

1 **Title:**

2 A Tool to Explore Discrete-Time Data: The Time Series Response Analyser

3

4 **Name(s) of Author(s):**

5 Benjamin J. Narang<sup>1</sup>, Greg Atkinson<sup>2</sup>, Javier T. Gonzalez<sup>1</sup>, James A. Betts<sup>1</sup>

6 **Institutional Affiliation:**

7 <sup>1</sup>Department for Health, University of Bath, Bath, UK.

8 <sup>2</sup>School of Health and Life Sciences, Teesside University, Middlesbrough, UK.

9

10

11 **Running Head:**

12 Time Series Response Analyser

13

14 **Corresponding Author:**

15 James A. Betts, Department for Health, University of Bath, BA2 7AY, United Kingdom.

16 E-mail: j.betts@bath.ac.uk

**Abstract**

The analysis of time series data is common in nutrition and metabolism research for quantifying the physiological responses to various stimuli. The reduction of many data from a time series into a summary statistic(s) can help quantify and communicate the overall response in a more straightforward way and in line with a specific hypothesis. Nevertheless, many summary statistics have been selected by various researchers, and some approaches are still complex. The time-intensive nature of such calculations can be a burden for especially large datasets and may, therefore, introduce computational errors, which are difficult to recognize and correct. In this short commentary, we introduce a newly-developed tool that automates many of the processes commonly used by researchers for discrete-time series analysis, with particular emphasis on how the tool may be implemented within nutrition and exercise science research.

**Keywords**

Incremental area under the curve; time series data; temporal response; post-prandial.

## 30 Introduction

31 It is common practice within the field of nutrition and metabolism research to analyse serial  
32 measurements made over time to determine the temporal pattern of a given response. Typical examples include  
33 metabolic control following nutritional challenges (i.e. oral glucose or fat tolerance tests; Berthiaume & Zinker,  
34 2002), monitoring of stable isotope enrichment in various body pools and associated substrate kinetics  
35 (Garlick et al., 1989), and markers of physiological response to exercise such as heart rate and oxygen  
36 consumption (Gore & Withers, 1990).

37  
38 Such analyses have become increasingly complex and necessary in recent years both due to technical  
39 advancements in measurement tools and due to our growing understanding of the interactions between various  
40 nutritional stimuli. Regarding the former, it is undoubtedly a mark of progress that modern technologies have  
41 enabled many measurements to be made with higher sampling frequency and thus with greater sensitivity to  
42 rapidly fluctuating responses over time. However, such high-resolution temporal data also bring certain  
43 analytical challenges (such as the control of type I and II error rates due to the number of multiple comparisons),  
44 which can complicate the elucidation and communication of clear conclusions.

45  
46 While early studies in many areas of nutrition science may have examined simple comparisons of  
47 treatments (e.g. 20 g carbohydrate *versus* water/placebo at a single time-point), the state of current  
48 understanding in many areas is now such that further progress requires more sophisticated factorial designs  
49 with multiple levels within each factor, to examine longer term effects and/or interactions between ingredients  
50 that work in concert (e.g. pre-post response to carbohydrate *versus* carbohydrate-protein *versus* water/placebo,  
51 *etc.*). This further evolution is necessary to detect more subtle and/or context specific effects but, again,  
52 introduces additional complicating factors, such as the reduced statistical power associated with quantifying  
53 interactive effects between all the additional independent variables (e.g. a 3-way ANOVA: 3 conditions\*pre-  
54 post\*multiple time-points), along with the complications arising when the data violate the assumption of  
55 sphericity (Huck & Cormier, 1995).

57 In all the above cases, condensing the time series data down to a summary statistic can simplify the  
58 analysis by removing the temporal element. In the above example, the 3-way ANOVA with multiple  
59 comparisons at many time-points becomes a 2-way condition\*time (pre, post) analysis. Beyond these  
60 advantages in relation to statistical analyses, this approach of using summary statistics facilitates the clear  
61 communication of the main findings both in simple terms for the general public and with complete reporting  
62 of individual responses for the scientific community. For example, graphical presentation of time series data  
63 on a line graph does not readily allow for individual or paired responses to be plotted, whereas this consistency  
64 of observed responses is easily presented as a histogram showing individual summary statistics (**Figure 1**).  
65 Measures of central tendency certainly have a place to illustrate group effects on graphs and figures but  
66 individual responses to each experimental condition should still be presented, particularly when sample sizes  
67 are relatively small, to facilitate critical evaluation of data (Weissgerber et al., 2015).

