



Contents lists available at ScienceDirect

Journal of Biomechanics

journal homepage: www.elsevier.com/locate/jbiomech
www.JBiomech.com

The probability of false positives in zero-dimensional analyses of one-dimensional kinematic, force and EMG trajectories

Todd C. Pataky^{a,*}, Jos Vanrenterghem^b, Mark A. Robinson^b

^a Institute for Fiber Engineering, Department of Bioengineering, Shinshu University, Tokida 3-15-1, Ueda, Nagano 386-8567, Japan

^b Research Institute for Sport and Exercise Sciences, Liverpool John Moores University, UK

ARTICLE INFO

Article history:

Accepted 13 March 2016

Keywords:

Statistical parametric mapping
Random field theory
Time series analysis
Kinematics
Ground reaction force
Three-dimensional analysis

ABSTRACT

A false positive is the mistake of inferring an effect when none exists, and although α controls the false positive (Type I error) rate in classical hypothesis testing, a given α value is accurate only if the underlying model of randomness appropriately reflects experimentally observed variance. Hypotheses pertaining to one-dimensional (1D) (e.g. time-varying) biomechanical trajectories are most often tested using a traditional zero-dimensional (0D) Gaussian model of randomness, but variance in these datasets is clearly 1D. The purpose of this study was to determine the likelihood that analyzing smooth 1D data with a 0D model of variance will produce false positives. We first used random field theory (RFT) to predict the probability of false positives in 0D analyses. We then validated RFT predictions via numerical simulations of smooth Gaussian 1D trajectories. Results showed that, across a range of public kinematic, force/moment and EMG datasets, the median false positive rate was 0.382 and not the assumed $\alpha=0.05$, even for a simple two-sample t test involving $N=10$ trajectories per group. The median false positive rate for experiments involving three-component vector trajectories was $p=0.764$. This rate increased to $p=0.945$ for two three-component vector trajectories, and to $p=0.999$ for six three-component vectors. This implies that experiments involving vector trajectories have a high probability of yielding 0D statistical significance when there is, in fact, no 1D effect. Either (a) explicit *a priori* identification of 0D variables or (b) adoption of 1D methods can more tightly control α .

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

In classical hypothesis testing p values represent the probability that a random process would produce an effect larger than the observed one. There are unfortunately many ways in which p value computations can go astray to yield “false positives” (Knudson, 2005; Kundson and Lindsey, 2014): the mistake of inferring an experimental effect when none exists in reality. This paper deals with one specific pitfall in p value computations which has not previously been quantified and is relevant to many branches of Biomechanics: zero-dimensional (0D) p values for one-dimensional (1D) (e.g. time varying) data.

Imagine a simple experiment which yields five scalar 1D force trajectories for each of two groups (Fig. 1b). Classical hypothesis testing can be conducted using either a “0D” or a “1D” approach (Pataky et al., 2015).

Zero-dimensional analysis: One could analyze the data using a 0D summary metric like the local maximum (Fig. 1a). In this case

a one-tailed two-sample t test of the maxima yields $t=2.357$, $p=0.023$ and one would reject the null hypothesis at $\alpha=0.05$. More completely, the 0D residuals (Fig. 1c) represent the variance about the group means, and one judges the effect size χ (Fig. 1a) against this variance. If the null hypothesis ($\chi=0$) were true, random 0D data with the same variance would produce a distribution of t values over an infinite number of identical experiments (Fig. 1e) and only 2.3% of those values would be greater than the observed $t=2.357$. The null hypothesis is rejected because the observed t value exceeds the threshold t_{0D}^* corresponding to α (Fig. 1e and g). This t_{0D}^* value can be rapidly computed in statistical software packages, or it can be computed by iteratively simulating thousands of t tests on randomly generated samples of 0D Gaussian data.

One-dimensional analysis: One could alternatively analyze the data using 1D methods (Lenhoff et al., 1999; Pataky et al., 2015) (Fig. 1, right panels). Analogous to the 0D procedure, the 1D residuals (Fig. 1d) embody the variance about mean trajectories, and the null hypothesis is the null difference trajectory: $\chi(q)=0$, where q is time or the 1D measurement domain. If the null hypothesis were true, random 1D data with the same variance and same smoothness would produce a distribution of t trajectory maxima

* Corresponding author. Tel.: +81 268 21 5609; fax: +81 268 21 5318.

E-mail address: tpataky@shinshu-u.ac.jp (T.C. Pataky).

