

pldist: Ecological Dissimilarities for Paired and Longitudinal Microbiome Association Analysis

Supplementary Materials

Anna M. Plantinga, Jun Chen, Robert R. Jenq, and Michael C. Wu

Supplementary Materials include tables describing the data transformations and dissimilarities, a proof that unweighted paired UniFrac is a proper distance, a counter-example demonstrating that generalized paired UniFrac is not a proper distance due to failure of the triangle inequality, and additional simulation results.

1 Summary of Data Transformations

In this section, we briefly summarize the binary and quantitative transformations for paired and longitudinal data (Table 1). The key difference between the paired version and the longitudinal version, even with $q = 2$ time points, is that the paired version takes direction of change into account (i.e., losses and gains of taxon presence or abundance are treated differently), whereas the longitudinal version looks only at magnitude of change between time points (a measure of variability of each taxon in the microbiota over time). Hence in the paired version, when comparing two subjects, if a taxon is lost between time points in one subject and gained in the other, that is the maximum difference possible between subjects for that taxon. In contrast, in the longitudinal version with $q = 2$, there is no difference between subjects with respect to the variability of that taxon because in each case, it was present at one time point and absent at the other.

For the longitudinal transformations, the time points (t_1, \dots, t_q) may vary between subjects in both value and quantity, so a more accurate representation would be $(t_1^A, \dots, t_{q_A}^A)$. For clarity this notation has been suppressed. The range of the longitudinally transformed data depends on the assumptions on t_k . For example, if we assume that $(t_k - t_{k-1}) \geq 1$ for all subjects and time points (i.e., the unit of time is at least as small as the two measurements closest in time), then the transformed data lie in $[0, 1]$ for both binary and quantitative versions.

Note that in the quantitative case, the normalized measures of change d_k^X only consider relative changes, e.g., a 50% decrease in relative abundance, not absolute changes in relative abundance, e.g., a decrease of 0.02 in relative abundance. The unnormalized measures of change consider absolute changes in relative abundance.

The centered log-ratio (CLR) transformation accounts for the compositionality of the

Table 1: Basic data transformations. “Norm” indicates whether the change in abundance is normalized to overall abundance of that taxon, and “CLR” indicates whether the CLR transformation is applied.

Times	Type	Norm?	CLR?	Transformation
2	Binary	N/A	N/A	$d_k^A(t_1, t_2) = \frac{1}{2} \cdot \left[I(p_k^{(A,t_1)} > 0) - I(p_k^{(A,t_2)} > 0) \right]$
2	Quant.	Yes	No	$d_k^A(t_1, t_2) = \frac{1}{2} \cdot \frac{p_k^{(A,t_2)} - p_k^{(A,t_1)}}{p_k^{(A,t_2)} + p_k^{(A,t_1)}}$
2	Quant.	Yes	Yes	$d_k^A(t_1, t_2) = \frac{1}{2} \cdot \frac{\text{clr}(p_k^{(A,t_2)}) - \text{clr}(p_k^{(A,t_1)})}{ \text{clr}(p_k^{(A,t_2)}) + \text{clr}(p_k^{(A,t_1)}) }$
2	Quant.	No	No	$d_k^A(t_1, t_2) = \frac{1}{2} \left(p_k^{(A,t_2)} - p_k^{(A,t_1)} \right)$
2	Quant.	No	Yes	$d_k^A(t_1, t_2) = \frac{1}{2} \left(\text{clr}(p_k^{(A,t_2)}) - \text{clr}(p_k^{(A,t_1)}) \right)$
≥ 2	Binary	N/A	N/A	$d_k^A(t_1, \dots, t_q) = \frac{1}{q-1} \sum_{l=1}^{q-1} \left(\frac{1}{t_{l+1} - t_l} \right) \cdot \left I(p_k^{A,t_{l+1}} > 0) - I(p_k^{A,t_l} > 0) \right $
≥ 2	Quant.	Yes	No	$d_k^A(t_1, \dots, t_q) = \frac{1}{q-1} \sum_{l=1}^{q-1} \left(\frac{1}{t_{l+1} - t_l} \right) \cdot \left \frac{p_k^{A,t_{l+1}} - p_k^{A,t_l}}{p_k^{A,t_{l+1}} + p_k^{A,t_l}} \right $
≥ 2	Quant.	Yes	Yes	$d_k^A(t_1, \dots, t_q) = \frac{1}{q-1} \sum_{l=1}^{q-1} \left(\frac{1}{t_{l+1} - t_l} \right) \cdot \frac{ \text{clr}(p_k^{A,t_{l+1}}) - \text{clr}(p_k^{A,t_l}) }{ \text{clr}(p_k^{A,t_{l+1}}) + \text{clr}(p_k^{A,t_l}) }$
≥ 2	Quant.	No	No	$d_k^A(t_1, \dots, t_q) = \frac{1}{q-1} \sum_{l=1}^{q-1} \left(\frac{1}{t_{l+1} - t_l} \right) \cdot \left p_k^{A,t_{l+1}} - p_k^{A,t_l} \right $
≥ 2	Quant.	No	Yes	$d_k^A(t_1, \dots, t_q) = \frac{1}{q-1} \sum_{l=1}^{q-1} \left(\frac{1}{t_{l+1} - t_l} \right) \cdot \left \text{clr}(p_k^{A,t_{l+1}}) - \text{clr}(p_k^{A,t_l}) \right $

data. For taxon proportion p_k^A for subject A , $\text{clr}(p_k^A)$ is defined as

$$\text{clr}(p_k^A) = \log \left(\frac{p_k^A}{\text{gm}(p^A)} \right)$$

where $\text{gm}(p_A)$ is the geometric mean of all abundances for that subject. The log transformation requires that all taxon proportions be nonzero. A variety of methods exist to account for zeros; replacing zeros with a small constant called a pseudocount is a common approach and the method used here. The differences $\text{clr}(p_k^{(A,t_2)}) - \text{clr}(p_k^{(A,t_1)})$ can be understood as log-ratios of (normalized) abundances,

$$\begin{aligned} \text{clr}(p_k^{(A,t_2)}) - \text{clr}(p_k^{(A,t_1)}) &= \log \left(\frac{p_k^{(A,t_2)}}{\text{gm}(p^{(A,t_2)})} \right) - \log \left(\frac{p_k^{(A,t_1)}}{\text{gm}(p^{(A,t_1)})} \right) \\ &= \log \left(\frac{p_k^{(A,t_2)} / \text{gm}(p^{(A,t_2)})}{p_k^{(A,t_1)} / \text{gm}(p^{(A,t_1)})} \right). \end{aligned}$$

2 Summary of Distances and Dissimilarities

In this section, we summarize and compare several distances and dissimilarities for single timepoint (S), paired (P), and longitudinal (L) samples (Table 2).

In binary versions of each metric, $|A|$ and $|B|$ are the numbers of species in compared sites (hence $|A \cap B|$ is the number of species in both sites and $|A \cup B|$ is the number at either site). The total number of OTUs (columns) is m .

In quantitative versions of the metrics, p_k^A refers to the proportion of taxon k for subject A , \bar{p}_k^A is the average proportion of taxon k for subject A across time points, and d_k^A is the transformed data using the appropriate transformation from Table 1. Many of the metrics are traditionally calculated on raw counts rather than proportions, but because the data transformations are defined for proportions, we have used proportions throughout for consistency (and simplified the metrics accordingly when possible).