68  
69 Despite the above benefits of summary statistics and the common use of time series experimental  
70 designs within the scientific literature, the general approaches and precise methods of analysis vary  
71 considerably between laboratories and experiments (Wolever, 2004; Matthan et al., 2016). In addition,  
72 calculations requiring multiple stages and various equations are time consuming and susceptible to human  
73 error. This short commentary introduces a downloadable spreadsheet, the Time Series Response Analyser  
74 (TSRA), designed specifically to automate and standardize many common processes, thus minimizing both  
75 the time spent analyzing data and the probability of computational errors. The TSRA is freely available under  
76 the 'Author Guidelines' section of the *IJSNEM* website  
77 (<https://journals.humankinetics.com/view/journals/ijsnem/ijsnem-overview.xml/>). This commentary will  
78 highlight a range of time series analysis procedures that can be computed with the tool, and briefly discuss  
79 their utility in the context of exercise and nutrition research.

### 81 **Area under the curve (AUC)**

82 The methodological approach to an AUC calculation is particularly variable (Wolever, 2004) and  
83 manual calculation is highly susceptible to human error. The AUC can be calculated using denominations of

84 the trapezoidal rule, where time series data are integrated to form a single value characterizing the overall  
85 response, representative of an area (e.g. blood glucose concentrations measured in  $\text{mmol}\cdot\text{l}^{-1}$  at serial time-  
86 points over a standard oral glucose tolerance test are expressed as the product of concentration and time;  
87  $\text{mmol}\cdot\text{l}^{-1}\cdot 120 \text{ min}$ ). **Figure 2** illustrates a range of AUC options, each of which is described in this section.  
88

89 Total AUC is the most straightforward approach, in which an area is calculated relative to the line  
90 representing an ordinate of zero (Matthews et al., 1990). This practice can provide a valid estimate of the  
91 overall exposure to the parameter of interest (i.e. including the value measured at baseline – e.g. if contrasting  
92 24 h plasma testosterone concentrations between males and females). However, by the same reasoning, total  
93 AUC can be limited by the variation commonly observed at baseline, despite the best efforts of researchers  
94 and participants to replicate experimental conditions (Altman, 1985). In cases where baseline differences are  
95 apparent and/or it is the response to a stimulus that is of primary interest, the incremental AUC relative to  
96 another nominal value (generally baseline) may be a more appropriate alternative (Wolever & Jenkins, 1986).  
97

98 Naturally, certain exposures can cause the dependent variable to drop below the value to which  
99 incremental AUC is being calculated. For example, the postprandial response to a standard oral glucose  
100 tolerance test is typically measured across two hours, as the blood glucose concentrations of healthy  
101 participants tend to return to baseline within this time period (Babraj et al., 2009). Therefore, the blood glucose  
102 concentrations of highly insulin sensitive individuals could feasibly fall below the value measured at baseline,  
103 which for an incremental AUC calculation provides multiple options for analysis. In this instance some  
104 researchers may choose to terminate the calculation at the time-point at which the measured value falls below  
105 the incremental reference value (Ha et al., 1992), while others will include any subsequent positive segments  
106 if the value returns above baseline. Within this latter approach, researchers could consider negative areas to  
107 equal zero (Hofman et al., 2004), or subtract them from the calculation (Gannon et al., 1989). It should be  
108 noted that, while the subtraction of negative areas follows the principle of mathematical integration, this  
109 process is rarely justified but may occasionally be applied in error. In theory, unless this subtractive process  
110 is rationalized, values representing AUC should always be positive. Moreover, some of the incremental AUC

111 variations can be applied to the nadir rather than the baseline value (Vorster et al., 1990), which may be of  
112 interest when variables tend to decrease in response to a stimulus, such as postprandial concentrations of non-  
113 esterified fatty acids (Bickerton et al., 2007), or the ‘hunger hormone’ ghrelin (le Roux et al., 2005).  
114 Alternatively, the AUC could be calculated relative to a pre-determined absolute value or clinical reference  
115 threshold that is indicative of a certain outcome (Monnier et al., 2003). It is beyond the scope of this  
116 commentary to discuss each of these methodologies in any greater detail as they ultimately depend on the  
117 context. Suffice to say, whilst some AUC calculations are relatively simple, others can become mathematically  
118 complex, particularly those that consider the intersection of certain thresholds. In these instances, the  
119 probability of conceptual and computational errors with manual calculations are increased, and the clarity with  
120 which the AUC values have been derived is reduced.

