arXiv: arXiv:0000.0000

SUPPLEMENTARY MATERIAL TO 'TWO-SAMPLE TESTS FOR MULTIVARIATE REPEATED MEASUREMENTS OF HISTOGRAM OBJECTS WITH APPLICATIONS TO WEARABLE DEVICE DATA'

By Jingru Zhang[1], Kathleen R. Merikangas[2], Hongzhe Li[1] and Haochang Shou[1]

[1] *University of Pennsylvania*
[2] *National Institutes of Health*

SUPPLEMENT A: PROOFS OF THEOREMS

A.1. Proof of Theorem 2.1.

Proof. According to the following definitions

$$R_{\mathrm{out},1} = \sum_{(u,v)\in G_{\mathrm{out}}} I(g_u = g_v = 1),$$
$$R_{\mathrm{out},2} = \sum_{(u,v)\in G_{\mathrm{out}}} I(g_u = g_v = 2),$$
$$R_{\mathrm{in},1} = \sum_{(i,j)\in G_{\mathrm{in}}} I(g_i = g_j = 1),$$

we have

$$\mathbf{E}(R_{\mathrm{out},1}) = \sum_{(u,v)\in G_{\mathrm{out}}} \mathbf{P}(g_u = g_v = 1) = |G_{\mathrm{out}}|\frac{n_1(n_1-1)}{N(N-1)},$$
$$\mathbf{E}(R_{\mathrm{out},2}) = \sum_{(u,v)\in G_{\mathrm{out}}} \mathbf{P}(g_u = g_v = 2) = |G_{\mathrm{out}}|\frac{n_2(n_2-1)}{N(N-1)},$$
$$\mathbf{E}(R_{\mathrm{in},1}) = \sum_{(i,j)\in G_{\mathrm{in}}} \mathbf{P}(g_i = g_j = 1) = |G_{\mathrm{in}}|\frac{n_1}{N}.$$

For $\mathbf{Var}(R_{\text{out},1})$, we only need to figure out $\mathbf{E}(R^2_{\text{out},1})$. We have

$$
\begin{aligned}
&\mathbf{E}(R^2_{\text{out},1}) \\
=&\sum_{(u,v)\in G_{\text{out}}} P(g_u = g_v = 1) + \sum_{\substack{(u,v),(u,s)\in G_{\text{out}}\\ v\neq s}} P(g_u = g_v = g_s = 1) \\
&+ \sum_{\substack{(u,v),(s,t)\in G_{\text{out}}\\ u,v,s,t \text{ are all different}}} P(g_u = g_v = g_s = g_t = 1) \\
=&|G_{\text{out}}| \frac{n_1(n_1-1)}{N(N-1)} + \sum_{u\neq v} \frac{D_{uv}(D_{uv}-1)}{2} \frac{n_1(n_1-1)}{N(N-1)} \\
&+ \sum_{u\neq v} D_{uv}(D_u - D_{uv}) \frac{n_1(n_1-1)(n_1-2)}{N(N-1)(N-2)} \\
&+ \left(|G_{\text{out}}|^2 + \sum_{u\neq v} \frac{D^2_{uv}}{2} - \sum_u D^2_u \right) \frac{n_1(n_1-1)(n_1-2)(n_1-3)}{N(N-1)(N-2)(N-3)} \\
=&\sum_{u\neq v} \frac{D^2_{uv}}{2} \frac{n_1(n_1-1)}{N(N-1)} + \sum_{u\neq v} D_{uv}(D_u - D_{uv}) \frac{n_1(n_1-1)(n_1-2)}{N(N-1)(N-2)} \\
&+ \left(|G_{\text{out}}|^2 + \sum_{u\neq v} \frac{D^2_{uv}}{2} - \sum_u D^2_u \right) \frac{n_1(n_1-1)(n_1-2)(n_1-3)}{N(N-1)(N-2)(N-3)}.
\end{aligned}
$$

The analytic expression of $\mathbf{Var}(R_{\text{out},2})$ can be derived in a similar way as that of $\mathbf{Var}(R_{\text{out},1})$.

For $\mathbf{Cov}(R_{\text{out},1}, R_{\text{out},2})$, we only need to figure out $\mathbf{E}(R_{\text{out},1}, R_{\text{out},2})$. We have

$$
\begin{aligned}
\mathbf{E}(R_{\text{out},1} R_{\text{out},2}) &= \sum_{(u,v)(s,t)\in G_{\text{out}}} P(g_u = g_v = 1, g_s = g_t = 2) \\
&= \left(|G_{\text{out}}|^2 + \sum_{u\neq v} \frac{D^2_{uv}}{2} - \sum_u D^2_u \right) \frac{n_1(n_1-1)n_2(n_2-1)}{N(N-1)(N-2)(N-3)}.
\end{aligned}
$$

Then we compute $\mathbf{E}(R^2_{\text{in},1})$, $\mathbf{E}(R_{\text{out},1}, R_{\text{in},1})$ and $\mathbf{E}(R_{\text{out},2}, R_{\text{in},1})$ so that the analytic expressions of $\mathbf{Var}(R_{\text{in},1})$, $\mathbf{Cov}(R_{\text{out},1}, R_{\text{in},1})$ and $\mathbf{Cov}(R_{\text{out},2}, R_{\text{in},1})$ can be derived immediately. We have

$$\begin{aligned}
&\mathbf{E}(R_{\text{in},1}^2)\\
=&\sum_{(i,j)\in G_{\text{in}}} P(g_i = g_j = 1) + \sum_{\substack{(i,j),(i,k)\in G_{\text{in}}\\ j\neq k}} P(g_i = g_j = g_k = 1)\\
&+\sum_{\substack{(i,j),(k,l)\in G_{\text{in}}\\ i,j,k,l \text{ are all different}}} P(g_i = g_j = g_k = g_l = 1)\\
=&\sum_{u=1}^{N} D_{uu}^2 \frac{n_1 n_2}{N(N-1)} + |G_{\text{in}}|^2 \frac{n_1(n_1-1)}{N(N-1)},\\
&\mathbf{E}(R_{\text{in},1}R_{\text{out},1})\\
=&\sum_{\substack{(i,j)\in G_{\text{in}}, i\in\mathcal{C}_s\\ (s,t)\in G_{\text{out}}}} P(g_i = g_j = g_t = 1) + \sum_{\substack{(i,j)\in G_{\text{in}}, i\notin\mathcal{C}_s, j\notin\mathcal{C}_t\\ (s,t)\in G_{\text{out}}}} P(g_i = g_j = g_s = g_t = 1)\\
=&\sum_{u=1}^{N} D_{uu}D_u \frac{n_1(n_1-1)}{N(N-1)} + \left(|G_{\text{in}}||G_{\text{out}}| - \sum_{u=1}^{N} D_{uu}D_u\right) \frac{n_1(n_1-1)(n_1-2)}{N(N-1)(N-2)},\\
&\mathbf{E}(R_{\text{in},1}R_{\text{out},2})\\
=&\sum_{\substack{(i,j)\in G_{\text{in}}, i\notin\mathcal{C}_s, j\notin\mathcal{C}_t\\ (s,t)\in G_{\text{out}}}} P(g_i = g_j = 1,\ g_s = g_t = 2)\\
=&\left(|G_{\text{in}}||G_{\text{out}}| - \sum_{u=1}^{N} D_{uu}D_u\right) \frac{n_1 n_2(n_2-1)}{N(N-1)(N-2)}.
\end{aligned}$$

□

A.2. Rewrite S_R.

Lemma A.1. *The statistic S_R can be rewritten in the following form*

$$S_R = (Z_{out,w}, Z_{out,d}, Z_{in})\mathbf{\Omega}^{-1}(Z_{out,w}, Z_{out,d}, Z_{in})^T,$$

where $\mathbf{\Omega} = \mathbf{Var}((Z_{out,w}, Z_{out,d}, Z_{in})^T)$.

Proof. Let

$$\mathbf{R} = \begin{pmatrix} R_{\text{out},1} - \mathbf{E}(R_{\text{out},1})\\ R_{\text{out},2} - \mathbf{E}(R_{\text{out},2})\\ R_{\text{in},1} - \mathbf{E}(R_{\text{in},1}) \end{pmatrix},$$

$$\begin{aligned}
\mathbf{Z} =& \begin{pmatrix} Z_{\text{out},m}\\ Z_{\text{out},d}\\ Z_{\text{in}} \end{pmatrix}\\
=& \begin{pmatrix} \frac{n_2-1}{\sqrt{\mathbf{Var}((n_2-1)R_{\text{out},1}+(n_1-1)R_{\text{out},2})}} & \frac{n_1-1}{\sqrt{\mathbf{Var}((n_2-1)R_{\text{out},1}+(n_1-1)R_{\text{out},2})}} & 0\\ \frac{1}{\sqrt{\mathbf{Var}(R_{\text{out},1}-R_{\text{out},2})}} & -\frac{1}{\sqrt{\mathbf{Var}(R_{\text{out},1}-R_{\text{out},2})}} & 0\\ 0 & 0 & \frac{1}{\sqrt{\mathbf{Var}(R_{\text{in},1})}} \end{pmatrix}\mathbf{R}\\
\triangleq& \mathbf{BR}.
\end{aligned}$$

It is easy to see that $\mathbf{B}$ is invertible. From the definition of S, it can be written as

$$S = \mathbf{R}^T\mathbf{\Sigma}^{-1}\mathbf{R} = (\mathbf{B}^{-1}\mathbf{Z})^T\mathbf{\Sigma}^{-1}(\mathbf{B}^{-1}\mathbf{Z}) = \mathbf{Z}^T(\mathbf{B}\mathbf{\Sigma}\mathbf{B}^T)^{-1}\mathbf{Z}.$$

Here, $\mathbf{B}\mathbf{\Sigma}\mathbf{B}^T$ is the variance of $(Z_{\text{out},w}, Z_{\text{out},d}, Z_{\text{in}})^T$. □

A.3. Proof of Theorem 3.1. Applying Stein's method, we prove $(R_{\text{out},1}, R_{\text{out},2}, R_{\text{in},1})^T$ converges in distribution to a trivariate Gaussian distribution as $N \to \infty$ first. Consider sums of the form $W = \sum_{i\in\mathcal{J}} \xi_i$, where $\mathcal{J}$ is an index set and ξ_i are random variables with $\mathbf{E}(\xi_i) = 0$, and $\mathbf{E}(W^2) = 1$. The following assumption restricts the dependence between $\{\xi_i : i \in \mathcal{J}\}$.

