Trials

METHODOLOGY

Open Access

Remote, bivariate prior elicitation for a Bayesian non-inferiority randomized controlled trial

Arlene Jiang[1*], Alex Aregbesola[2,3], Apoorva Gangwani[2], Terry P. Klassen[4], Amy C. Plint[5,6], Elisabete Doyle[3], William Craig[7], Mohamed Eltorki[8], Banke Oketola[2], Hoda Badran[2,3], Yongdong Ouyang[9†] and Anna Heath[1,10,11†]

Abstract

Background Prior distributions must be specified for the parameters of interest in a Bayesian clinical trial. When existing evidence on the effects of the trial interventions is limited or inconclusive, prior distributions can be constructed with expert elicitation. However, conventional elicitation requires face-to-face interactions and intensive pre-elicitation training, which can be infeasible and costly. Our remote elicitation was based on an established expert elicitation methodology, and we incorporated bivariate prior distributions to introduce dependencies between the elicited probabilities. We aimed to elicit a prior distribution for the Croup Dosing Trial, which assesses the efficacy of two separate doses of dexamethasone on the number of return visits to the emergency department within 7 days in children with croup. This trial evaluates the non-inferiority of 0.15 mg/kg of dexamethasone, compared to the standard dose of 0.60 mg/kg to treat croup.

Methods We conducted three remote workshops to elicit expert beliefs on the efficacy of the two doses of dexamethasone. Each workshop consisted of two survey rounds, separated by a group discussion. Prior to the workshop, experts reviewed the same current literature that was provided on the effects of the two doses of dexamethasone. Beliefs were aggregated using expert-specific bivariate distributions with latent effects. The aggregated distribution, along with the surveyed non-inferiority margin, determined the sample size for the Bayesian non-inferiority trial design.

Results Twelve emergency medicine physicians participated in our remote elicitation exercise. The elicitation generated a prior distribution centered at 6% for the 0.60 mg/kg dose and 8% for the 0.15 mg/kg dose. The aggregated prior distribution produced a sample size of 1850, based on a non-inferiority margin of 4%.

Conclusions We elicited a prior distribution that incorporated past evidence and expert opinion. The elicited prior is consistent with previous literature on the efficacy of the dexamethasone doses in treating croup. Our approach demonstrates the feasibility of remotely eliciting bivariate distributions to design clinical trials.

Trial registration NCT06272383 (Registered May 8, 2024).

[†]Yongdong Ouyang and Anna Heath contributed equally to this work.

*Correspondence:
Arlene Jiang
arlene.jiang@sickkids.ca
Full list of author information is available at the end of the article

Keywords Expert elicitation, Bayesian statistics, Randomized controlled trials, Trial design, Non-inferiority trial, Sample size determination, Prior probability distribution

Background

Clinical trials are increasingly designed and evaluated using Bayesian methods due to their interpretability and increased efficiency [1]. Bayesian methods require the specification of prior distributions for the parameters of interest that reflect the uncertainty of decision makers [2]. Existing evidence from previous trials is often incorporated into prior distributions, which are then combined with study data to generate trial conclusions. However, this approach can be limited when similar trials have not been previously conducted.

Alternatively, a prior distribution based on expert opinion can be formally obtained through expert elicitation [3]. Elicitation is the scientific process of converting content experts' opinions into statistical distributions to quantify the uncertainty on the parameters of interest [4]. Elicitation is often used to support decision-making in areas with limited or inconclusive evidence [5]. In the context of clinical trials, elicitation allows stakeholders to directly incorporate their uncertainty about the effects of the study interventions while planning the trial [6]. Over the past few decades, a growing number of frameworks and applications have been developed for the process of elicitation [5, 7–11].

Croup is a common respiratory disease in children, associated with frequent emergency department (ED) visits [12]. Usual symptoms of croup include a seal-like barking cough and a whistling noise while breathing [12]. Past studies have demonstrated that glucocorticoids are effective in managing symptoms and reducing healthcare costs (i.e., hospitalization, admission to intensive care unit, return visits) [13–15]. However, the course of glucocorticoids prescribed for croup is associated with adverse events such as gastrointestinal bleeding, pneumonia, and sepsis [16]. The Croup Dosing Trial will assess the efficacy of a lower dose of glucocorticoids for the treatment of croup in children. The trial will test the non-inferiority of a lower dose of dexamethasone compared to the standard dose on the number of revisits to the ED within 7 days. This trial aims to provide definitive evidence for how croup should be treated in children.

