Accepted in ***Neuropsychobiology***, special issue on "*Dimensional Psychiatry*"
Guest Editors: Manuel Morrens & Sebastian Walther

A brain imaging-based diagnostic biomarker for periodic catatonia:

Preliminary evidence using a Bayesian approach

Foucher Jack René[*,1,2], de Billy Clément[1,2], Jeanjean Ludovic Christophe[1,2], Obrecht Alexandre[2,3], Mainberger Olivier[1,2], Clauss Julie Marie Estelle[3,5], Schorr Benoit[3,4], Lupu Madalina Carmen[6], de Sousa Paulo Loureiro[1], Lamy Julien[1], Noblet Vincent[1], Sauleau Erik André[7], Landré Lionel[1], Berna Fabrice[3,4]

1. ICube – CNRS UMR 7357, Neurophysiology, FMTS, University of Strasbourg, France.
2. CEMNIS – Noninvasive Neuromodulation Center, University Hospital Strasbourg, France.
3. Physiopathologie et Psychopathologie Cognitive de la Schizophrénie – INSERM 1114, FMTS, University of Strasbourg, France.
4. Pôle de Psychiatrie, Santé Mentale et Addictologie, University Hospital Strasbourg, France.
5. SAGE – CNRS UMR 7363, FMTS, University of Strasbourg, France.
6. Pôle de psychiatrie G02, EPSAN, Brumath, France.
7. Biostatistical laboratory, iCube – CNRS UMR 7357, University of Strasbourg, France.

Abstract:

Periodic catatonia (PC) is a psychomotor phenotype with a progressive-remitting course. While it can fit any disorder-diagnosis of the schizoaffective spectrum, its core features consist in a mix of hypo- and hyperkinesia resulting in distortions of expressive movements such as grimacing and parakinesias. The replication of a cerebral blood flow (rCBF) increase in the left supplementary motor area (L-SMA) and lateral premotor cortex (L-LPM) in acute and remitted PC-patients makes the case that these could be used as diagnostic biomarkers.

In this proof-of-concept study, 2 different MRI sequences were repeated on 3 separated days to get reliable measures of rCBF in 9 PC and 26 non-PC patients during different cognitive tasks. Each patient was compared to 37 controls. In L-SMA [-9; +10; +60] and L-LPM [-46; -12; +43] a test was positive if the t-value > 2.02 ($\alpha < 0.05$; two-tails).

The measures had good analytical performances. Regarding the tests, their sensitivities and specificities were significantly different from chance level on both measures, except for L-SMA's sensitivities. When combining all the tests, among regions and methods, sensitivity = 98% (95% credible interval 76-100%) and specificity = 88% (72-97%). Bayesian inference of its negative predictive values for PC were > 95% regardless of the context, while its positive ones reached 94%, but only when used in combination with clinical criteria. The case-by-case analysis suggests that non-PC patients with neurological motor are at risk to be false positive.

Key words: Schizophrenia; Endogenous Psychosis; Periodic Catatonia; Cataphasia; Regional Cerebral Blood Flow; Functional Brain Imaging.

---

[*] **Corresponding Author:** Foucher, Jack René. CEMNIS (UF 4768) - Centre de neuromodulation Non Invasive de Strasbourg, Hôpitaux Universitaires de Strasbourg, 1 place de l'hôpital, BP 426, 67 091 Strasbourg cedex. Mail : jack.foucher@icube.unistra.fr / jack.foucher@chru-strasbourg.fr. Tel: +33 3 88 11 69 21. **Clinical trial:** NCT02868879, NCT02863380 and NCT03116425.

1. Introduction

At the turn of the century, doubt emerged about the DSM research program initiated two decades earlier with its third version. This triggered a surge of creative thinking ending with the proposal of many alternatives such as DSM-5 section 3 [1], HiTOP [2], RDoCs [3] and SyNoPsis [4]. These embrace very different epistemological assumptions in defining dimensions rather than categories like disorders in DSM-5 section II and ICD-11 [5]. Yet the latter were designed as proxies for ***phenotypes***, a concept that is backed up to the one of ***diseases*** as defined by the classical ***biomedical paradigm***. Yet, the DSM was not a fair implementation of the biomedical research program. Hence, its inadaptation to basic psychosis research cannot be interpreted as the failure of the biomedical framework (see supplementary material). At the opposite, the ***Wernicke-Kleist-Leonhard*** (WKL) research program [6] embraces the phenotype-disease epistemological foundations, which - it is worth reminding it - the concept of diagnostic biomarker is deeply anchored to. Recent development in brain imaging lead us to test for such diagnostic biomarker in ***periodic catatonia*** (PC), one of the WKL-phenotypes with the highest construct validity and for which a disease-model is already supported by some evidence.

As out of the ICD-DSM, PC is a poorly investigated phenotype with only 125 occurrences in PubMed (in March 2019). Only the points relevant for the paper are reminded in the following lines, but the interested reader could find a more detailed account in supplementary material. PC has a relatively high ***prevalence***, accounting for 10% of patients admitted for psychotic symptoms in Europe [7–10]. It overly has a ***progressive-remitting*** course (98.5%), with episodes having bipolar features in 86% of the cases on the long run [7]. Relapses lead to the progressive build-up of a residual state consisting in mild psychomotor anomalies, abulia and affective blunting [11]. The ***construct validity*** of PC is high. First, its diagnosis is reliable, even when compared to other WKL-catatonic phenotypes called system catatonias: inter-rater agreement $\kappa = 0.93$ [12]. It is also life-long consistent with $\kappa = 0.79$ at 30 years interval [13]. Second, PC also demonstrates interesting differential validities, especially relative to system catatonias. Its familiality is significantly higher: 27% of PC first-degree relatives are affected by the same phenotype without cross liability, vs. less than 5% in system catatonias [9]. PC is also significantly more responsive to antipsychotics [14,15], benzodiazepines [16] and electroconvulsive therapy [17], than system catatonias [14–17].

A plausible biomedical model for PC could be a defective lateral inhibition in sensorimotor and premotor cortices, i.e. supplementary motor area (SMA), or lateral premotor (LPM). This was inspired by PC-specific psychomotor manifestations, their response to GABAergic intervention and the report of a reduced binding for $GABA_A$ receptors radioligand in the left sensorimotor cortex of schizophrenic and affective patients having ≥3 catatonic symptoms (both ICD disorders can be diagnosed in PC-patients) [18]. New evidence came from the increase of cerebral blood flow (rCBF) in the SMA and the left LPM (L-LPM), when patients with acute DSM-schizophrenia with ≥2 catatonic symptoms were compared to those with none, whether current or past [19]. Considering the 10% prevalence of PC relative to the 20% prevalence of patients with ≥2 catatonic symptoms in the same population [20], it can reasonably be assumed that PC might have chiefly contributed to these results. This interpretation was further supported by a study specially focusing on the PC-phenotype, but this time during its residual state. Again, a large ***left-sided*** network including the sensorimotor cortex, L-LPM and L-SMA was hyper-perfused regardless of the comparison group, whether it was healthy controls or non-PC patients with schizophrenia spectrum diagnosis under the same medication regimen [21]. This suggests that left premotor hyper-perfusion, and perhaps defective lateral inhibition, might be ***trait*** characteristics of PC.

This proof-of-concept trial attempts to ***translate*** the disease-model for PC, i.e. defective lateral inhibition in motor-premotor cortices, into a ***diagnostic biomarker***. Beyond the ***trait*** characteristic of the L-SMA and L-LPM hyper-perfusions and their possible connection with PC pathophysiological model, this study takes advantage of the good analytical performances of rCBF measurements using arterial spin labeling or ASL [22]. To further increase their reproducibility, these were made independent of the cognitive state and the measurement method: each participant took part in 3 MRI sessions performed on separate days, during which 2 kinds of MRI sequences were acquired under various cognitive tasks. Patients with PC (n=9) and other psychiatric

diagnosis (n = 26) were compared to a group of 37 group-matched controls to get the departure of rCBF from the norm in the two relevant regions expressed in t-value. ***Bayesian statistics*** were used to assess the discriminating performances of the biomarkers. These were more adapted to our small samples and can be intuitively interpreted as a degree of belief in decision making, allowing the computation of positive and negative predictive values. The present estimation of the biomarkers performance characteristics started from a "null hypothesis", i.e. ***non-informative prior***, in order to be interpreted as an independent replication from a more familiar "frequentist" perspective. But these estimates could be used as priors in the analysis of an ongoing follow-up study. Methods and results are complying at best to PRISMA-DTA statement for reporting on diagnostic test [23].

Building on the "defective sensorimotor lateral inhibition hypothesis" for ***PC***, we wanted to test 1) whether ***L-SMA and L-LPM hyper-perfusions*** could be ***reliably*** assessed in each single individual patient and 2) whether they have good enough ***discriminating performances*** to support further investigations on their use as ***diagnostic biomarkers***.

