# The birthday problem: repeated sampling of animal populations and ethics of experimental design

D.M. Green [a,*], C.G. Mitchell [b]

[a] *Institute of Aquaculture, University of Stirling, United Kingdom FK9 4LA, United Kingdom*
[b] *WellFish Tech, University of the West of Scotland, F117 Henry Building West, Paisley PA1 2BE United Kingdom*

A B S T R A C T

Researchers who use animals in science must balance statistical power with the need to satisfy the three Rs, whereby researchers are required to *reduce* numbers of animals, *refine* what they experience, and use alternatives to (*replace*) higher animals where possible. In repeated sampling over time-series studies, there is potential loss of power as well as ethics implications posed by repeated sampling of individual animals, where this cannot be managed or avoided. Here, we consider the mathematics of repeated sampling from three perspectives: that of the population at large, from the experience of the individual, and the conditional probability of sampled individuals being sampled again. The calculations are illustrated using four theoretical case studies across veterinary epidemiology with different practical implications and a provided R Shiny tool for researchers. Despite the availability of exact calculations, it is necessary to also consider the biological factors which may affect capture and recapture rates in sampling studies such as animal personality and response to capture. Researchers must also choose their question carefully to avoid inappropriate framing of ethical concerns around repeated sampling.

© 2024 The Authors. Published by Elsevier B.V. on behalf of The Animal Consortium. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

## Implications

When animal populations are subject to repeated sampling in time-series studies, there is the possibility of sampling the same animal twice, potentially detrimental to welfare and a possible cause for ethical concern. In this paper, we show how often individual animals might be repeat sampled in several scenarios, and how many animals are likely to be repeat sampled overall, demonstrating the framing of the question is important for correct interpretation. An online tool is provided using the R Shiny library that researchers can use to help guide future experimental design for animal research that involves repeated sampling.

## Introduction

When sampling animal populations in research, theoretical questions arise in terms of sample size required to achieve adequate statistical power (Krzywinski and Altman, 2013). However, increased sample size comes with rapidly diminishing statistical returns, and sampling effort and reward must be balanced against the Three Rs of animal experimentation (Russell and Burch, 1992;

Tannenbaum and Bennett, 2015), under which auspices there is a need to *reduce* animal use, while maintaining statistically valid experiments, and to *refine* animal use to minimise suffering. The cost–benefit analysis of sample-size calculations (which should include benefits to animals more widely, not just the costs to the experimental population or sampled individuals) is frequently misinterpreted or ignored (Bacchetti et al., 2011).

Sampling may be performed with replacement – where after each sample the individual is replaced and may be sampled again – or without replacement – where the individual is retained or destructively sampled and cannot be sampled again. Sampling without replacement may be more difficult in well-mixed flocks such as farmed poultry or salmon, where individuals are not routinely marked, and opportunities for researchers to sample randomly or intentionally target a specific individual or class of individual (even if marked) may be limited (Marchewka et al., 2013).

From an ethical perspective, there may be circumstances in which recapture of individuals over repeated sampling is considered sub-optimal or may be restricted by animal experimentation protocols. For example, Hoskonen and Pirhonen (2006) found growth rate to be reduced in juvenile rainbow trout *Oncorhynchus mykiss* through repeat handling, although this effect was less for anaesthetised individuals. In Norway, the repeated use of animals and cumulative effect must be disclosed and severity considered

* Corresponding author.
  *E-mail address:* darren.green@stir.ac.uk (D.M. Green).

when applying for approval for experimental animal use (Landbruks- og matdepartementet, 2015), and repeated sampling of the same fish in aquaculture and related research is currently not permitted. The potential for repeated sampling has increased with the shift away from destructive sampling to satisfy *refinement*. There is also the concern of potential loss of statistical power by recapturing individuals previously studied, as the variance of samples taken with replacement (binomial) is larger than the equivalent variance for without-replacement sampling (hypergeometric; Wroughton and Cole, 2013), leading to larger standard errors of means.

