



Appendix

**This appendix was part of the submitted manuscript and has been peer reviewed.
It is posted as supplied by the authors.**

Appendix to: Beath A, Jones MP. Guided by the research design: choosing the right statistical test.
Med J Aust 2018; 208: 163-165. doi: 10.5694/mja17.00422.

The following example is used to demonstrate specific research questions that each statistical test would be used to address.

We have conducted a large-scale, representative study into the physical activity, lifestyle and health of Australian adolescents. This study has two parts:

- (1) A cross-sectional element, where the researchers collected data from adolescents to assess the state of health among Australian adolescents; and
- (2) A randomized trial of an exercise intervention aimed at improving the health of adolescents, particularly for those who were currently overweight.

Different kinds of research questions are outlined below, in order to demonstrate different kinds of statistical analyses.

Test	Design	Example research context	Example research question	Variables
One-sample t-test	Non-randomized, cross-sectional study	We have collected data on Australian adolescents' exercise habits: specifically, how much they exercise on a weekly basis. We wish to compare exercise habits of Australian adolescents to adolescents from the USA. There is existing information on how much USA adolescents exercise each week on average, taken from previously published research.	Is the average amount of exercise that Australian adolescents perform different to USA adolescents?	A single numeric variable
Pearson's chi-square goodness of fit test	Non-randomized, cross-sectional study	We have collected data on Australian adolescents' weight and height to be able to measure BMI. We wish to compare their BMI to what we know of BMI distribution of adolescents from the USA, taken from previously published research. As in the one-sample t-test above, the proportion of USA adolescents who are obese is already known from previously published research.	Is the proportion of Australian adolescents that are classified (according to their BMI) as obese different to the (known) proportion of USA adolescents who are obese?	A single categorical variable
Pearson's correlation	Non-randomized, cross-sectional study	We have collected information on adolescents' diet and exercise habits.	Is there a negative (that is, inverse) association between adolescents' exercise and amount of sugary drinks consumed per week?	Two numeric variables.
Independent samples t-test	Non-randomized, case-control study	We wish to compare the exercise habits of adolescents with Type-II diabetes (T2D, the clinical case group) and those without (the control group). The two groups have been matched on BMI.	Do adolescents with T2D exercise less than their (matched) controls?	One categorical grouping variable (predictor) with 2 or more levels. One numeric outcome.
	Randomized,	We have developed an exercise	Is the body fat percentage of	

Test	Design	Example research context	Example research question	Variables
	RCT with single time point outcome	intervention for overweight adolescents, who have been randomly allocated to either an exercise intervention or a education-only control condition. The primary outcome is body fat percentage, measured at the end of the 4-week treatment period.	those randomised to an exercise program lower than the control group?	
Pearson's chi-square test of independence / logistic regression	Non-randomized, case-control study	We wish to compare the incidence of problems with fatigue among adolescents with Type-II diabetes (T2D, the clinical case group) compared to adolescents without (the control group). The two groups have been matched on BMI.	Is the proportion of T2D adolescents who experience fatigue issues different to the proportion of control adolescents who experience fatigue issues?	One categorical grouping variable (with 2 or more levels). One categorical outcome (with 2 or more levels).
	Randomized, RCT with single time point outcome	We wish to evaluate the efficacy of a lifestyle intervention on BMI (specifically, normal vs overweight BMI). Adolescents are randomly allocated to one of three conditions: (a) combination exercise and diet program; (b) exercise-only; and (c) education-only control.	Are adolescents who complete the combination program more likely to have a normal BMI (compared to overweight), compared to (1) the control group and (2) the exercise-only group?	
Paired t-test / repeated measures ANOVA	Non-randomized, cohort study	In the context of an exercise program for overweight adolescents, their weight is taken at the start (baseline) and again at the end of the program (end-of-program). (There is no control or comparison group.)	Does weight decrease from baseline to end-of-program?	One numeric outcome, measured repeatedly. One categorical predictor, 'time', with 2 or more levels
Generalized estimating equations / generalized linear mixed model	Non-randomized, cohort study	Does the proportion of overweight (vs normal weight, as measured by BMI) adolescents decrease after an exercise intervention (from T1 to T2)?		One categorical outcome variable, measured repeatedly. One categorical predictor, 'time', with 2 or more levels

Test	Design	Example research context	Example research question	Variables
	Randomized, RCT with outcome over time	In the context of testing the efficacy of an exercise program on BMI (normal BMI vs overweight), BMI is measured immediately at the end of the program, and also 6 months later. Adolescents are randomly allocated to one of three conditions: (a) combination exercise and diet program; (b) exercise-only; and (c) education-only control.	Does the likelihood of being overweight decrease over time for the combination and exercise-only groups (compared to control), and is this decrease greater in the combination program compared to the exercise-only program?	One categorical outcome variable (with 2 or more levels), measured repeatedly. One categorical predictor, 'time' (with 2 or more levels). One categorical grouping variable (with 2 or more levels).
General linear mixed model	Randomized, RCT with outcome over time	In the context of testing the efficacy of an exercise program on body fat percentage, body fat is measured immediately at the end of the program, and also 6 months later. Adolescents are randomly allocated to one of three conditions: (a) combination exercise and diet program; (b) exercise-only; and (c) education-only control.	Does body fat decrease over time for the combination and exercise-only groups (compared to control), and is this decrease greater in the combination program compared to the exercise-only program?	One numeric outcome, measured repeatedly. One categorical predictor, 'time', with 2 or more levels. One categorical grouping variable (with two or more levels).