2. Material and methods

2.1. Participants

Patients and controls were recruited in two Strasbourg's expert centers for schizophrenia [24] and resistant depressions by WKL-trained psychiatrists in the context of two studies: iADAPT (NCT02863380) and RETONIC (NCT03116425). Both trials first aimed at demonstrating the superiority of personalized repetitive transcranial magnetic stimulation (rTMS) based on patient-specific brain perfusion anomalies. The analysis of the latter as potential diagnostic biomarkers was a secondary aim. iADAPT focused on treatment of resistant depressions and RETONIC on the treatment of chronic catatonias. The control group from which norms of rCBF measures could be computed was recruited at the Strasbourg's center for non-invasive neuromodulation in the context of another trial, i.e. Connect C3 (NCT02868879), matching them for age, gender and level of education with the group of schizophrenic patients included in this trial [21]. Nine further controls above 50 years of age (5 females) were added to match the patients of the iADAPT trial. Exclusion criteria for controls were left-handedness and significant personal or family neurological or psychiatric history including current drug abuse (except nicotine). Controls were clinically screened with DSM-5 self-rated level 1 cross-cutting symptom and excluded if above 1 on any item [1]. All studies' protocols comply to the declaration of Helsinki and were approved by the local ethic committee.

Commune inclusion criteria for patients were: being 18 - 65 years old, outpatient under stable medication regimen for ≥6 weeks, not being under involuntary commitment, nor having contraindication for MRI and rTMS. For iADAPT, patients had to fulfill the criteria for ***resistant depression***, i.e. DSM-5 criteria for ongoing major depressive episode unresponsive to ≥2 well-conducted courses of antidepressant and ≥1 psychotherapy. Importantly, as high-frequency rTMS delivered on the left dorsal-lateral prefrontal cortex was one of the treatment arm in this study, ***patients with significant psychotic manifestations were excluded*** to avoid the worsening of those symptoms [25]. Patients included in RETONIC had to fulfill one of the WKL-catatonic phenotypes, i.e. PC or one of the 7 system catatonia phenotypes. PC was diagnosed according to operationalized criteria for research (OCR-PC, suppl. material in ref [21]) while system-catatonia phenotypes were consensually diagnosed by two WKL trained psychiatrists according to the reference book (see supplementary material for more details) [26]. Each participant signed an informed consent as well as caregivers when appropriate. Only controls received a compensation, taking account that patients have benefited of a personalized treatment through their participation.

2.2. Clinical assessments

For each participants, handedness was appraised by the Edinburgh inventory [27] and an IQ was estimated using the French National Adult Reading Test or fNART [28]. For each patient, DSM-5 and WKL diagnoses were based on clinical assessment. The global assessment of functioning (GAF)[29] was evaluated, antipsychotics doses were converted to olanzapine equivalent (OLZ eq)[30], and benzodiazepines doses to diazepam equivalent (DZP eq)[31]. The 8 dimensions of the Clinician-Rated Dimension of Psychosis Symptom Severity (CRDPSS)[1] were rated for all patients who further filled the self-report version of the Inventory of Depressive Symptomatology (IDS-SR)[32]. For PC-

patients only, the psychotic psychopathology was assessed using the positive and negative syndrome scale (PANSS)[33], catatonic symptoms were evaluated using the Bush and Francis Catatonia Rating Scale (BFCRS)[34] and depressive symptoms were further assessed with the Calgary Depression Scale for Schizophrenia (CDSS)[35].

2.3. Imaging Protocol

All studies started by an automatic volume placement sequence in order to expand inter- and intra-subject repeatability (AutoAlign Head sequence, SIEMENS). All trials used the same cognitive tasks and the same functional imaging protocols. Each participant underwent 3 functional MRI sessions using a 3T Verio scanner (Siemens Healthcare, Erlangen, Germany) during which 2 different arterial spin labeling (ASL) sequences were performed: ASL-9.7 and ASL-21. Both are based on a PICORE Q2TIPS design (QUIPSS II with Thin-slice TI1 Periodic Saturation) with TR of 3000ms, TI1 = 600ms, TI2 = 1325ms, 90 ° flip angle, 4 x 4 x 4 mm resolution. Yet, the two sequences differed in TE: 9.7 ms (ASL-9.7) and 21ms (ASL-21) and were previously showed to have better analytical performances than BOLD fMRI activation measures [22]. A field map was acquired in each session and an anatomical 3D T1-weighted (MP-RAGE) volume was collected during the first one (0.7 x 0.7 x 0.7 mm resolution).

2.4. Cognitive tasks

ASL-9.7 were performed twice in resting state condition and once in a video watching condition (101 scans, 5 min each). ASL-21 were performed once in resting state condition, once in video watching condition (101 scans, 5 min) and once during a functional localizer protocol (405 scans, 20 min). For the resting state, participants were instructed to keep their eyes closed and to refrain from falling asleep. Videos were about real-life fall or car accidents without casualties. The functional localizer protocol consisted in six different tasks: mental calculation, reading, episodic memory, working memory, emotional clip and inner-feelings judgment tasks [21]. Each task was separated by short periods of rest and required participants to answer by pressing one of a three-button pad.

2.5. Image preprocessing and single subject analysis

All pre-processing and analyses were performed with Matlab12 (Mathworks, Natick, MA, USA) using the SPM12 Toolbox [36]. ASL-contrast was computed by subtracting tagged EPI to the non-tagged volumes which could be converted into equivalent rCBF after correction for the TE [37]. Functional volumes were corrected for magnetic distortions, spatially normalized using the deformation field obtained on the 3D-T1 image and smoothed (6 x 6 x 6 mm). The statistical analysis independently compared each patient to controls for each ASL-method (ASL-9.7 and ASL-21), for both active and rest conditions. The reasons for combining multiple measurements under multiple cognitive states are that we were interested in looking at ***trait features***, i.e. enduring differences that are intrinsic to the phenotype and not related to differences in momentary idiosyncratic thinking [21]. Resting conditions are at higher risk for mixing both intrinsic and idiosyncratic differences while active conditions reduce the inter-subject idiosyncratic variance which was not relevant for our purpose. All the 3 rCBF images (of 1 patient for 1 ASL method) were compared to their matched rCBF average of the 37 controls using a 2-sample t-test with age and global rCBF as co-variant of non-interest [38]. This can be viewed as a ***partial fixed-effect model*** preserving the patient's intra-subject variance without taking into account the intra-subject variance of the reference group [39].

2.6. Biomarkers performance characteristics

In order to avoid antipsychotic-related rCBF changes, the two voxels of interest (VOIs) were defined from our previous analysis of PC vs. other treated psychoses [21] by taking the most significant voxels in the same clusters than reported by Walther et al. (fig. 1c from ref. [19]). VOIs' [x; y; z]-coordinates according to the MNI reference frame were as follow: L-SMA [-9; +10; +60] and L-LPM [-46; -12; +43] (in mm). T-values were extracted from both VOIs for each ASL-method (ASL-9.7 and ASL-21).

The ***analytical performances*** of ASL-9.7- and ASL-21-derived t-values were assessed by looking at their correlation (r) and determination (r^2) coefficients for each VOI as proxies for their ***accuracy***. Their ***linearity*** was challenged by testing whether polynomial regressions up to the 6^{th} degree could better fit the data.

Cut-off for ***testing*** was defined *a priori* by setting the risk of false positive result at $\alpha < 0.05$ (two-tails). Yet, considering that only high t-values would be indicative of PC, the risk of false positive result was indeed $\alpha < 0.025$ one-tail, albeit not corrected for multiple comparisons, i.e. a test was positive if t-value > 2.02 (df = 38). Beyond the individual testing of each VOI (L-SMA, L-LPM) according to each measurement method (ASL-9.7, ASL-21), two kinds of combinations were further assessed (complex biomarkers): positive test in both methods, i.e. logical AND (ASL-9.7 ∧ ASL-21), positive test in at least one VOI, i.e. logical OR (L-SMA ∨ L-LPM) and a combination of both which we will refer to as the ***combo test***, i.e. ASL-9.7(L-SMA ∨ L-LPM) ∧ ASL-21(L-SMA ∨ L-LPM). Indicators of tests' ***discriminant performances*** were:

- Sensitivity (Se), specificity (Sp)
- Youden's J statistic = Se + Sp – 1. J value range from 0 (useless test) to 1 (perfect test).
- Area under the curve (AUC) driven from the receiver operating characteristics (ROC).

The Bayesian estimations of the posterior probability density function for Se and Sp were performed using the R Stats Package [40] running under R-Studio [41]. Distributions of *p(Se|[TP; FN])* and *p(Sp|[TN; FP])* were modeled by a beta function (TP = true positive, FN = false negative etc.). Starting with non-informative Jeffrey's priors, i.e. beta (½; ½), Se and Sp posterior distributions were computed using an optimized version of Kennedy's algorithm [42] (10^4 random number generations) from which their means could be calculated together with their 95% credible intervals and their probabilities to be above chance level, i.e. cumulative distribution ≥ 0.5.