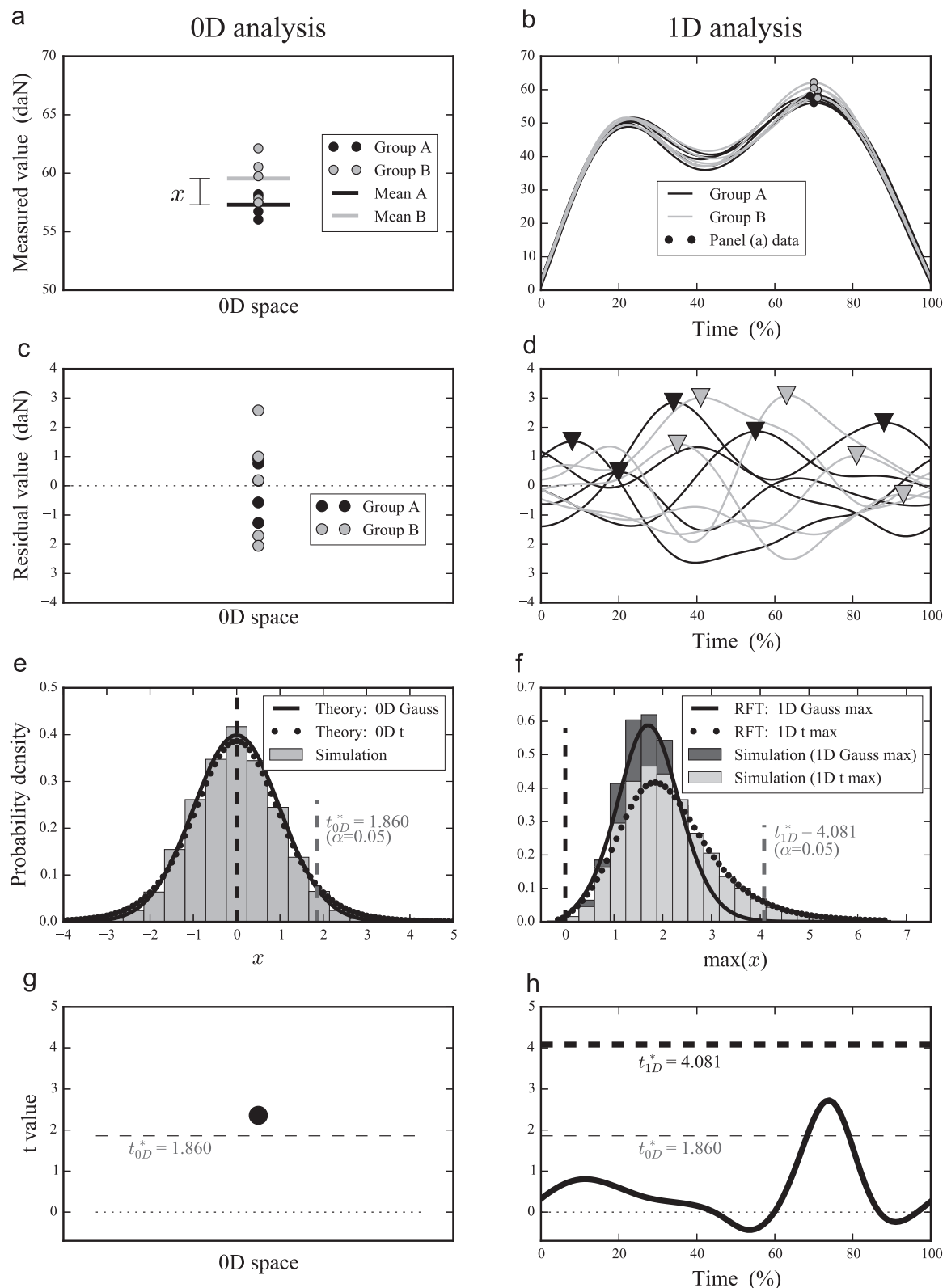


Fig. 1. Overview of 0D vs. 1D statistical analyses. (a,b) 0D data extracted from 1D dataset (unit: decanewton). (c,d) Residuals with 1D residual maxima highlighted. (e,f) Theoretical and simulated probability density functions; 1D densities correspond to the 1D maximum (RFT=random field theory). Simulations validate the theoretical predictions and involved 100,000 random residuals (Gaussian densities) and 100,000 two-sample t tests using Gaussian residuals (t densities). (g,h) Hypothesis testing results; the observed t value exceeds the α -defined t^* value for 0D analysis but not 1D analysis so the null hypothesis is rejected for 0D but not 1D analysis.

over an infinite number of identical experiments (Fig. 1f), and in this case random 1D data would produce the observed 0D effect of $t=2.357$ with a probability of approximately 55%, which is well above α . In other words, random 1D data produce particular t

values with generally much greater probability than do random 0D data. The null hypothesis is not rejected because the observed maximum t value, across the whole trajectory, does not exceed the threshold t_{1D}^* corresponding to α (Fig. 1f and h). The t_{1D}^* value can

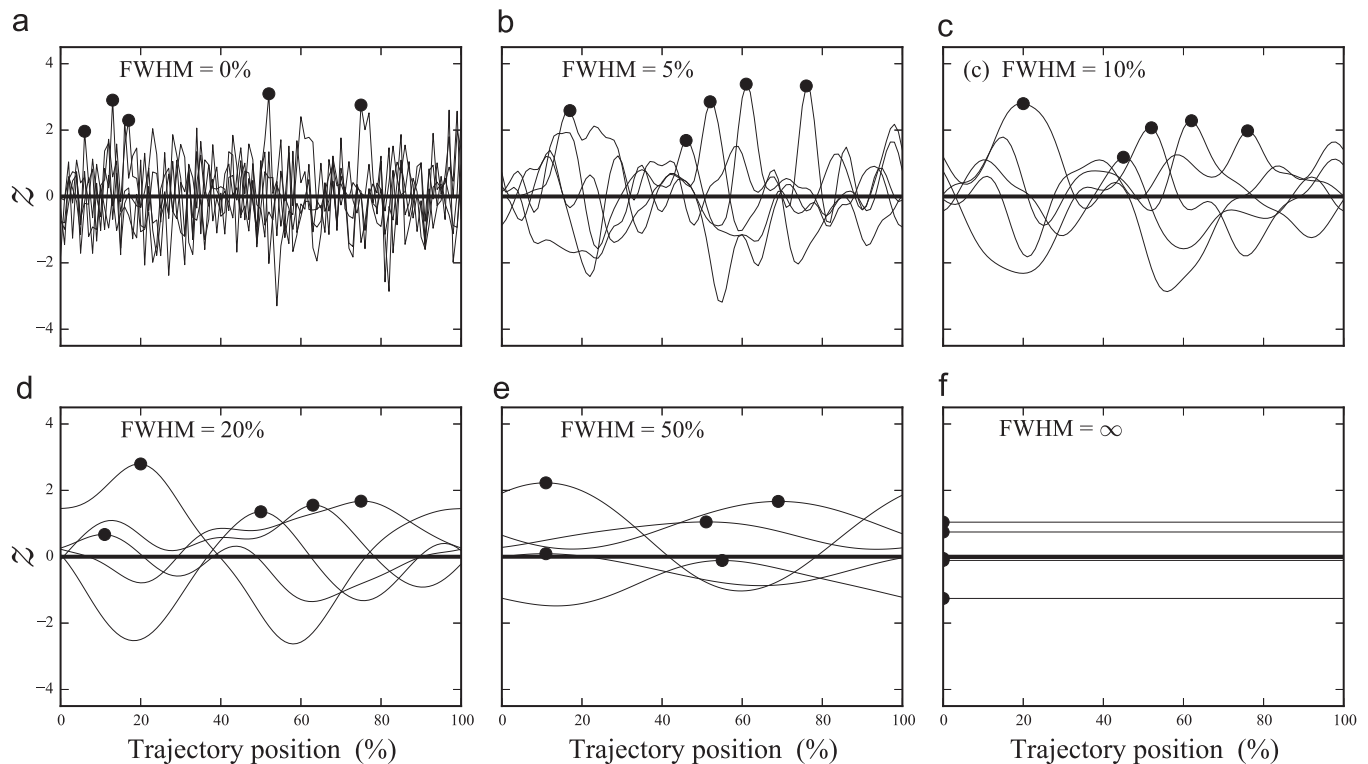


Fig. 2. Simulated 1D Gaussian random fields. (a) Uncorrelated Gaussian data. (b–e) Data from panel (a) but smoothed with Gaussian kernels of varying FWHM (see Appendix B). Circles depict trajectory maxima whose FWHM-dependent behavior random field theory (RFT) describes (Adler and Taylor, 2007). (f) Infinitely smooth 1D random fields are equivalent to 0D random scalars; RFT results converge to 0D results as FWHM approaches ∞ .

be analytically calculated using random field theory (RFT) (Adler and Taylor, 2007; Friston et al., 2007), or it can be computed by iteratively simulating thousands of t tests on randomly generated samples of smooth 1D Gaussian data (Fig. 2) (Pataky, 2016).