ASSUMPTION A.1. *[Chen and Shao (2005), p. 17] For each $i \in \mathcal{J}$ there exists $S_i \subset T_i \subset \mathcal{J}$ such that ξ_i is independent of $\xi_{S_i^c}$ and ξ_{S_i} is independent of $\xi_{T_i^c}$.*

We will use the following theorem.

THEOREM A.2. *[Chen and Shao (2005), Theorem 3.4] Under Assumption A.1, we have*

$$\sup_{h\in Lip(1)} |\mathbf{E}h(W) - \mathbf{E}h(Z)| \le \delta$$

where $Lip(1) = \{h : \mathbb{R} \to \mathbb{R},\ \| h' \| \le 1\}$, Z has $\mathcal{N}(0,1)$ distribution and

$$\delta = 2\sum_{i\in\mathcal{J}}(\mathbf{E}|\xi_i\eta_i\theta_i| + |\mathbf{E}(\xi_i\eta_i)|\mathbf{E}|\theta_i|) + \sum_{i\in\mathcal{J}}\mathbf{E}|\xi_i\eta_i^2|$$

with $\eta_i = \sum_{j\in S_i}\xi_j$ and $\theta_i = \sum_{j\in T_i}\xi_j$, where S_i and T_i are defined in Assumption A.1.

To prove Theorem 3.1, we take one step back to study the statistic under the bootstrap null distribution, which is defined as follows: For each subject, we assign it to be from sample 1 with probability n_1/N, and from sample 2 with probability n_2/N, independently of other subjects. Let n_1^B be the number of subjects that are assigned to be from sample 1. Then, conditioning on $n_1^B = n_1$, the bootstrap null distribution becomes the permutation null distribution. We use $\mathbf{P}_B, \mathbf{E}_B, \mathbf{Var}_B$ to denote the probability, expectation, and variance under the bootstrap null distribution, respectively.

Let

$$\mathbf{E}(R_{\text{out},1}) \triangleq \mu_1, \quad \mathbf{E}(R_{\text{out},2}) \triangleq \mu_2, \quad \mathbf{E}(R_{\text{in},1}) \triangleq \mu_3,$$
$$\mathbf{Var}(R_{\text{out},1}) \triangleq (\sigma_1)^2, \quad \mathbf{Var}(R_{\text{out},2}) \triangleq (\sigma_2)^2, \quad \mathbf{Var}(R_{\text{in},1}) \triangleq (\sigma_3)^2,$$
$$\mathbf{Cov}(R_{\text{out},1}, R_{\text{out},2}) \triangleq \sigma_{12}, \quad \mathbf{Cov}(R_{\text{out},1}, R_{\text{in},1}) \triangleq \sigma_{13}, \quad \mathbf{Cov}(R_{\text{out},2}, R_{\text{in},1}) \triangleq \sigma_{23}.$$

Let $p_n = n_1/N, q_n = n_2/N$, then $\lim_{n\to\infty} p_n = p, \lim_{n\to\infty} q_n = q$. Using the similar steps as in the Proof A.1, we have

$$\begin{aligned}
&\mathbf{E}_B(R_{\text{out},1}) = |G_{\text{out}}|p_n^2 \triangleq \mu_1^B,\\
&\mathbf{E}_B(R_{\text{out},2}) = |G_{\text{out}}|q_n^2 \triangleq \mu_2^B,\\
&\mathbf{E}_B(R_{\text{in},1}) = |G_{\text{in}}|p_n \triangleq \mu_3^B,\\
&\mathbf{Var}_B(R_{\text{out},1}) = p_n^2q_n^2\left\{\frac{1}{2}\sum_{u\ne v}D_{uv}^2 + \frac{p_n}{q_n}\sum_u D_u^2\right\} \triangleq (\sigma_1^B)^2,\\
&\mathbf{Var}_B(R_{\text{out},1}) = p_n^2q_n^2\left\{\frac{1}{2}\sum_{u\ne v}D_{uv}^2 + \frac{q_n}{p_n}\sum_u D_u^2\right\} \triangleq (\sigma_2^B)^2,\\
&\mathbf{Var}_B(R_{\text{in},1}) = p_nq_n\sum_u D_{uu}^2 \triangleq (\sigma_3^B)^2.
\end{aligned}$$

Let

$$W_1^B = \frac{R_{\text{out},1} - \mu_1^B}{\sigma_1^B}, \quad W_1 = \frac{R_{\text{out},1} - \mu_1}{\sigma_1},$$
$$W_2^B = \frac{R_{\text{out},2} - \mu_2^B}{\sigma_2^B}, \quad W_2 = \frac{R_{\text{out},2} - \mu_2}{\sigma_2},$$
$$W_3^B = \frac{R_{\text{in},1} - \mu_3^B}{\sigma_3^B}, \quad W_3 = \frac{R_{\text{in},1} - \mu_3}{\sigma_3},$$
$$W_4^B = \frac{n_1^B - Np_n}{\sigma_0}, \text{ where } \sigma_0^2 = Np_n(1-p_n).$$

Under the conditions of Theorem 3.1, as $N \to \infty$, we can prove the following results:

(1) $(W_1^B, W_2^B, W_3^B, W_4^B)$ becomes multivariate Gaussian distributed under the bootstrap null.

(2)

$$\frac{\sigma_1^B}{\sigma_1} \to c_1, \quad \frac{\mu_1^B - \mu_1}{\sigma_1^B} \to 0; \quad \frac{\sigma_2^B}{\sigma_2} \to c_2, \quad \frac{\mu_2^B - \mu_2}{\sigma_2^B} \to 0;$$
$$\frac{\sigma_3^B}{\sigma_3} \to c_3, \quad \frac{\mu_3^B - \mu_3}{\sigma_3^B} \to 0,$$

where c_1, c_2 and c_3 are constants.

(3)

$$|\lim_{N\to\infty} \mathbf{Cor}(W_1, W_2)| < 1, \quad |\lim_{N\to\infty} \mathbf{Cor}(W_1, W_3)| < 1,$$
$$|\lim_{N\to\infty} \mathbf{Cor}(W_2, W_3)| < 1.$$

From (1) and given that $\mathbf{Var}_B(W_4^B) = 1$, the conditional distribution of $(W_1^B, W_2^B, W_3^B)^T$ given W_4^B is a trivariate Gaussian distribution under the bootstrap null distribution as $N \to \infty$. Since the permutation null distribution is equivalent to the bootstrap null distribution given $W_4^B = 0$, $(W_1^B, W_2^B, W_3^B)^T$ follows a trivariate Gaussian distribution under the permutation null distribution as $N \to \infty$. Since

$$W_1 = \frac{\sigma_1^B}{\sigma_1}\left(W_1^B + \frac{\mu_1^B - \mu_1}{\sigma_1^B}\right), \quad W_2 = \frac{\sigma_2^B}{\sigma_2}\left(W_2^B + \frac{\mu_2^B - \mu_2}{\sigma_2^B}\right),$$
$$W_3 = \frac{\sigma_3^B}{\sigma_3}\left(W_3^B + \frac{\mu_3^B - \mu_3}{\sigma_3^B}\right),$$

given (2), we have $(W_1, W_2, W_3)^T$ follows a trivariate Gaussian distribution under the permutation null distribution as $N \to \infty$. Together with (3), we can conclude that $(R_{\text{out},1}, R_{\text{out},2}, R_{\text{in},1})^T$ converges in distribution to a trivariate Gaussian distribution as $N \to \infty$. In the following, we prove the results (1)—(3).

To prove (1), by Cramér-Wold device, we only need to show that $W = a_1 W_1^B + a_2 W_2^B + a_3 W_3^B + a_4 W_4^B$ is asymptotically Gaussian distributed for any combination of a_1, a_2, a_3, a_4 such that $\mathbf{Var}_B(W) > 0$.

We first define more notations. For an edge (u, v) of G_{out}, i.e., $uv \in \mathcal{J}_1 = \{uv : u < v, (u, v) \in G_{\text{out}}\}$, let

$$\xi_{uv} = a_1 \frac{I(g_u = g_v = 1) - p_n^2}{\sigma_1^B} + a_2 \frac{I(g_u = g_v = 2) - q_n^2}{\sigma_2^B}.$$

For an edge (k,l) of G_{in}, i.e., $kl \in \mathcal{J}_2 = \{kl : k < l, (k,l) \in G_{\text{in}}\}$, let

$$\xi_{kl} = a_3 \frac{I(g_k = g_l = 1) - p_n}{\sigma_3^B}.$$

And for a subject $s \in \mathcal{J}_3 = \{1, \dots, N\}$, let

$$\xi_s = a_4 \frac{I(g_s = 1) - p_n}{\sigma_0}.$$

Thus,

$$W = a_1 W_1^B + a_2 W_2^B + a_3 W_3^B + a_4 W_4^B = \sum_{i \in \mathcal{J}} \xi_i,$$

where $\mathcal{J} = \mathcal{J}_1 \cup \mathcal{J}_2 \cup \mathcal{J}_3$.