The ***agreement*** between both methods was evaluated for each VOI with the kappa coefficient (κ) between ASL-9.7 and ASL-21, but also appraised through the similarities in the spatial distribution of their discriminating performance (Youden's J statistic mapping).

The ***positive*** and ***negative predictive values*** (PPV, NPV) were estimated using Bayes theorem from the prevalence (Prev) of PC in various conditions and the Se and Sp of the combo test [43,44]. In this framework, PPV corresponds to the *a posteriori* (after the test) probability of being affected by PC. It is equal to the *a priori* probability of being affected by PC, i.e. the prevalence (known before the test) multiplied by the probability of having a positive test when being affected, i.e. Se, divided by the probability of having a positive test [43,44]:

$$PPV = \frac{Se \cdot Prev}{Se \cdot Prev + (1 - Sp) \cdot (1 - Prev)}$$

In the same vein, NPV corresponds to the *a posteriori* probability of not being affected by PC. It is equal to the *a priori* probability of not being affected by PC, multiplied by the probability of having a negative test when not being affected, i.e. Sp, divided by the probability of having a negative test [43,44]:

$$NPV = \frac{Sp \cdot (1 - Prev)}{(1 - Se) \cdot Prev + Sp \cdot (1 - Prev)}$$

PC prevalence is known for patients hospitalized for endogenous psychoses (10%)[7–10]. It can be further enriched using published Se and Sp values of 4 clinical criteria:

- ≥ 1 other affected first degree relative: Se = 60%; Sp = 70% [12],
- ≥ 2 relapses: Se = 98%; Sp = 25% [7],
- bipolar features in the WKL sense: Se = 86%; Sp = 51% [7],
- ≥ 2 catatonic symptoms during ≥ 1 episode: Se = 60%; Sp = 60% [20,45].

We combined them to recompute the prevalence of the tested sample [44]. Last, we independently considered the Se and Sp values of the operationalized criteria for research on PC (OCR-PC) as it already includes the above-mentioned clinical criteria: Se = 95%; Sp = 80% [21].

3. Results

3.1. Participants' characteristics

None of the patients had been included in a previous report (Tab. 1). The PC group consisted in 9 patients (8 from the RETONIC trial on chronic catatonia and 1 from the iADAPT study on resistant depression). The non-PC group gathered 26 patients (23 patients from iADAPT and 3 from RETONIC). As detailed in table 1, depressed patients were of various kinds according to the DSM or to WKL. The 3 non-PC patients having a WKL-diagnosis of system catatonia were diagnosed with DSM-5 schizophrenia and WKL pseudo-compulsive (or manneristic) catatonia (pCC), a more severe catatonic phenotype (see Tab. 1 and supplementary material). Regarding the depressive group, 6 patients had significant MRI anomalies

which were incidentally discovered in all but one (Tab. 2a). In 4 of them, the anomaly presumably accounted for treatment resistance. Last, 2 depressed patients had an abnormal neurological examination, essentially consisting in mild motor deficits related to their neurological history which was not an exclusion criterion for patients (unrelated to their depression, Tab. 2b). Regarding group-characteristics, PC patients were significantly younger, more likely to be left-handed and under antipsychotic treatment than non-PC patients (left handedness was not an exclusion criterion for patients). On the CRDPSS, the two groups significantly differed for the delusion, psychomotor and negative symptoms dimensions (all $p < 0.03$, see supplementary material).

		Patients			Controls	Significance
	Groups	PC	Non-PC		CTR	PC vs. non-PC
			Dep	pCC		
	Study	RETONIC (n = 8) iADAPT (n = 1)	iADAPT	RETONIC	Connect	
Characteristics	Number	9	23	3	37	
	Age (years)	40 ±13	52 ±12	35 ±9	43 ±11	0.03
	Gender (F/M)	3 / 6	14 / 9	1 / 2	19 / 18	ns
	IQ (fNART)	102 ±8	106 ±8	100 ±15	106 ±7	ns
	Education (Y)	15 ±2	16 ±2	15 ±1	14 ±3	ns
	Edinburgh (manuality)	2 ±92 %	84 ±39 %	97 ±6 %	89 ±20 %	0.03
	Sign. brain anomaly (Y/N)	0 / 9	6 / 17	0 / 3	0 / 37	ns
Clinical	Disorder-diagnosis (DSM-5) UD / BD / SZ-aff / SZ	0 / 1 / 1 / 7	14 / 8 / 1 / 0	0 / 0 / 0 / 3		0.0003
	Catatonia DSM-5 (Y/N) BFCRS ≥2 cutoff (Y/N)	0 / 9 8 / 1	0 / 23 0 / 23	1 / 2 2 / 1		ns $3{\cdot}10^{-6}$
	Phenotype-diag. (WKL) 2nd / Neu / PMD / Mis	-	4 / 4 / 7 / 8	-		
	Functioning (GAF)	42 ±13	50 ±11	35 ±17		ns
	Age at onset	22 ±7	31 ±15	17 ±2		[0.06]
	Duration (years)	17 ±12	21 ±14	18 ±7		ns
	Nb episodes	3 ±3	2 ±3	0 ±1		ns
	1rst deg relative (Y/N)	5 / 4	14 / 9	0 / 3		ns
	IDS-SR30	28 ±20	44 ±10	20 ±10		ns
	Calgary depression scale	8 ±8		3 ±1		
	PANSS total	67 ±9		62 ±9		
	PANSS positive	11 ±4		10 ±5		
	PANSS negative	19 ±2		16 ±4		
	BFCRS (catatonia)	3 ±2		8 ±9		
TTT	Antipsychotics (Y/N)	8 / 1	10 / 13	3 / 0		0.04
	All OLZ equivalent (mg)	17 ±14	5 ±8	10 ±5		[0.06]
	If OLZ equivalent (mg)	19 ±14	12 ±9	10 ±5		ns
	BdZ (Y/N)	7 / 2	14 / 9	3 / 0		ns
	All DZP equivalent (mg)	10 ±7	9 ±11	59 ±1		ns
	If DZP equivalent (mg)	13 ±5	15 ±11	59 ±1		[0.09]
	Antidepressant (Y/N)	5 / 4	13 / 10	3 / 0		ns
	Mood stabilizer (AE/Li/AP)	3 / 3 / 3	8 / 5 / 11	2 / 0 / 0		ns

Tab. 1. Participants characteristics. Significant differences are given only for the PC / non-PC comparison: non-significant ($\alpha > 0.1$, n.s.), trends are given in brackets ($0.1 \geq \alpha > 0.05$), the p-values are given when $\alpha \leq 0.05$ (t-test for quantitative variable, χ^2-test for qualitative variables). Patients with brain anomalies are described in table 2a, of which only one had an abnormal neurological examination (Tab. 2b). DSM-5 disorder-diagnoses: unipolar depression (UD), bipolar disorder (BD), schizophrenia and schizoaffective disorders (SZ-aff). WKL phenotype-diagnoses: secondary depression (2nd), neurotic depression, manic-depressive illness (PMD), pseudo-compulsive catatonia (previously named manneristic catatonia - pCC) and miscellaneous monopolar depressions (Mis). Regarding treatments, the proportion of treated and untreated patients is given for antipsychotics, benzodiazepines, and antidepressants. Olanzapine equivalent doses (OLZ eq) and diazepam equivalent doses (DZP eq) are given for the whole groups and for the subgroups of patients under the medication. For mood stabilizers the number of patients under antiepileptic (AE), lithium (Li), and antipsychotic (AP – for mood-stabilizing purpose) are given.

3.2. Analytical performances

Uncorrected rCBF values (ASL-9.7 and -21 average) were significantly higher in PC than non-PC and controls in L-SMA: 41 ±7 ml/min/100g (vs. 34 ±9; 32 ±9) and L-LPM: 68 ±8 ml/min/100g (vs. 53 ±12; 51 ±13) (all tests $p < 0.04$, see supplementary material). ASL-9.7 and ASL-21 rCBF measures converted in t-values corrected for age and global perfusion effects were well correlated to each other in both VOIs: $r = 0.84$ and $r = 0.9$ for the L-SMA and the L-LPM respectively ($p < 10^{-10}$), resulting in a determination coefficient (proportion of shared variance) of $r^2 = 0.71$ and $r^2 = 0.8$. In both VOIs, plots were not indicative for a deviation from a linear model:

$$t_{ASL\text{-}21} = \beta \cdot t_{ASL\text{-}9.7} + \mu$$

($t_{ASL\text{-}21}$ for ASL-21 t-value, idem for $t_{ASL\text{-}9.7}$ – see figure 1)

Constant 'μ' non-significantly differed from 0 and the data were not better fitted by polynomial regression up to the 6th degree. Whereas ASL-9.7 and ASL-21 had slope parameter $\beta = 0.97$ in L-LPM (n.s. different from 1), this was not the case in L-SMA which $\beta = 0.57$ significantly differed from 1 (post-hoc test, $p_{uc} = 1.1{\cdot}10^{-7}$).