The Croup Dosing Trial will take place in Canada, the USA, New Zealand, and Australia. Therefore, a remote elicitation involving experts from different countries would generate a more representative prior for the trial population. The goal of the elicitation is to produce a prior distribution that would be used for Bayesian sample size determination and the analysis of the trial data. This study describes the design and conduct of the elicitation to derive a prior distribution for the Croup Dosing Trial. Traditional elicitation requires intensive pre-elicitation training and face-to-face sessions [17]. This approach may be infeasible given the time, geographical, and cost restrictions in assembling a group of experts. Lan et al. successfully implemented a remote expert elicitation for a clinical trial in bronchiolitis in infants termed the BIPED trial [18], which addressed these limitations by holding multiple remote elicitation workshops. Conducting multiple workshops accommodated different time zones and schedules, which encouraged experts from Canada, the USA, Australia, and New Zealand to participate. This added to the heterogeneity of expert opinions, which is practically and computationally advantageous [19]. This approach is based on the validated IDEA (Investigate Discuss Estimate Aggregate) protocol [5], which we plan to leverage in the Croup Dosing Trial expert elicitation. The BIPED study assumed independence between the probability of efficacy in each treatment arm. However, experts likely used their estimate of one treatment arm to inform the other, also known as anchoring [20, 21]. One approach that avoids anchoring is to elicit two distinct quantities, such as the response in the control arm and the treatment effect [22, 23]. These quantities are considered independent, and thus these estimates are aggregated separately. However, a treatment effect such as an odds ratio is not directly observable and is therefore more difficult to elicit than the response in either study arm [7]. Because of this, we elicited the probability of the outcome in our study arms, addressing the dependence between these quantities with a bivariate prior distribution for each expert. Thall et al. implemented individual expert latent effects to introduce a correlation between the expert-specific prior distributions in each treatment arm [24]. This approach was used to produce a prior distribution for a trial studying pediatric idiopathic nephrotic syndrome. We planned to incorporate this bivariate approach to account for possible dependencies between treatment arm estimates and conduct remote expert elicitation for the Croup Dosing Trial.

Methods

Croup Dosing Trial

The Croup Dosing Trial is a phase III, double-blind, multicenter, randomized controlled non-inferiority trial comparing two treatment regimens for children presenting to the ED with croup within the Pediatric Emergency

Research Canada (PERC). Children between 6 months and 5 years with a clinical diagnosis of croup will be eligible. The trial will determine whether a 0.15 mg/kg dose of dexamethasone is non-inferior to the standard dose of 0.60 mg/kg. The primary outcome of this trial is the number of return visits within 7 days of initial presentation to the ED. Patients will be allocated in a ratio of 1:1 to receive either dose of dexamethasone.

Research ethics approval

The vanguard and elicitation components of the Croup Dosing Trial have received ethics approval of Health Canada, University of Manitoba Health Research Ethics Board, and Shared Health on behalf of the Children's Hospital of Winnipeg. Use of the elicitation data was approved by the ethics committee at the Hospital for Sick Children.

Designing the elicitation

Key parameters

In our elicitation, the key parameters were the probability of return visits to the ED within 7 days after initial presentation, following treatment with either 0.60 mg/kg or 0.15 mg/kg of dexamethasone. We denote these two probabilities as p1 and p2, respectively. For each expert, we assume that these probabilities follow beta distributions, a common distributional choice for modeling probabilities [25].

Key parameters were elicited using the workshop methodology described in Lan et al. [18] We provided experts with a case study of a child with croup who met inclusion/exclusion criteria for the Croup Dosing Trial (Supplementary File 1). The case study describes a 1-year-old boy with symptoms indicative of a mild to moderate croup diagnosis. Experts were asked to determine the number of patients, out of 100, with characteristics like this patient who would revisit the ED for the two doses of dexamethasone.

In addition to the elicited probability estimates, we considered the effect of each expert's professional experience on their judgement of p1 and p2. Therefore, we surveyed the professional experience of the experts (Supplementary File 2) and incorporated this information into aggregating expert opinions for the final prior distribution [24].

Online elicitation tool

To support our elicitation, we created an online, interactive tool using R and the shiny package (https://apoorva-gangwani.shinyapps.io/apoorva_elicitation/) [26, 27]. This tool allowed experts to visualize and send in their elicitation estimates. Experts were prompted to provide the lower and upper plausible values, followed by their most plausible value that subjectively described their uncertainty about the probability of ED revisits within 7 days of initial presentation to the ED at both doses of dexamethasone. The lower and upper plausible values were assumed to represent the 95% central credible interval of a beta distribution, while the most plausible value was assumed to represent the mode. These values were fit to beta distributions using betaExpert [28]. After selecting elicitation values, our tool provided experts with a summary plot of the corresponding beta distribution. This real-time visualization allowed experts to adjust the fitted beta distribution until they deemed their beliefs to be sufficiently elicited. Finally, experts were prompted to send their elicitation values on a submission page hosted using REDCap [29, 30].

