Nuclear Medicine and Molecular Imaging
https://doi.org/10.1007/s13139-026-01067-8

ORIGINAL ARTICLE

Deep Learning for Quantitative Perfusion ^{99m}Tc-ECD SPECT in Clinically Confirmed Neuropsychiatric Systemic Lupus Erythematosus: Comparison with Established Software

Luca Filippi[1,2] · Maria Silvia De Feo[3] · Shyqyri Samarxhiu[1] · Antonio Cairo[1,4] · Giuseppe De Vincentis[3] · Alessio Farcomeni[5] · Viviana Frantellizzi[3]

Received: 4 June 2026 / Revised: 26 August 2026 / Accepted: 27 August 2026

Abstract

Purpose We compared a novel deep learning (DL)-based quantification software (BTXBrain) with an established platform (Q.Brain) for quantitative perfusion single photon emission computed tomography (SPECT) in patients with clinically confirmed neuropsychiatric systemic lupus erythematosus (NPSLE).

Methods This retrospective exploratory inter-platform comparison study included patients with clinically established NPSLE who underwent brain SPECT with technetium-99m ethyl cysteinate dimer (^{99m}Tc-ECD). SPECT datasets were reprocessed using BTXBrain and Q.Brain. Regional perfusion estimates were compared across predefined cortical territories using paired analyses, variance ratio assessments, and Kendall's tau correlation analysis.

Results A total of 540 paired regional and hemispheric evaluations from 30 NPSLE patients were successfully processed and analyzed. BTXBrain and Q.Brain unveiled systematic, region-specific divergence in perfusion metrics. BTXBrain yielded higher standardized uptake value ratio values than Q.Brain in frontal (mean difference 0.081, $p<0.001$), posterior cingulate (0.090, $p<0.001$), medial temporal (0.333, $p<0.001$), and occipital regions (0.078, $p<0.001$). Conversely, a reversed trend was observed in the right lateral parietal cortex (-0.056, $p=0.009$). Variance ratio analysis confirmed that the DL-based approach introduces a significantly different data dispersion profile ($p<0.05$ across multiple macro-regions), while meaningful cross-platform correlation was restricted to specific territories, such as the left posterior cingulate cortex ($\tau=0.44$, $p=0.002$).

Conclusion BTXBrain and Q.Brain are not directly interchangeable in patients with NPSLE. Consequently, rather than a generic drop-in replacement, the routine clinical adoption of DL-driven SPECT quantification requires software-specific reference thresholds and further calibration before cross-platform use.

Keywords Neuropsychiatric systemic lupus erythematosus · Neurological disorders · ^{99m}Tc-ECD SPECT · Deep learning · Quantitative imaging

✉ Luca Filippi
luca.filippi@uniroma2.it

1 Department of Biomedicine and Prevention, University of Rome Tor Vergata, via Montpellier 1, 00133 Rome, Italy

2 Nuclear Medicine Research Unit, IRCCS San Raffaele Roma, Rome, Italy

3 Department of Radiological Sciences, Oncology and Anatomo Pathology, Sapienza University of Rome, Rome, Italy

4 Medical School, Catholic University "Our Lady of Good Counsel", Tiranë, Albania

5 Department of Enterprise Engineering, University of Rome Tor Vergata, Rome, Italy

Introduction

Cerebral perfusion single photon emission computed tomography (SPECT) using technetium-99m ethyl cysteinate dimer (^{99m}Tc-ECD) remains an established functional imaging technique for evaluating regional cerebral blood flow abnormalities across a broad range of neurological conditions, including neuropsychiatric systemic lupus erythematosus (NPSLE) [1, 2].

Neurological involvement arises from multiple pathogenic mechanisms, including autoantibody-mediated neuronal injury, blood–brain barrier disruption, microthrombosis

Published online: 05 September 2026

and ischemia related to antiphospholipid antibodies, as well as inflammatory vascular damage and cerebral vasculitis, making the assessment of cerebral perfusion by brain SPECT particularly relevant for detecting regional hypoperfusion and functional abnormalities [3].

Despite its recognized clinical value, quantitative interpretation of cerebral perfusion SPECT still largely depends on conventional image-processing approaches and remains susceptible to methodological variability [4, 5].

In routine clinical practice, the assessment of cerebral hypoperfusion on ^{99m}Tc-ECD SPECT frequently relies on visual image interpretation. However, visual analysis is inherently influenced by reader experience and may be particularly challenging when subtle or diffuse perfusion abnormalities are present [6]. To improve objectivity and reproducibility, several quantitative tools have been developed to support regional assessment of cerebral perfusion and assist diagnostic interpretation.

Among currently available approaches, Q.Brain (GE HealthCare, Milwaukee, WI, US) is a dedicated software platform widely implemented for quantitative analysis of cerebral perfusion SPECT studies. The software incorporates three-dimensional stereotactic surface projection (3D-SSP) processing and enables voxel-wise comparison of individual patient datasets against an age-matched normative database [7, 8]. Regional perfusion abnormalities are quantified through standardized z-score analysis, facilitating identification and visualization of areas demonstrating reduced tracer uptake.

More recently, artificial intelligence (AI), particularly deep learning (DL), has emerged as a promising approach for improving image processing and quantitative assessment in medical imaging [9, 10]. BTXBrain (v1.1.2, Brightonix Imaging Inc., Seoul, Korea) is an automated software platform developed for quantitative analysis of cerebral perfusion SPECT studies, as well as for the quantitative evaluation of positron emission tomography (PET) images with different tracers [11, 12]. Using an AI-based image normalization framework, the software automatically transforms individual datasets into a standardized anatomical space and performs regional quantitative analysis through a largely automated workflow, minimizing operator dependency and generating quantitative metrics including regional uptake values (standardized uptake value ratio/SUVR) measurements.

However, despite increasing interest in AI-based quantitative approaches, the reliability and agreement of DL-based tools for cerebral perfusion SPECT analysis remain insufficiently investigated. Therefore, the aim of this study was to compare a novel DL-based quantification software (BTXBrain) with an established quantitative platform (Q.Brain) in a clinically well-characterized cohort of patients with NPSLE.

Materials and Methods

Study Design

This retrospective single-center, exploratory inter-platform comparative study included consecutive patients with clinically confirmed NPSLE who underwent cerebral perfusion SPECT with ^{99m}Tc-ECD at a tertiary-care academic nuclear medicine center in Italy. Brain perfusion SPECT was performed as part of the routine multidisciplinary diagnostic work-up in patients with suspected neuropsychiatric involvement. Imaging was requested following multidisciplinary discussion whenever functional assessment was considered likely to provide clinically relevant information beyond conventional structural imaging.

The diagnosis of NPSLE was established through multidisciplinary clinical evaluation integrating neurological assessment, laboratory investigations, neuroradiological findings, and longitudinal follow-up data [13]. Based on the 1999 American College of Rheumatology nomenclature, NPSLE includes 19 neuropsychiatric syndromes affecting both the central and peripheral nervous systems, with the most frequent manifestations being headache, seizures, cerebrovascular events, cognitive impairment, mood disorders, psychosis, anxiety, and peripheral neuropathies [3, 14].

To improve diagnostic reliability in this heterogeneous condition, only patients with a confirmed diagnosis and a minimum clinical follow-up of three years were included.

