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

Todd C. Pataky[1], Jos Vanrenterghem[2], and Mark A. Robinson[2]

[1]Department of Bioengineering, Shinshu University, Japan
[2]Research Institute for Sport and Exercise Sciences, Liverpool John Moores University, UK

March 18, 2016

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 variance 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 and EMG datasets, the median false positive rate was 0.382 and not the assumed α=0.05, even for a simple two-sample t test involving N=10 trajectories per group. The median false positive rates 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 metrics or (b) adoption of 1D methods can more tightly control α.

Corresponding Author:
Todd Pataky, Ph.D., Institute for Fiber Engineering, Department of Bioengineering, Shinshu University
Tokida 3-15-1, Ueda, Nagano, Japan 386-8567
tpataky@shinshu-u.ac.jp, T.+81-268-21-5609, F.+81-268–21-5318

Keywords: *statistical parametric mapping; random field theory; time series analysis; kinematics; ground reaction force; three-dimensional analysis*

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 local maxima (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 α=0.05. More completely, the 0D residuals (Fig.1c) represent the variance about the group means, and one judges the effect size x (Fig.1a) against this variance. If the null hypothesis (x=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,g). This t^*_{0D} value can be rapidly computed in nearly all 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: $x(q) = 0$, 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 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,h). The t^*_{1D} value can be readily calculated using random field theory (RFT) procedures (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 is easy: both are correct but both cannot be simultaneously correct within a predefined α level. 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) which represent a variety of experimental tasks (walking, running, cutting, cycling) and data modalities (forces, kinematics, 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 (§4.3). The key theoretical concept we shall attempt to convey is that two parameters — the mean (μ) and standard deviation (σ) — completely describe 0D Gaussian behavior, and that only one additional parameter — 1D smoothness (FWHM) (Fig.2) (Appendix A, B)— is needed to completely describe 1D Gaussian behavior.

To clarify our "0D" and "1D" terminology we shall also employ the term: "nDmD", where n and m are the dimensionalities of the measurement domain and dependent variable, respectfully (Table 1). In nDmD datasets the physical nature of the variables changes across the m components but not across the nD measurement domain. Biomechanics studies often measure 1DmD 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 "1DmD", 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 1DmD data. We use only the two-sample t test because (a) this simple test sufficiently demonstrates the magnitude of the false positives problem and (b) 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:

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

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

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

$$P_{0\mathrm{D}}(t > t^*_{0\mathrm{D}}) = \alpha \tag{2}$$

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

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):

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

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 (Eqn.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 ($t^*_{0\mathrm{D}}$) is thus simply:

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

To our knowledge this false positive rate has not been previously reported. We thus calculated Eqn.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:

$$P_{\text{Bonf}}\left(t_{\max} > t^*_{0\text{D}}\right) = 1 - (1-\alpha)^{100} \tag{5}$$

Note that Eqn.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 demonstrated 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, respectfully. For example, m=6 could represent an experiment involving two joints' three rotations. For each case we assumed independently varying vector components, thereby yielding:

$$\text{False positive rate } \{\text{0D1D analysis, 1D}m\text{D data}\} = 1 - \left(1 - P_{1\text{D}}(t_{\max} > t^*_{0\text{D}})\right)^m \tag{6}$$

Note that Eqns.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 (c.f. 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 (Eqns.1, 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 (ν=4, ν=48), and for the 1D case we additionally repeated simulations for ten 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 ninety percent 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 through Fig.2e.

3.2 Theoretical false positive rates and validations

The 0D and 1D survival functions (Eqns.1 and 3, respectively) are depicted in Fig.4 for two different sample sizes (ν=4 and ν=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: $t^*_{0\mathrm{D}}$ =1.67 and 2.13 for ν=4 and 48, respectively (α=0.05). For these two sample sizes Fig.4 shows that the maximum t value produced by random 1D1D Gaussian trajectories will reach these critical heights ($t^*_{0\mathrm{D}}$) with much greater probability. For the median observed smoothness of FWHM=16.5%, that probability is p=0.431 and 0.376 for ν=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.

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 for 1D1D and 1D18D data, respectfully (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 α=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 was 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 nearly certain when conducting 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.7)a 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 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.7)b 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.7)c. The result is a unfortunately a procedure which contains critical scientific flaws. The fundamental problem is that hypothesis tests are 0D1D (see Table 1), but this is incommensurate with the dimensionality of the *a priori* DV which is 1DmD. In other words, models of 0D1D 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 0D1D dependent variable prior to conducting the experiment?

2. If yes to #1, then I must use 0D1D hypothesis testing methods and I mustn't 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 mustn't report 0D1D data except qualitatively because 0D1D data are irrelevant to the hypothesis.

4.3 Limitations

The false positive results (Table 4) pertain directly only to two-sample experiments involving ten 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 prevelance 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; Besier et al., 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'd nevertheless argue that choosing non-maxima is scientifically unjustified for the following reasons: if one has an *a priori* hypothesis regarding a 0D variable, then it is irrelevant whether that variable is a maximum, minimum, or intermediate value because only that 0D 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 signal.

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 almost exclusively 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 0DmD analyses (Fig.1) so sensitivity is generally lower in 1D procedures. However, in our experience 1D procedures usually have

sufficient sensitivity because effect sizes are usually large in biomechanics datasets, even for small datasets (Pataky et al., 2015). In fact it could be argued that sensitivity is much 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 nDmD 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 on independent subjects using 0D analysis of the identified variable (Table **??**).

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) procedures, but these contain many concepts which are unneeded in 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 nDmD 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 nDmD data analysis.

Perhaps the most important limitation of 1D techniques is software availability. No commercial statistical software package of which we aware implements 1D procedures. The most prominent open-source software packages to implement nD 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.

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.

Conflict of Interest

The authors report no conflict of interest, financial or otherwise.

References

Adler, R. J. and Taylor, J. E. 2007. *Random Fields and Geometry*, Springer-Verlag, New York.

Besier, T. F., Lloyd, D. G., Chochrane, J. L., and Ackland, T. R. 2001. External loading of the knee joint during running and cutting maneuvers, *Medicine & Science in Sports & Exercise* **33(7)**, 1168–1175.

Besier, T. F., Lloyd, D. G., and Ackland, T. R. 2003. Muscle activation strategies at the knee during running and cutting maneuvers, *Medicine & Science in Sports & Exercise* **35(1)**, 119–127.

Besier, T. F., Fredericson, M., Gold, G. E., Beaupre, G. S., and Delp, S. L. 2009. Knee muscle forces during walking and running in patellofemoral pain patients and pain-free controls, *Journal of Biomechanics* **42(7)**, 898–905, data: `https://simtk.org/home/muscleforces`.

Cavanagh, P. R. and Lafortune, M. A. 1980. Ground reaction forces in distance running, *Journal of Biomechanics* **13(5)**, 397–406.

Cao, J. and Worsley, K. J. 1999. The detection of local shape changes via the geometry of Hotelling's T2 fields, *Annals of Statistics* **27(3)**, 925–942.

Dorn, T. T., Schache, A. G., and Pandy, M. G. 2012. Muscular strategy shift in human running: dependence of running speed on hip and ankle muscle performance., *Journal of Experimental Biology* **215**, 1944–1956, data: `https://simtk.org/home/runningspeeds`.

Duhamel, A., Bourriez, J., Devos, P., Krystkowiak, P., Destee, A., Derambure, P., and Defebvre, L. 2004. Statistical tools for clinical gait analysis, *Gait and Posture* **20(2)**, 204–212.

Friston, K. J., Worsley, K. J., Frackowiak, R. S. J., Mazziotta, J. C., and Evans, A. C. 1994. Assessing the significance of focal activations using their spatial extent., *Hum Brain Mapp* **1**, 210–220.

Friston, K. J., Holmes, A., Poline, J. B., Price, C. J., and Frith, C. D. 1996. Detecting activations in PET and fMRI: levels of inference and power., *NeuroImage* **4(3)**, 223–235.

Friston, K. J. 1997. Testing for anatomically specified regional effects., *Hum Brain Mapp* **5**, 133–136.

Friston, K. J., Ashburner, J. T., Kiebel, S. J., Nichols, T. E., and Penny, W. D. 2007. *Statistical Parametric Mapping: The Analysis of Functional Brain Images*, Elsevier/Academic Press, Amsterdam.

Jones, E., Oliphant, T., and Peterson, P. 2001. *SciPy: Open Source Scientific Tools for Python*, `http://www.scipy.org`.

Kautz, S. A., Feltner, M. E., Coyle, E. F., and Baylor, A. M. 1991. The pedalin technique of elite endurance cyclists: changes with increasing workload at constant cadence, *International Journal of Sport Biomechanics* **7**, 29–53.

Kiebel, S. J., Poline, J., Friston, K. J., Holmes, A. P., and 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, *Proceedings of the 23rd International Conference on Biomechanics in Sports*, 811–814.

Kundson, D. V. and Lindsey, C. 2014. Type I and Type II errors in correlations of various sample sizes, *Comprehensive Psychology* **3(1)**, Article 1.

