VOXEL-BASED LESION-SYMPTOM MAPPING: A NONPARAMETRIC BAYESIAN APPROACH

Mame Diarra Fall ⋆, *Émilie Lavau* †, *Pascal Auzou* ‡

⋆ Institut Denis Poisson, Université d'Orléans, Université de Tours, CNRS, Route de Chartres BP 6759 45067 Orléans cedex 2, France.
† Université d'Orléans, Château de la Source, Avenue du Parc Floral, BP 6749 45067 Orléans cedex 2, France.
‡ Centre Hospitalier Régional d'Orléans, Service de Neurologie, 14 avenue de l'hôpital, CS 86709 45067 Orléans cedex 2, France.

ABSTRACT

The study of brain-injured patients (or lesion-based analysis) is a powerful paradigm for investigating structure–function relationships using neuroimaging techniques. Voxel-based Lesion-Symptom Mapping (VLSM) has been widely used to detect structure–function associations in neuroimaging studies. However this approach is based on Student t-test for which normality does not always hold. Our aim in the current study is twofold: 1) to confirm/refute the implication of the classical language areas using the Language Screening (LAST) test; and 2) to determine if it is possible to reduce the number of patients included in the VLSM study, using a different statistical approach. To achieve the second goal, we propose an alternative nonparametric and Bayesian test using Pólya trees. The approach is Bayesian, assigning prior distributions and computing the Bayes factor of H_0 (null hypothesis) to H_1 (alternative) ; and it is nonparametric since the priors are put on the unknown distribution functions under H_0 and H_1. Our results highlight that the Pólya tree prior provides a convenient and effective way for testing two sample differences in VLSM studies.

Index Terms— Hypothesis testing, Nonparametric Bayes, Pólya tree, VLSM, stroke, LAST test.

1. INTRODUCTION

The two sample comparison is a fundamental problem in statistics. The t-test is probably the most popular parametric statistical test [1], [2]. Popular nonparametric and frequentist approaches include Wilcoxon [3], [4], Kolmogorov-Smirnov [5], [6] and Cramer-von-Mises [7] tests. In this work, we consider the test problem from a Bayesian nonparametric (BNP) perspective. BNP is a relatively new and fast developing discipline, with a great number of applications [8]. The BNP approach has recently demonstrated its suitability in image reconstruction [9], [10], [11]. However, while there has been considerable interest in BNP estimation, the BNP hypothesis testing has received little attention. It is only recently that the hypothesis testing has been investigated from a nonparametric Bayes perspective using Pólya trees priors in the two sample case [12], [13], [14], and for more than two samples in [15]. There are a number of attractive properties in formulating the hypothesis testing in a BNP framework using Pólya trees as priors. Bayesian answers have a clear interpretability, as compared with the commonly used p-value. In the Bayesian framework, one simply finds and reports the probability that a particular hypothesis is true given the observations. The competing hypotheses are assigned probabilities and the one with the highest probability is selected. On the contrary, frequentist tests do not assign probabilities to hypotheses directly but rather to the statistic on which the test is based. Furthermore, since it is not based on asymptotics, the Bayesian approach can handle small sample sizes. The Pólya tree (PT) nonparametric model allows one to center the probability distribution at a given parametric model, thus embedding the parametric model in a larger encompassing nonparametric one. The PT has also a rich structure with numerous free parameters remaining after specification of the centering distribution. Last but not least, with the approach taken here, integrations can be handled and calculations of Bayes factors (BF) can be made explicitly. This is an important point since computation of BFs is typically considerably more challenging in Bayesian tests.

The rest of the paper is organised as follows. Section 2 reviews the VLSM approach and the definition and basic properties of the Pólya tree. The proposed BNP based Pólya trees test is described in Section 3. The clinical data, their treatment and the competing statistical tests are presented in Section 4. The numerical results are presented in Section 5. We conclude in Section 6 with a brief discussion and some pointers for future research.

