BMJ 2014;349:g5455 doi: 10.1136/bmj.g5455 (Published 5 September 2014)

ENDGAMES

STATISTICAL QUESTION

What is a factorial study design?

Philip Sedgwick *reader in medical statistics and medical education*

Institute for Medical and Biomedical Education, St George's, University of London, London, UK

Researchers investigated whether inclusion of glutamine or selenium in a standard isonitrogenous, isocaloric preparation of parenteral nutrition affected the occurrence of new infections in critically ill patients. A randomised double blind placebo controlled trial was conducted, using a full factorial study design. There were two factors—treatment with glutamine (20.2 g/day) or glutamine placebo and treatment with selenium (500 µg/day) or selenium placebo, each for up to seven days. Participants were 502 adults in intensive care units or high dependency units who had gastrointestinal failure and required parenteral nutrition.[1]

The outcome measure was a new clinically suspected infection within the first 14 days of treatment. The trial found no evidence of an interaction between glutamine and selenium in treatment effect (P=0.96). Selenium supplementation had no significant effect on the development of new infections (odds ratio 0.81, 95% confidence interval 0.57 to 1.15). Glutamine also had no significant effect on the development of new infections (odds ratio 1.07, 0.75 to 1.53). It was concluded that the supplementation of parenteral nutrition with glutamine or selenium had no effect on the development of new infections in critically ill patients.

Which of the following statements, if any, are true?

a) The trial used the double dummy method

b) There were four treatment groups

c) If an interaction existed between glutamine and selenium, the effects of glutamine would have differed between the selenium and selenium placebo treatment groups

d) The factorial design of the trial required a smaller sample size than investigating the effectiveness of glutamine and selenium in separate trials

Answers

Statements *a*, *b*, *c,* and *d* are all true.

The aim of the trial was to investigate whether inclusion of glutamine or selenium in parenteral nutrition had an effect on the development of new infections in critically ill patients. A randomised double blind placebo controlled trial that incorporated a full factorial study design was used. The factorial design enables a clinical trial to evaluate two or more interventions simultaneously. The use of a full factorial design in the above trial meant that the effects of glutamine and selenium could be evaluated both separately and combined.

In the above trial there were two separate interventions—treatment with glutamine (20.2 g/day) or glutamine placebo, and treatment with selenium (500 µg/day) or selenium placebo. In a factorial trial, the intervention is referred to as a factor. Each factor may have two or more so called levels. There were two levels for both the glutamine intervention (glutamine (20.2 g/day) or glutamine placebo) and the selenium intervention (selenium (500 µg/day) or selenium placebo). Two placebos (placebo glutamine and placebo selenium) were therefore needed to ensure that the trial was double blind, thereby minimising the potential of assessor and reporting biases.[2] The inclusion of two placebos is known as the double dummy method (*a* is true), described in a previous question.[3] This method was used because there were two active treatments but it was not possible to manufacture them so that they looked identical, although it was possible to manufacture an identical looking placebo for each active treatment.

In a factorial design it is useful to consider the allocation to treatment groups as resulting from more than one randomisation of the participants. In the above trial, participants would have been randomised, for example, first to glutamine or glutamine placebo and then to selenium or selenium placebo. There were 502 trial participants, of whom 250 were randomised to glutamine and 252 to glutamine placebo. Of the 250 participants randomised to glutamine, 124 were subsequently randomised to selenium (in addition to glutamine) and 126 to selenium placebo (in addition to glutamine). Of the 252 participants randomised to glutamine placebo, 127 were subsequently randomised to selenium (in addition to glutamine placebo), and 125 to selenium placebo (in addition to glutamine placebo). Therefore, about half of the participants were randomised to glutamine and half to glutamine placebo, whereas about half were randomised to selenium and half to selenium placebo (table⇓). There were four treatment groups (*b* is true)—glutamine and selenium placebo (described as glutamine alone), selenium and glutamine placebo (described as selenium

p.sedgwick@sgul.ac.uk

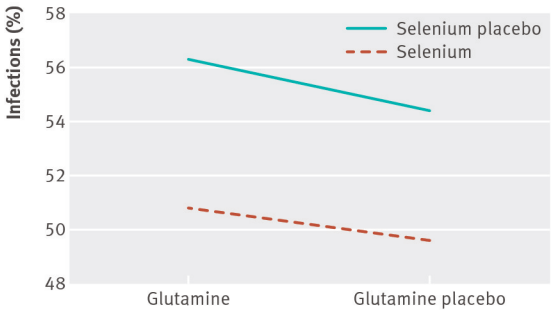
alone), glutamine and selenium combined, plus glutamine placebo and selenium placebo.

Sometimes a numbering notation is used to describe a factorial design. The above trial is described as a two by two (written as 2×2, or 2^2) design. Using this notation, the number of numbers indicates the number of factors, and the value of the number indicates the number of levels for the factor. The above trial was the simplest factorial design—it had the smallest number of factors and each factor had the smallest number of levels. A more complex design might, for example, include three factors with one factor having two levels and the other two factors having three levels. Such a design would be notated by 2×3×3, or 2×3^2. The order of the numbers is irrelevant, and this example could also be written using the notation of $3^2\times2$.