121 The TSRA generates AUC results from raw data consistently and instantaneously with a minimal risk  
122 of human error. The tool computes AUC for all treatments simultaneously and handles each of the  
123 aforementioned methodologies under the input of the user. In addition, the spreadsheet provides transparency  
124 by explicitly quantifying the segmental areas that combine to produce the chosen AUC (which can be valuable  
125 information in itself to retain some reference to the shape of the response curve despite reducing the individual  
126 time points into areas).

### 127 **Alternative summary statistics in discrete-time series analysis**

128 In addition to the AUC calculations computed by the TSRA, the peak and time-to-peak values for each trial  
129 are also included in the output. Errors and inconsistencies in the identification of these summary statistics are  
130 considerably less likely to occur when compared to AUC, as their definitions are more precise and their  
131 calculations are more straightforward. They can however be particularly informative within certain contexts,  
132 and they are therefore briefly discussed in this section. **Table 1** contains definitions, benefits, limitations and  
133 examples for each summary statistic included in the TSRA output.

134 *Peak*

138 Of the various alternative summary values that can describe a time series response, the absolute peak  
139 is an easily identifiable, interpretable and physiologically meaningful statistic. It is simply the highest value  
140 attained in the dependent variable across the time window through which it was measured. Therefore, rather  
141 than representing the totality of a response, as is the case with AUC, this value indicates the maximum  
142 *measured* value of the relevant outcome. Critically, this statistic should be determined separately for every  
143 distinct trial and individual, accepting that the peak value may occur at different time-points for different  
144 response curves. Thus, the contrast of maximum measured values cannot be ascertained from visual inspection  
145 of the data when plotted as a time series (i.e. it is possible that no single participant's maximum value occurred  
146 at the apex of the group mean line). The utility of a peak value during the response to a physiological challenge  
147 has been demonstrated in the diagnoses of various medical conditions such as growth hormone deficiency  
148 (Koppeschaar et al., 2004) and constitutional delay of puberty (Grinspon et al., 2010), and is practical in the  
149 application of diagnostic research due to the absence of any complex calculations. Despite the simplicity of  
150 this summary statistic representing a clear benefit of this approach, contextual limitations do exist. For  
151 example, measurement error is likely to be relatively high when a single data-point is used to summarize an  
152 overall response, and the accuracy is heavily influenced by the true location of a peak value relative to the  
153 frequency with which samples are collected (De Nicolao et al., 2000). The accuracy of this value may therefore  
154 be questioned when sampling frequency is insufficient and/or the random within-subjects variability or "noise"  
155 in the measurement of the dependent variable is high.

### 156 157 *Time-to-peak*

158 Alongside the reporting of the peak value, the time at which this peak occurs is typically reported and  
159 interpreted by authors. This "time-to-peak" summary statistic indicates the gradient of the response to the  
160 stimulus, demonstrating onset alongside magnitude. For example, both the AUC and peak values may be  
161 similar between treatments, yet the time-to-peak may still reveal important changes in the shape of the  
162 response curve (**Figure 3**). This may be useful when assessing the bioavailability of a nutrient or supplement,  
163 as it can indicate the net rate of appearance relative to an alternative condition (Matthews et al., 1990). For  
164 example, Vinson and Bose (1988) included a comparison of a time-to-peak summary statistic when

165 investigating ascorbic acid bioavailability, in response to the ingestion of equivalent doses of synthetic and  
166 naturally-occurring vitamin C. Importantly, unless a substance is not endogenously produced and maintains  
167 constant disappearance rates, or in the absence of isotopic tracer methodologies, this method provides fairly  
168 limited insight into substrate kinetics. However, the utility of the time-to-peak summary statistic as a  
169 diagnostic tool has been demonstrated in the context of insulin sensitivity. Specifically, risk-prediction models  
170 for prediabetes were shown to be reliably and independently enhanced by the addition of time-to-peak blood  
171 glucose concentration during an oral glucose tolerance test (Chung et al., 2017). Moreover, the use of this  
172 statistic in this context theoretically signified the early-phase insulin response, which may have provided  
173 additional mechanistic insight beyond alternative summary statistics (Cree-Green et al., 2018).

