Careless Statistics
Costs Lives

A tragi-comedy of errors
of the first and second kind,
both one- and two-sided

Richard Gill, Leiden University
8 May, 2008
Careless statistics does cost lives
Altman (1982): 50% of published medical statistics is wrong. Today: about 15%

This talk:

- Theory: Snapinn (1992; *Statistics in Medicine*) early stopping rule for randomized clinical trials

- Practice: the probiotica trial (infectious complication in acute pancreatitis)
• use of probiotics during admission

**Randomisation criteria**
After inclusion in the study, patients with predicted severe acute pancreatitis, represented by at least one of the following scores: 3 Imrie criteria, CRP 150 mg/L, APACHE II score 8, are randomised within the first 72 hours after the onset of abdominal pain. Patients with a predicted mild attack of acute pancreatitis do not receive the study product. They do give informed consent and are monitored.

**Ethics, informed consent**
This study is conducted in accordance with the principles of the Declaration of Helsinki and 'good clinical practice' guidelines. The independent ethics committee of all 15 participating hospitals approved the final protocol. Oral and written informed consent in form is obtained from the patient before inclusion in the trial.

**Safety**
All the probiotics used in this study have a long history of use in the food industry. Probiotics have been studied in many critical ill and immunocompromised patients without any serious adverse events being noted. There is one trial that studied probiotics in acute pancreatitis patients and no serious advents were noted. If an infection with one of the administered probiotics might occur, this could be treated with antibiotics. During administration of the

**Interim-analysis**
For ethical reasons it is desirable to end a therapeutic experiment once a statistical significant difference in treatment results has been reached. This study uses the stopping-rules according to Snapinn [32]. An interim-analysis will be performed after the data of the first 100 patients (50% fraction) is obtained. According to Snappin, the trial will be ended at this interim-analysis at p < 0,0081. The study will also be ended in case of adverse events without possibility of positive outcome, p > 0,382. The monitoring committee will discuss the results of the interim-analysis and advice the steering committee. The steering committee decides on the continuation of the trial.

**Sample size**
It is anticipated that probiotics will lead to a reduction of infectious complications from 50% (% of patients) to 30%. The sample size calculation is based on \( \alpha = 0.05 \), and a power of 80% This leads to a required sample size of 188 patients. Taking into account a 5% loss-to-follow up, a total of \( 2 \times 100 \) patients will be randomised. Based on hospital data of 2002 about 500 patients have to be included in order to randomise 200 patients with predicted severe acute pancreatitis. There is one post-discharge follow-up after three months. The expected study end is in 2006 (2 years inclusion period).
Snapinn stopping boundaries for one-sided interim p-value

Continuation region for A placebo B probiotica

Continuation region for A probiotica B placebo

Probiotica interim p-value

Probiotica final p-value

A probiotica B placebo bold. Vice-versa plain
simulated trial under H0 - - - H1 ... Probiotica p-values o o
Theory: a protoypical testing problem

\[ X_1, \ldots, X_N \sim \mathcal{N}(\mu, 1) \]

\[ H_0 : \mu = 0 \]
\[ H_1 : \mu > 0 \]
Data is obtained sequentially; time = \( n \)

\[
n = 1, \ldots, N
\]

\[
S_n = \sum_{i=1}^{n} X_i
\]
\[
\bar{X}_n = \frac{1}{n} S_n
\]
\[
T_n = n^{\frac{1}{2}} \bar{X}_n = n^{-\frac{1}{2}} S_n
\]
\[
p_n = 1 - \Phi^{-1}(T_n)
\]

*Interim* sum, average, t-statistic, p-value
Fixed sample size design parameters:

level $\alpha$ for testing null $\mu = 0$

power $1 - \beta$ against alternative $\mu = \delta$, given fixed $\delta > 0$

standard normal quantiles: $\Pr(\mathcal{N}(0, 1) \leq z_p) = \Phi(z_p) = p$

$$T_N \sim \mathcal{N}(\sqrt{N}\mu, 1)$$

rejection region: $T_N > z_{1-\alpha}$, equivalently $p_N < \alpha$

$$z_\beta = z_{1-\alpha} - \sqrt{N}\delta$$

$$\sqrt{N}\delta = z_{1-\alpha} - z_\beta$$

$\alpha = 0.025, \beta = 0.2; \; z_{1-\alpha} \approx 1.96, z_\beta \approx -0.84$
Interim analysis at time $n$