The 0D and 1D procedures have yielded opposite hypothesis testing results, so which is correct? The answer depends on the *a priori* hypothesis. If, prior to conducting an experiment, one explicitly identified that particular 0D metric as the sole metric of empirical interest then the empirical question is inherently 0D and the 0D result is correct. If, however, one measured 1D data and did not specify that particular 0D metric prior to the experiment then the empirical question is inherently 1D and the 1D result is correct. Failing to specify a 0D metric prior to a 1D experiment and then adopting 0D methods has been termed “regional focus bias” (Pataky et al., 2013) and is a potential source of false positives. False positive prevalence for 1D biomechanics datasets has previously been estimated for 0D procedures (Knudson, 2005) but not, to our knowledge, in the context of 0D vs. 1D procedures.

The purpose of this study was to quantify the false positive rates that could be expected in real 1D biomechanical datasets when employing 0D statistical inference. To that end we analyzed nine public datasets (Table 2) representing a variety of experimental tasks (walking, running, cutting, and cycling) and data modalities (forces, kinematics, and EMG). Based on the data’s temporal smoothness we estimated the likelihood of producing false positives using a simplified experimental design: a two-sample t test with $N=10$ for each group. Analyses for more complex designs like ANOVA are addressed in the Discussion (Section 4.3). The key theoretical concept we shall attempt to convey is that two parameters – the mean (μ) and standard deviation (σ) – describe 0D Gaussian behavior, and that only one additional parameter – 1D smoothness (FWHM) (Fig. 2) (Appendices A and B) – is needed to describe 1D Gaussian behavior.

Table 1

Dataset dimensionality terminology. An “ n D m D” dataset contains an m -dimensional vector or tensor measured over an n -dimensional spatiotemporal domain.

n	m	Example	
		n domain	m domain
0	1	Ground impact instant	Knee flexion
	3	Instant of max propulsion	Ground reaction force
1	1	Gait cycle time	Knee flexion
	2	Gait cycle time	Left and right knee flexion
	3	Stance phase	Ground reaction force
2	1	Foot contact surface	Pressure
3	1	Femur	von Mises stress
3	6	Femur	Strain tensor

To clarify our “0D” and “1D” terminology we shall also employ the term: “ n D m D”, where n and m are the dimensionalities of the measurement domain and dependent variable, respectively (Table 1). In n D m D datasets the physical nature of the variables changes across the m components but not across the n D measurement domain. Biomechanics studies often measure 1D m D data but use 0D1D models of randomness to define critical statistical thresholds, and this paper quantifies false positive rates associated with that approach. Throughout this paper “0D” and “1D” represent to “0D1D” and “1D m D”, respectively.

We emphasize that this paper focusses on just a single statistical issue: the probability of false positives in a single 0D1D two-sample t test conducted on 1D m D data. We use only the two-sample t test because this simple test sufficiently demonstrates the magnitude of the false positives problem and because the problem is exacerbated in more complex designs like ANOVA. We

acknowledge that many other issues must be considered when conducting statistical analyses including: small sample sizes, non-sphericity, normality, outliers, etc. Just as it is useful to consider each of these issues individually, we feel it is equally useful to consider 0D vs. 1D analysis individually because this issue is relevant to all 1D data analyses but has not been explicitly addressed in the literature.

2. Methods

All analyses were implemented in Python 2.7 (van Rossum, 2014) using Canopy 1.4 (Enthought Inc., Austin, USA) and the open-source software package “rft1d” (Pataky, 2016) (www.spm1d.org/rft1d). For readers unfamiliar with Python, MATLAB source code (The MathWorks, Natick, USA) replicating the study's main analyses and results is provided as Supplementary Material (Appendix C).

2.1. Experimental data smoothness estimations

Since 1D smoothness determines the height that random 1D trajectories reach (Fig. 2), we first estimated the smoothness of 1D biomechanical trajectories from nine public datasets (Table 2). Each dataset contained 1D scalar and/or vector time series spanning a normalized interval of 0–100%. All data were analyzed in their publicly available form and no extra signal processing was conducted. Detailed descriptions of the data are available in the original papers and in the public datasets themselves.

After computing residual trajectories (Fig. 1d) as the difference between group/subject means (as appropriate for each dataset), we then calculated the residuals' smoothness using a robust FWHM estimation procedure (Kiebel et al., 1999) (Appendix D). The “FWHM” is the full-width at half-maximum of a Gaussian kernel (Appendix A) which, when convolved with completely uncorrelated Gaussian data (Appendix B) yields random 1D trajectories with the same smoothness as the observed 1D residuals. Effectively the estimated FWHM is simply the mean temporal gradient normalized by residual magnitude. This FWHM estimation procedure has been validated elsewhere for 1D data (Pataky, 2016) and a similar validation is available in the MATLAB Supplementary Material (Appendix C). Subsequent analyses consider the full range of estimated FWHM values.

2.2. Theoretical false positive rates

2.2.1. False positive rate for random 0D1D data

The probability that random 0D1D data, or equivalently: 0D univariate Gaussian data, would produce a t value which exceeds an arbitrary height u is given by the survival function:

$$P_{0D}(t > u) = \int_u^\infty \left(\frac{\Gamma((\nu+1)/2)}{\sqrt{\nu\pi} \Gamma(\nu/2)} \left(1 + \frac{x^2}{\nu} \right)^{-(\nu+1)/2} \right) dx \quad (1)$$

where ν is the degrees of freedom and Γ is the gamma function. Note that, given a height u , $P_{0D}(t > u)$ is dependent on only sample size as manifested in the parameter ν .

Classical hypothesis testing on 0D1D data is conducted by setting Eq. (1) to α :

$$P_{0D}(t > t_{0D}^*) = \alpha \quad (2)$$

and then solving for the critical threshold t_{0D}^* . If the experimentally observed t value exceeds t_{0D}^* then the null hypothesis is rejected. To be clear, false positives occur in 0D analysis of 0D data when random 0D data exceed t_{0D}^* , and this occurs at a rate of α by definition.

Table 2

Dataset overview. N is the total of number of trajectories.