174  
175 A further application of time-to-peak has been to inform methodologies that seek to identify certain  
176 responses, such as the duration and sampling frequency of an oral fat tolerance test necessary to provide a  
177 holistic metabolic profile (Tentolouris et al., 2017). As with all considerations outlined in this paper, the  
178 precise calculations and reported outcomes should remain specific to the research question and will therefore  
179 depend heavily on the context in which time series data are being analyzed. Moreover, where the magnitude  
180 and/or timing of the peak is of interest, additional measurements should be taken throughout the time window  
181 within which it is expected to occur.

## 182 **Further considerations**

### 183 *Variability statistics*

184  
185 Another avenue for investigation of time series data is variability. For example, measures of variability  
186 in the continuous monitoring of glucose concentrations can be a useful parameter to describe glycemic control  
187 (Wijsman et al., 2013). A greater variability in glucose concentration could indicate a reduced ability to  
188 appropriately respond to nutritional stimuli, reflecting impaired homeostatic regulation and in the context of  
189 glucose metabolism, an increased risk of type-2 diabetes (Ceriello et al., 2008). Within this example, a variety  
190 of methods are available to characterize glycemic variability including overall standard deviation, standard  
191



deviation across fixed time windows (for variability changes across time), range, interquartile range, percentage coefficient of variation and time spent above/below certain thresholds (Akintola et al., 2015). Rodbard (2009) discussed these methods from a statistical standpoint and provided further context-specific options for alternative perspectives on time series data. Another context in which the variability in a measured marker is of interest within a certain time window is chronobiology. Whilst this is a particularly interesting avenue for time series data analysis in nutrition research, it is beyond the scope of the TSRA primarily because of the circular nature of chronobiological data measured over several biological rhythm periods. The intricacies of biological rhythm descriptions and summaries are discussed from a statistical perspective elsewhere (Landler, Ruxton & Malkemper, 2018). The appropriate application of variability statistics to time series data ultimately depends on the specific research question being addressed, and the information that each option can provide. Further key considerations may be the normality of data distribution, which can influence the appropriateness of certain measures of central tendency and variability, and the associated sensitivity of these approaches to more extreme values. The TSRA computes both the standard deviation and the coefficient of variation for each individual trial, and provides these simple variability statistics within the standard output. Alternative variability statistics are not calculated by the tool, as the provision of a finite number of complex options may influence the analytical approach taken by the user.

### *Missing values*

Missing values may be the result of missed or inappropriately handled samples, errors in a measurement technique or mistakes during data entry. These can be particularly common in time series data, as the probability of an error is increased when a large number of samples are collected (especially where humans and/or technology are involved!). Missing data pose a problem for the analysis of time series data as the intended temporal resolution within a given trial is transiently reduced. Key considerations include the amount, the pattern and the cause of missing data, each of which may influence the methods by which they are resolved. Regarding the cause, data could be *missing completely at random* (MCAR), where missing values are unrelated to any observed values and are therefore a totally random subset of the data. Alternatively, if missing values are related to observed data, or dependent on the unobserved values themselves, they are

219 considered to be *missing at random* (MAR) or *missing not at random* (MNAR), respectively (Little & Rubin,  
220 1987). Where data are MCAR, techniques typically aim to preserve the observed underlying parameters of  
221 the variables for which data are imputed (e.g. means, variances, covariances *etc.*). However, the systematic  
222 nature of data MAR and MNAR suggest potential bias may have been introduced in these parameter estimates  
223 due to the existence of the missing values. For example, if the accuracy of a measurement technique utilized  
224 during time series data collection is confounded outside a certain range, especially high and/or low values are  
225 likely to be missing more frequently, eliciting an unrepresentatively skewed distribution (an example of  
226 MNAR). Indeed, Bell, King and Fairclough (2014) demonstrated a greater level of bias in time series summary  
227 measures with data MAR or MNAR, compared with MCAR, using a simulated randomized controlled trial.  
228 Researchers are therefore recommended to identify the cause of missing time series data and handle this issue  
229 accordingly.