For an edge $e = (u,v) \in G_{\text{out}}$, let

$$C_{\text{out},e} = \{u,v\} \cup \{s \in \{1,\dots,N\} : s \text{ is connected to } u \text{ or } v \text{ in } G_{\text{out}}\}.$$

For an edge $e = (i,j) \in G_{\text{in}},\ i,j \in \mathcal{C}_u$, let

$$C_{\text{in},e} = \{u\} \cup \{s \in \{1,\dots,N\} : s \text{ is connected to } u \text{ in } G_{\text{out}}\}.$$

We define

$$C_e = C_{\text{out},e} I(e \in G_{\text{out}}) + C_{\text{in},e} I(e \in G_{\text{in}})$$

and introduce following index sets to satisfy Assumption A.1. For $i \in \mathcal{J}_1$ (i.e., i is an edge $(u,v) \in G_{\text{out}}$), let

$$S_i = A_i \cup \{u,v\},$$
$$T_i = B_i \cup C_i.$$

For $i \in \mathcal{J}_2 \cup \mathcal{J}_3$ (i.e., i is an edge $(k,l) \in G_{\text{in}},\ k,l \in \mathcal{C}_s$ or i is a subject $s \in \mathcal{J}_3$),

$$S_i = A_i \cup \{s\},$$
$$T_i = B_i \cup C_i.$$

Then S_i and T_i satisfy Assumption A.1.

Let $a = \max\{|a_1|,|a_2|,|a_3|,|a_4|\}$ and $\sigma = \min\{\sigma_1^B, \sigma_2^B, \sigma_3^B, \sigma_0\}$. Since

$$|\xi_i| \le \begin{cases} 2a/\sigma & \text{if } i \in \mathcal{J}_1, \\ a/\sigma & \text{if } i \in \mathcal{J}_2 \cup \mathcal{J}_3 \end{cases},$$

we have

$$\sum_{j \in S_i} |\xi_j| \le (|A_i| + 2) 2a/\sigma, \quad \sum_{j \in T_i} |\xi_j| \le (|B_i| + |C_i|) 2a/\sigma, \quad i \in \mathcal{J},$$

and the terms $\mathbf{E}|\xi_i \eta_i \theta_i|$, $|\mathbf{E}(\xi_i \eta_i)| \mathbf{E}|\theta_i|$ and $\mathbf{E}|\xi_i \eta_i^2|$ are all bounded by

$$\frac{32a^3}{\sigma^3} |A_i||B_i|.$$

So we have

$$\begin{aligned}\delta =&\frac{1}{\sqrt{(\mathbf{Var}_B(W))^3}}\left\{2\sum_{i\in\mathcal{J}}(\mathbf{E}_B|\xi_i\eta_i\theta_i| + |\mathbf{E}_B(\xi_i\eta_i)|\mathbf{E}_B|\theta_i|) + \sum_{i\in\mathcal{J}}\mathbf{E}_B|\xi_i\eta_i^2|\right\}\\ \leq&\frac{1}{\sqrt{(\mathbf{Var}_B(W))^3}}\frac{160a^3}{\sigma^3}\sum_{i\in\mathcal{J}}|A_i||B_i|\\ \leq&\frac{480a^3}{\sigma^3\sqrt{(\mathbf{Var}_B(W))^3}}\sum_{e\in G_{\text{out}}}|A_{\text{out},e}||B_{\text{out},e}|\end{aligned}$$

Since $480a^3/\sqrt{(\mathbf{Var}_B(W))^3}$ is of constant order and $\sigma = O(N^{0.5})$, we have $\delta \to 0$ when $\sum_{e\in G_{\text{out}}}|A_{\text{out},e}||B_{\text{out},e}| = o(N^{1.5})$.

Next we prove result (2). Since $|G_{\text{out}}| = O(N)$ and $\sum_u D_u^2 - 4|G_{\text{out}}|^2/N = O(N)$, let $\lim_{N\to\infty}|G_{\text{out}}|/N = b_1$ and $\lim_{N\to\infty}\sum_u D_u^2/N - 4|G_{\text{out}}|^2/N^2 = b_2$, $b_1, b_2 \in (0,\infty)$. Then $\lim_{N\to\infty}\sum_u D_u^2/N = 4b_1^2 + b_2$, i.e., $\sum_u D_u^2 = O(N)$. Since

$$2|G_{\text{out}}| = \sum_{u\neq v} D_{uv} \leq \sum_{u\neq v} D_{uv}^2 \leq \sum_u D_u^2,$$

we have $\sum_{u\neq v} D_{uv}^2 = O(N)$ and let $\lim_{N\to\infty}\sum_{u\neq v} D_{uv}^2/N = b_3 \in (0,\infty)$. Hence, as $N\to\infty$

$$\begin{aligned}&\frac{(\sigma_1^B)^2}{N} \longrightarrow p^2q^2\left\{\frac{1}{2}b_3 + \frac{p}{q}(4b_1^2 + b_2)\right\},\\ &\frac{(\sigma_1)^2}{N} \longrightarrow p^2q^2\left\{\frac{1}{2}b_3 + \frac{p}{q}b_2\right\},\\ &\frac{\sigma_1^B}{\sigma_1} \longrightarrow \sqrt{1 + \frac{8pb_1^2}{qb_3 + 2pb_2}}.\end{aligned}$$

Similarly, we have

$$\frac{\sigma_2^B}{\sigma_2} \longrightarrow \sqrt{1 + \frac{8qb_1^2}{pb_3 + 2qb_2}}.$$

Since $|G_{\text{in}}| = O(N)$ and $\sum_u D_{uu}^2 - |G_{\text{in}}|^2/N = O(N)$, let $\lim_{N\to\infty}|G_{\text{in}}|/N = b_4$ and $\lim_{N\to\infty}\sum_u D_{uu}^2/N - |G_{\text{in}}|^2/N^2 = b_5$, $b_4, b_5 \in (0,\infty)$. Then $\lim_{N\to\infty}\sum_u D_{uu}^2/N = b_4^2 + b_5$, and we have

$$\begin{aligned}&\frac{(\sigma_3^B)^2}{N} \longrightarrow pq(b_4^2 + b_5),\\ &\frac{(\sigma_3)^2}{N} \longrightarrow pqb_5,\\ &\frac{\sigma_3^B}{\sigma_3} \longrightarrow \sqrt{1 + \frac{b_4^2}{b_5}}.\end{aligned}$$

Also,

$$\mu_1^B - \mu_1 = |G_{\text{out}}|\frac{n_1n_2}{N^2(N-1)},$$

so

$$\lim_{N\to\infty}\frac{\mu_1^B - \mu_1}{\sigma_1^B} = 0.$$

Similarly, we have

$$\lim_{N\to\infty} \frac{\mu_2^B - \mu_2}{\sigma_2^B} = 0.$$

Since $\mu_3^B - \mu_3 = 0$,

$$\lim_{N\to\infty} \frac{\mu_3^B - \mu_3}{\sigma_3^B} = 0.$$

Last, we prove result (3). As $N \to \infty$,

$$\frac{\sigma_{12}}{N} \longrightarrow p^2 q^2 \left\{ \frac{1}{2} b_3 - b_2 \right\}.$$

We have

$$\begin{aligned} \lim_{N\to\infty} \mathbf{Cor}(W_1, W_2) &= \lim_{N\to\infty} \frac{\sigma_{12}}{\sqrt{(\sigma_1)^2(\sigma_2)^2}} \\ &= \frac{b_3 - 2b_2}{\sqrt{(b_3 - 2b_2)^2 + 2b_2 b_3/pq}}, \end{aligned}$$

Strictly positive $2b_2b_3/pq$ implies $|\lim_{N\to\infty} \mathbf{Cor}(W_1, W_2)| < 1$. Note that

$$|G_{\text{in}}| = \sum_u D_{uu} \le \sum_u D_{uu} D_u \le \sqrt{\sum_u D_{uu}^2} \sqrt{\sum_u D_u^2},$$

$|G_{\text{in}}| = O(N)$, $\sum_u D_{uu}^2 = O(N)$ and $\sum_u D_u^2 = O(N)$. We have $\sum_u D_{uu} D_u = O(N)$ and let $\lim_{N\to\infty} \sum_u D_{uu} D_u / N = b_6 \in (0, \infty)$. Thus,

$$\frac{\sigma_{13}}{N} \longrightarrow p^2 q (b_6 - 2b_1 b_4),$$

and

$$\begin{aligned} \lim_{N\to\infty} \mathbf{Cor}(W_1, W_3) &= \lim_{N\to\infty} \frac{\sigma_{13}}{\sqrt{(\sigma_1)^2(\sigma_3)^2}} \\ &= \frac{p^2 q (b_6 - 2b_1 b_4)}{\sqrt{\frac{1}{2} p^3 q^3 b_3 b_5 + p^4 q^2 b_2 b_5}}, \end{aligned}$$

Note that

$$\left| \frac{\sum_{u=1}^N D_{uu} D_u - \frac{2}{N} |G_{\text{in}}||G_{\text{out}}|}{\sqrt{(\sum_{u=1}^N D_u^2 - \frac{4|G_{\text{out}}|^2}{N})(\sum_{u=1}^N D_{uu}^2 - \frac{|G_{\text{in}}|^2}{N})}} \right| = |\mathbf{Cor}(Z_{\text{out},d}, Z_{\text{in}})| \le 1,$$

so

$$\left| \frac{b_6 - 2b_1 b_4}{\sqrt{b_2 b_5}} \right| = \left| \lim_{N\to\infty} \frac{\frac{\sum_{u=1}^N D_{uu} D_u}{N} - \frac{2}{N^2} |G_{\text{in}}||G_{\text{out}}|}{\sqrt{(\sum_{u=1}^N \frac{D_u^2}{N} - \frac{4|G_{\text{out}}|^2}{N^2})(\sum_{u=1}^N \frac{D_{uu}^2}{N} - \frac{|G_{\text{in}}|^2}{N^2})}} \right| \le 1.$$

We have $|\lim_{N\to\infty} \mathbf{Cor}(W_1, W_3)| < 1$, since $p^3 q^3 b_3 b_5 / 2$ is strictly positive. Similarly, we obtain $|\lim_{N\to\infty} \mathbf{Cor}(W_2, W_3)| < 1$.

