



ARTICLE OPEN

A potential biomarker for treatment stratification in psychosis: evaluation of an [¹⁸F] FDOPA PET imaging approachMattia Veronese¹, Barbara Santangelo^{1,2}, Sameer Jauhar², Enrico D'Ambrosio^{2,3}, Arsime Demjaha², Hugh Salimbeni⁴, Jin Huajie⁵, Paul McCrone⁶, Federico Turkheimer¹ and Oliver Howes^{2,7,8}

[¹⁸F]FDOPA PET imaging has shown dopaminergic function indexed as K_i^{cer} differs between antipsychotic treatment responders and non-responders. However, the theragnostic potential of this biomarker to identify non-responders has yet to be evaluated. In view of this, we aimed to evaluate this as a theragnostic test using linear and non-linear machine-learning (i.e., Bernoulli, support vector, random forest and Gaussian processes) analyses and to develop and evaluate a simplified approach, standardised uptake value ratio (SUV_{Rc}). Both [¹⁸F]FDOPA PET approaches had good test-retest reproducibility across striatal regions (K_i^{cer} ICC: 0.68–0.94, SUV_{Rc} ICC: 0.76–0.91). Both our linear and non-linear classification models showed good predictive power to distinguish responders from non-responders (receiver operating curve area under the curve for region-of-interest approach: $K_i^{cer} = 0.80$, SUV_{Rc} = 0.79; for voxel-wise approach using a linear support vector machine: 0.88) and similar sensitivity for identifying treatment non-responders with 100% specificity (K_i^{cer} : ~50%, SUV_{Rc}: 40–60%). Although the findings were replicated in two independent datasets, given the total sample size ($n = 84$) and single setting, they warrant testing in other samples and settings. Preliminary economic analysis of [¹⁸F]FDOPA PET to fast-track treatment-resistant patients with schizophrenia to clozapine indicated a potential healthcare cost saving of ~£3400 (equivalent to \$4232 USD) per patient. These findings indicate [¹⁸F]FDOPA PET dopamine imaging has potential as biomarker to guide treatment choice.

Neuropsychopharmacology (2021) 46:1122–1132; <https://doi.org/10.1038/s41386-020-00866-7>

INTRODUCTION

Schizophrenia and related psychotic disorders are common mental disorders and amongst the leading causes of global disability [1–3]. Antipsychotic drugs are central to their treatment [4, 5]. However, about one-third of patients show limited response to first-line antipsychotic treatment, often from illness onset [6, 7]. Poor response is associated with increased health burden, higher costs and longer hospital stays [8]. Clozapine is an alternative treatment that is effective in people resistant to first-line antipsychotic drugs [9]. Moreover, clozapine treatment is associated with reduced mortality, healthcare costs and functional outcomes [10–14]. However, clinical guidelines around the world recommend its use is restricted to non-responders because of the risk of side effects and the need for blood monitoring [15]. As there is currently no way to differentiate non-responders from responders, clinical guidelines recommend a series of empirical treatment trials to determine if a patient is a non-responder, before then initiating clozapine. In practice this leads to long delays, on average over 4 years, before initiation of clozapine [16]. There is thus a clinical need for a biomarker to identify non-responders early to guide treatment choice.

Both the role of dopamine hyperactivity in the pathoetiology of psychosis [17–21], and findings that antipsychotic drugs act by blocking dopamine [4, 22] indicate that molecular imaging of the striatal dopamine innervation is a candidate biomarker for predicting treatment response. A number of studies have shown that striatal dopamine synthesis capacity, as measured by [¹⁸F] FDOPA PET imaging, is elevated in schizophrenia [23], schizophreniform psychoses [24] and people at clinical high risk for psychosis (UHR) [24–26], and linked to subsequent development of psychosis [24–26]. Moreover, dopamine synthesis capacity has been found to distinguish patients who have responded to standard antipsychotic drugs from non-responders [27] and both this and synaptic dopamine levels have been shown to predict response to antipsychotic treatment [20, 28, 29]. This evidence suggests that [¹⁸F]FDOPA PET imaging of the striatal dopamine innervation could be used as a neurochemical basis to stratify patients into those likely to respond and those unlikely to respond to first-line antipsychotic drugs. Timely stratification would enhance patient welfare by enabling more rapid treatment responses, while bringing a considerable economy to the healthcare system.

¹Department of Neuroimaging, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK; ²Department of Psychosis Studies, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK; ³Psychiatric Neuroscience Group, Department of Basic Medical Sciences, Neuroscience and Sense Organs, University of Bari "Aldo Moro", Bari, Italy; ⁴Faculty of Engineering, Department of Computing, Imperial College London, South Kensington Campus, London, UK; ⁵King's Health Economics, Health Service and Population Research Department, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK; ⁶Faculty of Education, Health and Human Sciences, University of Greenwich, London, UK; ⁷MRC London Institute of Medical Sciences, Hammersmith Hospital, London W12 0NN, UK and ⁸Institute of Clinical Sciences (ICS), Faculty of Medicine, Imperial College London, Du Cane Road, London W12 0NN, UK

Correspondence: Mattia Veronese (mattia.veronese@kcl.ac.uk) or Oliver Howes (oliver.howes@kcl.ac.uk)

These authors contributed equally: Mattia Veronese, Barbara Santangelo

Received: 27 May 2020 Revised: 27 July 2020 Accepted: 19 August 2020

Published online: 22 September 2020

However, the theragnostic potential of this method to identify patients who will respond to first-line treatment has yet to be evaluated. In view of this, our first aim was to determine the sensitivity and specificity of [¹⁸F]FDOPA PET imaging to distinguish treatment responders from non-responders. For a diagnostic test to be clinically useful it also needs to be practical. One potential limitation of the [¹⁸F]FDOPA PET imaging used in prior studies is the long duration of the scans, about 95 min. To address this, we aimed to evaluate a shorter, simplified protocol [¹⁸F]FDOPA PET imaging protocol and compare its accuracy with the full 95-min scans. Finally, we conducted a preliminary cost analysis using [¹⁸F]FDOPA PET to identify non-responders to first-line antipsychotic treatment.

METHODS AND MATERIALS

The most common approach to [¹⁸F]FDOPA PET imaging uses continuous dynamic acquisition, with the scanning beginning with the tracer injection and lasting for 90–95 min, during which the participant is required to lie still in the PET scanner. A 60-min acquisition had also been considered; however, a correlation analysis indicated that the 95-min acquisition is more reliable (Supplementary Tables 1 and 2).

Compared to this dynamic [¹⁸F]FDOPA PET acquisition, the simplified protocol consists of a brief single-frame acquisition (10–15 min) and a simplified index of FDOPA uptake (standardised uptake value ratio (SUV_{RC})), defined as the ratio of the tracer activity in the striatum to that of the cerebellum, as a proxy for dopamine synthesis capacity using the standard dynamic approach. This approach is similar to simplified [¹⁸F]FDOPA PET scanning methods that have already been used to distinguish patients with early Parkinson disease (PD) from healthy volunteers [30, 31].

To explore the generalisability of the simplified [¹⁸F]FDOPA PET imaging beyond the specifics of the datasets included in this study, we also tested its sensitivity to important experimental variables (the tracer injected dose and specific activity (SA) as well as the length of PET scan acquisition).

Datasets

The data presented are a new analysis of two different [¹⁸F]FDOPA PET imaging datasets that have been previously published [28, 32]. *Dataset1* [28] consists of 26 first-episode psychosis patients, who were scanned prior to antipsychotic treatment and three minimally treated for <2 weeks, and 14 age matched healthy volunteers to enable normative comparisons. Following the scans, all patients began treatment with a first-line antipsychotic medication, selected upon consultation with their psychiatrist and without reference to PET results. Treatment choice was made by the patient in consultation with their psychiatrist [28]. Treatment was titrated to a therapeutic dose (based on the Maudsley Prescribing Guidelines) and patients all received follow-up to at least 6 months to determine response status. To be included patients were required to show good adherence to treatment defined as taking more than 80% of prescribed antipsychotic doses in line with recommendations and other studies [15, 33, 34]. To assess concordance with antipsychotic medication, a multisource approach was used. This required evidence of adequate adherence on at least two of the following: antipsychotic plasma levels, pharmacy and electronic medical dispensing records and report from the patient and an independent source (family member/caregiver or healthcare professional). Adequate concordance was defined as taking a minimum of 80% of prescribed doses, in line with consensus recommendations [34].

Treatment response was defined as >50% reduction in PANSS total symptom severity rating from baseline to follow-up and response sustained over at least 6 months, in line with

recommendations for response in early course of the illness [5, 15]. Treatment non-response was defined as <50% reduction in PANSS total symptom severity rating from baseline and no evidence of response over at least 6 months. Of the 26 patients, 13 met criteria for treatment response and 13 met criteria for non-response [35, 36].