2. THEORETICAL FRAMEWORK

2.1. VLSM approach

The Voxel-based Lesion-Symptom Mapping (VLSM) approach involves mapping the relationship between tissue damage and behavioural performance, on a voxel-by-voxel basis. Let X denote a score representing task performance. In this paper, X represents the outcome of the Language

Screening Test (LAST) [16]. Lesion-symptom mapping aims at detecting differences in the distributions of lesions across groups. To this aim, lesions are first delineated (manually or automatically) for each subject; this produces a binary lesion map. The lesions maps of all subjects are registered to a common stereotaxic space. Second, in voxel-based lesion-symptom mapping, statistical hypothesis testing is performed for each voxel. Let $\mathbf{x} = (x_1, \ldots, x_n)$, denote the LAST scores of n patients and V the number of voxels. For each voxel $v \in \{1, \ldots, V\}$ and each patient i $(i = 1, \ldots, n)$, let z_i^v be a binary variable such that

$$z_i^v = \begin{cases} 1 \text{ if voxel } j \text{ in patient } i \text{ is lesioned,} \\ 0 \text{ otherwise.} \end{cases} \tag{1}$$

In the LAST test, there are two subscores, namely an expression score (with 8 points being the maximum) and a receptive one (maximum of 7 points). Only the total sum score (maximum of 15) is considered here. If the score is less than 15, the patient is ranked as aphasic. Let us now define

$$x_i \begin{cases} \in \{0, \ldots, 14\} \text{ if patient } i \text{ is aphasic,} \\ = 15 \text{ if patient } i \text{ is not aphasic.} \end{cases} \tag{2}$$

In each voxel v, patients can be partitioned into two groups, group 1 and 2, with 1 consisting of those who have voxel v lesioned. We can then define two sets of scores according to group membership. Let

$$\mathbf{x}_1^v = \{x_i, i : z_i^v = 1\} \quad \text{and} \quad \mathbf{x}_2^v = \{x_i, i : z_i^v = 0\},$$

and define

$$n_1^v = \textstyle\sum_{i=1}^n \mathbf{1}(z_i^v = 1) \quad \text{and} \quad n_2^v = \sum_{i=1}^n \mathbf{1}(z_i^v = 0),$$

where $\mathbf{x}_1^v$, with cardinality n_1^v, contains the scores of patients whose v-th voxel is lesioned; $\mathbf{x}_2^v$,with cardinality n_2^v, stands for $\mathbf{x}_1^v$ relative complement. In lesion-symptom mapping, one is interested in whether groups 1 and 2 are significantly different. That is given the sets of samples $\mathbf{x}_1^v \overset{iid}{\sim} F_1$ and $\mathbf{x}_2^v \overset{iid}{\sim} F_2$, with F_1 and F_2 unknown[1], we consider in each voxel v the two competing hypotheses

$$H_0 : F_1 = F_2 \quad \textit{versus} \quad H_1 : F_1 \neq F_2.$$

The significant voxels are considered as being involved in language disorder. This is a two sample comparison problem, usually performed via t-tests as in [17]. In this paper, we propose to handle this via a fully BNP framework with Pólya tree priors. We now recall the basics of Pólya trees.

[1]The test is voxel-based and performed voxel-by-voxel. To simplify the notation, we do not include the voxel index in F_1 and F_2.

2.2. Pólya trees

The Pólya tree is a prior on a random probability distribution F on some domain $\mathcal{X}$. The famous Dirichlet process (DP) introduced by [18] is an example of a PT, although unlike the DP the Pólya tree can select continuous distributions with positive probability and, if necessary, even with probability one. Reference papers on Pólya trees are [19], [20], and [21], [22], whose notations we follow. Let $E = \{0, 1\}$, E^m the m-fold Cartesian product $E \times \cdots \times E$ with $E^0 = \emptyset$. Set $E^\star = \cup_{m=0}^\infty E^m$ and let $\pi_m = \{B_\epsilon : \epsilon \in E^m\}$ a partition of $\mathcal{X}$ such that $\pi_0 = \{\mathcal{X}\}$ and $\Pi = \cup_{m=0}^\infty \pi_m$. Π denotes a collection of subsets $\{B_0, B_1, B_{00}, B_{01}, \ldots\}$. Let $\mathcal{A} = \{\alpha_\epsilon : \epsilon \in E^\star\} = \{\alpha_0, \alpha_1, \alpha_{00}, \alpha_{01}, \ldots\}$ denote a set of positive numbers and $\mathcal{W} = \{W_\epsilon : \epsilon \in E^\star\}$ a collection of random variables.

