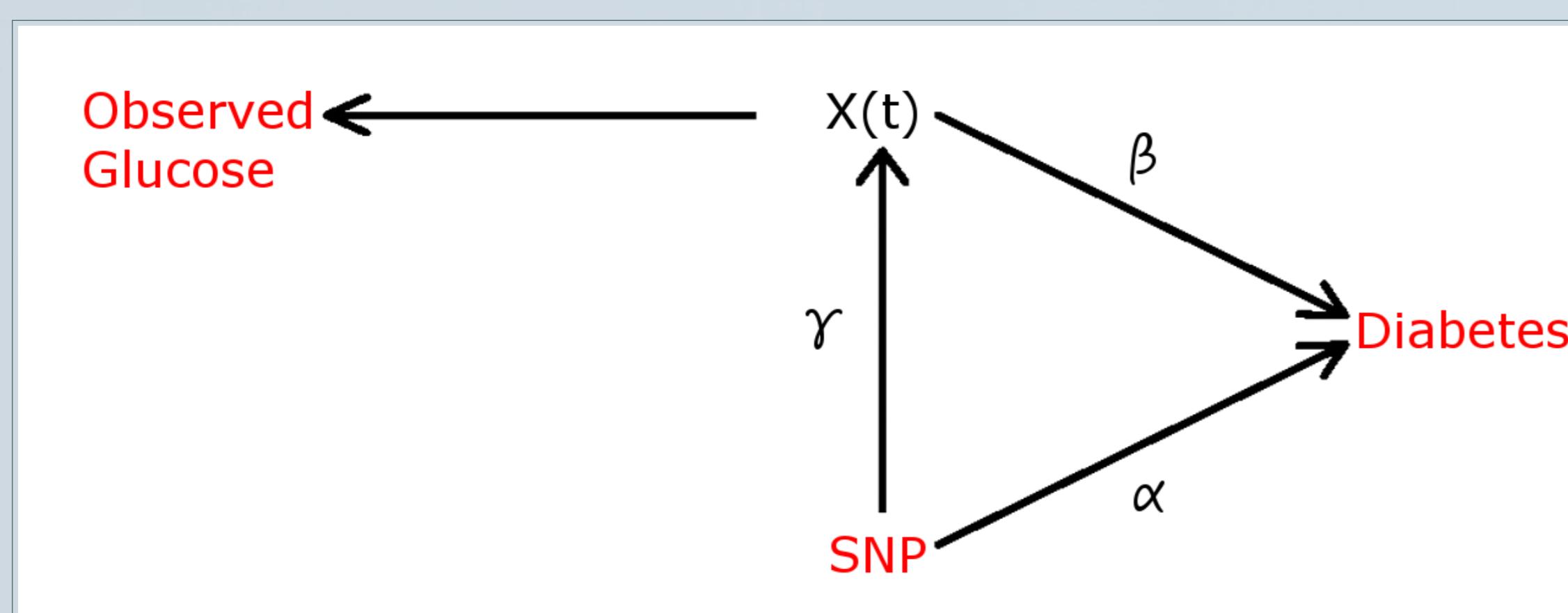


Figure 1: Causal diagram for joint modelling (adapted from Ibrahim et al. [2010]).

The longitudinal component typically consists of a linear mixed model:

$$Y_{ij} = X_{ij} + \epsilon_{ij}$$

Y_{ij} : observed value;

$$X_{ij} = \theta_{0i} + \theta_{1i} \times t_{ij} + \gamma \times SNP_i$$

X_{ij} : true (unobserved) value of the longitudinal variable;

$$\theta \sim \mathcal{N}_2(\mu, \Sigma); \epsilon_{ij} \sim \mathcal{N}(0, \sigma^2)$$

γ : SNP effect on the trajectory function (Figure 1).

Using the Cox model for the time-to-event (survival) component, we define:

$$\lambda_i(t) = \lambda_0(t) \exp(\beta X_i(t) + \alpha SNP_i)$$

$\lambda_i(t)$: hazard function at time t

α : SNP effect on time-to-event (Figure 1).

$\lambda_0(t)$: (unspecified) baseline hazard function.

β : association between the trajectory function and time-to-event (Figure 1).

Application: The D.E.S.I.R. Cohort

As expected, we confirm some findings from GWAS, especially the strong association between FG and some SNP ($\gamma \neq 0$) in G6PC2 and MTNR1B genes (Table 1).

References

- Chen, L. M., Ibrahim, J. G., and Chu, H. (2011). Sample size and power determination in joint modeling of longitudinal and survival data. *Statistics in Medicine*, 30(18):2295–2309.
 Ibrahim, J. G., Chu, H., and Chen, L. M. (2010). Basic concepts and methods for joint models of longitudinal and survival data. *J. Clin. Oncol.*, 28(16):2796–2801.
 Rizopoulos, D. (2010). JM: An R package for the joint modelling of longitudinal and time-to-event data. *Journal of Statistical Software*, 35(9):1–33.
 Tsiatis, A. A. and Davidian, M. (2004). Joint modeling of longitudinal and time-to-event data: an overview. *Statistica Sinica*, 14:809–834.