Dataset2 [32] consists of 12 treatment non-responsive patients with schizophrenia, 12 treatment-responsive patients with schizophrenia and 12 age matched healthy volunteers for normative comparisons. Treatment non-response was defined as meeting modified Kane criteria for treatment resistance in schizophrenia [37]. Treatment response was defined as meeting the Remission in Schizophrenia Working Group criteria for treatment remission [38]. Full details of inclusion criteria, medication status of the patients and clinical assessment are reported in the original references [28, 32]. For *Dataset1* and *Dataset2*, both treatment responders and non-responder patients show no significant difference for age, gender, weight, ethnicity and cigarette smoking with the corresponding control groups.

An additional test-retest dataset (*Dataset3*), consisting of eight healthy controls (mean age 23.6 ± 3.5 years, 5 male, injected dose of ~150 MBq) scanned twice, ~2 years apart, was also used for the reliability analysis of the proposed simplified [¹⁸F]FDOPA PET imaging protocol. Full details on the research protocol and subject inclusion criteria are reported in the original reference [39].

A summary of subject demographics for all the three datasets is reported in Supplementary Table 3. All studies were approved by the local research ethics committee. After full description of the respective studies, all participants gave written informed consent to participate, and consent for data to be used in further analyses.

PET imaging

The experimental protocol for the three datasets used the same standard approach [34, 40–42], although the target injected radioactivity was ~150 MBq for *Dataset1* and *Dataset3*, and ~180 MBq for *Dataset2*. All participants received carbidopa (150 mg) and entacapone (400 mg) orally 1 h before imaging. Both drugs are used to increase the signal-to-noise ratio (SNR) of the tracer uptake in brain tissue by reducing the peripheral formation of radiolabelled dopamine and 3-O-methyl-[¹⁸F]fluorodopa, the brain-penetrating metabolite, respectively [43–45].

The [¹⁸F]FDOPA tracer was administered by intravenous bolus injection after acquisition of a brain CT scan for attenuation correction. For *Dataset1* and *Dataset2*, [¹⁸F]FDOPA PET imaging was performed dynamically in 3-dimensional mode. Data were acquired using a Siemens Biograph 6 HiRez PET scanner (Siemens, Erlangen, Germany) for *Dataset1*, a Siemens/CTI ECAT HR + 962 PET scanner (Siemens, Erlangen, Germany) for *Dataset2* and a Siemens/CTI ECAT/EXACT3D (Knoxville, Tennessee) for *Dataset3*. The three machines have similar spatial resolution (4.5 ± 0.24, 4.8 ± 0.2 and 4.5 ± 0.2 mm respectively) and comparable sensitivity (4.2, 4.2 and 4.5 cps/kBq). In *Dataset1* PET data were binned in 32 frames of increasing duration over the 95 min scans (frame intervals in seconds: 8 × 15, 3 × 60, 5 × 120, 16 × 300). Emission data for *Dataset2* were obtained as 26 frames of increasing duration over 90 min (frame intervals in seconds: 1 × 30, 4 × 60, 3 × 120, 3 × 180, 15 × 300). In *Dataset3* PET data were acquired in list mode for 95 min, re-binned into 26 frames (frame intervals in seconds: 1 × 30, 4 × 60, 3 × 120, 3 × 180, 15 × 300).

Image analysis

[¹⁸F]FDOPA PET imaging analyses used the approach reported in previous papers [34, 40–42]. For all datasets, motion correction was performed first realigning frame-to-frame nonattenuated dynamic images to a single reference frame. The transformation parameters were then applied to the corresponding attenuated-corrected frames. The realigned frames were finally summed to create motion-corrected dynamic images. A tracer-specific

template [32, 41] and atlas defining the striatum and cerebellum (see [46]) were co-registered onto each subject's PET image using a combination of Statistical Parametric Mapping 8 (<https://www.fil.ion.ucl.ac.uk/spm/>) and in-house Matlab-based scripts. The striatal atlas included limbic, associative and sensorimotor subdivisions based on the predominant origin of projections to sub-regions of the striatum and given evidence that dopaminergic alterations are most marked in the associative striatum in schizophrenia [47]. The main outcome parameter was K_i^{cer} (min^{-1}), calculated using the Patlak–Gjedde graphical approach with the cerebellum as reference region, as this region shows negligible dopamine synthesis [48]. K_i^{cer} parametric images of the brain were constructed from motion-corrected images using a wavelet-based approach [49]. The parametric image for each participant was then normalised into Montreal Neurological Institute standard space using the participant's PET summation image and the [^{18}F] FDOPA PET template. This analysis pipeline is completely automated and operator independent, leading to reproducible ($\text{ICC} > 0.8$ [39]) and replicable results [27, 28, 50].

The simplified index of FDOPA uptake (SUVRc) was then defined by the ratio of the tracer activity in the striatum and its functional subdivisions to that in the reference region, using the same atlas as used for the dynamic scan. Consistent with the dynamic analysis, the cerebellum was used as the reference region. To test the sensitivity of this metric to time point after tracer injection, SUVRc was generated for the striatum and each functional striatal subdivision for a 10-min time frame at 60, 75 and 90 min after injection. These timepoints were chosen because the Gjedde–Patlak plot is linear within this time window, and it is the linear portion of the Gjedde–Patlak plot that is used to derive K_i^{cer} [48].

Classification analysis was limited to the whole striatum and its associative subdivision. Previous studies have identified the associative striatum as the striatal region with the most marked dopaminergic alteration in psychosis [51]. The associative striatum was also reported to be the main locus for differentiating responders from non-responders [32].

We investigated the impact of reducing or increasing the length of the simplified PET scan acquisition by 5 min using *Dataset3*. The functional striatal atlas was co-registered with the pseudo-static PET images obtained from taking the mean signal of the original dynamic scans over 5, 10 and 15 min, respectively starting at 75 min from the tracer injection. The co-registration was performed using Statistical Parametric Mapping 8 (<https://www.fil.ion.ucl.ac.uk/spm/>). The SNR was defined by the ratio of the mean signal in the striatum and the standard deviation of the brain signal outside the striatal regions [52].

Preliminary health economic analysis

We investigated the economic sustainability of the biomarker if it was offered as screening test in a representative population of 1000 individual with first-episode psychosis.

We used the following assumptions in the calculation:

- (1) Cost per scan £3000 (\$3900) (<https://www.xe.com/currencyconverter/>), for a static 10–15 min FDOPA scan in a clinical nuclear medicine setting inclusive of tracer delivery. This is based on the cost of commercial delivery of tracer plus scanning time and all other imaging centre costs as provided by current providers, although for national health system it can be cheaper [53].
- (2) In total, 33% of patients do not respond to conventional antipsychotics [54].
- (3) Potential annual healthcare savings of ~ £24,000/year based on the healthcare costs of a patient with active psychosis of £39,141/year and the healthcare costs of a patient with no active psychotic symptoms of £15,086 [55–58].

- (4) Cost of treating patients with clozapine (costs of medication, the monitoring service and management of neutropenia occurring in 3% of patients) for all the patients classified by the biomarker as non-responders [59, 60].
- (5) A mean delay to clozapine of 4 years in current clinical practice [61].
- (6) In total, 50% of patients with treatment-resistant schizophrenia achieve clinical response with clozapine treatment [62].

Based on these assumptions we calculated the potential annual savings of identifying a patient who is a non-responder to first-line antipsychotic treatment and starting them on clozapine within a month from the diagnosis and the minimal statistical performance of the biomarker to be cost effective in clinical use.

Statistical analysis

Statistical analyses were performed using Prism, version 7 (<https://www.graphpad.com/scientific-software/prism/>), and SPSS, version 24 (<https://www.ibm.com/analytics/spss-statistics-software>). Normality of distribution was assessed using Shapiro–Wilkes test. Test-retest reliability of SUVRc was calculated using the intraclass correlation coefficient (ICC) [63]. This model estimates the correlation between individual [^{18}F]FDOPA SUVRc values between scan sessions using a two-way ANOVA with random subject effects and fixed session effects. This model was chosen over a one-way random model because SUVRc values were ordered into two sessions (test and retest scan), and the two-way random model additionally accounts for systematic sources of variance associated with session effects [64]. Test-retest reproducibility was calculated as the percentage test-retest difference (VAR absolute value):

$$\text{VAR} = 2 \times |\text{Retest} - \text{Test}| / (\text{Test} + \text{Retest}) \times 100.$$

The agreement between SUVRc values and K_i^{cer} values determined using a full scan was studied with a correlation analysis. In this analysis, K_i^{cer} and SUVRc of all participants of *Dataset1* and *Dataset2* were correlated within each functional striatal ROI by computing Pearson's product moment correlations for normally distributed data and Spearman's rank correlation coefficients for non-normally distributed data. Secondly, we computed the responder vs. non-responder Cohen's d effect size to identify the degree of group difference in dopamine synthesis capacity as returned by SUVRc in comparison to K_i^{cer} .