Definition. *A random probability measure F on $\mathcal{X}$ is said to have a Pólya tree distribution, or a Pólya tree prior, with parameters $(\Pi, \mathcal{A})$ and denoted $F \sim PT(\Pi, \mathcal{A})$, if for every $m = 1, 2, \ldots,$ and every $\epsilon = \epsilon_1, \ldots, \epsilon_m \in E^m$*

$$F(B_{\epsilon_1 \cdots \epsilon_m}) = \prod_{j=1}^m W_{\epsilon_1 \cdots \epsilon_j}, \tag{3}$$

where the conditional probabilities $W_{\epsilon_1 \cdots \epsilon_{j-1} 0}$ are mutually independent Beta random variables,

$$W_{\epsilon_1 \cdots \epsilon_{j-1} 0} \sim \textit{Beta}(\alpha_{\epsilon_1 \ldots \epsilon_{j-1} 0}, \alpha_{\epsilon_1 \ldots \epsilon_{j-1} 1}),$$

with $W_{\epsilon_1 \cdots \epsilon_{j-1} 1} = 1 - W_{\epsilon_1 \cdots \epsilon_{j-1} 0}$, and the terms for $j = 1$ are W_0 and $1 - W_0$.

The parameters α_ϵ determine the smoothness of the realizations of F. Most of the time, one does not assign a different α_ϵ for each ϵ but instead takes α_ϵ to be constant in a level m: $\alpha_\epsilon = a_m$, $\forall m$. Smoothness of generated measures is controlled by the rate of increase of the a_m as one descends the PT. Taking $\alpha_\epsilon = cm^2$, $c > 0$ is a sufficient condition to guarantee that the PT assigns probability one to the set of distributions which have densities. We shall consider such Pólya trees. In practice, it can be difficult to elicit the partitions Π. As for Dirichlet processes, the Pólya tree can be centered at a chosen distribution F_0 on $\mathcal{X}$, so that $\mathbb{E}[F] = F_0$. The simplest way to do this is to take at each level m, the partitioning subsets to coincide with the quantiles F_0^{-1}. For example, if $\mathcal{X} = \mathbb{R}$, partition elements are at each level m,

$$[F_0^{-1}(j/2^m),\ F_0^{-1}((j+1)/2^m)),\ \text{for } j = 0, 1, \ldots, 2^m - 1,$$

where $F_0^{-1}(0) = -\infty$ and $F_0^{-1}(1) = +\infty$. Again, the parameters α_ϵ control how closely the distribution of F is concentrated around its prior mean F_0. Choosing $\alpha_\epsilon = c \times m^2$, the positive parameter c controls how much weight is placed on the centering distribution F_0, playing a similar role to the precision parameter of the Dirichlet process prior. Hanson and Johnson [23] proved two results that indicate broadly the effect c has on inference. As $c \to \infty$, the parametric model F_0 is obtained regardless of n, whereas small values of c allow $\mathbb{E}[F]$ to closely follow the empirical distribution function.

Pólya trees enjoy the following conjugacy property. That is, given a PT prior $F \sim PT(\Pi, \mathcal{A})$ and a sample $\mathbf{x} = (x_1, \ldots, x_n)|F \overset{iid}{\sim} F$, the posterior distribution $F|\mathbf{x}$ is still a PT with updated parameters: $F|X \sim \text{PT}(\mathcal{A}^*, \Pi)$, where

$$\mathcal{A}^* = \{\alpha^*_\epsilon = \alpha_\epsilon + n_\epsilon : \epsilon \in E^*\},$$

and $n_\epsilon = \sum_{i=1}^{n} \mathbf{1}(x_i \in B_\epsilon)$ denotes the number of observations in $\mathbf{x}$ that lie in the partition B_ϵ. This conjugacy allows for an exact closed-form expression of the marginal likelihood. Therefore, problems traditionally encountered when computing Bayes factors are avoided and computation can be made efficient.

3. METHODOLOGY

In this section, we consider the case in which two independent and identically distributed (i.i.d.) samples are observed and we are interested in testing the potential difference between the two underlying distributions.

