

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

Benjamin J. Narang Greg Atkinson
University of Bath Teesside University

Javier T. Gonzalez and James A. Betts
University of Bath

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 data sets and may, therefore, introduce computational errors, which are difficult to recognize and correct. In this short commentary, the authors 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, postprandial, temporal response, time series data

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

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

While early studies in many areas of nutrition science may have examined simple comparisons of treatments (e.g., 20 g carbohydrate vs. water/placebo at a single time point), the state of current understanding in many areas is now such that further progress requires more sophisticated factorial designs with multiple levels within each factor, to examine longer term effects and/or interactions between ingredients that work in concert (e.g., pre-post

response to carbohydrate vs. carbohydrate-protein vs. water/placebo, etc.). This further evolution is necessary to detect more subtle and/or context-specific effects but, again, introduces additional complicating factors, such as the reduced statistical power associated with quantifying interactive effects between all the additional independent variables (e.g., a three-way analysis of variance: three Conditions \times Pre-Post \times Multiple time points), along with the complications arising when the data violate the assumption of sphericity ([Huck & Cormier, 1995](#)).

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

Despite the above benefits of summary statistics and the common use of time series experimental designs within the scientific literature, the general approaches and precise methods of analysis vary considerably between laboratories and experiments ([Matthan et al., 2016](#); [Wolever, 2004](#)). In addition, calculations requiring multiple stages and various equations are time consuming

Narang, Gonzalez, and Betts are with the Centre for Nutrition, Exercise & Metabolism, Department for Health, University of Bath, Bath, United Kingdom. Atkinson is with the School of Health & Life Sciences, Teesside University, Middlesbrough, United Kingdom. Betts (j.betts@bath.ac.uk) is corresponding author.