Inclusion criteria were: (1) age $\geq$ 18 years; (2) clinically established diagnosis of NPSLE; (3) availability of cerebral perfusion SPECT performed using ^{99m}Tc-ECD; (4) availability of complete clinical and imaging data; (5) minimum clinical follow-up of three years confirming diagnosis and disease course. Conversely, exclusion criteria included: (1) major neurological disorders unrelated to NPSLE potentially affecting cerebral perfusion; (2) significant cerebrovascular disease or structural brain abnormalities; (3) active infectious diseases or severe systemic comorbidities potentially influencing cerebral blood flow; (4) incomplete clinical or imaging datasets; (5) SPECT studies with suboptimal image quality, significant motion artifacts, or technical limitations impairing image interpretation; (6) inability to successfully process image datasets using both quantitative software tools.

All SPECT datasets were retrieved from the centralized archive and retrospectively reprocessed using a novel deep learning (DL)-based software (BTXBrain) and an

established software routinely used in clinical practice (Q.Brain). Quantitative inter-software agreement was dynamically mapped across 18 regional and hemispheric cortical metrics per patient.

The primary study objective was comparison of regional standardized uptake value ratio (SUVR) estimates generated by the two software packages. Secondary objectives included assessment of inter-method agreement.

SPECT Acquisition Protocol

Cerebral perfusion SPECT studies were performed following intravenous administration of 740 MBq of ^{99m}Tc-ECD (Neurolite, Bristol-Myers Squibb Pharma). Prior to radiotracer injection, patients were instructed to avoid excessive sensory stimulation and were positioned in a quiet, dimly lit environment to minimize external influences on cerebral perfusion. After an adequate resting period, the radiotracer was intravenously administered under standardized resting conditions. Image acquisition was performed approximately 30–45 min after tracer injection using a dual-head, large field of view gamma camera (Infinia®, GE Healthcare, Milwaukee, USA) equipped with a couple of high-resolution low-energy parallel holes collimators (LEHR). SPECT acquisition was performed using the smallest possible radius of rotation, with 120 projections over 360°, an angular step of 3°, zoom factor of 1.3 and an acquisition time of 25 s per projection using a 128 × 128 matrix. A 10% energy window was centered on the 140-keV photopeak of ^{99m}Tc. SPECT images were reconstructed using the iterative ordered subset expectation-maximization (OSEM) algorithm (5 iterations, 10 subsets), with a Butterworth filter (cutoff frequency 0.48, power 10) and attenuation correction based on Chang's method (linear attenuation coefficient $\mu = 0.11\ cm^{-1}$). Reconstructed transaxial datasets were reformatted into coronal and sagittal planes.

Quantitative analysis was subsequently performed using both Q.Brain and BTXBrain software packages, allowing regional assessment of cerebral perfusion and comparative evaluation across different analytical platforms.

Image Analysis

Each patient underwent quantitative processing with both Q.Brain and BTXBrain on the same perfusion SPECT data. In the current study, we employed for the analysis the Q.Brain software integrated into the Xeleris 5 Workstation (GE Healthcare). In Q.Brain, images were spatially normalized and quantified against a global cortical reference, yielding regional uptake indices and z-scores across multiple anatomical territories; the report provided bilateral values for prefrontal lateral and medial, sensorimotor, anterior and posterior cingulate, precuneus, superior and inferior parietal, lateral occipital, primary visual, lateral temporal, and medial temporal cortices, with cerebellum and pons reported separately [8].

BTXBrain was then applied using global cortex as the reference region, producing regional SUVRs for the macro-regions frontal, PCC-precuneus, lateral temporal, lateral parietal, medial temporal, occipital, and basal ganglia. Because BTXBrain does not generate z-scores, inter-software comparison was performed exclusively on SUVR values, defined as the ratio of the mean tracer concentration within a target region of interest (ROI) to the mean concentration within the reference region, as follows:

$$SUVR = mean\ SUV\ target\ ROI / mean\ SUV\ reference\ region$$

For both platforms, global cortical uptake was selected as the reference region. To harmonize the two parcellation schemes, Q.Brain subregions that corresponded to a single BTXBrain compartment were averaged prior to comparison,

Table 1 Patient clinical and demographic characteristics

Variable	Overall cohort ($n = 30$)
Age at SPECT, years	46.7 ± 13.9 (range 23–76)
Female sex	26 (86.7%)
Male sex	4 (13.3%)
Disease duration at SPECT, years	5.5 ± 6.0 (range 1–28)
MRI available	12 (40.0%)
MRI negative	5 (16.7%)
MRI positive	7 (23.3%)
Headache	10 (33.3%)
Mood disorders	10 (33.3%)
Cognitive impairment	22 (73.3%)
Sensorimotor syndrome	22 (73.3%)
Other clinical features	11 (36.7%)
Epilepsy	1 (3.3%)

MRI magnetic resonance imaging; *SPECT* single-photon emission computed tomography