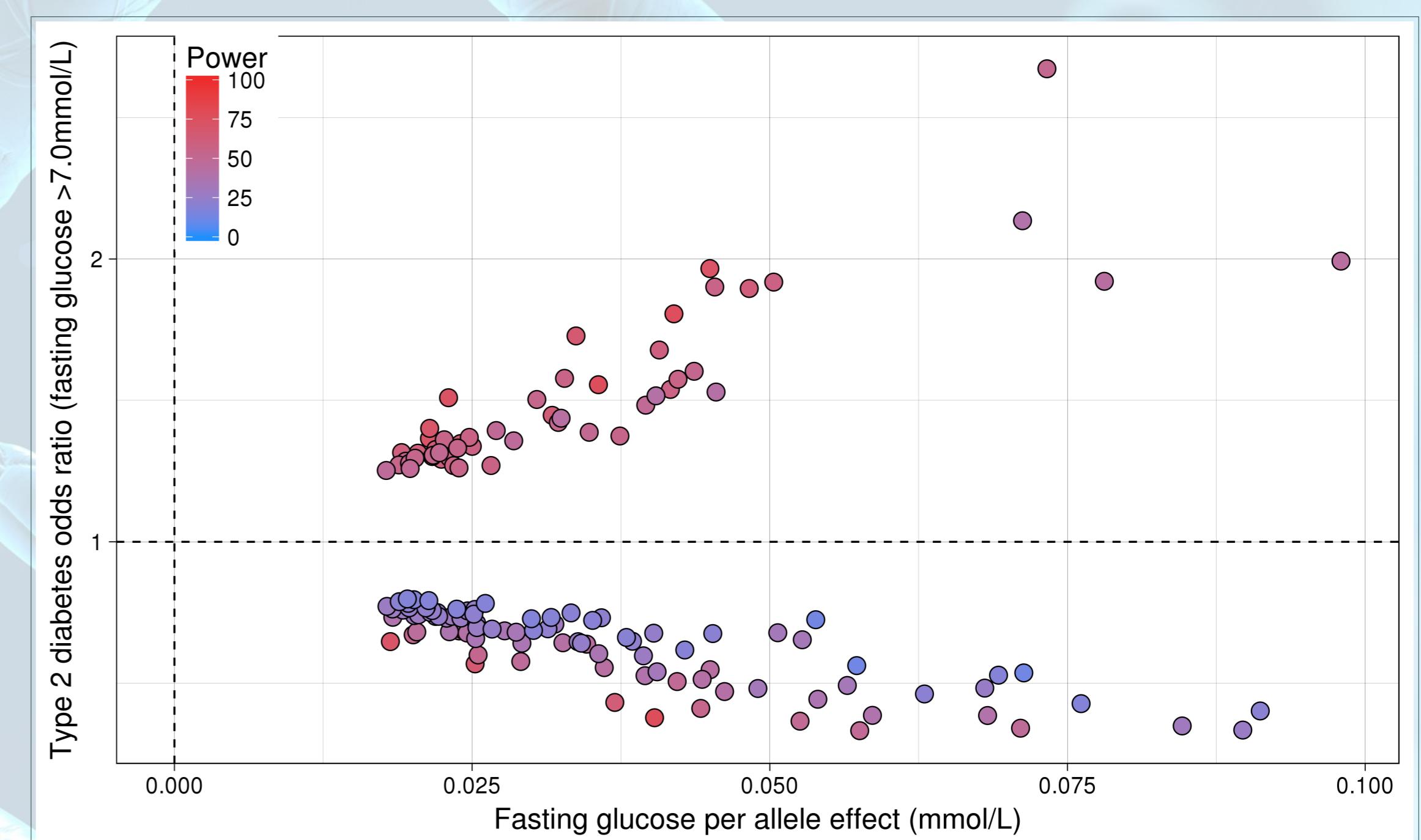


Figure 2: Results from Joint Model (JM [Rizopoulos, 2010]). Power from the formula in Chen et al. [2011] for joint effect.

The association between FG and T2D risk (β) is highly significant ($p < 10^{-40}$) as a result of the T2D diagnostic ($FG > 7\text{mM/L}$).

SNP	α	γ	β
rs1942873 G (MC4R)	0.412	0.023	3.15
rs55899248 C (TCF7L2)	0.291	0.025	3.49
rs10830962 G (MTNR1B)	-0.388	0.0507	3.24
rs12475693 C (G6PC2)	-0.392	0.0452	3.17

Table 1: Parameter estimates from R package JM [Rizopoulos, 2010]. In blue: p-value < 0.05, in red: p-value < 5×10^{-8} .

Statistical Power: Joint Model VS. Cross-sectional Model

SNP	α	γ
rs1942873 G (MC4R)	47.7 / 71.6	64.9 / 52.8
rs55899248 C (TCF7L2)	66.6 / 35.3	91.6 / 81.0
rs10830962 G (MTNR1B)	55.1 / 32.1	60.8 / 47.1
rs12475693 C (G6PC2)	76.2 / 56.5	56.7 / 44.5

Table 2: Statistical power through simulations. In blue, when Joint Model (JM) is more powerful than Cross-sectional Model (CM) displayed JM/CM. (Type 1 error are $5.81\% \pm 0.83$ / $5.47\% \pm 0.64$ for JM and CM.)

Highlights

- The joint modelling approach showed higher power to detect effects (α and γ) than the usual linear or logistic regression.
- SNPs in MC4R or TCF7L2 showed simultaneous associations with FG trajectory and T2D risk.