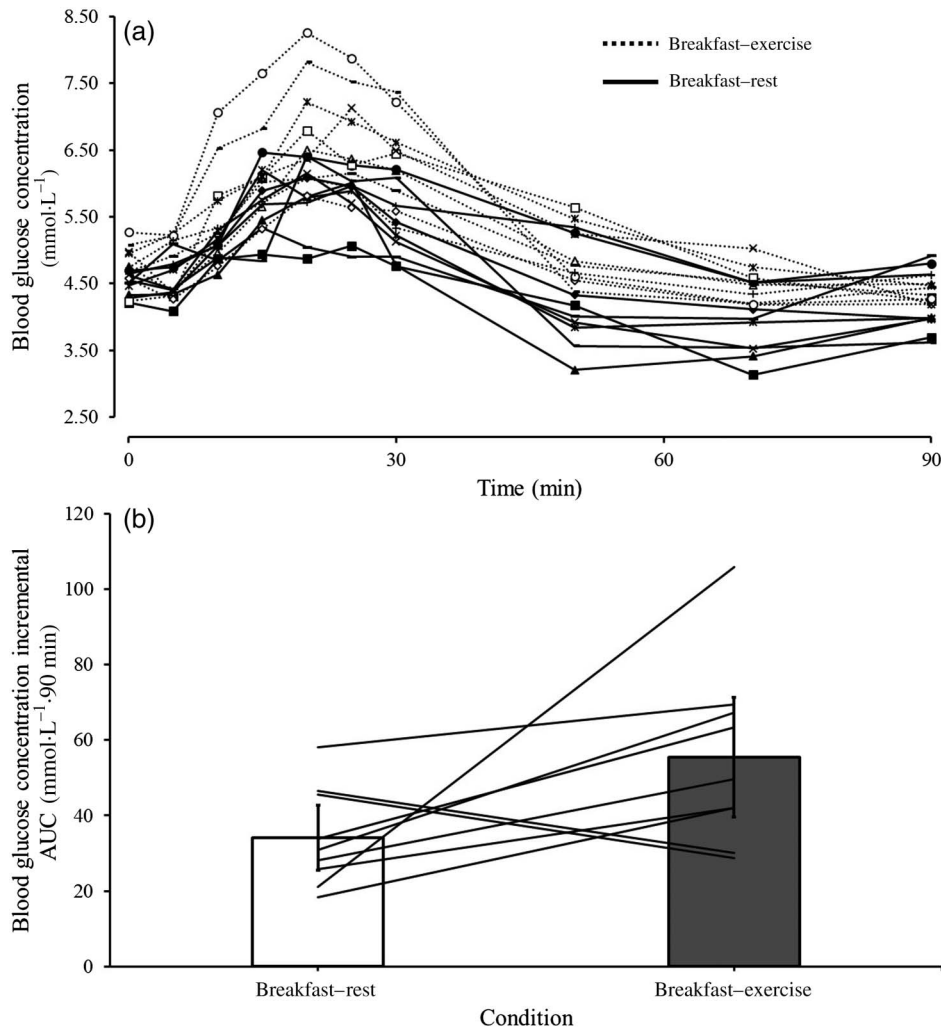


Figure 1 — The 90-min blood glucose concentration response to milkshake ingestion under two conditions (breakfast–rest vs. breakfast–exercise). Data are presented as individual measured responses across time (a) and using the incremental AUC summary statistic displayed as mean \pm 95% CI with individual measured responses (b). Real experimental data for nine participants extracted from Gonzalez et al. (2013). AUC = area under the curve; CI = confidence interval.

and susceptible to human error. This short commentary introduces a downloadable spreadsheet, the Time Series Response Analyser (TSRA) (Supplementary Material 1), designed specifically to automate and standardize many common processes, thus minimizing both the time spent analysing data and the probability of computational errors. The TSRA is freely available under the “Author Guidelines” section of the *International Journal of Sport Nutrition Exercise Metabolism* website (<https://journals.human-kinetics.com/view/journals/ijnsnem/ijnsnem-overview.xml/>). This commentary will highlight a range of time series analysis procedures that can be computed with the tool, and briefly discuss their utility in the context of exercise and nutrition research.

Area Under the Curve

The methodological approach to an area under the curve (AUC) calculation is particularly variable (Wolever, 2004) and manual calculation is highly susceptible to human error. The AUC can be calculated using denominations of the trapezoidal rule, where time

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

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

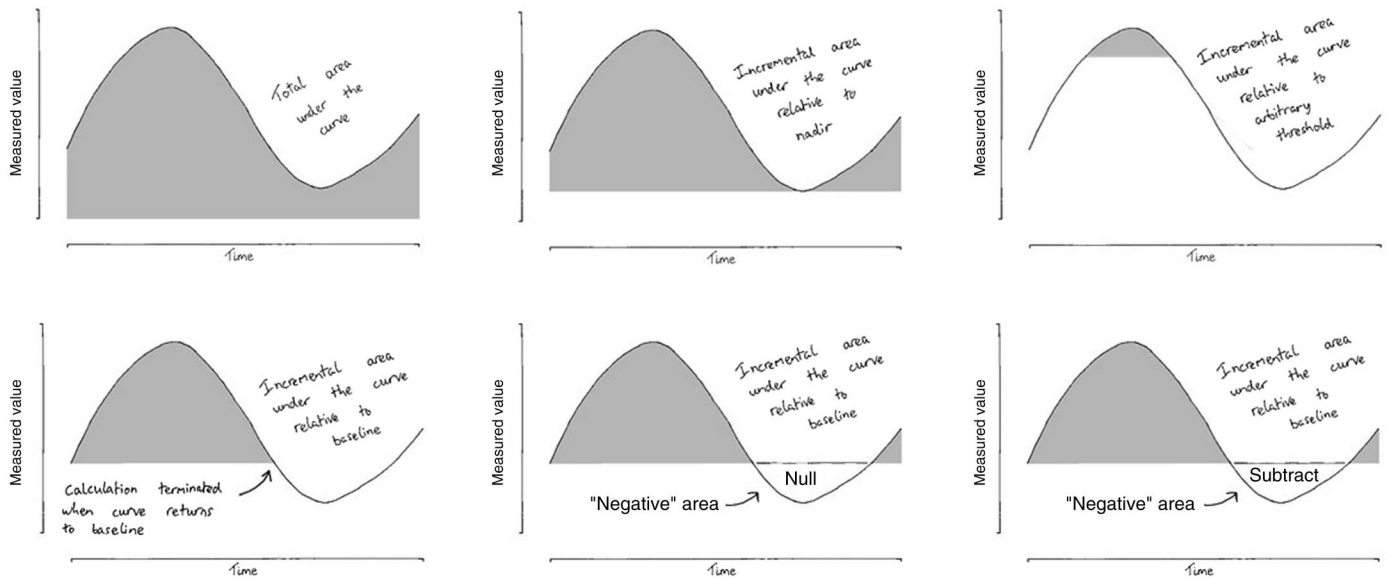


Figure 2 — Illustrations of the range of AUC definitions used throughout the literature. See text for descriptions and examples for each. AUC = area under the curve.