3.1. Formulation of the two sample test

Suppose we have samples of sizes n_1 and n_2 drawn from distributions F_1 and F_2, respectively. Let $\mathbf{x} = (x_1, \ldots, x_{n_1})$ and $\mathbf{y} = (y_1, \ldots, y_{n_2})$ be the two samples. We wish to test $H_0 : F_1 = F_2$ versus the alternative $H_1 : F_1 \neq F_2$. Let $\mathbf{z}$ denote the combined sample $\mathbf{z} = (z_1, \ldots, z_{n_1+n_2})$. Under the null hypothesis, both samples come from a common distribution: $F_1 = F_2 \equiv F_0$, with F_0 unknown. Under the alternative, $\mathbf{x} \sim F_1$, $\mathbf{y} \sim F_2, F_1 \neq F_2$, where F_1 and F_2 are also unknown. We assume independent PT priors: $F_0, F_1, F_2 \overset{iid}{\sim} PT(\Pi, \mathcal{A})$. To choose between H_0 and H_1, a Bayesian test evaluates the Bayes factor, which is the subject of the next subsection.

3.2. Derivation of the Bayes factor

The test of H_0 versus H_1 is performed by computing the Bayes factor (BF) of H_0 to H_1:

$$BF_{01} = \frac{\text{posterior odds}}{\text{prior odds}} = \frac{P(H_0|\mathbf{z})/P(H_1|\mathbf{z})}{P(H_0)/P(H_1)}, \quad (4)$$

where $P(H_i)$ is the prior probability of H_i, $0 < P(H_i) < 1$, $P(H_0) + P(H_1) = 1$. We assign equal priors to H_0 and H_1, thus the BF is given by the ratio of the posterior odds. Since

$$P(H_i|\mathbf{z}) = \frac{P(H_i)P(\mathbf{z}|H_i)}{\sum_{i=0}^{1} P(H_i)P(\mathbf{z}|H_i)}, \quad i = 0, 1,$$

the BF is reduced to

$$BF_{01} = \frac{P(H_0|\mathbf{z})}{P(H_1|\mathbf{z})} = \frac{P(\mathbf{z}|H_0)}{P(\mathbf{z}|H_1)},$$

and equals the ratio of the marginal likelihoods of $\mathbf{z}$ under H_0 to that under H_1 :

$$BF_{01} = \frac{\int \prod_{i=1}^{n_1+n_2} dF_0(z_i) dPT(F)}{\int \prod_{i=1}^{n_1} dF_1(x_i) \prod_{i=1}^{n_2} dF_2(y_i) dPT(F)}. \quad (5)$$

Authors in [24], [23] and many others follow Lavine [21] and use a Markov expansion of the marginal density. Holmes *et al.* [12] noted that the marginal density is the product of independent Binomials-Beta trials and introduced an efficient way to compute BF_{01} given by the following:

$$\prod_j \frac{\Gamma(\alpha_{j0} + n^{(1)}_{j0} + n^{(2)}_{j0})\Gamma(\alpha_{j1} + n^{(1)}_{j1} + n^{(2)}_{j1})\Gamma(\alpha_{j0})\Gamma(\alpha_{j1})}{\Gamma(\alpha_{j0} + n^{(1)}_{j0} + n^{(2)}_{j0} + \alpha_{j1} + n^{(1)}_{j1} + n^{(2)}_{j1})\Gamma(\alpha_{j0} + \alpha_{j1})}$$
$$\times \frac{\Gamma(\alpha_{j0} + n^{(1)}_{j0} + \alpha_{j1} + n^{(1)}_{j1})\Gamma(\alpha_{j0} + n^{(2)}_{j0} + \alpha_{j1} + n^{(2)}_{j1})}{\Gamma(\alpha_{j0} + n^{(1)}_{j0})\Gamma(\alpha_{j1} + n^{(1)}_{j1})\Gamma(\alpha_{j0} + n^{(2)}_{j0})\Gamma(\alpha_{j1} + n^{(2)}_{j1})}.$$

We next provide results demonstrating the suitability of this approach to VLSM studies.