Interim fraction $n/N = f$

Given $T_n = t_n$

$$T_N = \frac{\sqrt{n}}{\sqrt{N}} t_n + \frac{1}{\sqrt{N}} \mathcal{N}\left((N - n)\mu, N - n\right)$$

$$= \sqrt{f} t_n + \mathcal{N}\left(\sqrt{N}(1 - f)\mu, 1 - f\right)$$
Snapinn lower boundary: *stopping for significance*: early rejection

*Reject early* if given $T_n$ probability of ultimate rejection is *larger* than some critical (large) $p_{\text{rej}}$ (e.g., $=0.90$).

$$\Pr_\mu(\text{ultimately reject}|T_n = t_n) = 1 - \Phi\left(\frac{1}{\sqrt{1-f}}\left(z_{1-\alpha} - \sqrt{f} t_n - \sqrt{N(1-f)} \mu\right)\right)$$

Choice for $\mu$: compromise between neutral best guess, and least favourable scenario for *early rejection*

$$\hat{\mu} = f \cdot \bar{x}_n + (1 - f) \cdot 0$$
\[ \hat{\Pr}(\text{ultimately reject } \mid T_n = t_n) \]

\[ = 1 - \Phi\left( \frac{z_{1-\alpha} - \sqrt{f} t_n - \sqrt{N(1-f)f} t_n / \sqrt{n}}{\sqrt{1-f}} \right) \]

\[ = 1 - \Phi\left( \frac{z_{1-\alpha} - \sqrt{f} t_n - (1-f) \sqrt{f} t_n}{\sqrt{1-f}} \right) \]

\[ = 1 - \Phi\left( \frac{z_{1-\alpha} - \sqrt{f} (2-f) t_n}{\sqrt{1-f}} \right) \]

We reject early if this probability is larger than \( p_{\text{rej}} \)
Reject early if

\[ p_{\text{rej}} < 1 - \Phi \left( \frac{z_{1-\alpha} - \sqrt{f} (2 - f) t_n}{\sqrt{1 - f}} \right) \]

\[ \Leftrightarrow \quad z_{1-p_{\text{rej}}} > \frac{z_{1-\alpha} - \sqrt{f} (2 - f) t_n}{\sqrt{1 - f}} \]

\[ \Leftrightarrow \quad t_n > \frac{z_{1-\alpha} - \sqrt{1 - f} z_{1-p_{\text{rej}}}}{\sqrt{f} (2 - f)} \]

\[ \Leftrightarrow \quad p_n < 1 - \Phi \left( \frac{z_{1-\alpha} - \sqrt{1 - f} z_{1-p_{\text{rej}}}}{\sqrt{f} (2 - f)} \right) \]
Snapinn upper boundary: *stopping for futility*: early acceptance

Accept early if given $T_n$ probability of ultimate rejection is smaller than some critical (small) $p_{\text{acc}}$ (e.g., $=0.20$).

\[
\Pr_{\mu} (\text{ultimately reject} | T_n = t_n) = 1 - \Phi \left( \frac{1}{\sqrt{1-f}} \left( z_{1-\alpha} - \sqrt{ft_n - \sqrt{N(1-f)\mu}} \right) \right)
\]

Choice for $\mu$: compromise between neutral best guess, and least favourable scenario for early acceptance

\[
\hat{\mu} = f \cdot \bar{x}_n + (1-f) \cdot \delta
\]
\[
\hat{\text{Pr}}\left(\text{ultimately reject} \mid T_n = t_n\right)
\]
\[
= 1 - \Phi\left(\frac{z_{1-\alpha} - \sqrt{f} t_n - \sqrt{N}(1 - f)f t_n/\sqrt{n} - \sqrt{N} (1 - f)^2 \delta}{\sqrt{1 - f}}\right)
\]
\[
= 1 - \Phi\left(\frac{z_{1-\alpha} - \sqrt{f} t_n - (1 - f)\sqrt{f} t_n - (1 - f)^2 (z_{1-\alpha} - z_\beta)}{\sqrt{1 - f}}\right)
\]
\[
= 1 - \Phi\left(\frac{z_{1-\alpha} - \sqrt{f} (2 - f)t_n - (1 - f)^2 (z_{1-\alpha} - z_\beta)}{\sqrt{1 - f}}\right)
\]

