Randomized response surface pathway design with odd response outcomes in a Latin Square designed study

Trond Holand1
Kristian Ellingsen2
Sagita Dewi1,3
Stig Larsen1

1Centre for Epidemiology and Biostatistics, Department of Production Animal Clinical Sciences, Faculty of Veterinary Science, Norwegian University of Life Science, 1Norwegian Veterinary Institute, Oslo, Norway; 2Bali Indera Hospital, Government of Bali Province, Denpasar, Bali, Indonesia

Background: Response surface pathway (RSP) design has been recently developed and described for dose-finding studies. The aim of this paper was to introduce and demonstrate additional procedures to strengthen and generalize the design and combine RSP with classical study design.

Materials and methods: Nine bull calves and six heifer calves were included in an unbalanced 2 × 2 Latin Square (LSQ) designed study with large and small aperture bottle teats. The two LSQ sequences were performed with independent randomized three-level between-patient RSP design with an odd number of response classifications. The milk temperature window was 8°C–38°C with a mid temperature of 23°C. X-rays of the abdominal cavity were taken before, during, and immediately after intake of milk and recorded as “milk”, “trace”, or “no milk”. Based on the results of the first design level, the milk temperatures for five calves in the second design level were obtained by a randomization procedure. A similar procedure was performed for seven calves in the third design level. Adjustment of the dose from one design level to the next was based on a k-adjustment factor estimated to ensure coverage of the entire predefined dose window.

Results: Starting with a low number of subjects and increasing this number with increasing design levels reduces the sample size without reducing the power. The suggested randomization procedure worked as expected. No milk in Rumen was recorded and minimum milk temperature (MMT) for large bottle teat was estimated to be ≤8°C for both teats. The odd number of response categories increases the flexibility of RSP, enabling category redefinition in hindsight from “trace” to “uncertain”. After category redefinition, MMT for large bottle teat was estimated to be 14.3°C (95% confidence interval: 8.3–20.3°C), but 8°C for small.

Conclusion: The suggested changes and additional procedures increase the strength and flexibility of the RSP design.

Keywords: k-adjustment factor, Latin Square design, patient reduction in clinical trials, randomized response surface pathway design

Background

During the first few weeks of life, milk is the primary source of nutrition for the dairy calf. Over the milk feeding period, the calf has an anatomical feature called the esophageal groove or sulcus reticuli. The groove, when activated, turns into a tube allowing the milk to pass by the reticularum straight into the abomasum where the milk is digested.1 Activation of the reticular groove is caused by many factors, including the smell and taste of milk, milk temperature, sucking behavior, and body posture while drinking.2 If the esophageal groove is not properly closed, milk can enter Rumen, the...
first number of outcomes, and combinations with other clas-
cal up-and-down procedure with common response
rusum. In very young animals, this may not pose a problem
the Rumen is emptied into the abomasum within a few
nicians, the methodological aim behind increasing patients with increasing design level is intuitively
at least sufficient feeding condition, very good general health
s were classified with at least sufficient feeding condition, very good general health
resents a challenge for researchers worldwide.10 Eth-
al level based on the outcome from the previous study? 

Choice of study design for an interventional uncontrolled
ical up-and-down procedure with common response
sibility. These opposing goals make it difficult to choose
s in a laboratory setting, but has not yet been tested in
equal number strategy with patient-increase procedure intro-
duces the need for randomization procedure. Randomization
is one of the cornerstones of the randomized controlled trial
(RCT) and a true random allocation procedure should be

The aim of this paper is to introduce and demonstrate RSP
with increasing patient number along with increasing design
levels, randomization procedure between design levels, odd
number of study outcomes, and RSP used in an LSQ designed
study to estimate MMT for young calves using small and
large aperture bottle teats.