2a	Subject	Finding	Brain anomaly
	RD05	Incidental	Idiopathic hypertrophic pachymeningitis. Probable secondary resistant depression.
	RD08	Incidental	Dysembryoplastic neuroepithelial tumor (DNET) of the left dorso-lateral prefrontal cortex. Probable secondary resistant depression.
	RD09	Incidental	FLAIR hyperintensity in the brainstem with lactate peak with MR-spectroscopy in a patient with previous history bipolar disorder. Unknown relation.
	RD10	Incidental	Large right orbitofrontal hypoperfusion secondary to an emergency surgery of an anterior cerebral artery aneurysm rupture in a patient with history of bipolar disorder. Resistance probably accounted for by the surgery.
	RD18	Incidental	Large arachnoid cyst of the right temporal pole with orbitofrontal hypoperfusion in a patient with history of bipolar disorder. Unknown relation.
	RD23	Expected	Secondary progressive multiple sclerosis with multiple demyelinations. Probable secondary resistant depression.
2b			**Abnormal neurologic examination**
	RD13		Mild motor deficit of the left leg secondary to infantile poliomyelitis
	RD23		Fully ambulatory without aid, moderate disability (EDSS = 3.5)

Tab. 2. Individuals with MRI and neurological anomalies. a) Patients with significant MRI anomalies. For four patients, these were considered as responsible for the resistance of their depression. Except for the patient with multiple sclerosis, none of these findings were anticipated. b) Patients with abnormal neurological examination. This includes the patient with multiple sclerosis, but also one other patient with normal MRI.

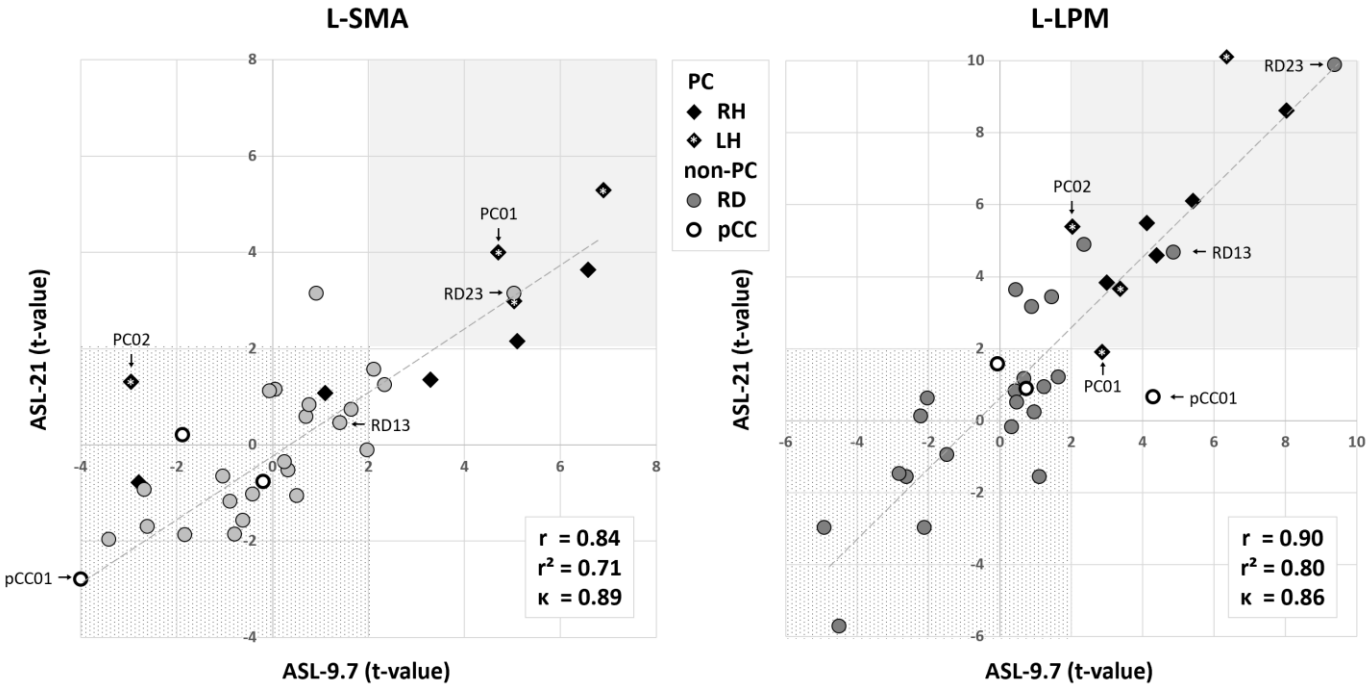

Fig. 1. Analytical performances of the measurements. The patients' p-value for ASL-9.7 are plotted against their ones for ASL-21: L-SMA (on the left) and L-LPM (on the right). The shaded parts of the graphs indicate where both measures are tested positive while the dotted part indicates where both measures are tested negative. Patients within the white areas have incoherent test-results. PC-patients are marked by black diamonds (◆), added with a white star when the subject was left-handed (LH). Non-PC patients are marked by circles either filled in gray for resistant depression (RD - ●) or non-filled for patients with pseudo-compulsive catatonia, previously named manneristic catatonia (pCC - ○). Special cases discussed in the result section are indicated by an arrow. Values for correlation (r), determination (r^2) and kappa (κ) coefficients of each VOI are given in the boxes for each measurement method.

3.3. Discriminant performances

Se and Sp were significantly different from chance level for all the tests except for Se in the L-SMA regardless of the method (Tab. 3, Fig. 2). The combo test was positive for all PC-patients (TP = 9/9, FN = 0/9) while it was only the case for 3 non-PC (FP = 3/26, TN = 23/26), hence Se = 98% and Sp = 88%. Considering the good analytical performances of the

	ASL-9.7							
	Se	CI	p(Se > 0.5)	Sp	CI	p(Sp > 0.5)	Youden's J	AUC
L-SMA	66 %	35 - 90 %	0.84	88 %	72 - 100 %	1	0.54	0.76
L-LPM	98 %	76 - 100 %	1	84 %	67 - 95 %	1	0.82	0.91
L-SMA V L-LPM	98 %	76 - 100 %	1	77 %	59 - 90 %	1	0.75	0.91
	ASL-21							
L-SMA	55 %	25 - 83 %	0.63	92 %	77 - 98 %	1	0.47	0.86
L-LPM	87 %	58 - 98 %	0.99	77 %	59 - 90 %	1	0.64	0.92
L-SMA V L-LPM	87 %	59 - 99 %	0.99	80 %	63 - 92 %	1	0.67	0.91
	ASL-9.7 ∧ ASL-21							
L-SMA	55 %	25 - 83 %	0.63	96 %	83 - 100 %	1	0.51	0.86
L-LPM	98 %	76 - 100 %	1	84 %	67 - 95 %	1	0.82	0.92
L-SMA V L-LPM	98 %	76 - 100 %	1	88 %	72 - 97 %	1	0.86	0.91

Tab. 3. Discriminant characteristics of the biomarkers given for L-SMA, L-LPM and [L-SMA ∨ L-LPM] (∨ = logical OR). These are coming from ASL-9.7 and ASL-21 measures, either in isolation or in combination, i.e. significant in both (∧ = logical AND). The performances of the combo test correspond to the last line. Specificity (Se) and sensitivity (Sp) are given in % with their 95% credible interval (CI) and their probability to differ from chance level, i.e. p(Se/Sp > 0.5). The latter can be interpreted as the reverse of the probability of H_0, i.e. probability $p \geq 0.95$ means that the probability to get the result by chance would have had a p-value ≤ 0.05 using a non-parametric frequentist approach. The combined measures of discriminant performances are the Youden's J statistic and the area under the curve (AUC) derived from the receiver operating characteristic (Fig. 2).

