NMR-based prediction of cardiovascular risk in diabetes.

Ronan Roussel, France Mentré, Nadia Bouchemal, Samy Hadjadj, Michel Lièvre, Gilles Chatellier, Joël Menard, Xavière Panhard, Anne Le Hénanff, Michel Marre, et al.

NMR-based prediction of cardiovascular risk in diabetes

To the editor:

It has been proposed that nuclear magnetic resonance (NMR) plasma analysis can improve lipoprotein subclass discrimination and predict coronary artery disease (CAD)[1]. In the recent paper by Kirschenlohr *et al.*[2], it was concluded that proton NMR spectroscopy shows only weak discrimination for detecting angiography-defined CAD. An even more important question is whether proton NMR spectroscopy is able to predict cardiovascular events. Indeed, the mortality and morbidity for CAD in type 2 diabetes patients is two to four times higher than in nondiabetic subjects[3], making cardiovascular disease the first cause of death and disability amongst this population. Moreover, CAD is often undetected in diabetic patients. The traditional risk factors do not fully explain the level of the cardiovascular risk in these patients, and new tools are being developed to better identify high-risk patients, who need invasive coronary evaluation and/or aggressive treatments. Indeed, the strategy must have a high degree of accuracy for detecting CAD, owing to the potential morbidity associated with angiographic procedures.

To assess whether NMR spectroscopy can predict cardiovascular disease, we performed a case-control study nested in the placebo arm of a larger clinical trial, DIABHYCAR (ref. 4). DIABHYCAR, a randomized clinical trial, involved high-risk type 2 diabetic patients who were followed for 4 years and who, during that period, had fatal or nonfatal acute myocardial infarction. The study protocol was approved by the ethical committee of Angers University, France. All participants provided written informed consent. This nested case-control study included 190 patients (150 men and 40 women). We acquired proton NMR spectra of samples collected from patients upon their inclusion in DIABHYCAR. Patients of the ‘case’ group were defined as those who presented either a fatal or nonfatal acute myocardial infarction or who underwent sudden death during the 4-year follow-up. Control patients were defined as patients not presenting any cardiovascular events. Control patients were matched to the case patients for age ± 2 years, sex and geographical origin, and their follow-up lasted for at least as long as that of the case patients (**Table 1**). We obtained blood samples after an overnight fast, using tubes without anticoagulant. We stored sera at –80 °C before NMR analysis. We acquired spectra at 500 MHz (Varian Inova) with a 1.5-s presaturation of the water signal[5]. The acquisition sequence consisted of a 4-s-long relaxation delay (including water suppression delay) and a 90° excitation pulse. The signal was acquired on 16,384 data points. Fourier transform was applied without zero filling, and

with a 0.1 Hz line-broadening factor. The chemical shift scale was calibrated on the fumaric acid signal at 6.53 p.p.m.

We used two distinct methods to analyze the spectrum: determination of lipid subclasses by deconvolution in the frequency domain (method 1) and whole-spectra analysis (method 2). More specifically, for method 1, lipoprotein deconvolution was performed with WINNMR software as described previously[5]. Briefly, the region between 0.5 and 1.1 p.p.m. was deconvoluted into 11 Lorentzians whose areas were used for multivariate analysis. In method 2, the spectra were divided into bins of width 0.04 p.p.m., defined from 10 to 0.4 p.p.m., excluding the water signal between 4.9 to 4.5 p.p.m. This data reduction resulted in 207 elements for each spectrum[1].

To the results of each method, we then applied orthogonal signal correction (OSC) followed by partial least-squares discriminant analysis (PLS-DA), a similar approach to the one used previously[1]. This statistical analysis was performed twice: first on the complete set of 190 patients of the nested case-control study and then, to evaluate the model predictability on an independent population, on a randomly chosen training group (70% of the subjects) and validation group (30%). To classify the patients of this validation group, we applied the model obtained from the OSC/PLS-DA analysis of the training group to the spectra of the validation group.

The OSC/PLS-DA model, which was based on the 11 elements (Lorentzian) obtained by lipid signal deconvolution of the 190 subject sera, gave predictions that were 58% correct. When individuals were randomly assigned to either the training group (70% of subjects) or the validation group (30%) and the model applied, the predictions were correct for 62% of patients in the former group and for only 53% of patients in the latter group. This is barely more than the 50% that would be obtained through random prediction. When whole-spectrum analysis was conducted on the 207 elements according to a previously described protocol[1,2], the correctness of the prediction was 82%. However, it was 87% on the training group but only 53% in the independent validation group.