4. EXPERIMENTAL SETUP

4.1. Data and image processing

We analyzed data collected from 58 participants (47 men, 11 women) who had suffered a single left-hemispheric stroke in the acute phase (< 7 days). All patients, regardless of the arterial distribution of their stroke, were included. The average age is 66.1 years (S.D.=13.4, range=19-91). All participants were evaluated with the LAST test, with a total score between 0 and 15. Patient's lesions were imaged with 3D T1 scans, diffusion-weighted imaging (DWI) and Flair images within the first week following the stroke. The lesions were delineated directly on each patient's DWI or Flair digital MRI images (choosing the best contrasted image) using OSIRIX software. DWI or FLAIR images were yoked to the T1 images so that the extent of the lesion could be verified on these image sequences. Then, we obtain a mask for each patient. MRI images were registered into MNI space (standard template of the Montreal Neurological Institute) using the standard nonlinear spatial normalization procedure from Statistical Parametric Mapping (SPM12) (Wellcome Trust Centre for Neuroimaging) running under Matlab 2017a. We re-aligned and co-registered 3D images with a 5-th Degree B-Spline interpolation method in SPM12 and then averaged them. Masks were re-sliced and normalized to the native space of the averaged 3D images with trilinear interpolation by voxels of 1mm^3.

4.2. Statistical analysis

We now examine the performance of the proposed BNP-PT testing method compared to some other frequentist tests: the t-test (used in the original VLSM of [17]), and two common frequentist nonparametric procedures, namely the Kolmogorov-Smirnov and the Mann-Whitney tests. The total number of analysed voxels is $V = 7,109,137$.

The original VLSM technique has been implemented using the nonparametric mapping (NPM) software, distributed as part of the MRIcron toolset. In VLSM, t-tests are run at

each voxel. It is usual to confine tests to voxels in which there are at least five patients with a lesion and five patients without a lesion; this is the approach we have taken in the current study for all frequentist tests. The significance level is 5%.

Regarding the proposed BNP test, the procedure described in Section 3 has been applied to each voxel, with $\mathbf{x} = \mathbf{x}_1^v$, $\mathbf{y} = \mathbf{x}_2^v$, $n_1 = n_1^v$ and $n_2 = n_2^v$. The prior parameters values of the PT are set as follows: $\mathcal{A} = \{\alpha_{\epsilon_1 \ldots \epsilon_m} = c \times m^2\}$ with $c = 0.1$. As already mentioned, the parameter c does have an impact on inference. In order to investigate the sensitivity of our results to the choice of c, we have looked at the results on simulated data (not shown here) for different values of c. We have found that small values of c perform well in practice. The joint data have been first standardized and the partitions Π set as quantiles of a standard normal density. The centering of the PT to a standard Gaussian distribution was not critical. Other experiments (not reported here) with a partition centered on a uniform distribution showed little difference compared to a partition centered on a standard Gaussian. We reject H_0 if $P(H_0|\mathbf{x}_1^v, \mathbf{x}_2^v) < 0.5$.

5. NUMERICAL RESULTS

We first investigate the results obtained using the total number of patients, $n = 58$. Results not shown here, because of space limitations, showed that all competing methods located more or less the classical language zones known as Broca's and Wernicke's area. The former is involved mostly in the production of speech, while receptive speech has traditionally been associated with the latter. Second, since one of our goals was to reduce the number of patients involved in VLSM studies, we tried smaller sample sizes for which the uncertainty would be greater. The most remarkable results are the ones for $n = 34$, shown in Fig. 1 for the four competitors. Panel a) depicts the outcome of the classical VLSM (t-test); b) the Mann-Whitney; c) the Kolmogorov-Smirnov; and d) the proposed BNP-PT. The colour scale, from black to white, indicates increasing levels of p-values for frequentist tests, or levels of $P(H_0|\mathbf{x}_1^v, \mathbf{x}_2^v)$ for the BNP-PT one. Columns illustrate different views of the brain. Language areas are located in the left hemisphere of the brain (at the right hand side in images of Fig. 1), from the anterior part (Broca) to the posterior one (Wernicke). Table 1 shows the number of significant voxels according to the number n of patients. As one may notice, only the proposed BNP-PT test produces stable regions when we reduce the sample size. The frequentist tests fail in recovering the total Wernicke's area (t-test and Mann-Whitney), and both areas (Kolmogorov). Furthermore, it is worth mentioning that since we perform hypothesis testing for each voxel, we are faced to the multiple-comparison problem for frequentist tests. In order to preserve the overall type 1 error rate, the p-value of each individual test must be adjusted. We have then performed FDR correction [25], but no voxel was found significant for the three frequentist tests.