To investigate the statistical power of SUVRc to identify non-responder patients from the entire patient pool, an analysis of receiver operating characteristics (ROC) curves was performed. *Dataset1* and *Dataset2* were analysed separately and the SUVRc metrics computed at different timepoints were compared with K_i^{cer} performances. The ROC area under the curve (AUC) was used as the performance index. As an additional performance metric, we extracted the biomarker sensitivity when the classification threshold was set at 100% specificity, corresponding to the fraction of non-responders correctly identified without any mislabelling of the responder group as non-responders. The classification threshold was set at this high level of specificity to avoid treatment with clozapine in someone who would respond to an alternative antipsychotic, given the monitoring and potential side effect burden associated with clozapine [65].

Machine-learning approaches applied to voxel-level data

The primary analyses of the discriminative power used K_i^{cer} values from the atlas-based region-of-interests for the whole striatum and its functional subdivisions, which represents an average value of all voxels within the region-of-interest. However, it is possible that other characteristics of the data, such as K_i^{cer} values in individual voxels and/or non-linear relationships within the data, are important in discriminating between responders and non-

responders; i.e., there might be sub-anatomical areas in the striatum that have better predictive power for the treatment response.

In view of this, we also investigated the potential of using machine-learning models to predict the non-responder status using the voxel-based PET data across the striatum to improve the discriminatory power compared to using mean striatal values from the region-of-interest analyses. We used the 2847 voxels from the striatum segmentation mask, with no additional pre-processing, and applied a leave-one-out cross validation approach to train and evaluate the classifiers.

We evaluated a variety of linear and non-linear models, using leave-one-out cross validation to estimate the generalisation of performance. The predictor, x_i , is the voxel-level dopamine synthesis values (flattened to a vector representation) and the target, y_i , is the responder status, with 1 for non-responder and -1 for responders. We considered two linear models with loss functions of the form:

$$\sum_{i=1}^N l(y_i, f_i) - \alpha \|W\|^2 \quad \text{with} \quad f_i = Wx_i + b,$$

where N is the number of the subjects, α is a regularisation parameter (fixed to 1, as good trade-off between data fitting and regularisation [66]), W is a vector of weights to be learned and b is a bias parameter to be learned.

The first was a Bernoulli observations model with a logistic link function [67]. The second was a linear Support Vector Machine [68], which uses a high loss to maximise the margin between classes. We also evaluated two non-linear models. First, we used a Random Forests approach [69], which avoids overfitting by averaging base classifiers in an ensemble technique known as bootstrap aggregating. The base classifiers are decision trees, which are de-correlated through data subsampling and feature subsampling. Second, we evaluated a Gaussian processes approaches [66], which uses kernelized Bayesian linear models, with the Squared Exponential kernel and Bernoulli loss with probit link function, fit using variational inference. All models were implemented in sklearn [70], except the Gaussian process, which was implemented in gpflow [71].

Supplementary analyses: sensitivity to experimental variables

To explore the generalisability of SUVRc beyond the specifics of the datasets included in this study, we tested its sensitivity to the tracer injected dose and SA by using Spearman's correlation analysis.

RESULTS

Static vs. dynamic FDOPA

For both datasets, the standard dynamic measure of dopamine synthesis capacity (K_i^{cer}) was significantly correlated with the simplified index of FDOPA uptake (SUVRc) in whole striatum (Spearman's rho from 0.60 to 0.89, all p values < 0.001). The correlation between striatal SUVRc and K_i^{cer} at 75 min after injection was the highest (Spearman's rho from 0.79 to 0.89), while it was lowest for SUVRc calculated at 90 min (Spearman's rho from 0.75 to 0.85). Significant correlations were also observed when considering only the patient groups (both responders and non-responders) in both datasets, with Spearman's rho ranging from 0.74 to 0.87 at 60 min, from 0.74 to 0.82 at 75 min and from 0.60 to 0.83 at 90 min after tracer injection. Full details of the correlation analysis results are reported in Supplementary Tables 4 and 5 and Supplementary Fig. 1. Mean values of K_i^{cer} and SUVRc of *Dataset1* and *Dataset2* are reported in Figs. 1 and 2.

For *Dataset1* (in which one of the three minimally treated patients was a responder) we observed significant differences in SUVRc between responders and non-responders with small to

very large Cohen's d effect sizes (mean \pm SD: 0.77 ± 0.37), with the maximum obtained for the associative striatal subdivision at 75 min from tracer injection (Cohen's d : 1.32) and the minimum obtained for the sensorimotor striatal subdivision at 90 min from tracer injection (Cohen's d : 0.17). For *Dataset2*, we also observed significant differences in SUVRc between responders and non-responders, with moderate to very large Cohen's d effect sizes (mean \pm SD: 0.69 ± 0.08). As for *Dataset1*, the highest values were obtained for the striatum and the associative striatal subdivision at 75 min from tracer injection (both with Cohen's d effect sizes of 0.79), while the smallest difference was obtained for the sensorimotor striatal subdivision at 60 min from tracer injection (Cohen's d : 0.57). Overall the whole striatum and its associative striatal subdivision were the areas where the effect sizes for the differences between responders and non-responders were the highest (Cohen's d values in the striatum: from 0.62 to 1.09; Cohen's d values in the associative striatum: from 0.64 to 1.32).

Scan timing did have an impact on effect size for SUVRc. The lowest values were obtained at 90 min from tracer injection (mean \pm SD: 0.55 ± 0.22) while the highest ones were obtained at 75 min from tracer injection (mean \pm SD: 0.81 ± 0.24) in both datasets. Full details of the Cohen's d effect size are reported in Supplementary Table 6.

Test-retest analysis

Mean percent test-retest differences and reliability for SUVRc values in the striatum and in the functional striatal subdivisions are presented in Supplementary Table 7. Overall, the reproducibility of SUVRc was excellent (%VAR < 4), regardless of the time after injection and the striatal area in which it was measured. We also observed good test-retest reliability values in all the functional striatal subdivisions (ICC: 0.76–0.91).

Sensitivity to experimental variables

There was no correlation between SUVRc and either injected dose ($p = 0.89$) or SA ($p = 0.97$) (Supplementary Fig. 2). In contrast, there was a significant effect of scan duration on SUVRc maps (Supplementary Fig. 3): the signal-to-noise ratio increased as the acquisition length was extended (mean \pm SD: for 5 min acquisition 0.76 ± 0.048 ; for 10 min acquisition 0.87 ± 0.49 ; for 15 min acquisition 0.88 ± 0.55).

Classification of treatment response: region-of-interest analyses

The ROC curves for SUVRc from 75 min after tracer injection and K_i^{cer} values in the whole striatum and in its associative striatal subdivision are shown in Fig. 3, while the ROC curves for SUVRc from 60 and 90 min after tracer injection are reported in Supplementary Fig. 4. AUC estimates and the sensitivity at 100% specificity are reported in Table 1.

For *Dataset1* we observed ROC AUC estimates varying from 0.66 to 0.80 (mean \pm SD: 0.74 ± 0.05). For *Dataset2* the AUC estimates varied from 0.74 to 0.87 (mean \pm SD: 0.84 ± 0.03). The best estimates were observed at 75 min after tracer injection (mean \pm SD: 0.82 ± 0.04) for both the striatum and the associative striatal subdivision.

For *Dataset1* the K_i^{cer} sensitivity to identify non-responders at 100% specificity varied from 46 to 61% (mean \pm SD: 53.5 ± 10.6), whereas the SUVRc sensitivity varied from 23 to 46% (mean \pm SD: 32.2 ± 9.0). For *Dataset2* we observed K_i^{cer} sensitivity at 100% specificity of 50% and SUVRc sensitivity varying from 36 to 79% (mean \pm SD: 60.8 ± 17.9).

Classification of treatment response: machine-learning voxel-based analyses

We also investigated the potential of using machine-learning approaches applied to striatal voxel-wise data to distinguish responders from non-responders. For both the linear and non-linear methods the AUC estimates are given in Table 2. A