Kruschke, J. K. 2013. Bayesian estimation supersedes the t test, *Journal of Experimental Psychology: General* **142(2)**, 573–603.

Lenhoff, M. W., Santer, T. J., Otis, J. C., Peterson, M. G., Williams, B. J., and Backus, S. I. 1999. Bootstrap prediction and confidence bands: a superior statistical method for analysis of gait data, *Gait and Posture* **9**, 10–17.

Neptune, R. R., Wright, I. C., and van den Bogert, A. J. 1999. Muscle coordination and function during cutting movements, *Medicine & Science in Sports & Exercise* **31(2)**, 294–302, data: `http://isbweb.org/data/rrn/`.

Nichols, T. E. and Holmes, A. P. 2002. Nonparametric permutation tests for functional neuroimaging a primer with examples, *Human Brain Mapping* **15(1)**, 1–25.

Pataky, T. C . 2012. One-dimensional statistical parametric mapping in Python, *Computer Methods in Biomechanics and Biomedical Engineering* **15(3)**, 295–301.

Pataky, T. C., Robinson, M. A., and Vanrenterghem, J. 2013. Vector field statistical analysis of kinematic and force trajectories., *Journal of Biomechanics* **46(14)**, 2394–2401.

Pataky, T. C. 2016. RFT1D: Smooth one-dimensional random field upcrossing probabilities in Python, *Journal of Statistical Software*, in press.

Pataky, T. C., Vanrenterghem, J., and 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., *Journal of Biomechanics* **48(7)**, 1277–1285.

Ramsay, J. O. and Silverman, B. W. 2005. *Functional Data Analysis*, Springer, New York.

Rencher, A. C. and Chistensen, W. F. 2012. *Methods of Multivariate Analysis*, Wiley, New Jersey.

Robinson, M. A., Vanrenterghem, J., and Pataky, T. C. 2015. Statistical parametric mapping for alpha-based statistical analyses of multi-muscle EMG time-series, *Journal of Electromyography and Kinesiology* **25(1)**, 14–19.

Schwartz, M. H., Trost, J. P., and Wervey, R. A. 2004. Measurement and management of errors in quantitative gait data, *Gait and Posture* **20(2)**, 196–203.

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., and Evans, A. C. 1999. Detecting changes in nonisotropic images, *Human Brain Mapping* **8**, 98–101.

Worsley, K. J., Taylor, J. E., Tomaiuolo, F., and 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.

Table 1: Dataset dimensionality terminology. An "nDmD" dataset contains an m-dimensional vector or tensor measured over an n-dimensional spatiotemporal domain.

		Example	
n	m	n domain	m domain
0	1	ground impact instant	knee flexion
0	3	instant of max propulsion	ground reaction force
1	1	gait cycle time	knee flexion
1	2	gait cycle time	left and right knee flexion
1	3	stance phase	ground reaction force
2	1	foot contact surface	pressure
3	1	femur	von Mises stress
3	6	femur	strain tensor

Table 2: Dataset overview. N is the total of number of trajectories. *Note: the Kautz et al. (1991) dataset contains 26 trials but six are duplicates. Variables: COP = center of pressure, EMG = electromyography, GRF = ground reaction force.

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 et al. (2008)	161	Walking	Kinematics	ncbi.nlm.nih.gov/pubmed/18565753

Table 3: Smoothness estimation results, summary across all datasets. FWHM values represent 1D residual smoothness (Appendix F) and higher FWHM values reflect smoother residuals. The FWHM represents 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 ± SD
Kinematics	10.8	33.1	67.0	34.7 ± 14.1
Forces	6.2	14.3	32.4	16.1 ± 7.1
EMG	7.2	11.8	15.6	12.1 ± 2.5
Overall	6.2	16.5	67.0	21.9 ± 13.6

Table 4: False positive results for two-sample t tests with N=10 per group; summary of results from Figs.5&6. The three FWHM columns represent the minimum, median, and maximum value observed across all datasets (Table **??**). 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	$p = 1.000$	$p = 0.999$	$p = 0.940$

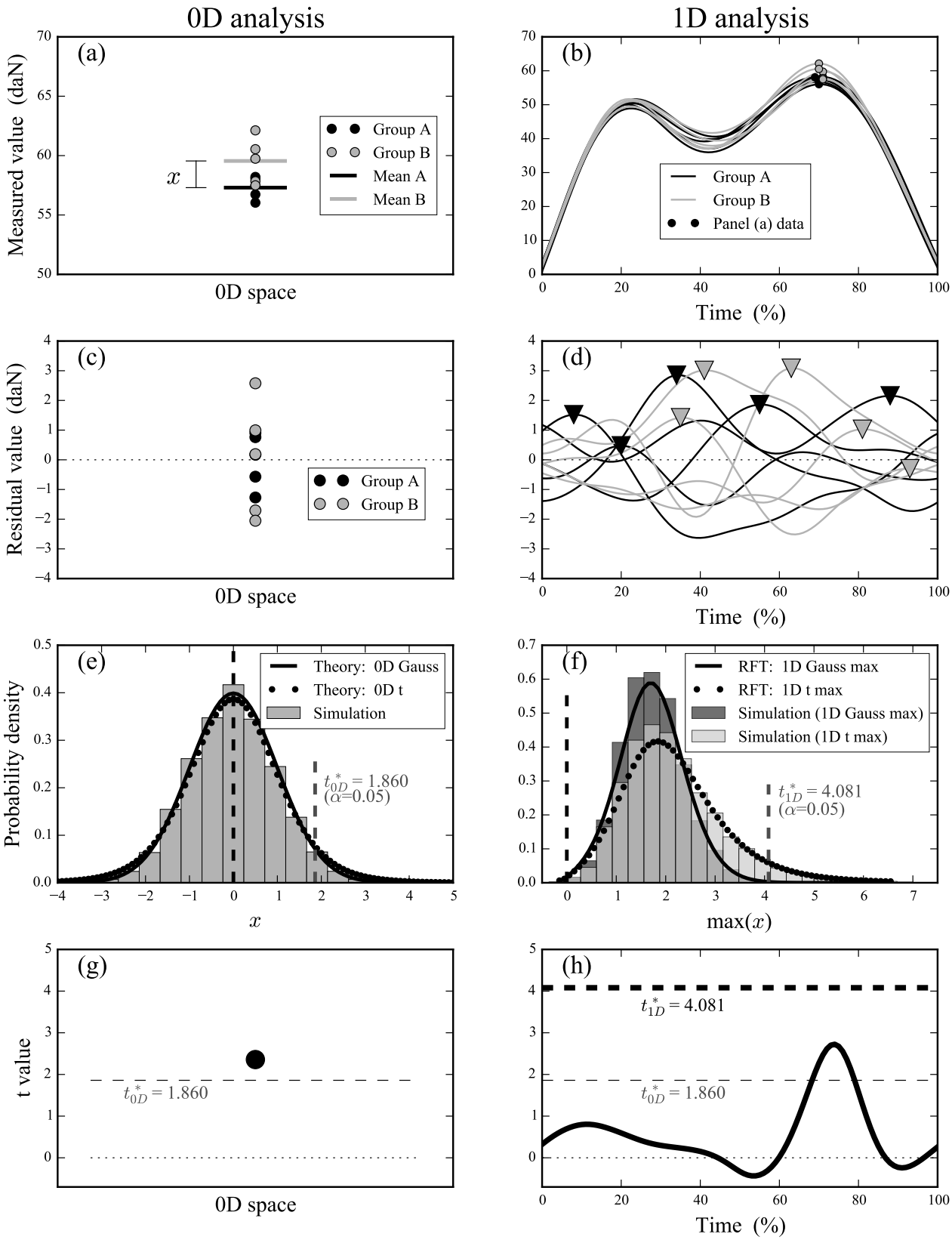

Figure 1: Overview of 0D vs. 1D statistical analyses. (a,b) 0D data extracted from 1D dataset (unit: decanewton). (c,d) Residuals with 1D residual maxima highlighted. (e,f) Theoretical and simulated probability density functions; 1D densities correspond to the 1D maximum (RFT= random field theory). Simulations 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.