Materials and methods
Nine bull calves and six heifer calves were used in the devel-
opment of this randomized RSP designed study with odd
response classification. The animals were classified with at
least sufficient feeding condition, very good general health
condition, and no abnormalities. The mean age of the calves
was 18.5 days (range: 9–27 days) with a mean body weight
of 54.9 kg (range: 45.5–71.0 kg). After being evaluated by the
Norwegian Animal Research Authority, the study was exempt
from having to apply for approval because the participating
animals served as a model in a non-invasive feeding experi-
ment with no expected side effects.

Study performance
The study is a randomized single-center study with an unbal-
canced 2 × 2 LSQ design. Each calf participated twice and was
offered 4 L of whole milk per trial, once via a small (S) and
once via a large (L) aperture bottle teat. In accordance with
LSQ, the calves randomly received the milk in the sequence
S–L or L–S. To control the influence of a possible sequence
effect and to allow potential milk in the Rumen to be passed on, washout period of 2 days between the two feedings was included.

A between-patient RSP design with milk temperature as the interventional variable and milk in Rumen as response was independently used for the small and the large aperture teat arms. The milk temperature with the given bottle teat to be used at one design level is based on the results obtained on the previous design level. The start temperature on the first design level was 23°C with a lower limit of 8°C and an upper limit of 38°C. Each calf was offered 4 L of milk with the given temperature. Milk consumption was recorded in liters, and the duration of the feeding in seconds. X-rays of the abdominal cavity were taken before, during, and immediately after feeding, and milk in the Rumen was recorded as “milk”, “trace”, or “no milk”. If no difficulties were encountered, the second feeding with the other bottle teat could start. Adverse events were registered the following 2 days.

Ethics approval and consent to participate
Norwegian national authority-approved protocols were followed with regard to the animal welfare during our study following our publication ethics standard. Consent for publication is not applicable.

RSP design
The RSP methodology has previously been presented both for within and between different groups of patients with one interventional- and one response variable. In order to capture, utilize, and adjust according to accumulated clinical information during the study, the number of included subjects is reduced to a minimum in the first design level and increases with increasing level. In the between-patient RSP design, dimension groups of three patients are allocated into sequences with the same number of patients as the number of levels in the design. In this between-patient model, the RSP design consists of n levels and r independent patient sequences. The first patient in the sequence receives the pre-defined starting dose. The response from this patient determines the dose for the second patient in the same sequence. In general, the response from the patient at design level i (1 ≤ i ≤ n) determines the dose for the patient at the next design level i + 1. The sequences develop a random walk and describe a pathway with equal number of patients at each design level.

Development of the methodology in laboratory animals and simulations demonstrated that this equal number strategy is not an optimal solution. By starting with a low number of patients at the first design level and increasing this number with increased level, the sample size reduces without reduction in accuracy. This is possible in the between-patient RSP design using randomization.

Randomized between-patient RSP design
To optimize the RSP model, the number of patients reduces to a minimum of ≥2 in the first design level and increases with increasing level. The previously recommended patient-increase procedure starts with three patients at level 1 increasing to five, seven, nine etc., at level 2 and upward. The three calves allocated to the first design level receive 4 L of milk with a temperature of 23°C. Assume that a represents the number of calves on the first design level indicating milk temperature A, for the second design level, a pointing out milk temperature A, and a pointing out milk temperature A. In general, a weighted randomization in ratio (a:1:2:2:3) will be appear. If two doses are equal (a=a) and the third dose is deviating, the calf on the second level will be weighted randomized (2:1). This means that the probability for a calf to be allocated to milk temperature A is 3/2 and 1/3 for milk temperature A (Table 1).

Based on the milk in Rumen classification obtained in the five calves on the second design level, theoretically five new milk temperatures (B1B2B3B4B5) allocate to the seven patients in the third design level. Assume that b1 calf on the second design level calculated by the RSP procedure to receive intervention B, b1 to receive dose B, b1 to receive dose B, b1 to receive dose B, and b1 to receive dose B. The milk temperature to be used on the third design level for the seven new calves will then be allocated by weighted complete randomization (b,b,b,b,b) following the same procedure as explained for the second design level. By replacing the previously described sequence-procedure in allocation to design levels with this randomization procedure, while keeping the other part of the RSP, a randomized between-patient RSP designed RCT is introduced.