Selecting the experts

We identified experts from Canada, the USA, Australia, and New Zealand, to determine a prior representative of the regions where the Croup Dosing Trial will be conducted. Experts were primarily recruited from PERC, a national network, and the Pediatric Emergency Research Network (PERN) [31], an international organization. Participants could be selected if they (i) were identified as experts in croup and its treatment and (ii) had experience in pediatric emergency medicine. Participants were ineligible if they had (i) prior involvement with the Croup Dosing Trial or (ii) authored any studies that preferred dexamethasone at 0.60 mg/kg or 0.15 mg/kg. We aimed to recruit 10 to 15 experts for a diversity in geographical representation and experience [5, 32].

Aggregating the prior distribution

To derive the final prior distribution, we aggregated individual distributions with methods developed by Thall et al. [24] These methods construct a bivariate prior for each expert, introducing a correlation between responses for the two doses. We first derived marginal distributions for each expert's estimate for p1 and p2. Covariates related to physician's expertise were added during the fitting process to explain the variability between experts' responses. An expert-specific correlation between two marginal distributions was constructed using latent expert effects. For this regression, we used noninformative priors for the covariate effects (N (0, 10)), the intercepts (N (0, 100)), the standard deviation of the latent effects (IG (1, 1)), and the correlation between the latent effects (Unif (− 1, 1)). Each physician's bivariate prior was weighted evenly in the final aggregate.

The elicitation

Pre-workshop materials

Before the workshop, experts were sent a set of readings to prepare for the workshop [33, 34]. These readings included one clinical trial and one meta-analysis, presenting past and latest evidence on the efficacy of 0.60 mg/kg and 0.15 mg/kg of dexamethasone in the treatment of croup in children. These studies were sent to complement the experts' existing knowledge on this topic.

Remote, real-time expert elicitation workshop

We conducted a 90-min workshop, standardized with a script (Supplementary File 3) [35], over Zoom in April 2024. The same workshop was held a total of three times to accommodate different time zones and the schedules of working physicians. Three to four statistical or medical experts facilitated each workshop (AJ, AG, AA, AH). The workshop began with an introduction to the context of the trial and Bayesian statistics. We then demonstrated the elicitation tool with an example dummy question (Supplementary File 3).

The workshop format was adapted from the IDEA protocol 5. The IDEA protocol outlines four stages to an elicitation: (1) participants individually form estimates on the quantities of interest, (2) participants discuss their reasoning, (3) participants individually re-estimate the quantities of interest, and (4) individual estimates are aggregated. Accordingly, each of our elicitation workshops included two rounds of questions, separated by a group discussion. During each round, experts individually submitted their prior distributions on the probability of revisiting the ED in each treatment arm. During the group discussion, experts were presented with deidentified boxplots containing their answers from round 1 (Supplementary File 4). Experts were given 15 min to discuss why their answers were different. After reflecting on the discussion at workshop 1, the facilitating team prepared a set of questions to guide the discussion at the other two workshops (Supplementary File 3). The group discussion provided experts with an opportunity to reconsider their initial beliefs.

Post-workshop

Post-workshop, experts were sent their individual, workshop-aggregate, and aggregated distributions for the probability of ED revisits in each treatment arm. Experts were invited to comment on these findings over email.

Bayesian sample size determination

Non-inferiority margin

The non-inferiority margin was estimated by surveying PERC physicians at the 11 prospective sites for the Croup Dosing Trial in December 2025 (Supplementary File 5). The median of the survey answers was taken as the margin.

Sample size determination

Our sample size was determined using assurance, which required simulation in R using INLA [26, 36]. We generated binomial data, representing the number of return visits in each arm, with probabilities given by each pair of samples in the aggregated p1 and p2 distributions. These simulated data were fit to general linear models with a prior distribution for the difference between p2 and p1. We declared non-inferiority when 95% of the posterior mean differences were smaller than the non-inferiority margin. To evaluate the null scenario (analogous to type I error), we repeated this process with binomial data generated with probabilities from p1 and with probabilities from p1 plus the non-inferiority margin. We chose a sample size that provided 80% of the maximum possible assurance, i.e., relative assurance, and 5% assurance in the null scenario. Simulations were repeated 1000 times.

Results

Elicitation workshop

Baseline characteristics

We invited a total of 33 experts from Australia ($n=5$), Canada ($n=17$), New Zealand ($n=4$), and the USA ($n=7$). Experts declined participation because they were on vacation, had scheduling conflicts, or had large time differences. We recruited 12 clinicians from Canada ($n=8$) and the USA ($n=4$) to participate in our workshops. Table 1 displays the characteristics of these experts from our survey of their professional experience. Experts had 11 to 38 years of experience in practicing medicine. In the past year, all physicians had prescribed 0.60 mg/kg of dexamethasone, while only two had prescribed 0.15 mg/kg of dexamethasone. Most clinicians had received training in clinical trial methodology and statistics.