Source	N	Tasks	Variables	Link
Besier et al. (2009)	43	Walking, running	GRF, muscle forces	simtk.org/home/muscleforces
Caravaggi et al. (2010)	30	Walking	Plantar arch deformation	spm1d.org/Downloads.html
Dorn et al. (2012)	8	Running	GRF	simtk.org/home/runningspeeds
Fregley et al. (2012)	19	Walking	GRF, COP, knee implant forces	simtk.org/home/kneeloads
Kautz et al. (1991)	20	Cycling	Pedal dynamics	isbweb.org/data/kautz/
Murley et al. (2014)	5	Walking	EMG	ncbi.nlm.nih.gov/pubmed/24618372
Neptune et al. (1999)	20	Cutting	Kinematics, EMG	isbweb.org/data/rrn/index.html
Pataky et al. (2008)	600	Walking	GRF	spm1d.org/Downloads.html
Schwartz and Rozumalski. (2008)	161	Walking	Kinematics	ncbi.nlm.nih.gov/pubmed/18565753

Note: The Kautz et al. (1991) dataset contains 26 trials but six are duplicates. Variables: COP=center of pressure, EMG=electromyography, and GRF=ground reaction force.

2.2.2. False positive rate for random 1D1D data

The probability that random 1D1D data, or equivalently 1D univariate Gaussian trajectories, produce t trajectories which reach arbitrary heights u is (Worsley et al., 2004; Friston et al., 2007):

$$P_{1D}(t_{\max} > u) = 1 - \exp \left[-P_{0D}(t > u) - \frac{S}{W} \frac{\sqrt{4 \log 2}}{2\pi} \left(1 + \frac{u^2}{\nu} \right)^{-(\nu-1)/2} \right] \quad (3)$$

where t_{\max} is the trajectory maximum, S is the trajectory length (constant for all trajectories in one dataset, usually $S=100$) and W is the FWHM representing trajectory smoothness. Note that, relative to the 0D case (Eq. 1), just one additional parameter (S/W) is needed to describe the probabilistic behavior of t_{\max} .

The probability that 1D1D data will reach the 0D threshold for significance (t_{0D}^*) is thus simply:

$$\text{False positive rate (0D1D analysis, 1D1D data)} = P_{1D}(t_{\max} > t_{0D}^*) \quad (4)$$

To our knowledge this false positive rate has not been previously reported. We thus calculated Eq. (4) as a function of ν for three α values: 0.01, 0.05 and 0.10, and also over the range of the experimentally estimated FWHM values.

Last, we computed the false positive rate for the case of completely uncorrelated data (Fig. 2a) using the Bonferroni correction:

$$P_{\text{Bonf}}(t_{\max} > t_{0D}^*) = 1 - (1 - \alpha)^{100} \quad (5)$$

Note that Eq. (5) is accurate only for the case of 100 independent tests and is therefore inaccurate when 1D data are smooth. We nevertheless use this Bonferroni result as a reference to demonstrate that 1D results converge to this result as 1D trajectories become increasingly rough.

2.2.3. False positive rate for random 1DmD data

We last considered three separate cases of 1DmD data: one, two and six vector trajectories where each vector has three components. Equivalently, these are 1DmD multivariate Gaussian trajectories with: $m=3$, $m=6$ and $m=18$, respectively. For example, $m=6$ could represent an experiment involving two joints' three rotations. For each case we assumed independently varying vector components, thereby yielding:

$$\text{False positive rate \{0D1D analysis, 1DmD data\}} = 1 - (1 - P_{1D}(t_{\max} > t_{0D}^*))^m \quad (6)$$

Note that Eqs. (4) and (6) are equivalent when $m=1$.

2.3. Theoretical validations

We validated all theoretical results using 0D and 1D Gaussian random data generators as implemented in SciPy (Jones et al., 2001) and rft1d (Pataky, 2016), respectively (cf. Appendix E). Specifically, we generated 100,000 random datasets then conducted one t test for each dataset, thereby yielding 100,000 t values or 100,000 t trajectories. To validate theoretical predictions (Eqs. (1) and (3)) we calculated the percentage of the 0D t values and 1D t_{\max} values to exceed arbitrary thresholds u .

For 0D data we repeated these simulations for two sample sizes ($\nu=4$ and $\nu=48$), and for the 1D case we additionally repeated simulations for 10 different FWHM values to span the range of experimentally observed smoothness values. Last, we repeated all 1D simulations for the aforementioned multivariate cases of one-, two- and six three-component vector trajectories which could represent the analysis of a single joint in three dimensions, the synergistic function of the hip and knee and all the three main joints of the lower limb, respectively.

3. Results

3.1. Experimental data smoothness

Residual 1D trajectories from all datasets, two of which are depicted in Fig. 3, were qualitatively consistent with simulated smooth Gaussian 1D trajectories (Fig. 2). Quantitative consistency between 1D residuals and Gaussian 1D trajectories has been demonstrated elsewhere (Pataky et al., 2015).

Residual smoothness estimates yielded minimum, median and maximum FWHM values of: 6.2%, 16.5% and 67.0%, respectively, across all datasets (Table 3). Kinematic residuals were smoothest on average, followed by force and then EMG residuals. Half of all trajectory variables analyzed lay within a range of FWHM=[11.9%, 29.5%] and 90% of all trajectory variables lay within a range of FWHM=[9.4%, 36.5%] (Appendix F). Most of the experimental trajectories investigated therefore lay in the smoothness range depicted in Fig. 2b–e.

3.2. Theoretical false positive rates and validations

The 0D and 1D survival functions (Eqs. (1) and (3), respectively) are depicted in Fig. 4 for two different sample sizes ($\nu=4$ and $\nu=48$). Considering first the 0D survival functions, it is clear that 50% of tests on random 0D data yield t values larger than zero, implying that 50% are less than zero. Next, the critical threshold is sample-size dependent: $t_{0D}^* = 1.67$ and 2.13 for $\nu=4$ and 48,

respectively ($\alpha=0.05$). For these two sample sizes Fig. 4 shows that the maximum t value produced by random 1D Gaussian trajectories will reach these critical heights (t_{0D}^*) with much greater probability. For the median observed smoothness of FWHM=16.5%, that probability is $p=0.431$ and 0.376 for $\nu=4$ and 48, respectively. Simulating random Gaussian 0D and 1D data (depicted as dots in Fig. 4) validated all results.

For two-sample t tests with $N=10$ in each group ($\nu=2N-2=18$), false positive rates were greater than α for all smoothness values (Fig. 5), and this rate approached the Bonferroni rate (depicted as stars in Fig. 5) as the FWHM approached zero. The 1D false positive rate converged to the 0D α only for FWHM= ∞ because an infinitely smooth 1D trajectory is equivalent to a 0D scalar.

Table 3

Smoothness estimation results, summary across all datasets. FWHM values represent 1D residual smoothness (Appendix F) and higher FWHM values reflect smoother residuals. FWHM data unit: % of the 1D trajectory length. Refer to Fig. 2 for a qualitative representation of similar smoothness values.