Naturally, certain exposures can cause the dependent variable to drop below the value to which incremental AUC is being calculated. For example, the postprandial response to a standard oral glucose tolerance test is typically measured across 2 hr, as the blood glucose concentrations of healthy participants tend to return to baseline within this time period (Babraj et al., 2009). Therefore, the blood glucose concentrations of highly insulin-sensitive individuals could feasibly fall below the value measured at baseline, which for an incremental AUC calculation provides multiple options for analysis. In this instance, some researchers may choose to terminate the calculation at the time point at which the measured value falls below the incremental reference value (Ha et al., 1992), while others will include any subsequent positive segments if the value returns above baseline. Within this latter approach, researchers could consider negative areas to equal zero (Hofman et al., 2004), or subtract them from the calculation (Gannon et al., 1989). It should be noted that, while the subtraction of negative areas follows the principle of mathematical integration, this process is rarely justified but may occasionally be applied in error. In theory, unless this subtractive process is rationalized, values representing AUC should always be positive. Moreover, some of the incremental AUC variations can be applied to the nadir rather than the baseline value (Vorster et al., 1990), which may be of interest when variables tend to decrease in response to a stimulus, such as postprandial concentrations of nonesterified fatty acids (Bickerton et al., 2007), or the “hunger hormone” ghrelin (le Roux et al., 2005). Alternatively, the AUC could be calculated relative to a predetermined absolute value or clinical reference threshold that is indicative of a certain outcome (Monnier et al., 2003). It is beyond the scope of this commentary to discuss each of these methodologies in any greater detail as they ultimately depend on the context. Suffice to say, while some AUC calculations are relatively simple, others can become mathematically complex, particularly those that consider the intersection of certain thresholds. In these instances, the probability of conceptual and computational errors with manual calculations are increased, and the clarity with which the AUC values have been derived is reduced.

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

Alternative Summary Statistics in Discrete Time Series Analysis

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

Peak

Of the various alternative summary values that can describe a time series response, the absolute peak is an easily identifiable, interpretable, and physiologically meaningful statistic. It is simply the highest value attained in the dependent variable across the time window through which it was measured. Therefore, rather than representing the totality of a response, as is the case with AUC, this value indicates the maximum *measured* value of the relevant outcome. Critically, this statistic should be determined separately for every distinct trial and individual, accepting that the peak value may occur at different time points for different response curves. Thus, the contrast of maximum measured values cannot be

Table 1 Summary of the Various Summary Statistics Available in the Output of the Time Series Response Analyser

Summary statistic	Definition/inference	Advantages	Limitations	Examples in nutrition and exercise science
Area under the curve	A value representative of the magnitude of the total response to a stimulus across a given time period calculated using the trapezoidal rule.	A single value that takes into account the two-dimensionality of time series data (e.g., both the magnitude and the duration of the response are accounted for)	Inconsistent definitions throughout the literature Mathematical complexity increases probability of human/computational error	Blood glucose and insulin concentration responses to an oral glucose tolerance test Appetite hormone responses to certain meals
Peak	The maximum measured value attained in response to the stimulus.	Simple identification of the highest measured value Clearly indicative of the maximum instantaneous exposure to the stimulus	Validity dependent on measurement frequency relative to true peak, and error associated with the measurement technique	Diagnosis of diabetes during an oral glucose tolerance test Exogenous glucose oxidation rates during exercise, when comparing carbohydrate-based sports drinks
Time-to-peak	The time taken to reach the maximum measured value. The onset of a given exposure.	Simple identification of the time at which the highest measured value was sampled May provide insight into the early phase response to a stimulus	Validity dependent on measurement frequency relative to true peak, and error associated with the measurement technique Mechanistic inference may be confounded by contributing rates of appearance and disappearance	Early phase insulin response to an oral glucose tolerance test Oxygen uptake kinetics at the onset of steady-state exercise Enhancing postexercise glycogen resynthesis rates
Minimum	The minimum value attained in response to a stimulus.	Simple identification of the lowest measured value	Validity dependent on measurement frequency relative to true nadir, and error associated with the measurement technique	Analysis of variables that are known to decrease in response to a stimulus, such as plasma nonesterified fatty acid or glucagon-like peptide-1 responses to carbohydrate ingestion
Variability statistics	The degree to which a measured marker varies throughout a given period of time.	Calculations can be relatively straightforward (e.g., standard deviation, coefficient of variation, etc.) Provides insight into holistic homeostatic control mechanisms	Wide range of variability statistics available Susceptible to confounding by the existence of outliers	Glycemic variability with continuous glucose monitoring data Exercise intensity variability during endurance events (e.g., heart rate or perceived exertion during a cycling road race)