Prior distributions

Figure 1 and Table 2 show the individual's prior beliefs on the efficacy of each dose of dexamethasone before and after discussion. Individual responses varied and tended to decrease in variability from round 1 to round 2. Furthermore, many experts modified their best estimate from round to round. In general, experts believed that there would be a higher probability of revisiting the ED in children treated with 0.15 mg/kg of dexamethasone. Figure 2 shows the aggregated prior distributions from all experts for each round of questions. Similar to the previous figure, these graphs indicate the decrease in variability of the answers from round 1 to round 2. The aggregated distributions from round 2 were taken to be the final prior distributions used for sample size

Table 1 Baseline characteristics of the expert elicitation participants

Country of practice	**Canada**	**8**
	USA	**4**
How many years have you practiced medicine?	11–15	3
	16–20	2
	21–25	3
	26–30	1
	31–35	2
	36–40	1
Have you prescribed 0.60 mg/kg of dexamethasone for croup over the last year?	Yes	12
	No	0
Have you prescribed 0.15 mg/kg of dexamethasone for croup over the last year?	Yes	2
	No	10
What's the maximum dose (mg) of dexamethasone that you have prescribed for croup in the last year?	6	1
	10	5
	12	2
	15	1
	16	3
Have you received any training in clinical trial methodology?	Yes	10
	No	2
Have you received any training in statistics?	Yes	11
	No	1

determination (Table 3). Figure 3 presents the risk difference between the probability of revisits under 0.15 mg/kg versus 0.60 mg/kg of dexamethasone from the joint distribution. The narrow central tendency demonstrates the correlation between p1 and p2. The posterior estimates have a correlation of 0.40, indicating a moderate positive association between the two elicited quantities. Experts did not provide additional comments over email.

Bayesian sample size determination

Our survey of 60 physicians found a non-inferiority margin of 4%. To determine the trial's sample size, binomial data was generated with probabilities of 6.44% (p1) and 8.19 (p2) (Table 3). We approximated the prior distribution for the difference between p2 and p1 with a Pearson IV distribution (m = 1.91, nu = −0.36, location = 0.01, scale = 0.08) to best capture the shape of the distribution presented in Fig. 3, Table 3). A sample size of 1850 provided 80% relative assurance and 5% assurance in the null scenario.

Discussion

In this study, we implemented a remote expert elicitation exercise to develop a bivariate prior distribution for the Croup Dosing Trial. This prior distribution was used to determine the sample size and will be used to analyze the trial. Our elicited prior had median ED revisit probabilities of 6.4% for the 0.60 mg/kg dose and 8.2% for the 0.15 mg/kg dose. These probabilities are consistent with Parker and Cooper [34], who found that 5.9% and 8.8% of patients revisited the ED when treated with 0.60 mg/kg and 0.15 mg/kg of dexamethasone, respectively. Our elicited probabilities are not directly comparable to the meta-analysis [33], which accounted for return visits to both the ED and family doctors. Consistent with the meta-analysis, however, which reported return visit rates of 19% and 21% for 0.60 mg/kg and 0.15 mg/kg, we found a comparable risk difference between treatment arms. The similarities between our elicited probabilities and past evidence demonstrate the reliability of our prior distribution.

During our elicitation, we found a shift in expert beliefs from round to round. Between survey rounds, the group discussions provided experts with an opportunity to comment on how they chose their estimates. Experts re-evaluated their interpretation of the lower and upper plausible values, as well as how the primary outcome relates to the clinical progression of croup. Physicians commented on factors that would impact the efficacy of the doses of dexamethasone such as time of year and virulence of the virus. Another consideration was how regional differences in access to primary care would affect the number of children who would visit their family doctor rather than return to the ED. Experts subsequently provided less varied answers during the second survey round. The topics discussed were consistent with