To understand fully the ethical and practical concerns arising from repeat or non-lethal sampling, an important distinction must be made between the experience of the individual – essentially the likelihood of being selected again as a sample – and the event itself, the likelihood of a repeated sampling occurring in a population (Fig. 1). We can refer to the former as a recapture experience and the latter as a recapture event. From an ethical perspective, if a recapture experience, which might produce a bad welfare outcome for that individual, is very unlikely and the benefit to the population significant, then a cost–benefit ethical trade-off may be favourably balanced.

In this paper, we consider the mathematics of repeated sampling in terms of recaptures. The numbers that emerge from such calculations can be a little surprising, varying over many orders of magnitude depending upon whether the question asked relates to a recapture experience or a recapture event. Thus, the research question must be posed with care: Certain research questions are more "population focussed" and consider the chance that something happens globally (recapture event). These can be useful where any recapture of sample targets is critical and must be avoided, a precautionary principle is being applied, or where the welfare implications of repeated sampling are poorly understood, e.g. on the mucosa of fish (Kelly et al., 2024). Alternatively, we can focus more on the experiences of individuals within the population, which may be more important from an ethical perspective (recapture experience). Nevertheless, the shared experiences of many individuals across a large population can still be important, in the same way that a small improvement in health outcomes may be unnoticeable by a single person, but significant at the population level, e.g. in excess death statistics. A third perspective is that of an individual which has just been sampled: what has been their prior experience? These conditional probabilities differ from global or individual-level values.

Our code is made available to the general user in the form of a deployed R (R Core Team, 2022) Shiny (Chang et al., 2024) application that is instantly usable at https://pinkmongoose.shinyapps.io/ShinyRepeatSampling providing a tool that researchers can use while planning experimental trials.

## Methods

### Overall probability of any animal being sampled more than once

The potential of repeated sampling of the same individuals across a series of sampling events within the population (below, "flock") is akin to the well-known birthday problem (Borja and Haigh, 2007). This is typically stated as the probability that a group of $K$ people contains at least two people sharing a birthday (below, a "recapture" to keep the same terminology as above, assuming 365 days per year and all birthdays equally likely). The relatively high probability of a recapture amongst quite a small population can feel unintuitive, with only 23 people required for a greater than 50% chance of a shared birthday (recapture event). We can show that repeated sampling with replacement is a related problem: The birthday problem considers distributing $K$ random birthdays with a range 1–365 and counting recaptures amongst those random birthdays. This is the same mathematics as labelling our population with unique ID tags 1–$N$ (where $N$ is the flock size), sampling with replacement $K$ random individual IDs, and counting recaptures amongst those random IDs.

The exact answer to the classic birthday problem can be obtained by considering the fraction $1 - p_{\text{bday}}$ of possible permutations of samples which are favourable: i.e. without recaptures (Borja and Haigh, 2007), where $_xP_y$ denotes permutation: $_xP_y = x!/(x-y)!$:

$$p_{\text{bday}} = 1 - \frac{_NP_K}{N^K} \tag{1}$$

Next, we consider our focal problem: a repeated sampling study on an animal population, with repeated sampling of $k$ individuals over $n$ separate sampling events (below, "repeats"), where sampling within repeats is without replacement (the most likely situation), but replacement occurs between repeats (the most likely situation
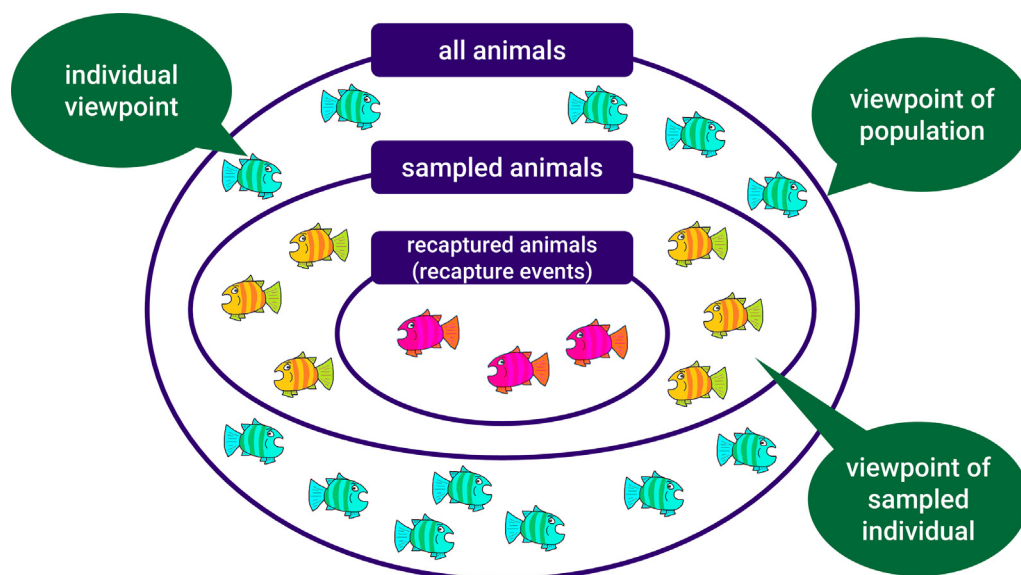


**Fig. 1.** Viewpoints for consideration of recapture probabilities for repeated sampling, which can be from the perspective of the entire population (recapture events), specific individuals (recapture experiences), or conditional on individuals having been sampled. We can also consider single sample repeats, or a whole series of sampling events.

in large flocks). We therefore have total sample attempts $K = nk$ but not necessarily $K$ unique individuals sampled. Plugging $K$ into Eq. (1) overestimates the probability of a recapture, as can be seen where $n = 1$, when no recapture should be possible. Instead, we need to reformulate the denominator of the equation to reflect the possible permutations of repeat samples:

$$p_{\mathrm{rep}} = 1 - \frac{{}_N P_{nk}}{({}_N P_k)^n} \tag{2}$$

A problem with these equations is that even for quite small flock size $N$, factorials such as $N!$ exceed representation on even 64-bit floating-point arithmetic. Instead, we can use the equivalent gamma function, taking advantage of the log-gamma function (Wolfram, 2024), which is widely found in office and statistical software, for example in Microsoft Excel as the GAMMALN.PRECISE function or R's lgamma() function. Products and quotients of large factorials become additions and subtractions on the log scale, and such mathematics is used routinely for built-in functions in software that handle probability distributions:

$$\ln \Gamma(x + 1) = \ln(x!)$$

We thus obtain, after replacing binomial coefficients with log-gamma functions and simplifying,

$$p_{\mathrm{rep}} = 1 - \exp((1-n)\ln\Gamma(N+1) + n\ln\Gamma(N-k+1) - \ln\Gamma(N-K+1)) \tag{3}$$

*Consecutive repeats*

We may also be interested in the easier problem of the probability that a recapture occurs in two consecutive repeats. This is simply Eq. (2) above, with $n = 2$, which simplifies to the hypergeometric distribution, corresponding to sampling without replacement. The notation used here, $\mathrm{Hyp}(x; a, b, k)$, denotes the probability of obtaining $x$ target items in a sample of $k$ items taken *without* replacement from a population containing $a$ target items and $b$ non−targets.

$$p_{\mathrm{seq2}} = 1 - \mathrm{Hyp}(0; k, N-k, k) \tag{4}$$

And from this, we can calculate the chance this occurs across any of the $n - 1$ pairs of adjacent repeats:

$$p_{\mathrm{seq}} = 1 - (\mathrm{Hyp}(0; k, N-k, k))^{n-1} \tag{5}$$

noting that when recaptures are rare,

$$p_{\mathrm{seq}} \approx (n-1)p_{\mathrm{seq2}}$$

due to $1 - x^n \approx n(1-x)$ when $x$ is close to one.