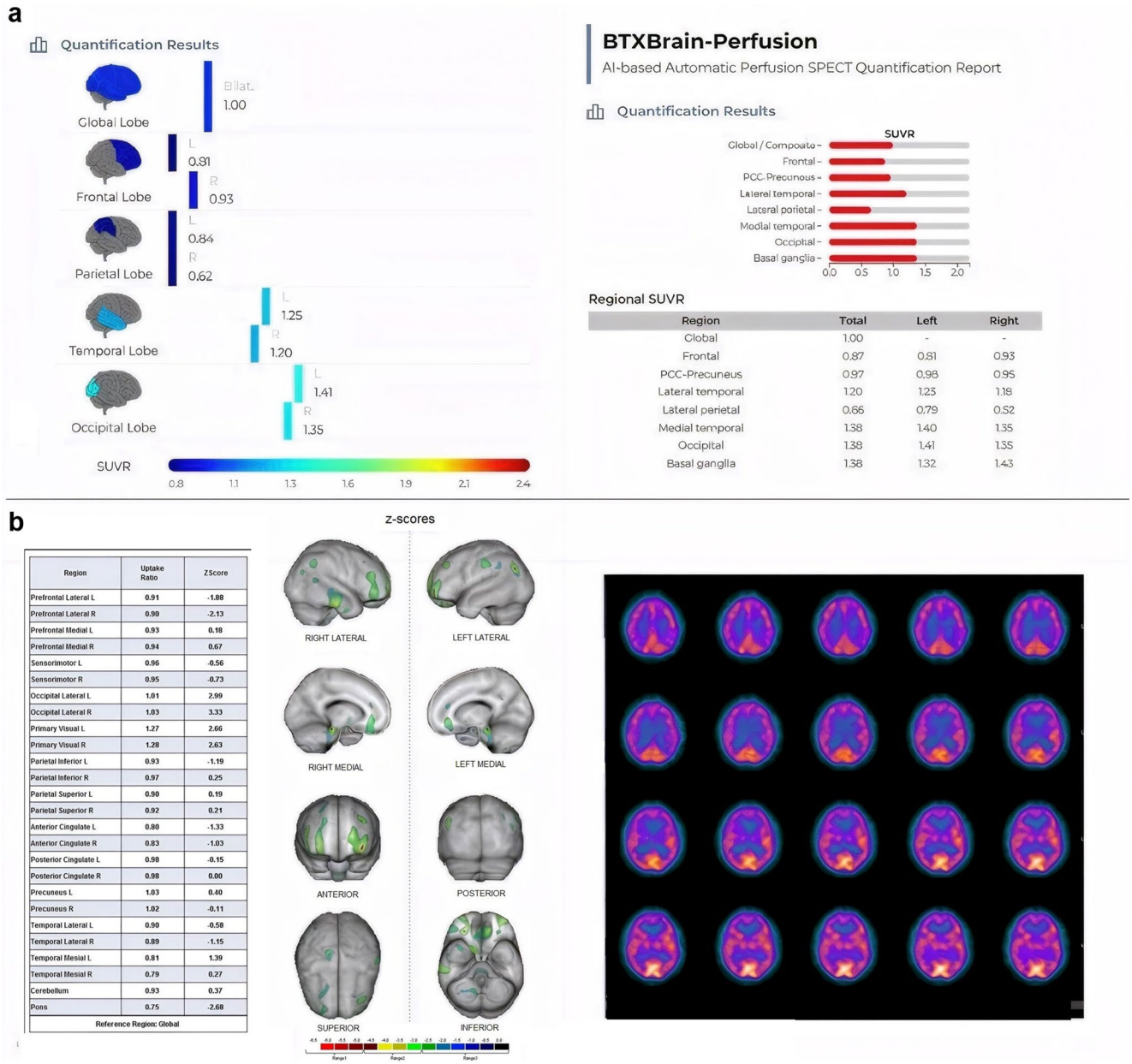

Fig. 1 Representative case of a 52-year-old male with NPSLE, presenting with spatial disorientation, balance impairment, and vertigo. Comparison of brain perfusion SPECT quantification software. (**a**) Automated, deep learning (DL)-based perfusion SPECT quantification report (BTXBrain) displaying global and regional Standardized Uptake Value Ratios (SUVR). (**b**) Standard analysis performed with an established quantification software (Q.Brain), showing regional uptake ratios, Z-score tables, 3D statistical surface projection maps, and transaxial SPECT brain slices

specifically prefrontal lateral and medial for the frontal cortex, posterior cingulate and precuneus for PCC-precuneus, superior and inferior parietal for lateral parietal, lateral occipital and primary visual for occipital cortex. A detailed mapping scheme linking each Q.Brain subregion to its corresponding BTXBrain compartment is provided in Supplementary Data (Table S1).

Statistical Analysis

Continuous variables are expressed as mean ± standard deviation or median (range), as appropriate, and categorical variables as counts and percentages. Statistical analyses were performed using R software, version 4.3.3. Paired regional values were compared using paired t-tests, after

Table 2 Paired comparison of BTXBrain and Q.Brain perfusion estimates Mean difference = BTXBrain − Q.Brain; 95% CI = confidence interval

Region	Mean difference	95% CI lower	95% CI upper	*p*-value
FRONTAL TOTAL	0.0810	0.0643	0.0977	< 0.001
FRONTAL L	0.0685	0.0457	0.0913	< 0.001
FRONTAL R	0.0936	0.0784	0.1088	< 0.001
PCC TOTAL	0.0904	0.0732	0.1077	< 0.001
PCC L	0.0955	0.0785	0.1125	< 0.001
PCC R	0.0995	0.0744	0.1246	< 0.001
LATERAL TEMP TOTAL	0.0342	0.0048	0.0636	0.0243
LATERAL TEMP L	0.0260	-0.0070	0.0590	0.1178
LATERAL TEMP R	0.0430	0.0143	0.0717	0.0046
LATERAL PARIET TOTAL	-0.0260	-0.0526	0.0007	0.0555
LATERAL PARIET L	0.0050	-0.0169	0.0269	0.6440
LATERAL PARIET R	-0.0563	-0.0972	-0.0154	0.0086
MEDIAL TEMP TOTAL	0.3335	0.3039	0.3631	< 0.001
MEDIAL TEMP L	0.3053	0.2770	0.3337	< 0.001
MEDIAL TEMP R	0.3620	0.3244	0.3996	< 0.001
OCC TOTAL	0.0784	0.0523	0.1045	< 0.001
OCC L	0.1158	0.0913	0.1404	< 0.001
OCC R	0.0410	0.0085	0.0735	0.0153

SUVR standardized uptake value ratio; *CI* confidence interval; *PCC* posterior cingulate cortex; *TEMP* temporal; *PARIET* parietal; *OCC* occipital; *L* left; *R* right

assessment of normality of within subjects' differences, using the Kolmogorov-Smirnov test. Dispersion between methods was assessed by variance ratio analysis. Significance levels were computed using Levene's test. Agreements between individual measurements at each segment and overall were evaluated using Kendall's tau correlation coefficients. To further assess inter-method agreement, intraclass correlation coefficients (ICC) with 95% confidence intervals were estimated using a two-way mixed-effects model for single measurements [ICC(A,1)]. Agreement was also evaluated by Bland–Altman analysis. For this analysis, bias was defined as BTXBrain minus Q.Brain, so that positive values indicate higher SUVR estimates obtained with the DL-based software. Mean bias and the corresponding 95% limits of agreement were calculated for each cortical region. Two-sided p values < 0.05 were considered statistically significant. Given the exploratory inter-platform design of the study, unadjusted p-values are reported to describe region-specific measurement behavior across all 18 evaluated regional metrics.

Ethics

The study was approved by the local Ethics Committee. Owing to the retrospective design and the use of anonymized clinical and imaging data, the requirement for written informed consent was waived. All procedures performed in studies involving human participants were in accordance with the Declaration of Helsinki as revised in 2024 and its later amendments.

Results

Demographic and Clinical Characteristics

A total of 540 paired regional and hemispheric evaluations from 30 NPSLE patients who underwent brain SPECT with ^{99m}Tc-ECD were successfully processed and analyzed. The analyzed cohort was predominantly female (26/30, 86.7%) and had a mean age at SPECT of 46.7 ± 13.9 years (median 47.5; range 23–76). The mean disease duration at the time of SPECT was 5.5 ± 6.0 years (median 3.0; range 1–28). MRI data were available in 12 patients (40.0%): 5 had negative findings, whereas 7 showed positive abnormalities, including chronic periventricular vascular gliosis, nonspecific subcortical gliosis, diffuse gliotic changes, and focal lesions compatible with vasculitic sequelae. From a clinical standpoint, the most frequent manifestations were cognitive impairment and sensorimotor syndrome (each 22/30, 73.3%), followed by headache and mood disorders (each 10/30, 33.3%). Additional neurological or systemic features were reported in 11/30 patients (36.7%), while epilepsy was uncommon and observed in one case (3.3%). Patients' characteristics are summarized in Table 1. A representative clinical case showing quantification results from both BTXBrain and Q.Brain is shown in Fig. 1.