Since $(Z_{\text{out},w}, Z_{\text{out},d}, Z_{\text{in}})^T$ is the linear transformation of $(R_{\text{out},1}, R_{\text{out},2}, R_{\text{in},1})^T$, the proof above implies $(Z_{\text{out},w}, Z_{\text{out},d}, Z_{\text{in}})^T$ converges in distribution to a trivariate Gaussian distribution with mean $\mathbf{0}$ and covariance matrix Γ. As shown in Proof A.2, the linear transformation is nondegenerate. Thus Γ is invertible and $\Gamma = \left(\begin{smallmatrix} 1 & 0 & 0 \\ 0 & 1 & \rho_Z \\ 0 & \rho_Z & 1 \end{smallmatrix}\right)$, where

$$\rho_Z = \lim_{N\to\infty} \frac{\sum_{u=1}^{N} D_{uu}D_u - \frac{2}{N}|G_{\text{in}}||G_{\text{out}}|}{\sqrt{(\sum_{u=1}^{N} D_u^2 - \frac{4|G_{\text{out}}|^2}{N})(\sum_{u=1}^{N} D_{uu}^2 - \frac{|G_{\text{in}}|^2}{N})}} = \frac{b_6 - 2b_1 b_4}{\sqrt{b_2 b_5}}$$

and $|\rho_Z| < 1$.

SUPPLEMENT B: ADDITIONAL SIMULATION RESULTS FOR DATA WITH AN EXPONENTIALLY DECAYED WITHIN-SUBJECT CORRELATION

(1) Simulations for one-dimensional density, $p = 1$.

TABLE 1
Parameter values for 5 different simulation settings for comparison one-dimensional density functions.

A1': null model.
$\rho_1 = 0.6,\ \beta_1 = 0,\ \epsilon_1 = 1,\ \nu_{11} = 1,\ \nu_{12} = 2;$
$\rho_2 = 0.6,\ \beta_2 = 0,\ \epsilon_2 = 1,\ \nu_{21} = 1,\ \nu_{22} = 2;\ \sigma = 1.$
A2': within-subject variability difference in ρ.
$\rho_1 = 0,\ \beta_1 = 0,\ \epsilon_1 = 1,\ \nu_{11} = 1,\ \nu_{12} = 1.2;$
$\rho_2 = 0.9,\ \beta_2 = 0,\ \epsilon_2 = 1,\ \nu_{21} = 1,\ \nu_{22} = 1.2;\ \sigma = 1.$
A3': between-subject mean difference in β and $\nu_{\cdot 1} + \nu_{\cdot 2}$.
$\rho_1 = 0,\ \beta_1 = 0,\ \epsilon_1 = 1,\ \nu_{11} = 1,\ \nu_{12} = 1.2;$
$\rho_2 = 0,\ \beta_2 = 0.7,\ \epsilon_2 = 1,\ \nu_{21} = 0.96,\ \nu_{22} = 1.16;\ \sigma = 1.$
A4': between-subject variability difference in ϵ and $\nu_{\cdot 2} - \nu_{\cdot 1}$.
$\rho_1 = 0,\ \beta_1 = 0,\ \epsilon_1 = 1,\ \nu_{11} = 1,\ \nu_{12} = 1.3;$
$\rho_2 = 0,\ \beta_2 = 0,\ \epsilon_2 = 1.1,\ \nu_{21} = 0.97,\ \nu_{22} = 1.33;\ \sigma = 1.$
A5': within-subject variability difference in ρ, between-subject mean difference in β and $\nu_{\cdot 1} + \nu_{\cdot 2}$, variance difference in ϵ and $\nu_{\cdot 2} - \nu_{\cdot 1}$.
$\rho_1 = 0,\ \beta_1 = 0,\ \epsilon_1 = 1,\ \nu_{11} = 1,\ \nu_{12} = 1.3;$
$\rho_2 = 0.35,\ \beta_2 = 0.5,\ \epsilon_2 = 1.1,\ \nu_{21} = 0.97,\ \nu_{22} = 1.36;\ \sigma = 1.$

TABLE 2
Empirical power of the proposed test statistics in the first 6 columns, generalized edge-count test ($S1$, $S2$) and Fréchet test (Fretest1, Fretest2) at 0.05 significance level under the five scenarios denoted by A1'–A5'. Those above 95 percentage of the best power under A2'–A5' are in bold.

	T_{in}	$Z_{\text{out},w}$	$T_{\text{out},d}$	M_{out}	S_R	M	$S1$	$S2$	Fretest1	Fretest2
					Null model					
A1'	0.057	0.055	0.046	0.061	0.055	0.063	0.056	0.047	0.048	0.054
					Alternative model					
A2'	**0.966**	0.050	0.149	0.111	0.851	0.901	0.153	**1.000**	0.463	0.056
A3'	0.036	**0.970**	0.071	**0.960**	0.914	**0.949**	0.640	0.555	0.293	0.288
A4'	0.053	0.186	**0.908**	0.859	0.790	0.814	0.140	0.062	0.304	0.277
A5'	0.067	0.697	**0.993**	**0.995**	**0.986**	**0.991**	0.398	0.391	0.467	0.333

(2) Simuations for moderate-dimensional density, $p = 30$.

Table 3

Parameter values for 5 different simulation settings for comparison 30-dimensional density functions.

B1': null model.
$\rho_1 = 0.3,\ \beta_1 = \mathbf{0}_p,\ \epsilon_1 = 1,\ \nu_{11} = 1,\ \nu_{12} = 2;$
$\rho_2 = 0.3,\ \beta_2 = \mathbf{0}_p,\ \epsilon_2 = 1,\ \nu_{21} = 1,\ \nu_{22} = 2;\ \sigma = 1.$
B2': within-subject variability difference in ρ.
$\rho_1 = 0,\ \beta_1 = \mathbf{0}_p,\ \epsilon_1 = 1,\ \nu_{11} = 1,\ \nu_{12} = 1.3;$
$\rho_2 = 0.25,\ \beta_2 = \mathbf{0}_p,\ \epsilon_2 = 1,\ \nu_{21} = 1,\ \nu_{22} = 1.3;\ \sigma = 1.$
B3': between-subject mean difference in β and $\nu_{\cdot 1} + \nu_{\cdot 2}$.
$\rho_1 = 0,\ \beta_1 = \mathbf{0}_p,\ \epsilon_1 = 1,\ \nu_{11} = 1,\ \nu_{12} = 1.3;$
$\rho_2 = 0,\ \beta_2 = 0.1\mathbf{1}_p,\ \epsilon_2 = 1,\ \nu_{21} = 1.2,\ \nu_{22} = 1.5;\ \sigma = 1.$
B4': between-subject variability difference in ϵ and $\nu_{\cdot 2} - \nu_{\cdot 1}$.
$\rho_1 = 0,\ \beta_1 = \mathbf{0}_p,\ \epsilon_1 = 1,\ \nu_{11} = 1,\ \nu_{12} = 1.3;$
$\rho_2 = 0,\ \beta_2 = \mathbf{0}_p,\ \epsilon_2 = 1.1,\ \nu_{21} = 0.8,\ \nu_{22} = 1.5;\ \sigma = 1.$
B5': within-subject variability difference in ρ, between-subject mean difference in β and $\nu_{\cdot 1} + \nu_{\cdot 2}$, variance difference in ϵ and $\nu_{\cdot 2} - \nu_{\cdot 1}$.
$\rho_1 = 0,\ \beta_1 = \mathbf{0}_p,\ \epsilon_1 = 1,\ \nu_{11} = 1,\ \nu_{12} = 1.3;$
$\rho_2 = 0.22,\ \beta_2 = 0.2\mathbf{1}_p,\ \epsilon_2 = 1.03,\ \nu_{21} = 1,\ \nu_{22} = 1.4;\ \sigma = 1.$

Table 4

Empirical power of the proposed test statistics in the first 6 columns, generalized edge-count test ($S1$, $S2$) and Fréchet test (Fretest1, Fretest2) at 0.05 significance level under the five scenarios denoted by B1'–B5'. Those above 95 percentage of the best power under B2'–B5' are in bold.

	T_{in}	$Z_{\text{out},w}$	$T_{\text{out},d}$	M_{out}	S_R	M	$S1$	$S2$	Fretest1	Fretest2
					Null model					
B1'	0.041	0.058	0.039	0.045	0.049	0.051	0.047	0.050	0.050	0.066
					Alternative model					
B2'	**0.934**	0.054	0.046	0.043	0.838	0.852	0.212	0.054	0.485	0.079
B3'	0.057	**0.945**	0.049	**0.916**	0.812	0.893	0.772	0.717	0.118	0.181
B4'	0.144	0.241	**0.882**	0.835	0.774	0.806	0.720	0.800	**0.859**	**0.883**
B5'	**0.892**	0.668	0.131	0.596	**0.911**	**0.906**	0.756	0.483	0.853	0.461

SUPPLEMENT C: ADDITIONAL RESULTS ON COMPARISON OF ASYMPTOTIC P-VALUES AND PERMUTATION P-VALUES

We next examine whether the asymptotic p-values are close to the p-values obtained based on 10,000 permutations. We consider the data generating procedure in Section 4.1 with the parameters set as those in model (A5). In particular, under the chosen sample sizes n_1=n_2=25, 50, 75 and 100, data were generated over 100 simulation runs. For each run, we estimate the asymptotic p-values and permutation p-values over 10,000 permutations.

Figure 4 shows the empirical powers of each test estimated by the asymptotic and permutation p-values over the 100 runs for n_1=n_2=25, 50, 75, 100. The results show that the power obtained by the asymptotic p-value is very close to that based on the permutation p-value for all the proposed test statistics. As sample size increases, the results are almost identical as expected. We also provide the comparison of asymptotic p-value and permutation p-value over 100 simulation runs in the Supplementary Material.