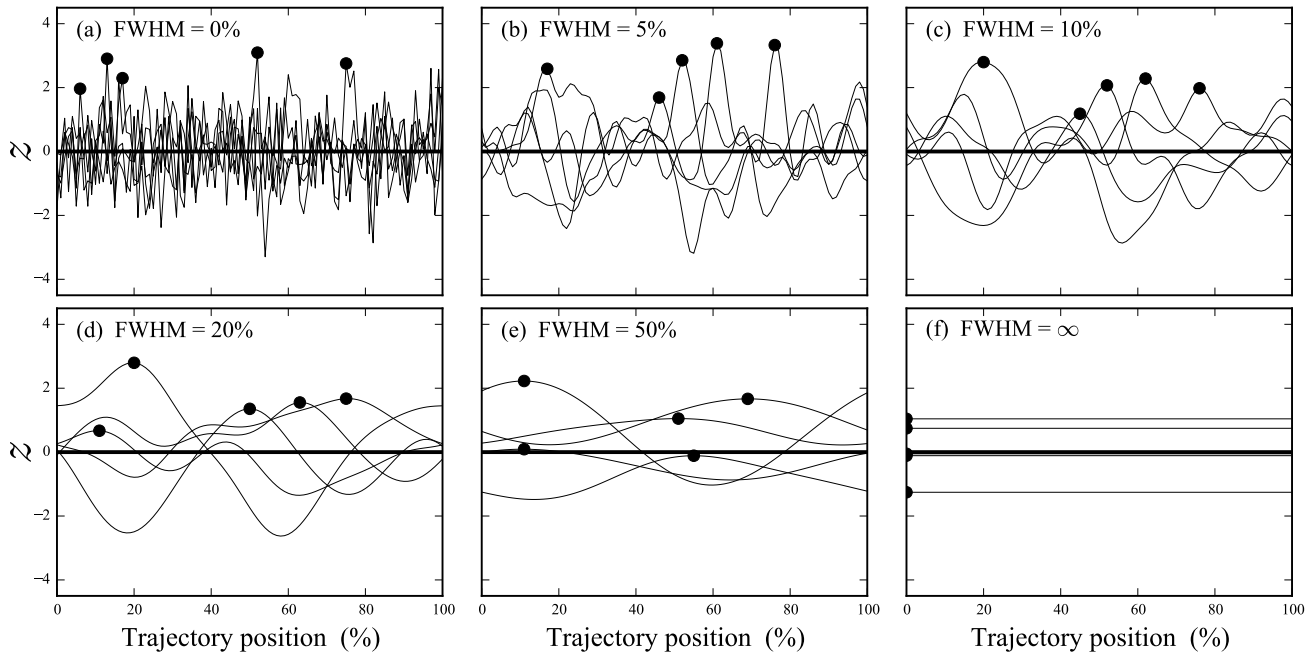

Figure 2: Simulated 1D Gaussian random fields. (a) Uncorrelated Gaussian data. (b-e) Data from panel (a) but smoothed with Gaussian kernels of varying FWHM (see Appendix A). Circles depict trajectory maxima whose FWHM-dependent probabilistic behavior Random Field Theory (RFT) describes (Adler and Taylor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

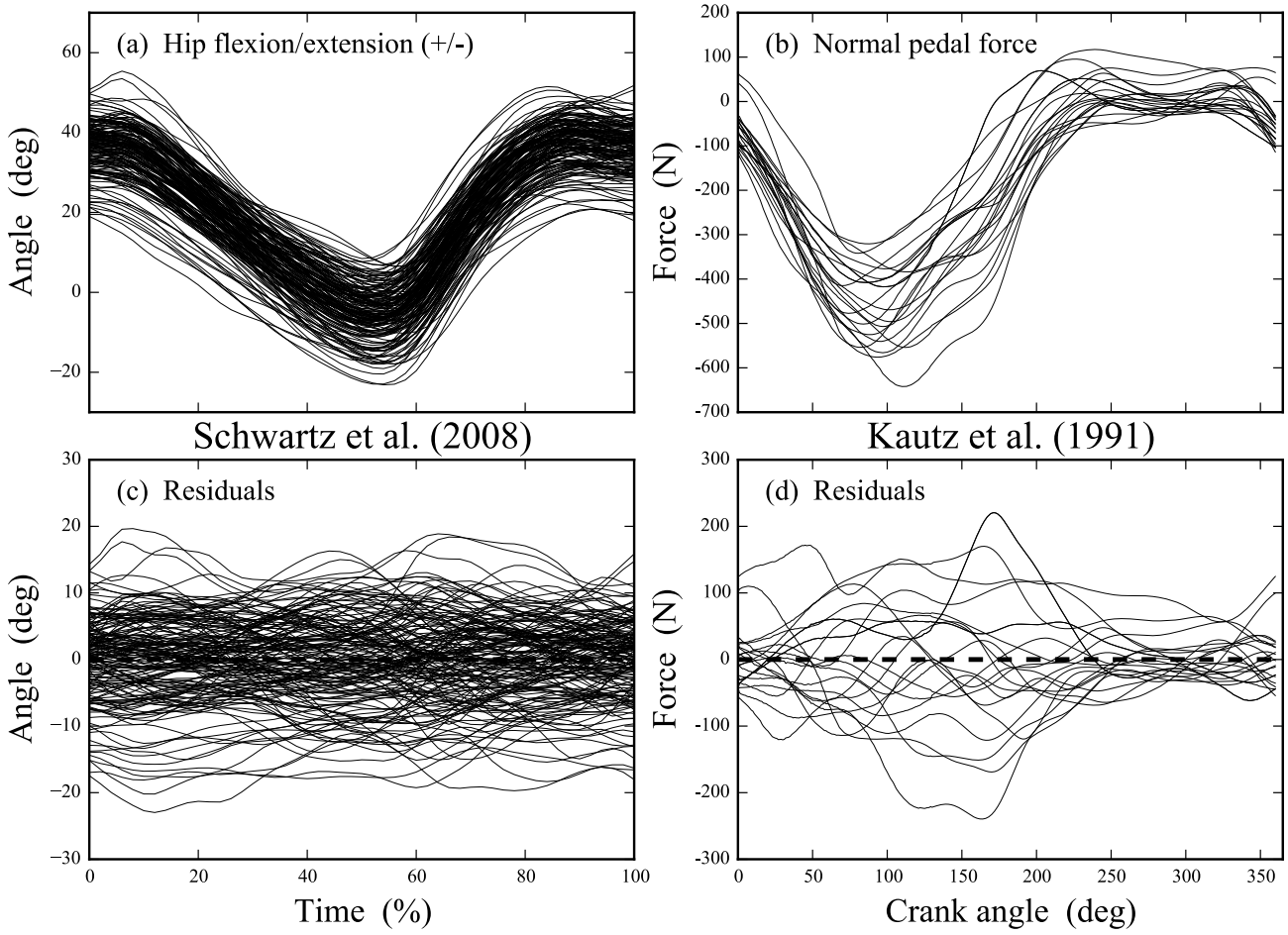

Figure 3: Example residuals from two public datasets. (a,b) Original data. (c,d) Residuals; here the differences between individual trajectories and the mean trajectory. The estimated smoothness values for (c) and (d) were FWHM=19.6% and FWHM=18.9%, respectively.

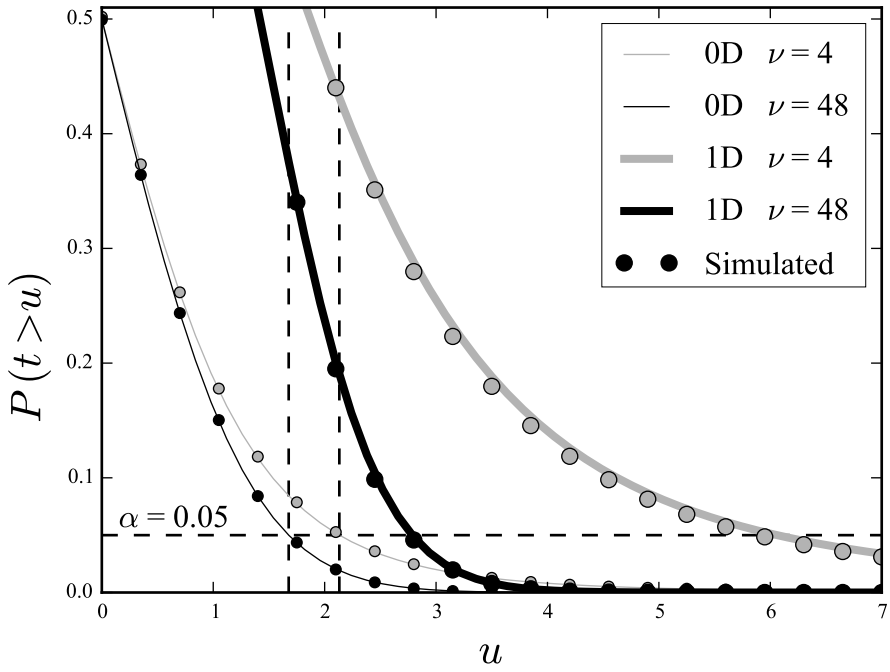

Figure 4: Survival functions for the *t* statistic for different degrees of freedom (*v*=4 and *v*=48) when the underlying data are (i) 0D Gaussian scalars (Eqn.1) or (ii) smooth 1D Gaussian trajectories with the median observed smoothness of FWHM=16.5% (Eqn.3). Solid lines depict theoretical probabilities 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 those 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).