Combination of randomization results in a 2 × 2 LSQ design
Let N1, N2, and N3 denote the RSP-calculated milk temperature for the small teat of the bottle opening and M1, M2, and M3 denote the same for the large teat. The three calves in level 1 indicate the three milk temperature combinations: “N1–M1”, “N2–M2”, and “N3–M3”.

These three milk temperature combinations for small and large teats of the bottle openings will be used as the basis.
for allocation of the five calves participating in the second design level.

A similar procedure is performed for the third design level. Let N4–N8 denote the RSP-calculated milk temperature recommended for the small teat on the bottle, while M4–M8 for large bottle teat on the third design level. The basis for randomization of the seven calves in design level 3 will be “N4–M4”, “N5–M5”, “N6–M6”, “N7–M7”, and “N8–M8”. This procedure builds on and considers the dependency within calves.

### Dose adjustment procedure in RSP design

Let m denote the starting dose, mi the dose at design level i, and k the dose adjustment factor. The dose at design level i is given by equation (1). Let Du denote the upper limit of the interventional variable and n the number of design levels, then Du is given as the sum of a geometric series in equation (2).15

\[
\begin{align*}
  m_i &= m_{i-1} \pm \frac{m}{k^{i-1}} \\
  D_u &= \sum_{i=1}^{n} m \left( \frac{1}{k} \right)^{i-1} \\
  D_u &= \frac{m(k^n - 1)}{(k^n - k^{n-1})}
\end{align*}
\]

With known upper limit of the interventional variable, the starting value m, and design level n, the k-adjustment factor is calculated from equation (2).
Escalation and de-escalation procedure

The response variable is multinomial with unequal number of categories denoted as \(2c+1\). Of these possible \(2c+1\) response values, one of the \(c\)'s gives escalation, one \(c\) de-escalation, and one remaining unchanged value of the interventional variable for the patients in the next design level. It may be convenient to use the mid value of the predefined dose window as the starting value denoted as \(m\). To ensure coverage of the dose window, a dose adjustment procedure is established.

Assume that the response variable has a sample space \(\{1, 2, \ldots, 2c, 2c+1\}\), the median value will then be \(c + 1\). Let \(i\) represent the new design level, \(j\) the outcome from the previous dose level \(i - 1\), and \(h\) the outcome from the dose level \(i - 1\). For calculation of the first dose level \(h = 1\) per definition. Let \(m_0 = m\). In this case, there are only three outcome values giving \(c = 1\); then for design level \(i \geq 2\) the interventional value is given as:

\[
\begin{align*}
\hat{m}_1 &= m_1 + m / k^{i+j+h-2c-1} &; j \leq c &\text{&} h \leq c \\
\hat{m}_1 &= m_1 &; j = c+1 &\text{&} h \leq c \\
\hat{m}_1 &= m_1 - m / k^{i+j+h-2c} &; j > c+1 &\text{&} h \leq c \\
\end{align*}
\]

\[
\begin{align*}
\hat{m}_2 &= m_2 + m / k^{i+j+h-2c} &; j \leq c &\text{&} h = c+1 \\
\hat{m}_2 &= m_2 &; j = c+1 &\text{&} h = c+1 \\
\hat{m}_2 &= m_2 - m / k^{i+j+h-2c} &; j > c+1 &\text{&} h = c+1 \\
\end{align*}
\]

The equations above calculate the RSP dose adjustment procedure for the next design level.