Data modality	Min	Median	Max	Mean \pm SD
Kinematics	10.8	33.1	67.0	34.7 \pm 14.1
Forces	6.2	14.3	32.4	16.1 \pm 7.1
EMG	7.2	11.8	15.6	12.1 \pm 2.5
Overall	6.2	16.5	67.0	21.9 \pm 13.6

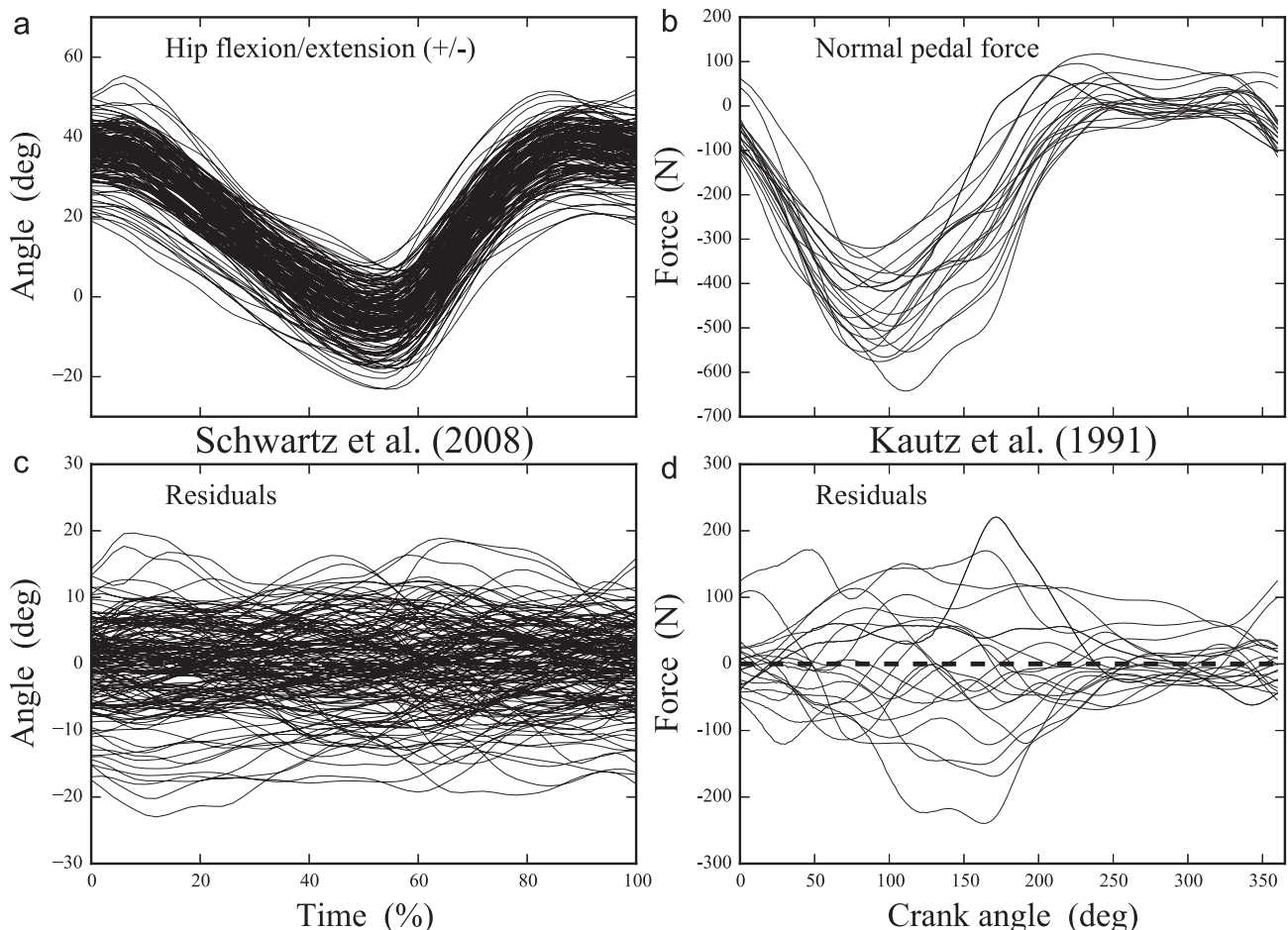


Fig. 3. Example residuals from two public datasets. (a,b) Original data. (c,d) Residuals; here the differences between individual trajectories and the mean trajectory. The estimated smoothness values for (c) and (d) were FWHM=19.6% and FWHM=18.9%, respectively.

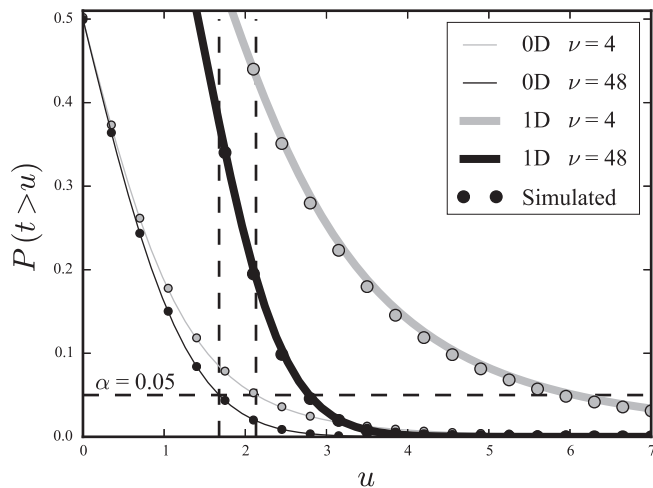


Fig. 4. Survival functions for the t statistic for two different degrees of freedom ($\nu=4$ and $\nu=48$) when the underlying data are (i) random 0D Gaussian scalars (Eq. (1)) or (ii) smooth 1D Gaussian trajectories (Eq. (3)) with the median observed smoothness of FWHM=15.1%. Solid lines depict theoretical probabilities and dots depict validation results. The broken horizontal line depicts the classical hypothesis testing threshold for significance, and broken vertical lines depict the critical t values for the two 0D cases. The intersection of these vertical lines with the 1D curves gives the true false positive rate (approximately 37% and 47%, respectively) if one regards the residuals as 1D rather than 0D (Fig. 1c,d).