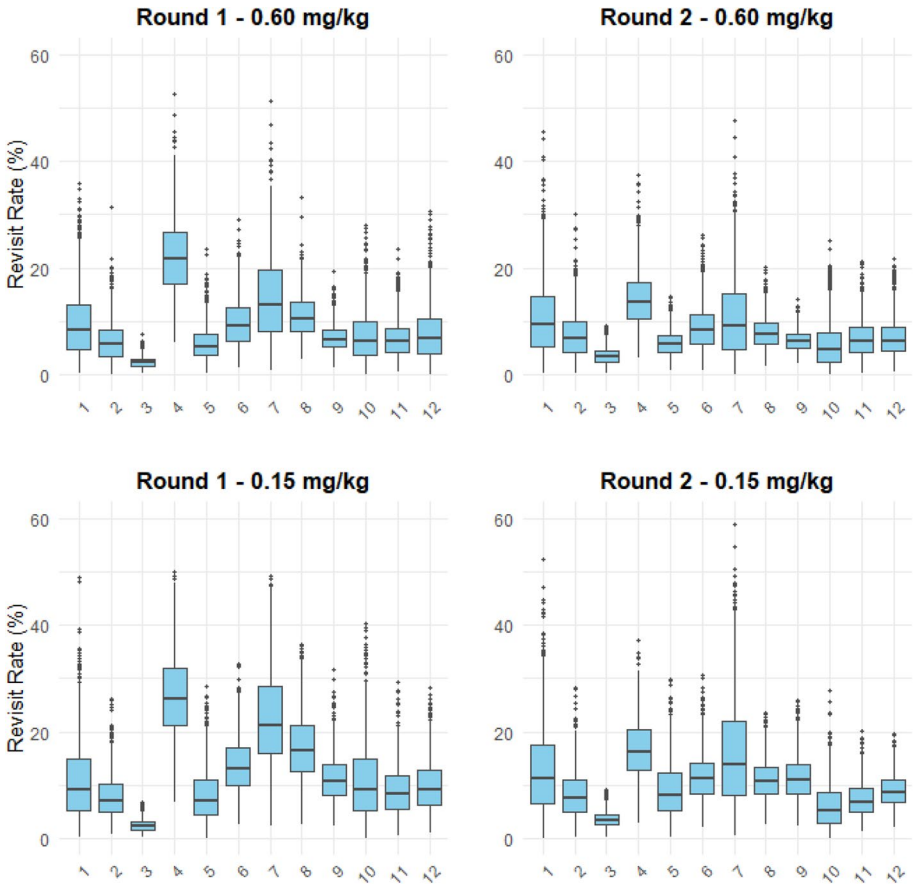

Fig. 1 Individual prior distributions for the probability of ED revisit under 0.60 mg/kg and 0.15 mg/kg of dexamethasone by round

Table 2 Mean and standard deviations (SD) of individual prior distributions for the probability of ED revisit under 0.60 mg/kg and 0.15 mg/kg of dexamethasone

	Round 1		Round 2		Round 1		Round 2	
	0.60 mg/kg		0.60 mg/kg		0.15 mg/kg		0.15 mg/kg	
Expert	Mean	SD	Mean	SD	Mean	SD	Mean	SD
1	9.43%	6.33%	11.37%	7.61%	10.79%	7.05%	12.74%	9.34%
2	6.16%	3.65%	7.53%	4.36%	7.90%	3.92%	8.63%	5.66%
3	2.47%	1.08%	3.61%	1.49%	2.47%	1.08%	3.61%	1.73%
4	21.75%	6.94%	14.58%	5.10%	26.64%	7.88%	16.68%	4.57%
5	5.91%	3.36%	5.81%	2.21%	8.32%	5.21%	9.12%	5.24%
6	9.87%	4.40%	8.74%	3.98%	13.71%	5.10%	11.67%	4.39%
7	14.36%	7.98%	10.88%	7.71%	22.74%	8.77%	15.80%	13.03%
8	11.24%	3.91%	8.01%	2.92%	16.92%	6.13%	11.24%	2.84%
9	6.85%	2.47%	6.53%	1.91%	11.43%	4.21%	11.43%	4.21%
10	7.51%	5.05%	5.38%	4.16%	10.88%	7.71%	6.16%	3.03%
11	6.78%	3.50%	6.78%	3.50%	9.26%	4.64%	7.45%	2.62%
12	8.05%	4.93%	6.78%	3.50%	10.16%	4.78%	8.97%	2.03%