For the UniFrac family, “Binary” corresponds to unweighted UniFrac and “Quantitative” corresponds to generalized UniFrac. We are disregarding weighted UniFrac. Unlike the single timepoint UniFrac metrics, longitudinal generalized UniFrac with $\gamma = 1$ is not in general equivalent to longitudinal weighted UniFrac because the abundance weights depend on \bar{p}_k^X and the measures of change on d_k^X . Specifically, longitudinal generalized UniFrac is

$$D_{AB}^{(\gamma)} = \frac{\sum_i b_i (\bar{p}_i^A + \bar{p}_i^B)^\gamma |d_i^A - d_i^B|}{\sum_i b_i (\bar{p}_i^A + \bar{p}_i^B)^\gamma}$$

whereas the longitudinal version of weighted UniFrac would be

$$D_{AB} = \frac{\sum_i b_i |d_i^A - d_i^B|}{\sum_i b_i (d_i^A + d_i^B)}$$

in which the denominator depends on the changes in composition d_i rather than the average taxon abundance \bar{p}_i .

Unless otherwise specified, the \sum_k notation is shorthand for $\sum_{k=1}^m$ (summing over OTUs/taxa). \sum_i is used instead for the UniFrac metrics because the sum is over branches, not tips (OTUs).

Table 2: Summary of distances and dissimilarities. The normalized quantitative definitions of d_k are used for UniFrac dissimilarities; the unnormalized quantitative definitions of d_k are used for all other dissimilarities.

Metric	Design	Binary (Qualitative)	Quantitative
Bray-Curtis	S	$D_{AB} = 1 - \frac{2 A \cap B }{ A + B }$	$D_{AB} = \frac{1}{2} \sum_k p_k^A - p_k^B $
	P, L	$D_{AB} = \frac{1}{m} \sum_k d_k^A - d_k^B $	$D_{AB} = \frac{1}{m} \sum_k d_k^A - d_k^B $
Jaccard	S	$D_{AB} = 1 - \frac{ A \cap B }{ A \cup B }$	$D_{AB} = 1 - \frac{\sum_k \min(p_k^A, p_k^B)}{\sum_k \max(p_k^A, p_k^B)}$
	P	$D_{AB} = 1 - \frac{\sum_k I(d_k^A = d_k^B) I(d_k^A \neq 0)}{\sum_k [I(d_k^A \neq 0) \text{ OR } d_k^B \neq 0]}$	$D_{AB} = 1 - \frac{\sum_k \min(d_k^A , d_k^B) I\{\text{sgn}(d_k^A) = \text{sgn}(d_k^B)\}}{\sum_k \max(d_k^A , d_k^B)}$
	L	$D_{AB} = 1 - \frac{\sum_k \min(d_k^A, d_k^B)}{\sum_k \max(d_k^A, d_k^B)}$	$D_{AB} = 1 - \frac{\sum_k \min(d_k^A, d_k^B)}{\sum_k \max(d_k^A, d_k^B)}$
Kulczynski	S	$D_{AB} = 1 - \frac{1}{2} \left(\frac{ A \cap B }{ A } + \frac{ A \cap B }{ B } \right)$	$D_{AB} = 1 - \sum_k \min(p_k^A, p_k^B)$
	P	$D_{AB} = 1 - \frac{1}{m} \sum_k I(d_k^A = d_k^B)$	$D_{AB} = 1 - \frac{1}{2} \sum_k \left(\frac{1}{\sum_k d_k^A } + \frac{1}{\sum_k d_k^B } \right) \times \min(d_k^A , d_k^B) I\{\text{sgn}(d_k^A) = \text{sgn}(d_k^B)\}$

(continued on next page)

Table 2 – continued from previous page

Metric	Design	Binary (Qualitative)	Quantitative
	L	$D_{AB} = 1 - \frac{1}{2} \sum_k \left(\frac{1}{\sum_k d_k^A} + \frac{1}{\sum_k d_k^B} \right) \min(d_k^A, d_k^B)$	$D_{AB} = 1 - \frac{1}{2} \sum_k \left(\frac{1}{\sum_k d_k^A} + \frac{1}{\sum_k d_k^B} \right) \min(d_k^A, d_k^B)$
Gower	S	$D_{AB} = \frac{ A + B -2 A \cap B }{m}$	$D_{AB} = \frac{1}{m} \sum_k \frac{ p_k^A - p_k^B }{\max p_k - \min p_k}$
	P, L	$D_{AB} = \frac{1}{m} \sum_k \frac{ d_k^A - d_k^B }{\max d_k - \min d_k}$	$D_{AB} = \frac{1}{m} \sum_k \frac{ d_k^A - d_k^B }{\max d_k - \min d_k}$
UniFrac	S	$D_{AB} = \frac{\sum_i b_i I(p_i^A > 0) - I(p_i^B > 0) }{\sum_i b_i}$	$D_{AB}^{(\gamma)} = \frac{\sum_i b_i (p_i^A + p_i^B)^\gamma \left \frac{p_i^A - p_i^B}{p_i^A + p_i^B} \right }{\sum_i b_i (p_i^A + p_i^B)^\gamma}$
	P, L	$D_{AB} = \frac{\sum_i b_i d_i^A - d_i^B }{\sum_i b_i}$	$D_{AB}^{(\gamma)} = \frac{\sum_i b_i (\bar{p}_i^A + \bar{p}_i^B)^\gamma d_i^A - d_i^B }{\sum_i b_i (\bar{p}_i^A + \bar{p}_i^B)^\gamma}$

3 Proofs of Distance

Unweighted Paired UniFrac is a Distance

To prove that the unweighted paired UniFrac measure is a true distance metric, we need to show that:

1. $D_{XY} \geq 0$ (nonnegativity)
2. $D_{XY} = 0$ iff $X = Y$ (identity of indiscernables)
3. $D_{XY} = D_{YX}$ (symmetry)
4. $D_{XZ} \leq D_{XY} + D_{YZ}$ (triangle inequality)

Nonnegativity: By definition, branch lengths $b_k > 0$. The remaining term in the metric is an absolute value term ($|d_k^X - d_k^Y| \geq 0$). Therefore, $D_{XY} \geq 0$.

Identity: In order to demonstrate this item, we must first define what we mean by “indiscernables.”

Let X and Y be two microbial communities measured at two time points each, so we have $X(t_1)$ and $X(t_2)$, and similarly, $Y(t_1)$ and $Y(t_2)$. The mapping from $X_k(t_1)$ and $X_k(t_2)$ to d_k^X is not 1-1. However, the quantity of interest for this metric is not the microbial community itself, it is changes in microbial communities. Therefore, in this case the relevant “indiscernables” are d^X and d^Y rather than X and Y . That is, we must demonstrate that $D_{XY} = 0$ iff $d^X = d^Y$, regardless of the original communities $X(t_1)$, $X(t_2)$, $Y(t_1)$, and $Y(t_2)$,

Suppose $d_k^X = d_k^Y$ for all taxa k . Then $d_k^X - d_k^Y = 0 \forall k$, and so $\sum_{k=1}^p b_k |d_k^X - d_k^Y|/2 = 0$. Therefore $X = Y \implies D_{XY} = 0$.

Now suppose $D_{XY} = 0$. Because each term is nonnegative and branch lengths b_k are strictly positive ($b_k > 0 \forall k$), this implies that $|d_k^X - d_k^Y| = 0 \forall k$. This only holds if $d_k^X = d_k^Y \forall k$. Therefore $D_{XY} = 0 \implies d^X = d^Y$.

Symmetry: Here, we note that

$$\begin{aligned} D_{XY} &= \frac{\sum_{k=1}^p b_k |d_k^X - d_k^Y|/2}{\sum_{k=1}^p b_k} = \frac{\sum_{k=1}^p b_k |(-1) \times (d_k^X - d_k^Y)|/2}{\sum_{k=1}^p b_k} \\ &= \frac{\sum_{k=1}^p b_k |d_k^Y - d_k^X|/2}{\sum_{k=1}^p b_k} = D_{YX} \end{aligned}$$

Triangle Inequality: Using the triangle inequality for the absolute value ($|x + y| \leq |x| + |y|$):

$$\begin{aligned}
 2 \sum_{k=1}^p b_k \times (D_{XY} + D_{YZ}) &= \sum_{k=1}^p b_k |d_k^X - d_k^Y| + \sum_{k=1}^p b_k |d_k^Y - d_k^Z| \\
 &= \sum_{k=1}^p b_k [|d_k^X - d_k^Y| + |d_k^Y - d_k^Z|] \\
 &\geq \sum_{k=1}^p b_k (|d_k^X - d_k^Y| + |d_k^Y - d_k^Z|) \\
 &= 2 \sum_{k=1}^p b_k \times D_{XZ}
 \end{aligned}$$

that is, $D_{XY} + D_{YZ} \geq D_{XZ}$.