231 Individual time-points for continuous time series data are inherently not mutually exclusive, so it seems  
232 appropriate to estimate missing values using known data for a given trial. The precise method by which this  
233 process has been conducted may however be ambiguous. As AUC calculations follow the trapezoidal rule,  
234 this summary statistic would typically use simplistic linear interpolation to estimate missing values. Briefly,  
235 existing points either side of missing values are connected with a straight line, and these are imputed as a  
236 function of time using the resulting linear equation (**Figure 4A**). It should be noted that this approach has  
237 limitations, particularly if missing values occur where the true response is likely to have reached a peak, as a  
238 linear connection would undercut this value (**Figure 4B**). An alternative approach may be to fit a polynomial  
239 curve of appropriate order to the known data and impute missing values using the resulting polynomial  
240 equation. In the context of time series data, imputing missing values using alternative trials for the same  
241 treatment or the same individual are not recommended, as these approaches are likely confounded by inter-  
242 individual variability and the effect of treatments, respectively. For a comprehensive review of missing value  
243 handling in the context of randomized controlled trials in nutrition, the reader is directed to Li and Stuart  
244 (2019).

## 246 *Outliers*

247 Another contentious topic in the initial screening of data is the identification and subsequent handling  
248 of outliers. Outlier identification typically uses statistical approaches, such as Tabachnick and Fidell (2007)  
249 defining values  $\geq 3.29$  standard deviations above or below the mean as outliers (the probability of obtaining a  
250 true sample this extreme is 0.1%). However, similar to missing values, continuous time series data are unique  
251 in that an outlier may be identifiable by its magnitude in relation to the rest of the response curve. This  
252 viewpoint may however lead to the exclusion of certain values simply because they don't follow a relatively  
253 smooth pattern which, as measurement error is likely to exist in all samples, may be too subjective an approach.  
254 de Souza and colleagues (2015) advocate for data analyses to be conducted with and without suspected outliers,  
255 to assess whether the main analysis is robust to these extreme cases. Comprehensive reporting of this  
256 sensitivity analysis may then be the most transparent approach to the handling of outliers.

## 257

## 258 **Conclusion**

259 The TSRA has been specifically designed to speed up and standardize the calculation of summary  
260 statistics from time series data. Therefore, this tool can be used to validate calculations, and can then be cited  
261 in publications to provide transparency and to verify that the reported summary statistics are free from error.  
262 In turn, readers can have greater confidence in the reported conclusions.

273

274 **References**

- 275 Akintola, A. A., Noordam, R., Jansen, S. W., de Craen, A. J., Ballieux, B.E., Cobbaert, C. M., Mooijaart, S.  
276 P., Pijl, H., Westendorp, R. G., & van Heemst, D. (2015). Accuracy of continuous glucose monitoring  
277 measurements in normo-glycemic individuals. *PLoS One*, *10*(10), e0139973.
- 278 Altman, D. G. (1985). Comparability of randomised groups. *The Statistician*, *34*(1), 125-36.
- 279 [Babraj, J. A., Vollaard, N. B. J., Keast, C., Guppy, F. M., Cottrell, G., & Timmons, J. A. \(2009\). Extremely](#)  
280 [short duration high intensity interval training substantially improves insulin action in young healthy males.](#)  
281 [\*BMC Endocrine Disorders\*, \*9\*\(1\), 3.](#)
- 282 Bell, M. L., King, M. T., & Fairclough, D. L. (2014). Bias in area under the curve for longitudinal clinical  
283 trials with missing patient reported outcome data: summary measures versus summary statistics. *SAGE Open*,  
284 *4*(2), 1-12.
- 285 Berthiaume, N., & Zinker, B. A. (2002). Metabolic responses in a model of insulin resistance: comparison  
286 between oral glucose and meal tolerance tests. *Metabolism: Clinical and Experimental*, *51*(5), 595-8.
- 287 Bickerton, A. S. T., Roberts, R., Fielding, B. A., Hodson, L., Black, E. E., Wagenmakers, A. J. M., Gilbert,  
288 M., Karpe, F., & Frayn, K. N. (2007). Preferential uptake of dietary fatty acids in adipose tissue and muscle  
289 in the postprandial period. *Diabetes*, *56*(1), 168-76.
- 290 [Ceriello, A., Esposito, K., Piconi, L., Ihnat, M. A., Thorpe, J. E., Testa, R., Boemi, M., & Giugliano, D. \(2008\).](#)  
291 [Oscillating glucose is more deleterious to endothelial function and oxidative stress than mean glucose in](#)  
292 [normal and type 2 diabetic patients. \*Diabetes\*, \*57\*\(5\), 1349-54.](#)
- 293 Chung, S. T., Ha, J., Onuzuruike, A. U., Kasturi, K., Galvan-De La Cruz, M., Bingham, B. A., Baker, R. L.,  
294 Utumatwishima, J. N., Mabundo, L. S., Ricks, M., Sherman, A. S., & Sumner, A. E. (2017). Time to glucose  
295 peak during an oral glucose tolerance test identifies prediabetes risk. *Clinical Endocrinology*, *87*(5), 484-91.
- 296 Cree-Green, M., Xie, D., Rahat, H., Garcia-Reyes, Y., Bergman, B. C., Scherzinger, A., Behn, C. D., Chan,  
297 C. L., Kelsey, M. M., Pyle, L., & Nadeau, K. J. (2018). Oral glucose tolerance test glucose peak time is most  
298 predictive of prediabetes and hepatic steatosis in obese girls. *Journal of the Endocrine Society*, *2*(6), 547-62.