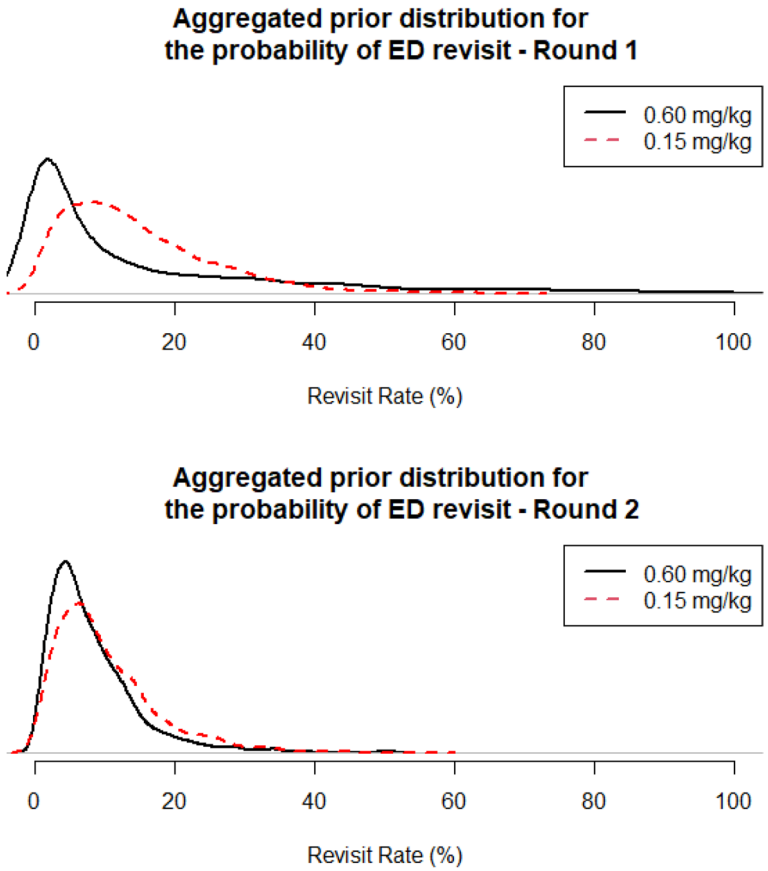

Fig. 2 Aggregated prior distributions from individual expert's elicited beta prior distribution for the probability of ED revisit by round

Table 3 Aggregated prior distributions from individual expert's elicited beta prior distribution for the probability of ED revisit from round 2

	Mean	Median	SD
p1	8.01%	6.44%	6.22%
p2	1.00%	8.19%	7.27%
p2–p1	1.98%	1.52%	7.42%

the goals of the discussion stage from the IDEA protocol of encouraging critical thinking, discussing evidence, and clarifying linguistic ambiguity. A change in expert beliefs from each round was also reported in the BIPED study [18]. This demonstrates the value of the real-time discussion component to our elicitation.

Prior to the elicitation, there was existing evidence on the efficacy of the Croup Dosing Trial interventions. As noted by one of our experts, the meta-analysis could have been used as the prior distribution for the trial [33]. While this meta-analysis presents the best evidence on the efficacy of the two doses of dexamethasone on return visits, there are only three studies in this area [34, 37, 38]. This relatively small number of studies may result in imprecise estimates [39, 40]. Therefore, we opted to use expert opinion to develop priors. As demonstrated during the group discussion, experts have a broader understanding of the context of their practice and how it differs from other settings. This may allow experts to produce estimates that generalize better than past single-center trials [34, 37, 38]. By consulting with multiple physicians, we also included perspectives from heterogeneous settings differing in case mix [19]. Furthermore, our elicitation exercise asked physicians to consider past empirical evidence, which effectively combined data with expert knowledge. Accordingly, five of our experts brought up using past data and their own experience to inform their estimates during the group discussion. Taken together, our expert elicitation allowed us to generate a more robust prior that accounted for both existing data and physicians' expertise.

The duration of the workshops is a strength and limitation of our methodology. Our remote 90-min workshops could accommodate the schedules of multiple physicians. Conversely, the length of these workshops

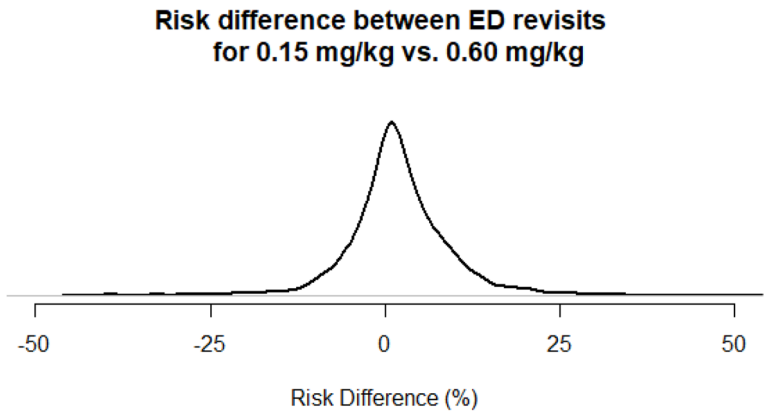

Fig. 3 Risk difference between ED revisits for 0.15 mg/kg vs. 0.60 mg/kg of dexamethasone

limited how much training we provided to our experts. For instance, we did not review potential biases or subjective probability, as recommended by many frameworks [7]. Future elicitations should aim to include such training to derive high-quality prior distributions. Another limitation to our elicitation exercise is that it is unclear why experts changed their opinions following the group discussion. While group discussions support the calibration of expert judgment, they may also introduce biases. For instance, an expert may uncritically adopt the opinions of others, also known as groupthink [41]. This would result in a prior distribution that does not truly reflect the collective experts' opinion, but rather overrepresents certain group members' opinions. To reduce the effects of groupthink, future studies should document expert rationale during the second round of elicitation to increase transparency into each expert's thought process.