Statistical analysis

The sample space of the milk temperature may be expressed as \(\Omega = \{D_L \leq \ldots \leq D_U\}\). Let \(\mu_d\) represent MMT for newborn calves and assume \(\mu_d\) covers by \(\Omega_d\). Milk in Rumen is ordinal in interventional variable, and the probability increases monotonically over the interventional levels. Isotonic regression is the suggested model for analyzing the material.16 Continuously distributed variables are expressed as mean values, standard deviation (SD) in brackets, and 95% confidence interval.17

Results

At 23°C milk temperature in design level 1, no milk was found in Rumen, neither with small nor large aperture teat of the bottle (Figure 2). However, one calf failed to consume more than two of the four prescribed liters with large teat. No randomization was needed, and all the five calves at design level 2 received 13°C milk with both teat sizes. Two calves were not able to consume the 4 L at this temperature using large teat. No milk in Rumen was obtained with neither of the two teat sizes and all the seven calves at design level 3 received the prescribed 4 L of milk with the temperature of 8°C (Figure 2). No milk in Rumen was detected in these seven calves neither with small nor large aperture teat of the bottle. However, two of the calves were not able to consume >2 and 3 L using small teat and one calf consumed 2.5 L with large teat. Without any adjustments for \(<4\) L milk consumption, the estimated MMT independent of teat size is 8°C.

The classification used in the RSP procedure was only related to the X-ray findings without considering the deviation from the prescribed 4 L of milk. Four calves deviated from the prescribed amount when using large aperture teat and two calves when using small aperture teat. By expanding the classification “Trace” to “Uncertain”, including the situation when a calf was not able to consume the prescribed 4 L, the results change. For small teat, the pattern will be nearly unchanged except from design level 3. Instead of ending with seven calves without milk in Rumen at 8°C, the results change to five calves without milk and two classified as “Uncertain” (Figure 3). This results in an estimated MMT of 8°C (95% confidence interval [CI]: 6.2–9.8°C) for small teat.

For large teat, the change in results was more pronounced (Figure 4). In design level 1 with milk temperature of 23°C, two calves were classified as “No milk” and one as “Uncertain”. The randomization of five calves to the second design level will be 13°C and 23°C in a 2:1 ratio. The outcome with the largest probability will be three calves to 13°C and two to 23°C. Assuming the two calves receiving milk with temperature 23°C at design 2 level will consume all 4 L without any milk in Rumen, this gives two times 21°C as contribution to the randomization for design level 3. Based on the study results, the probability for “No milk” with a temperature of 13°C is 3/5 and 2/5 for “Uncertain”. The outcome with highest probability will be two “No milk” and one “Uncertain”. The background for randomization to design level 3 will then be (21°C, 21°C, 13°C, 8°C, and 8°C). The expected number of calves randomized to 8°C and
21°C is 2.8, and the expected number for 13°C is 1.4. The most probable outcome of the randomization will be three calves receiving milk with the temperature 21°C, one with milk temperature 13°C, and three with milk temperature 8°C. From the obtained study results, the most probable outcome with large teat is “No milk” for all seven calves for the given set of milk temperatures. This results in an estimated MMT of 14.3°C (95% CI: 12.1–16.5°C) for large teat. In the LSQ analysis, the difference between small and large teat will reach the 5% significance level.

**Discussion**

Despite missing the targeted 4 L milk consumption in six of the 30 milk feedings, the study design was with sufficient accuracy able to predict likelihood of “milk” within an 8–38°C temperature window while indicating influence of small and large bottle teat sizes. The study design worked well, presented the opportunity to reinvestigate obtained results after reclassification of the mid outcome category, and demonstrated an efficient pathway toward the area of clinical interest.

RSP methodology in patient-related dose–response study has earlier used an equal number of included patients at each design level. Development of the concept in laboratory animals and computer simulations demonstrated that this is not an optimal solution. Starting with a low number of subjects and increasing this number with succeeding design levels reduces the sample size without reducing the accuracy of the estimates. This procedure is not possible in the within-patient situation, but can be implemented in the
between-patient RSP design by introducing a randomization procedure between design levels. Advantages from increasing the patient number, accompanied with an optimized randomization procedure between increasing design levels, have earlier been demonstrated.