ascertained from visual inspection of the data when plotted as a time series (i.e., it is possible that no single participant's maximum value occurred at the apex of the group mean line). The utility of a peak value during the response to a physiological challenge has been demonstrated in the diagnoses of various medical conditions, such as growth hormone deficiency (Koppeschaar et al., 2004) and constitutional delay of puberty (Grinspon et al., 2010), and is practical in the application of diagnostic research due to the absence of any complex calculations. Despite the simplicity of this summary statistic representing a clear benefit of this approach, contextual limitations do exist. For example, measurement error is likely to be relatively high when a single data point is used to summarize an overall response, and the accuracy is heavily influenced by the true location of a peak value relative to the frequency with which samples are collected (De Nicolao et al., 2000). The accuracy of this value may therefore be questioned when sampling frequency is insufficient and/or the random within-subjects variability or "noise" in the measurement of the dependent variable is high.

Time-to-Peak

Alongside the reporting of the peak value, the time at which this peak occurs is typically reported and interpreted by authors. This "time-to-peak" summary statistic indicates the gradient of the response to the stimulus, demonstrating onset alongside

magnitude. For example, both the AUC and peak values may be similar between treatments, yet the time-to-peak may still reveal important changes in the shape of the response curve (Figure 3). This may be useful when assessing the bioavailability of a nutrient or supplement, as it can indicate the net rate of appearance relative to an alternative condition (Matthews et al., 1990). For example, Vinson and Bose (1988) included a comparison of a time-to-peak summary statistic when investigating ascorbic acid bioavailability in response to the ingestion of equivalent doses of synthetic and naturally occurring vitamin C. Importantly, unless a substance maintains constant disappearance rates and is not endogenously produced, or in the absence of isotopic tracer methodologies, this method provides fairly limited insight into substrate kinetics. However, the utility of the time-to-peak summary statistic as a diagnostic tool has been demonstrated in the context of insulin sensitivity. Specifically, risk prediction models for prediabetes were shown to be reliably and independently enhanced by the addition of time-to-peak blood glucose concentration during an oral glucose tolerance test (Chung et al., 2017). Moreover, the use of this statistic in this context theoretically signified the early phase insulin response, which may have provided additional mechanistic insight beyond alternative summary statistics (Cree-Green et al., 2018).

A further application of time-to-peak has been to inform methodologies that seek to identify certain responses, such as the duration and sampling frequency of an oral fat tolerance test