Generalized Paired UniFrac is Not a Distance

While the generalized paired UniFrac dissimilarity trivially satisfies nonnegativity, identity of indistinguishable elements, and symmetry, it does not satisfy the triangle inequality. We demonstrate this by counterexample. Consider a phylogenetic tree with three taxa, as displayed in Figure 1. Suppose the relative abundance of these three taxa has been observed for three subjects at two time points each, as shown in Table 3. The resulting weighted paired UniFrac dissimilarity matrix is shown in Table 4. From the table, $D_{12} + D_{23} = 0.108 + 0.132 = 0.239 < 0.278 = D_{13}$. That is, this dissimilarity does not satisfy the triangle inequality because it does not hold that $D_{AB} + D_{BC} \geq D_{AC}$ for every combination of subjects A, B, C . Therefore, the generalized paired UniFrac dissimilarity is not a proper distance.

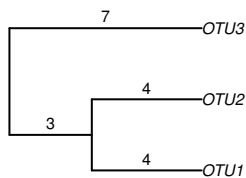


Figure 1: Simple phylogenetic tree with three taxa. Numbers indicate branch lengths.

Table 3: Relative abundances of taxa in counterexample.

Subject	Time	OTU1	OTU2	OTU3
Subj1	1	0.40	0.10	0.50
Subj1	2	0.05	0.65	0.30
Subj2	1	0.30	0.05	0.65
Subj2	2	0.25	0.25	0.5
Subj3	1	0.55	0.2	0.25
Subj3	2	0.45	0.15	0.4

Table 4: Generalized paired UniFrac dissimilarity for taxon abundances in Table 3.

	Subj1	Subj2	Subj3
Subj1	0	0.108	0.278
Subj2	0.108	0	0.132
Subj3	0.278	0.132	0

4 Additional Simulation Results

Tables 5 and 6 provide empirical Type 1 errors for kernels based on each new distance or dissimilarity. Since kernel choice does not affect the size of kernel machine regression-based tests, these results serve mainly as a basic check that none of the new kernels are systematically exhibiting concerning behavior.

Figures 2-7 show results from additional power simulations with $n = 100$. Figure 2 is similar to the figure in the main text, but includes more dissimilarities and is in color for ease of reading and comparison to the other figures. The figures are:

2. Continuous outcome, 2 time points.
3. Continuous outcome, 4 time points.
4. Dichotomous outcome, 2 time points.
5. Dichotomous outcome, 4 time points.
6. Time-to-event outcome, 2 time points.
7. Time-to-event outcome, 4 time points.

In each plot, the kernels are:

- K_{omni} : omnibus test
- K_1 : generalized paired/longitudinal UniFrac kernel with $\gamma = 1$
- $K_{0.5}$: generalized paired/longitudinal UniFrac kernel with $\gamma = 0.5$
- K_U : unweighted paired/longitudinal UniFrac kernel
- K_{JQ} : quantitative paired/longitudinal Jaccard kernel

- K_{JB} : binary paired/longitudinal Jaccard kernel
- K_{JC} : quantitative paired/longitudinal Jaccard kernel with CLR-transformed data
- K_{1C} : generalized paired/longitudinal UniFrac kernel with $\gamma = 1$ and CLR-transformed data
- $K_{0.5C}$: generalized paired/longitudinal UniFrac kernel with $\gamma = 0.5$ and CLR-transformed data

In each figure, the outcome is associated with:

- Panel A: Change in presence of a rare cluster.
- Panel B: Change in presence of 60 randomly selected taxa.
- Panel C: Change in abundance of a moderately common cluster.
- Panel D: Change in abundance of the 10 most abundant taxa.
- Panel E: Change in CLR-transformed abundance of a moderately common cluster.
- Panel F: Change in CLR-transformed abundance of the 10 most abundant taxa.

The same general pattern of relative performance holds across outcome types, although power is highest for continuous outcomes (in the plots, notice that the range of true effect sizes considered is smaller for continuous outcomes). For the paired settings in Panel D (change in abundance of top 10), the quantitative Jaccard kernel has highest power; in contrast, for longitudinal settings in Panel D, the weighted UniFrac kernel has highest power. This is likely because the formulation of the weighted UniFrac kernel allows the weight for each OTU to depend on its abundance, and so by choosing $\gamma = 1$, the most common OTUs are very highly weighted in the distance metric.

Figures 8 and 9 display power based on each nonphylogenetic kernel for the two nonphylogenetic settings. Panel A displays power for detecting changes in presence of 60 randomly selected taxa, Panel B displays power for detecting changes in abundance of the 10 most common taxa, and Panel C displays power for detecting changes in CLR-transformed abundance of the 10 most common taxa. As expected, generally tests based on the binary metrics (dashed lines) have higher power when changes in presence of rare taxa are associated with the outcome (Panel A); tests based on the quantitative metrics (solid lines) tend to have higher power when changes in abundance of common taxa are associated with the outcome (Panel B); and tests based on the CLR-transformed metrics (dotted lines) have higher power when changes in CLR-transformed abundance are associated with the outcome. The differences in power across nonphylogenetic metrics is smallest in Panel A of the longitudinal setting (changes in rare taxon presence) and largest in Panels B and C of the longitudinal setting (changes in abundance). The differences in how metrics measure dissimilarity is most notable in the latter setting (and in particular, metrics such as the quantitative Gower distance divide changes for each taxon by the maximum change in that taxon across all subjects, so changes in abundant taxa for each subject tend to be similar across subjects after

dividing by max - min for each taxon. While different metrics have highest power in different situations, the Bray-Curtis family (quantitative, binary, and CLR-transformed) tend to perform quite well across settings, as does the quantitative Jaccard metric across quantitative settings. Interestingly, CLR-transformed Jaccard tends to have lower power than quantitative Jaccard even when the true association is with CLR-transformed abundances.

Table 5: Empirical size with nominal $\alpha = 0.05$. Type is the outcome type, n is the sample size, and times is the number of observed time points, where the paired version was used for 2 time points and the longitudinal version for 4 time points. Each kernel's subscript includes first an indicator of which distance was used (B = Bray-Curtis, G = Gower, K = Kulczynski, and J = Jaccard) and then an indicator of whether the binary/qualitative (B), untransformed quantitative (Q), or CLR-transformed quantitative (C) version was used. K_U is the unweighted UniFrac, $K_{0.5}$ the generalized UniFrac with $\gamma = 0.5$, and K_1 the generalized UniFrac dissimilarity with $\gamma = 1$.