Another limitation is that we did not evaluate experts' experience with our elicitation. Future studies should collect whether experts feel adequately trained and what experts think about the elicitation exercise to identify areas that could be improved. Overall, this study successfully demonstrates the remote elicitation of a bivariate prior for the Croup Dosing Trial. The elicited prior is consistent with previous literature on the efficacy of the dexamethasone doses in treating croup. This prior distribution combined empirical evidence with expert opinion and determined the sample size of the Croup Dosing Trial. Our study supports the use of prior elicitation for designing future clinical trials.

Abbreviations

IDEA	Investigate Discuss Estimate Aggregate
ED	Emergency department
PERN	Pediatric Emergency Research Network
PERC	Pediatric Emergency Research Canada
SD	Standard deviation

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13063-026-09828-8.

Supplementary Material 1.
Supplementary Material 2.
Supplementary Material 3.
Supplementary Material 4.
Supplementary Material 5.

Acknowledgements
We would like to thank all elicitation workshop participants, James Chamberlain, Todd A. Florin, Jocelyn Gravel, Lise Nigrovic, Martin Osmond, Tania Principi, Vikram Sabhaney, Graham Thompson, Jennifer Thull-Freedman, Leah Tzimenatos, Rob Woods, Bruce Wright. The authors also like to thank the administrative staff of the Children's Hospital Research Institute of Manitoba for their support of our research work and appreciate the financial contributions from the Children's Hospital Foundation of Manitoba toward REACH (Research into the Enhancement of Acute care for Children's Health).

Authors' contributions
AH designed the elicitation exercise with the support of TPK. AG created the elicitation tool. ACP, ED, WC, ME, and TPK helped develop the elicitation case scenario. AG and AA managed expert recruitment. AJ, AA, AG, and AH facilitated the elicitation. AG collected data and managed communication with the experts. BO and HB participated in the elicitation workshops. AJ analyzed the data and drafted the manuscript. AA, AG, and YO contributed substantive revisions. YO supervised the writing process. All authors approved the final manuscript.

Funding
AA is supported by a Research Manitoba Operating Grant. AJ's time was also supported by the Canadian Network of Statistical Training in Trials. TPK is funded by a Canada Research Chair in Clinical Trials and AH is funded by a Canada Research Chair in Statistical Trial Design.

Data availability
All elicited data generated during this study are included in this published article. Individual survey data are not publicly available due to the ease of identifying individuals but are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate
The elicitation of the Croup Dosing Trial has ethics approval from Health Canada (# CDS0923) as well as University of Manitoba Health Research Ethics

Board & Shared Health (# HS26166 (B2023:092)) approval. The Hospital for Sick Children Research Ethics Board approved a secondary use of the elicitation data (# 1000081720). Implied consent was used in this study.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Child Health Evaluative Sciences, The Hospital for Sick Children, 686 Bay Street, Toronto, ON M5G 0A4, Canada. [2]Children's Hospital Research Institute of Manitoba, Winnipeg, MB, Canada. [3]Department of Pediatrics and Child Health, Max Rady College of Medicine, Rady Faculty of Health Sciences, University of Manitoba, Winnipeg, Canada. [4]College of Medicine, University of Saskatchewan and Jim Pattison Children's Hospital, Saskatchewan Health Authority, Saskatoon, SK, Canada. [5]Department of Pediatrics, University of Ottawa, Ottawa, ON, Canada. [6]Division of Emergency Medicine, Children's Hospital of Eastern Ontario, Ottawa, ON, Canada. [7]Department of Medicine and Dentistry, University of Alberta, Edmonton, AB, Canada. [8]Department of Pediatrics, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada. [9]Department of Biostatistics and Bioinformatics, Roswell Park Comprehensive Cancer Center, Buffalo, NY, USA. [10]Division of Biostatistics, Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada. [11]Department of Statistical Science, University College London, London, UK.

