suddengains: An R package to identify sudden gains in longitudinal data

Milan Wiedemann[1,2], Graham R Thew[1,2,3], Richard Stott[4], & Anke Ehlers[1,2,4]

[1] Department of Experimental Psychology, University of Oxford, Oxford, UK

[2] Oxford Health NHS Foundation Trust, Oxford, UK

[3] Oxford University Hospitals NHS Foundation Trust, Oxford, UK

[4] King's College London, London, UK

This manuscript has not yet been peer-reviewed.

Author Note

Earlier versions of this manuscript are available at https://psyarxiv.com/2wa84/.

Correspondence concerning this article should be addressed to Milan Wiedemann, Department of Experimental Psychology, University of Oxford. E-mail: milan.wiedemann@gmail.com

Abstract

Sudden gains are large and stable changes on an outcome variable between consecutive measurements, for example during a psychological intervention with multiple assessments. Researching these occurrences could help understand individual change processes in longitudinal data. Tang and DeRubeis (1999) suggested three criteria to define sudden gains in psychological interventions. However, identifying sudden gains based on these criteria can be time consuming and prone to errors if not fully automated. Further, methodological decisions such as how missing data, or multiple gains, are handled vary across studies and are reported with different levels of detail. These problems limit the comparability of individual studies and make it hard to understand or replicate the exact methods used. The R package *suddengains* provides a set of tools to facilitate sudden gains research. This article illustrates how to use the package to identify sudden gains or sudden losses within longitudinal data, while accounting for missing data, and how to extract descriptive statistics as well as exportable data files for further analysis. The *suddengains* package therefore offers significant scope to improve the efficiency, reporting, and reproducibility of sudden gains research.

Keywords: sudden gains, sudden losses, R, mechanisms of change, processes of change, psychological therapy, longitudinal analysis, single case designs

suddengains: An R package to identify sudden gains in longitudinal data

Introduction

A sudden gain is a large and stable change in an outcome variable experienced by an individual participant between two consecutive measurement points of a longitudinal data series. They were first defined and investigated by Tang and DeRubeis (1999), who examined session to session changes in depression symptoms among participants undertaking cognitive behavioural therapy. The majority of sudden gains studies to date have been in relation to psychological therapies (see Aderka, Nickerson, Boe, & Hofmann, 2012), but the analytic approach can be applied to any longitudinal dataset with regular repeated measurement, usually monitoring the impact of an intervention. In a meta-analysis of 16 studies of psychological therapies (total $N = 1104$), Aderka et al. (2012) found that experiencing a sudden gain was associated with better overall clinical outcomes at the end of treatment and at follow-up compared to those who did not experience gains. Given this potential significance of sudden gains, examining such events specifically may be informative in understanding when and why such large improvements occur, which could help to increase the efficacy and efficiency of the intervention.

Rates of sudden gains within published clinical studies vary quite considerably (e.g. 17.8% to 52.2% in Aderka et al. (2012)), which may partly be due to differences in the methods used to identify them. However, such differences are hard to examine given that sufficient methodological details to permit a comparison are not always reported. In addition, some studies have raised concerns about the validity of sudden gains identified through current methods, demonstrating that they can be found in placebo interventions and simulated datasets (Vittengl, Clark, & Jarrett, 2005; Vittengl, Clark, Thase, & Jarrett, 2015), and thus may be less meaningful than is often assumed. The *suddengains* R package is the first software program to offer explicit and reproducible methods to automatically identify sudden gains, which may be valuable in improving methodological reporting and

consistency across studies. It may also facilitate closer examination of the methods used to identify sudden gains, to help improve their validity and ensure that they more accurately reflect meaningful events. This article aims to provide an accessible overview of how sudden gains are calculated, describe the principal functions of the package, and give instructions on how to use these with longitudinal datasets. It is hoped that using this package will facilitate improvements in the efficiency, reporting, and reproducibility of sudden gains research.

Identification of sudden gains

Following Tang and DeRubeis (1999), the three criteria used to identify sudden gains are (see also Tang, DeRubeis, Beberman, & Pham, 2005):

1. The gain must be large in absolute terms. While this was originally operationalised as a decrease of at least 7 points on the Beck Depression Inventory (BDI; Beck & Steer, 1993), subsequent studies have generally used the reliable change index (RCI; Jacobson & Truax, 1991) to define an appropriate cutoff for other scales (see Stiles et al., 2003). Further details are discussed below.