type	n	times	K_{BB}	K_{BQ}	K_{BC}	K_{GB}	K_{GQ}	K_{GC}	K_{KB}	K_{KQ}	K_{KC}	K_{JB}	K_{JQ}	K_{JC}	K_U	$K_{0.5}$	K_1	$K_{0.5C}$	K_{1C}	K_{omni}		
Cont.	50	2	0.046	0.046	0.050	0.049	0.047	0.049	0.045	0.046	0.051	0.050	0.046	0.052	0.052	0.050	0.043	0.047	0.049	0.048		
		4	0.042	0.058	0.046	0.042	0.049	0.044	0.047	0.053	0.046	0.043	0.043	0.052	0.045	0.045	0.053	0.047	0.054	0.045		
		100	2	0.041	0.055	0.050	0.043	0.052	0.055	0.050	0.053	0.046	0.042	0.046	0.049	0.052	0.054	0.051	0.061	0.058	0.050	
			4	0.052	0.056	0.054	0.055	0.062	0.054	0.055	0.058	0.057	0.057	0.057	0.056	0.054	0.048	0.048	0.052	0.054	0.050	
	200	2	0.046	0.041	0.044	0.048	0.047	0.048	0.048	0.051	0.047	0.049	0.051	0.042	0.047	0.054	0.052	0.051	0.050	0.047	0.049	
		4	0.041	0.045	0.042	0.042	0.039	0.040	0.039	0.050	0.047	0.039	0.038	0.047	0.039	0.043	0.048	0.050	0.046	0.045	0.047	
		50	2	0.050	0.042	0.045	0.046	0.039	0.045	0.049	0.042	0.043	0.043	0.043	0.036	0.040	0.042	0.045	0.049	0.043	0.041	0.045
			4	0.043	0.042	0.044	0.046	0.042	0.047	0.045	0.041	0.041	0.044	0.043	0.037	0.043	0.042	0.044	0.046	0.040	0.042	0.048
	100	2	0.046	0.0395	0.041	0.043	0.043	0.039	0.047	0.045	0.041	0.041	0.044	0.042	0.044	0.050	0.043	0.046	0.045	0.047	0.045	
		4	0.046	0.041	0.053	0.046	0.046	0.053	0.041	0.042	0.051	0.045	0.045	0.045	0.051	0.046	0.049	0.057	0.048	0.050	0.047	
		200	2	0.044	0.044	0.053	0.043	0.044	0.052	0.048	0.050	0.053	0.045	0.045	0.049	0.048	0.051	0.052	0.054	0.045	0.051	0.052
			4	0.047	0.040	0.049	0.046	0.043	0.048	0.051	0.040	0.053	0.049	0.049	0.045	0.051	0.054	0.046	0.048	0.055	0.046	0.043
Surv.	50	2	0.051	0.056	0.049	0.053	0.055	0.051	0.051	0.053	0.049	0.050	0.050	0.054	0.052	0.054	0.048	0.051	0.054	0.053	0.051	
		4	0.051	0.054	0.050	0.053	0.050	0.049	0.052	0.056	0.050	0.055	0.055	0.054	0.051	0.052	0.057	0.051	0.052	0.046	0.053	
		100	2	0.043	0.045	0.048	0.045	0.046	0.043	0.048	0.047	0.045	0.044	0.044	0.050	0.046	0.047	0.052	0.047	0.045	0.055	0.047
			4	0.045	0.050	0.051	0.050	0.046	0.051	0.044	0.049	0.046	0.047	0.047	0.050	0.049	0.043	0.042	0.040	0.041	0.047	0.045
	200	2	0.047	0.050	0.047	0.046	0.054	0.045	0.049	0.051	0.051	0.051	0.055	0.053	0.050	0.057	0.051	0.047	0.053	0.049	0.048	
		4	0.055	0.049	0.049	0.056	0.055	0.053	0.055	0.050	0.050	0.050	0.051	0.048	0.050	0.051	0.051	0.049	0.052	0.046	0.051	

Table 6: Empirical size with nominal $\alpha = 0.01$. Type is the outcome type, n is the sample size, and times is the number of observed time points, where the paired version was used for 2 time points and the longitudinal version for 4 time points. Each kernel's subscript includes first an indicator of which distance was used (B = Bray-Curtis, G = Gower, K = Kulczynski, and J = Jaccard) and then an indicator of whether the binary/qualitative (B), untransformed quantitative (Q), or CLR-transformed quantitative (C) version was used. K_U is the unweighted UniFrac, $K_{0.5}$ the generalized UniFrac with $\gamma = 0.5$, and K_1 the generalized UniFrac dissimilarity with $\gamma = 1$.

type	n	times	K_{BB}	K_{BQ}	K_{BC}	K_{GB}	K_{GQ}	K_{GC}	K_{KB}	K_{KQ}	K_{KC}	K_{JB}	K_{JQ}	K_{JC}	K_U	$K_{0.5}$	K_1	$K_{0.5C}$	K_{1C}	K_{omni}	
Cont.	50	2	0.006	0.009	0.009	0.006	0.008	0.013	0.007	0.009	0.011	0.006	0.007	0.009	0.005	0.008	0.009	0.007	0.008	0.006	
		4	0.010	0.011	0.011	0.010	0.014	0.011	0.010	0.011	0.009	0.010	0.011	0.009	0.011	0.011	0.011	0.011	0.009	0.012	0.010
	100	2	0.009	0.011	0.012	0.010	0.010	0.011	0.008	0.008	0.009	0.008	0.009	0.011	0.010	0.012	0.008	0.008	0.010	0.010	0.008
		4	0.008	0.013	0.009	0.010	0.009	0.008	0.008	0.008	0.016	0.008	0.007	0.016	0.008	0.006	0.012	0.010	0.011	0.012	0.011
	200	2	0.009	0.009	0.010	0.009	0.009	0.009	0.009	0.011	0.008	0.011	0.010	0.008	0.010	0.012	0.012	0.009	0.013	0.013	0.010
		4	0.005	0.010	0.009	0.006	0.008	0.009	0.009	0.004	0.008	0.009	0.005	0.009	0.008	0.004	0.010	0.011	0.010	0.009	0.008
Dich.	50	2	0.007	0.006	0.008	0.008	0.004	0.006	0.007	0.006	0.008	0.007	0.007	0.006	0.008	0.006	0.010	0.008	0.006	0.012	0.008
		4	0.008	0.007	0.010	0.008	0.009	0.009	0.009	0.008	0.007	0.011	0.008	0.006	0.010	0.005	0.008	0.009	0.007	0.007	0.010
	100	2	0.006	0.007	0.008	0.007	0.007	0.008	0.008	0.007	0.008	0.009	0.007	0.006	0.010	0.009	0.007	0.009	0.009	0.006	0.008
		4	0.010	0.011	0.010	0.010	0.010	0.010	0.010	0.011	0.012	0.009	0.011	0.010	0.008	0.013	0.009	0.010	0.010	0.010	0.009
	200	2	0.012	0.008	0.009	0.011	0.008	0.010	0.010	0.012	0.008	0.010	0.014	0.006	0.010	0.011	0.011	0.009	0.009	0.008	0.007
		4	0.011	0.008	0.010	0.011	0.008	0.010	0.010	0.013	0.008	0.007	0.012	0.011	0.009	0.008	0.014	0.009	0.012	0.012	0.012
Surv.	50	2	0.009	0.009	0.007	0.009	0.012	0.010	0.010	0.007	0.008	0.006	0.006	0.008	0.008	0.009	0.009	0.008	0.009	0.010	0.012
		4	0.008	0.011	0.010	0.010	0.010	0.010	0.010	0.011	0.009	0.010	0.010	0.010	0.008	0.006	0.011	0.015	0.009	0.011	0.011
	100	2	0.008	0.009	0.010	0.009	0.011	0.010	0.010	0.009	0.012	0.010	0.010	0.013	0.010	0.007	0.010	0.009	0.009	0.009	0.010
		4	0.007	0.007	0.006	0.006	0.009	0.006	0.006	0.005	0.008	0.005	0.004	0.007	0.006	0.008	0.008	0.012	0.005	0.007	0.007
	200	2	0.011	0.010	0.009	0.011	0.008	0.011	0.008	0.011	0.009	0.012	0.008	0.013	0.008	0.010	0.011	0.008	0.008	0.008	0.011
		4	0.011	0.007	0.013	0.009	0.012	0.009	0.012	0.012	0.010	0.012	0.012	0.007	0.013	0.012	0.006	0.009	0.014	0.007	0.009