- 299 de Souza, R. J., Eisen, R. B., Perera, S., Bantoto, B., Bawor, M., Dennis, B. B., Samaan, Z., & Thabane, L.  
300 (2016). Best (but oft-forgotten) practices: sensitivity analyses in randomized controlled trials. *American*  
301 *Journal of Clinical Nutrition*, 103(10), 5-17.
- 302 Di Nicolao, G., Liberati, D., & Sartorio, A. (2000). Stimulated secretion of pituitary hormones in normal  
303 humans: a novel direct assessment from blood concentrations. *Annals of Biomedical Engineering*, 28(9), 1136-  
304 45.
- 305 Gannon, M. C., Nuttall, F. Q., Westphal, S. A., Neil, B. J., & Seaquist, E. R. (1989). Effects of dose of ingested  
306 glucose on plasma metabolite and hormone responses in type II diabetic subjects. *Diabetes Care*, 12(8), 544-  
307 52.
- 308 Garlick, P. J., Wernerman, J., McNurlan, M. A., Essen, P., Lobley, G. E., Milne, E., Calder, G. A., & Vinnars,  
309 E. (1989). Measurement of the rate of protein synthesis in muscle of postabsorptive young men by injection  
310 of a 'flooding dose' of [ $1-^{13}\text{C}$ ] leucine. *Clinical Science (London, England: 1979)*, 77(3), 329-36.
- 311 Gonzalez, J. T., Veasey, R. C., Rumbold, P. L. S., & Stevenson, E. J. (2013). Breakfast and exercise  
312 contingently affect postprandial metabolism and energy balance in physically active males. *British Journal of*  
313 *Nutrition*, 110(4), 721-32.
- 314 Gore, C. J., & Withers, R. T. (1990). The effect of exercise intensity and duration on the oxygen deficit and  
315 excess post-exercise oxygen consumption. *European Journal of Applied Physiology*, 60(3), 169-74.
- 316 Grinspon, R. P., Ropelato, M. G., Gottlieb, S., Keselman, A., Martínez, A., Ballerini, M. G., Domené, H. M.,  
317 & Rey, R. A. (2010). Basal follicle-stimulating hormone and peak gonadotropin levels after gonadotropin-  
318 releasing hormone infusion show high diagnostic accuracy in boys with suspicion of hypogonadotropic  
319 hypogonadism. *The Journal of Clinical Endocrinology & Metabolism*, 95(6), 2811-18.
- 320 Ha, M. A., Mann, J. I., Melton, L. D., & Lewisbarned, N. J. (1992). Calculation of the glycemic index.  
321 *Diabetes, Nutrition & Metabolism*, 5(2), 137-9.
- 322 Hofman, Z., van Drunen, J. D. E., de Later, C., & Kuipers, H. (2004). The effect of different nutritional feeds  
323 on the postprandial glucose response in healthy volunteers and patients with type II diabetes. *European*  
324 *Journal of Clinical Nutrition*, 58(11), 1553-6.