Comparison of Regional SUVR Estimates

Paired comparison of cerebral perfusion estimates (SUVR) obtained with BTXBrain and Q.Brain showed region-specific

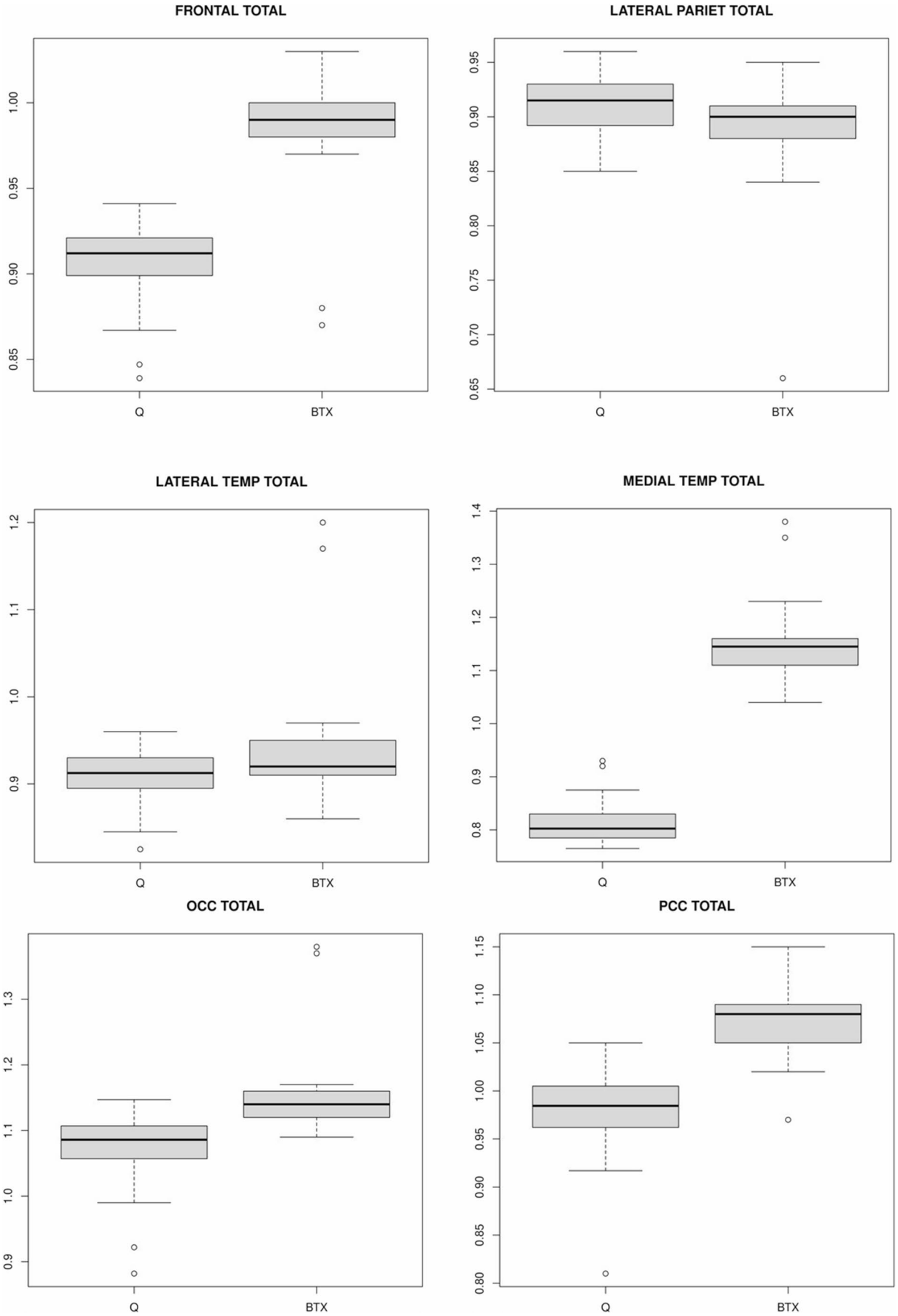
FRONTAL TOTAL
1.00
0.95
0.90
0.85
Q
BTX
LATERAL PARIET TOTAL
0.95
0.90
0.85
0.80
0.75
0.70
0.65
Q
BTX
LATERAL TEMP TOTAL
1.2
1.1
1.0
0.9
Q
BTX
MEDIAL TEMP TOTAL
1.4
1.3
1.2
1.1
1.0
0.9
0.8
Q
BTX
OCC TOTAL
1.3
1.2
1.1
1.0
0.9
Q
BTX
PCC TOTAL
1.15
1.10
1.05
1.00
0.95
0.90
0.85
0.80
Q
BTX

Fig. 2 Box-and-whisker plots comparing total regional SUVR values obtained with Q.Brain (Q) and BTXBrain (BTX) across the analyzed cortical territories. The plots illustrate differences in central tendency and distribution between the two software packages for aggregated cortical regions

differences in patients with NPSLE (Table 2; Figs. 2 and 3). BTXBrain yielded significantly higher values than Q.Brain in the frontal cortex (total, left, and right), posterior cingulate cortex (total, left, and right), medial temporal cortex (total, left, and right), occipital cortex (total, left, and right), and right lateral temporal cortex (all $p < 0.05$). The largest positive mean difference was observed in the medial temporal region, indicating substantially higher BTXBrain estimates in that area. By contrast, the right lateral parietal cortex showed a significant negative mean difference, with BTXBrain values lower than those obtained with Q.Brain. No statistically significant differences were observed for the left lateral temporal cortex, total lateral parietal cortex, or left lateral parietal cortex, indicating insufficient evidence for systematic differences between the two software packages in these regions.

Comparative Analysis of Value Dispersion Between Q.Brain and BTXBrain

Variance ratio analysis was performed to compare the dispersion of perfusion values obtained with the conventional Q.Brain software and the DL-based BTXBrain approach across all investigated brain regions (Table 3). Overall, Q.Brain showed significantly lower variability than BTXBrain in several regions, including the frontal cortex (total and left hemisphere), lateral temporal regions (total, left, and right), medial temporal regions (left and right), and occipital regions (total and right), with variance ratios ranging from 0.177 to 0.839. In contrast, BTXBrain demonstrated significantly lower dispersion only in the right frontal region (variance ratio = 1.96, $p = 0.0025$). No statistically significant differences in dispersion were observed for PCC regions, total and lateral parietal regions, medial temporal total values, or the left occipital cortex; however, these findings do not support equivalence between methods and merely reflect insufficient evidence for a detectable difference.