The number of potential treatment groups in a factorial design can be determined by multiplying through the number notation described. The above trial is described as a 2×2 factorial design, which gives four possible groups. It had a full factorial design and therefore all possible combinations of treatment were included. Factorial designs can sometimes include a potentially large number of treatment groups. For example a $3^2\times2$ full factorial design would involve 18 treatment groups. If there are limited resources or it is not necessary to include all treatment groups to answer the research question, then a subset or fraction of the treatment groups needed for a full factorial design may be carefully selected. Such a design would be described as a partial or fractional factorial design.

The effects of glutamine and selenium were assessed by undertaking a marginal analysis. The effects of treatment with glutamine were assessed by comparing the marginal totals—that is, the proportion of participants allocated to glutamine that developed an infection was compared with that for those allocated to glutamine placebo, regardless of whether participants were also allocated to selenium or selenium placebo. The marginal totals for glutamine and glutamine placebo are shown (table). The proportion of the glutamine group that developed an infection was higher than for the glutamine placebo group, although this difference was not significant (53.6% *v* 52.0%; odds ratio 1.07, 95% confidence interval 0.75 to 1.53). Similarly, the effects of selenium were investigated by comparing the marginal total for all participants allocated to selenium with that for all those allocated to selenium placebo, regardless of whether participants were also allocated to glutamine or glutamine placebo. The proportion of the selenium group that developed an infection was lower than for the selenium placebo group, although the difference was not significant (50.2% *v* 55.4%; 0.81, 0.57 to 1.15).

It was essential that the analysis of the factorial trial also looked for an interaction between glutamine and selenium in treatment effects. An interaction would have existed if the effects of glutamine were modified by whether participants received selenium or selenium placebo (*c* is true). Equally, an interaction would have existed if the effects of selenium were modified by whether participants received glutamine or glutamine placebo—the statements are analogous. The existence of an interaction can be investigated graphically using a profile plot (figure). The figure shows the proportion of participants allocated to the glutamine and glutamine placebo groups that developed an infection, plotted separately for selenium and selenium placebo. Alternatively, the profile plot could have shown the proportion of participants allocated to selenium and selenium placebo that developed an infection, plotted separately for glutamine and glutamine placebo.



Profile plot displaying the proportions of participants allocated to glutamine and glutamine placebo with an infection, plotted for the selenium intervention

The profile plot indicates that although the proportion of participants that developed an infection was lower with selenium supplements than with selenium placebo, the effects of glutamine (the difference between glutamine and glutamine placebo in the proportion of participants that developed an infection) did not differ much, if at all, between those participants allocated to selenium and selenium placebo. This suggests that there was no interaction between glutamine and selenium in treatment effect. The presence of an interaction was investigated more formally using traditional statistical hypothesis testing, with a null and alternative hypothesis as described in a previous question.[4] The null hypothesis stated that in the population from which the participants were selected, no interaction existed between glutamine and selenium in the treatment effect of proportion of infections in the first 14 days. The alternative hypothesis indicates that an interaction existed. The resulting P value was 0.96, and because P was greater than 0.05 the null hypothesis was not rejected in favour of the alternative. The conclusion was that there was no evidence of an interaction between glutamine and selenium, confirming the observation from the profile plot. If an interaction had existed, it would have suggested that the combined effects of glutamine and selenium were not simply the sum of the effects of each treatment but were greater or smaller than expected. If a significant interaction had existed, then the marginal estimates of the effects of glutamine and selenium would have been biased.

The advantage of the factorial design is that the treatment effects of glutamine and selenium could be evaluated simultaneously. The design was a more efficient use of resources because a smaller sample size was needed than for two separate placebo controlled trials (*d* is true) or one trial with glutamine, selenium, and placebo treatment arms. The design also provided information about the potential interaction between glutamine and selenium. In the absence of an interaction between glutamine and selenium, it was possible to investigate the treatment effects of glutamine and selenium using marginal analysis.

Competing interests: None declared.

1 Andrews PJD, Avenell A, Noble DW, Campbell MK, Croal BL, Simpson WG, et al, and the SIGNET (Scottish Intensive care Glutamine or seleNium Evaluative Trial) Trials Group. Randomised trial of glutamine, selenium, or both, to supplement parenteral nutrition for critically ill patients. *BMJ* 2011;342:d1542.
2 Sedgwick P. Bias in clinical trials. *BMJ* 2011;342:d4176.
3 Sedgwick P. Double dummy trials. *BMJ* 2011;343:d7294.
4 Sedgwick P. Understanding statistical hypothesis testing. *BMJ* 2014;348:g3557.

Cite this as: *BMJ* 2014;349:g5455

Table

Table 1| Infection outcomes in 502 intensive care patients randomised to combinations of the trial parenteral nutrition formulations. Values are numbers (percentages) of infections*

	Selenium	Selenium placebo	Marginal total
Glutamine	63 (50.8) (n=124)	71 (56.3) (n=126)	134 (53.6) (n=250)
Glutamine placebo	63 (49.6) (n=127)	68 (54.4) (n=125)	131 (52.0) (n=252)
Marginal total	126 (50.2) (n=251)	139 (55.4) (n=251)	

*n=total number treated.