*Probability of a specific individual being sampled more than once*

The probability of a specific individual being sampled in any one repeat is simply given by $k/N$. Knowing this, we can assume that the number of times a specific individual is sampled across all repeats is binomially distributed where $B(x; k, q)$ denotes the probability of obtaining $x$ target items in a sample of $k$ items taken *with* replacement from a population containing a fraction $q$ target items. From this distribution, all terms other than zero or one repeat correspond to recapture experiences. Thus, the probability of a recapture experience is as follows:

$$p_{\mathrm{ind}} = 1 - B(0; n, k/N) - B(1; n, k/N) \tag{6}$$

Below, we discuss viewpoints such as the relative recapture rate for animals already sampled. This leads us to consider conditional probability: Taking the perspective of individuals known to have been sampled, the conditional probability of at least one recapture for a sampled individual is as follows:

$$p_{\mathrm{cond}} = \frac{1 - B(0; n, k/N) - B(1; n, k/N)}{1 - B(0; n, k/N)} \tag{7}$$

Again, we can ask some related simpler questions, such as the probability of a specific individual being sampled again, conditional on it having been sampled in the previous repeat, which is simply

$$p_{\mathrm{cond2}} = k/N \tag{8}$$

where sampling events are independent.

*Expected number of unique individuals sampled amongst all sample repeats*

Finally, we consider the ratio between total sample attempts $K$ and the estimated number of unique individuals covered by sampling across all repeats $U$, limited to the range $k \le U \le K$. We address this by iteratively considering the distribution of the number of unique sampled individuals $P_u$ at repeat $t + 1$, from the convolution of the distribution at repeat $t$ and the hypergeometric distribution for the number of new unique individuals sampled, starting from an unsampled population at repeat zero through to time $t = n$.

$$P_{u,t+1} = \sum_{i=0}^{k} P_{u-i,t} \cdot \mathrm{Hyp}(i; N-u+i, u-i, k)$$

(Impossible combinations in the equation above implicitly have zero probability, but some statistics functions may return error codes such as *NA*.) From this, we calculate the expected number of unique individuals sampled, most usefully expressed a proportion of total sample size, $q$:

$$U_t = \sum_{u=k}^{K} u \cdot P_{u,t}$$

$$q_t = \frac{U_t}{K} \tag{9}$$

*App deployment*

The mathematical formulae introduced above are implemented in the R Shiny platform for deployment, at https://pinkmongoose.shinyapps.io/ShinyRepeatSampling. The design is responsive and outputs are recalculated as the user changes the inputs. The app was used to generate the results for the scenarios explored in the results section below, and can be applied by other users to aid their own experimental design. There are just three parameters to be specified by the user: the total or effective population size of the whole population from which samples are drawn ($N$), the number of repeated sampling events ($n$), and the number of units sampled from the population, without replacement, at each repeated sampling event ($k$). This app does not determine a sampling regime or experimental design, but instead may be used alongside such calculators and tools such as the NC3Rs Experimental Design Assistant (du Sert

et al., 2017), when assessing what a particular design would mean in terms of the potential for recaptures.

## Results

We consider several scenarios to contextualise the mathematics above.

Scenario 1: **Salmon**: 100 000 Atlantic salmon *Salmo salar* in a sea pen are sampled by net each month over 14 months of a production cycle, with a sample size of 20. The salmon may or may not be marked upon sampling, e.g. by Panjet tattoo (Dietrich and Cunjak, 2006), but selecting individuals is not possible here. In this scenario, the potential concern is how often recapture might occur, but populations are large and recapture is expected to be rare.

Scenario 2: **Poultry**: 5 000 broilers in a unit are sampled each week for a 6-week production cycle, with a sample size of 30. As above, the broilers may or may not be marked upon sampling, but selecting individuals is assumed not possible. Here, the question is similar to **Salmon** above, but recaptures are likely to be more frequent and the frequency of sampling is also more intense.