- 325 Huck, S. W., & Cormier, W. H. (1995). *Reading statistics and research* (2nd ed.). New York, USA:  
326 HarperCollins.
- 327 Koppeschaar, H. P. F., Popovic, V., Leal, A., Otero, X. L., Torres, E., Paramo, C., Micic, D., Garcia-Mayor,  
328 R. V., Sartorio, A., Dieguez, C., & Casanueva, F. F. (2004). Growth hormone (GH) peaks versus areas under  
329 the curve in the diagnosis of adult GH deficiency: analysis of the variables provided by the GHRH + GHRP-  
330 6 test. *Pituitary*, 7(1), 15-20.
- 331 [Landler, L., Ruxton, G. D., & Malkemper, E. P. \(2018\). Circular data in biology: advice for effectively](#)  
332 [implementing statistical procedures. \*Behavioral Ecology and Sociobiology\*, 72, 128.](#)
- 333 le Roux, C. W., Patterson, M., Vincent, R. P., Hunt, C., Ghatei, M. A., & Bloom, S. R. (2005). Postprandial  
334 plasma ghrelin is suppressed proportional to meal calorie content in normal-weight but not obese subjects.  
335 *The Journal of Clinical Endocrinology & Metabolism*, 90(2), 1068-71.
- 336 Li, P., & Stuart, E. A. (2019). Best (but oft-forgotten) practices: missing data methods in randomized  
337 controlled nutrition trials. *American Journal of Clinical Nutrition*, 109(3), 504-8.
- 338 Little, R. J. A., & Rubin, D. B. (1987). *Statistical analysis with missing data*. New York, NY: John Wiley.
- 339 Matthan, N. R., Ausman, L. M., Meng, H., Tighiouart, H., Lichtenstein A. H. (2016). Estimating the reliability  
340 of glycemic index values and potential sources of methodological and biological variability. *American Journal*  
341 *of Clinical Nutrition*, 104(4), 1004-13.
- 342 Matthews, J. N., Altman, D. G., Campbell, M. J., & Royston, P. (1990). Analysis of serial measurements in  
343 medical research. *British Medical Journal*, 300(6719), 230-5.
- 344 Monnier, L., Lapinski, H., & Colette, C. (2003). Contributions of fasting and postprandial plasma glucose  
345 increments to the overall diurnal hyperglycemia of type 2 diabetic patients. *Diabetes Care*, 26(3), 881-5.
- 346 Rodbard, D. (2009). New and improved methods to characterize glycemic variability using continuous glucose  
347 monitoring. *Diabetes Technology & Therapeutics*, 11(9), 551-65.
- 348 Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5<sup>th</sup> ed.). Boston, MA: Allyn and Bacon.
- 349 Tentolouris, N., Kanellos, P. T., Siami, E., Athanasopoulou, E., Chaviaras, N., Kolovou, G., Sfikakis, P. P.,  
350 & Katsilambros, N. (2017). Assessment of the validity and reproducibility of a novel standardised test meal  
351 for the study of postprandial triacylglycerol concentrations. *Lipids*, 52(8), 675-86.

352 [Vinson, J., & Bose, P. \(1988\). Comparative bioavailability to humans of ascorbic acids alone of in Citrus](#)  
353 [extract. \*American Journal of Clinical Nutrition\*, 48\(3\), 601-4.](#)

354 Vorster, H. H., Venter, C. S., & Silvis, N. (1990). The glycaemic index of foods: a critical evaluation. *South*  
355 *African Journal of Clinical Nutrition*, 2(1), 13-17.

356 Weissgerber, T. L., Milic, N. M., Winham, S. J., & Garovic, V. D. (2015). Beyond bar and line graphs: time  
357 for a new data presentation paradigm. *PLoS Biology*, 13(4), e1002128.

358 Wijsman, C. A., van Heemst, D., Hoogeveen, E. S., Slagboom, P. E., Maier, A. B., de Craen, A. J. M., van  
359 der Ouderaa, F., Pijl, H., Westendorp, R. G. J., & Mooijaart, S. P. (2013). Ambulant 24-h glucose rhythms  
360 mark calendar and biological age in apparently healthy individuals. *Aging cell*, 12(2), 207-13.

361 Wolever, T. M. (2004). Effect of blood sampling schedule and method of calculating the area under the curve  
362 on validity and precision of glycaemic index values. *British Journal of Nutrition*, 91(2), 295-301.

363 Wolever, T. M., & Jenkins, D. J. A. (1986). The use of the glycemic index in predicting the blood glucose  
364 response to mixed meals. *American Journal of Clinical Nutrition*, 43(1), 167-72.