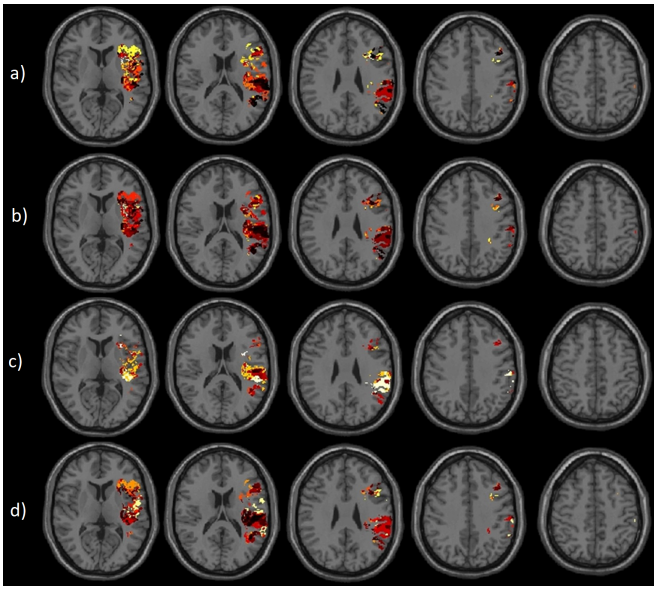

Fig. 1. Panel of images corresponding to: a) t-test, b) Mann-Whitney, c) Kolmogorov and d) proposed BNP-PT ($n = 34$).

n	58	34
t-test	49805	43661
Mann-Whitney	64085	49754
Kolmorov-Smirnov	57032	28308
BNP-PT	45027	44523

Table 1. Number of significant voxels according to n.

This result can be explained by the small effect size and the fact that the number of subjects with lesions at a given voxel was small. The proposed BNP-PT approach obviates multiple-comparison procedures.

6. CONCLUDING REMARKS

Voxel-based lesion mapping (VLSM) is useful to determine the relationship between behavioral measures and the location of brain lesions. However one of the main limitations of classical VLSM studies is that they require a great number of patients. In this work, we have proposed and implemented an alternative Pólya tree-based test that can be applied irrespective of the sample size. The reason why we use Pólya tree priors instead of mixtures of Pólya trees [23] is that we obtain an explicit expression of the Bayes factor. This alleviates the potentially difficult task of computing Bayes factors as in mixtures of Pólya trees. Given the number of tests to perform (several millions), computations should be made fast.

In this paper, all the competing methods analyse a voxel independently of its neighbors. However, a more realistic spatial model for imaging data should include spatial interactions among neighboring voxels. Current work is undergoing in this direction.

7. REFERENCES

[1] R. A. Fisher, "The use of multiple measurements in taxonomic problems," *Annals of Eugenics*, vol. 7, pp. 179–188, 1936.

[2] B.L. Welch, "The Generalization of "Student's" Problem when Several Different Population Variances are Involved," *Biometrika*, vol. 34, pp. 28–35, 1947.

[3] F. Wilcoxon, "Individual Comparisons by Ranking Methods," *Biometrics Bulletin*, vol. 1, no. 6, pp. 80–83, 1945.

[4] H. Mann and D. Whitney, "On a test of whether one of two random variables is stochastically larger than the other," *Annals of Mathematical Statistics*, vol. 18, pp. 50–60, 1947.

[5] A. Kolmogorov, "Sulla determinazione empirica di una legge di distribuzione," *Giornale dell' Istituto Italiano degli Attuari*, vol. 4, pp. 83–91, 1933.

[6] N. Smirnov, "Table for estimating the goodness of fit of empirical distributions," *Ann. Math. Statist.*, vol. 19, pp. 279–281, 1948.