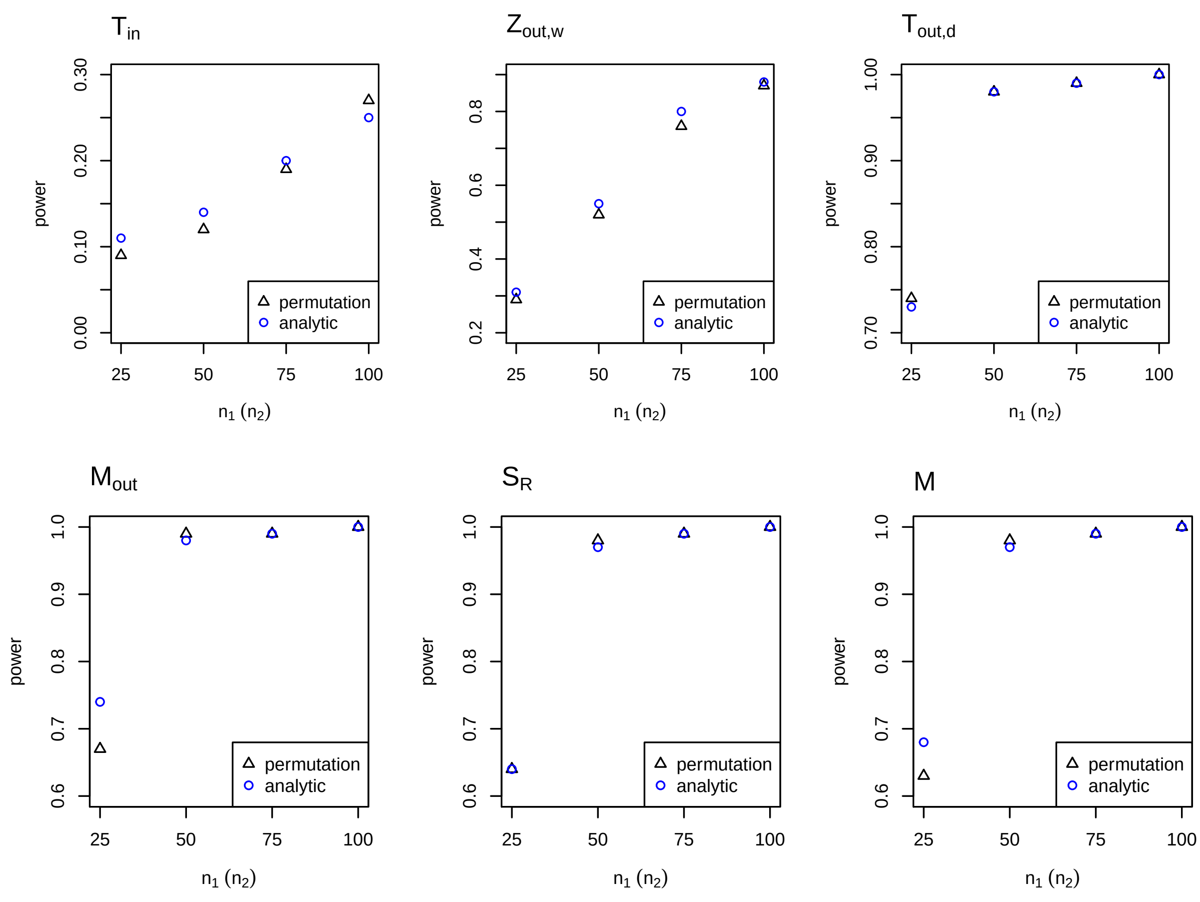

FIG 1. *Comparison of the empirical power estimated by the asymptotic p-value and by the p-value calculated from 10,000 permutations for different proposed test statistics.*

Figures 2 and 3 show the comparison of asymptotic p-value and permutation p-value over 100 simulation runs. The results show that the asymptotic p-values are very similar to that based on the permutation p-value for all the proposed test statistics. As sample size increases, the results are almost identical as expected.

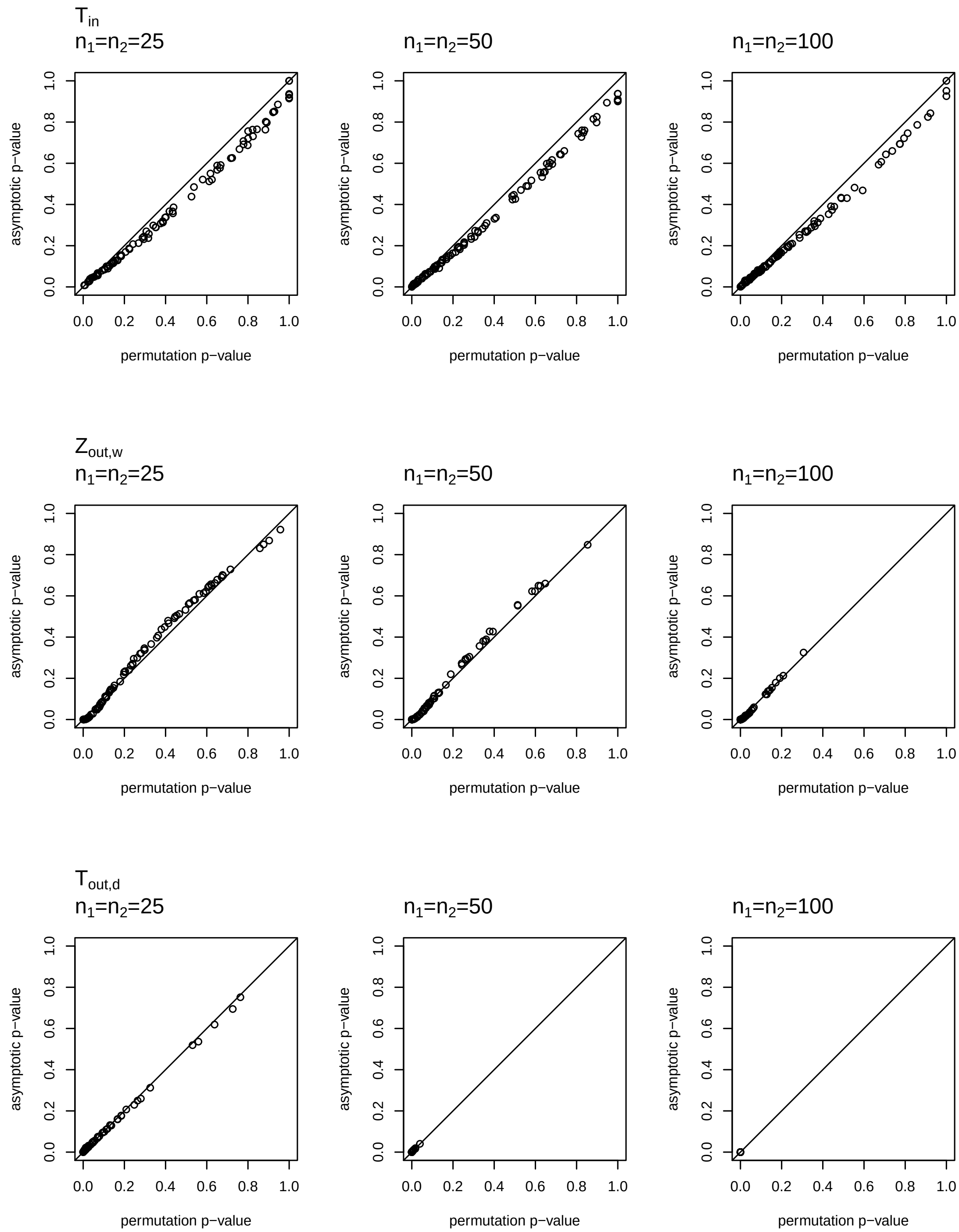

FIG 2. *Compare the asymptotic p-value with the p-value calculated from 10,000 permutations with 100 simulations for test statistics* T_{in}, $Z_{out,w}$ *and* $T_{out,d}$.