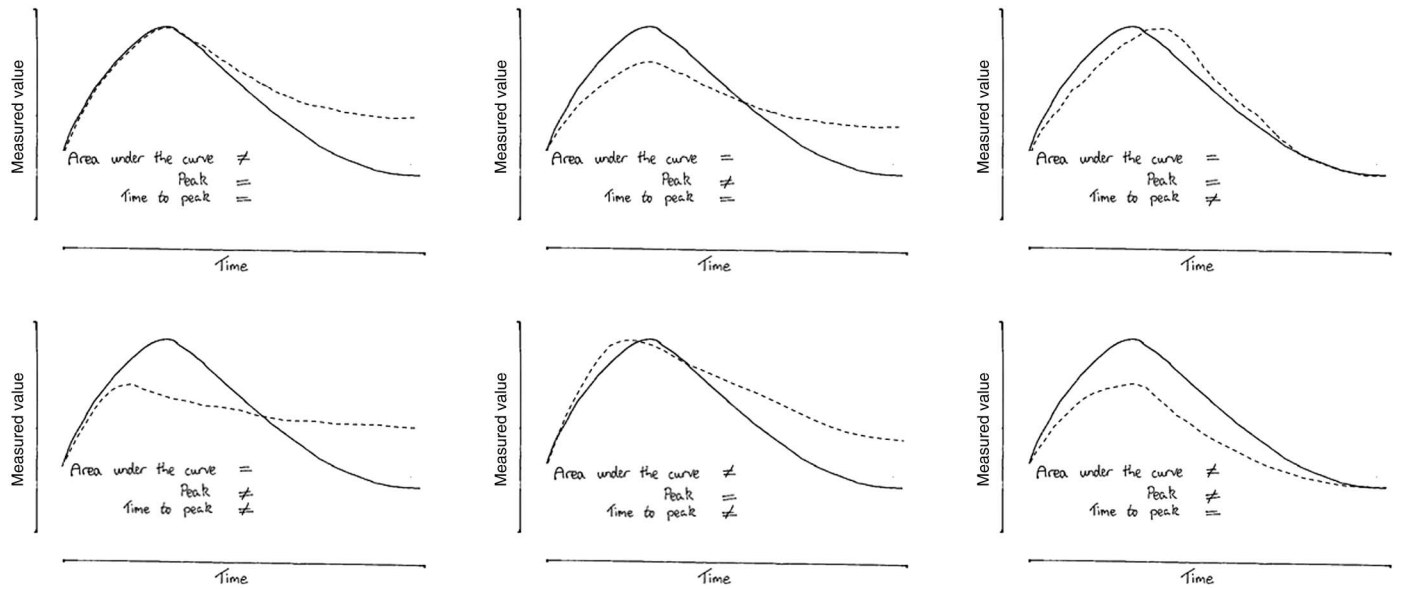


Figure 3 — Hypothetical illustration of an individual measured response to a stimulus across time (solid lines). The alternative measured responses on each panel (dashed lines) demonstrate when the area under the curve, peak, and time-to-peak summary statistics each provide different inferences, requiring cautious and contextual interpretation.

necessary to provide a holistic metabolic profile (Tentolouris et al., 2017). As with all considerations outlined in this paper, the precise calculations and reported outcomes should remain specific to the research question and will therefore depend heavily on the context in which time series data are being analyzed. Moreover, where the magnitude and/or timing of the peak is of interest, additional measurements should be taken throughout the time window within which it is expected to occur.

Further Considerations

Variability Statistics

Another avenue for investigation of time series data is variability. For example, measures of variability in the continuous monitoring of glucose concentrations can be a useful parameter to describe glycemic control (Wijsman et al., 2013). A greater variability in glucose concentration could indicate a reduced ability to appropriately respond to nutritional stimuli, reflecting impaired homeostatic regulation and in the context of glucose metabolism, an increased risk of Type 2 diabetes (Ceriello et al., 2008). Within this example, a variety of methods are available to characterize glycemic variability including overall SD , SD 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. While 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 et al., 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 SD 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*, 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 considered to be *missing at random* or *missing not at random* (MNAR), respectively (Little & Rubin, 1987). Where data are *missing completely at random*, techniques typically aim to preserve the observed underlying parameters of the variables for which data are imputed (e.g., means, variances, covariances, etc.). However, the systematic nature of data *missing at random* and MNAR suggest potential bias may have been introduced in these parameter estimates due to the existence of the missing values. For example, if the accuracy of a measurement technique

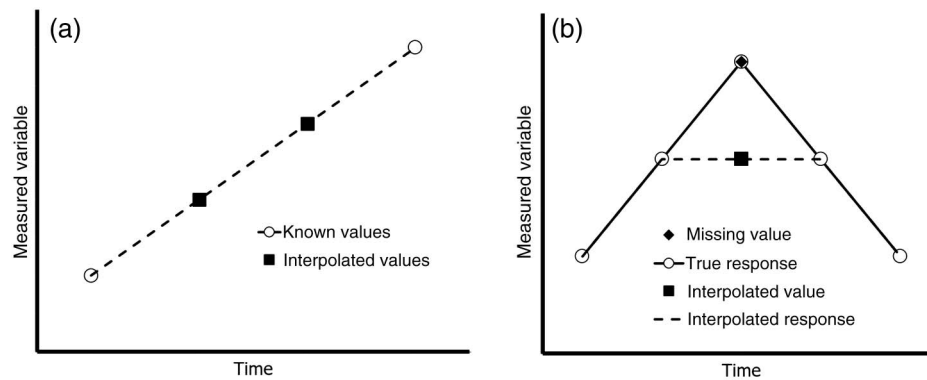


Figure 4 — (a) Simple representation of linear interpolation to impute missing data, and (b) a hypothetical time series response demonstrating a key limitation of linear interpolation.