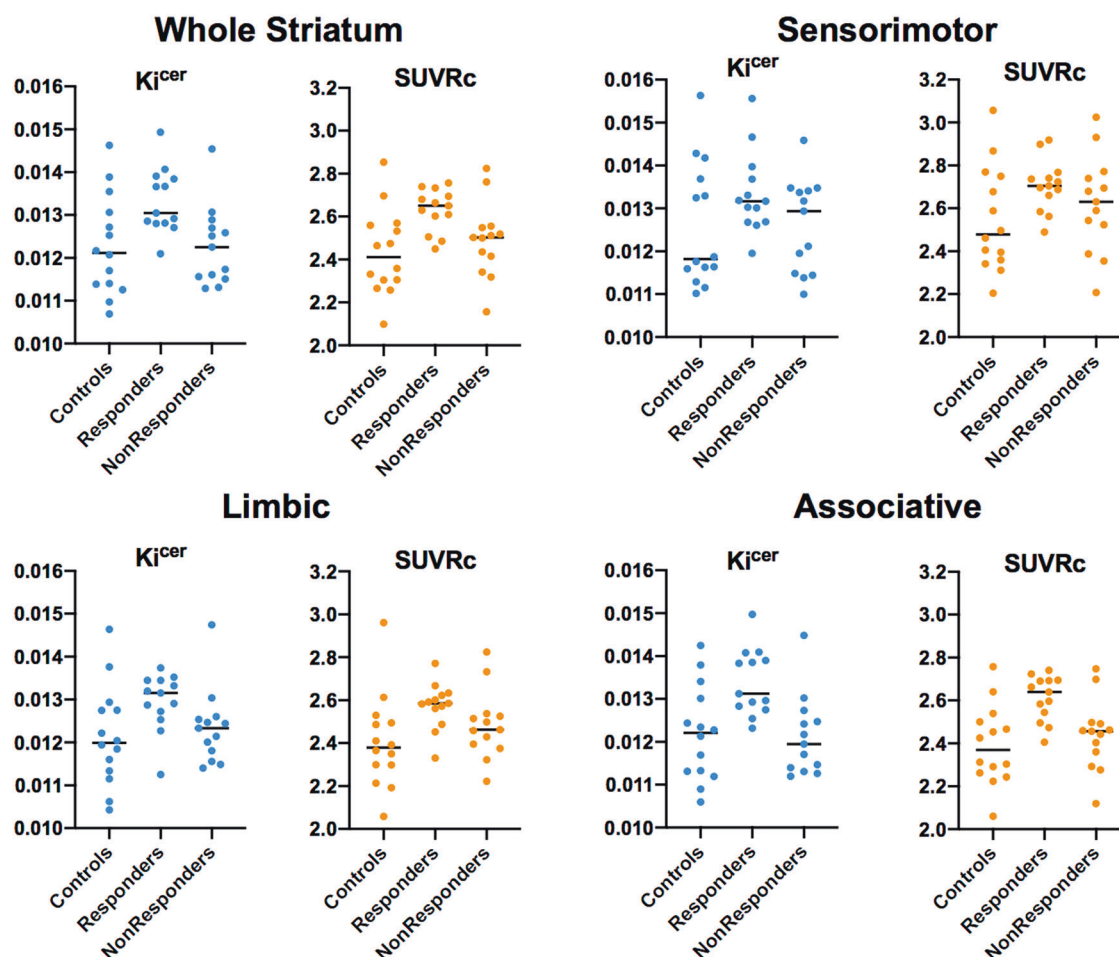


Fig. 1 Gold standard index of dopamine synthesis capacity (K_i^{cer} using 95 min acquisition) and the simplified index of FDOPA uptake (SUVRc with a 15 min static acquisition 75 min after tracer injection) in striatum and its subdivisions. These values are from each group (controls, responders and non-responders) of Dataset1.

consistent pattern was a higher score in *Dataset2* than *Dataset1*. For *Dataset1* none of the models surpassed the ROI classifier as measured by ROC AUC (mean \pm SD AUC for linear SVM = 0.71 ± 0.001 , for SVM with radial basis function = 0.45 ± 0.025 , for Random Forest = 0.70 ± 0.008 , Linear = 0.74 , for Kernel K-nearest with 2 neighbours Knn(2) = 0.63 , for Kernel K-nearest with 3 neighbours Knn(3) = 0.68 , and for Gaussian process approach = 0.64). For *Dataset2* the linear model and linear Support Vector Machine model provided superior classification than the ROI classifier as measured by ROC AUC (mean \pm SD AUC for linear SVM = 0.88 ± 0.001 , for SVM with radial basis function = 0.80 ± 0.007 , for Random Forest = 0.83 ± 0.006 , Linear = 0.89 , for Kernel K-nearest with 2 neighbours Knn(2) = 0.75 , for Kernel K-nearest with 3 neighbours Knn(3) = 0.79 , and for Gaussian process approach = 0.83). We also included results from combining both datasets to a single one and obtained the following results in term of ROC AUC: linear SVM = 0.89 ± 0.001 , SVM with radial basis function = 0.80 ± 0.001 , Random Forest = 0.76 ± 0.004 , Linear = 0.87 , Kernel K-nearest with 2 neighbours Knn(2) = 0.74 , Kernel K-nearest with 3 neighbours Knn(3) = 0.74 , Gaussian process = 0.81 .

Illustrative health economic model

By averaging the classification performance of the SUVRc from linear and non-linear analyses (both datasets) we estimated it is capable of identifying nearly one half of the (33% of) patients destined to prove unresponsive to first-line treatment.

By combining this estimate with our model cost assumptions (see "Methods"), and assuming a specificity of 95% (i.e., 5% misclassification of responders), we obtained a total saving of £3400/patient. With these assumptions, the breakeven sensitivity of the classification test is 26% (Supplementary Fig. 5). While sensitivity modulates the potential savings of the biomarker, specificity has a strong effect on costs (Supplementary Fig. 6). By reducing the specificity of the test to 80%, the breakeven sensitivity will increase to 34.2%, and, at a sensitivity of 50%, there will be a saving per patient of \sim £2000. Note that with both specificity and sensitivity at 50% (chance level), the biomarker would stop being economically favourable (Supplementary Table 8).

DISCUSSION

Our first main finding is that a dynamic [18 F]FDOPA PET scan is able to identify 40–60% of treatment non-responsive patients with a specificity of 100%. Our second main finding is that a simplified 10 min [18 F]FDOPA PET imaging scan acquisition shows good test-retest reliability (ICC = 0.76 – 0.91), good agreement with the dynamic [18 F]FDOPA PET, and is able to identify 40–60% of treatment non-responsive patients with a specificity of 100%. Based on this, our economic modelling indicates the use of the simplified [18 F]FDOPA PET imaging to guide early use of clozapine has a potential cost saving of £3.4 million per 1000 patients

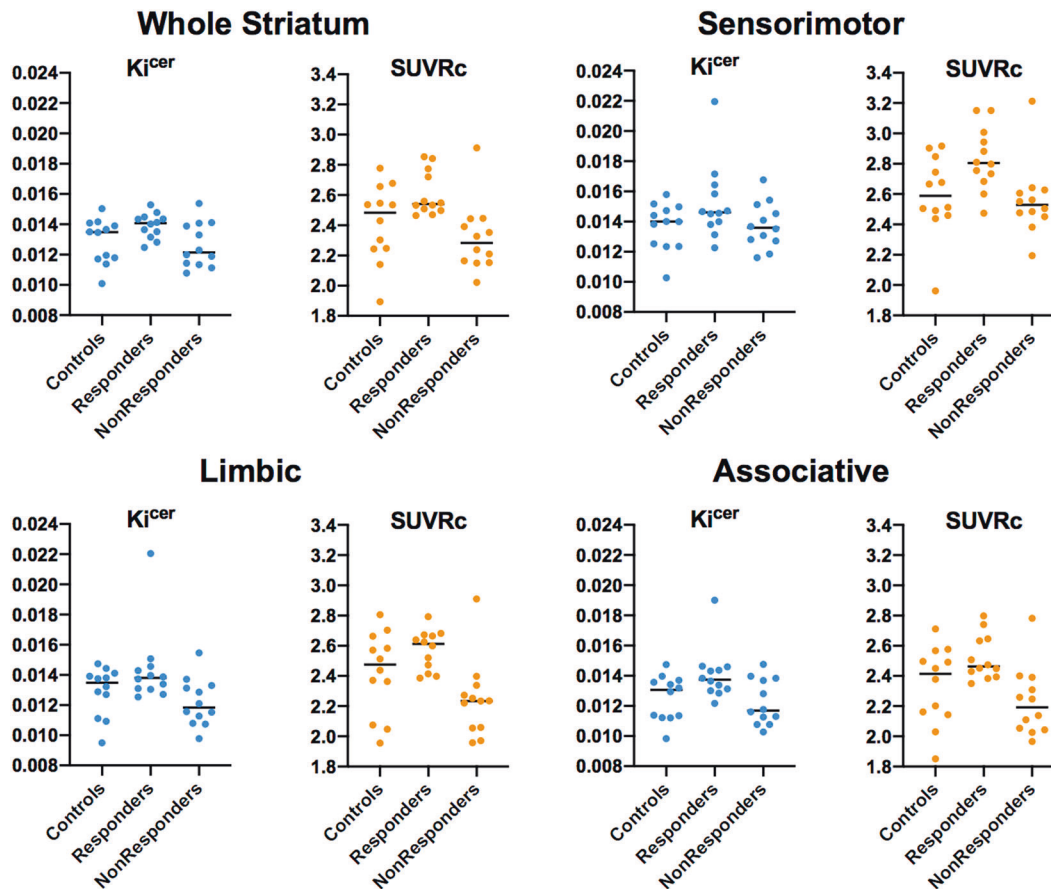


Fig. 2 Gold standard index of dopamine synthesis capacity (K_i^{cer} using 95 min acquisition) and the simplified index of FDOPA uptake (SUVRc with a 15 min static acquisition 75 min after tracer injection) in striatum and its subdivisions. These values are from each group (controls, responders and non-responders) of Dataset2. For consistency, dataset 2 was reprocessed with the same analysis pipeline used for dataset 1, which may explain differences between data points shown here and the original publication [32].

relative to current practice. However, our illustrative health economic analysis considered only direct health service costs; if social costs were included (e.g., criminal justice service costs), the potential cost savings is likely to be greater [72]. These findings extend previous molecular imaging studies reporting presynaptic striatal dopamine measures are associated with treatment response [20, 25, 27, 28, 32] by showing the clinical potential of these approaches if used in clinical practice.