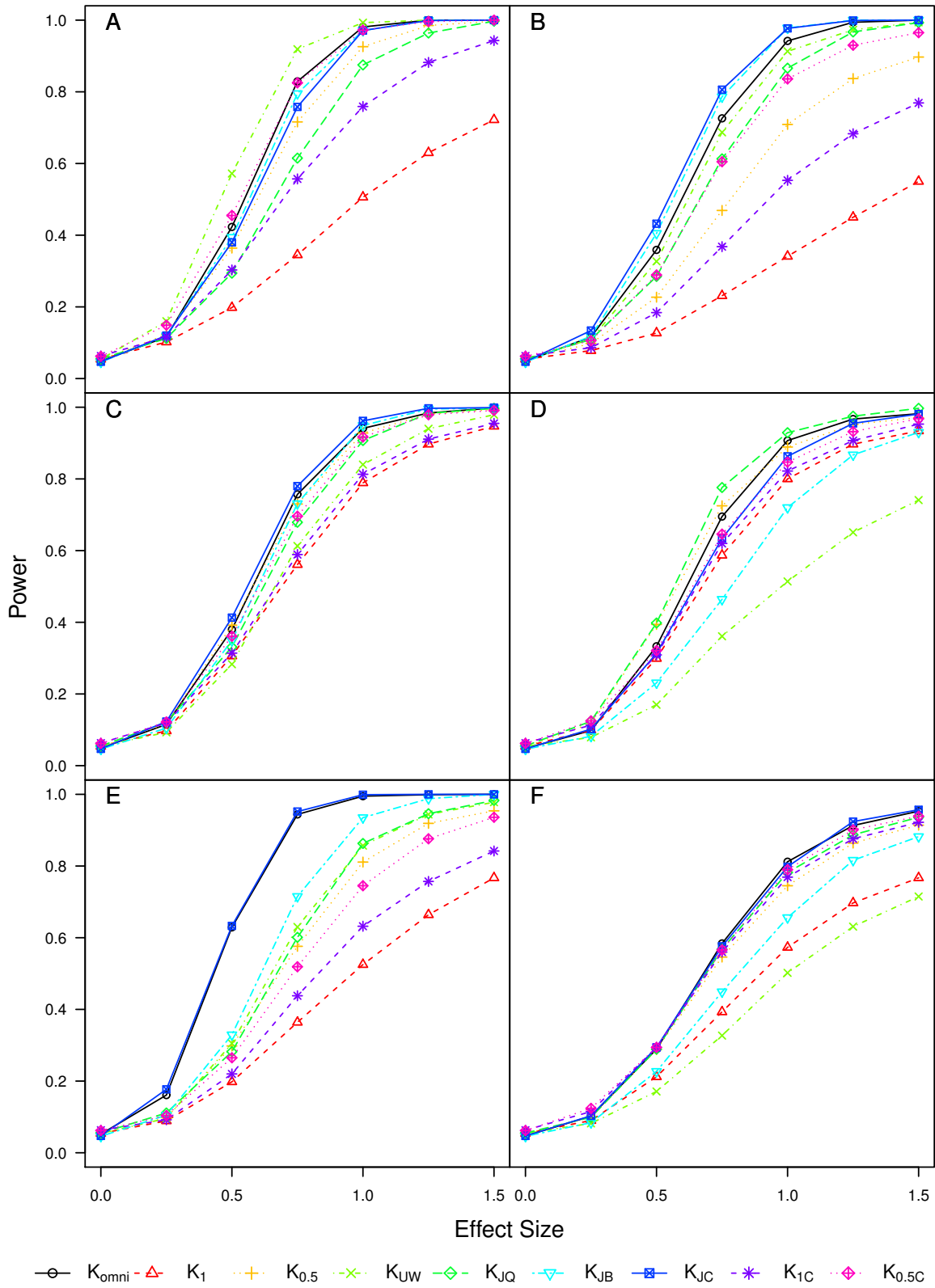


Figure 2: Continuous outcome, 2 time points.

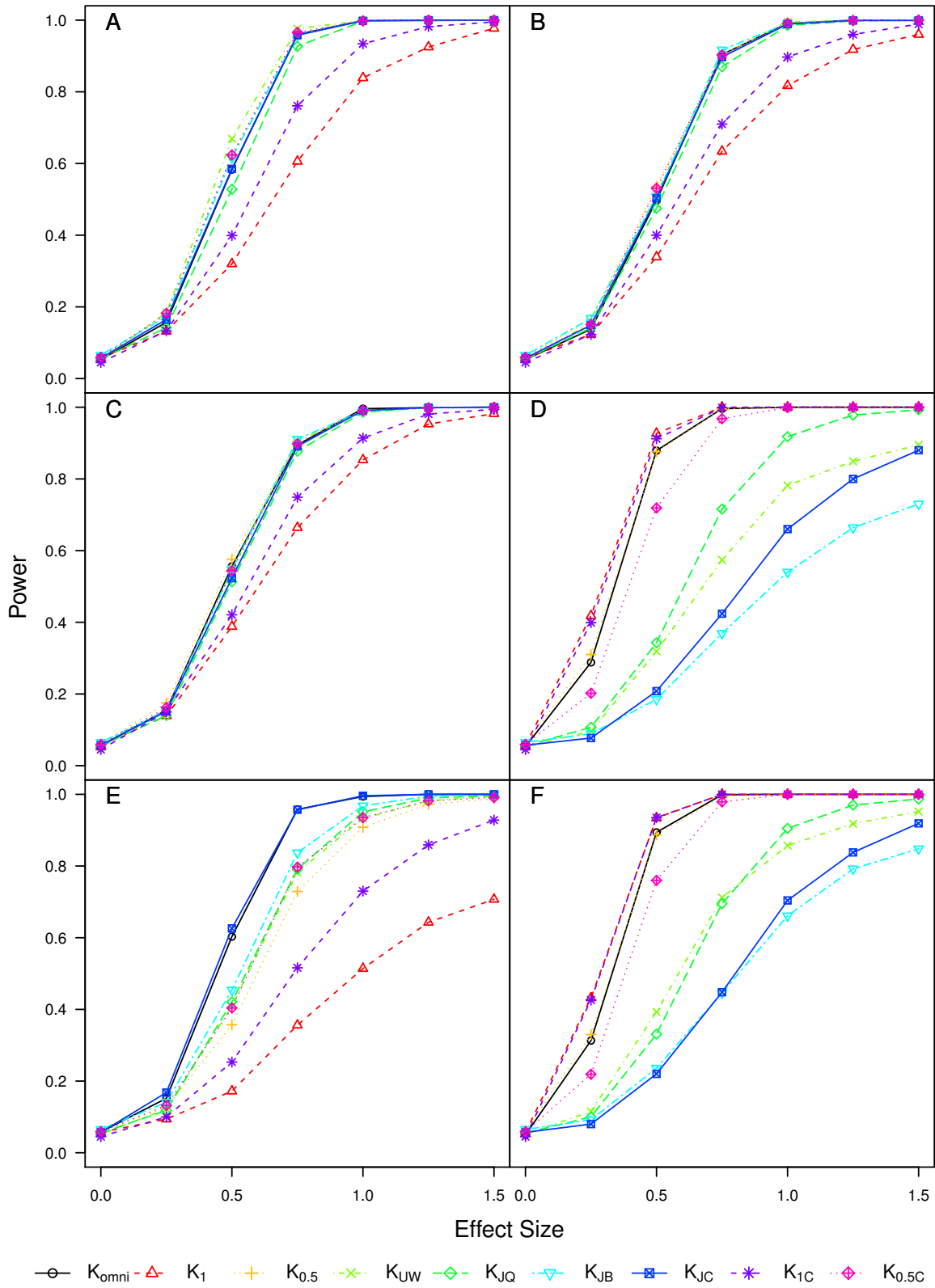


Figure 3: Continuous outcome, 4 time points.