utilized during time series data collection is confounded outside a certain range, especially high and/or low values are likely to be missing more frequently, eliciting an unrepresentatively skewed distribution (an example of MNAR). Indeed, Bell et al. (2014) demonstrated a greater level of bias in time series summary measures with data *missing at random* or MNAR, compared with *missing completely at random*, using a simulated randomized controlled trial. Researchers are therefore recommended to identify the cause of missing time series data and handle this issue accordingly.

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

Outliers

Another contentious topic in the initial screening of data is the identification and subsequent handling of outliers. Outlier identification typically uses statistical approaches, such as Tabachnick and Fidell (2007) defining values ≥ 3.29 SDs above or below the mean as outliers (the probability of obtaining a true sample this extreme is 0.1%). However, similar to missing values, continuous time series data are unique in that an outlier may be identifiable by its magnitude in relation to the rest of the response curve. This viewpoint may however lead to the exclusion of certain values simply because they

don't follow a relatively smooth pattern which, as measurement error is likely to exist in all samples, may be too subjective an approach. de Souza et al. (2016) advocate for data analyzes to be conducted with and without suspected outliers, to assess whether the main analysis is robust to these extreme cases. Comprehensive reporting of this sensitivity analysis may then be the most transparent approach to the handling of outliers.

Conclusion

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

Notes

1. Detailed user instructions can be found in the TSRA User Guide (Supplementary Material 2 [available online]), and details about the formulae used for the various area under the curve calculation options available in the TSRA can be found in Supplementary Material 3 (available online).

References

- Akintola, A.A., Noordam, R., Jansen, S.W., de Craen, A.J., Ballieux, B.E., Cobbaert, C.M., . . . van Heemst, D. (2015). Accuracy of continuous glucose monitoring measurements in normo-glycemic individuals. *PLoS One*, 10(10), e0139973. PubMed ID: 26445499 doi:10.1371/journal.pone.0139973
- Altman, D.G. (1985). Comparability of randomised groups. *The Statistician*, 34(1), 125–36. doi:10.2307/2987510
- Babraj, J.A., Vollaard, N.B.J., Keast, C., Guppy, F.M., Cottrell, G., & Timmons, J.A. (2009). Extremely short duration high intensity interval training substantially improves insulin action in young healthy males. *BMC Endocrine Disorders*, 9(1), 3. PubMed ID: 19175906 doi:10.1186/1472-6823-9-3
- Bell, M.L., King, M.T., & Fairclough, D.L. (2014). Bias in area under the curve for longitudinal clinical trials with missing patient reported