We accept early if this probability is smaller than \( p_{\text{acc}} \)
Accept early if

\[ p_{\text{acc}} > 1 - \Phi \left( \frac{z_{1-\alpha} - \sqrt{f} (2 - f) t_n - (1 - f)^2 (z_{1-\alpha} - z_\beta)}{\sqrt{1-f}} \right) \]

\[ \iff \quad z_{1-p_{\text{acc}}} < \frac{z_{1-\alpha} - \sqrt{f} (2 - f) t_n - (1 - f)^2 (z_{1-\alpha} - z_\beta)}{\sqrt{1-f}} \]

\[ \iff \quad t_n < \frac{z_{1-\alpha} - (1 - f)^2 (z_{1-\alpha} - z_\beta) - \sqrt{1-f} z_{1-p_{\text{acc}}}}{\sqrt{f}(2 - f)} \]

\[ \iff \quad p_n > 1 - \Phi \left( \frac{z_{1-\alpha} - (1 - f)^2 (z_{1-\alpha} - z_\beta) - \sqrt{1-f} z_{1-p_{\text{acc}}}}{\sqrt{f}(2 - f)} \right) \]
Putting it all together

Treat $p_{\text{acc}}$ and $p_{\text{rej}}$ as *arbitrary tuning parameters*

Demand actual significance level = nominal level

equivalently: under null
probability accept early but ultimately reject (function of $p_{\text{acc}}$) =
probability reject early but ultimately accept (function of $p_{\text{rej}}$)

determines $p_{\text{acc}}$ as function of $p_{\text{rej}}$

final tuning of $p_{\text{rej}}$:
maintain power, versus increased early stopping
Snapinn, empirical findings:

1) For given $\alpha$ and $\beta$, given $p_{\text{rej}}$, $p_{\text{acc}}$ is almost constant in $f$

2) Sig. level is maintained under *multiple* interim analyses

3) Method works also for comparing binary fractions ...

1) Similar shape of curves: match for one $f$, matches all (illustration, next slide)

2) Probability of crossing both boundaries once is already small, probability of multiple crossings is negligible; most action at $f \approx 0.7$ (illustration)

3) Universality/invariance: $n$ times estimated treatment effect asymptotically a random walk (under the alternative, a random walk with drift)
Figure 1. Probabilities of false early rejection and false early acceptance, with $\alpha = 0.025$ and $\beta = 0.05$

Snapinn, 1992: tuning $p_{\text{acc}}$ to $p_{\text{rej}}$ simultaneously for all $f$. 
Stopping boundaries
(Probiotica parameters)

\[ \text{stop for futility: early accept} \]
\[ \Pr(\text{ultim. rej.}|T_n) < p_{\text{acc}} \]

\[ \text{stop for significance: early reject} \]
\[ \Pr(\text{ultim. rej.}|T_n) > p_{\text{rej}} \]

continue the trial
Stopping boundaries

(Probiotica parameters; p-value in logarithmic scale)

stop for futility: early accept
\[ \Pr(\text{ultim. rej.}|T_n) < p_{\text{acc}} \]

stop for significance: early reject
\[ \Pr(\text{ultim. rej.}|T_n) > p_{\text{rej}} \]

continue the trial
Stopping boundaries
(Probiotica parameters; zoom in on smaller p-values)

- **stop for significance: early reject**
  \[ \Pr(\text{ultim. rej.}|T_n) > p_{\text{rej}} \]

- **stop for futility: early accept**
  \[ \Pr(\text{ultim. rej.}|T_n) < p_{\text{acc}} \]

- **continue the trial**

(sample fraction)
The probiotica trial
test use of probiotics against infectious complication in acute pancreatitis

Compare two Bernoulli probabilities

\[ H_0 : p_1 = p_2, \quad H_1 : p_1 \neq p_2 \]

Design:

\[ H_0 : p_1 = p_2 = 0.3, \quad H_1 : p_1 = 0.3, p_2 = 0.16 \]
\[ \alpha = 0.05 \text{ two-sided}, \quad \beta = 0.20, \quad N = 300 \]

Snapinn according to Schouten (1995, *Klinische Statistiek*):

\[ f = 0.50, p_{\text{rej}} = 0.90, p_{\text{acc}} = 0.20 \]