Scenario 3: **Sheep FEC**: A flock of 250 sheep is sampled by 10 3-week faecal-egg-count (**FEC**) samples, using collection of fresh faeces from the field, with a sample size of 15, where this sample size is suggested by the Moredun Research Institute (Moredun, 2024). We assume the faeces are sufficiently fresh that sampling is effectively without replacement (there may be some aspect of replacement sampling), but we do not know the individual that produced each faecal sample. Here, there is no particular ethical issue, but a concern might be the degree to which recapture reduces our effective sample size of independent sheep.

Scenario 4: **Sheep saliva**: Saliva from a flock of 50 sheep is randomly sampled as part of a study on stress, on three occasions, with a sample size of 15. Sampling can be performed without replacement, and in a small flock could be individually focussed, but random sheep are assumed here. Nevertheless, the sheep are ear-tagged so individual IDs are known. Here, a concern might be that recaptures could affect the stress state of the animals, affecting the experimental conditions themselves, and therefore, the degree of recapture is of interest.

The probabilities of recaptures taken across the three viewpoints for each scenario are shown in Table 1. As expected, given the general birthday problem results, the probability of a global recapture tends towards unity quite quickly with decreasing population size, and recaptures are inevitable in both **Sheep FEC** and **Sheep saliva** scenarios, and nearly so for the **Poultry** scenario. Even for the **Salmon** scenario, they are sufficiently likely to occur

as to raise attention if any recapture were disallowed or the precautionary principle applied.

As the number of repeats increases, the options for global recaptures rise faster than the options for adjacent recaptures, and this is partly reflected in the greater difference between (3) and (5) for scenarios with higher $n$ such as **Salmon**. Sampling is sufficiently infrequent in the **Salmon** scenario that global adjacent sampling-period recaptures rather than global recaptures may be of interest to the researchers; however, protocol may also depend on any additional disturbance experienced by either husbandry or research purposes, such as grading or initial stocking. For the **Poultry** scenario, the relatively frequent global adjacent recaptures may still be of concern, depending on the severity of the procedures applied.

For the individual recapture probabilities, the larger difference compared with global recapture can be seen clearly where the populations are large, such as the **Salmon** and **Poultry** scenarios. This shows that at the individual-animal level, the probability of a recapture is vanishingly small, whereas it is noticeable in the smaller population of the **Sheep saliva** scenario. Nevertheless, these probabilities do not consider the cumulative impact across the whole population, which is where the conditional recapture calculations are helpful.

The conditional recapture (single) viewpoint (7) provides magnitudes of probabilities intermediate between the global recapture and individual recapture values, and is possibly the most easily interpretable statistic, answering the question of the likelihood of the recapture experience. As with the global probabilities, restricting to adjacent repeats is more conservative and provides lower probabilities (Table 1) particularly where recaptures are rare and the numbers of repeats are higher. For the **Salmon** scenario, recapture from the perspective of a captured fish is still rare, which might be considered during experimental design where the alternative is destructive sampling. For the **Sheep saliva** scenario, recapture is still seen to be a frequent experience for the animals.

Finally, Table 1 provides the proportion of uniquely sampled individuals $q$ (9). This is useful in terms of diagnosing if our sampling schedule is inefficient at capturing different individuals (values substantially lower than one), but less easily interpreted as a diagnostic for the proportion of recaptures amongst sampled individuals. We can see that this is close to one for the **Salmon** and **Poultry** scenarios, but noticeably less efficient for both sheep scenarios. We can consider here how repeated sampling effectively creates pseudoreplicates rather than true replicates where population statistics are the goal: In both cases, the coverage of unique individuals is only about three−quarters of the sample size.