2. The gain must be large in relative terms. This is defined as a drop of at least 25% of the previous score.

3. The gain must be large relative to symptom fluctuation. The difference between the mean scores of three measurements before the gain (M_{pre}), and the three measurements after the gain (M_{post}), must be greater than the pooled standard deviation of these two groups multiplied by 2.776 (i.e. the two-tailed critical value from the Student's t distribution for $\alpha = 0.05$ and $df = 4$). The formula for criterion 3 is therefore:

$$M_{pre} - M_{post} > \text{critical value} * \sqrt{\frac{(n_{pre} - 1) * SD_{pre}^2 + (n_{post} - 1) * SD_{post}^2}{n_{pre} + n_{post} - 2}}; \text{ where}$$

$$\text{critical value} = t_{\left(df,\ 1-\frac{\alpha}{2}\right)}$$

Defining a cutoff for the first criterion

Tang and DeRubeis (1999) originally defined a 7 point cutoff on the BDI for the first criterion based on frequency distribution plots of session to session change scores on the BDI in clinical trials. The authors reported that 7 BDI points approximately reflected one standard deviation in clinical samples (see Tang, 2015). Stiles et al. (2003) noted that 7 BDI points was close to the reliable change value reported in Barkham et al. (1996) and therefore used the RCI formula to define a cutoff for a new measure. Subsequent studies have generally adopted this approach. Jacobson and Truax (1991) proposed the following formula to test whether the observed change on a measure reflects more than just fluctuation due to measurement error:

$$\frac{\text{pre - post}}{S_{\text{diff}}} = \text{RCI};$$

Following Jacobson and Truax (1991) reliable change on a measure is present when:

$$\frac{\text{pre - post}}{S_{\text{diff}}} > 1.96; \text{ therefore}$$

$$\text{reliable change} > 1.96 \times S_{\text{diff}};$$

where S_{diff} is the standard error of the difference between pre and post scores. Using the standard error of measurement (S_E), S_{diff} can be expressed as:

$$S_{\text{diff}} = \sqrt{2 \times (S_E)^2};$$

where S_E is calculated using the standard deviation of the control group or normal population (s_1) and the test-retest reliability of the measure (r_{xx}):

$$S_E = s_1\sqrt{1 - r_{xx}};$$

Some studies have adapted this formula by replacing the test-retest reliability with the internal consistency (following suggestions from Martinovich, Saunders, & Howard, 1996) and replacing the standard deviation of the normal population (s_1) with the standard deviation of the clinical sample at baseline (SD_{pre}) so that all statistics can be extracted from the sample data (e.g. König, Karl, Rosner, & Butollo, 2014):

$$S_E = SD_{pre}\sqrt{1 - \alpha}$$

In the sudden gains literature different approaches have been used to define a cutoff for the first criterion using the RCI formula. Some studies have used the standard error of the difference (S_{diff}; e.g. Doane, Feeny, & Zoellner, 2010; Jun, Zoellner, & Feeny, 2013) while others have used the reliable change value ($1.96 \times S_{\text{diff}}$; e.g. Lutz et al., 2013; Zilcha-Mano, Eubanks, & Muran, 2019). When considering methods for defining a cutoff it is important to ensure that this value reflects a meaningful change (large in absolute terms) that is realistic in a session by session context for the intervention.

Missing data

Missing data around the period of a potential sudden gain need to be considered carefully, as depending on the number and distribution of missing datapoints for an individual, it may not be possible to calculate sudden gains. Specifically, in order to estimate the standard deviation values in criterion 3, at least two of the three measurements immediately prior to the gain must be present, as well as at least two of the three measurements immediately following the gain. The critical t values used in criterion 3 can be adjusted for missingness as follows (e.g., Lutz et al., 2013): Where no data are missing

$t_{(4;97.5\%)} > 2.776$; where one datapoint is missing either before or after the gain $t_{(3;97.5\%)} > 3.182$; and where one datapoint is missing both before and after the gain $t_{(2;97.5\%)} > 4.303$.

Terminology

The naming of specific sessions (or measurement points) around the gain follows the convention that the session immediately prior to the gain is session *N* (also known as the *pregain* session), and the session immediately after is session *N+1* (or *postgain* session). Other sessions are referred to in relation to session *N* (e.g. *N-2*, *N+3*).

Reversals

According to Tang and DeRubeis (1999) a sudden gain is counted as reversed if 50% of the improvement made during the gain was lost at any subsequent point. For example, where the sudden gain represents a drop in score from 40 to 30 points, the gain is classed as having reversed if a score of 35 or more is observed at any later session. As discussed by Wucherpfennig, Rubel, Hofmann, and Lutz (2017) a reversal might not necessarily be a stable phenomenon. These authors modified this criterion by suggesting that a stable reversal is present when a reversal is also classified as a sudden loss (see below).