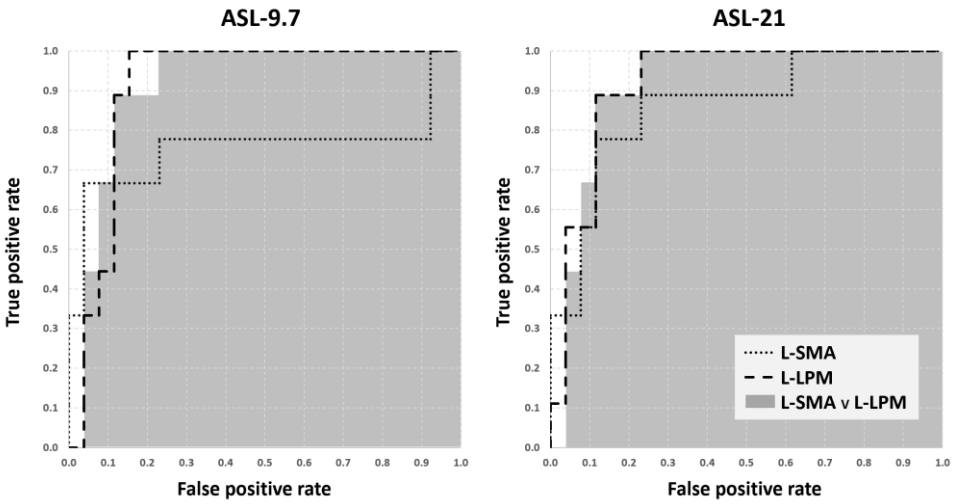

Fig. 2. Receiver operating characteristics (ROC curves). False positives (1 - specificity) are plotted against true positives (sensitivity) for each measurement: ASL-9.7 (on the left) and ASL-21 (on the right). The ROCs for the L-SMA are the dotted lines, the ones of the L-LPM are the dashed lines and the surface is for the OR combination of the two VOIs (L-SMA ∨ L-LPM). The latter are similar for the two measurement methods.

measurements, some remarkable cases were worth to mentioned (Fig. 1):

- Astonished by the proportion of left handers in the PC group, we marked their data points with a white star, but these did not show up as having a distinctive distribution. We further computed the correlation between VOIs p-values with handedness scores according to the Edinburg inventory after having regressed the group effect: none were significant ($r \in [-0.09, 0.14]$).
- None of the 3 non-PC positive cases at the combo test belonged to the pCC group (the system catatonia phenotype). Even when looking at individual tests, none were positive except pCC01 in L-LPM for ASL-9.7. The patient was far from threshold in all the other tests.
- Within the group of false positive, RD23 and RD13 were the two patients having moderate neurological motor impairments unrelated to their depression but related to CNS lesions (Tab. 2b). Considering that in neurological cases, premotor hyper-perfusions may have been adaptive reactions to the motor impairment, we recomputed the test performances without the 5 non-PC patients with significant neurological

anomalies (combo: Se = 98%, Sp = 95%, see table 1 in supplementary material). Considering that this might have also affected our pCC patients, test performances we recomputed without them and without the 5 depressed patients with significant MRI or neurological anomalies (combo: Se = 98%, Sp = 98%, see table 3 in supplementary material).

- PC01 was the patient with a phenotype-diagnosis of PC included in the resistant depression trial having a disorder-diagnosis of bipolar disorder type 1. He was also neuroleptic free. Despite that, he was positive at the combo test (positive for all individual tests except L-LPM; ASL-21 for which the t-value = 1.91 would have only been significant at $\alpha = 0.05$ one-tail).
- One PC-patient, PC02, was far from the regression line due to an outlier value in one test (L-SMA; ASL-9.7). By looking back at the data, the patient did move during two ASL-9.7 measurements which could also account for his marginal significance in the other VOI (L-LPM; ASL-9.7).

3.4. Agreement between testing methods

ASL-9.7 and ASL-21 had a very good level of agreement with $\kappa = 0.89$ in the L-SMA and $\kappa = 0.86$ in the L-LPM. Moreover, the two measures were also concordant in the regions with high discriminating performances (Youden's J map, Fig. 3): L-SMA and L-LPM, L-operculum and the striatum from both sides.

3.5. Inferences for PPV and NPV

It is clear from table 4 that negative test could be confidently interpreted as indicative of diagnostic exclusion whatever the prevalence of PC in the sample (all NPVs ≥ 95%). On the other hand, PPVs remained insufficient in most cases (< 80%) except when the test was used in patients fulfilling at least 3 clinical criteria, PPV = 81% or 86% depending on the criteria. The use of the combo test on a sample preselected with the operationalized criteria (OCR-PC)[21] increased the PPV to 94%, which is close to clinical usability. Yet, the test appeared even more promising when using the Se-Sp values computed after the exclusion of the 5 neurologically impaired patients: PPV = 98% (see supplementary material).

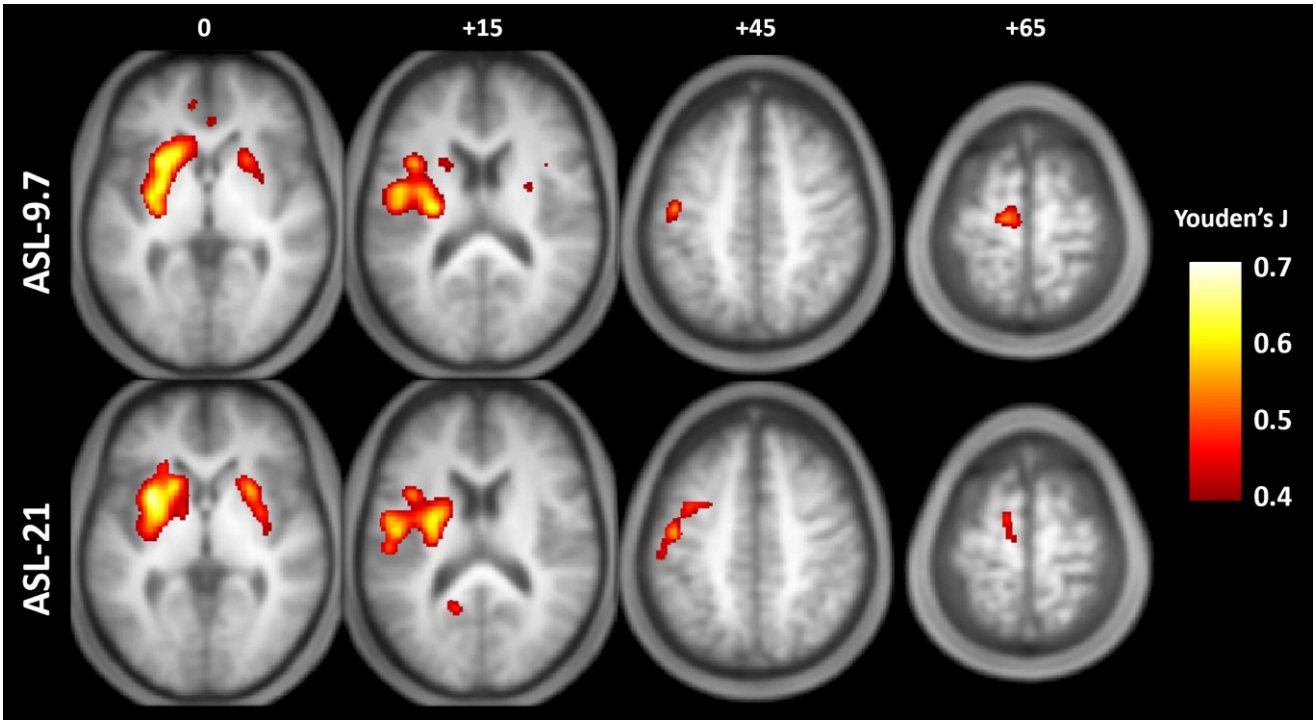

Fig. 3. Youden's J statistic maps. The threshold was set to J ≥ 0.4 and k ≥ 150 vx (1.2 cm^3). Horizontal slices are presented in neurological orientation (left is left) for the MNI-z coordinates given in the upper line (in mm). Results using ASL-9.7 method are shown in the upper row and the ones using ASL-21 method in the lower row. This combined performance index shows that the same regions have the highest discriminant power: the two used as biomarkers (L-SMA and L-LPM), but also the left operculum and both striatum.

For Se = 98% and Sp = 88% + Conditions	*a priori* PC prev.	*a posteriori* PPV	NPV
Psychotic inpatient	*10%*	48%	100%
↳ No affected 1st deg. relative	*6%*	34%	100%
↳ ≤ 1 relapse	*1%*	8%	100%
↳ ≥ 2 relapses	*8%*	42%	100%
↳ Bipolar	*13%*	55%	100%
↳ ≥ 2 catatonic Σ	*18%*	64%	100%
↳ Affected 1st deg. relative	*18%*	64%	100%
↳ ≤ 1 relapse	*2%*	14%	100%
↳ ≥ 2 relapses	*23%*	71%	99%
↳ Bipolar	*34%*	81%	99%
↳ ≥ 2 catatonic Σ	*43%*	86%	98%
↳ OCR - PC	*64%*	94%	96%

Tab. 4. Inference of predictive values. Positive and negative predictive values (PPV and NPV respectively) depend on the prevalence of PC in the tested population (PC prev.). Starting from PC prevalence in a sample of patients hospitalized for psychosis (10%), the combo test (last line Tab. 3) would only have a PPV of 48% while the NPV would be of 100%. The existence of an affected 1st degree relative would increase the prevalence to 18%, hence allowing PPV to raise up to 64%. Yet, this must be combined (↳) with a relapsing course (at least 2 episodes) and with bipolar or catatonic features in order to enrich the sample in PC to 34 and 43% resulting in the PPV to raise up to 81 and 86% respectively. Only the use of operationalized criteria for research on PC (OCR-PC, see ref. [21]) allows the sample to be sufficiently enriched (estimated prevalence of 64%) to get a PPV = 94% and NPV = 96% for the combo-test.