## Discussion

Though we derive exact calculations for the probabilities above, biological factors not accounted for in these equations may have a

**Table 1**

Recapture probabilities for repeated sampling where sampling is without replacement within repeats, and with replacement between repeats. Equation references are given in parentheses. Four scenarios are considered as described in the main text: Salmon, Poultry, Sheep FEC (faecal egg count) and Sheep.

| Item | Salmon | Poultry | Sheep FEC | Sheep saliva |
|---|---|---|---|---|
| Sample size $k$ | 20 | 30 | 15 | 15 |
| Sample repeats $n$ | 14 | 6 | 10 | 3 |
| Population size $N$ | 100 000 | 5 000 | 250 | 50 |
| Total sample attempts $K = nk$ | 280 | 180 | 150 | 45 |
| Global recapture $p_{rep}$ (3) | 0.305 | 0.935 | 1.00 | 1.00 |
| Global adjacent recapture $p_{seq2}$ (single)(4) | 0.00399 | 0.166 | 0.616 | 0.999 |
| — all repeats $p_{seq}$ (5) | 0.0507 | 0.596 | 0.9998 | 1.000 |
| Individual recapture $p_{ind}$ (6) | $3.63 \times 10^{-6}$ | $531 \times 10^{-6}$ | 0.118 | 0.216 |
| Conditional recapture (single $p_{cond}$) (7) | 0.0013 | 0.0150 | 0.255 | 0.329 |
| Conditional adjacent recapture (single $p_{cond2}$) (8) | $200 \times 10^{-6}$ | 0.0060 | 0.060 | 0.300 |
| Proportion of uniquely sampled individuals $q$ (9) | 0.9987 | 0.985 | 0.769 | 0.730 |

strong impact in practice. First, individuals within the flock might have unequal likelihood of being sampled, for example due to their individual personalities or states of health. Where heterogeneity correlates capture rate to the target variables of the study (e.g. health status), this is especially problematic for experimental design and may lead to bias (de Blas et al., 2020). Clustering likely increases the frequency of recaptures compared with the equations presented above, heightening ethical concerns: Most forms of heterogeneity will likely increase the probability of recaptures, also demonstrated in generalised forms of the standard birthday problem (Munford, 1977).

Clustering might occur in the **Salmon** scenario, depending upon the sampling procedure used: Net sampling is likely to favour the selection of some individuals over others. Such effects on sample biases have been previously noted in aquaculture (Nilsson and Folkedal, 2019); however, alleviating such bias can be challenging. It may be sufficient for precautionary reasons to use a coverage-adjusted population size, e.g. by assuming only a fraction of the population can be captured (Aebischer, 1986) and substitute this value for $N$. Parallel concerns occur in wildlife sampling, and multiple authors have considered models for estimating density from mark–recapture data (e.g. McClintock, 2015), and clustering could occur in the Sheep FEC scenario if the sheep do not evenly use their living space.

Under some circumstances, the sampling process itself may alter the probability of recapture on subsequent samplings, in either direction. For example, sampled individuals may become "trap shy" and be less likely to be sampled later, or alternatively may become "trap happy" (Ghosh and Norris, 2005), increasing the probability of a recapture. Such statistical behaviour may also result from physiological rather than behavioural factors. Where this is not known in advance, sampling studies themselves may provide some answers to these questions.

In this paper, we have considered a range of scenarios and points of view involving potentially small probabilities, combined with large numbers. As with any analysis of risk, it is worth considering the statistical fallacies and cognitive biases that may present when farmers or researchers are presented with such information, that might lead to misinterpretation of outputs of experimental design software or decision-making that does not fully address all concerns. *Scope neglect* occurs where a risk is evaluated without due consideration of both the frequency and the severity of a hazard, in particular by failing to appreciate the cumulative impact of small effects at the individual level across very large numbers of individuals. For example, a small inconvenience to a single sampled animal may equate to an observable problem when considered at population level (Dickert et al., 2015). Without a mathematical treatment, potentially through a risk-analysis framework, e.g. impact = frequency × effect, humans may struggle to cope with considering risks to "thousands", "millions", or "billions" of individuals, leading to flawed decisions. This is more difficult for studies involving large population sizes such as our **Salmon** or **Poultry** scenarios.