[7] T. W. Anderson, "On the distribution of the two-sample Cramer-von Mises criterion," *Ann. Math. Statist.*, vol. 33, no. 3, pp. 1148–1159, 1962.

[8] P. Müller and R. Mitra, "Bayesian Nonparametric Inference–Why and How," *Bayesian Analysis*, vol. 8, no. 2, pp. 269–302, 2013.

[9] M. D. Fall, É. Barat, C. Comtat, T. Dautremer, T. Montagu, and A. Mohammad-Djafari, "A discrete-continuous Bayesian model for Emission Tomography," in *IEEE International Conference on Image Processing (ICIP)*, 2011, pp. 1373–1376.

[10] M. D. Fall, É. Barat, C. Comtat, T. Dautremer, T. Montagu, and S. Stute, "Continuous space-time reconstruction in 4D PET," in *IEEE Nuclear Science Symposium Conference Record*, 2011, pp. 2581–2586.

[11] M. D. Fall, É. Barat, C. Comtat, T. Dautremer, T. Montagu, and S. Stute, "Dynamic and clinical PET data reconstruction: A nonparametric Bayesian approach," in *IEEE International Conference on Image Processing*, 2013, pp. 345–349.

[12] C.C Holmes, F Caron, J.E Griffin, and D.A Stephens, "Two-sample Bayesian Nonparametric hypothesis testing," *Bayesian Analysis*, pp. 297–320, 2015.

[13] L. Ma and W. H. Wong, "Coupling optional Pólya trees and the two sample problem," *Journal of the American Statistical Association*, vol. 106, no. 496, pp. 1553–1565, 2011.

[14] L. Huang and M. Ghosh, "Two-sample hypothesis testing under Lehmann alternatives and Polya tree priors," *Statistica Sinica*, vol. 24, 10 2014.

[15] Y. Chen and T. E. Hanson, "Bayesian nonparametric k-sample tests for censored and uncensored data," *Computational Statistics & Data Analysis*, vol. 71, pp. 335–346, 03 2014.

[16] C. Flamand-Roze, B. Felissard, E. Roze, L. Maintigneux, J. Beziz, A. Chacon, C. Join-Lambert, D. Adams, and C. Denier, "Validation of a new language screening tool for patients with acute stroke : the Language Screening Test (LAST)," *Stroke*, pp. 1224–1229, 2011.

[17] E. Bates, S. M Wilson, A. Saygin, F. Dick, M. Sereno, R. T Knight, and N. Dronkers, "Voxel-based lesion-symptom mapping," *Nature Neuroscience*, vol. 6, no. 5, pp. 448–450, 2003.

[18] T. S. Ferguson, "A Bayesian analysis of some nonparametric problems," *Ann. Statist.*, vol. 1, pp. 209–230, 1973.

[19] T. S. Ferguson, "Prior distributions on spaces of probability measures.," *Ann. Statist.*, vol. 2, pp. 615–629, 1974.

[20] R. D. Mauldin, W. D. Sudderth, and S. C. Williams, "Pólya trees and random distributions," *Ann. Statist.*, vol. 20, pp. 1203–1221, 1992.

[21] M. Lavine, "Some aspects of Pólya tree distributions for statistical modelling," *Ann. Statist.*, vol. 20, pp. 1222–1235, 1992.

[22] M. Lavine, "More aspects of Pólya tree distributions for statistical modelling," *The Annals of Statistics*, vol. 22, no. 3, pp. 1161–1176, 1994.

[23] T. Hanson and W. O. Johnson, "Modeling regression error with a mixture of Pólya trees.," *J. Am. Stat. Assoc.*, vol. 97, pp. 1020–1033, 2002.

[24] J.O. Berger and A. Guglielmi, "Bayesian and conditional frequentist testing of a parametric model versus nonparametric alternatives," *Journal of the American Statistical Association*, vol. 96, pp. 174–184, 2001.

[25] Y. Benjamini and Y. Hochberg, "Controlling the false discovery rate: a practical and powerful approach to multiple testing," *Journal of the royal statistical society. Series B (Methodological)*, pp. 289–300, 1995.