3.6. Post-hoc tests

In order to assess the influence of antipsychotic intake in the results, two post-hoc tests were performed: a correlation of t-values with OLZeq and a t-test between patients taking and those not-taking antipsychotics. After having regressed the group effect (PC / non-PC), both tests were significant in L-SMA ($r = -0.46$; $t = -0.45$; both $p < 0.01$), but not in the L-LPM ($r = -0.15$; $t = -0.47$; both n.s.): taking antipsychotic decreased t-values and the higher the doses, the lower the t-values.

4. Discussion

This proof-of-concept translational study supports L-SMA and L-LPM hyper-perfusions as potential biomarkers for PC. In line with our previous report [22], both ASL methods provided measurements of good analytical characteristics. These resulted in a diagnostic test of very good reproducibility and good discriminant performances. Yet, even the combo test did not have enough discriminating power to be used on an unselected psychotic population: while its NPV would be sufficient to reject the diagnosis (≥ 95%), its PPV remains too low except when used in combination with OCR-PC (96%).

4.1. Confounding factors: age, handedness and treatment

The significant difference in ***age*** between PC and non-PC is related to the younger age of onset of PC compared to most non-PC phenotypes, i.e. 22 ±7 vs. 29 ±15, and the shorter duration of the illness in PC compared to non-PC, i.e. 17 ±12 vs. 21 ±13. It is however unlikely that this might have significantly contributed to the results since the effect of age was regressed out in the computation of each subject's t-values.

More problematic might be the significant difference in ***handedness*** between PC and non-PC patients that came out as a surprise. We are not aware of a previous report about hand dominance in catatonia and this couldn't have been observed in previous brain imaging studies in which left-handedness was an exclusion criterion in both [19,21].

Could left-handedness accounts for the left-lateralization of premotor hyper-perfusions in PC? Three lines of argument run counter to this hypothesis. First, there were no relation between handedness and VOIs p-values. Second, the literature would have predicted the opposite since left-handers have lesser left and greater right premotor activation during motor tasks [46]. Last, the previous studies showed the same left-lateralization even though the patients were exclusively right-handed [19,21].

The other way around, could left-handedness be an adaptive reaction to the primary pathogenic process? If true, this would set the beginning of the pathogenic process long before the first psychotic outburst. This would challenge the "neurodegenerative" hypothesis suggested by the typical progressive-remitting course of PC which should be at least enriched by a neurodevelopmental perspective. In any case, handedness merits to be further investigate in PC.

Last the groups also significantly differed in the proportion of patients under ***antipsychotics***. Yet the treatment was unlikely to have contributed to this effect since it was correlated with lower t-values in

L-SMA. This result should be taken with caution as it was a post-hoc test and as significance only came out after having regressed out the group effect. Yet, if confirmed it raises the possibility that antipsychotics could correct PC-related L-SMA hyperactivity. This would be in line with evidence showing that they partially correct for the defective inhibition in the motor-cortex [47]. Considering the larger corrective effect of clozapine [48], this might further explain why this treatment is credited to be especially more effective in PC [49].

4.2. Challenging the disease-model of PC

These preliminary results are mainly replicating previous studies showing that a spatially overlapping pattern of left-sided premotor hyper-perfusion can discriminate PC from non-PC patients [19,21]. The only difference with previous studies is the high discriminant performance of the ***L-striatum*** hyper-perfusion according to the Youden'J statistical map. This was unexpected since most of the evidence ascribe this change to antipsychotic treatments: striatal rCBF is increased under medication [50] and this increase might even be predictive of the therapeutic response [51]. Moreover, this effect is congruent with the increases in glucose consumption [52] and gray mater volume [53] of the striatum after antipsychotic intake. Hence there was no significant striatal difference in rCBF when comparing PC to another medicated group having similar response to antipsychotic in the previous trials [19,21], whereas it was observed in the medicated group when compared to controls [21]. L-striatum hyper-perfusion deserves to be further examined in non-PC psychotic and affective groups to understand its good discriminant performance in this trial.

The single subject approach of this study allowed two interesting observations. First, the biomarker appeared to be ***specific for PC*** as a categorical ***phenotype*** regardless of the main clinical presentation or the treatment, in contradistinction with the predictions of dimensional paradigms, i.e. DSM-5 section III [1], SyNoPsis [4], RDoCs [3] or HiTOP [2]. On the one side, L-SMA and L-LPM hyper-perfusions were unrelated to catatonic symptoms *per se* since none of the pCC patients presented these hyper-perfusions despite their high scores at the BFCRS. These observations would fit the ***PC-trait marker*** better than the ***catatonic symptoms state marker*** hypothesis. On the other side, these rCBF increases remained even in the PC patient diagnosed with depressive episode of a bipolar disorder type 1, who never presented psychotic symptoms and was free of antipsychotic medication (PC01). Our point is not to argue for WKL categories against dimensional paradigms. In fact, the latter remain more supported by current evidence [54,55]. Yet, this illustrates how categorical and dimensional research programs are making sufficiently different predictions to be confronted on functional brain measures. We already called for the implementation of such crucial experimental series [56]. This is the primary aim of our follow-up study for which we are looking for collaborations especially in an adversarial perspective [57].

The second observation elicited by the possibility to draw inference from single-cases was the significant premotor hyper-perfusions in the two depressed patients with ***neurological motor impairment***: RD23 and RD13, both positive at the combo test. This is suggestive of the same sensorimotor-premotor compensatory hyperactivities than reported after subcortical strokes [58]. Such interpretation makes the case for neurological ***motor*** impairment to be an unanticipated situation of diagnostic uncertainty which should conservatively lead us to exclude this confounding factor in forthcoming validation studies to improve the interpretability of the biomarker [59]. Besides, this observation also questions the current understanding of these measures as a ***pathogenic biomarker***, suggesting that they could result from an adaptive reaction to the ***psychomotor*** disturbances and hence be of ***non-pathogenic*** nature (see supplementary material). Such distinction is critical from a therapeutic perspective since only the correction of a pathogenic process has any chance to be effective whereas impairing a compensatory process might even be detrimental. Preliminary evidence of improvements after the inhibition of L-SMA and L-LPM by rTMS were suggestive of the pathogenetic nature of these hyper-perfusions, though the number of subject remained small [60].

4.3. Pitfalls and strategies for future improvements

While encouraging, this study only provides a ***proof of feasibility*** since its design, implementation and population samples are far to comply with current standards [23]. The most critical drawback is the composition of the ***non-PC group*** which significantly differs from diagnostically challenging patients in

clinical practice, i.e. those suffering from another phenotype of endogenous psychoses. Indeed, these only accounted for 11 of the 26 non-PC patients (42%): 3 pCC, 1 confusion psychosis and 7 manic-depressive illnesses (according to WKL-phenotypes classification).

Putting aside the problem of scanning time, some methodological flaws merit to be addressed. First, the case of patient PC02 raised the issue of ***quality control***. Basic indicators should be automatically computed and considered when interpreting the results, e.g. signal-noise ratio, head movements or variance of EPI time series. For the test deployability [59], minimal quality values should be defined together with appropriate artifact correction methods to reduce the risk of missing data or scan repetition [61,62]. Second, the standard ***spatial normalization*** procedure is known to be suboptimal and could lead to detrimental consequences if the patient's brain is misaligned relative to controls [63]. Future works should take advantage of more accurate brain registration methods [64], perhaps even considering realignments based on functional connectivity [65]. Last, new ***individual vs. group statistics*** remain to be evaluated for our purpose since they were recently shown to outperform the univariate approach used in the present analysis [66]. These are taking advantage of the multivariate nature of brain imaging by involving the surrounding voxels in the comparison. While departing from our model-based approach, supervised linear or non-linear classifiers could take this advantage even further [67]. Yet, the switch to unsupervised approaches [68] would be in clear-cut opposition with our hypothesis-driven strategy: automatic classifiers are designed to identify ***new biotypes*** while we intend to validate ***biomarkers*** of ***known phenotypes*** [56].

4.4 Limitations

These results are just a proof of concept of poor generalizability. Beyond the ***low number*** of PC-patients, we already mentioned that the non-PC group was ***not much diagnostically challenging***, since none except the 3 pCC patients had psychotic and psychomotor symptoms. It belongs to the follow-up study to test it against more difficult differential diagnoses, e.g. PC vs. other WKL psychotic phenotypes belonging to the same schizo-affective spectrum. Moreover, Bayesian inferences of PPVs and NPVs show that even if the discriminant performances do not abate, they still are insufficient for the test to be used on a non-preselected population, i.e. a group in which PC prevalence is already high (> 50%). Hence, the biomarker will not exempt investigators from using demanding operationalized clinical criteria (OCR-PC)[21].