Earlier studies using RSP methodology have used an odd number of patient-increase procedure starting with three patients at level 1, increasing to 5, 7, 9, and so on at level 2 and upward. In the present study, 15 calves were exposed twice and an odd number of three outcome categories recorded. Because of uncertainty regarding exact interpretation of X-ray findings, categories “Yes”, “Trace”, and “No” were chosen. The basic RSP model was introduced in an LD$_{50}$ study requiring an odd number of mice in each design level, because LD$_{50}$ was between the escalation and de-escalation pathways. The RSP methodology can, however, also be used in an even number of patient-increase procedure coupled with even or odd outcomes. Introducing odd outcomes can in this way unlock statistical strength and usability when an even number of included patients in each design level is feasible. Instead of an RSP sample size 3 + 5 + 7 in design level 1, 2, and 3 including 15 subjects, a 2 + 4 + 6 or 4 + 6 + 8 sample size with odd outcomes for 12 or 18 subjects might be sufficient to maneuver the pathway to statistical and clinical relevance. Applied in the present trial, an RSP sample size of 12 subjects with the given outcomes, the same study results with no X-ray findings would have been reached. Introduction of an “Uncertain” category into study design can address
unexpected or new findings during clinical trials. In many clinical trials, for instance when clear cut interpretation of clinical outcomes is complex, use of an odd outcome category “Uncertain” between “Yes” and “No” could improve the study design flexibility. In the present study, we used this developed RSP methodology to handle odd outcomes coupled with an odd number of patient-increase procedure. The obtained results with no X-ray findings in any of the 30 feedings of the 15 calves indicate a large statistical outcome accuracy. 

Regarding the recorded odd outcomes, the results show interesting opportunities for how the RSP method can be further developed to handle study outcome uncertainty. Lack of X-ray findings in any of the feedings without considering the uncertainty arising from the 4 L milk consumption deviations introduces new questions around interpretation of the present study outcomes. In six of the 30 feedings, the calves did not manage to drink all the prescribed milk. Regardless of the many potential physiological reasons for and relevance of this protocol deviation, it exposes an observed clinical finding that needs to be addressed. Reclassification of the category “Trace” to “Uncertain” includes the uncertainty around the observed <4L milk consumption. This unlocks an alternative RSP pathway and is one way of addressing uncertainty of achieved outcomes in hindsight. Analyzing how an outcome

Figure 4 The adjusted RSP pathway for large test after reclassifying “Trace” to “Uncertain”.

Notes: The category (blue) is based on milk temperature in °C (yellow) for large bottle test. The end results are given in green. “Trace” outcome category reclassification to “Uncertain” in the case of <4 L milk consumption.

Abbreviation: RSP, response surface pathway.
Maneuvering one continuous outcome variable like milk temperature or drug dose window in RSP together with odd number of ordinal outcomes like yes/trace/no or yes/uncertain/no requires a randomization procedure between design levels. For RSP, the three temperature outcomes \( a_{1-3} \) in design level 1 indicates at the next dose \( k \)-factor-adjusted temperature \( T_{A_{1-3}} \) to undergo randomization and they are selected to pinpoint the respective five temperatures clinically used in design level 2. Similarly, the five temperature-adjusted \( A_{1-3} \) for small and large teats, respectively, give rise to the five next temperatures which in turn in design level 2 produce the clinical outcomes \( b_{1-5} \). Furthermore, the five temperature outcomes in design level 2 will undergo randomization and in turn point out the selected pathway for the seven next temperatures to be used in design level 3. This randomization procedure was found appropriate in the present material.