Sudden losses

Although less frequently studied than sudden gains, sudden losses represent the inverse phenomenon, where a participant shows a large and stable increase of scores on the outcome variable. While some authors use the inverse criteria of the sudden gains (e.g., Krüger et al., 2014; Lutz et al., 2013), others adjust the percentage threshold of the second criterion (e.g. 33% in König et al., 2014).

Why is a package needed?

As indicated by the criteria above, identifying sudden gains requires the application of each of the three criteria to each session to session interval, and that this is performed for

each individual in a given dataset. A large number of calculations and extensive manipulation of data is therefore involved, particularly in larger datasets. Doing these data manipulations manually (e.g. in Excel) can be extremely time consuming and lead to errors. It also means that certain methodological decisions, such as the handling of missing data, or of participants with multiple gains, may not be addressed sufficiently or in a consistent way across studies. It is hoped that the use of the *suddengains* package will provide faster and more accurate calculations, as well as offering a transparent and consistent method to address these methodological considerations.

Functions of the *suddengains* package

The *suddengains* package provides a set of functions to calculate the presence of sudden gains (and sudden losses) within a longitudinal dataset, and to provide basic plots and descriptive statistics of the gains. It can also extract scores on secondary outcome or process measures around the period of each gain. Output files (in SPSS, Excel, or CSV formats) arranged by individual gain, or by person can be generated for further analyses in other programs. Table 1 lists and describes the main functions.

Worked example

This demonstration uses a dataset (`sgdata`) that was created to illustrate the functions of this package. The data show self-report weekly questionnaire scores for 43 participants who have received psychological therapy for depression. The intervention lasted for 12 sessions, and each participant completed a set of outcome measures at the beginning of each session, including the BDI (Beck & Steer, 1993) and a fictional secondary measure assessing rumination (RQ).

Preparation of data

The data to be analysed for sudden gains are arranged in wide format i.e. one row per participant, and one column for each questionnaire score at each measurement point. A

unique identifier variable also needs to be included. The optional **`select_cases()`** function can be used to determine cases that may need to be excluded, for example where outcome measures were completed infrequently, or where the pattern of missing data is such that it would not be possible to identify any sudden gains (see Table 2).

Identification of sudden gains

The **`identify_sg()`** function applies the sudden gains criteria as specified by the user to each session to session interval in the dataset. As shown below, the user specifies: `data`, the dataset to use in wide format; `sg_crit1_cutoff`, the cutoff value to use for criterion 1 (which can be entered manually or calculated using the **`define_crit1_cutoff()`** function); `sg_crit2_pct`, the percentage change value to use for criterion 2 (0.25 by default); `sg_crit3`, whether or not to apply the third criterion (`TRUE` by default); `sg_crit3_alpha`, the alpha value to use when calculating the criterion 3 critical value (0.05 by default); `id_var_name`, the name of the unique identifier variable within the dataset; and `sg_var_list`, a list of the variables representing the span of sessions to be analysed, which is sessions 1 to 12 in this example. By default all functions that identify sudden gains adjust for missingness as described earlier. To turn off this adjustment and instead use a critical value of 2.776 across all session to session intervals, the argument `sg_crit3_adjust = FALSE` can be included. Additional options to customise this analysis are discussed in the package documentation. An alternative function, **`identify_sl()`**, is identical to **`identify_sg()`** but applies the criteria in the inverse direction to calculate "sudden losses". The function **`check_interval()`** can be used to examine whether a specific session to session interval is a sudden gain/loss.

```
# First, load the suddengains R package
# See author notes for a link to package installation instructions
library(suddengains)
# Identify sudden gains in the dataset 'sgdata':
```

```
identify_sg(data = sgdata,
            sg_crit1_cutoff = 7,
            sg_crit2_pct = 0.25,
            sg_crit3 = TRUE,
            id_var_name = "id",
            sg_var_list = c("bdi_s1", "bdi_s2", "bdi_s3", "bdi_s4",
                            "bdi_s5", "bdi_s6", "bdi_s7", "bdi_s8",
                            "bdi_s9", "bdi_s10", "bdi_s11", "bdi_s12"))
```