Schouten: table for one-sided level 5% testing, table for two-sided level 5% testing

Snapinn: table for one-sided level 2.5% testing only
Schouten: “a two sided test can be thought of as a combination of two one-sided tests”

But: *early stopping* is not symmetric when we compare a new treatment to a standard

Giving an experimental treatment *which you can’t prove is beneficial* to the still-to-be-treated “treatment group” patients is *not* the same as giving a standard treatment *which you believe can be improved* to the still-to-be-treated “control group” patients
Interim analysis \( n = 184, \ f \approx 0.60, \ t_n \approx \pm 1 \)

If A=placebo, B=probiotica \( p_n = 0.16 \)

Snapinn one-sided \( \alpha = 0.025 \): continue

if A=probiotica, B=placebo \( p_n = 0.84 \)

Snapinn one-sided \( \alpha = 0.025 \): stop and accept
Snapinn stopping boundaries for one-sided interim p-value

Schouten two-sided continuation region = union of two one-sided continuation regions

Continuation region for A placebo B probiotica

Probiotica interim p-value

Probiotica final p-value

Continuation region for A probiotica B placebo

A probiotica B placebo bold. Vice-versa plain
Probiotica p-values o o
Digression: Schouten recommends smaller $p_{\text{rej}}$ than Snapinn – consequences?

$\Pr(\text{early acc late rej} \mid \text{null})$ and vice-versa

(Probiotica parameters)

Null errors balancing act:
$p_{\text{acc}}$ should have been lower, to bring the curves closer together

Loss of power:
Error of second kind with interim analysis at $f = 0.6$ is increased from 0.20 to 0.2865

Sample fraction probiotica trial parameters; higher curve early acc late rej
Data distributed at press conference, 13 February 2008

<table>
<thead>
<tr>
<th></th>
<th>Groep A n = 94</th>
<th>Groep B n = 90</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Man</td>
<td>58</td>
<td>48</td>
<td>0.297</td>
</tr>
<tr>
<td>Binnen 3 dagen naar IC</td>
<td>23</td>
<td>14</td>
<td>0.145</td>
</tr>
<tr>
<td>Meer dan 30% necrose</td>
<td>15</td>
<td>13</td>
<td>0.837</td>
</tr>
<tr>
<td><strong>Alle infecties</strong></td>
<td><strong>29</strong></td>
<td><strong>22</strong></td>
<td><strong>0.323</strong></td>
</tr>
<tr>
<td>geinf pancreas necrose</td>
<td>12</td>
<td>8</td>
<td>0.478</td>
</tr>
<tr>
<td>luchtweginfectie</td>
<td>11</td>
<td>5</td>
<td>0.189</td>
</tr>
<tr>
<td>urineweginfectie</td>
<td>9</td>
<td>4</td>
<td>0.248</td>
</tr>
<tr>
<td>positieve bloedkweek</td>
<td>13</td>
<td>16</td>
<td>0.545</td>
</tr>
<tr>
<td>Pos bloedkweek en IPN</td>
<td>20</td>
<td>17</td>
<td>0.714</td>
</tr>
<tr>
<td>Operatie</td>
<td>14</td>
<td>8</td>
<td>0.256</td>
</tr>
<tr>
<td>Mortaliteit</td>
<td>14</td>
<td>6</td>
<td>0.097</td>
</tr>
</tbody>
</table>

Thanks to Hester van Santen, NRC
What went wrong? Possibly, both of:

- The monitoring committee was blinded and optimistic about the treatment, saw the $p$-value of the one-sided test as if A=placebo

- The monitoring committee saw the $p$-value of the two-sided test and used Schouten’s table for two-sided testing
• The DMC should have considered both scenarios A, B = ... with corresponding one-sided $p$-values 0.16, 0.84

• A=placebo: $p_n = 0.16$; continue the trial since there is a good chance of getting a significant positive result

• A=probiotica: $p_n = 0.84$; stop the trial since there is almost no chance of getting a significant positive result

• The DMC should have de-blinded the data to determine which action to take

• Instead, the trial was in effect continued because there was a good chance of proving probiotica bad almost no chance of proving it good
Conclusions, recommendations

- “Early stopping” of RCTs raises complex issues and requires professional statistical expertise

- Data monitoring, safety and ethics committees should include an independent statistician who is not blind to the actual treatment assignments