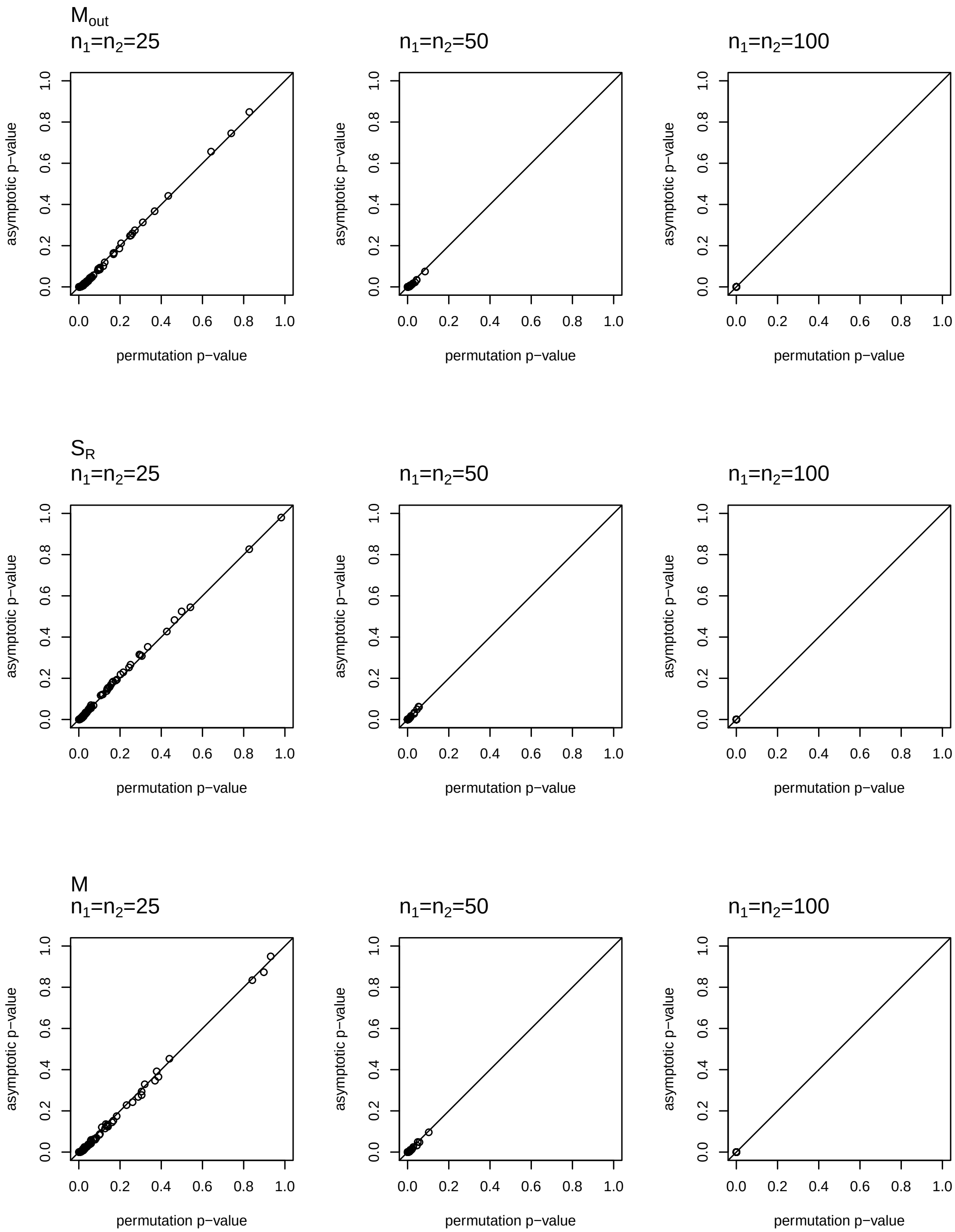

FIG 3. *Compare the asymptotic p-value with the p-value calculated from 10,000 permutations with 100 simulations for test statistics* $M_{out}(1.14)$*,* S_R *and* $M(1, 1.14)$*.*

SUPPLEMENT D: DETAILED P-VALUES IN REAL DATA ANALYSIS

Table 5

Comparisons of activity distributions among the controls, MDD, BPI and BPII patients. The p-values are presented for the proposed test statistics, the generalized edge-count tests ($S1$, $S2$) and Fréchet tests (Fretest1, Fretest2) (bold for those <0.05).

weekdays, $n_1 = 117,\ n_2 = 106,\ n_3 = 26,\ n_4 = 32.$										
	T_{in}	$Z_{\text{out},w}$	$T_{\text{out},d}$	M_{out}	S_R	M	$S1$	$S2$	Fretest1	Fretest2
HC vs MDD	0.729	0.625	0.307	0.428	0.567	0.566	0.472	0.171	0.381	0.305
HC vs BPII	0.777	0.442	0.238	0.345	0.550	0.488	0.305	0.053	0.661	0.231
HC vs BPI	0.770	0.061	0.867	0.140	0.405	0.205	0.734	0.734	0.262	0.105
MDD vs BPII	0.708	0.464	0.067	0.116	0.155	0.169	0.243	0.133	0.708	0.769
MDD vs BPI	0.746	0.119	0.617	0.272	0.539	0.377	0.693	0.680	0.703	0.491
BPII vs BPI	0.718	0.055	0.119	0.107	0.088	0.156	0.286	0.151	0.123	0.127
weekends, $n_1 = 116,\ n_2 = 98,\ n_3 = 30,\ n_4 = 33.$										
	T_{in}	$Z_{\text{out},w}$	$T_{\text{out},d}$	M_{out}	S_R	M	$S1$	$S2$	Fretest1	Fretest2
HC vs MDD	0.404	0.623	0.700	0.780	0.695	0.632	0.471	0.710	0.061	0.060
HC vs BPII	0.652	0.618	0.658	0.747	0.817	0.781	0.498	0.404	0.539	0.515
HC vs BPI	0.287	**0.003**	0.068	**0.005**	**0.005**	**0.006**	0.063	**0.018**	**0.014**	**0.002**
MDD vs BPII	0.714	0.451	0.848	0.813	0.885	0.828	0.482	0.728	0.274	0.317
MDD vs BPI	0.744	0.085	0.317	0.161	0.297	0.237	0.154	0.133	0.480	0.288
BPII vs BPI	0.699	**0.002**	0.250	**0.004**	**0.017**	**0.005**	**0.005**	0.054	**0.007**	**0.005**

Table 6

Comparisons of activity distrbutions in different age groups, where C1, C2 and C3 denoting young, middle-aged and older age groups. The p-values of the proposed test statistics, the generalized edge-count tests ($S1$, $S2$) and Fréchet tests (Fretest1, Fretest2) are presented for different comparisons (bold for those <0.05).

weekdays										
	T_{in}	$Z_{\text{out},w}$	$T_{\text{out},d}$	M_{out}	S_R	M	$S1$	$S2$	Fretest1	Fretest2
HC 36, 55, 25										
C1 vs C2	**0.020**	0.139	0.513	0.314	**0.028**	0.054	0.812	0.060	0.718	0.547
C1 vs C3	**0.004**	**<1e-3**	0.867	**<1e-3**	**<1e-3**	**<1e-3**	**<1e-3**	**<1e-3**	0.078	0.116
C2 vs C3	0.201	**0.001**	0.781	**0.002**	**0.007**	**0.003**	0.332	**0.008**	0.272	0.369
MDD 21, 62, 20										
C1 vs C2	0.383	0.295	0.879	0.680	0.639	0.596	0.543	0.617	0.053	0.060
C1 vs C3	0.160	**0.012**	0.841	**0.024**	**0.044**	**0.033**	**0.048**	0.144	**0.011**	**0.033**
C2 vs C3	0.546	0.464	0.725	0.769	0.850	0.746	0.745	0.819	0.862	0.674
weekends										
	T_{in}	$Z_{\text{out},w}$	$T_{\text{out},d}$	M_{out}	S_R	M	$S1$	$S2$	Fretest1	Fretest2
HC 36, 53, 26										
C1 vs C2	0.722	0.069	0.349	0.147	0.261	0.215	0.120	0.241	0.655	0.159
C1 vs C3	0.317	**0.001**	**0.013**	**<1e-3**	**<1e-3**	**0.001**	**0.047**	**0.004**	0.558	0.083
C2 vs C3	0.478	**0.027**	**0.042**	**0.035**	**0.017**	**0.050**	0.671	0.090	0.770	0.483
MDD 18, 55, 22										
C1 vs C2	0.408	0.468	0.276	0.375	0.394	0.408	0.426	0.210	0.358	0.604
C1 vs C3	0.533	**0.025**	0.155	**0.047**	0.065	0.070	0.145	0.110	0.073	0.097
C2 vs C3	0.633	0.526	0.684	0.726	0.767	0.726	0.868	0.691	0.543	0.355

Table 7

Comparisons of activity distrbutions in different BMI groups, where D1 and D2 denoting lean and obese people, respectively. The p-values of the proposed test statistics, the generalized edge-count tests ($S1$, $S2$) and Fréchet tests (Fretest1, Fretest2) are presented for different comparisons (bold for those <0.05).

					weekdays					
	T_{in}	$Z_{\text{out},w}$	$T_{\text{out},d}$	M_{out}	S_R	M	$S1$	$S2$	Fretest1	Fretest2
HC 32, 48										
D1 vs D2	0.815	0.837	0.487	0.621	0.645	0.776	0.809	0.788	0.529	0.390
OTHER 43, 75										
D1 vs D2	0.756	**0.009**	0.163	**0.017**	**0.040**	**0.022**	0.384	0.614	0.826	0.858
					weekends					
	T_{in}	$Z_{\text{out},w}$	$T_{\text{out},d}$	M_{out}	S_R	M	$S1$	$S2$	Fretest1	Fretest2
HC 30, 49										
D1 vs D2	0.579	0.718	0.366	0.493	0.490	0.566	0.636	0.375	0.262	0.168
OTHER 40, 74										
D1 vs D2	0.715	**0.030**	0.223	0.063	0.133	0.094	0.372	0.803	0.533	0.673

SUPPLEMENT E: SENSIVITY ANALYSIS RESULTS OF REAL APPLICATION WHEN USING 5-MST AND 15-MST AS THE SIMILARITY GRAPH

E.1 5-MST with the 2-Wasserstein distance as the similarity graph

TABLE 8
Comparisons of activity distributions among the controls, MDD, BPI and BPII patients. The p-values are presented for the proposed test statistics, the generalized edge-count tests ($S1$, $S2$) and Fréchet tests (Fretest1, Fretest2) (bold for those <0.05).

				weekdays, $n_1 = 117,\ n_2 = 106,\ n_3 = 26,\ n_4 = 32.$						
	T_{in}	$Z_{\text{out},w}$	$T_{\text{out},d}$	M_{out}	S_R	M	$S1$	$S2$	Fretest1	Fretest2
HC vs MDD	0.701	0.803	0.658	0.769	0.674	0.807	0.426	0.212	0.384	0.307
HC vs BPII	0.728	0.380	0.392	0.485	0.751	0.643	0.164	0.056	0.661	0.230
HC vs BPI	0.592	0.192	0.730	0.433	0.657	0.523	0.611	0.754	0.262	0.104
MDD vs BPII	0.834	0.390	0.134	0.210	0.354	0.303	0.176	0.181	0.711	0.767
MDD vs BPI	0.773	0.263	0.544	0.496	0.734	0.636	0.620	0.577	0.705	0.485
BPII vs BPI	0.712	0.137	0.260	0.250	0.265	0.348	0.508	0.116	0.124	0.127
				weekends, $n_1 = 116,\ n_2 = 98,\ n_3 = 30,\ n_4 = 33.$						
	T_{in}	$Z_{\text{out},w}$	$T_{\text{out},d}$	M_{out}	S_R	M	$S1$	$S2$	Fretest1	Fretest2
HC vs MDD	0.541	0.655	0.838	0.877	0.788	0.779	0.640	0.676	0.052	0.060
HC vs BPII	0.405	0.491	0.725	0.723	0.691	0.592	0.596	0.403	0.542	0.515
HC vs BPI	0.206	**0.021**	0.092	**0.031**	**0.023**	**0.036**	0.275	0.077	**0.014**	**0.002**
MDD vs BPII	0.759	0.431	0.830	0.782	0.905	0.845	0.530	0.718	0.279	0.322
MDD vs BPI	0.654	0.176	0.276	0.237	0.378	0.334	0.354	0.196	0.477	0.288
BPII vs BPI	0.720	**0.030**	0.670	0.065	0.209	0.095	**0.026**	0.090	**0.007**	**0.005**