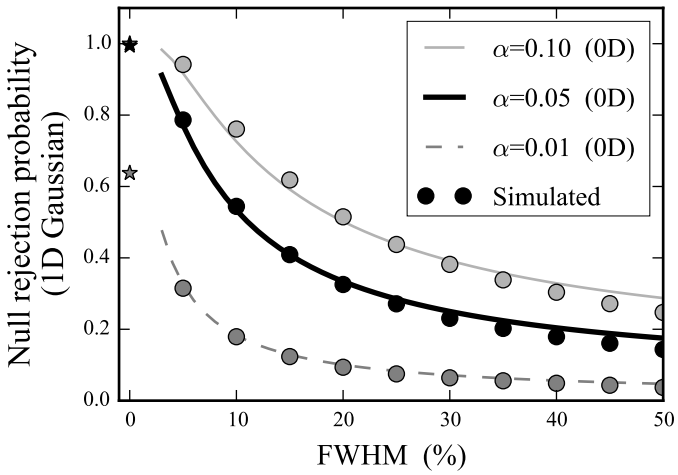

Figure 5: Probability of false positives for two-sample t tests (N=10) if using a 0D randomness model and if analyzing the maximum 1D t value. Star symbols (FWHM=0) depict the case of completely uncorrelated trajectories (Eqn.4).

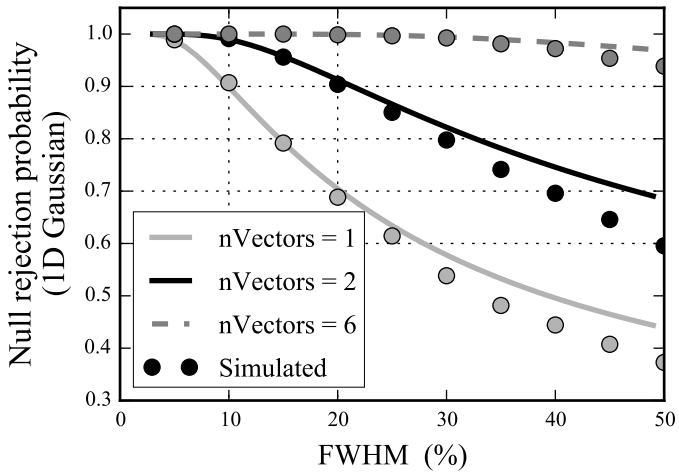

Figure 6: Probability of false positives for two-sample t tests (N=10) for three-component vectors with independently varying components.

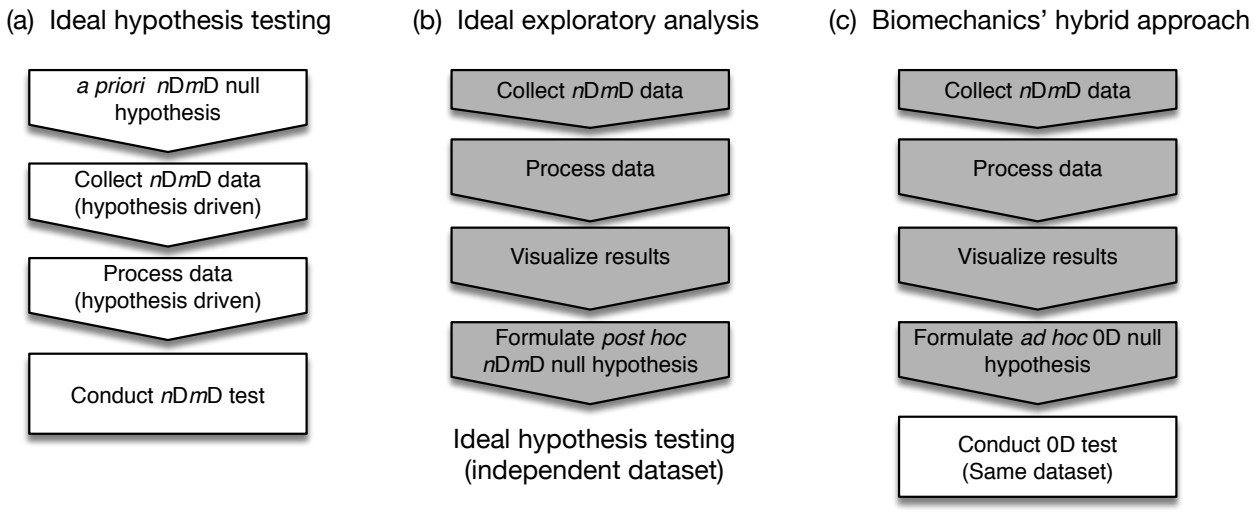

Figure 7: General procedures for hypothesis testing and exploratory analysis. In Biomechanics the dimensionalities of the test and the data are incommensurate.

Appendix A. Full width at half maximum (FWHM)

The FWHM parameter can be used to describe the smoothness of experimentally observed 1D residuals (Fig.1d, main manuscript). Most generally, the FWHM describes the shape of a Gaussian kernel (Fig.A1), which is typically defined as:

$$g(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{A.1}$$

Here μ and σ are its mean and standard deviation, respectively. Gaussian kernels can alternatively be expressed in terms of the FWHM (Fig.A2) through the following identity:

$$\text{FWHM} = 2\sigma\sqrt{2\log 2} \approx 2.4\sigma \tag{A.2}$$

The FWHM is somewhat more-intuitive than σ for describing 1D field smoothness because it is linked directly to kernel height: the kernel loses half of its maximum height over a distance of 0.5 FHWM units (Fig.A2). More specifically, the FWHM represents the width of a Gaussian kernel which, when convolved with uncorrelated Gaussian data (Appendix B), yields 1D Gaussian trajectories with the same smoothness as the observed 1D residuals.

Random field theory (RFT) (Adler & Taylor, 2007; Friston et al. 2007) regards experimentally observed 1D residuals as 1D random fields, and uses the estimated FWHM value to describe the probabilistic behavior of an infinite number of identically smooth fields. Thus, once one knows or estimates the FWHM, one can use RFT to calculate the maximum 1D differences / effect that random 1D fields would produce in arbitrary experiments.

While σ could be used in place of the FWHM to describe field smoothness, the main manuscript uses the FWHM parameter preferentially over σ because: (i) σ is typically used to represent population standard deviation, and (ii) the literature convention is to use FWHM (Friston et al. 2007).

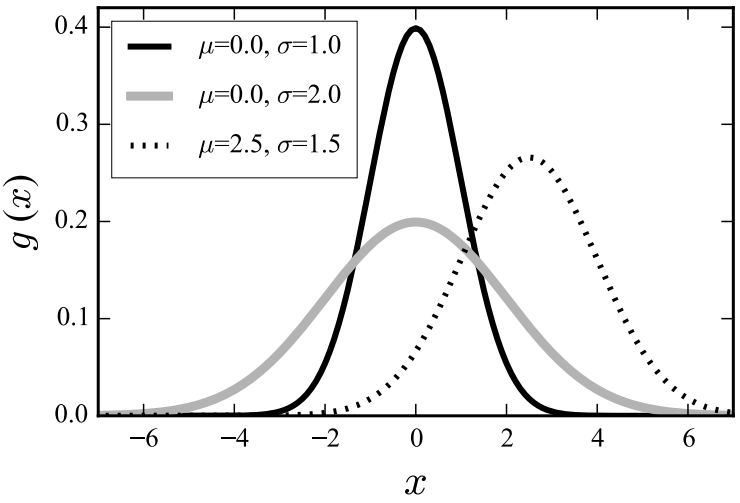

Figure A1: Gaussian kernels.

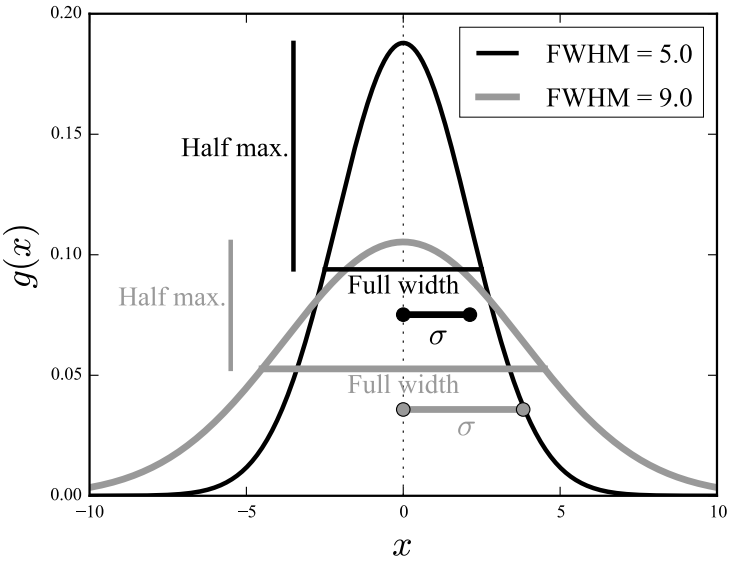

Figure A2: Breadth parameters for Gaussian kernels: σ and FWHM.