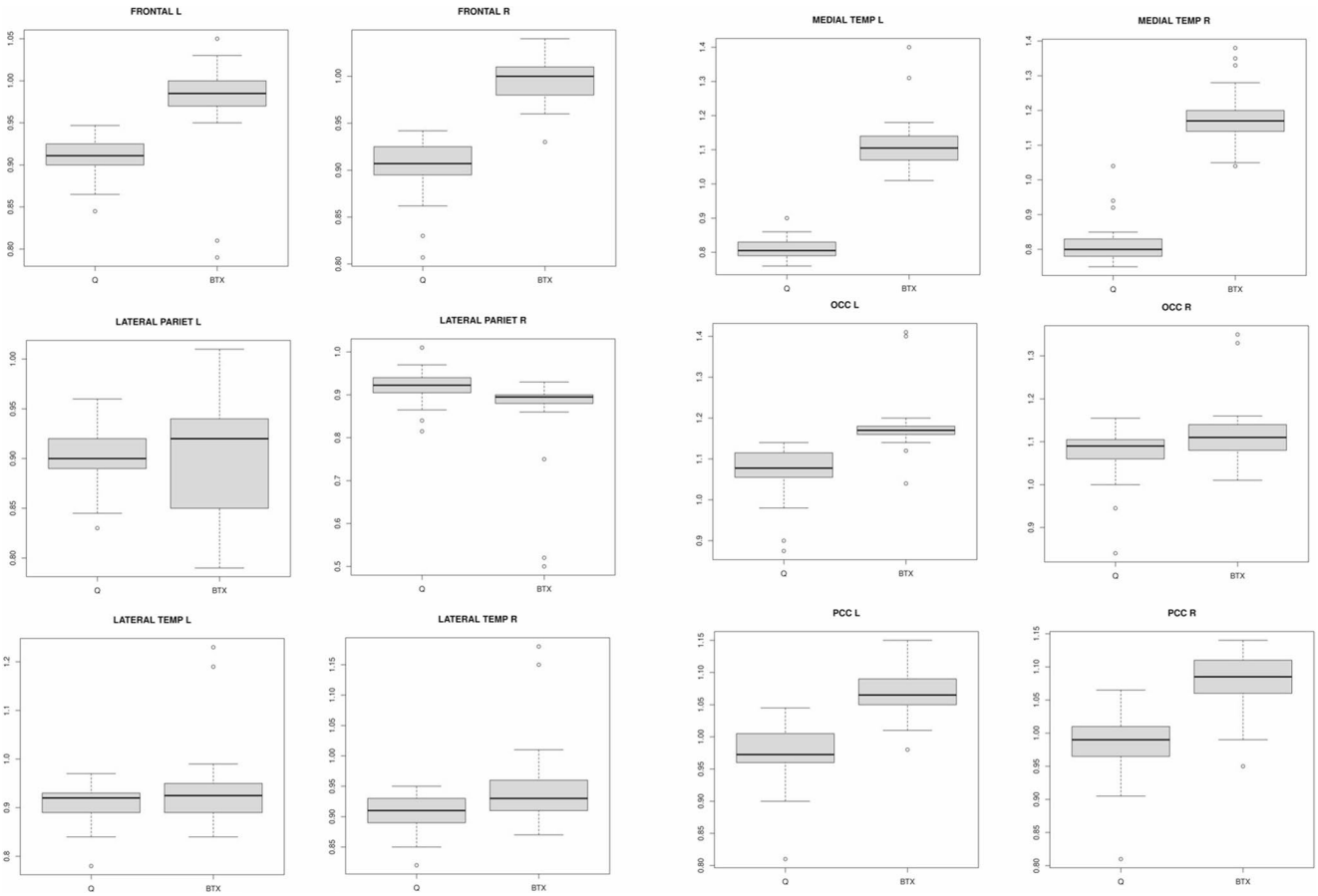

Fig. 3 Hemispheric comparison of regional SUVR estimates generated by Q.Brain (Q) and BTXBrain (BTX) across left and right cortical territories. Box-and-whisker plots display hemisphere-specific distributions and highlight regional asymmetries and systematic differences between the two quantification approaches

Table 3 Variance ratio analysis between Q.Brain and BTXBrain

Brain Area	Variance Ratio (VarQ/VarBTX)	*p*-value
FRONTAL TOTAL	0.5678	0.0436
FRONTAL L	0.1974	0.0424
FRONTAL R	1.9600	0.0025
PCC TOTAL	1.3707	0.1530
PCC L	1.6678	0.0696
PCC R	1.3985	0.0891
LATERAL TEMP TOTAL	0.1770	0.0032
LATERAL TEMP L	0.2021	0.0065
LATERAL TEMP R	0.1831	0.0031
LATERAL PARIET TOTAL	0.1666	0.1921
LATERAL PARIET L	0.2236	0.0604
LATERAL PARIET R	0.1611	0.5131
MEDIAL TEMP TOTAL	0.3116	0.1658
MEDIAL TEMP L	0.1973	0.0003
MEDIAL TEMP R	0.5461	<0.0001
OCC TOTAL	0.7447	0.0113
OCC L	0.8382	0.3346
OCC R	0.6627	0.0065

VarQ variance of Q.Brain values; *VarBTX* variance of BTXBrain values; *PCC* posterior cingulate cortex; *TEMP* temporal; *PARIET* parietal; *OCC* occipital; *L* left; *R* right

Agreement Analysis

Agreement analysis between the conventional Q.Brain software and BTXBrain platform for SUVR quantification revealed a heterogeneous pattern across cortical regions (Table 4). Statistically significant correlations were observed in the left posterior cingulate cortex (PCC; $\tau = 0.44$, $p = 0.002$), the occipital cortex both at the total and left hemispheric level ($\tau = 0.27$ and 0.35, respectively), and in the right frontal cortex, where a weak inverse correlation emerged ($\tau = -0.29$, $p = 0.029$). Most of the remaining cortical areas showed weak or non-significant concordance between the two quantification approaches.

ICC analysis showed a consistent pattern of poor agreement across all investigated territories (Supplementary Data, Table S2). Most cortical regions demonstrated near-zero or

Table 4 Kendall's tau correlation analysis between Q.Brain and BTXBrain PCC, posterior cingulate cortex; TEMP, temporal; PARIET, parietal; OCC, occipital; L, left; R, right; τ, Kendall's tau correlation coefficient

Region	Kendall's τ	*p*-value
FRONTAL TOTAL	-0.27	0.045
FRONTAL L	-0.09	0.461
FRONTAL R	-0.29	0.029
PCC TOTAL	0.21	0.118
PCC L	0.44	0.002
PCC R	0.03	0.881
LATERAL TEMP TOTAL	0.06	0.689
LATERAL TEMP L	0.14	0.321
LATERAL TEMP R	-0.11	0.401
LATERAL PARIET TOTAL	0.14	0.309
LATERAL PARIET L	0.07	0.637
LATERAL PARIET R	-0.1	0.417
MEDIAL TEMP TOTAL	0.18	0.203
MEDIAL TEMP L	0.16	0.254
MEDIAL TEMP R	0.18	0.192
OCC TOTAL	0.27	0.039
OCC L	0.35	0.01
OCC R	0.14	0.318

negligible ICC values with confidence intervals spanning zero. Even in regions where p-values reached statistical significance, the point estimates remained exceptionally low with wide confidence intervals, including the left posterior cingulate cortex (ICC = 0.10, 95% CI: -0.07 to 0.36), the overall occipital cortex (ICC = 0.19, 95% CI: -0.10 to 0.49), and the left occipital cortex (ICC = 0.18, 95% CI: -0.09 to 0.51). These findings further confirm the absence of meaningful quantitative agreement between BTXBrain and Q.Brain.

Bland–Altman Analysis

Bland–Altman analysis demonstrated slight systematic bias for most cortical regions (Figs. 4 and 5; Table 5). Small positive biases were observed in the right frontal cortex and in the posterior cingulate cortex (total and left), indicating slightly higher SUVR values obtained with BTXBrain. The largest systematic difference involved the medial temporal cortex, where BTXBrain consistently yielded higher values than Q.Brain, with an average bias of approximately 0.33 SUVR units across total, left, and right measurements. Despite low average bias in the analyzed areas, the wide limits of agreement across multiple cortical regions reflect substantial individual-level variation, confirming that the two platforms cannot be considered clinically interchangeable. Because Bland–Altman estimates are sensitive to individual observations in relatively small samples, these findings should be interpreted cautiously.

Discussion

To the best of our knowledge, this is the first study investigating BTXBrain in cerebral perfusion SPECT. Although PET is technically and translationally appealing, SPECT remains widely available and clinically established [15], making validation of AI-based quantification on SPECT relevant. This study should therefore be interpreted primarily as a technical and methodological assessment of measurement behavior rather than as a diagnostic accuracy or outcome study. Its central question was whether a novel DL-based quantification software reproduces or meaningfully diverges from an established platform in a real-world cohort with clinically confirmed NPSLE.