TABLE 9
Comparisons of activity distrbutions in different age groups, where C1, C2 and C3 denoting young, middle-aged and older age groups. The p-values of the proposed test statistics, the generalized edge-count tests ($S1$, $S2$) and Fréchet tests (Fretest1, Fretest2) are presented for different comparisons (bold for those <0.05).

						weekdays				
	T_{in}	$Z_{\text{out},w}$	$T_{\text{out},d}$	M_{out}	S_R	M	$S1$	$S2$	Fretest1	Fretest2
HC 36, 55, 25										
C1 vs C2	0.109	0.074	0.663	0.169	**<1e-3**	0.142	0.473	**0.044**	0.709	0.559
C1 vs C3	**0.005**	**<1e-3**	0.723	**<1e-3**	**<1e-3**	**<1e-3**	**0.005**	**<1e-3**	0.076	0.116
C2 vs C3	0.143	**<1e-3**	0.725	**0.002**	**0.004**	**0.002**	0.395	**0.044**	0.272	0.369
MDD 21, 62, 20										
C1 vs C2	0.418	0.320	0.868	0.710	0.674	0.623	0.649	0.527	0.053	0.060
C1 vs C3	0.182	**0.029**	0.800	0.062	0.100	0.079	0.126	0.182	**0.011**	**0.033**
C2 vs C3	0.611	0.368	0.746	0.697	0.842	0.730	0.746	0.778	0.861	0.675
						weekends				
	T_{in}	$Z_{\text{out},w}$	$T_{\text{out},d}$	M_{out}	S_R	M	$S1$	$S2$	Fretest1	Fretest2
HC 36, 53, 26										
C1 vs C2	0.656	0.093	0.587	0.201	0.356	0.278	0.317	0.215	0.661	0.157
C1 vs C3	0.597	**<1e-3**	**0.027**	**<1e-3**	**<1e-3**	**0.002**	**0.024**	**0.003**	0.553	0.082
C2 vs C3	0.688	0.109	**0.048**	0.068	0.078	0.100	0.471	0.106	0.767	0.480
MDD 18, 55, 22										
C1 vs C2	0.558	0.542	0.308	0.416	0.471	0.484	0.340	0.197	0.348	0.603
C1 vs C3	0.490	0.069	0.424	0.143	0.197	0.186	0.200	0.087	0.075	0.099
C2 vs C3	0.716	0.580	0.447	0.553	0.656	0.640	0.801	0.588	0.553	0.359

TABLE 10

Comparisons of activity distrbutions in different BMI groups, where D1 and D2 denoting lean and obese people, respectively. The p-values of the proposed test statistics, the generalized edge-count tests ($S1$, $S2$) and Fréchet tests (Fretest1, Fretest2) are presented for different comparisons (bold for those <0.05).

				weekdays						
	T_{in}	$Z_{\mathrm{out},w}$	$T_{\mathrm{out},d}$	M_{out}	S_R	M	$S1$	$S2$	Fretest1	Fretest2
HC 32, 48										
D1 vs D2	0.803	0.920	0.242	0.354	0.258	0.485	0.796	0.726	0.533	0.395
OTHER 43, 75										
D1 vs D2	0.793	**0.011**	0.128	**0.020**	**0.034**	**0.029**	0.439	0.451	0.826	0.857
				weekends						
	T_{in}	$Z_{\mathrm{out},w}$	$T_{\mathrm{out},d}$	M_{out}	S_R	M	$S1$	$S2$	Fretest1	Fretest2
HC 30, 49										
D1 vs D2	0.631	0.332	0.401	0.444	0.542	0.544	0.595	0.424	0.257	0.163
OTHER 40, 74										
D1 vs D2	0.692	0.105	0.121	0.142	0.204	0.208	0.464	0.680	0.504	0.657

E.2 15-MST with the 2-Wasserstein distance as the similarity graph

Table 11

Comparisons of activity distributions among the controls, MDD, BPI and BPII patients. The p-values are presented for the proposed test statistics, the generalized edge-count tests ($S1$, $S2$) and Fréchet tests (Fretest1, Fretest2) (bold for those <0.05).

	T_{in}	$Z_{\text{out},w}$	$T_{\text{out},d}$	M_{out}	S_R	M	$S1$	$S2$	Fretest1	Fretest2
weekdays, $n_1 = 117,\ n_2 = 106,\ n_3 = 26,\ n_4 = 32.$										
HC vs MDD	0.801	0.596	0.231	0.341	0.597	0.482	0.644	0.158	0.388	0.309
HC vs BPII	0.773	0.248	0.089	0.148	0.237	0.219	0.706	0.057	0.657	0.226
HC vs BPI	0.759	0.056	0.886	0.055	0.222	0.081	0.781	0.623	0.257	0.103
MDD vs BPII	0.540	0.451	0.055	0.064	0.059	0.094	0.574	0.089	0.717	0.762
MDD vs BPI	0.787	0.058	0.755	0.107	0.333	0.154	0.746	0.778	0.696	0.482
BPII vs BPI	0.354	0.051	0.147	0.051	0.058	0.059	0.113	0.189	0.122	0.126
weekends, $n_1 = 116,\ n_2 = 98,\ n_3 = 30,\ n_4 = 33.$										
	T_{in}	$Z_{\text{out},w}$	$T_{\text{out},d}$	M_{out}	S_R	M	$S1$	$S2$	Fretest1	Fretest2
HC vs MDD	0.329	0.390	0.392	0.468	0.513	0.443	0.459	0.699	0.051	0.059
HC vs BPII	0.754	0.442	0.414	0.502	0.718	0.650	0.481	0.393	0.538	0.509
HC vs BPI	0.508	**0.001**	**0.031**	**0.001**	**0.001**	**0.002**	**0.007**	**0.005**	**0.013**	**0.002**
MDD vs BPII	0.319	0.348	0.848	0.711	0.609	0.499	0.545	0.748	0.276	0.317
MDD vs BPI	0.620	**0.038**	0.218	0.073	0.122	0.107	0.072	0.078	0.482	0.288
BPII vs BPI	0.768	< 1**e-3**	0.168	**0.001**	**0.003**	**0.001**	**0.003**	**0.006**	**0.007**	**0.005**

Table 12

Comparisons of activity distrbutions in different age groups, where C1, C2 and C3 denoting young, middle-aged and older age groups. The p-values of the proposed test statistics, the generalized edge-count tests ($S1$, $S2$) and Fréchet tests (Fretest1, Fretest2) are presented for different comparisons (bold for those <0.05).

	T_{in}	$Z_{\text{out},w}$	$T_{\text{out},d}$	M_{out}	S_R	M	$S1$	$S2$	Fretest1	Fretest2
weekdays										
HC 36, 55, 25										
C1 vs C2	**0.023**	0.073	0.368	0.165	**0.020**	0.055	0.859	0.083	0.717	0.551
C1 vs C3	**0.001**	< 1**e-3**	0.577	< 1**e-3**	< 1**e-3**	< 1**e-3**	< 1**e-3**	< 1**e-3**	0.076	0.115
C2 vs C3	0.074	< 1**e-3**	0.726	**0.001**	**0.001**	**0.001**	**0.011**	**0.003**	0.268	0.366
MDD 21, 62, 20										
C1 vs C2	0.250	0.259	0.885	0.613	0.467	0.456	0.470	0.755	0.053	0.061
C1 vs C3	0.059	**0.006**	0.885	**0.010**	**0.006**	**0.013**	**0.005**	**0.020**	**0.011**	**0.033**
C2 vs C3	0.399	0.401	0.815	0.781	0.772	0.644	0.705	0.863	0.858	0.677
weekends										
	T_{in}	$Z_{\text{out},w}$	$T_{\text{out},d}$	M_{out}	S_R	M	$S1$	$S2$	Fretest1	Fretest2
HC 36, 53, 26										
C1 vs C2	0.455	0.059	0.145	0.091	0.089	0.129	0.189	0.307	0.670	0.158
C1 vs C3	**0.050**	**0.002**	**0.005**	**0.002**	< 1**e-3**	**0.003**	0.187	**0.021**	0.556	0.083
C2 vs C3	0.311	**0.042**	0.058	0.056	**0.026**	0.076	0.722	0.096	0.769	0.481
MDD 18, 55, 22										
C1 vs C2	0.181	0.395	0.295	0.395	0.285	0.289	0.837	0.452	0.350	0.607
C1 vs C3	0.370	**0.039**	0.116	0.068	0.057	0.096	**0.005**	**0.021**	0.071	0.095
C2 vs C3	0.624	0.369	0.865	0.762	0.856	0.776	0.906	0.699	0.545	0.359

TABLE 13

Comparisons of activity distrbutions in different BMI groups, where D1 and D2 denoting lean and obese people, respectively. The p-values of the proposed test statistics, the generalized edge-count tests ($S1$, $S2$) and Fréchet tests (Fretest1, Fretest2) are presented for different comparisons (bold for those <0.05).

				weekdays						
	T_{in}	$Z_{\text{out},w}$	$T_{\text{out},d}$	M_{out}	S_R	M	$S1$	$S2$	Fretest1	Fretest2
HC 32, 48										
D1 vs D2	0.843	0.717	0.649	0.767	0.874	0.891	0.886	0.826	0.531	0.394
OTHER 43, 75										
D1 vs D2	0.793	**0.006**	0.263	**0.011**	**0.043**	**0.016**	0.496	0.833	0.825	0.856
				weekends						
	T_{in}	$Z_{\text{out},w}$	$T_{\text{out},d}$	M_{out}	S_R	M	$S1$	$S2$	Fretest1	Fretest2
HC 30, 49										
D1 vs D2	0.609	0.861	0.496	0.628	0.480	0.689	0.672	0.340	0.256	0.167
OTHER 40, 74										
D1 vs D2	0.763	**0.031**	0.422	0.065	0.195	0.098	0.690	0.673	0.482	0.643