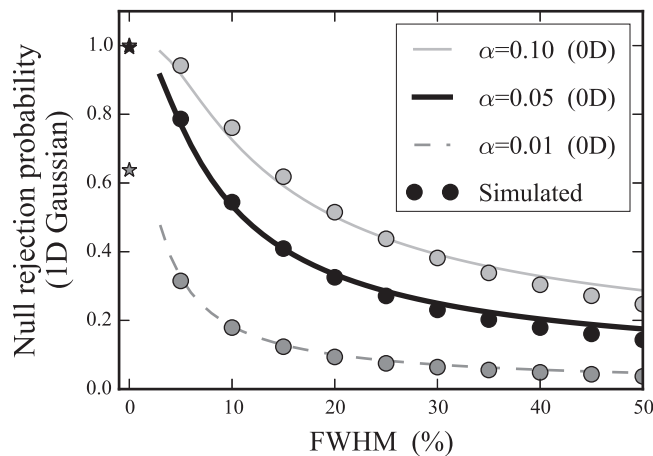


Fig. 5. Probability of false positives for two-sample t tests ($N=10$) if using a 0D randomness model and if analyzing the maximum 1D t value. Star symbols for FWHM=0 depict the case of completely uncorrelated trajectories (Eq. (5)).

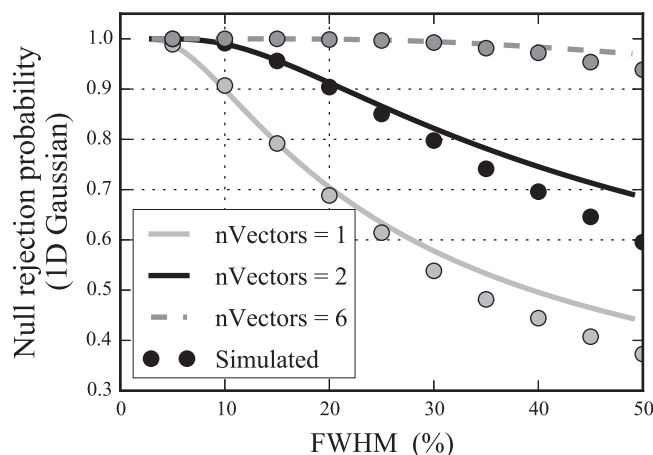


Fig. 6. Probability of false positives for two-sample t tests ($N=10$) for three-component vectors with independently varying components.

Table 4

False positive results for two-sample t tests with $N=10$ per group; summary of results from Figs. 5 and 6. The three FWHM columns represent the minimum, median, and maximum value observed across all datasets (Table 3). Vector data are assumed to contain three independent components each.

Type of 1D data	Reference	FWHM=6.2	FWHM=16.5	FWHM=67.0
One scalar	Fig. 5	$p=0.699$	$p=0.382$	$p=0.145$
One vector	Fig. 6	$p=0.973$	$p=0.764$	$p=0.374$
Two vectors	Fig. 6	$p=0.999$	$p=0.945$	$p=0.609$
Six vectors	Fig. 6	$p=1.000$	$p=0.999$	$p=0.940$

Last, Fig. 6 depicts results for two-sample experiments involving 1DmD data (i.e. m -component vector trajectories). For 1D3D data, Fig. 6 suggests that the probability of a false positive is approximately $p=0.90$ when FWHM=10%. For 1DmD data where $m \geq 6$, Fig. 6 suggests that false positives are nearly certain irrespective of trajectory smoothness. For the smoothest observed data (FWHM=67.0), the probability of a false positive was estimated to be $p=0.145$ and $p=0.940$ for 1D1D and 1D18D data, respectively (Table 4). This implies that methods which control 1D false positive rates should also control false positive rates associated with multivariate data when multivariate data are measured and analyzed.

4. Discussion

4.1. Main implications

The convention of $\alpha=0.05$ implies that one accepts a 5% false positive rate when conducting classical hypothesis testing. The main result of this study was that smooth, random 1D trajectories generally produce false positives in 0D analyses with a probability much higher than α . Even for the best case – maximum smoothness (FWHM=67.0) and one scalar trajectory – false positive rates were nearly three times greater than α ($p=0.145$, Table 4). For the median smoothness observed across all datasets (FWHM=16.5), the false positive rates for three-component vector trajectories were greater than $p=0.76$. In the worst case – maximum roughness (FWHM=6.2) and two or more three-component vectors – the false positive rate exceeded $p=0.999$. Since Biomechanics studies often measure and analyze multiple three-component vector trajectories (e.g. 3D joint angles and moments at the hip, knee and ankle during gait), these results imply that false positives are highly probable in 0D analyses of typical 1D datasets. Inflated false positive rates in 0D analyses of 1D data have previously been suggested (Lenhoff et al., 1999; Pataky et al., 2015; Robinson et al., 2015) but to our knowledge have not been previously quantified.

We stress that these results in no way invalidate published 0D analyses of 1D data most obviously because large effects are generally discoverable irrespective of the analysis procedure; clearly a somewhat stronger signal in Fig. 1h would cross both 0D and 1D thresholds for significance. Second, as a general limitation of classical hypothesis testing: significance does not imply practical meaning and vice versa. Last, all results are valuable regardless of their magnitude if they lead to subsequent independent scrutiny and verification through repeated experimentation.

4.2. Context of ideal hypothesis testing

An ideal hypothesis-driven experiment (Fig. 7a) involves formulating a testable null hypothesis, from which the independent variables (IVs), dependent variables (DVs) and often the experiment itself directly emerge. When dealing with complex systems

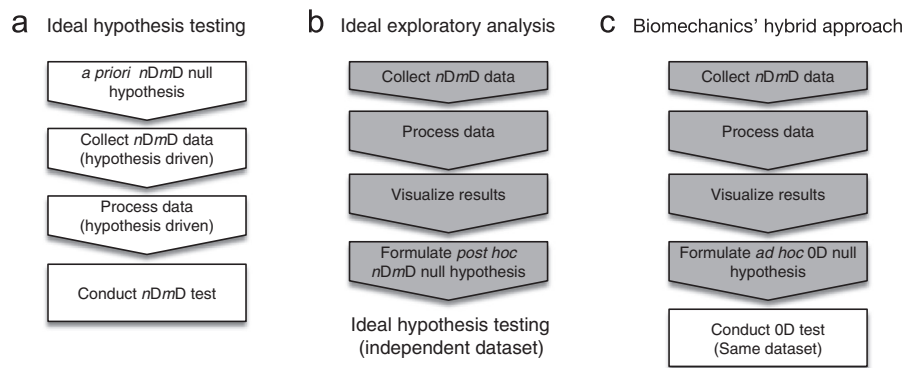


Fig. 7. General procedures for hypothesis testing and exploratory analysis. In Biomechanics the dimensionalities of the test and the data may be incommensurate.

it can be difficult to follow that procedure because it may be difficult to formulate specific hypotheses prior to conducting the experiment. In this case, exploratory analyses (Fig. 7b) provide an extended framework for valid hypothesis testing. In particular, the investigator can freely choose arbitrary measurements, processing procedures, IVs, DVs and tests provided the reported results are confirmed by applying the identical procedures to an independent dataset. In other words, exploratory analyses are useful for narrowing the empirical scope of the study and to generate specific null hypotheses to be tested on independent data.