The output shows each session to session interval, for example `sg_2to3`, with sudden gains indicated by a value of 1. To permit further analysis of our data, we wish to obtain an output dataset containing both the original data and the newly identified sudden gains. As participants may experience more than one gain, as in the present example, and to allow for different subsequent analyses, the package provides two options for output datasets: The **`create_bysg()`** function creates a dataset structured with one row per sudden gain, and the **`create_byperson()`** function creates a dataset structured with one row per person, indicating whether or not they experienced a sudden gain. The **`create_bysg()`** function is shown below. The `tx_start_var_name` and `tx_end_var_name` arguments are used to specify the start and end of treatment (tx) variables, and `sg_measure_name` specifies the name of the measure used to calculate sudden gains. To continue working in another program (e.g. SPSS, STATA, Excel) the functions **`write_bysg()`** and **`write_byperson()`** can be used to export the datasets created in R (R Core Team, 2018) as *.sav*, *.dta*, *.xlsx*, or *.csv* files.

```
# Create output dataset with one row per sudden gain
# and save as an object called "bysg" to use later
bysg <- create_bysg(data = sgdata,
                    sg_crit1_cutoff = 7,
                    sg_crit2_pct = 0.25,
```

```
                    sg_crit3 = TRUE,
                    id_var_name = "id",
                    tx_start_var_name = "bdi_s1",
                    tx_end_var_name = "bdi_s12",
                    sg_var_list = c("bdi_s1", "bdi_s2", "bdi_s3",
                                    "bdi_s4", "bdi_s5", "bdi_s6",
                                    "bdi_s7", "bdi_s8", "bdi_s9",
                                    "bdi_s10", "bdi_s11", "bdi_s12"),
                    sg_measure_name = "bdi")
```

Analysis of sudden gains

In this example, we have calculated sudden gains based on depression scores using the BDI. In analysing these gains, we are interested in how rumination scores on the fictional RQ measure change around the period of the sudden gains in depression. The **`extract_values()`** function extracts the RQ values from the three sessions before (*N-2*, *N-1*, *N*) and the three sessions after (*N+1*, *N+2*, *N+3*) each depression sudden gain. In the dataset that gets returned by this function we refer to these sessions as sg_bdi_**2n**, sg_bdi_**1n**, sg_bdi_**n**, sg_bdi_**n1**, sg_bdi_**n2**, and sg_bdi_**n3**, respectively. This function can be applied to either the `bysg` or `byperson` dataset. By default the extracted values will be added as new variables to the dataset used. Here we demonstrate applying this function to the `bysg` dataset, as shown in the code below. First, the RQ variables are added to the `bysg` dataset. Second, the **`extract_values()`** function is applied. Note that the list of RQ variables included in the `extract_var_list` argument must match those used for the `sg_var_list` argument used previously in the **`create_bysg()`** function. This means that the number of variables in these lists has to be identical and measured at the same timepoints. The output can be saved as a new object, or the existing `bysg` object can be overwritten, as in this example. The RQ scores now in the `bysg` dataset can be examined,

for example to look at the temporal relationship between changes in rumination and changes in depression symptoms.

```r
# 1. Select the ID and variables from a second measure
sgdata_rq <- dplyr::select(sgdata,
                           "id",
                           "rq_s1", "rq_s2", "rq_s3",
                           "rq_s4", "rq_s5", "rq_s6",
                           "rq_s7", "rq_s8", "rq_s9",
                           "rq_s10", "rq_s11", "rq_s12")

# 2. Add the variables in 'sgdata_rq' to the 'bysg' dataset created earlier
bysg <- dplyr::left_join(bysg, sgdata_rq, by = "id")

# 3. Extract values on the second measure around the sudden gain
bysg <- extract_values(data = bysg,
                       id_var_name = "id_sg",
                       extract_var_list = c("rq_s1", "rq_s2", "rq_s3",
                                            "rq_s4", "rq_s5", "rq_s6",
                                            "rq_s7", "rq_s8", "rq_s9",
                                            "rq_s10", "rq_s11", "rq_s12"),
                       extract_measure_name = "rq",
                       add_to_data = TRUE)
```

The **`describe_sg()`** function provides descriptive statistics about the sudden gains based on the variables from the `bysg` and `byperson` datasets. For the present example, this function indicates that 16 of the 43 participants experienced a sudden gain, and 9 experienced more than one gain, leading to a total of 26 sudden gains within the data.

Information on the mean gain magnitude and reversals is also provided.