- Primary endpoint should include all “serious adverse events” so that early stopping rule can incorporate safety concerns

- Snapinn parameters $p_{acc}, p_{rej}$ need to be tuned to actual trial parameters $\alpha, \beta$

- Reporting of “serious adverse events” should not be left to the discretion of the doctors at local centres, in a multi-centre trial
References

- Besselink et al. (2004) *BMC Surgery*
- Besselink et al. (2008) *Lancet*
- Gill (2008) *Statistica Neerlandica*
- Snapinn (1992) *Statistics in Medicine*
- Schouten (1995) *Klinische Statistiek*
Study protocol

Probiotic prophylaxis in patients with predicted severe acute pancreatitis (PROPATRIA): design and rationale of a double-blind, placebo-controlled randomised multicenter trial [ISRCTN38327949]

Marc GH Besselink,* 1, Harro M Timmerman1, Erik Buskens2, Vincent B Nieuwenhuijs1, Louis MA Akkermans1, Hein G Gooszen1 and the members of the Dutch Acute Pancreatitis Study Group

Address: 1Department of Surgery, University Medical Center Utrecht PO Box 85500, 3508 GA Utrecht, The Netherlands and 2Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht PO Box 85060, 3500 AB Utrecht, The Netherlands

Email: Marc GH Besselink* - m.besselink@chir.azu.nl; Harro M Timmerman - h.timmerman@chir.azu.nl; Erik Buskens - e.buskens@umcutrecht.nl; Vincent B Nieuwenhuijs - v.b.nieuwenhuijs@btem.nl; Louis MA Akkermans - l.m.a.akkermans@chir.azu.nl; Hein G Gooszen - h.gooszen@chir.azu.nl; the members of the Dutch Acute Pancreatitis Study Group - info@pancreatitis.nl

* Corresponding author

Abstract

Background: Infectious complications are the major cause of death in acute pancreatitis. Small bowel bacterial overgrowth and subsequent bacterial translocation are held responsible for the vast majority of these infections. Goal of this study is to determine whether selected probiotics are capable of preventing infectious complications without the disadvantages of antibiotic prophylaxis; antibiotic resistance and fungal overgrowth.

Methods/design: PROPATRIA is a double-blind, placebo-controlled randomised multicenter trial in which 200 patients will be randomly allocated to a multispecies probiotic preparation (Ecologic 641) or placebo. The study is performed in all 8 Dutch University Hospitals and 7 non-University hospitals. The study-product is administered twice daily through a nasojejunal tube for 28 days or until discharge. Patients eligible for randomisation are adult patients with a first onset of predicted severe acute pancreatitis: Imrie criteria 3 or more, CRP 150 mg/L or more, APACHE II score 8 or more. Exclusion criteria are post-ERCP pancreatitis, malignancy, infection/sepsis caused by a second disease, intra-operative diagnosis of pancreatitis and use of probiotics during the study. Administration of the study product is started within 72 hours after onset of abdominal pain. The primary endpoint is the total number of infectious complications. Secondary endpoints are mortality, necrosectomy, antibiotic resistance, hospital stay and adverse events. To demonstrate that probiotic prophylaxis reduces the proportion of patients with infectious complications from 50% to 30%, with alpha 0.05 and power 80%, a total sample size of 200 patients was calculated.

Conclusion: The PROPATRIA study is aimed to show a reduction in infectious complications due to early enteral use of multispecies probiotics in severe acute pancreatitis.
Probiotic prophylaxis in predicted severe acute pancreatitis: a randomised, double-blind, placebo-controlled trial


Summary

Background Infectious complications and associated mortality are a major concern in acute pancreatitis. Enteral administration of probiotics could prevent infectious complications, but convincing evidence is scarce. Our aim was to assess the effects of probiotic prophylaxis in patients with predicted severe acute pancreatitis.

Methods In this multicentre randomised, double-blind, placebo-controlled trial, 288 patients with predicted severe acute pancreatitis (Acute Physiology and Chronic Health Evaluation [APACHE II] score >35, 1min score >3, or C-reactive protein >150 mg/L) were randomly assigned within 72 h of onset of symptoms to receive a multispecies probiotic preparation (n=153) or placebo (n=145), administered enterally twice daily for 28 days. The primary endpoint was the composite of infectious complications—ie, infected pancreatic necrosis, bacteraemia, pneumonia, urosepsis, or infected abscesses—during admission and 90-day follow-up. Analyses were by intention to treat. This study is registered, number ISRCTN38327949.