4.5. Conclusion

In view of these encouraging results, a follow-up study was initiated to recruit a larger and more challenging samples and to reduce the MRI scanning time. Despite some methodological differences, the Bayesian framework will allow the use of current Se and Sp estimates as priors in forthcoming analyses. This might well accelerate the buildup of significant results even from small samples, and thus speed-up decision making in the research process. However, even if the diagnostic accuracy of the current biomarker is confirmed and generalized across other laboratories, it is not intended to be used in clinical practice. It primary aims at providing a simple and reliable process for making a PC-diagnosis, a mandatory step for the phenotype to be studied and challenged by researchers unfamiliar with the WKL framework

5. Appendix

See **supplementary material**.

6. Statements

6.1. Acknowledgements

The authors express their gratitude to Mrs Annabel Schneider for her tireless work to plan, organize, scan, evaluate the patients, to Mrs Corinne Marrer for the MRI acquisitions, to Mr Gregory Johais for securing and managing behavioral data, to Mr Jean-Paul Armspach for his continuous support as well as to the other members of the CEMNIS (Noninvasive Neuromodulation Center): Mrs Mandy Haumesser, Mrs Céline Berhendt and Mrs Frédérique Ostré. We are grateful to the IRIS platform of the iCube Laboratory for providing us the scanning slots. Last, we thank Mrs Hélène Soavelo (project manager), Mrs Madoé Julians (clinical research assistant) and the Hopitaux Universitaires de Strasbourg for their help in the regulatory processes and the administration of these studies. We also thank the two reviewers of our paper for their highly relevant

remarks that have helped us improve the manuscript.

6.2. Statement of Ethics

All subjects included in these trials together as their caregivers when appropriate have given their written informed consent. All study protocols comply to the Code of Ethics of the World Medical Association (Declaration of Helsinki) and were approved by the local ethic committee. Studies were previously registered on clinicaltrials.gov: Connect C3 (NCT02868879), iADAPT (NCT02863380), RETONIC (NCT03116425).

6.3. Disclosure Statement

The authors have no conflicts of interest to declare. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

6.4. Funding Sources

Connect C3 was funded by “Projet hospitalier de Recherche Clinique” (DGS 2005/0543); iADAPT was supported by the “Fond Unique Interministeriel” (F1111009 A, rTMS robot); RETONIC was funded by the “Fondation de l’Avenir” (AP-RMA-2015-006). Costs of the regulatory procedures were supported by the Hôpitaux universitaires de Strasbourg which also paid 1-year salary for the “année-recherche” of C.d.B.

6.5. Author Contributions

J.F. conceived, planned and directed, the 3 studies, he drafted the manuscript and designed figures and tables. J.F., C.d.B., L.J., O.M., J.M.E.C., B.S., M.C.L. and F.B. contributed to patients’ selection, inclusion and evaluation. P.L.d.S. implemented the MRI sequences, J.L. provided key contribution in image quality control and conversion. J.F., C.d.B., J.L., V.N., E.A.S., L.L. and F.B. helped in data analysis and interpretation. All authors provided critical feedback and significantly contributed to the writing of this article.

7. References

1. American Psychiatric Association. DSM 5. American Journal of Psychiatry. 2013.
2. Kotov R, Waszczuk MA, Krueger RF, Forbes MK, Watson D, Clark LA, et al. The hierarchical taxonomy of psychopathology (HiTOP): A dimensional alternative to traditional nosologies. J Abnorm Psychol. 2017;126: 454–477.
3. Cuthbert BN. Research Domain Criteria: Toward future psychiatric nosologies. Dialogues Clin Neurosci. 2015;17: 89–97.
4. Strik W, Stegmayer K, Walther S, Dierks T. Systems Neuroscience of Psychosis: Mapping Schizophrenia Symptoms onto Brain Systems. Neuropsychobiology. 2017: 100–116.
5. WHO. ICD-11: International Classification of Diseases 11th Revision [Internet]. 2018 [cited 19 Mar 2019]. Available: https://icd.who.int/
6. Shorter E. Wernicke-Kleist-Leonhard Pathway. A historical dictionary of psychiatry. New-York: Oxford University Press; 2005. pp. 300–3.
7. Leonhard K. Age of Onset, Sex Incidence, Course. Classification of Endogenous Psychoses and their Differentiated Etiology. Vienna: Springer Vienna; 1999. pp. 250–274.
8. Leonhard K. Lassen sich die Schizophrenien klinisch und ätiologisch trennen? In: Seidel K, Neumärker K-J, Schulze HAF, editors. Zur Klassifikation endogener Psychosen. Vienna: Springer Vienna; 1986. pp. 26–42.
9. Stöber G, Franzek E, Lesch KP, Beckmann H. Periodic catatonia: a schizophrenic subtype with major gene effect and anticipation. Eur Arch Psychiatry Clin Neurosci. 1995;245: 135–41.
10. Krüger S, Bräunig P. Studies on course outcome and genetic loading of periodic catatonia. In: Beckmann H, Neumärker K-J, editors. Endogenous psychoses: Leonhard’s impact on modern psychiatry. Berlin, Germany: Ullstein-Mosby; 1995. pp. 181–185.
11. Leonhard K. Periodic catatonia. Classification of Endogenous Psychoses and their Differentiated Etiology. Vienna: Springer Vienna; 1999. pp. 104–112.
12. Beckmann H, Franzek E, Stöber G. Genetic heterogeneity in catatonic schizophrenia: A family study. Am J Med Genet. 1996;67: 289–300.
13. Pethő B, Tolna J, Tusnády G, Farkas M, Vizkeleti G, Vargha A, et al. The predictive validity of the Leonhardean classification of endogenous psychoses. Eur Arch Psychiatry Clin Neurosci. 2008;258: 324–334.
14. Fish F. The influence of the tranquilizers on the leonhard schizophrenic syndromes. Encephale. 1964;53: Suppl:245-9.
15. Astrup C. The effects of ataraxic drugs on schizophrenic subgroups related to experimental findings. Acta Psychiatr Scand Suppl. 1959;34: 388–93.
16. Ungvari GS, Leung CM, Wong MK, Lau J. Benzodiazepines in the treatment of catatonic syndrome. Acta Psychiatr Scand. 1994;89: 285–8.
17. Ungvari GS, Chiu HF, Chow LY, Lau BS, Tang WK. Lorazepam for chronic catatonia: a randomized,