Biomechanics studies often adopt a hybrid of the aforementioned hypothesis testing and exploratory procedures (Fig. 7c). The result is unfortunately a procedure which contains critical scientific flaws. The fundamental problem is that hypothesis tests are OD1D (see Table 1), but this is incommensurate with the dimensionality of the *a priori* DV which is 1DmD. In other words, models of OD1D randomness cannot describe the probabilistic behavior of 1DmD data (Fig. 1), so statistical tests cannot pertain to the collected 1DmD data. The relative absence of 1DmD procedures in the Biomechanics literature is, in our opinion, a major oversight. Access to 1DmD theory and procedures would help future studies converge to ideal hypothesis testing and exploratory procedures.

To summarize, the recipe below will ensure that ideal classical hypothesis testing is conducted when analyzing 1DmD data (Table 1):

1. Did I formulate a specific null hypothesis regarding a specific OD1D dependent variable prior to conducting the experiment?
2. If yes to #1, then I must use OD1D hypothesis testing methods and I must not report 1DmD data except qualitatively because 1DmD data are irrelevant to the hypothesis.
3. If no to #1, then I must use 1DmD methods and must not report OD1D data except qualitatively because OD1D data are irrelevant to the hypothesis.

4.3. Limitations

The false positive results (Table 4) pertain directly only to two-sample experiments involving 1 trajectories in each group, and these values will change for different sample sizes and different designs. We nevertheless found qualitatively identical trends after repeating these analyses for a large variety of designs (one-sample, regression, one-way and two-way ANOVA and MANOVA). In particular, false positive rates increased with the number of measured trajectories in all designs, and numerical values were only slightly different from those in Table 4. For example, one-way ANOVA with three groups and ten trajectories per group yielded a false positive rate of $p=0.433$, which is just slightly higher than the two-sample case ($p=0.382$). We report only two-sample results for brevity.

A second consideration is that we conducted univariate analysis of multivariate (vector) trajectories using two-sample *t* tests even though we should have conducted multivariate analysis using Hotelling's T^2 test (Cao and Worsley, 1999; Pataky et al., 2013). We report univariate results to be consistent with the Biomechanics literature's high prevalence of univariate analyses of multivariate data (Knudson, 2005).

A third apparent limitation is that one is not obliged to analyze maxima (Fig. 1d). For example, the maximum or “peak” does not necessarily lie within *a priori* temporal windows of empirical interest in sports maneuvers (Besier et al., 2001, 2003) so other metrics may be chosen to represent those movement phases. False positive rates for non-maxima would be lower than those reported here. We would nevertheless argue that choosing non-maxima is scientifically unjustified for the following reasons: if one has an *a priori* hypothesis regarding a OD variable, then it is irrelevant whether that variable is a maximum, minimum, or intermediate value because only that OD variable must be analyzed. The issue of maxima vs. non-maxima therefore pertains only to 1D analyses. For 1D analysis the statistical goal is to quantify the probability that random 1D trajectories would produce an effect as large as the observed effect, and this by definition pertains to the maximum effect.

A fourth, and real limitation is the broad issue of classical vs. Bayesian inference. This paper considered only classical inference. Bayesian inference affords a much broader class of statistical inferences and is generally regarded to supersede classical inference as a more objective means of scientific inquiry (Kruschke, 2013). Since the Biomechanics literature mainly adopts classical inference (Knudson, 2005) we leave Bayesian inference for future work.

A fifth limitation is that the present analyses assumed isotropically smooth residual trajectories (Fig. 3). This may not be a good assumption for some datasets which have mixed-frequency signals. As an example, ground reaction forces during running typically have high-frequency impact signals followed by comparatively low-frequency propulsive forces (Cavanagh and LaFortune, 1980). Fortunately, there are two factors mitigating potential problems associated with anisotropic smoothness. First, it is possible to correct for anisotropic smoothness simply by estimating the trajectory length which yields constant smoothness (Worsley et al., 1999). Second, the unbiased smoothness estimation approach adopted herein (Kiebel et al., 1999) assures that FWHM estimates are robust to minor anisotropy. Regardless, research is needed to estimate the prevalence and seriousness of smoothness anisotropy in biomechanical trajectories.

A general limitation of 1D methods is sensitivity, which is the ability of a statistical test to detect a true effect. Statistical thresholds are naturally higher for 1DmD analyses than for ODmD

analyses (Fig. 1) so sensitivity is generally lower in 1D procedures. However, in our experience 1D procedures usually have sufficient sensitivity because effect sizes tend to be large in biomechanics datasets, even for small datasets (Pataky et al., 2015). In fact it could be argued that sensitivity is lower in 0D vs. 1D analyses because 0D analyses fail to consider the entire dataset and therefore cannot detect all signals. Regardless, in order to ensure adequate sensitivity one should generally conduct power analysis before proceeding with a full experiment. Procedures for 0D1D power analysis are well known and procedures for n DmD power analysis are described elsewhere (Friston et al., 2007). Additionally, a simple way to boost sensitivity when dealing with 1DmD data is to first conduct 1D analyses, then identify a 0D variable of interest and last conduct a separate experiment using 0D analysis of the identified variable (Fig. 7).

A second general limitation of 1D methods is complexity. 1D procedures are naturally more complex than 0D procedures and are usually more difficult both to learn and to deploy. Exacerbating the complexity problem is a lack of suitable reference material. While many textbooks detail the differences between 0D1D and 0DmD procedures (Rencher and Chistensen, 2012), none of which we are aware details 1DmD procedures. Some books detail subsets of 1D1D procedures (Ramsay and Silverman, 2005; Zhang, 2013), and others thoroughly summarize 3D3D and 4D1D procedures (Friston et al., 2007), but these contain a variety of concepts which are not directly relevant to 1DmD data analysis. Regardless of reference material availability, we feel that complexity is an unavoidable necessity: 1DmD methods are the least complex methods for controlling false positives in 1DmD datasets. A side benefit of learning 1DmD methods is that all aspects of simpler analyses become clearer; from an n DmD perspective typical 0D1D procedures are simply the special case: $n=0$, $m=1$, and all concepts relevant to 0D1D analysis (e.g. ANOVA, normality, outliers, non-sphericity, etc.) also apply directly to n DmD data analysis.

Perhaps the most important limitation of 1D techniques is software availability. No commercial statistical software package of which we are aware implements 1D procedures. The most prominent open-source software packages to implement n D procedures, including packages like SPM8 (Friston et al., 2007), are all tailored for $n=3$ or $n=4$, making them somewhat bulky and overly complex for 1D datasets. A few open-source packages tailored specifically for 1D analysis exist including: FDA (Ramsay and Silverman, 2005), *spm1d* (Pataky, 2012) and *rft1d* (Pataky, 2016), but all are still at relatively early stages of development. Clear interfaces to 1D procedures in commercial packages would remove a major barrier to 1D procedure accessibility in Biomechanics.