*Base-rate neglect* and the *prosecutor fallacy* are caused by the failure to appreciate the background rate of events when considering a particular event (Bar-Hillel, 1980). Above, we contrast the frame of reference of animals already subject to a prior sampling event (recapture experiences) with those of the general population (recapture events). Misusing these frames can lead to some incorrect interpretations, e.g. where the chance of a recapture event is reported as one in a hundred, where one in ten is the corresponding rate of a recapture experience. Again, this is particularly noteworthy for the **Salmon** and **Poultry** scenarios with their large population sizes: Focusing on **Salmon**, the probability of a recapture event (0.305) is several orders of magnitude higher than the

probability of a capture experience ($3.63 \times 10^{-6}$), with the conditional probability of recapture intermediate (0.0013).

Finally, we must avoid the cognitive biases of *framing* by focusing only on the negatives of sampling where these are perceived to be present (Tversky and Kahneman, 1981), and consider both the costs and the benefits to sampling a population, i.e., to balance the potentially small impact of recaptures amongst the sampled individuals against the wider ethical benefits posed to the whole population in monitoring population health and welfare. After all, without an assumption of benefit, we would not be sampling at all, but different stakeholders in animal experimentation research may have varied perceptions of these benefits (Cabaret and Fortin, 2022).

To conclude, some simple mathematics can help predict when researchers are likely to recapture the same individuals while sampling a population, which can assist them in planning trials and considering ethics. However, such models need to be interpreted with care, with the right question asked prior to experimental design selection. Beyond mathematics, biological factors can affect recapture probability, both through innate physiological and behavioural differences amongst the sampled individuals, and the effects of sampling itself.

## Ethics approval

This study was approved under the University of Stirling General University Ethical Panel (GUEP) review reference EC 2024 18301 13155.

## Data and model availability statement

The models constructed in this project are publicly available for use and accessible from ShinyApps.IO via https://pinkmongoose. shinyapps.io/ShinyRepeatSampling. Information can be made available from the authors upon request.

## Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the author(s) did not use any AI and AI-assisted technologies.

## Author ORCIDs

**Darren Green:** https://orcid.org/0000-0001-9026-5675.

## CRediT authorship contribution statement

**D.M. Green:** Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Formal analysis, Conceptualization. **C.G. Mitchell:** Writing – review & editing, Writing – original draft, Conceptualization.

## Declaration of interest

None.

## Acknowledgements

## Financial support statement

## References

Aebischer, N.J., 1986. Estimating the proportion of uncatchable animals in a population by double-sampling. Biometrics 42, 973–979. https://doi.org/10.2307/2530713.

Bacchetti, P., Deeks, S.G., McCune, J.M., 2011. Breaking free of sample size dogma to perform innovative translational research. Science Translational Medicine 3, 87ps24. https://doi.org/10.1126/scitranslmed.3001628.

Bar-Hillel, M., 1980. The base-rate fallacy in probability judgments. Acta Psychologica 44, 211–233. https://doi.org/10.1016/0001-6918(80)90046-3.

Borja, M.C., Haigh, J., 2007. The Birthday Problem. Significance 4, 124–127. https://doi.org/10.1111/j.1740-9713.2007.00246.x.

Cabaret, J., Fortin, L., 2022. Varying degrees of animal reification by stakeholders in experimental research. Animals 12, 190. https://doi.org/10.3390/ani12020190.

Chang W., Cheng, J., Allaire, J., Sievert, C., Schloerke, B., Xie, Y., Allen, J., McPherson, J., Dipert, A., & Borges, B. (2024). *Shiny: Web Application Framework for R*. R package version 1.8.0.9000. Available at: https://github.com/rstudio/shiny https://shiny.posit.co/ Accessed 2024-02-10.

de Blas, I., Muniesa, A., Vallejo, A., Ruiz-Zarzuela, I., 2020. Assessment of sample size calculations used in aquaculture by simulation techniques. Frontiers in Veterinary Science 7, 253. https://doi.org/10.3389/fvets.2020.00253.

Dickert, S., Västfjäll, D., Kleber, J., Slovic, P., 2015. Scope insensitivity: the limits of intuitive valuation of human lives in public policy. Journal of Applied Research in Memory and Cognition 4, 248–255. https://doi.org/10.1016/j.jarmac.2014.09.002.