double-blind, placebo-controlled cross-over study. Psychopharmacology (Berl). 1999;142: 393–8.
18. Northoff G, Steinke R, Czcervenka C, Krause R, Ulrich S, Danos P, et al. Decreased density of GABA-A receptors in the left sensorimotor cortex in akinetic catatonia: investigation of in vivo benzodiazepine receptor binding. J Neurol Neurosurg Psychiatry. 1999;67: 445–50.
19. Walther S, Schäppi L, Federspiel A, Bohlhalter S, Wiest R, Strik W, et al. Resting-State Hyperperfusion of the Supplementary Motor Area in Catatonia. Schizophr Bull. 2017;43: 972–981.
20. Walther S, Strik W. Catatonia. CNS Spectr. 2016;21: 341–348.
21. Foucher JR, Zhang YF, Roser M, Lamy J, De Sousa PL, Weibel S, et al. A double dissociation between two psychotic phenotypes: Periodic catatonia and cataphasia. Prog Neuro-Psychopharmacology Biol Psychiatry. 2018;86: 363–9.
22. Foucher J, Roquet D, Bruns A, Kindo C, Kletzl H, Squassante L, et al. What could be the best method for pharmacological MRI: BOLD or ASL? Human Brain Mapping. Hambourg; 2014.
23. McInnes MDF, Moher D, Thombs BD, McGrath TA, Bossuyt PM, Clifford T, et al. Preferred Reporting Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy Studies. JAMA. 2018;319: 388.
24. Schürhoff F, Fond G, Berna F, Bulzacka E, Vilain J, Capdevielle D, et al. A National network of schizophrenia expert centres: An innovative tool to bridge the research-practice gap. Eur Psychiatry. 2015;30: 728–735.
25. Lefaucheur JP, André-Obadia N, Poulet E, Devanne H, Haffen E, Londero A, et al. Recommandations françaises sur l'utilisation de la stimulation magnétique transcrânienne répétitive (rTMS): Règles de sécurité et indications thérapeutiques. Neurophysiologie Clinique. 2011. pp. 221–295.
26. Leonhard K. Classification of Endogenous Psychoses and their Differentiated Etiology. Beckmann H, editor. Vienna: Springer Vienna; 1999.
27. Oldfield RC. The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia. 1971;9: 97–113.
28. Mackinnon A, Mulligan R. [The estimation of premorbid intelligence levels in French speakers]. Encephale. 2005;31: 31–43.
29. American Psychiatric Association. Task Force on DSM-IV. DSM-IV. American Psychiatric Association; 1994.
30. Leucht S, Samara M, Heres S, Davis JM. Dose Equivalents for Antipsychotic Drugs: The DDD Method. Schizophr Bull. 2016;42: S90–S94.
31. Brett J, Murnion B. Management of benzodiazepine misuse and dependence. Aust Prescr. NPS MedicineWise; 2015;38: 152–5.
32. Rush AJ, Carmody T, Reimitz PE. The Inventory of Depressive Symptomatology (IDS): Clinician (IDS-C) and Self-Report (IDS-SR) ratings of depressive symptoms. Int J Methods Psychiatr Res. 2000;9: 45–59.
33. Kay SR, Fiszbein A, Opler LA. The positive and negative syndrome scale (PANSS) for schizophrenia. Schizophr Bull. 1987;13: 261–76.
34. Bush G, Fink M, Petrides G, Dowling F, Francis A. Catatonia. I. Rating scale and standardized examination. Acta Psychiatr Scand. 1996;93: 129–36.
35. Addington D, Addington J, Maticka-Tyndale E. Specificity of the Calgary Depression Scale for schizophrenics. Schizophr Res. 1994;11: 239–44.
36. Friston KJ, Ashburner J, Kiebel S, Nichols T, Penny WD. Statistical parametric mapping : the analysis of funtional brain images. Elsevier, Academic Press; 2007.
37. Foucher JR, Roquet D, Marrer C, Pham BT, Gounot D. Correcting for the echo-time effect after measuring the cerebral blood flow by arterial spin labeling. J Magn Reson Imaging. 2011;34: 785–790.
38. Signorini M, Paulesu E, Friston K, Perani D, Colleluori A, Lucignani G, et al. Rapid Assessment of Regional Cerebral Metabolic Abnormalities in Single Subjects with Quantitative and Nonquantitative [18F]FDG PET: A Clinical Validation of Statistical Parametric Mapping. Neuroimage. 1999;9: 63–80.
39. Maumet C, Maurel P, Ferré J-C, Carsin B, Barillot C. Patient-specific detection of perfusion abnormalities combining within-subject and between-subject variances in Arterial Spin Labeling. Neuroimage. 2013;81: 121–130.
40. Team RDC, R Development Core Team R. R: A Language and Environment for Statistical Computing. R Found Stat Comput. 2016; doi:10.1007/978-3-540-74686-7
41. RStudio. RStudio: Integrated development environment for R. Boston, MA; 2018.
42. Hung Y-C, Balakrishnan N, Lin Y-T. Evaluation of Beta Generation Algorithms. Commun Stat - Simul Comput. Taylor & Francis Group ; 2009;38: 750–770.
43. Meyer N, Vinzio S, Goichot B. La statistique bayésienne : une approche des statistiques adaptée à la clinique. La Rev Médecine Interne. 2009;30: 242–249.
44. Krzych LJ. Use of Bayesian statistical approach in diagnosing secondary hypertension. Pol Arch Med Wewn. 2008;118: 132–42.
45. Krause A. Symptom- und Verlaufscharakteristika bei periodischer Katatonie. Doctoral Thesis, Würzburg, Germany. 2012.
46. Siebner HR, Limmer C, Peinemann A, Drzezga A, Bloem BR, Schwaiger M, et al. Long-term consequences of switching handedness: a positron emission tomography study on handwriting in

"converted" left-handers. J Neurosci. 2002;22: 2816–25.
47. Daskalakis ZJ, Christensen BK, Chen R, Fitzgerald PB, Zipursky RB, Kapur S. Evidence for impaired cortical inhibition in schizophrenia using transcranial magnetic stimulation. Arch Gen Psychiatry. 2002;59: 347–54.
48. Daskalakis ZJ, Christensen BK, Fitzgerald PB, Moller B, Fountain SI, Chen R. Increased cortical inhibition in persons with schizophrenia treated with clozapine. J Psychopharmacol. 2008;22: 203–209.
49. Stöber G. Different etiological backgrounds in periodic catatonia and systematic catatonia. In: Franzek E, Ungvari GS, Rüther E, Beckmann H, editors. Progress in differentiated psychopathology. Würzburg, Germany: International Wernicke-Kleist-Leonhard society; 2000. pp. 280–291.
50. Handley R, Zelaya FO, Reinders AATS, Marques TR, Mehta MA, O'Gorman R, et al. Acute effects of single-dose aripiprazole and haloperidol on resting cerebral blood flow (rCBF) in the human brain. Hum Brain Mapp. 2013;34: 272–282.
51. Lahti AC, Weiler MA, Holcomb HH, Tamminga CA, Cropsey KL. Modulation of limbic circuitry predicts treatment response to antipsychotic medication: a functional imaging study in schizophrenia. Neuropsychopharmacology. 2009;34: 2675–90.
52. Potkin SG, Buchsbaum MS, Jin Y, Tang C, Telford J, Friedman G, et al. Clozapine effects on glucose metabolic rate in striatum and frontal cortex. J Clin Psychiatry. 1994;55 Suppl B: 63–6.
53. Brandt GN, Bonelli RM. Structural neuroimaging of the basal ganglia in schizophrenic patients: a review. Wiener Medizinische Wochenschrift. 2008;158: 84–90.
54. Cantisani A, Stegmayer K, Bracht T, Federspiel A, Wiest R, Horn H, et al. Distinct resting-state perfusion patterns underlie psychomotor retardation in unipolar vs. bipolar depression. Acta Psychiatr Scand. John Wiley & Sons, Ltd (10.1111); 2016;134: 329–338.
55. Walther S, Höfle O, Federspiel A, Horn H, Hügli S, Wiest R, et al. Neural correlates of disbalanced motor control in major depression. J Affect Disord. 2012;136: 124–133.
56. Foucher JR, de Billy C, Obrecht A, Mainberger O, Schorr B, Bonah C, et al. Bases épistémologiques de la recherche sur les psychoses Quelle solution pour le choc des cadres paradigmatiques ? Ann Médico-psychologiques, Rev Psychiatr. In press.
57. Kahneman D. Experiences of Collaborative Research. Am Psychol. 2003;58: 723–730.
58. Garraux G, Hallett M, Talagala SL. CASL fMRI of subcortico-cortical perfusion changes during memory-guided finger sequences. Neuroimage. 2005;25: 122–132.
59. Woo C-W, Wager TD. Neuroimaging-based biomarker discovery and validation. Pain. NIH Public Access; 2015;156: 1379–81.
60. Foucher JR, de Billy C, Mainberger O, Schorr B, Clauss J, Berna F. Personalized rTMS improves chronic and treatment resistant catatonias – A proof of concept study. 3rd European Conference on Brain Stimulation in Psychiatry "From mechanisms to medicine" Lyon 18th to 20th October 2018.
61. Roquet DR, Pham BT, Foucher JR. Manual selection of spontaneous activity maps derived from independent component analysis: Criteria and inter-rater reliability study. J Neurosci Methods. 2014;223: 30–34.
62. Sourty M, Thoraval L, Roquet D, Armspach J-P, Foucher J. Towards an automated selection of spontaneous co-activity maps in functional magnetic resonance imaging. Progress in Biomedical Optics and Imaging - Proceedings of SPIE. 2015.
63. Presotto L, Ballarini T, Caminiti SP, Bettinardi V, Gianolli L, Perani D. Validation of 18F–FDG-PET Single-Subject Optimized SPM Procedure with Different PET Scanners. Neuroinformatics. 2017;15: 151–163.
64. Tustison NJ, Cook PA, Klein A, Song G, Das SR, Duda JT, et al. Large-scale evaluation of ANTs and FreeSurfer cortical thickness measurements. Neuroimage. 2014;99: 166–179.
65. Conroy BR, Singer BD, Guntupalli JS, Ramadge PJ, Haxby J V. Inter-subject alignment of human cortical anatomy using functional connectivity. Neuroimage. 2013;81: 400–411.
66. Maumet C, Maurel P, Ferré J-C, Barillot C. An a contrario approach for the detection of patient-specific brain perfusion abnormalities with arterial spin labelling. Neuroimage. 2016;134: 424–433.
67. Mahmoudi A, Takerkart S, Regragui F, Boussaoud D, Brovelli A. Multivoxel Pattern Analysis for fMRI Data: A Review. Comput Math Methods Med. 2012;2012: 1–14.
68. Drysdale AT, Grosenick L, Downar J, Dunlop K, Mansouri F, Meng Y, et al. Resting-state connectivity biomarkers define neurophysiological subtypes of depression. Nat Med. 2017;23: 28–38.