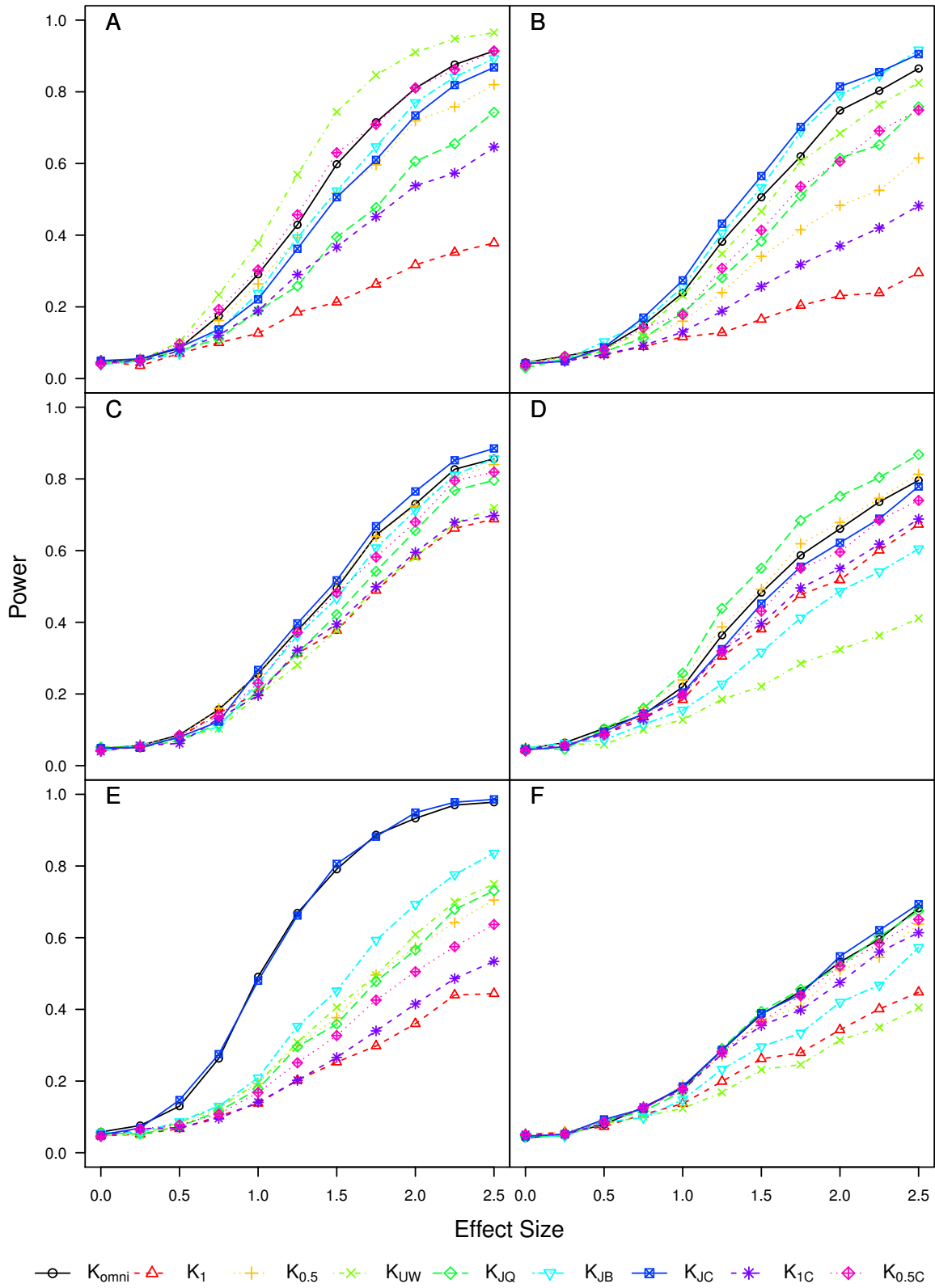


Figure 4: Dichotomous outcome, 2 time points.

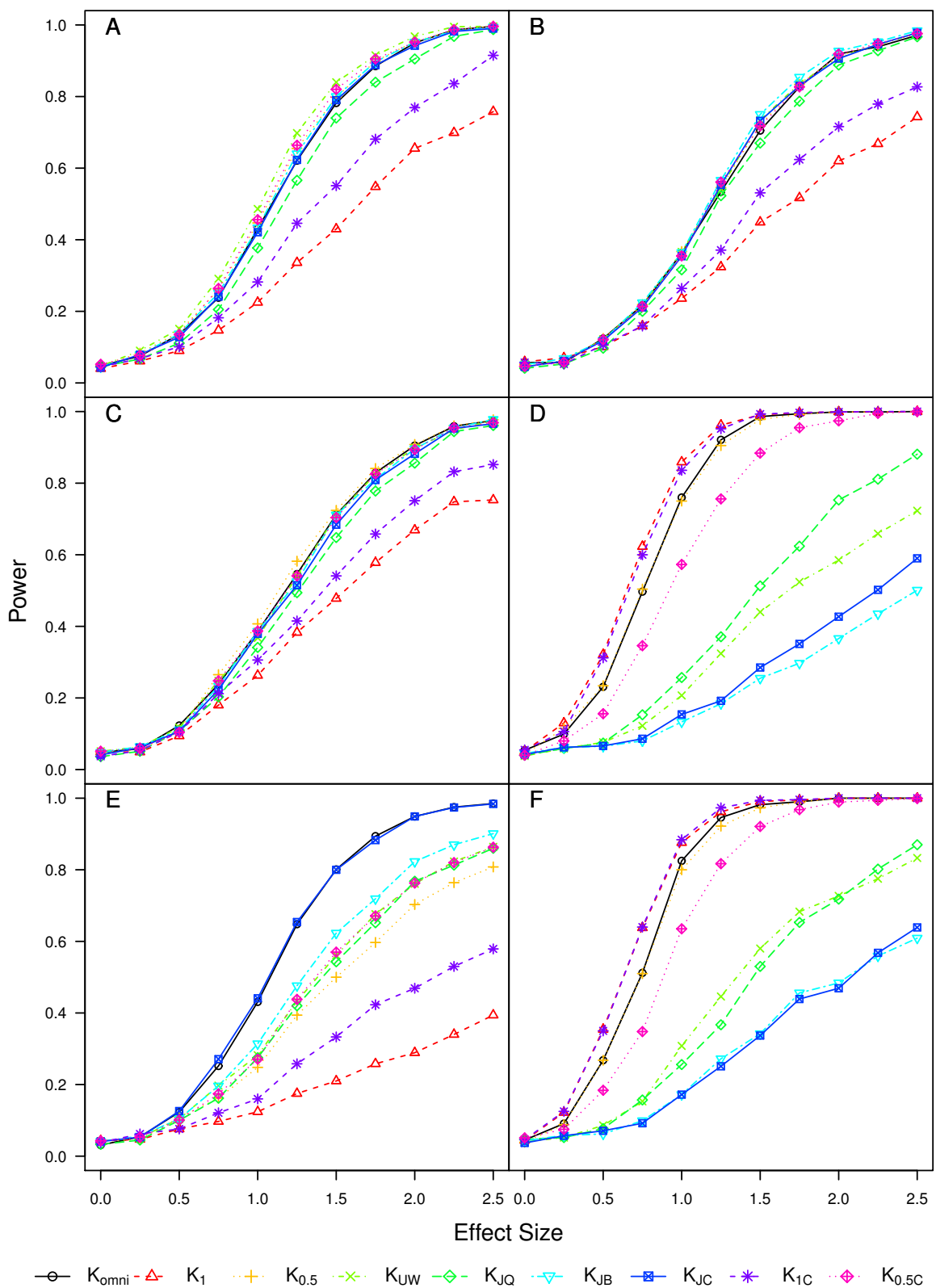


Figure 5: Dichotomous outcome, 4 time points.