Appendix B. Convolution and random 1D Gaussian fields

This Appendix describes one procedure for generating smooth 1D Gaussian fields. Consider the functions $f(q)$ and $g(q')$ which are defined on one-dimensional (1D) domains q and q', respectively (Fig.B1a,b). Convolution is a procedure which slides $g(q')$ over $f(q)$ (Fig.B1c) to yield an 'overlapping area' function $h(q)$ (Fig.B1d). Convolving $f(q)$ with a Gaussian kernel — also called "Gaussian filtering" — yields a similar but smoother result (Fig.B2). In the context of the main paper the functions $f(q)$, $g(q')$ and $h(q)$ represent the experimental data, a smoothing kernel, and the smoothed data, respectively.

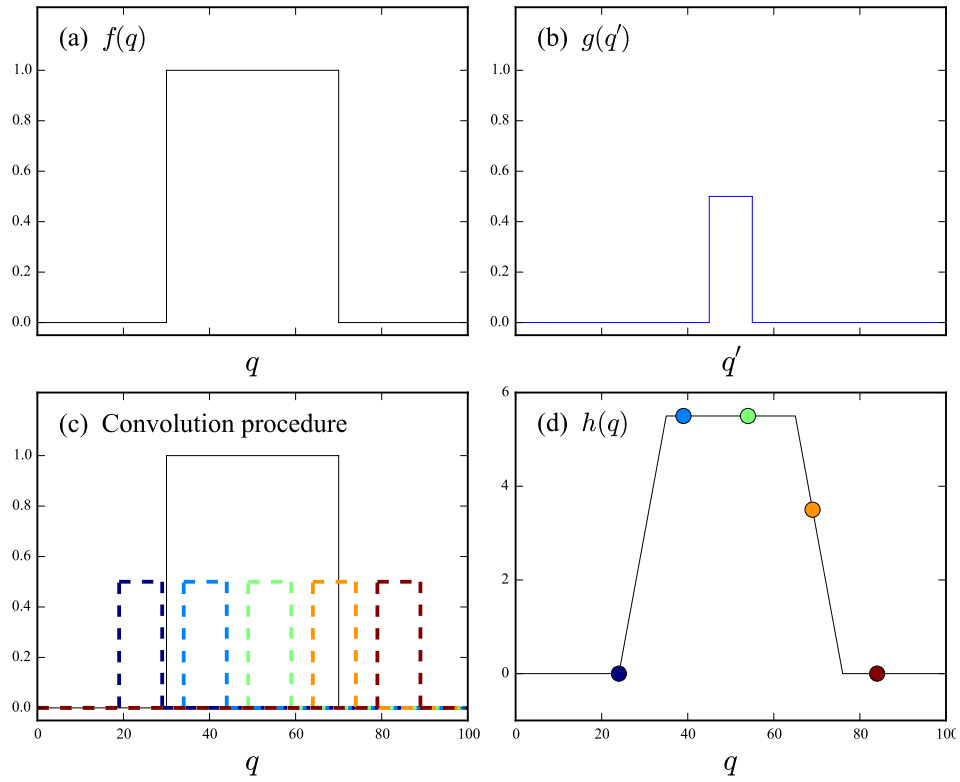

Figure B1: Convolution of two square waves. (a) Stationary function. (b) Moving function. (c) Depiction of $g(q')$ moving across $f(q)$. (d) Convolution result: colored circles depict the overlapping area between $f(q)$ and $g(q')$ when the right edge of $g(q')$ reaches the position q.

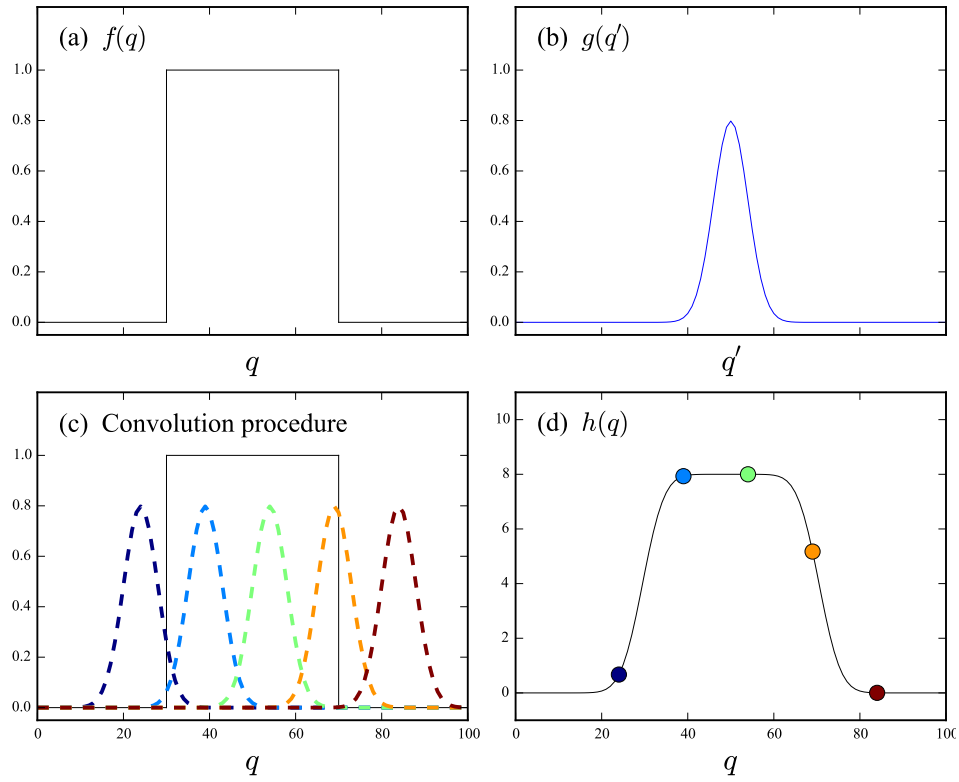

Figure B2: Convolution of a square wave with a Gaussian pulse.

When the data $f(q)$ are more like an experimental time series but consist of completely uncorrelated random Gaussian values (Fig.B3a) then convolving with a Gaussian kernel (Fig.B3b,c) yields a smooth Gaussian random field (Adler & Taylor, 2007) (Fig.B3d). The broader the smoothing kernel, the smoother the resulting random field (Fig.B4). Kernel breadth is parameterized by its full-width-at-half-maximum (FWHM) (Appendix A) and the FWHM parameter is central to 1D probability results (Friston et al. 2007).

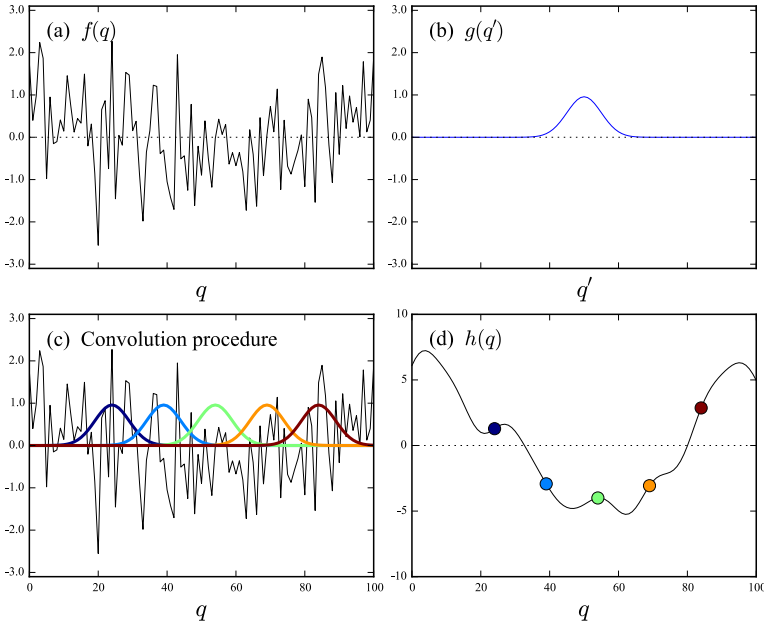

Figure B3: Convolution of uncorrelated Gaussian data (a) with a Gaussian kernel (b–c) yields a Gaussian random field (d).

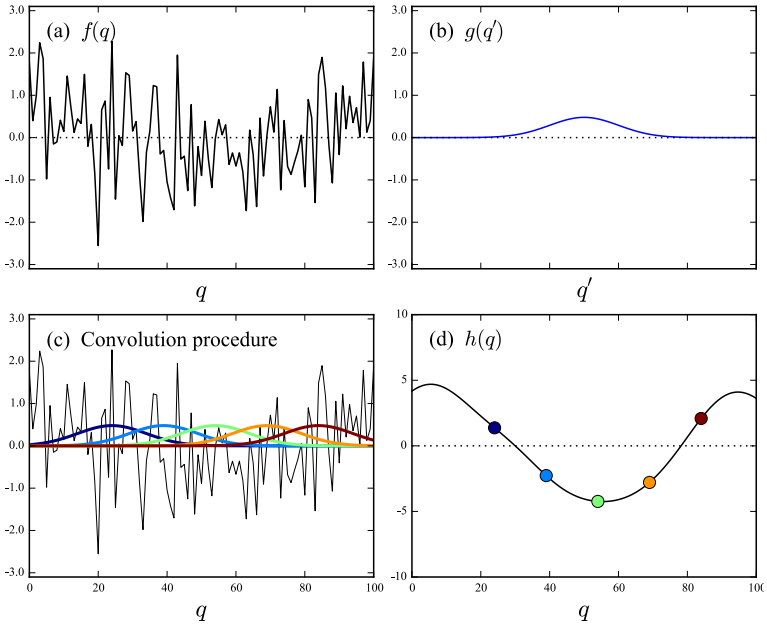

Figure B4: Identical to Fig.B3, but with a broader kernel (b) which yields a smoother random field (d).