E.3 5-MST with the maximum mean discrepancy as the similarity graph

TABLE 14

Comparisons of activity distributions among the controls, MDD, BPI and BPII patients. The p-values are presented for the proposed test statistics, the generalized edge-count tests ($S1$, $S2$) and Fréchet tests (Fretest1, Fretest2) (bold for those <0.05).

	weekdays, $n_1 = 117,\ n_2 = 106,\ n_3 = 26,\ n_4 = 32.$									
	T_{in}	$Z_{\text{out},w}$	$T_{\text{out},d}$	M_{out}	S_R	M	$S1$	$S2$	Fretest1	Fretest2
HC vs MDD	0.730	0.810	0.603	0.714	0.629	0.774	0.718	0.504	0.382	0.307
HC vs BPII	0.735	0.062	0.744	0.133	0.302	0.182	0.466	0.322	0.656	0.230
HC vs BPI	0.644	0.205	0.774	0.435	0.634	0.507	0.768	0.758	0.259	0.103
MDD vs BPII	0.801	0.115	0.616	0.221	0.426	0.310	0.549	0.321	0.712	0.768
MDD vs BPI	0.706	0.343	0.738	0.608	0.782	0.684	0.761	0.759	0.700	0.483
BPII vs BPI	0.708	0.053	0.672	0.071	0.177	0.101	0.600	0.306	0.121	0.124
	weekends, $n_1 = 116,\ n_2 = 98,\ n_3 = 30,\ n_4 = 33.$									
	T_{in}	$Z_{\text{out},w}$	$T_{\text{out},d}$	M_{out}	S_R	M	$S1$	$S2$	Fretest1	Fretest2
HC vs MDD	0.511	0.912	0.733	0.824	0.447	0.696	0.614	0.743	0.062	0.060
HC vs BPII	0.593	0.593	0.716	0.755	0.745	0.726	0.636	0.316	0.553	0.523
HC vs BPI	0.456	0.066	0.190	0.087	0.119	0.112	0.144	0.093	**0.014**	**0.002**
MDD vs BPII	0.391	0.600	0.753	0.778	0.585	0.561	0.728	0.439	0.267	0.310
MDD vs BPI	0.689	0.243	0.257	0.236	0.378	0.329	0.296	0.143	0.485	0.284
BPII vs BPI	0.409	**0.021**	0.359	**0.041**	0.078	0.055	0.127	0.340	**0.007**	**0.005**

TABLE 15

Comparisons of activity distrbutions in different age groups, where C1, C2 and C3 denoting young, middle-aged and older age groups. The p-values of the proposed test statistics, the generalized edge-count tests ($S1$, $S2$) and Fréchet tests (Fretest1, Fretest2) are presented for different comparisons (bold for those <0.05).

	weekdays									
	T_{in}	$Z_{\text{out},w}$	$T_{\text{out},d}$	M_{out}	S_R	M	$S1$	$S2$	Fretest1	Fretest2
HC 36, 55, 25										
C1 vs C2	0.409	**0.036**	0.725	0.082	0.156	0.106	0.708	0.495	0.713	0.550
C1 vs C3	**0.029**	**< 1e-3**	0.270	**< 1e-3**	**< 1e-3**	**< 1e-3**	**0.002**	**< 1e-3**	0.076	0.115
C2 vs C3	0.206	**0.002**	0.669	**0.005**	**0.010**	**0.006**	0.154	0.062	0.272	0.369
MDD 21, 62, 20										
C1 vs C2	0.396	0.121	0.801	0.281	0.313	0.278	0.425	0.332	0.053	0.060
C1 vs C3	0.206	**0.002**	0.742	**0.005**	**0.012**	**0.006**	0.191	0.230	**0.011**	**0.032**
C2 vs C3	0.744	0.183	0.624	0.380	0.602	0.482	0.779	0.793	0.863	0.681
	weekends									
	T_{in}	$Z_{\text{out},w}$	$T_{\text{out},d}$	M_{out}	S_R	M	$S1$	$S2$	Fretest1	Fretest2
HC 36, 53, 26										
C1 vs C2	0.696	**0.012**	0.721	**0.027**	0.079	**0.037**	0.291	0.353	0.657	0.157
C1 vs C3	0.612	**0.002**	0.538	**0.004**	**0.011**	**0.005**	**0.014**	**0.009**	0.554	0.081
C2 vs C3	0.664	0.091	0.492	0.162	0.278	0.223	0.259	0.154	0.771	0.482
MDD 18, 55, 22										
C1 vs C2	0.480	0.270	0.413	0.344	0.481	0.406	0.468	0.237	0.358	0.604
C1 vs C3	0.646	**0.041**	0.721	0.091	0.220	0.126	0.094	0.104	0.073	0.096
C2 vs C3	0.715	0.379	0.236	0.281	0.418	0.377	0.754	0.700	0.546	0.353

TABLE 16

Comparisons of activity distrbutions in different BMI groups, where D1 and D2 denoting lean and obese people, respectively. The p-values of the proposed test statistics, the generalized edge-count tests ($S1$, $S2$) and Fréchet tests (Fretest1, Fretest2) are presented for different comparisons (bold for those <0.05).

					weekdays					
	T_{in}	$Z_{\text{out},w}$	$T_{\text{out},d}$	M_{out}	S_R	M	$S1$	$S2$	Fretest1	Fretest2
HC 32, 48										
D1 vs D2	0.770	0.798	0.501	0.614	0.624	0.716	0.774	0.713	0.533	0.400
D1 vs D2	0.794	**0.008**	0.284	**0.015**	**0.038**	**0.021**	0.319	0.235	0.826	0.861
					weekends					
	T_{in}	$Z_{\text{out},w}$	$T_{\text{out},d}$	M_{out}	S_R	M	$S1$	$S2$	Fretest1	Fretest2
HC 30, 49										
D1 vs D2	0.630	0.467	0.559	0.584	0.658	0.635	0.697	0.681	0.258	0.168
D1 vs D2	0.697	0.134	0.592	0.293	0.490	0.404	0.551	0.455	0.531	0.697

SUPPLEMENT F: SENSITIVITY ANALYSES OF RANDOM SUBSETTING L DAYS

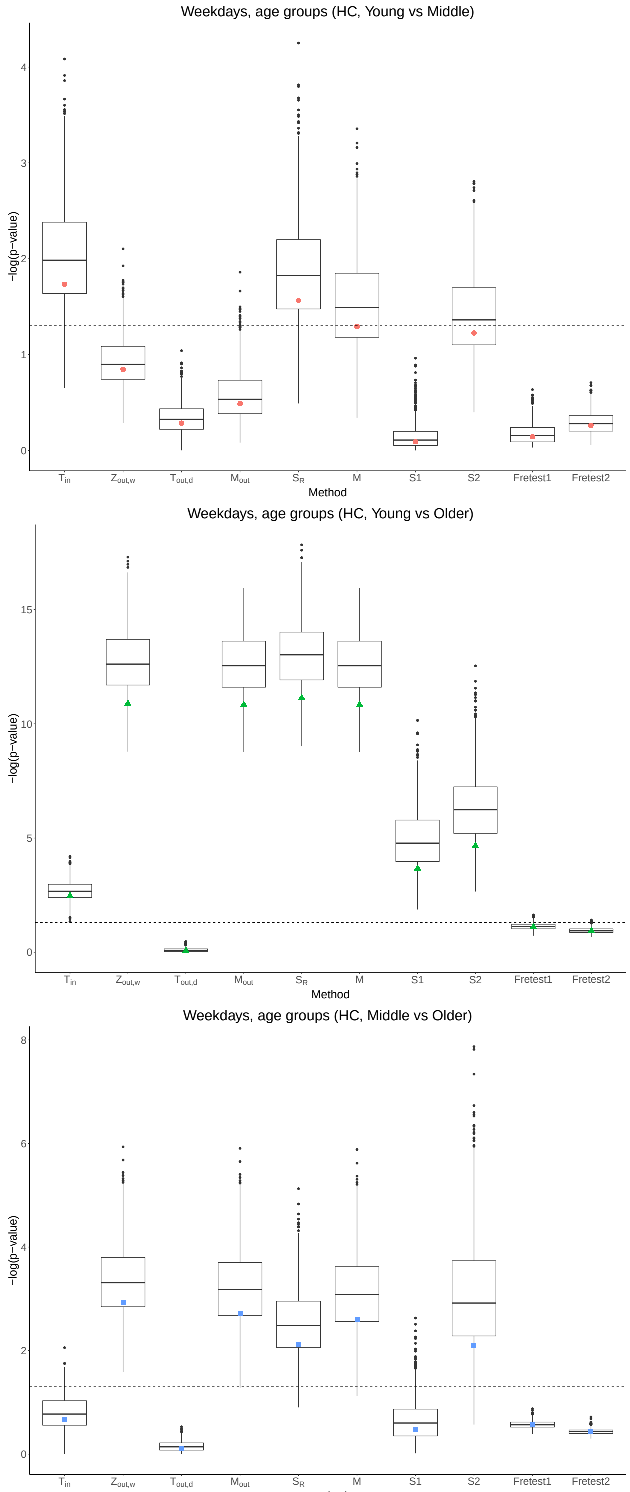

FIG 4. *Boxplots of -log(p-value) for age group comparison over 1000 random subsetting l days of data to be included in the analysis. Results are shown for each test statistics using weekday data. Overall the pvalues were shown to be robust towards choices of days.*

REFERENCES

Chen, L. H. and Shao, Q.-M. (2005). Stein's method for normal approximation. *An introduction to Stein's method* **4** 1–59.