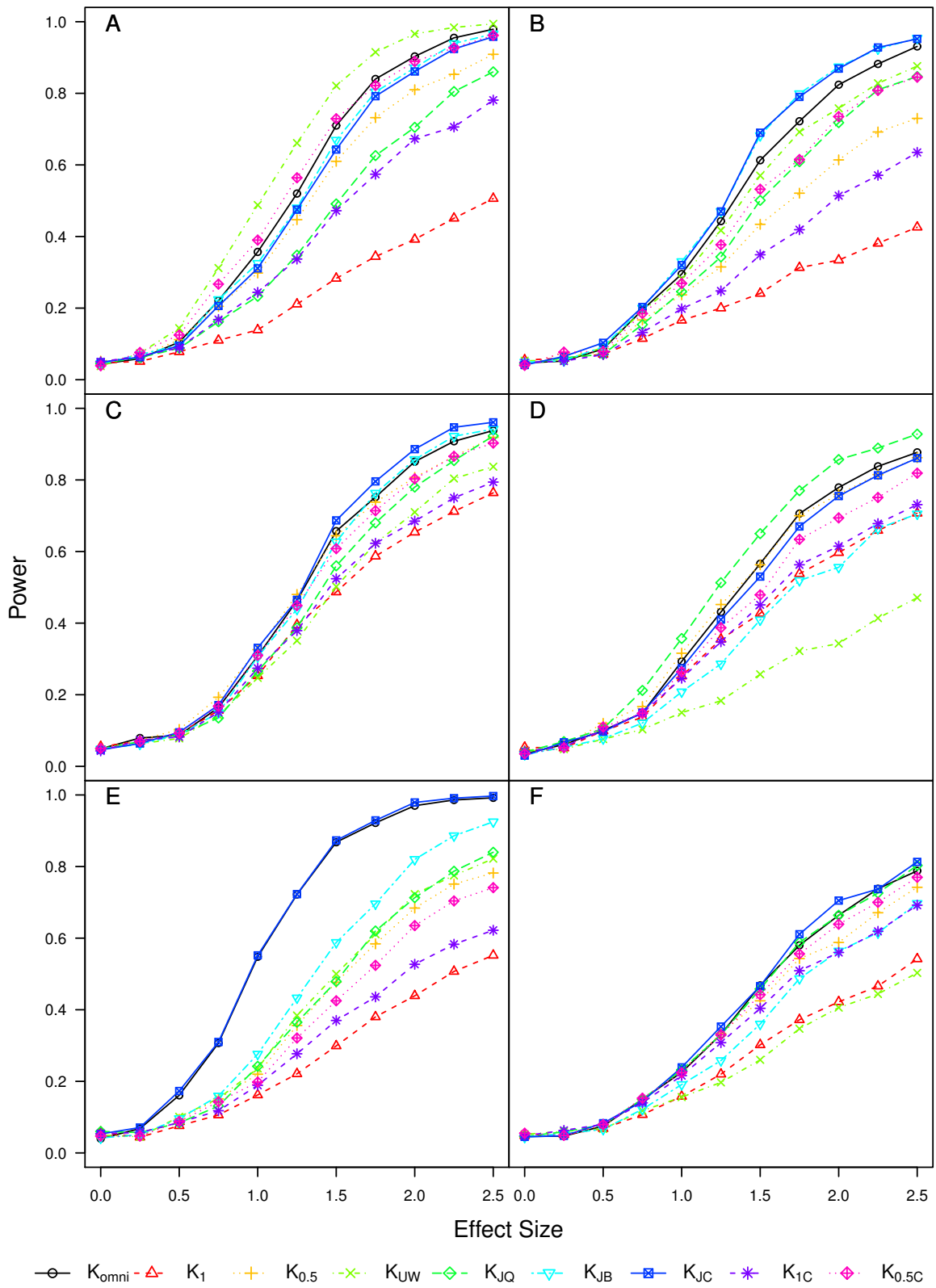


Figure 6: Time-to-event outcome, 2 time points.

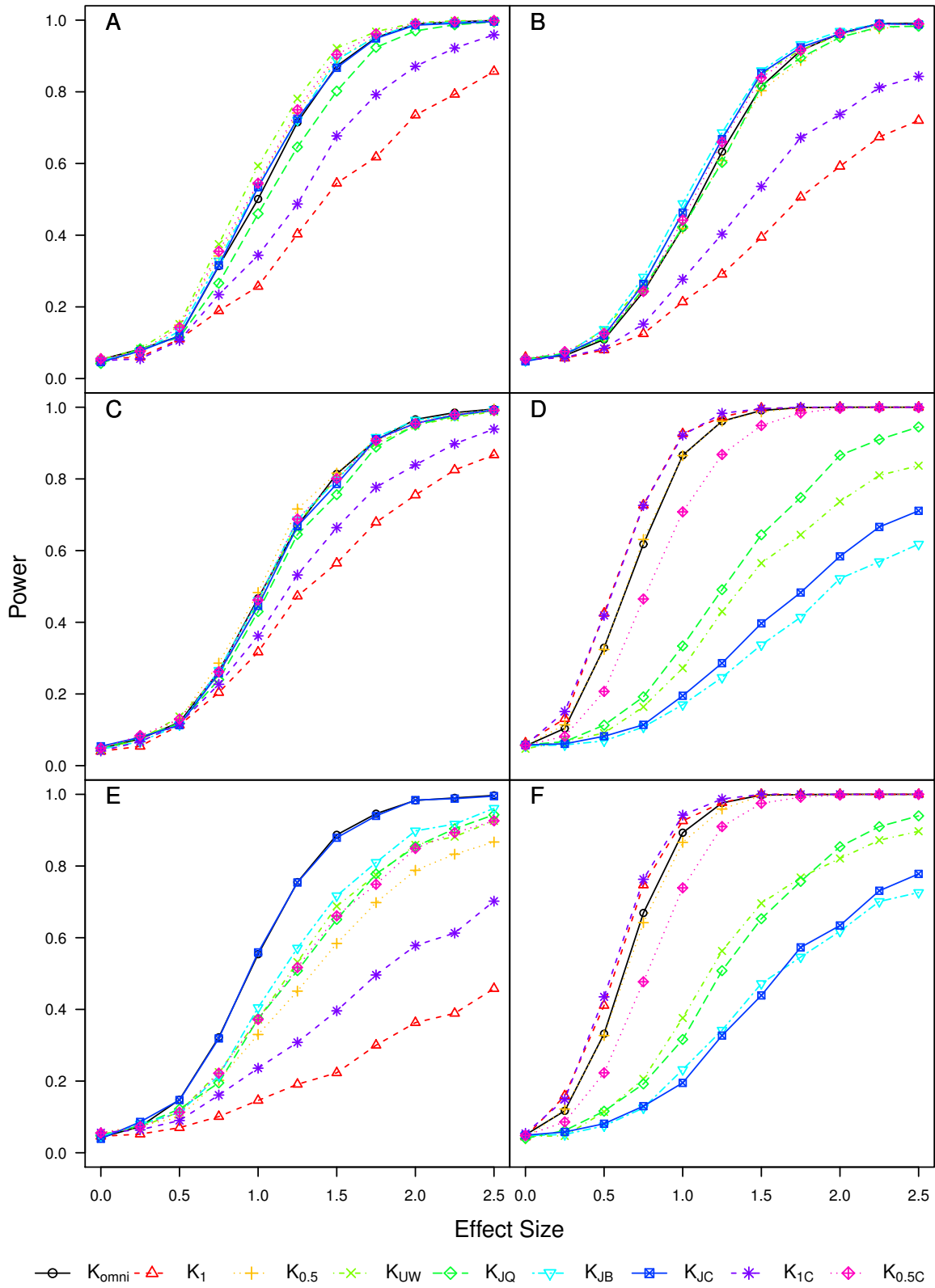


Figure 7: Time-to-event outcome, 4 time points.