NPSLE provides a biologically coherent setting for evaluating cerebral perfusion SPECT, as ^{99m}Tc-ECD is sensitive to regional perfusion abnormalities related to microvascular injury, endothelial dysfunction, and impaired cerebral blood flow [16–18]. This differs from ^{18}F-FDG PET, which primarily reflects glucose metabolism and downstream neuronal dysfunction [19, 20]. Thus, the present study evaluates quantification in a disease context in which regional perfusion abnormalities are pathophysiologically relevant.

A methodological strength was the inclusion of patients with clinically confirmed NPSLE, which minimized diagnostic uncertainty and allowed direct comparison of the two quantification methods within a clinically relevant pathological setting. However, the absence of healthy or disease-control groups precludes assessment of specificity and limits generalizability beyond NPSLE [21, 22].

Our results show that the two platforms are not fully interchangeable. Interestingly, several of the regions showing the greatest quantitative differences between BTXBrain and Q.Brain—including the frontal cortex, posterior cingulate/precuneus, medial temporal structures, and occipital cortex—overlap with cortical territories that have repeatedly been described as vulnerable in NPSLE by both perfusion SPECT and MRI studies [16–18]. Although this study was not designed to assess diagnostic performance or disease localization, the findings indicate that software-dependent variability may affect clinically relevant regions and potentially influence longitudinal interpretation or the application of quantitative thresholds derived from another platform. Software-specific validation is therefore warranted before routine clinical implementation.

BTXBrain produced higher values than Q.Brain in several cortical regions, while the right lateral parietal cortex showed the opposite direction. Agreement was also uneven, with meaningful concordance only in a subset of regions. This pattern argues against treating the DL output as a simple drop-in replacement for the established software. Instead, BTXBrain appears to apply a different normalization/quantification logic that changes both central tendency and dispersion. For clinical use, this matters: a software package can be promising and still require software-specific reference thresholds, region-specific interpretive rules, and further calibration before cross-platform use [23]. The observed heterogeneity also suggests that regional anatomy, cortical thickness, local perfusion gradients, and normalization behavior may affect the two algorithms differently.

The fact that BTXBrain, as other DL-based platforms, is essentially a black box deserves explicit discussion [24]. Unlike older analytical tools, whose processing chain is more familiar and whose behavior has been observed over long clinical use, a DL pipeline cannot be readily "opened up" in the same way to understand exactly why one region shifts upward or another downward [25]. For this reason, validation has to rely on external performance, reproducibility, and agreement rather than on interpretability alone. The proper scientific question is not whether the model is transparent in the classical sense, but whether its outputs are stable, clinically plausible, and sufficiently aligned with accepted methods to justify further use.

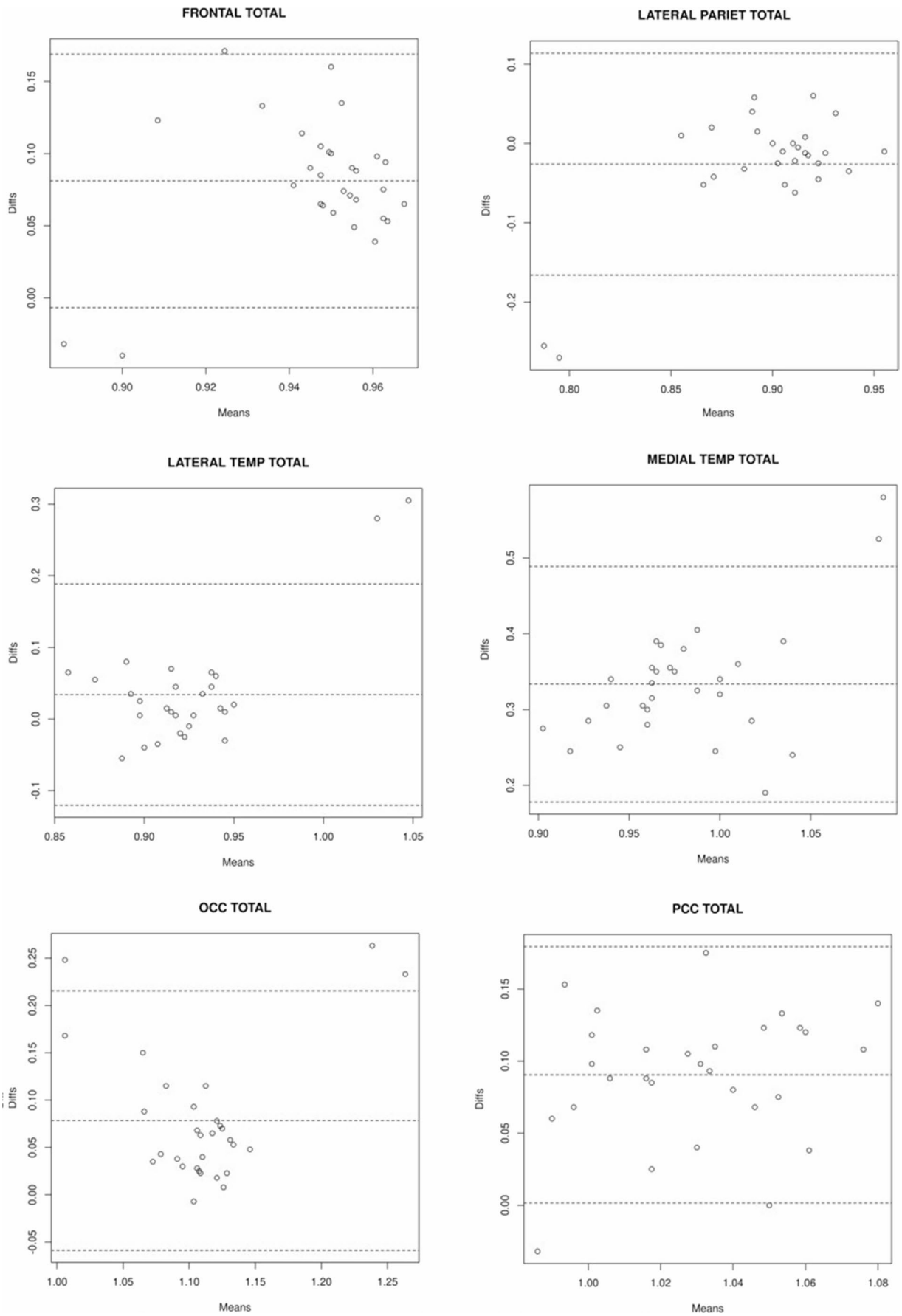
FRONTAL TOTAL
Diffs
Means
LATERAL PARIET TOTAL
Diffs
Means
LATERAL TEMP TOTAL
Diffs
Means
MEDIAL TEMP TOTAL
Diffs
Means
OCC TOTAL
Diffs
Means
PCC TOTAL
Diffs
Means

Fig. 4 Bland–Altman plots comparing total regional SUVR values obtained with Q.Brain and BTXBrain across the analyzed cortical territories. The central dashed line indicates the mean bias, and the outer dashed lines represent the 95% limits of agreement (± 1.96 SD). Differences (y-axis) are calculated as BTXBrain minus Q.Brain

Another point worth highlighting is the asymmetry between methodological promise and clinical readiness [26, 27]. A DL-based platform may reduce operator dependence, shorten processing time, and simplify workflow. However, the present data indicate that workflow simplicity should not be mistaken for equivalence. A new software tool may be operationally attractive yet still need formal calibration before it can be used interchangeably with an established platform. In practical terms, this means that clinicians should be cautious about transferring cutoffs, visual interpretation habits, or severity judgments from one system to another without validation. That is especially true in NPSLE, where subtle regional hypoperfusion may carry diagnostic and follow-up value.