In absolute terms the machine-learning analyses showed similar performance to the ROI analyses, although the linear and support vector machine classifiers were marginally superior to the ROI classifier for the medicated patient dataset (*Dataset2*). Recent meta-analyses of machine-learning methods have shown no or limited improvements over logistic regression in a medical context [73]. Our findings are generally consistent with these findings but extend them by showing that a logistic regression model shows no appreciable improvement in classification relative to the simple averaging approach. In our case the small size of the datasets compared with the high dimensionality of the features is likely to cause machine-learning and other data-driven methods to fail to capture reliable signal without further assumptions [74]. Further work to validate decision algorithms should be repeated considering larger samples following the guidelines of Steyerberg [75] and the PRoGRESS group [76].

The test-retest reproducibility and reliability of the simplified index of FDOPA uptake are in line with the ICC (0.68–0.94) and % VAR (0.7) values for K_i^{cer} in the striatal subdivisions reported in Egerton et al. using a full dynamic acquisition [39]. These values

indicate that the simplified imaging approach does not sacrifice reliability or increase variability compared to the standard, dynamic method.

$[^{18}F]$ FDOPA PET imaging is not the first method that has been proposed for early identification of treatment response in psychosis. Peripheral biomarkers (e.g., insulin levels, metabolite levels in blood and urine) have been shown to be linked to treatment response in psychosis [77, 78], but have not been consistent among studies [79]. Moreover, these findings were not validated using a cohort of responders and non-responders [78]. Mondelli et al. showed that lower cortical awakening response and increased levels of inflammatory markers are associated with poor treatment response in first-episode psychosis patients [80]. However, this study does not report any classification analysis and it has not been replicated to date. In the domain of neuroimaging, structural techniques have shown that differences in gyrification, cortical thickness and asymmetry are associated with subsequent response to antipsychotic treatments [81, 82]. Reduced white matter integrity has also been linked to non-response to treatment in patients with first-episode psychosis [83], and functional connectivity of the ventral tegmental area has been associated with treatment response [84]. Elevated anterior cingulate cortex glutamate levels have also been reported to be associated with treatment resistance compared to treatment-responsive schizophrenia patients [85]. However, none of these studies reported a validation analysis of the predictive power of the proposed measure, AUC values nor sensitivity/specificity classification. Moreover, these studies include one dataset only.

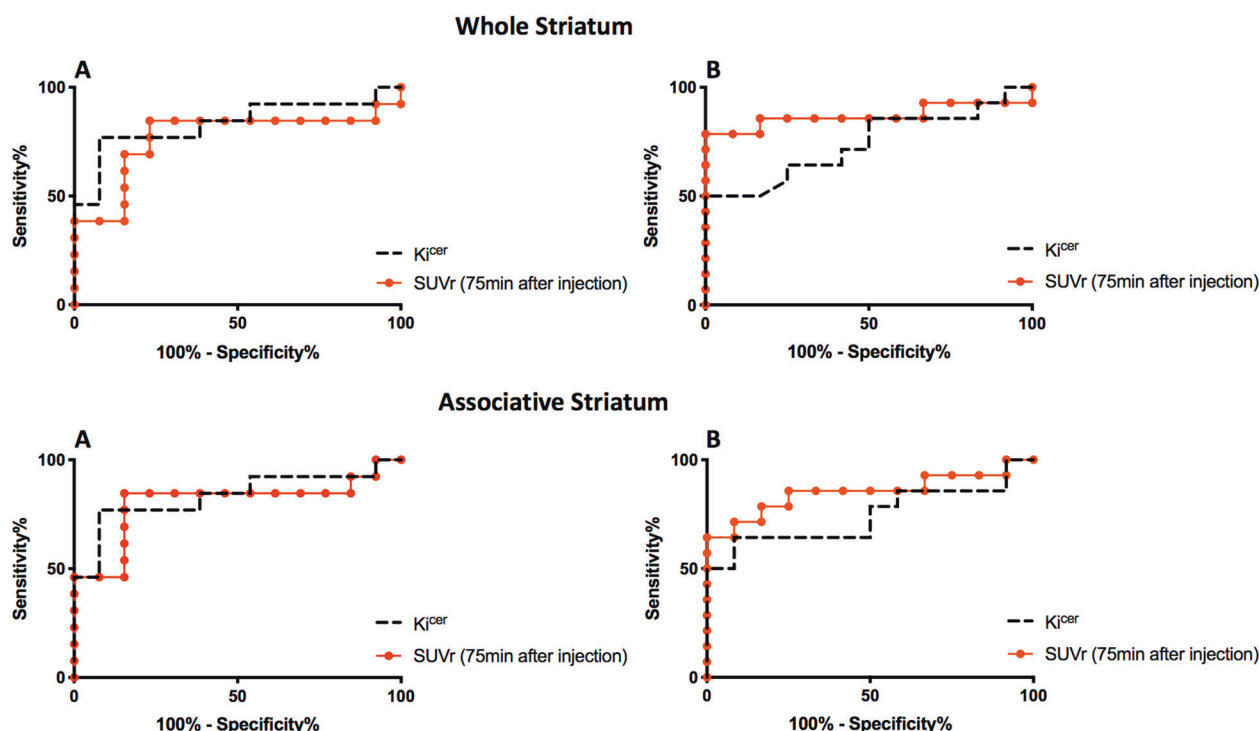


Fig. 3 Receiver Operating Characteristic (ROC) curves for the classification of patients into responders and non-responder groups using [¹⁸F]FDOPA PET imaging analysed using the simplified index of FDOPA uptake (SUVRc) or the standard dynamic measure of dopamine synthesis capacity (K_i^{cer}) in the whole striatum and associative striatal subdivision. This analysis shows the results from Dataset1 (a) and Dataset2 (b).

Table 1. Receiver Operating Characteristic (ROC) curves and sensitivity analyses for the classification of patients into antipsychotic responders and non-responders using the simplified index of FDOPA uptake (SUVRc) and the gold standard (K_i^{cer}) index of dopamine synthesis capacity in the striatum and associative striatal subdivision for Dataset1 and Dataset2.

	AUC ROC				Sensitivity at 100% specificity cut-off			
	Whole striatum		Associative		Whole striatum		Associative	
	Dataset1	Dataset2	Dataset1	Dataset2	Dataset1 (%)	Dataset2 (%)	Dataset1 (%)	Dataset2 (%)
SUVRc 60 min	0.70*	0.86*	0.74*	0.83*	31	79	31	64
SUVRc 75 min	0.77*	0.87*	0.80*	0.85*	39	79	46	64
SUVRc 90 min	0.66 ^{NS}	0.81*	0.74*	0.81*	23	43	23	36
K_i^{cer}	0.83*	0.74*	0.87*	0.75*	46	50	61	50

The *p* value tests the null hypothesis that the area under the curve really equals 0.50 (chance).

AUC area under the curve.

**p* < 0.05; NS *p* > 0.05.

Thus, our study extends these previous peripheral biomarker and neuroimaging studies by including these analyses and considering two independent datasets.

A recent study provided evidence that individual differences in striatal functional connectivity predict response to antipsychotic treatment in acutely psychotic patients (AUC: 0.78, sensitivity: 80% and specificity: 75%) [86]. Our potential biomarker showed predictive power that was higher than this (AUC of 0.89). However, as they did not report sensitivity at 100% specificity or cost analyses, it is not clear how it performs on these metrics. It would be useful to directly compare both approaches in a future study.

Strengths and limitations

Strengths of our work include the use of an in vivo measure of the pathophysiology of psychosis relevant to the mechanism of action of antipsychotic treatments, and that we were able to test the [¹⁸F]

FDOPA PET imaging approach in two independent datasets and test reliability of the simplified protocol in a third dataset. However, by the very nature of the cohorts recruited (first-episode psychosis, people with established illness) the criteria for treatment response were different in the two datasets, as outlined in "Methods". This is reflected in the difference in sensitivity, 39% vs. 79%. It has also been shown that demographics and environmental factors, such as age, gender, smoking, ethnicity and childhood trauma, may influence dopamine synthesis capacity [87, 88]. The relative contribution of state and trait related factors to dopamine alterations in psychosis remains to be determined [89]. Nevertheless, the pathological mechanisms underlying the disorder should not be confused with the capacity of the biomarker to predict a clinical response to treatment. An ideal biomarker should return accurate results irrespective of the history of the patients to which it is applied [90]. Moreover, whilst

Table 2. AUC results for machine-learning methods, averaged over 100 random seeds (1 standard error).