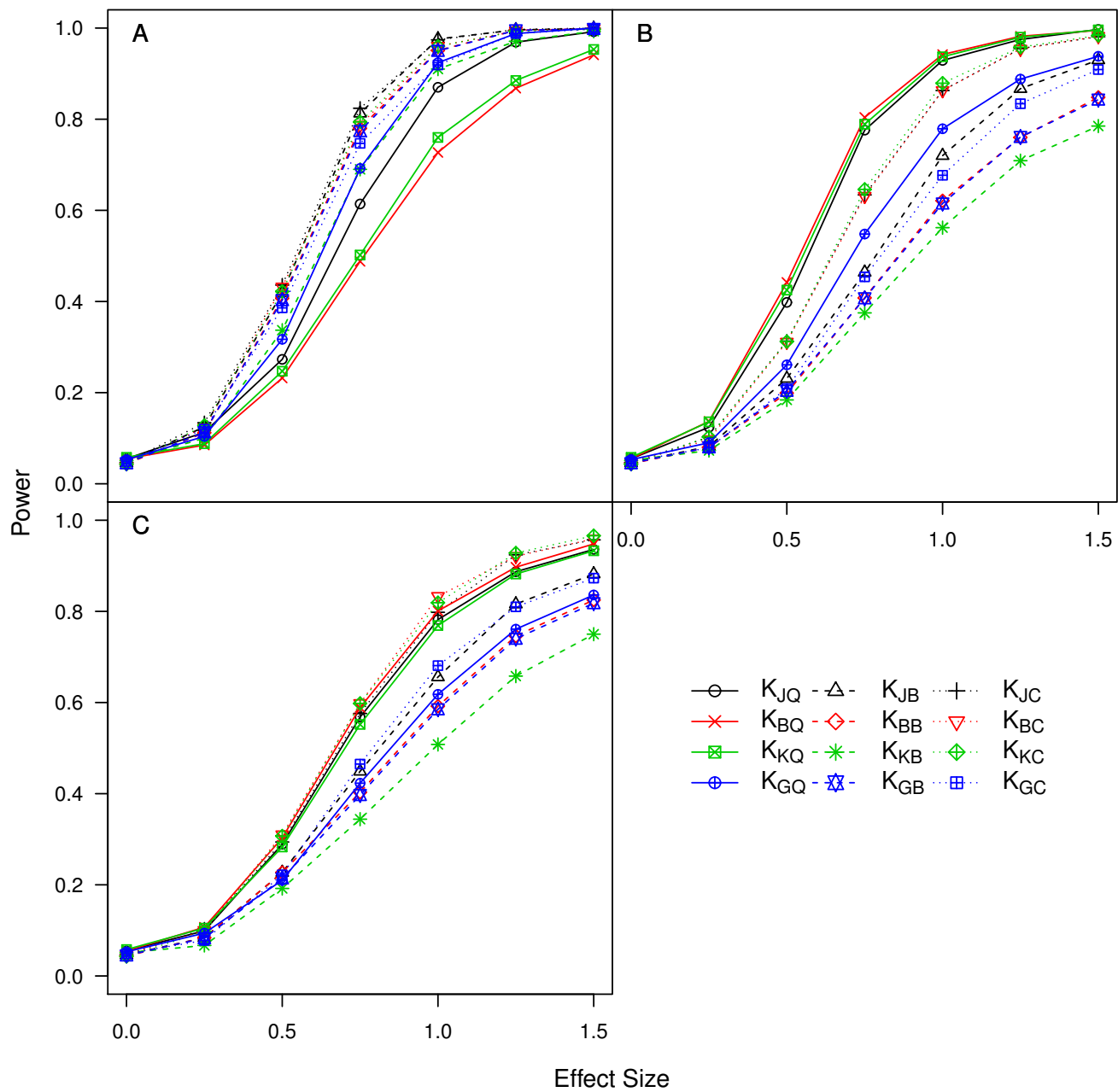


Figure 8: Continuous outcome, all nonphylogenetic distances, 2 time points.

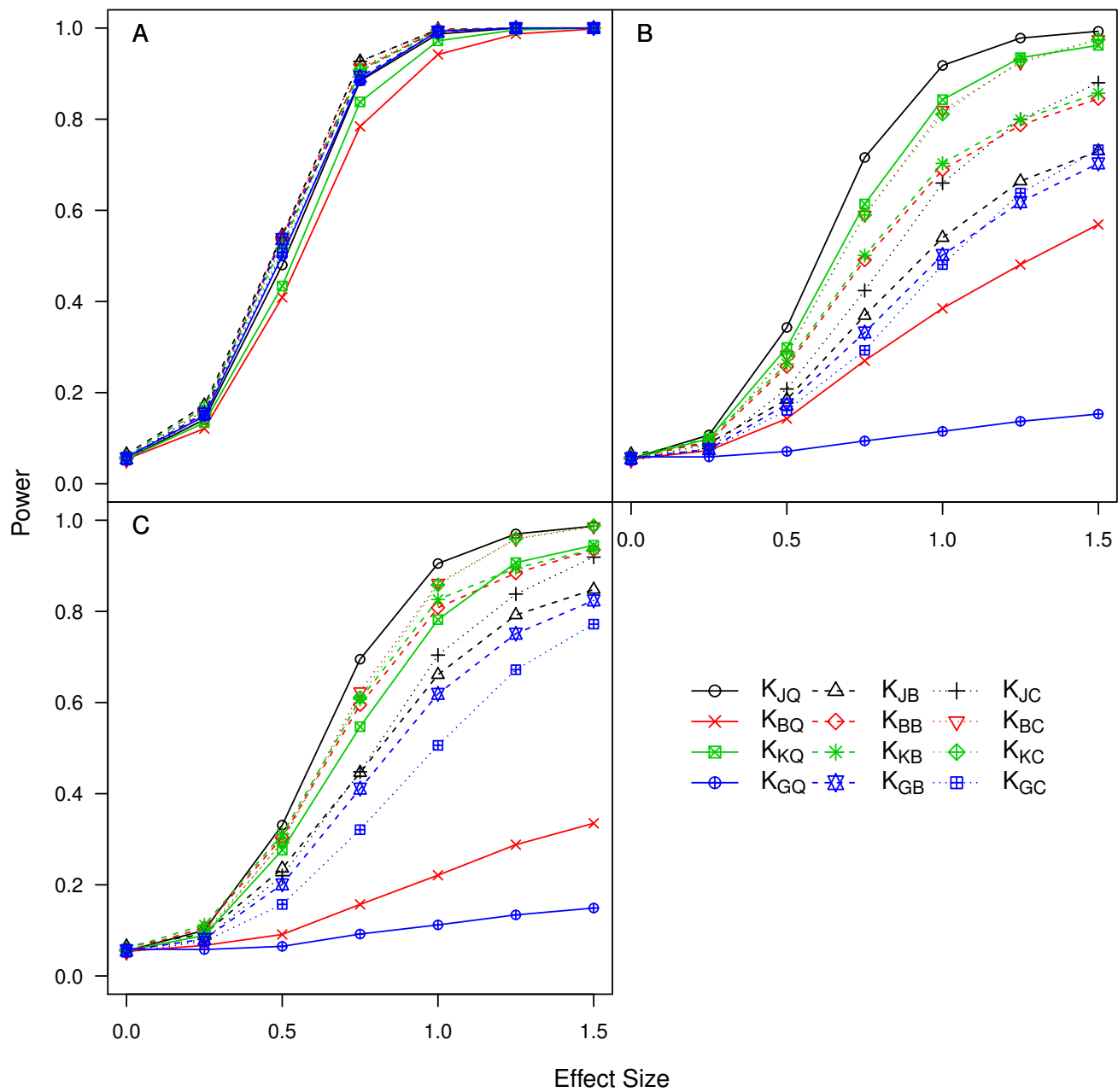


Figure 9: Continuous outcome, all nonphylogenetic distances, 4 time points.