Conclusion

This exploratory study shows that BTXBrain and Q.Brain present systematic and region-specific differences in quantitative perfusion estimates in patients with clinically confirmed NPSLE, indicating that the two platforms should not be used interchangeably. BTXBrain appears to be a promising automated DL-based tool, but its outputs require disease-specific validation, calibration, and interpretive thresholds before routine clinical adoption. Further prospective studies in larger cohorts, ideally including control populations and additional reference standards, are warranted to define its clinical role more precisely.

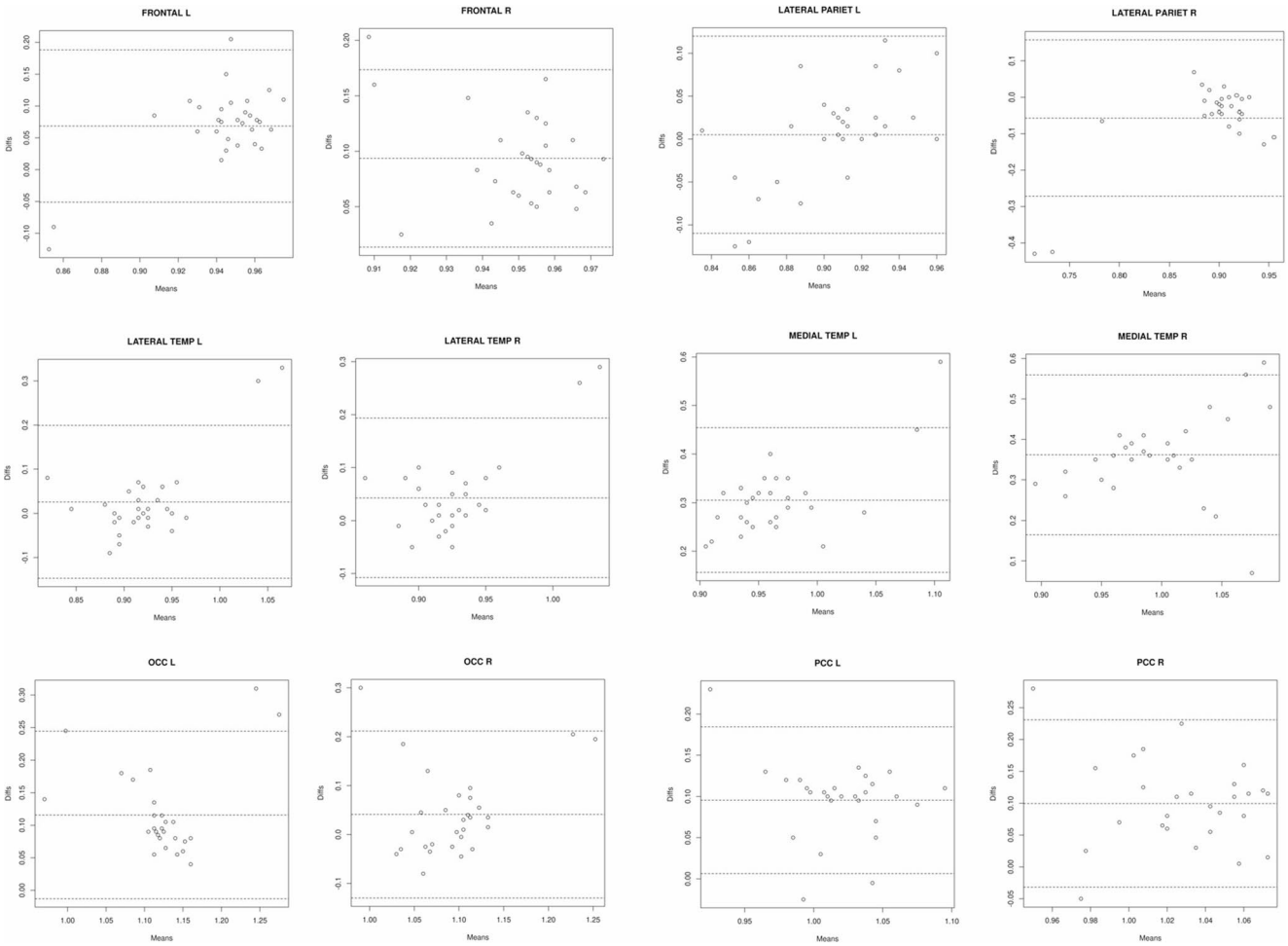

Fig. 5 Bland–Altman plots comparing hemispheric regional SUVR values obtained with Q.Brain and BTXBrain across the analyzed cortical territories, detailed for the left (L) and right (R) hemispheres. The central dashed line indicates the mean bias, and the outer dashed lines represent the 95% limits of agreement (± 1.96 SD). Differences (y-axis) are calculated as BTXBrain minus Q.Brain

Table 5 Bland–Altman analysis between Q.Brain and BTXBrain

Region	Mean bias	Lower 95% LoA	Upper 95% LoA
FRONTAL TOTAL	0.08	-0.01	0.17
FRONTAL L	0.07	-0.05	0.19
FRONTAL R	0.09	0.01	0.17
PCC TOTAL	0.09	0.00	0.18
PCC L	0.10	0.01	0.18
PCC R	0.10	-0.03	0.23
LATERAL TEMP TOTAL	0.03	-0.12	0.19
LATERAL TEMP L	0.03	-0.15	0.20
LATERAL TEMP R	0.04	-0.11	0.19
LATERAL PARIET TOTAL	-0.03	-0.17	0.11
LATERAL PARIET L	0.00	-0.11	0.12
LATERAL PARIET R	-0.06	-0.27	0.16
MEDIAL TEMP TOTAL	0.33	0.18	0.49
MEDIAL TEMP L	0.31	0.16	0.45
MEDIAL TEMP R	0.36	0.16	0.56
OCC TOTAL	0.08	-0.06	0.22
OCC L	0.12	-0.01	0.24
OCC R	0.04	-0.13	0.21

LoA limits of agreement; *PCC* posterior cingulate cortex; *TEMP* temporal; *PARIET* parietal; *OCC* occipital; *L* left; *R* right

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s13139-026-01067-8.

Acknowledgements None

Authors' Contributions Luca Filippi, Maria Silvia De Feo and Viviana Frantellizzi participated in the study design, drafting of the manuscript, and data acquisition and analysis. Shyqyri Samarxhiu and Antonio Cairo participated in data acquisition and data analysis. Alessio Farcomeni participated in statistical analysis and drafting of the manuscript. Giuseppe De Vincentis participated in the study conception, in data analysis, and approval of the final content of the manuscript. All authors read and approved the final manuscript.

Funding Open access funding provided by Università degli Studi di Roma Tor Vergata within the CRUI-CARE Agreement. This work was supported by funding of the Italian Ministry of Health [Ricerca corrente].

Data Availability The datasets generated during and/or analyzed during the current study are not publicly available, but are available from the corresponding author on reasonable request.