	SVM linear (mean ± SD)	SVM rbf (mean ± SD)	Random Forest (mean ± SD)	Linear	Knn (2n)	Knn (3n)	GP
Dataset1	0.71 ± 0.001	0.45 ± 0.025	0.70 ± 0.008	0.74	0.63	0.68	0.64
Dataset2	0.88 ± 0.001	0.80 ± 0.007	0.83 ± 0.006	0.89	0.75	0.79	0.83
Both datasets	0.89 ± 0.001	0.80 ± 0.001	0.76 ± 0.004	0.87	0.74	0.74	0.81

SVM Support Vector Machine, SVM rbf Support Vector Machine with radial basis function, *Kernels*. Knn (2n) K-nearest with 2 neighbours, Knn (3n) K-nearest with 3 neighbours, GP Gaussian Process.

our sample sizes are large for PET imaging studies ($N = 84$ individuals pooled together from 3 different studies), it is important to recognise that the samples are relatively modest, and our findings will require testing in other samples and settings to determine their generalisability, and if clinical-demographic factors influence results, before they can be translated into routine clinical practice. Assuming a ROC AUC = 0.7, a marginal error of 10%, a conservative frequency of non-responding patients of 20% and a drop-out rate of 10%, a power calculation in easyROC (v1.3.1) [91] indicates it would take 84 first-episode patients to test the biomarker accuracy to predict treatment response in a prospective study with power = 0.80, and $\alpha = 0.05$.

Of particular interest might be the use of PET-MR scanners to obtain hybrid samples that could provide additional information to include in the presented machine-learning tool. This could be particularly relevant for those MRI modalities that provided complementary information on the dopamine system, like neuromelanin-MRI [92] or dopamine enriched functional connectivity [93, 94]. Finally, we calculated SUVRc retrospectively from a 10-min time frame of a dynamic PET scan and this warrants testing in a prospective study. The AUC in the receiver operating curve analyses were between 0.66 and 0.87 and statistically higher than chance performance.

Performance of this simplified FDOPA PET is dependent on the experimental design, including the co-administration of carbidopa and entacapone. These are recommended to enhance the specific brain FDOPA signal [27] by reducing metabolite production [95], but future studies would be useful to determine the importance of this for classification accuracy. All the datasets were collected by the same research group and used the same experimental design, imaging acquisition and analysis pipeline. Future studies should include multi-centre acquisitions to investigate the generalisability of the biomarker across settings.

In our preliminary economic analysis, we set specificity at 95% and found substantial cost savings. A prospective study is now needed to determine if the biomarker can achieve 95% specificity in routine clinical practice. Nevertheless, our further analyses show that the biomarker has the potential to be cost effective with specificity and sensitivity values as low as 0.6. This is in line with the findings of a recent modelling study of a test to identify patients for clozapine [59], which showed that a stratified test with 60% sensitivity and specificity is still cost effective compared to no test for people with first-episode psychosis who failed a first-line antipsychotic. Whilst we accounted for the cost of managing neutropenia, we did not factor in the costs of other side effects. However, in contrast with the risk of neutropenia, the comparative risks of other side effects of clozapine and first-line antipsychotics are relatively similar [65, 96, 97], suggesting that there is unlikely to be a large cost differential. Notwithstanding this, a limitation of our economic analysis is that the costs of side effects have only been partially considered.

Implications of our findings

Our results have two main implications. The first is that in the two examined samples we were able to identify 50% of non-

responders without overlap with responders; these results need to be replicated in future studies. This could enable early use of clozapine, which is associated with improved outcomes [98]. Studies have reported that early intervention for psychosis via therapeutic biomarkers could lead also to long-term economic benefit [72, 99–101]. Hospitalisation costs are the main component of the higher total healthcare costs associated with treatment-resistant schizophrenia [102], with approximately half of these costs attributable to hospital admissions [8]. Health economic evidence suggests that the use of clozapine has the potential to improve the use of health and social service resources in treatment-resistant schizophrenia [103], through shorter inpatient admissions [11, 104], lower rates of relapse and reduced rehospitalization [105–108]. A cost-benefit model study within the Veterans Health Administration in the USA compared costs with clozapine use and non-use in treatment-resistant schizophrenia. It identified that, if 20% of treatment-resistant cases started clozapine, this would be associated with an average annual reduction in health costs of \$22,444 per patient treated with clozapine [109].

Our preliminary economic analysis indicates that screening test with [^{18}F]FDOPA PET would lead to an average healthcare saving of £3400 per patient with schizophrenia. However, these potential benefits need to be tested in a prospective study and assume that all non-responders identified go on to clozapine. Another consideration for translation is the practicalities of [^{18}F]FDOPA PET scanning. Practicability of PET in first-episode psychosis was demonstrated both in research samples [34, 40–42], suggesting that 10-min [^{18}F]FDOPA PET scan could be potentially well tolerated, but this needs evaluating in clinical settings. In support of [^{18}F]FDOPA PET practicability in a clinical environment the key evidence is:

- (1) FDOPA is already approved by the European Medicines Agency as a diagnostic agent (EMA/729548/2018), meaning its manufacture is standardised, facilitating introduction to health services.
- (2) FDOPA imaging is already approved by the FDA for the diagnosis of Parkinson's Disease (2019, ID: 4504654). As such, the extension to another indication would be relatively straightforward.
- (3) FDOPA has a sufficiently long half-life to enable it to be produced in a distribution centre, and then delivered to hospitals. This is well established for clinical oncology PET imaging with other tracers like FDG.

If this approach is to be implemented, parameter values and operating procedures would need to be established for a range of scanners and sites. The former could be done using normative data from healthy volunteers or brain phantoms scanned across different centres and scanners. The second implication is that [^{18}F]FDOPA PET imaging could be used to identify non-responders for inclusion in trials of novel treatments for non-response. This could be used to enrich trials for patients who might respond to a novel therapy.

CONCLUSIONS

Our study showed that a short [¹⁸F]FDOPA PET imaging protocol provides reliable and reproducible measures of dopamine synthesis and that it and the full dynamic [¹⁸F]FDOPA PET imaging protocol can be used to distinguish patients with schizophrenia who are unlikely to respond to first-line antipsychotic treatments from those who will respond at first episode. The classification accuracy of the proposed simplified imaging method is comparable to that from a dynamic [¹⁸F]FDOPA PET scan. These results support the development of the short [¹⁸F]FDOPA scan as a pathophysiologically relevant biomarker to guide therapy choice for patients with psychosis.

FUNDING AND DISCLOSURE

This study was funded by Medical Research Council-UK (no. MC-A656-5QD30), Maudsley Charity (no. 666), Brain and Behaviour Research Foundation, and Wellcome Trust (no. 094849/Z/10/Z) grants to OH and the National Institute for Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health. MV is funded by the National Institute for Health Research Biomedical Research Centre at South London and Maudsley National Health Service Foundation Trust and King's College London, by the Wellcome Trust Digital Award 215747/Z/19/Z. SJ is funded by the National Institute for Health Research Biomedical Research Centre at South London and Maudsley National Health Service Foundation Trust and King's College London, and a JMAS SIM Fellowship from the Royal College of Physicians, Edinburgh. OH has received investigator-initiated research funding from and/or participated in advisory/speaker meetings organised by Angellini, Astra-Zeneca, Autifony, Biogen, Boehringer-Ingelheim, Eli Lilly, Heptares, Invicro, Jansenn, Lundbeck, Lyden-Delta, Mylan, Neurocrine, Otsuka, Sunovion, Rand, Recordati, and Roche. Neither OH nor his family have been employed by or have holdings/a financial stake in any pharmaceutical company. MV has received consulting honoraria from GSK. OH and MV hold a patent application for the use of dopamine imaging as a prognostic tool. SJ has received honoraria for educational talks from Sunovion and his employer (KCL) has received honoraria for educational talks he has given for Lundbeck. SJ is also Co-PI for a drug trial for a product manufactured by Alkermes. All the other authors have nothing to disclose.

AUTHOR CONTRIBUTIONS

BS, MV, SJ and OH contributed to the written recording and interpretation of data. BS and MV contributed also to the analysis of data. BS, MV, SJ, FT and OH contributed to the conception and design of the work. EDA and AD contributed to the acquisition of the data. HS contributed to the design and the written recording of the machine-learning approaching. JH and PM contributed to the design and the written recording of the economic model.

ADDITIONAL INFORMATION

Supplementary Information accompanies this paper at (<https://doi.org/10.1038/s41386-020-00866-7>).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES

1. Howes OD, Murray RM. Schizophrenia: an integrated sociodevelopmental-cognitive model. *Lancet*. 2014;383:1677–87.
2. Murray CJ, Lopez AD. Global mortality, disability, and the contribution of risk factors: global burden of disease study. *Lancet*. 1997;349:1436–42.