The **`plot_sg()`** function plots the "average" sudden gain, and can be used to show the primary or secondary outcome measure data (Figures 1A and 1B). The `sg_pre_post_var_list` argument specifies the pregain and postgain variables to be plotted, namely sessions *N-2* to *N+3*. This function is built using the R-package *ggplot2* (Version 3.2.1; Wickham, 2016) and additional *ggplot2* functions can be added to the plot. It is also possible to plot the average gain magnitude of different groups (e.g. two treatment arms in a trial) in one figure by using the optional `group` argument (see Figure 1C). An additional function, **`plot_sg_trajectories()`**, is available to plot the trajectories of a selection of individual cases within the dataset (see Figure 2A). This function can be paired with a filter command (e.g. **`filter()`** from R-package *dplyr* (Version 0.8.3; Wickham, François, Henry, & Müller, 2018)) to visualise trajectories of specific groups of participants. For example, all participants with more than one sudden gain, or all participants with a sudden gain between sessions 3 and 4 (see Figure 2B).

```
# Create average sudden gain plot for BDI data (see Figure 1A):
plot_sg(data = bysg,
        id_var_name = "id",
        tx_start_var_name = "bdi_s1",
        tx_end_var_name = "bdi_s12",
        sg_pre_post_var_list = c("sg_bdi_2n", "sg_bdi_1n", "sg_bdi_n",
                                 "sg_bdi_n1", "sg_bdi_n2", "sg_bdi_n3"),
        ylab = "BDI")
```

Discussion

The analysis of sudden gains provides a detailed examination of individuals' trajectories of change during the course of an intervention, and may help to understand

processes of change. The *suddengains* package aims to facilitate the computation of gains, which can be laborious and error-prone. It also aims to address common methodological issues, for example by making adjustments to the critical value for the third criterion in the presence of missing data, and highlighting participants with multiple gains. It is hoped that this package will permit faster and more transparent examination of sudden gains within a range of longitudinal datasets, and that it could provide a valuable tool to explore how the criteria might be refined or adapted to better identify gains that reflect meaningful change processes. The present examples refer to package version 0.4.0. Instructions for installing the package along with all code, materials, data, and further technical details can be found at https://github.com/milanwiedemann/suddengains and https://CRAN.R-project.org/package=suddengains. Readers are also welcome to contact the authors via email or GitHub at https://github.com/milanwiedemann/suddengains/issues.

Author contributions

RS wrote an initial script to identify sudden gains in SPSS syntax. MW planned the R package as part of his DPhil and together with GRT conceptualised the functionality. MW wrote the R code for the functions. GRT and MW wrote the function documentation. GRT and RS both reviewed and tested the functions. GRT and MW wrote the first draft of the manuscript. AE supervised MW and provided consultation throughout. All authors reviewed and approved the final manuscript.

Acknowledgements

This paper was written using the R package *papaja* (Aust & Barth, 2018).

Funding

This project was supported by a Mental Health Research UK studentship (MW), the Wellcome Trust [102176 (GRT); 069777 and 200976 (AE, RS)], the Oxford Health NIHR Biomedical Research Centre (MW, GRT, AE), and the NIHR Oxford Biomedical Research Centre (GRT). The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

Aderka, I. M., Nickerson, A., Boe, H. J., & Hofmann, S. G. (2012). Sudden gains during psychological treatments of anxiety and depression: A meta-analysis. *Journal of Consulting and Clinical Psychology*, *80*(1), 93–101. doi:10.1037/a0026455

Aust, F., & Barth, M. (2018). *papaja: Create APA manuscripts with R Markdown*. Retrieved from https://github.com/crsh/papaja

Barkham, M., Rees, A., Stiles, W. B., Shapiro, D. A., Hardy, G. E., & Reynolds, S. (1996). DoseEffect relations in time-limited psychotherapy for depression. *Journal of Consulting and Clinical Psychology*, *64*(5), 927–935. doi:10.1037/0022-006X.64.5.927

Beck, A. T., & Steer, R. A. (1993). *Beck Depression Inventory Manual*. San Antonio, TX: The Psychological Corporation.

Doane, L. S., Feeny, N. C., & Zoellner, L. A. (2010). A preliminary investigation of sudden gains in exposure therapy for PTSD. *Behaviour Research and Therapy*, *48*(6), 555–560. doi:10.1016/j.brat.2010.02.002

Jacobson, N. S., & Truax, P. A. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, *59*(1), 12–19. doi:10.1037/0022-006X.59.1.12

Jun, J. J., Zoellner, L. A., & Feeny, N. C. (2013). Sudden gains in prolonged exposure and sertraline for chronic PTSD. *Depression and Anxiety*, *30*(7), 607–613. doi:10.1002/da.22119