365

366

367

368

369

370

371

372

373

374

375

376

377

378

**Table 1.** Summary of the various summary statistics available in the output of the TSRA.

<u>Summary</u> <u>Statistic</u>	<u>Definition/Inference</u>	<u>Advantages</u>	<u>Limitations</u>	<u>Examples</u> <u>in</u> <u>Nutrition</u> <u>and</u> <u>Exercise Science</u>
<u>Area</u> <u>under the</u> <u>curve</u>	<u>A value</u> <u>representative of the</u> <u>magnitude of the</u> <u>total response to a</u> <u>stimulus across a</u> <u>given time period,</u> <u>calculated using the</u> <u>trapezoidal rule.</u>	<u>A single value that</u> <u>takes into account</u> <u>the two-</u> <u>dimensionality of</u> <u>time-series data (e.g.</u> <u>both the magnitude</u> <u>and the duration of</u> <u>the response are</u> <u>accounted for)</u>	<u>Inconsistent definitions</u> <u>throughout the literature</u> <u>Mathematical</u> <u>complexity increases</u> <u>probability of</u> <u>human/computational</u> <u>error</u>	<u>Blood glucose and</u> <u>insulin</u> <u>concentration</u> <u>responses to an</u> <u>oral glucose</u> <u>tolerance test</u> <u>Appetite hormone</u> <u>responses to</u> <u>certain meals</u>
<u>Peak</u>	<u>The maximum</u> <u>measured value</u> <u>attained in response</u> <u>to the stimulus.</u>	<u>Simple identification</u> <u>of the highest</u> <u>measured value</u> <u>Clearly indicative of</u> <u>the maximum</u> <u>instantaneous</u> <u>exposure to the</u> <u>stimulus</u>	<u>Validity dependent on</u> <u>measurement frequency</u> <u>relative to true peak,</u> <u>and error associated</u> <u>with the measurement</u> <u>technique</u>	<u>Diagnosis of</u> <u>diabetes during an</u> <u>oral glucose</u> <u>tolerance test</u> <u>Exogenous</u> <u>glucose oxidation</u> <u>rates during</u> <u>exercise, when</u> <u>comparing</u> <u>carbohydrate-</u> <u>based sports</u> <u>drinks</u>



<b><u>Time to Peak</u></b>	<u>The time taken to reach the maximum measured value. The onset of a given exposure.</u>	<u>Simple identification of the time at which the highest measured value was sampled</u> <u>May provide insight into the early-phase response to a stimulus</u>	<u>Validity dependent on measurement frequency relative to true peak, and error associated with the measurement technique</u> <u>Mechanistic inference may be confounded by contributing rates of appearance and disappearance</u>	<u>Early-phase insulin response to an oral glucose tolerance test</u> <u>Oxygen uptake kinetics at the onset of steady-state exercise</u> <u>Enhancing post-exercise glycogen resynthesis rates</u>
<b><u>Minimum</u></b>	<u>The minimum value attained in response to a stimulus.</u>	<u>Simple identification of the lowest measured value</u>	<u>Validity dependent on measurement frequency relative to true nadir, and error associated with the measurement technique</u>	<u>Analysis of variables that are known to decrease in response to a stimulus, such as plasma non-esterified fatty acid or glucagon-like peptide-1 responses to carbohydrate ingestion</u>
<b><u>Variability Statistics</u></b>	<u>The degree to which a measured marker</u>	<u>Calculations can be relatively straightforward (e.g.</u>	<u>Wide range of variability statistics available</u>	<u>Glycemic variability with continuous</u>

---

varies throughout a      standard deviation,      Susceptible to      glucose  
given period of time.      coefficient of      confounding by the      monitoring data  
variation etc.)      existence of outliers      Exercise intensity  
Provides insight into      variability during  
holistic homeostatic      endurance events  
control mechanisms      (e.g. heart rate or  
perceived exertion  
during a cycling  
road race)

---

379

380

381

382

383

384

385

386

387

388

389

390

391

392

393

394

395

396

## 397 **Figure Legends**

398  
399 **Figure 1.** 90-minute blood glucose concentration response to milkshake ingestion under two conditions  
400 (breakfast-rest vs. breakfast-exercise). Data are presented as individual measured responses across time (A),  
401 and using the incremental area under the curve (AUC) summary statistic displayed as mean  $\pm$  95% confidence  
402 intervals with individual measured responses (B). Real experimental data for nine participants extracted from  
403 Gonzalez et al. (2013).

404  
405 **Figure 2.** Illustrations of the range of area under the curve definitions used throughout the literature. See text  
406 for descriptions and examples for each.

407  
408 **Figure 3.** Hypothetical illustration of an individual measured response to a stimulus across time. The  
409 alternative measured responses on each panel demonstrate when area under the curve, peak and time-to-peak  
410 summary statistics all provide different inferences, requiring cautious and contextual interpretation

411  
412 **Figure 4.** Simple representation of linear interpolation to impute missing data (A), and a hypothetical time  
413 series response demonstrating a key limitation of linear interpolation (B).