The additional RSP methodology procedures introduced in this trial also include combination with LSQ study design to compare the influence of small and large teats for milk feeding. Despite an unbalanced \( 2 \times 2 \) LSQ in the present study, the use of RSP in each LSQ teat sequence can give increased flexibility, reduced sample size, and new opportunities in RCT compared to alternative study design approaches. An improvement of this combined RSP-LSQ study design could be to use an even number of included patients. By changing the number of patients at the three design levels from 3, 5, 7 to 2, 4, 8, the present unbalanced \( 2 \times 2 \) LSQ could be performed balanced and hence increase the power of the study. In the present study, the skewness due to the odd number of included subjects may only give a small and limited influence on the results, but including an even number of subjects would improve the RSP-LSQ design.

The combination of the randomization performed separately for small and large bottle teats may have been developed differently. Due to the use of LSQ in the present study, randomization between small and large teats to give rise to three different temperature-teat combinations. These combinations were used as basis for allocation of the five calves in the second design level. By the same strategy, the five calves in design level 2 give rise to five combinations for allocation in design level 3. In this way, both milk temperature and teat size are randomized between each study design level, each individual calf serving as its own control in the LSQ teat sequence, creating a combined randomized RSP in a LSQ designed trial.

Reduction in sample size while increasing statistical accuracy is an important goal for developing and improving clinical trial methods. Stipulating outcome accuracy and calculating sample size in alternative study designs with the same study aims will of course be influenced by the chosen model. From both an ethical and statistical point of view, the sample size must be calculated in the planning stage of the study. In the present \( 2 \times 2 \) LSQ design, the number of included subjects for each outcome category would have to be 12 or 18 to obtain clinically relevant differences within a 95% or 99% power. Dividing the temperature window into conservative \( 5 \)°C incremental temperature intervals would give six outcome categories. To cover all these categories in a classical \( 2 \times 2 \) LSQ design with a significance level of 5% and a power of 95% would require inclusion of \( 12 \times 6=72 \) subjects. Furthermore, a significance level of 1% and a power of 99% would require inclusion of 108 subjects in a traditional \( 2 \times 2 \) LSQ design. Even though it is difficult to quantify the outcome accuracy differences between study designs without simulations, this comparison indicates statistical and clinical superiority of the RSP-LSQ combination over a traditional LSQ strategy.

Combination of RSP in an LSQ designed study for maneuvering one continuous and one discrete ordinal outcome variable demonstrates how RSP can be developed to an RCT. Inclusion of two or more interventional variables in higher dimensional RSP designs, maneuvering continuous outcome variables for instance in dose finding studies, could further improve the methodology.

**Conclusion**

The randomized between-patient RSP in an LSQ-designed RCT detected its power and predicted the outcome variable with high accuracy. The study demonstrates advantages of RSP methodology with increasing patient number along with increasing design levels, use of odd number of outcomes to handle uncertainty, randomization procedure between design levels, and use of RSP in an LSQ designed study.

**Acknowledgments**

The authors would like to thank clinicians and staff members at the Norwegian University of Life Science, Faculty of Veterinary Science, Department of Production Animal Clinical
Sciences and the Norwegian Veterinary Institute for their participation in the study. The data sets used and analyzed during the current study are available from the corresponding author on reasonable request. The study was funded by the Norwegian University of Life Science, Faculty of Veterinary Science, Department of Production Animal Clinical Sciences, Centre for Epidemiology and Biostatistics and the Norwegian Veterinary Institute, Oslo, Norway.

**Author contributions**
KE was responsible for study protocol, clinical trial conduct, X-ray interpretations, and data gathering. SD performed the methodological response surface pathway (RSP) review interpretation and gave suggestions for improvements in the dose adjustment procedures. TH performed the overall manuscript write up, RSP methodological development and descriptions. SL developed, analyzed, and interpreted the statistical procedures for the RSP methodology. All authors contributed toward data analysis, drafting and critically revising the paper and agree to be accountable for all aspects of the work.

**Disclosure**
The authors report no conflicts of interest in this work.

**References**
2. Lateur-Rowet HJ, Breukink HJ. The failure of the esophageal groove reflex, when fluids are given with an esophageal feeder to newborn and young calves. *Vet Q*. 1983;5:68–74.