Findings One person in each group was excluded from analyses because of incorrect diagnoses of pancreatitis; thus, 275 individuals in the probiotics group and 274 in the placebo group were analysed. Groups were much the same in terms of patients’ characteristics and disease severity. Infectious complications occurred in 46 (30%) patients in the probiotics group and 41 (28%) of those in the placebo group (relative risk 1.06, 95% CI 0.75–1.51). 24 (16%) patients in the probiotics group died, compared with nine (6%) in the placebo group (relative risk 2.53, 95% CI 1.22–5.25). Nine patients in the probiotics group developed bowel ischaemia (eight with fatal outcome), compared with none in the placebo group (p=0.004).

Interpretation In patients with predicted severe acute pancreatitis, probiotic prophylaxis with this combination of probiotic strains did not reduce the risk of infectious complications and was associated with an increased risk of mortality. Probiotic prophylaxis should therefore not be administered in this category of patients.

Introduction

The incidence of acute pancreatitis in Europe and the USA is increasing by about 3% per year, mainly owing to an increase in biliary pancreatitis.1–3 About a fifth of patients will develop necrotising pancreatitis, which is associated with a 10–30% mortality rate, mostly attributed to infectious complications and infection of (peri)pancreatic necrotic tissue in particular.4 These infections are thought to be the sequelae of a cascade of events starting with small-bowel bacterial overgrowth, mucosal barrier failure, and a proinflammatory response leading to bacterial translocation of intestinal bacteria.5–7 Systemic antibiotic prophylaxis has long been studied as a measure to prevent secondary infection in acute pancreatitis.1 However, two double-blind, placebo-controlled trials32 and two meta-analyses53–55 have failed to show a beneficial effect, and many clinicians have abandoned this strategy. In the two antibiotic trials, the incidence of extrapancreatic infections (eg, bacteraemia, pneumonia) and pancreatic infection remained high.56 Consequently, there is a clear need for other strategies to prevent infectious complications in patients with acute pancreatitis.

Probiotics, as an adjunct to enteral nutrition, have raised high expectations and are currently gaining worldwide popularity for their presumed health-promoting effects.10–13 Certain strains of probiotic bacteria might prevent infectious complications by reducing small-bowel bacterial overgrowth, restoring gastrointestinal barrier function, and modulating the immune system.57–59 A reduction of infectious complications has been reported in several clinical studies with probiotics in patients undergoing elective abdominal operations60–63 and in patients with acute pancreatitis.63 However, because of their small size and methodological quality, these studies do not justify global implementation of probiotics as a preventive measure in acute pancreatitis. Accordingly, we embarked on a nationwide multicentre randomised, double-blind, placebo-controlled trial—the PRObitics in Pancreatitis TRIAl (PROPATRIA)—to assess the effects of probiotic prophylaxis in patients with predicted severe acute pancreatitis.

Methods

Patients The design and rationale of the study have been described in detail elsewhere.64 Adult patients admitted with a first...
Abstract

A randomized clinical trial comparing an experimental new treatment to a standard therapy for a life-threatening medical condition should be stopped early on ethical grounds, in either of the following scenarios: (1) it has become overwhelmingly clear that the new treatment is better than the standard; (2) it has become overwhelmingly clear that the trial is not going to show that the new treatment is any better than the standard. The trial is continued in the third scenario: (3) there is a reasonable chance that the new treatment will finally turn out to be better than the standard, but we are not sure yet.

However, the (blinded) data monitoring committee in the “PROPATRIA” trial of an experimental probiotica treatment for patients with acute pancreatitis allowed the trial to continue at the half way interim analysis, in effect because there was still a good chance of proving that the probiotica treatment was very harmful to their patients. The committee did not know whether treatment A was the probiotica treatment or the placebo. In itself this should not have caused a problem, since it could easily have determined the appropriate decision under both scenarios. Were the decisions in the two scenarios different, then the data would have to be de-blinded, in order to determine the appropriate decision. And the decisions in this trial were indeed different: were A the probiotica treatment, the monitoring committee would have found themselves in scenario (2) and would have stopped the trial; were A the placebo, the monitoring committee would have found themselves in scenario (3) and would have let the trial continue.