Appendix C. MATLAB functions and scripts

The attached MATLAB functions and scripts (Table C1) replicate the paper's main results. Note that these MATLAB files are meant primarily for exploring concepts of 1D analyses through Random Field Theory (RFT) (Adler and Taylor, 2007). We encourage readers interested in actually conducting 1D analyses to access our free-and-open-source software at `www.spm1d.org`. The **spm1d** software package contains 1D functionality for standard tests including t tests, regression and ANOVA, with multivariate and repeated-measures functionality currently in-development.

Table C1: Overview of attached MATLAB files.

Category	File name	Summary
Functions (1D)	`estimate_fwhm.m`	Unbiased estimates of 1D field smoothness
	`randn1d.m`	Random 1D Gaussian field generator
	`rft_Tsf.m`	RFT survival function for the t statistic
Functions (0D)	`spm_invBcdf.m`	Inverse β cumulative distribution function
	`spm_invTcdf.m`	Inverse t cumulative distribution function
	`spm_Tcdf.m`	Cumulative distribution function for the t statistic
Data	`Schwartz2008.mat`	[1]Public dataset from Schwartz et al. (2008)
Scripts (FWHM)	`sA0_fwhm_verbose.m`	Verbose smoothness estimation
	`sA1_fwhm.m`	Smoothness estimation using `estimate_fwhm`
	`sA2_fwhm_validate.m`	Validation of `estimate_fwhm` using known FWHM
Scripts (random)	`sB0_randn1d_verbose.m`	Verbose random field generation
	`sB1_randn1d.m`	Random field generation using `randn1d`
	`sB2_randn1d_meaning.m`	The meaning of 'Gaussian 1D random field'
Scripts (results)	`sC0_falsepos.m`	Verbose estimate of false positive rate.
	`sC1_validate_onesample.m`	Validation of `rft_Tsf.m` for one-sample tests
	`sC2_validate_twosample.m`	Validation of `rft_Tsf.m` for two-sample tests

The first three functions embody the core functionality of RFT-based 1D t tests. The second three functions embody 0D functionality as implemented in the open-source software package SPM8 (`www.fil.ion.ucl.ac.uk/spm`); MATLAB users who have access to the Statistics Toolbox may which to use the equivalent 0D three functions: "`betaincinv`", "`tinv`" and "`tcdf`", respectively. Neither MATLAB nor any other commercial software package of which

[1]Redistributed with author's permission.

we are aware implements the first three (1D) functions.

The attached dataset from Schwartz et al. (2008) is redistributed with permission of the author; if you use these data in your own work please cite the original Schwartz et al. (2008) article. This dataset is attached to demonstrate how we computed the paper's smoothness results (Appendix F). All other datasets can be accessed using the links provided in Table 1 (main manuscript).

Last, the scripts are grouped in three sets of files labeled "sA∗", "sB∗" and "sC∗" which correspond to smoothness, 1D randomness and false positive results, respectively. All scripts and functions are commented for clarity. If any section of any of these scripts and functions is unclear, please contact us for support through our site at `www.spm1d.org`.

Appendix D. Estimating 1D residual smoothness

This Appendix summarizes the smoothness estimation procedures of Kiebel et al. (1999). While that paper describes the procedure for 3D data, the procedure is conceptually identical for 1D data. The 1D domain could be time, space or any other continuous variable, but for writing convenience we shall consider only 1D temporal trajectories.

The ultimate goal is to estimate the temporal smoothness of experimentally observed 1D residuals (Fig.1d, main manuscript) using a single scalar parameter: the FWHM (Appendix A). That single FWHM value represents the breadth of a Gaussian kernel which, when convolved with uncorrelated Gaussian time series (Appendix B) would yield random trajectories which have the same smoothness as the observed residuals. The procedure is depicted in Fig.D1 and is described in detail below.

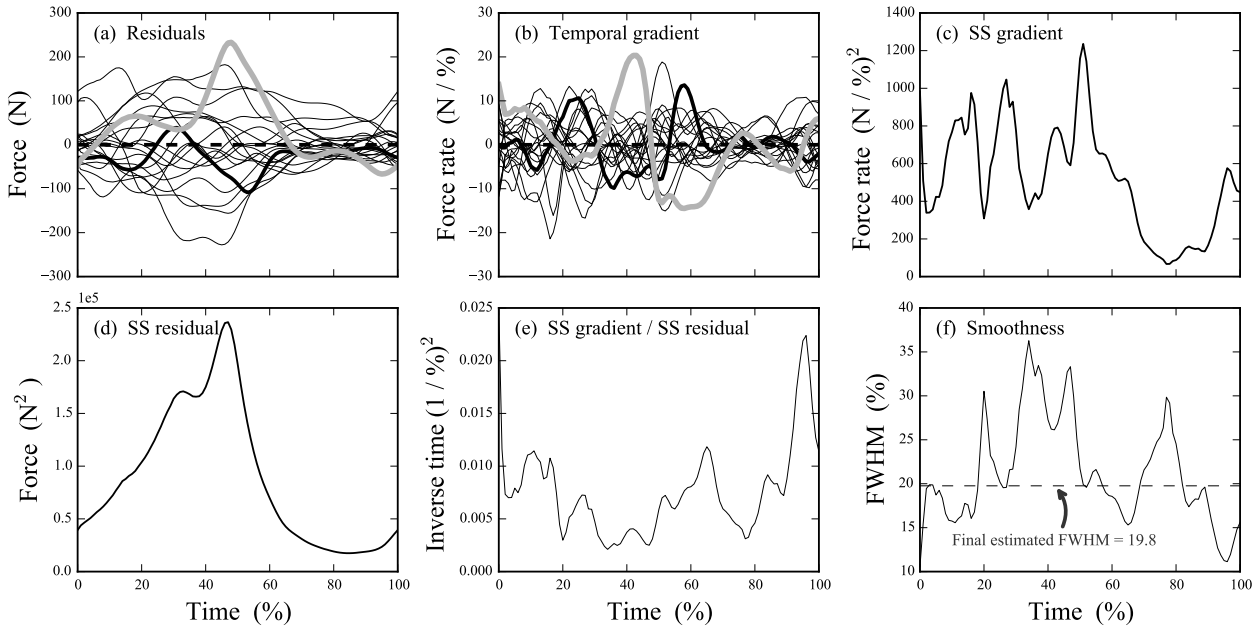

Figure D1: Overview of the FWHM estimation procedure of Kiebel et al. (1999). The residual data (a) are from the cycling normal pedal force of Kautz et al. (1991). SS = sum-of-squares.

Imagine that there are J residual trajectories which are each sampled at Q discrete time points and that the jth residual trajectory is denoted "$r_j(q)$". The first step is to compute the temporal gradient at each point q for each of the J trajectories (Fig.D1b):

$$r'_j(q) \equiv \frac{dr_j(q)}{dq} \tag{D.1}$$

The gradients can be estimated most easily using the differences between adjacent samples (i.e. $r_j(q+1) - r_j(q)$), but could also be done using alternative procedures. Practically, differences in gradient estimation procedures will likely have negligible effects on the ultimately estimated FWHM value. Regardless of the procedure, gradient estimation yields a total of J gradient trajectories.

Next, the true gradient magnitude at point q is estimated as the sum-of-squares of the observed gradients (Fig.D1c):

$$SS[r'](q) = \sum_{j=1}^{J} \Big(r'_j(q) \Big)^2 \tag{D.2}$$

In order to normalize across datasets and experiments the sum-of-squared residual values is also needed (Fig.D1d):

$$SS[r](q) = \sum_{j=1}^{J} \Big(r_j(q) \Big)^2 \tag{D.3}$$

The estimated gradients are then normalized by the residual magnitudes (Fig.D1e):

$$\lambda(q) = \frac{SS[\lambda](q)}{SS[r](q)} \tag{D.4}$$

Last, the FWHM trajectory (Fig.D1f) is given as:

$$\mathrm{FWHM}(q) = \sqrt{\frac{4 \log 2}{\lambda(q)}} \tag{D.5}$$

where an unbiased estimate of the true FWHM is simply the mean of the FWHM trajectory:

$$\hat{\text{FWHM}} = \frac{1}{Q}\sum_{q=1}^{Q}\text{FWHM}(q) \tag{D.6}$$

Note that estimated FWHM values are generally different at each point q in the 1D field (Fig.D1f). Smoothness which is non-constant across the field is termed 'anisotropic'. In order to deal with this issue let's consider 'apparent' vs. 'real' anisotropy. If the anisotropy is merely apparent, then Eqn.D.6 is valid. To understand why, consider a 0D random variable x which comes from a population with constant variance. Although the population variance is constant, random samples of x will yield different sample variances. Nevertheless, sample variance is an unbiased estimate of the true population variance. Similarly, even when the true FWHM is constant, randomly sampled 1D data will yield sample FWHM estimates which vary not only from sample to sample but also from point to point in the 1D field. Thus the mean FWHM value is an unbiased estimate of the true population FWHM when the anisotropy is merely apparent. Note also this FWHM estimation procedure has been validated for 1D data elsewhere (Pataky, 2015).