4.4. Summary

Conducting scalar 0D analyses of 1D data without clear *a priori* specification of the 0D scalar produces false positives at relatively high rates ($p > 0.38$ and $p > 0.76$ for 1D scalar trajectories and 1D three-component vector trajectories, respectively). The most robust protection against false positive results is good experimental design focussing on the testing of a small number of specific hypotheses whose variables are clearly defined *a priori*. Since 0D analysis of ambiguous 0D variables fails to consider 1D randomness, 0D techniques cannot control false positive rates (α) in experiments whose hypotheses pertain to 1D data. To solve the problem one should either (i) explicitly identify 0D variables prior to conducting an experiment or (ii) adopt 1D procedures.

Conflict of interest

None declared.

Acknowledgments

This work was supported by Wakate A Grant 15H05360 from the Japan Society for the Promotion of Science. We also wish to thank Cyril J. Donnelly for helpful discussions and continued support.

Appendix A. Supplementary data

Supplementary data associated with this paper can be found in the online version at <http://dx.doi.org/10.1016/j.jbiomech.2016.03.032>.

References

- Adler, R.J., Taylor, J.E., 2007. *Random Fields and Geometry*. Springer-Verlag, New York.
- Besier, T.F., Lloyd, D.G., Chocrane, J.L., Ackland, T.R., 2001. External loading of the knee joint during running and cutting maneuvers. *Med. Sci. Sports Exercise* 33 (7), 1168–1175.
- Besier, T.F., Lloyd, D.G., Ackland, T.R., 2003. Muscle activation strategies at the knee during running and cutting maneuvers. *Med. Sci. Sports Exercise* 35 (1), 119–127.
- Besier, T.F., Fredericson, M., Gold, G.E., Beaupre, G.S., Delp, S.L., 2009. Knee muscle forces during walking and running in patellofemoral pain patients and pain-free controls. *J. Biomech.* 42 (7), 898–905 (<https://simtk.org/home/muscleforces>).
- Cavanagh, P.R., LaFortune, M.A., 1980. Ground reaction forces in distance running. *J. Biomech.* 13 (5), 397–406.
- Cao, J., Worsley, K.J., 1999. The detection of local shape changes via the geometry of Hotelling's T2 fields. *Ann. Stat.* 27 (3), 925–942.
- Dorn, T.T., Schache, A.G., Pandey, M.G., 2012. Muscular strategy shift in human running: dependence of running speed on hip and ankle muscle performance. *J. Exp. Biol.* 215, 1944–1956 (<https://simtk.org/home/runningspeeds>).
- Friston, K.J., Ashburner, J.T., Kiebel, S.J., Nichols, T.E., Penny, W.D., 2007. *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. Elsevier/Academic Press, Amsterdam, p. 2007 (ISBN 9780123725608).
- Jones, E., Oliphant, T., Peterson, P., 2001. *SciPy: Open Source Scientific Tools for Python* (<http://www.scipy.org>).
- Kautz, S.A., Feltner, M.E., Coyle, E.F., Baylor, A.M., 1991. The pedalin technique of elite endurance cyclists: changes with increasing workload at constant cadence. *Int. J. Sport Biomech.* 7, 29–53.
- Kiebel, S.J., Poline, J., Friston, K.J., Holmes, A.P., Worsley, K.J., 1999. Robust smoothness estimation in statistical parametric maps using standardized residuals from the general linear model. *NeuroImage* 10 (6), 756–766.
- Knudson, D.V., 2005. Statistical and reporting errors in applied biomechanics research. In: *Proceedings of the 23rd International Conference on Biomechanics in Sports*, pp. 811–814.
- Kundson, D.V., Lindsey, C., 2014. Type I and type II errors in correlations of various sample sizes. *Compr. Psychol.* 3(1) (Article 1).
- Kruschke, J.K., 2013. Bayesian estimation supersedes the *t* test. *J. Exp. Psychol.: Gen.* 142 (2), 573–603.
- Lenhoff, M.W., Santer, T.J., Otis, J.C., Peterson, M.G., Williams, B.J., Backus, S.I., 1999. Bootstrap prediction and confidence bands: a superior statistical method for analysis of gait data. *Gait Posture* 9, 10–17.
- Neptune, R.R., Wright, I.C., van den Bogert, A.J., 1999. Muscle coordination and function during cutting movements. *Med. Sci. Sports Exercise* 31 (2), 294–302 (Data: (<http://isbweb.org/data/rnn/>)).
- Pataky, T.C., 2012. One-dimensional statistical parametric mapping in Python. *Comput. Methods Biomech. Biomed. Eng.* 15 (3), 295–301.
- Pataky, T.C., Robinson, M.A., Vanrenterghem, J., 2013. Vector field statistical analysis of kinematic and force trajectories. *J. Biomech.* 46 (14), 2394–2401.
- Pataky, T.C., 2016. Smooth one-dimensional random field upcrossing probabilities in Python. *J. Stat. Softw.*, in press.
- Pataky, T.C., Vanrenterghem, J., Robinson, M.A., 2015. Zero- vs. one-dimensional, parametric vs. non-parametric, and confidence interval vs. hypothesis testing procedures in one-dimensional biomechanical trajectory analysis. *J. Biomech.* 48 (7), 1277–1285.
- Ramsay, J.O., Silverman, B.W., 2005. *Functional Data Analysis*. Springer, New York.
- Rencher, A.C., Chistensen, W.F., 2012. *Methods of Multivariate Analysis*. Wiley, New Jersey.

- Robinson, M.A., Vanrenterghem, J., Pataky, T.C., 2015. Statistical parametric mapping for alpha-based statistical analyses of multi-muscle EMG time-series. *J. Electromyogr. Kinesiol.* 25 (1), 14–19.
- Schwartz, M.H., Rozumalski, A., 2008. The gait deviation index: a new comprehensive index of gait pathology. *Gait Posture* 28, 351–357.
- van Rossum, G., 2014. The Python Library Reference Release 2.7.8 (<https://docs.python.org/2/library/>).
- Worsley, K.J., Andermann, M., Koulis, T., MacDonald, D., Evans, A.C., 1999. Detecting changes in nonisotropic images. *Hum. Brain Mapp.* 8, 98–101.
- Worsley, K.J., Taylor, J.E., Tomaiuolo, F., Lerch, J., 2004. Unified univariate and multivariate random field theory. *NeuroImage* 23, S189–S195.
- Zhang, J.T., 2013. *Analysis of Variance for Functional Data*. CRC Press, London.