Dietrich, J., Cunjak, R.A., 2006. Evaluation of the impacts of Carlin Tags, fin clips, and Panjet tattoos on juvenile Atlantic salmon. North American Journal of Fisheries Management 26, 163–169. https://doi.org/10.1577/M05-032.1.

du Sert, P.N., Bamsey, I., Bate, S.T., Berdoy, M., Clark, R.A., Cuthill, I., Fry, D., Karp, N.A., Macleod, M., Moon, L., Stanford, S.C., Lings, B., 2017. The experimental design assistant. PLoS Biology 15, e2003779. https://doi.org/10.1371/journal.pbio.2003779.

Ghosh, S.K., Norris, J.L., 2005. Bayesian capture-recapture analysis and model selection allowing for heterogeneity and behavioral effects. Journal of Agricultural, Biological, and Environmental Statistics 10, 35–49. https://doi.org/10.1198/108571105X28651.

Hoskonen, P., Pirhonen, J., 2006. Effects of repeated handling, with or without anaesthesia, on feed intake and growth in juvenile rainbow trout, *Oncorhynchus mykiss* (Walbaum). Aquaculture Research 37, 409–415. https://doi.org/10.1111/j.1365-2109.2005.01448.x.

Kelly, L.A., Yost, C.K., Cooke, S.J., 2024. Opportunities and challenges with transitioning to non-lethal sampling of wild fish for microbiome research. Journal of Fish Biology 104, 912–919. https://doi.org/10.1111/jfb.15650.

Krzywinski, M., Altman, N., 2013. Power and sample size. Nature Methods 10, 1139–1140. https://doi.org/10.1038/nmeth.2738.

Landbruks- og matdepartementet (2015). Forskrift om bruk av dyr i forsøk. Available at: https://lovdata.no/dokument/SF/forskrift/2015-06-18-761 Accessed 2024-08-08.

Marchewka, J., Watanabe, T.T.N., Ferrante, V., Estevez, I., 2013. Welfare assessment in broiler farms: transect walks versus individual scoring. Poultry Science 92, 2588–2599. https://doi.org/10.3382/ps.2013-03229.

McClintock, B.T., 2015. Multimark: an R package for analysis of capture–recapture data consisting of multiple "noninvasive" marks. Ecology and Evolution 5, 4920–4931. https://doi.org/10.1002/ece3.1676.

Moredun (2024). *FEC Check*. Available at: https://app.moredun.org.uk/fec/check Accessed 2024-02-12.

Munford, A.G., 1977. A note on the uniformity assumption in the birthday problem. The American Statistician 31, 119.

Nilsson, J., Folkedal, O., 2019. Sampling of Atlantic salmon *Salmo salar* from tanks and sea cages is size-biased. Aquaculture 502, 272–279. https://doi.org/10.1016/j.aquaculture.2018.12.053.

R Core Team (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Available at: https://www.R-project.org/ Accessed 2024-02-10.

Russell, W.M.S., & Burch, R.L. (1992) *The Principles of Humane Experimental Technique* (Special Edition). UFAW, Wheathampstead, UK. ISBN: 0-900767-78-2. Available at: https://caat.jhsph.edu/the-principles-of-humane-experimental-technique/ Accessed 2024-02-15.

Tannenbaum, J., Bennett, B.T., 2015. Russell and Burch's 3Rs then and now: the need for clarity in definition and purpose. Journal of the American Association of Laboratory Animal Science 54, 120–132.

Tversky, A., Kahneman, D., 1981. The framing of decisions and the psychology of choice. Science 211, 453–458. https://doi.org/10.1126/science.7455683.

Wolfram (2024). *Log Gamma Function*. Available at: https://mathworld.wolfram.com/LogGammaFunction.html Accessed 2024-02-10.

Wroughton, J., Cole, T., 2013. Distinguishing between binomial, hypergeometric and negative binomial distributions. Journal of Statistics Education 21, 1. https://doi.org/10.1080/10691898.2013.11889663.