'True anisotropy' is a theoretically less trivial problem which exists when different field regions actually do have different population-level smoothnesses. As an example, consider impacts in running ground reaction forces: the initial impact phase is generally associated with higher signal frequencies than the midstance and push-off phases, so in this situation the true population FHWM is likely different in the different phases. There is fortunately an easy solution to the problem (Worsley et al. 1999). If there are Q points in the field, one simply computes the field length Q' for which smoothness is isotropic. Since the procedures are validated in Worsley et al. (1999), and since this anisotropy correction has no effect on the main paper's conclusions regarding 0D vs. 1D false positives, we leave the issue of anisotropic smoothness for future projects where assuming isotropic smoothness may have less trivial effects on analyses' results.

Appendix E. Generating smooth 1D Gaussian trajectories

This Appendix describes how to generate smooth 1D Gaussian trajectories in MATLAB (The MathWorks, Natick, USA) by summarizing the mathematical theory and code of Penny (2008). Interested readers are encouraged to consult Penny (2008) for a variety of additional background details regarding Random Field Theory and its applications. The target audience for this Appendix is anyone with (a) familiarity with linear algebra basics and MATLAB programming, and (b) a desire to generate their own random 1D trajectories for validating and/or exploring Random Field Theory predictions.

Penny W (2008) Mathematics for Brain Imaging (Course Notes), Chapter 3: "Random Field Theory". Retrieved on 12 Feb 2015 from `http://www.fil.ion.ucl.ac.uk/~wpenny/mbi/`

Smooth Gaussian trajectories can be generated using the convolution procedure described in Appendix B but in most programming languages it is considerably easier to use a single matrix multiplication to achieve the same result:

$$\boldsymbol{y} = \boldsymbol{C}\boldsymbol{z} \tag{E.1}$$

where, if Q is the number of trajectory nodes:

- $\boldsymbol{y}$ is the $(Q \times 1)$ smooth 1D Gaussian trajectory
- $\boldsymbol{C}$ is the $(Q \times Q)$ convolution matrix which embodies the correlation between neighboring points in the 1D trajectory
- $\boldsymbol{z}$ is a $(Q \times 1)$ trajectory of uncorrelated Gaussian values (Fig.B3a)

In MATLAB $\boldsymbol{z}$ can be generated as follows:

```
>>  Q = 101;
>>  z = randn(Q,1);
```

The convolution matrix $\boldsymbol{C}$ requires a bit more work. First note that a Gaussian covariance function with unit power is defined as:

$$r(d) = \exp\left(-\frac{d^2}{2s^2}\right) \tag{E.2}$$

where:

- d is the distance between the current node and a reference node in the trajectory. If Q=101 then d=0.01 for adjacent nodes, and d=0.01n for two nodes which are n nodes apart.
- s is the trajectory smoothness which linearly scales with the FWHM (Appendix A) as:

$$FWHM = s(Q-1)\sqrt{4\log 2} \tag{E.3}$$

The convolution matrix $\boldsymbol{C}$ embodies this covariance (Eqn.E.2) simultaneously for all points in the trajectory. To assemble $\boldsymbol{C}$ in MATLAB first define the number of trajectory nodes and the smoothness:

```
>>  Q = 101;              % number of trajectory nodes
>>  FWHM = 10;         % smoothness (FWHM units)
>>  s = FWHM / (  (Q-1) * sqrt( 4*log(2) )  )   %smoothness (s=0.060 for FWHM=10)
```

and then build a distance matrix $\boldsymbol{D}$ as follows:

```
>>  dq = 1 / (Q -1);      % inter-node distance
>>  x = dq * (1:Q);         % trajectory position
>>  X = repmat(x, Q, 1);  % temporary holder of trajectory positions
>>  D =  X - repmat(x', 1, Q);   % matrix of distances from diagonal nodes
```

The (i, j)th element of the resulting matrix represents the distance between nodes i and j:

```
>>  D(1:5, 1:5)

ans  =

        0    0.0100    0.0200    0.0300    0.0400
   -0.0100         0    0.0100    0.0200    0.0300
   -0.0200   -0.0100         0    0.0100    0.0200
   -0.0300   -0.0200   -0.0100         0    0.0100
   -0.0400   -0.0300   -0.0200   -0.0100         0
```

Next apply Eqn.E.2 to yield the covariance matrix $\boldsymbol{A}$:

```
>>  A = exp( -0.5 * D.^2 / s^2);
```

This covariance matrix can be visualized using the 'pcolor' command (Fig.E1):

```
>>  pcolor(A)
>>  colorbar
```

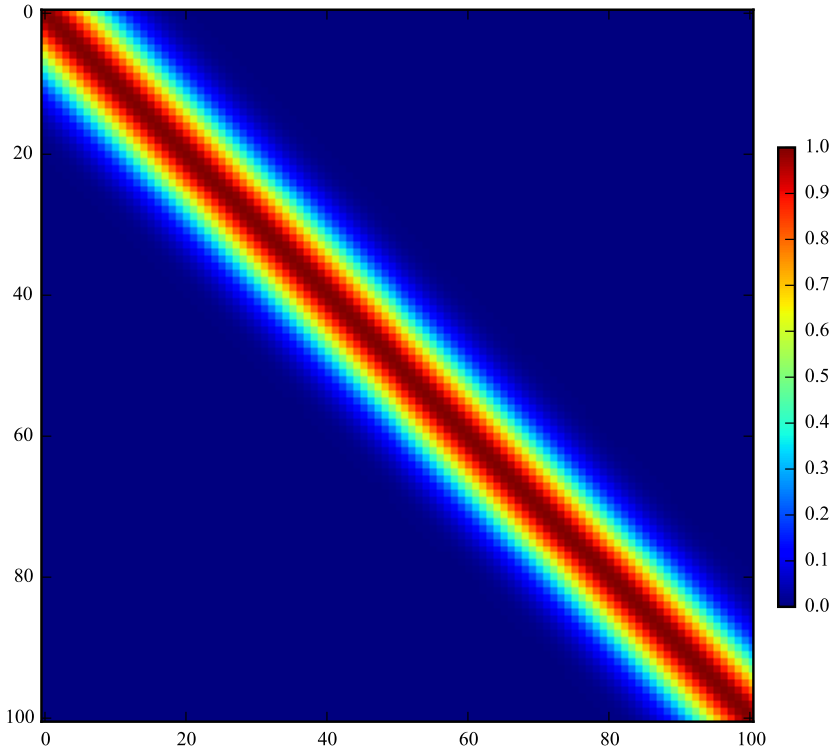

Figure E1: Visualization of the covariance matrix $\boldsymbol{A}$ for Q=101 and FWHM=10.0. The value of the (i, j)th matrix element represents the correlation values at 1D trajectory nodes i and j. When i=j the correlation is defined as 1.0, and since the data are smooth adjacent nodes' values are positively correlated. The farther apart the nodes the weaker their correlation.