König, J., Karl, R., Rosner, R., & Butollo, W. (2014). Sudden gains in two psychotherapies for posttraumatic stress disorder. *Behaviour Research and Therapy*, *60*, 15–22. doi:10.1016/j.brat.2014.06.005

Krüger, A., Ehring, T., Priebe, K., Dyer, A. S., Steil, R., & Bohus, M. (2014). Sudden losses and sudden gains during a DBT-PTSD treatment for posttraumatic stress disorder following childhood sexual abuse. *European Journal of Psychotraumatology*, *5*, 24470. doi:10.3402/ejpt.v5.24470

Lutz, W., Ehrlich, T., Rubel, J., Hallwachs, N., Rottger, M.-A., Jorasz, C., ... Tschitsaz-Stucki, A. (2013). The ups and downs of psychotherapy: Sudden gains and sudden losses identified with session reports. *Psychotherapy Research*, *23*(1), 14–24. doi:10.1080/10503307.2012.693837

Martinovich, Z., Saunders, S., & Howard, K. (1996). Some Comments on "Assessing Clinical Significance". *Psychotherapy Research*, *6*(2), 124–132. doi:10.1080/10503309612331331648

R Core Team. (2018). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Stiles, W. B., Leach, C., Barkham, M., Lucock, M., Iveson, S., Shapiro, D. A., ... Hardy, G. E. (2003). Early sudden gains in psychotherapy under routine clinic conditions: Practice-based evidence. *Journal of Consulting and Clinical Psychology*, *71*(1), 14–21. doi:10.1037/0022-006X.71.1.14

Tang, T. Z. (2015). Sudden gains. In R. L. Cautin & S. O. Lilienfeld (Eds.), *The Encyclopedia of Clinical Psychology* (pp. 2745–2751). Hoboken, New Jersey: Wiley-Blackwell. doi:10.1002/9781118625392.wbecp378

Tang, T. Z., & DeRubeis, R. J. (1999). Sudden gains and critical sessions in cognitive-behavioral therapy for depression. *Journal of Consulting and Clinical Psychology*, *67*(6), 894–904. doi:10.1037/0022-006X.67.6.894

Tang, T. Z., DeRubeis, R. J., Beberman, R., & Pham, T. (2005). Cognitive changes, critical sessions, and sudden gains in cognitive-behavioral therapy for depression. *Journal of Consulting and Clinical Psychology*, *73*(1), 168–172. doi:10.1037/0022-006X.73.1.168

Vittengl, J. R., Clark, L. A., & Jarrett, R. B. (2005). Validity of sudden gains in acute phase treatment of depression. *Journal of Consulting and Clinical Psychology*, *73*(1), 173–182. doi:10.1037/0022-006X.73.1.173

Vittengl, J. R., Clark, L. A., Thase, M. E., & Jarrett, R. B. (2015). Detecting sudden gains during treatment of major depressive disorder: Cautions from a monte carlo analysis. *Current Psychiatry Reviews*, *11*(1), 19–31. doi:10.2174/1573400510666140929195441

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York. Retrieved from http://ggplot2.org

Wickham, H., François, R., Henry, L., & Müller, K. (2018). *dplyr: A grammar of data manipulation*. Retrieved from https://CRAN.R-project.org/package=dplyr

Wucherpfennig, F., Rubel, J. A., Hofmann, S. G., & Lutz, W. (2017). Processes of change after a sudden gain and relation to treatment outcome - Evidence for an upward spiral. *Journal of Consulting and Clinical Psychology*, *85*(12), 1199–1210. doi:10.1037/ccp0000263

Zilcha-Mano, S., Eubanks, C. F., & Muran, J. C. (2019). Sudden gains in the alliance in cognitive behavioral therapy versus brief relational therapy. *Journal of Consulting and Clinical Psychology*, *87*(6), 501–509. doi:10.1037/ccp0000397

Table 1

Main functions of the suddengains R package.

Function	Description
Identify sudden gains	
define_crit1_cutoff()	Uses modified RCI formula to determine a cutoff value for criterion 1
check_interval()	Checks if a given interval is a sudden gain/loss
identify_sg(), identify_sl()	Identifies sudden gains/losses
Create datasets	
create_bysg(), create_byperson()	Creates a dataset with one row for each sudden gain/loss or one row for each person
extract_values()	Extracts values on a secondary measure around the sudden gain/loss
Describe sudden gains	
describe_sg()	Generates summary descriptive statistics
plot_sg(), plot_sg_trajectories()	Creates plots of the average sudden gain, or plots of individual case trajectories
Additional functions	
select_cases()	Selects cases to be included in the sudden gains analysis based on different criteria
write_bysg(), write_byperson()	Exports CSV, SPSS, Excel, or STATA files of the sudden gains datasets

Note. More details of each function can be found in the package documentation or using the help() function in R.