- outcome data: Summary measures versus summary statistics. *SAGE Open*, 4(2), 1–12. doi:10.1177/2158244014534858
- Berthiaume, N., & Zinker, B.A. (2002). Metabolic responses in a model of insulin resistance: Comparison between oral glucose and meal tolerance tests. *Metabolism: Clinical and Experimental*, 51(5), 595–598. doi:10.1053/meta.2002.31989
- Bickerton, A.S.T., Roberts, R., Fielding, B.A., Hodson, L., Black, E.E., Wagenmakers, A.J.M., ... Frayn, K.N. (2007). Preferential uptake of dietary fatty acids in adipose tissue and muscle in the postprandial period. *Diabetes*, 56(1), 168–176. PubMed ID: 17192479 doi:10.2337/db06-0822
- Ceriello, A., Esposito, K., Piconi, L., Ihnat, M.A., Thorpe, J.E., Testa, R., ... Giugliano, D. (2008). Oscillating glucose is more deleterious to endothelial function and oxidative stress than mean glucose in normal and type 2 diabetic patients. *Diabetes*, 57(5), 1349–1354. PubMed ID: 18299315 doi:10.2337/db08-0063
- Chung, S.T., Ha, J., Onuzuruike, A.U., Kasturi, K., Galvan-De La Cruz, M., Bingham, B.A., ... Sumner, A.E. (2017). Time to glucose peak during an oral glucose tolerance test identifies prediabetes risk. *Clinical Endocrinology*, 87(5), 484–491. PubMed ID: 28681942 doi:10.1111/cen.13416
- Cree-Green, M., Xie, D., Rahat, H., Garcia-Reyes, Y., Bergman, B.C., Scherzinger, A., ... Nadeau, K.J. (2018). Oral glucose tolerance test glucose peak time is most predictive of prediabetes and hepatic steatosis in obese girls. *Journal of the Endocrine Society*, 2(6), 547–562. PubMed ID: 29942919 doi:10.1210/js.2018-00041
- De Nicolao, G., Liberati, D., & Sartorio, A. (2000). Stimulated secretion of pituitary hormones in normal humans: A novel direct assessment from blood concentrations. *Annals of Biomedical Engineering*, 28(9), 1136–1145. PubMed ID: 11132197 doi:10.1114/1.1312185
- de Souza, R.J., Eisen, R.B., Perera, S., Bantoto, B., Bawor, M., Dennis, B.B., ... Thabane, L. (2016). Best (but oft-forgotten) practices: Sensitivity analyses in randomized controlled trials. *American Journal of Clinical Nutrition*, 103(1), 5–17. PubMed ID: 26675766 doi:10.3945/ajcn.115.121848
- Gannon, M.C., Nuttall, F.Q., Westphal, S.A., Neil, B.J., & Seaquist, E.R. (1989). Effects of dose of ingested glucose on plasma metabolite and hormone responses in type II diabetic subjects. *Diabetes Care*, 12(8), 544–552. PubMed ID: 2673694 doi:10.2337/diacare.12.8.544
- Garlick, P.J., Wernerman, J., McNurlan, M.A., Essen, P., Lobley, G.E., Milne, E., ... Vinnars, E. (1989). Measurement of the rate of protein synthesis in muscle of postabsorptive young men by injection of a 'flooding dose' of [1-13C] leucine. *Clinical Science*, 77(3), 329–336. PubMed ID: 2680231 doi:10.1042/cs0770329
- Gonzalez, J.T., Veasey, R.C., Rumbold, P.L.S., & Stevenson, E.J. (2013). Breakfast and exercise contingently affect postprandial metabolism and energy balance in physically active males. *British Journal of Nutrition*, 110(4), 721–732. PubMed ID: 23340006 doi:10.1017/S0007114512005582
- Gore, C.J., & Withers, R.T. (1990). The effect of exercise intensity and duration on the oxygen deficit and excess post-exercise oxygen consumption. *European Journal of Applied Physiology*, 60(3), 169–174. PubMed ID: 2347316 doi:10.1007/BF00839153
- Grinson, R.P., Ropelato, M.G., Gottlieb, S., Keselman, A., Martínez, A., Ballerini, M.G., ... Rey, R.A. (2010). Basal follicle-stimulating hormone and peak gonadotropin levels after gonadotropin-releasing hormone infusion show high diagnostic accuracy in boys with suspicion of hypogonadotropic hypogonadism. *The Journal of Clinical Endocrinology and Metabolism*, 95(6), 2811–2818. PubMed ID: 20371659 doi:10.1210/jc.2009-2732
- Ha, M.A., Mann, J.I., Melton, L.D., & Lewisbarned, N.J. (1992). Calculation of the glycemic index. *Diabetes, Nutrition & Metabolism*, 5(2), 137–139.
- Hofman, Z., van Drunen, J.D.E., de Later, C., & Kuipers, H. (2004). The effect of different nutritional feeds on the postprandial glucose response in healthy volunteers and patients with type II diabetes. *European Journal of Clinical Nutrition*, 58(11), 1553–1556. PubMed ID: 15173856 doi:10.1038/sj.ejcn.1602007
- Huck, S.W., & Cormier, W.H. (1995). *Reading statistics and research* (2nd ed.). New York, NY: HarperCollins.
- Koppeschaar, H.P.F., Popovic, V., Leal, A., Otero, X.L., Torres, E., Paramo, C., ... Casanueva, F.F. (2004). Growth hormone (GH) peaks versus areas under the curve in the diagnosis of adult GH deficiency: Analysis of the variables provided by the GHRH + GHRP-6 test. *Pituitary*, 7(1), 15–20. PubMed ID: 15638293 doi:10.1023/B:PITU.0000044629.10484.40
- Landler, L., Ruxton, G.D., & Malkemper, E.P. (2018). Circular data in biology: Advice for effectively implementing statistical procedures. *Behavioral Ecology and Sociobiology*, 72(8), 128. PubMed ID: 30100666 doi:10.1007/s00265-018-2538-y
- le Roux, C.W., Patterson, M., Vincent, R.P., Hunt, C., Ghatei, M.A., & Bloom, S.R. (2005). Postprandial plasma ghrelin is suppressed proportional to meal calorie content in normal-weight but not obese subjects. *The Journal of Clinical Endocrinology and Metabolism*, 90(2), 1068–1071. PubMed ID: 15522935 doi:10.1210/jc.2004-1216
- Li, P., & Stuart, E.A. (2019). Best (but oft-forgotten) practices: Missing data methods in randomized controlled nutrition trials. *American Journal of Clinical Nutrition*, 109(3), 504–508. PubMed ID: 30793174 doi:10.1093/ajcn/nqy271
- Little, R.J.A., & Rubin, D.B. (1987). *Statistical analysis with missing data*. New York, NY: John Wiley.
- Matthan, N.R., Ausman, L.M., Meng, H., Tighiouart, H., & Lichtenstein, A.H. (2016). Estimating the reliability of glycemic index values and potential sources of methodological and biological variability. *American Journal of Clinical Nutrition*, 104(4), 1004–1013. PubMed ID: 27604773 doi:10.3945/ajcn.116.137208
- Matthews, J.N., Altman, D.G., Campbell, M.J., & Royston, P. (1990). Analysis of serial measurements in medical research. *British Medical Journal*, 300(6719), 230–235. PubMed ID: 2106931 doi:10.1136/bmj.300.6719.230
- Monnier, L., Lapinski, H., & Colette, C. (2003). Contributions of fasting and postprandial plasma glucose increments to the overall diurnal hyperglycemia of type 2 diabetic patients. *Diabetes Care*, 26(3), 881–885. PubMed ID: 12610053 doi:10.2337/diacare.26.3.881
- Rodbard, D. (2009). New and improved methods to characterize glycemic variability using continuous glucose monitoring. *Diabetes Technology & Therapeutics*, 11(9), 551–565. PubMed ID: 19764834 doi:10.1089/dia.2009.0015
- Tabachnick, B.G., & Fidell, L.S. (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Allyn and Bacon.
- Tentolouris, N., Kanellos, P.T., Siami, E., Athanasopoulou, E., Chavias, N., Kolovou, G., ... Katsilambros, N. (2017). Assessment of the validity and reproducibility of a novel standardised test meal for the study of postprandial triacylglycerol concentrations. *Lipids*, 52(8), 675–686. PubMed ID: 28653085 doi:10.1007/s11745-017-4275-9
- Vinson, J., & Bose, P. (1988). Comparative bioavailability to humans of ascorbic acids alone of in citrus extract. *American Journal of Clinical Nutrition*, 48(3), 601–604. PubMed ID: 3414575 doi:10.1093/ajcn/48.3.601