Our results confirmed that, as suggested by Kirschenlohr *et al.*[2], the achievable accuracy of the prediction of the NMR plasma results makes the technique unfeasible for clinical use. Kirschenlohr *et al.*[2] obtained a low prediction accuracy predictability when they compared NMR data analyses to angiographic findings performed at the same time[2], but, as shown here, it is even worse for the prediction of clinical events from baseline samples in type 2 diabetic patients with high cardiovascular risk. We were not able to achieve patient

discrimination in the validation sample by using whole-spectrum analysis as described by Brindle *et al.*[1], nor could we obtain it from Lorentzian lipid deconvolution. This may be due to over-fitting of the data, as Kirschenlohr *et al.*[2] have suggested, but 190 samples represent a relatively large number of patients when compared to the population size used by Brindle *et al.*[1] We applied the same experimental and analytical protocol and similar statistical analysis to spectra obtained for 120 hyperlipidemic subjects in samples collected before and after statin therapy. These analyses were part of a short-term clinical trial involving hypercholesterolemic men, as described in Le Moyec *et al.*[5] The ability to discriminate samples before and after treatment was better with OSC/PLS-DA of Lorentzians areas than with whole-spectra analysis (correctness of the prediction in the 30% validation sample was 86% and 72%, respectively), suggesting that only very stringent characteristics may be discriminated by NMR analysis, if the size of the population is adequate with respect to the number of parameters to be analyzed.

Our work was not a replication study but was aimed at evaluating the relevance of this technique for the prediction of severe cardiovascular events in high-risk type 2 diabetic patients. Our conclusions are very similar to those of Kirschenlohr *et al.*[2] As these authors concluded, the discriminatory capacity of multivariate analyses of NMR spectroscopic data is not high enough to render the technique feasible for clinical use. With respect to cardiovascular risk prediction, global statistical analyses, including classical biochemical and clinical risk factors at baseline, should be considered in addition to NMR data.

Ronan Roussel[1–3], France Mentré[4,5], Nadia Bouchemal[6], Samy Hadjadj[7,8], Michel Lièvre[9], Gilles Chatellier[10], Joël Menard[10], Xavière Panhard[4,5], Anne Le Hénanff [4,5], Michel Marre [1–3] & Laurence Le Moyec[6], for the DIABHYCAR Study Group

[1]Institut National de la Santé et de la Recherche Médicale, U695, 75018 Paris, France. [2]Université Paris 7-Denis Diderot, Unité de Formation et de Recherche de Médecine Xavier-Bichat, 75018 Paris, France. [3]Assistance Publique-Hôpitaux de Paris, Hôpital Bichat, Département d'Endocrinologie, Diabétologie et Nutrition, 75018 Paris, France. [4]Institut National de la Santé et de la Recherche Médicale, U738, 75018 Paris, France. [5]Assistance Publique-Hôpitaux de Paris, Hôpital Bichat, Département d'Epidémiologie, Biostatistique et Recherche Clinique, 75018 Paris, France. [6]Unité Mixte de Recherche-Centre National de la Recherche Scientifique 7033, Université Paris 13 and Université Paris 6, 93000 Bobigny, France. [7]Hôpital Universitaire, Department d'Endocrinologie et Diabétologie, 86021 Poitiers, France. [8]Institut National de la Santé et de la Recherche Médicale, Equipe de Recherche Mixte 324, 86021 Poitiers, France. [9]Université Claude Bernard Lyon 1, Faculté de Médecine Laënnec, 69000 Lyon, France. [10]Assistance Publique-Hôpitaux de Paris, Georges Pompidou European Hospital, Department of Biostatistics, 75015 Paris, France.

e-mail: ronan.roussel@bch.aphp.fr

COMPETING INTERESTS STATEMENT

The authors declare no competing financial interests.

1. Brindle, J.T. *et al. Nat. Med.* **8**, 1439–1444 (2002).

2. Kirschenlohr, H.L. *et al. Nat. Med.* **12**, 705–710 (2006).

3. Wei, M., Gaskill, S.P., Haffner, S.M. & Stern, M.P. *Diabetes Care* **21**, 1167–1172 (1998).

4. Marre, M. *et al. Br. Med. J.* **328**, 495–500 (2004).

5. Le Moyec, L., Valensi, P., Charniot, J.C., Hantz, E. & Albertini, J.P. *NMR Biomed.* **18**, 421–429 (2005).

Table 1 Characteristics of subjects

	Cases	Controls	*P*-values
Number of subjects	95	95	–
Male/female	75/20	75/20	–
Age (years)	69 ± 9	69±9	0.97
BMI (kg/m²)	29.3 ± 4.6	29.2 ± 4.7	0.90
Active smoker	16%	19%	0.61
Percentage of patients with myocardial infarction history	7%	4%	0.35
Fasting glycemia (mM)	9.9 ± 3.1	9.2 ± 2.8	0.12
Glycated haemoglobin (HbA1c, %)	7.8 ± 1.8	7.7 ± 1.6	0.80
Total cholesterol (mM)	5.8 ± 1.2	5.6 ± 1.0	0.28
HDL (mM)	1.2 ± 0.3	1.3 ± 0.4	0.30
LDL (mM)	3.6 ± 0.9	3.4 ± 0.8	0.26

Percentage of patients taking cholesterol-lowering drugs	33%	36%	0.65
Systolic blood pressure (mm Hg)	146.5 ± 14.0	143.8 ± 14.2	0.19
Diastolic blood pressure (mm Hg)	82.8 ± 7.9	81.9 ± 8.8	0.48
Hypertension	60%	54%	0.38

Clinical and biological characteristics (mean ± s.d. or frequency) of the type 2 diabetic patients included in the case-control study nested in the placebo arm of the DIABHYCAR trial, and comparisons between the patients who either presented a fatal/nonfatal acute myocardial infarction or underwent sudden death during the 4-year follow-up, and patients who did not present any such event (*P*-values).