Declarations

Clinical trial number Not applicable

Ethics approval and consent to participate The study was approved by the local Ethics Committee (Comitato Etico Territoriale Lazio – Area 5, n. 462/SR/26). Owing to the retrospective design and the use of anonymized clinical and imaging data, the requirement for written informed consent was waived. All procedures performed in studies involving human participants were in accordance with the Declaration of Helsinki as revised in 2024 and its later amendments.

Consent for publication The Institutional Review Board at our institute approved this retrospective study, and the requirement to obtain informed consent was waived

Competing interests Luca Filippi, Maria Silvia De Feo, Shyqyri Samarxhiu, Antonio Cairo, Giuseppe De Vincentis, Alessio Farcomeni and Viviana Frantellizzi declare that they have no competing interests

References

1. Driver CB, Wallace DJ, Lee JC, Forbess CJ, Pourrabbani S, Minoshima S, et al. Clinical validation of the watershed sign as a marker for neuropsychiatric systemic lupus erythematosus. Arthritis Rheum. 2008;59:332–7.
2. Appenzeller S, Amorim BJ, Ramos CD, Rio PA, de Etchebehere C, Camargo ECS. Voxel-based morphometry of brain SPECT can detect the presence of active central nervous system involvement in systemic lupus erythematosus. Rheumatology (Oxford). 2007;46:467–72.
3. Carrión-Barberà I, Salman-Monte TC, Vílchez-Oya F, Monfort J. Neuropsychiatric involvement in systemic lupus erythematosus: A review. Autoimmun Rev. 2021;20:102780.
4. Bailey DL, Willowson KP. An evidence-based review of quantitative SPECT imaging and potential clinical applications. J Nucl Med. 2013;54:83–9.
5. Michopoulou SK, Dickson JC, Gardner GG, Gee TR, Fenwick AJ, Melhuish T, et al. Brain PET and SPECT imaging and

quantification: a survey of the current status in the UK. Nucl Med Commun. 2023;44:834–42.
6. Sugiura A, Shibutani T, Onoguchi M, Nagaki A, Tsuboi K, Ito T, et al. Relationship between physical parameters and visual analysis for assessment of image quality: a multi-center and multi-vendor phantom study in brain SPECT. Asia Ocean J Nucl Med Biol. 2025;13:42–52.
7. Warwick JM. Imaging of brain function using SPECT. Metab Brain Dis. 2004;19:113–23.
8. Seeley M-C, O'Brien H, Wilson G, Coat C, Smith T, Hickson K, et al. Novel brain SPECT imaging unravels abnormal cerebral perfusion in patients with postural orthostatic tachycardia syndrome and cognitive dysfunction. Sci Rep. 2025;15:3487.
9. Carriero S, Cannella R, Cicchetti F, Angileri A, Bruno A, Biondetti P, et al. AI Revolution in Radiology, Radiation Oncology and Nuclear Medicine: Transforming and Innovating the Radiological Sciences. J Med Imaging Radiat Oncol. 2025;69:649–59.
10. Filippi L, Bianconi F, Frantellizzi V, Ferrari C, Marongiu A, De Feo MS, et al. Machine learning for automated differentiation of parkinson's disease and its mimics using 123I-mIBG scintigraphy: insights from a multicentre real-world cohort (ITA-mIBG study). Eur J Nucl Med Mol Imaging. 2026;53:3469–80.
11. Chong A, Ha J-M, Chung JY. Does Q.Clear Processing Change PET Ratios? Quantitative Evidence Using BTXBrain-DAT. Brain Sci. 2025;15:1036.
12. Kang Y-K, Min JW, Kwon SJ, Ha S. Reliability of Automated Amyloid PET Quantification: Real-World Validation of Commercial Tools Against Centiloid Project Method. Tomography. 2025;11:86.
13. Zirkzee EJ, Steup-Beekman GM, van der Mast RC, Bollen EL, van der Wee NJ, Baptist E, et al. Prospective study of clinical phenotypes in neuropsychiatric systemic lupus erythematosus; multidisciplinary approach to diagnosis and therapy. J Rheumatol. 2012;39:2118–26.
14. The American College. of Rheumatology nomenclature and case definitions for neuropsychiatric lupus syndromes. Arthritis Rheum. 1999;42:599–608.
15. Israel O, Pellet O, Biassoni L, De Palma D, Estrada-Lobato E, Gnanasegaran G, et al. Two decades of SPECT/CT - the coming of age of a technology: An updated review of literature evidence. Eur J Nucl Med Mol Imaging. 2019;46:1990–2012.
16. Sahebari M, Rezaieyazdi Z, Khodashahi M, Abbasi B, Ayatollahi F. Brain Single Photon Emission Computed Tomography Scan (SPECT) and functional MRI in Systemic Lupus Erythematosus Patients with Cognitive Dysfunction: A Systematic Review. Asia Ocean J Nucl Med Biol. 2018;6:97–107.
17. Huang WS, Chiu PY, Tsai CH, Kao A, Lee CC. Objective evidence of abnormal regional cerebral blood flow in patients with systemic lupus erythematosus on Tc-99m ECD brain SPECT. Rheumatol Int. 2002;22:178–81.
18. Oku K, Atsumi T, Furukawa S, Horita T, Sakai Y, Jodo S, et al. Cerebral imaging by magnetic resonance imaging and single photon emission computed tomography in systemic lupus erythematosus with central nervous system involvement. Rheumatology (Oxford). 2003;42:773–7.
19. Mackay M, Tang CC, Vo A. Advanced neuroimaging in neuropsychiatric systemic lupus erythematosus. Curr Opin Neurol. 2020;33:353–61.
20. Berndorfler BD, Warwick JM, Doruyter AGG. Role of F-18 FDG PET-CT in neuropsychiatric systemic lupus erythematosus. Compr Psychiatry. 2024;132:152480.
21. Choi H. Deep Learning in Nuclear Medicine and Molecular Imaging: Current Perspectives and Future Directions. Nucl Med Mol Imaging. 2018;52:109–18.
22. Apostolopoulos ID, Papandrianos NI, Feleki A, Moustakidis S, Papageorgiou EI. Deep learning-enhanced nuclear medicine SPECT imaging applied to cardiac studies. EJNMMI Phys. 2023;10:6.
23. Bradshaw TJ, Brosch-Lenz J, Uribe C, Karakatsanis N, Bruce R, Strigari L, et al. Recommendations for Standardizing Nuclear Medicine Terminology and Data in the Era of Theranostics and Artificial Intelligence. J Nucl Med. 2025;66:1471–9.
24. Urso L, Manco L, Filippi L. Synthetic imaging for research and education in nuclear medicine: Who's afraid of the black box? Eur J Nucl Med Mol Imaging. 2025;52:3071–2.
25. Yang G, Ye Q, Xia J. Unbox the black-box for the medical explainable AI via multi-modal and multi-centre data fusion: A mini-review, two showcases and beyond. Inf Fusion. 2022;77:29–52.
26. Currie G, Rohren E. Intelligent Imaging in Nuclear Medicine: the Principles of Artificial Intelligence, Machine Learning and Deep Learning. Semin Nucl Med. 2021;51:102–11.
27. Yang F, Lei B, Zhou Z, Song T-A, Balaji V, Dutta J. AI in SPECT Imaging: Opportunities and Challenges. Semin Nucl Med. 2025;55:294–312.