3. Wittchen HU, et al. The need for a behavioural science focus in research on mental health and mental disorders. *Int J Methods Psychiatr Res*. 2014;23:28–40.
4. Howes O, et al. Mechanisms underlying psychosis and antipsychotic treatment response in schizophrenia: insights from PET and SPECT imaging. *Curr Pharm Des*. 2009;15:2550–9.
5. Leucht S, et al. Second-generation versus first-generation antipsychotic drugs for schizophrenia: a meta-analysis. *Lancet*. 2009;373:31–41.
6. Howes OD, et al. The role of genes, stress, and dopamine in the development of schizophrenia. *Biol Psychiatry*. 2017;81:9–20.
7. Lally J, et al. Augmentation of clozapine with electroconvulsive therapy in treatment resistant schizophrenia: a systematic review and meta-analysis. *Schizophrenia Res*. 2016;171:215–24.
8. Kennedy JL, et al. The social and economic burden of treatment-resistant schizophrenia: a systematic literature review. *Int Clin Psychopharmacol*. 2014;29:63–76.
9. Howes OD, et al. The nature of dopamine dysfunction in schizophrenia and what this means for treatment: meta-analysis of imaging studies. *Arch Gen Psychiatry*. 2012;69:776–86.
10. Buchanan RW, et al. Positive and negative symptom response to clozapine in schizophrenic patients with and without the deficit syndrome. *Am J Psychiatry*. 1998;155:751–60.
11. Hayhurst K, Brown P, Lewis S. The cost-effectiveness of clozapine: a controlled, population-based, mirror-image study. *J Psychopharmacol*. 2002;16:169–75.
12. National Collaborating Centre for Mental Health (UK). Schizophrenia: core interventions in the treatment and management of schizophrenia in primary and secondary care (update) [Internet]. Leicester (UK): British Psychological Society; 2009 Mar. PMID: 20704054.
13. Tiihonen J, et al. A nationwide cohort study of oral and depot antipsychotics after first hospitalization for schizophrenia. *Am J Psychiatry*. 2011;168:603–9.
14. Tiihonen J, et al. 11-year follow-up of mortality in patients with schizophrenia: a population-based cohort study (FIN11 study). *Lancet*. 2009;374:620–7.
15. Howes OD, et al. Treatment-resistant schizophrenia: treatment response and resistance in psychosis (TRRIP) working group consensus guidelines on diagnosis and terminology. *Am J Psychiatry*. 2016;174:216–29.
16. Howes O, (ed). *Treatment Response and Resistance in Schizophrenia*. Oxford University Press; 2018.
17. Davis KL, Kahn RS. Dopamine in schizophrenia: a review and reconceptualization. *Am J Psychiatry*. 1991;148:1474.
18. Howes O, McCutcheon R, Stone J. Glutamate and dopamine in schizophrenia: an update for the 21st century. *J Psychopharmacol*. 2015;29:97–115.
19. Howes OD, Kapur S. The dopamine hypothesis of schizophrenia: version III—the final common pathway. *Schizophrenia Bull*. 2009;35:549–62.
20. Abi-Dargham A, et al. Increased baseline occupancy of D2 receptors by dopamine in schizophrenia. *Proc Natl Acad Sci USA*. 2000;97:8104–9.
21. Mizrahi R, et al. Increased stress-induced dopamine release in psychosis. *Biol Psychiatry*. 2012;71:561–7.
22. Kapur S, et al. Relationship between dopamine D2 occupancy, clinical response, and side effects: a double-blind PET study of first-episode schizophrenia. *Am J Psychiatry*. 2000;157:514–20.
23. Jauhar S, et al. Regulation of dopaminergic function: an [18 F]-DOPA PET apomorphine challenge study in humans. *Transl Psychiatry*. 2017;7:e1027.
24. Howes O, et al. Progressive increase in striatal dopamine synthesis capacity as patients develop psychosis: a PET study. *Mol Psychiatry*. 2011;16:885.
25. Demjaha A, et al. Antipsychotic treatment resistance in schizophrenia associated with elevated glutamate levels but normal dopamine function. *Biol Psychiatry*. 2014;75:e11–3.
26. Egerton A, et al. Presynaptic striatal dopamine dysfunction in people at ultra-high risk for psychosis: findings in a second cohort. *Biol Psychiatry*. 2013;74:106–12.
27. Kim E, et al. Presynaptic dopamine capacity in patients with treatment-resistant schizophrenia taking clozapine: an [18 F] DOPA PET Study. *Neuropsychopharmacology*. 2017;42:941.
28. Jauhar S, et al. Determinants of treatment response in first-episode psychosis: an 18 F-DOPA PET study. *Mol Psychiatry*. 2018;24:1.
29. Jauhar S, Howes OD. Understanding and predicting variability in response to treatment in psychotic disorders: in vivo findings. *Clin Pharmacol Ther*. 2019;105:1079.
30. Dhawan V, et al. Comparative analysis of striatal FDOPA uptake in Parkinson's disease: ratio method versus graphical approach. *J Nucl Med*. 2002;43:1324.
31. Jokinen P, et al. Simple ratio analysis of 18F-fluorodopa uptake in striatal sub-regions separates patients with early Parkinson disease from healthy controls. *J Nucl Med*. 2009;50:893.
32. Demjaha A, et al. Dopamine synthesis capacity in patients with treatment-resistant schizophrenia. *Am J Psychiatry*. 2012;169:1203–10.