- Vorster, H.H., Venter, C.S., & Silvis, N. (1990). The glycaemic index of foods: A critical evaluation. *South African Journal of Clinical Nutrition*, 2(1), 13–17.
- Weissgerber, T.L., Milic, N.M., Winham, S.J., & Garovic, V.D. (2015). Beyond bar and line graphs: Time for a new data presentation paradigm. *PLoS Biology*, 13(4), e1002128. PubMed ID: 25901488 doi:10.1371/journal.pbio.1002128
- Wijsman, C.A., van Heemst, D., Hoogeveen, E.S., Slagboom, P.E., Maier, A.B., de Craen, A.J.M., . . . Mooijaart, S.P. (2013). Ambulant 24-h glucose rhythms mark calendar and biological age in apparently healthy individuals. *Aging Cell*, 12(2), 207–213. PubMed ID: 23279694 doi:10.1111/ace1.12042
- Wolever, T.M. (2004). Effect of blood sampling schedule and method of calculating the area under the curve on validity and precision of glycaemic index values. *British Journal of Nutrition*, 91(2), 295–301. PubMed ID: 14756916 doi:10.1079/BJN20031054
- Wolever, T.M., & Jenkins, D.J.A. (1986). The use of the glycemic index in predicting the blood glucose response to mixed meals. *American Journal of Clinical Nutrition*, 43(1), 167–172. PubMed ID: 3942088 doi:10.1093/ajcn/43.1.167