Last, to generate data which have the covariance structure depicted above in Fig.E1, the convolution matrix $\boldsymbol{C}$ (Fig.E2) is assembled by first calculating the covariance matrix's eigen-solution:

```
>>  [V,U] =  eig(A);
```

where $\boldsymbol{V}$ and $\boldsymbol{U}$ are $\boldsymbol{A}$'s eigenvectors and eigenvalues, respectively, and then projecting the positive eigenvalues as follows:

```
>>  U = diag( U );
>>  U( U < 0 ) = 0;
>>  C = V * diag( sqrt( U ) ) * V' ;
```

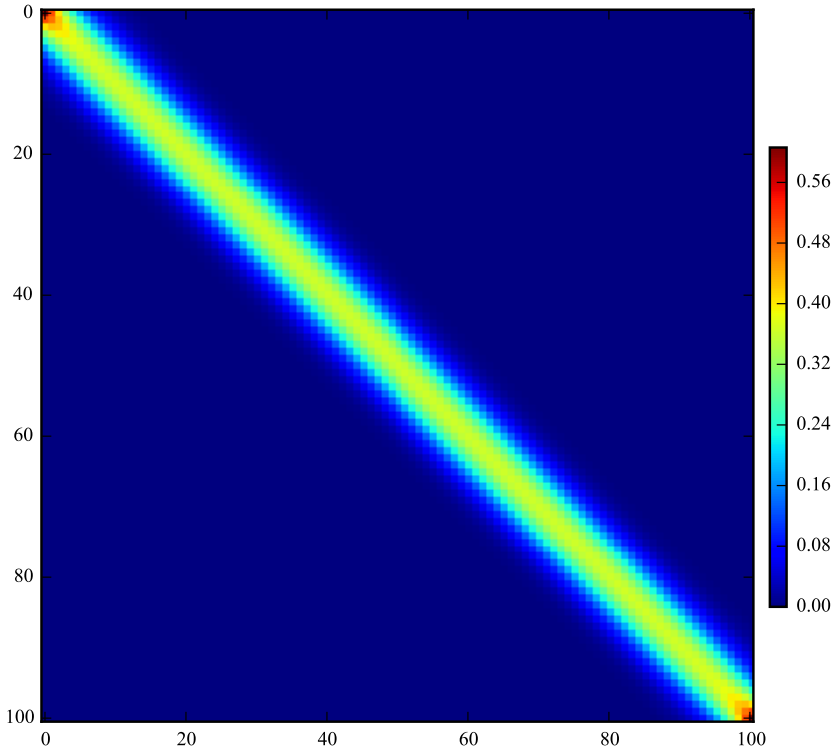

Figure E2: Visualization of the convolution matrix (Eqn.E.2) for Q=101 and FWHM=10.0.

This convolution matrix $\boldsymbol{C}$ needs to be computed only once for each smoothness value. It can then be used to generate single smooth Gaussian trajectories (Fig.E3) as follows :

```
>>  y0 = C * randn(Q,1);
>>  y1 = C * randn(Q,1);
>>  y2 = C * randn(Q,1);
>>  plot( y0 )
>>  plot( y1 )
>>  plot( y2 )
```

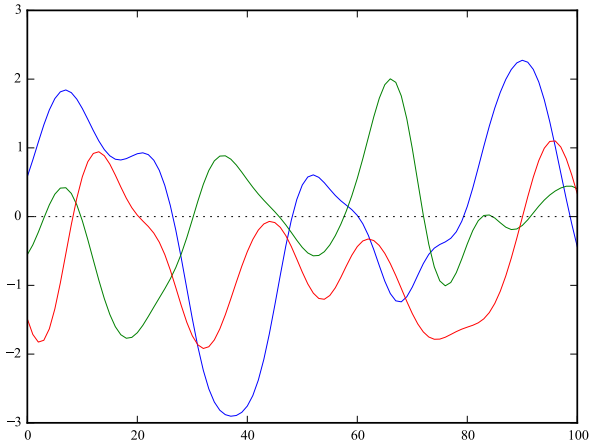

Figure E3: Three example 1D Gaussian trajectories.

Alternatively many random trajectories (Fig.E4) can be generated like this:

```
>>  J = 50;              % number of trajectories
>>  Z = randn(Q, J);     % uncorrelated Gaussian data
>>  Y = C * Z;           % (Q x J) array containing J separate random trajectories
>>  plot( Y )
```

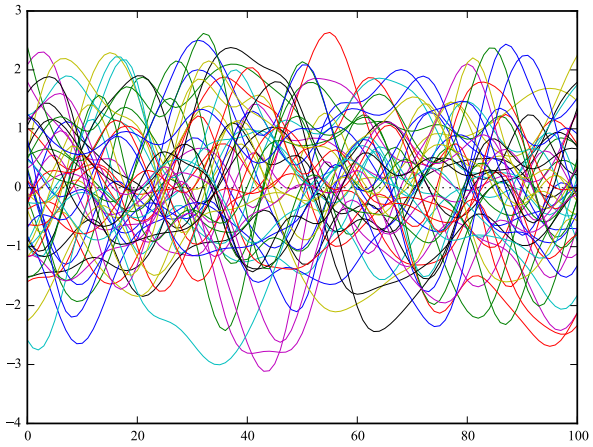

Figure E4: Fifty example 1D Gaussian trajectories.

For the analyses in the main manuscript we generated J=10 trajectories for each of two groups, computed the two-sample t trajectory using the usual definition of the two-sample t statistic, then repeated 100,000 times for each smoothness (FWHM) value. This yielded 100,000 t trajectories for each FWHM value. Last, we validated Eqns.3&4 (main manuscript) by extracting the maximum t value ($t_{\max}$) for each trajectory and then counting the number of $t_{\max}$ values which exceeded particular heights u.

Appendix F. FWHM estimation results

This Appendix lists the smoothness estimation results for all nine datasets (Table 1, main manuscript). First the 1D residuals for each dataset were calculated by subtracting the mean 1D trajectory as depicted in Fig.3, main manuscript. Next 1D residual smoothness was estimated as the FWHM (Appendix A) using the procedures of Kiebel et al. (1999) as summarized in Appendix D. Results are organized below into kinematic, force and EMG variables in Tables F1, F2 & F3, respectively.

Table F1: Smoothness estimates for the kinematics datasets. Smoothness was quantified using the FWHM (Appendix A)

Category	Source	Task	Variable	FWHM (%)
Joint rotations	Besier et al. (2009)	Walking	Hip flexion	51.5
			Knee flexion	32.6
			Angle dorsiflexion	30.8
	Besier et al. (2009)	Running	Hip flexion	64.4
			Knee flexion	33.6
			Angle dorsiflexion	33.5
	Neptune et al. (1999)	Cutting	Ankle supination/pronation	23.9
			Ankle dorsi/plantar flexion	36.4
			Knee extension/flexion	36.7
			Knee adduction/abduction	21.4
			Knee internal/external rotation	46.1
			Hip extension/flexion	40.9
			Hip adduction/abduction	33.6
			Hip internal/external rotation	10.8
	Schwartz et al. (2008)	Walking	Pelvic Up/Dn	14.2
			Pelvis Ant/Pst	33.5
			Pelvic Int/Ext	15.6
			Hip Flx/Ext	19.6
			Hip Add/Abd	13.9
			Hip Int/Ext	17.8
			Knee Flx/Ext	10.5
			Ankle Dor/Pla	9.2
			Foot Int/Ext	15.1
Center of pressure	Fregley et al. (2012)	Walking	Anterior/posterior	24.4
			Medial/lateral	65.2
Other	Kautz et al. (1991)	Cycling	Pedal angle	33.1
	Caravaggi et al. (2010)	Walking	Plantar arch angle	18.8

Table F2: Smoothness estimates for the dynamics datasets. Results from Pataky et al. (2008) are based on unsmoothed data.

Category	Source	Task	Variable	FWHM (%)
Ground reaction force	Dorn et al. (2012)	Running	Anterior / Posterior	8.8
			Medial / Lateral	9.5
			Vertical	11.1
	Fregley et al. (2012)	Walking	Anterior / Posterior	8.4
			Medial / Lateral	8.2
			Vertical	6.2
	Neptune et al. (1999)	Cutting	Vertical	11.9
	Pataky et al. (2008)	Walking	Vertical	6.2
Muscle forces	Besier et al. (2009)	Walking	Semimembranosus	16.4
			Semitendinosus	15.1
			Biceps femoris (long head)	16.7
			Biceps femoris (short head)	13.7
			Rectus femoris	9.3
			Vastus medialis	12.5
			Vastus intermedius	12.9
			Vastus lateralis	12.9
			Medial gastrocnemius	15.1
			Lateral gastrocnemius	13.4
	Besier et al. (2009)	Running	Semimembranosus	29.1
			Semitendinosus	29.3
			Biceps femoris (long head)	32.4
			Biceps femoris (short head)	27.8
			Rectus femoris	17.9
			Vastus medialis	20.5
			Vastus intermedius	21.6
			Vastus lateralis	21.8
			Medial gastrocnemius	25.0
			Lateral gastrocnemius	27.2
Joint implant forces	Fregley et al. (2012)	Walking	Knee: posterior-medial	14.7
			Knee: anterior-medial	11.7
			Knee: anterior-lateral	12.3
			Knee: posterior-lateral	11.9
Other	Kautz et al. (1991)	Cycling	Pedal normal force	19.8
			Pedal tangental force	15.3
			Crank torque	12.7

Table F3: Smoothness estimates for the EMG datasets. Results from Murley et al. (2014) are based on average time series and FWHM values were estimated relative to the average cross-task trajectory.

Source	Task	Variable	FWHM (%)
Murley et al. (2014)	Walking	Tibialis posterior	10.3
		Peroneus longus	10.8
		Tibialis anterior	7.2
		Medial gastrocnemius	7.9
Neptune et al. (1999)	Cutting	Vastus lateralis	13.0
		Rectus femoris	15.5
		Biceps femoris	15.6
		Medial hamstring	14.3
		Tibialis anterior	14.2
		Medial gastrocnemius	11.8
		Gluteus maximus	13.8
		Gluteus medius	11.8
		Adductor magnus	11.6
		Vastus medialis	12.8
		Peroneus longus	10.7