Table 2

Minimum available data patterns required to identify sudden gains.

	x_i	x_{i+1}	x_{i+2}	x_{i+3}	x_{i+4}	x_{i+5}
Pattern 1	●	$●_n$	●	●	○	○
Pattern 2	●	$●_n$	●	○	●	○
Pattern 3	●	○	$●_n$	●	●	○
Pattern 4	●	○	$●_n$	●	○	●

Note. x_i to x_{i+5} represent the scores on the measure used to identify sudden gains at time point i and the subsequent five measurement points. ● represents available data; $●_n$ represents available data to be considered as the possible pregain session; ○ represents missing data.

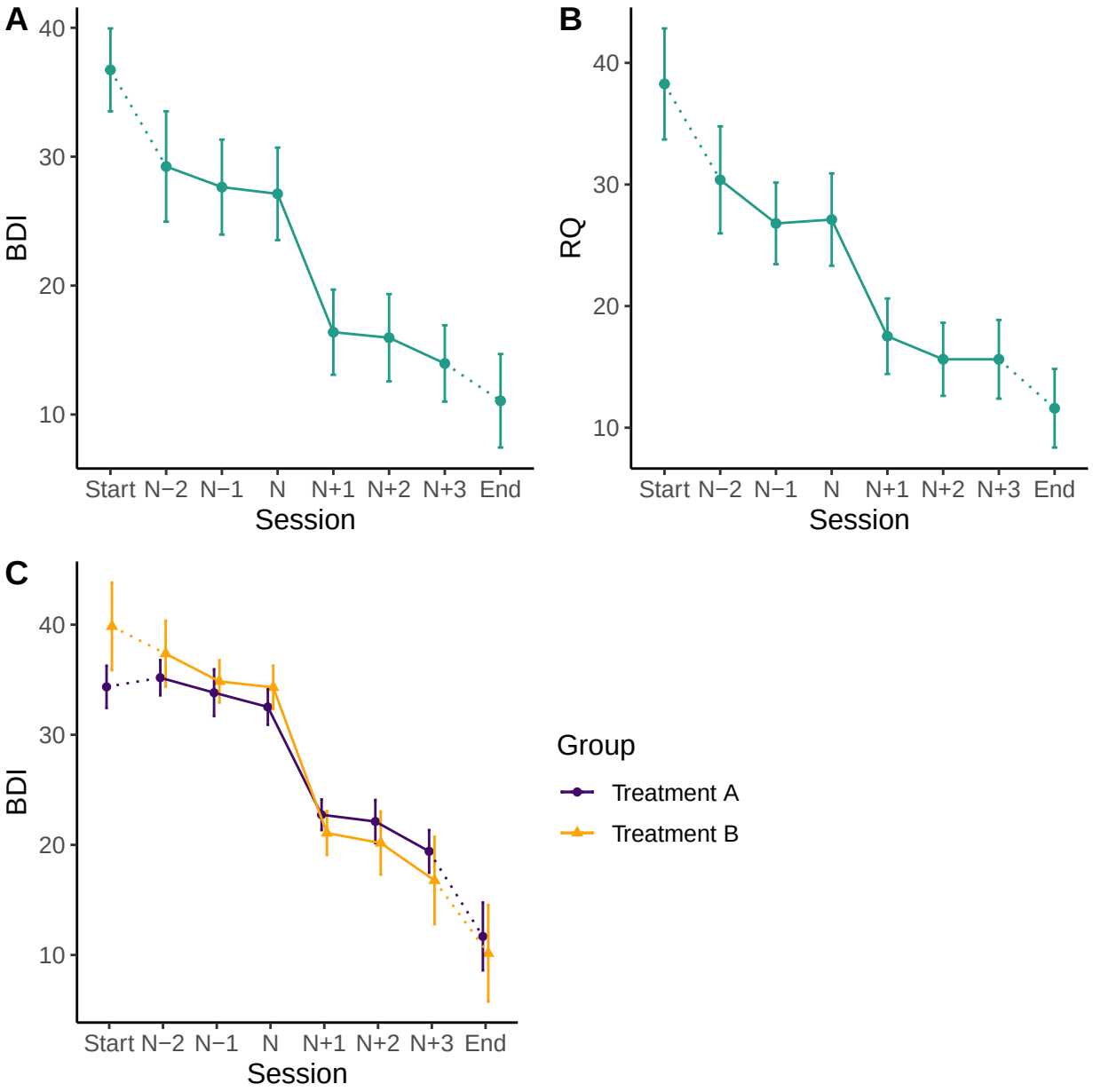

Figure 1. (A) Average gain magnitute on the BDI and (B) average change in rumination (RQ) around the time of the sudden gain on BDI for all sudden gains. (C) Average gain magnitude on the BDI for two different treatments.