33. Subotnik KL, et al. Risperidone nonadherence and return of positive symptoms in the early course of schizophrenia. *Am J Psychiatry*. 2011;168:286–92.
34. Jauhar S, et al. The effects of antipsychotic treatment on presynaptic dopamine synthesis capacity in first episode psychosis: a positron emission tomography study. *Biol Psychiatry*. 2018;85:79–87.
35. Busner J, Targum SD. The clinical global impressions scale: applying a research tool in clinical practice. *Psychiatry*. 2007;4:28.
36. Leucht S, et al. Definitions of response and remission in schizophrenia: recommendations for their use and their presentation. *Acta Psychiatr Scand*. 2009;119:7–14.
37. Conley RR, Kelly DL. Management of treatment resistance in schizophrenia. *Biol Psychiatry*. 2001;50:898–911.
38. Andreasen NC, et al. Remission in schizophrenia: proposed criteria and rationale for consensus. *Am J Psychiatry*. 2005;162:441–9.
39. Egerton A, et al. The test–retest reliability of 18F-DOPA PET in assessing striatal and extrastriatal presynaptic dopaminergic function. *Neuroimage*. 2010;50:524–31.
40. Howes OD, et al. Dopamine synthesis capacity before onset of psychosis: a prospective [18F]-DOPA PET imaging study. *Am J Psychiatry*. 2011;168:1311–7.
41. Howes OD, et al. Elevated striatal dopamine function linked to prodromal signs of schizophrenia. *Arch Gen Psychiatry*. 2009;66:13–20.
42. Howes OD, et al. Dopaminergic function in the psychosis spectrum: an [18F]-DOPA imaging study in healthy individuals with auditory hallucinations. *Schizophrenia Bull*. 2012;39:807–14.
43. Cumming P, et al. Pharmacokinetics of plasma 6-[18F] fluoro-L-3, 4-dihydroxyphenylalanine ([18F] Fdopa) in humans. *J Cereb Blood Flow Metab*. 1993;13:668–75.
44. Hoffman JM, et al. The effects of carbidopa administration on 6-[18F] fluoro-L-dopa kinetics in positron emission tomography. *J Nucl Med*. 1992;33:1472–7.
45. Sawle G, et al. The effect of entacapone (OR-611) on brain [18F]-6-L-fluorodopa metabolism Implications for levodopa therapy of Parkinson's disease. *Neurology*. 1994;44:1292.
46. Martinez D, et al. Imaging human mesolimbic dopamine transmission with positron emission tomography. Part II: amphetamine-induced dopamine release in the functional subdivisions of the striatum. *J Cereb Blood Flow Metab*. 2003;23:285–300.
47. McCutcheon R, et al. Defining the locus of dopaminergic dysfunction in schizophrenia: a meta-analysis and test of the mesolimbic hypothesis. *Schizophrenia Bull*. 2018;44:1301–11.
48. Kumakura Y, Cumming P. PET studies of cerebral levodopa metabolism: a review of clinical findings and modeling approaches. *Neuroscientist*. 2009;15:635–50.
49. Turkheimer FE, et al. Multi-resolution Bayesian regression in PET dynamic studies using wavelets. *Neuroimage*. 2006;32:111–21.
50. Cheng PWC, et al. The role of dopamine dysregulation and evidence for the transdiagnostic nature of elevated dopamine synthesis in psychosis: a positron emission tomography (PET) study comparing schizophrenia, delusional disorder, and other psychotic disorders. *Neuropsychopharmacology*. 2020;45:1–7.
51. Kegeles LS, et al. Increased synaptic dopamine function in associative regions of the striatum in schizophrenia. *Arch Gen Psychiatry*. 2010;67:231–9.
52. Bushberg JT, Boone JM. The essential physics of medical imaging. Lippincott Williams & Wilkins; 2011.
53. NHS. 2019. <https://improvement.nhs.uk/resources/national-tariff/#h2-tariff-documents>.
54. Demjaha A, et al. Antipsychotic treatment resistance in first-episode psychosis: prevalence, subtypes and predictors. *Psychological Med*. 2017;47:1981–9.
55. Akhtar W, Chung Y. Saving the NHS one blood test at a time. *BMJ Open Qual*. 2014;2:u204012.w1749.
56. Briggs A, et al. Impact of schizophrenia and schizophrenia treatment-related adverse events on quality of life: direct utility elicitation. *Health Qual Life Outcomes*. 2008;6:105.
57. Curtis LA, Burns A. Unit costs of health and social care 2015. Personal Social Services Research Unit; 2015.
58. Taylor M, Perera U. NICE CG178 psychosis and schizophrenia in adults: treatment and management—an evidence-based guideline?. *Br J Psychiatry*. 2015;206:357–9.
59. Jin H, McCrone P, MacCabe JH. Stratified medicine in schizophrenia: how accurate would a test of drug response need to be to achieve cost-effective improvements in quality of life?. *Eur J Health Econ*. 2019;20:1425–35.
60. Atkin K, et al. Neutropenia and agranulocytosis in patients receiving clozapine in the UK and Ireland. *Br J Psychiatry*. 1996;169:483–8.
61. John AP, Ko EKF, Dominic A. Delayed initiation of clozapine continues to be a substantial clinical concern. *Can J Psychiatry*. 2018;63:526–31.
62. Lally J, MacCabe JH. Antipsychotic medication in schizophrenia: a review. *Br Med Bull*. 2015;114:169–79.
63. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychological Bull*. 1979;86:420.
64. McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. *Psychological Methods*. 1996;1:30.
65. Kaar SJ, et al. Antipsychotics: mechanisms underlying clinical response and side-effects and novel treatment approaches based on pathophysiology. *Neuropharmacology*. 2019;77:107704.
66. Rasmussen CE. Gaussian processes in machine learning. In: Summer school on machine learning. pp. 63–71. Springer, Berlin, Heidelberg, 2003.
67. Nelder JA, Wedderburn RW. Generalized linear models. *J R Stat Soc: Ser A (Gen)*. 1972;135:370–84.
68. Suykens JA, Vandewalle J. Least squares support vector machine classifiers. *Neural Process Lett*. 1999;9:293–300.
69. Ho TK. Random decision forests. In: Proceedings of 3rd international conference on document analysis and recognition. Vol. 1, pp. 278–282 IEEE; 1995.
70. Pedregosa F, et al. Scikit-learn: machine learning in Python. *J Mach Learn Res*. 2011;12:2825–30.
71. Matthews DG, et al. GPflow: a Gaussian process library using TensorFlow. *J Mach Learn Res*. 2017;18:1299–304.
72. National Institute for Clinical Excellence. Psychosis and schizophrenia in adults: prevention and management. NICE guidelines. 2014.
73. Jie M, et al. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *J Clin Epidemiol*. 2019;110:12–22.
74. Wang Y, et al. High-dimensional pattern regression using machine learning: from medical images to continuous clinical variables. *Neuroimage*. 2010;50:1519–35.
75. Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. *Eur Heart J*. 2014;35:1925–31.
76. Hemingway H, et al. Prognosis research strategy (PROGRESS) 1: a framework for researching clinical outcomes. *BMJ*. 2013;346:e5595.
77. Cai H-L, et al. Metabolomic analysis of biochemical changes in the plasma and urine of first-episode neuroleptic-naïve schizophrenia patients after treatment with risperidone. *J Proteome Res*. 2012;11:4338–50.
78. Schwarz E, et al. Identification of blood-based molecular signatures for prediction of response and relapse in schizophrenia patients. *Transl Psychiatry*. 2012;2:e82.
79. Fond G, et al. The promise of biological markers for treatment response in first-episode psychosis: a systematic review. *Schizophrenia Bull*. 2015;41:559–73.
80. Mondelli V, et al. Cortisol and inflammatory biomarkers predict poor treatment response in first episode psychosis. *Schizophrenia Bull*. 2015;41:1162–70.
81. Palaniyappan L, et al. Cortical folding defects as markers of poor treatment response in first-episode psychosis. *JAMA Psychiatry*. 2013;70:1031–40.
82. Szeszko PR, et al. Magnetic resonance imaging predictors of treatment response in first-episode schizophrenia. *Schizophrenia Bull*. 2010;38:569–78.
83. Reis Marques T, et al. White matter integrity as a predictor of response to treatment in first episode psychosis. *Brain*. 2013;137:172–82.
84. Hadley JA, et al. Ventral tegmental area/midbrain functional connectivity and response to antipsychotic medication in schizophrenia. *Neuropsychopharmacology*. 2014;39:1020.
85. Mouchlianitis E, et al. Treatment-resistant schizophrenia patients show elevated anterior cingulate cortex glutamate compared to treatment-responsive. *Schizophrenia Bull*. 2015;42:744–52.
86. Sarpal DK, et al. Baseline striatal functional connectivity as a predictor of response to antipsychotic drug treatment. *Am J Psychiatry*. 2015;173:69–77.
87. Egerton A, et al. Adversity in childhood linked to elevated striatal dopamine function in adulthood. *Schizophrenia Res*. 2016;176:171–6.
88. Jauhar S, et al. The relationship between cortical glutamate and striatal dopamine in first-episode psychosis: a cross-sectional multimodal PET and magnetic resonance spectroscopy imaging study. *Lancet Psychiatry*. 2018;5:816–23.
89. McCutcheon RA, Krystal JH, Howes OD. Dopamine and glutamate in schizophrenia: biology, symptoms and treatment. *World Psychiatry*. 2020;19:15–33.
90. Jain KK, Jain KK. The handbook of biomarkers. New York: Springer; 2010.
91. Goksuluk D, et al. easyROC: an interactive web-tool for ROC curve analysis using R language environment. *R J*. 2016;8:213.
92. Cassidy CM, et al. Neuromelanin-sensitive MRI as a noninvasive proxy measure of dopamine function in the human brain. *Proc Natl Acad Sci USA*. 2019;116:5108–17.
93. Dipasquale O, et al. Receptor-enriched analysis of functional connectivity by targets (REACT): a novel, multimodal analytical approach informed by PET to study the pharmacodynamic response of the brain under MDMA. *Neuroimage*. 2019;195:252–60.
94. Li A, et al. A neuroimaging biomarker for striatal dysfunction in schizophrenia. *Nat Med*. 2020;26:558–65.

95. Ruottinen H, et al. Striatal [18 F] fluorodopa utilization after COMT inhibition with entacapone studied with PET in advanced Parkinson's disease. *J Neural Transm—Parkinson's Dis Dement Sect.* 1995;10:91–106.
96. Pillinger T, et al. Comparative effects of 18 antipsychotics on metabolic function in patients with schizophrenia, predictors of metabolic dysregulation, and association with psychopathology: a systematic review and network meta-analysis. *Lancet Psychiatry.* 2020;7:64–77.
97. Howes OD, et al. A prospective study of impairment in glucose control caused by clozapine without changes in insulin resistance. *Am J Psychiatry.* 2004;161:361–3.
98. Mizuno Y, et al. Heterogeneity and efficacy of antipsychotic treatment for schizophrenia with or without treatment resistance: a meta-analysis. *Neuropsychopharmacology.* 2019;45:1–11.
99. McCrone PR, Dhanasiri S, Patel A, Knapp M, Lawton-Smith S. Paying the price: the cost of mental health care in England to 2026. King's Fund; 2008.
100. Trachtenberg M, et al. Peer support in mental health care: is it good value for money? Centre for Mental Health, London, UK, 2013.
101. Psychiatrists R.C.o. Report of the National Audit of Schizophrenia (NAS) 2012. Healthcare Quality Improvement Partnership London; London, 2012.
102. Knapp M, Mangalore R, Simon J. The global costs of schizophrenia. *Schizophrenia Bull.* 2004;30:279–93.
103. Aitchison KJ, Kerwin RW. Cost-effectiveness of clozapine: a UK clinic-based study. *Br J Psychiatry.* 1997;171:125–30.
104. Kirwan P, et al. The impact of switching to clozapine on psychiatric hospital admissions: a mirror-image study. *Ir J Psychol Med.* 2017;1–5.
105. Essock SM, et al. Clozapine's effectiveness for patients in state hospitals: results from a randomized trial. *Psychopharmacol Bull.* 1996;32:683–97.
106. Land R, et al. The impact of clozapine on hospital use: a systematic review and meta-analysis. *Acta Psychiatr Scand.* 2017;135:296–309.
107. Stroup TS, et al. Comparative effectiveness of clozapine and standard antipsychotic treatment in adults with schizophrenia. *Am J Psychiatry.* 2015;173:166–73.
108. Taipale H, Mehtälä J, Tanskanen A, Tiihonen J. Comparative effectiveness of antipsychotic drugs for rehospitalization in schizophrenia—a nationwide study with 20-year follow-up. *Schizophrenia bulletin.* 2018;44:1381–7.
109. Gören JL, et al. The business case for expanded clozapine utilization. *Psychiatr Serv.* 2016;67:1197–205.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020