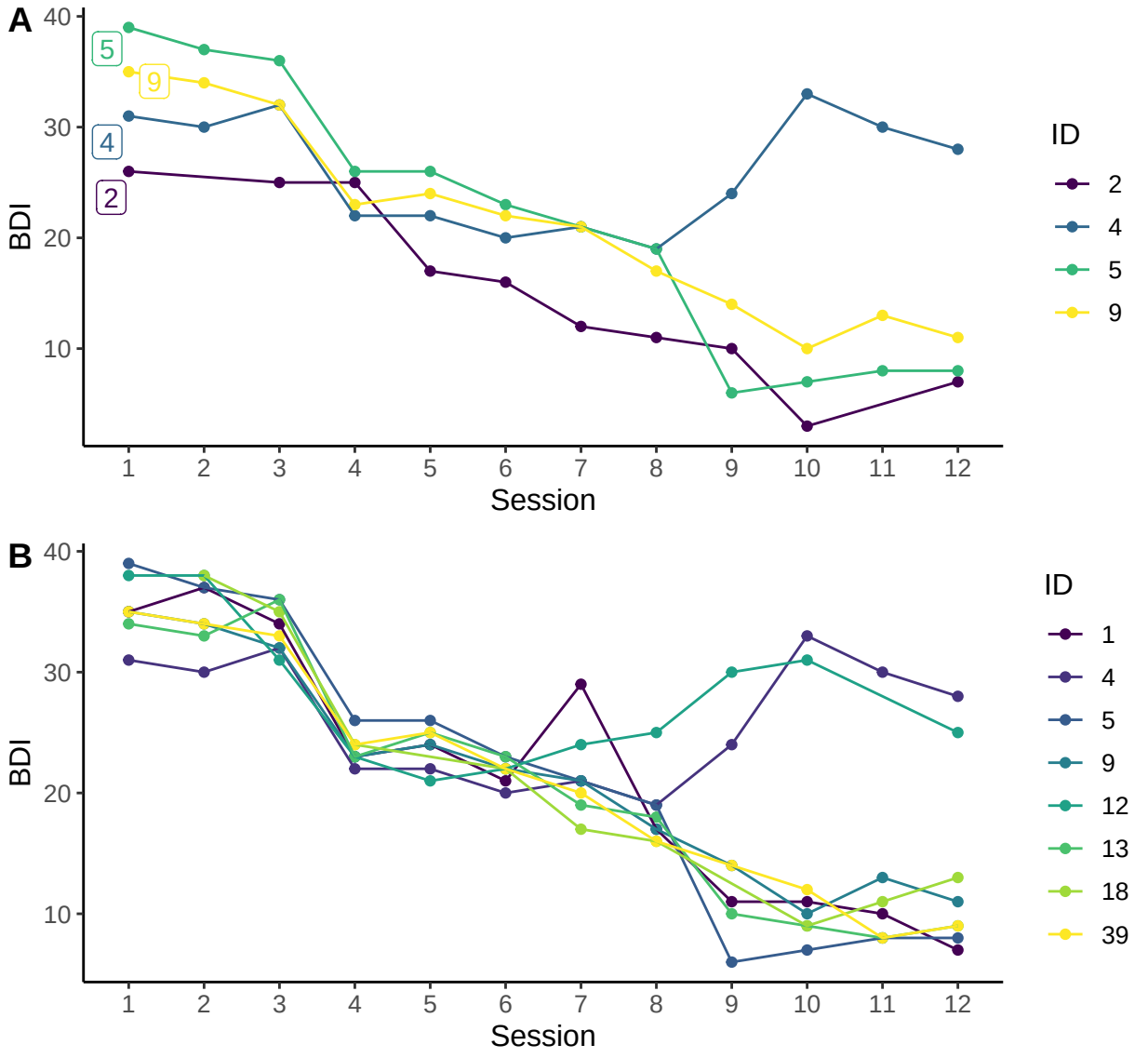

Figure 2. (A) Trajectories for a selection of individual cases. (B) Trajectories of BDI scores for all participants with a sudden gain between sessions 3 and 4.