Received: 15 November 2024 Accepted: 28 May 2026
Published online: 05 June 2026

References

1. Berry DA. Bayesian clinical trials. Nat Rev Drug Discov. 2006;5(1):27–36.
2. Ruberg SJ, Beckers F, Hemmings R, Honig P, Irony T, LaVange L, et al. Application of Bayesian approaches in drug development: starting a virtuous cycle. Nat Rev Drug Discov. 2023;22(3):235–50.
3. van de Schoot R, Depaoli S, King R, Kramer B, Märtens K, Tadesse MG, et al. Bayesian statistics and modelling. Nat Rev Methods Primers. 2021;1(1):1–26.
4. Dias LC, Morton A, Quigley J, editors. Elicitation: the science and art of structuring judgement. 1st ed. Springer Cham; 2018. 542 p. Available from: https://doi.org/10.1007/978-3-319-65052-4. Cited 2024 June 20.
5. Hemming V, Burgman MA, Hanea AM, McBride MF, Wintle BC. A practical guide to structured expert elicitation using the IDEA protocol. Methods Ecol Evol. 2017;9(1):169–80.
6. Azzolina D, Berchialla P, Gregori D, Baldi I. Prior elicitation for use in clinical trial design and analysis: a literature review. Int J Environ Res Public Health. 2021;18(4):1833.
7. Bojke L, Soares M, Claxton K, Colson A, Fox A, Jackson C, et al. Developing a reference protocol for structured expert elicitation in health-care decision-making: a mixed-methods study. Health Technol Assess. 2021;25(37):1–124.
8. Gosling JP. SHELF: The Sheffield elicitation framework. In: Elicitation. Springer; 2017. Available from: https://doi.org/10.1007/978-3-319-65052-4_4. Cited 2024 July 22.
9. James A, Choy SL, Mengersen K. Elicitator: an expert elicitation tool for regression in ecology. Environ Model Softw. 2010;25(1):129–45.
10. Johnson SR, Tomlinson GA, Hawker GA, Granton JT, Grosbein HA, Feldman BM. A valid and reliable belief elicitation method for Bayesian priors. J Clin Epidemiol. 2010;63(4):370–83.
11. Morris DE, Oakley JE, Crowe JA. A web-based tool for eliciting probability distributions from experts. Environ Model Softw. 2014;52:1–4.
12. Bjornson CL, Johnson DW. Croup. Lancet. 2008;371(9609):329–39.
13. Brown JC. The management of croup. Br Med Bull. 2002;61:189–202.
14. Kairys SW, Olmstead EM, O'Connor GT. Steroid treatment of laryngotracheitis: a meta-analysis of the evidence from randomized trials. Pediatrics. 1989;83(5):683–93.
15. Geelhoed GC, Turner J, Macdonald WB. Efficacy of a small single dose of oral dexamethasone for outpatient croup: a double blind placebo controlled clinical trial. BMJ. 1996;313(7050):140–2.
16. Yao TC, Wang JY, Chang SM, Chang YC, Tsai YF, Wu AC, et al. Association of oral corticosteroid bursts with severe adverse events in children. JAMA Pediatr. 2021;175(7):723–9.
17. Bojke L, Soares MO, Claxton K, Colson A, Fox A, Jackson C, et al. Reference case methods for expert elicitation in health care decision making. Med Decis Making. 2022;42(2):182–93.
18. Lan J, Plint AC, Dalziel SR, Klassen TP, Offringa M, Heath A, et al. Remote, real-time expert elicitation to determine the prior probability distribution for Bayesian sample size determination in international randomised controlled trials: Bronchiolitis in Infants Placebo Versus Epinephrine and Dexamethasone (BIPED) study. Trials. 2022;23(1):279.
19. Clemen RT, Winkler RL. Combining probability distributions from experts in risk analysis. Risk Anal. 1999;19(2):187–203.
20. European Food Safety Authority. Guidance on expert knowledge elicitation in food and feed safety risk assessment. EFSA J. 2014;12(6):3734.
21. Werner C, Hanea AM, Morales-Nápoles O. Eliciting multivariate uncertainty from experts: considerations and approaches along the expert judgement process. In: Dias LC, Morton A, Quigley J, editors. Elicitation: The Science and Art of Structuring Judgement. Cham: Springer International Publishing; 2018. p. 171–210. Available from: https://doi.org/10.1007/978-3-319-65052-4_8. Cited 2024 July 22.
22. Hiance A, Chevret S, Lévy V. A practical approach for eliciting expert prior beliefs about cancer survival in phase III randomized trial. J Clin Epidemiol. 2009;62(4):431-437.e2.
23. Hampson LV, Whitehead J, Eleftheriou D, Tudur-Smith C, Jones R, Jayne D, et al. Elicitation of expert prior opinion: application to the MYPAN trial in childhood Polyarteritis nodosa. PLoS ONE. 2015;10(3):e0120981.
24. Thall PF, Ursino M, Baudouin V, Alberti C, Zohar S. Bayesian treatment comparison using parametric mixture priors computed from elicited histograms. Stat Methods Med Res. 2019;28(2):404–18.
25. Curran JM, Bolstad WM. Introduction to Bayesian statistics. 3rd ed. Wiley; 2016.
26. R Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2024. Available from: https://www.R-project.org/. Cited 2024 July 24.
27. Chang W, Cheng J, Allaire J, Sievert C, Schloerke B, Xie Y, et al. shiny: web application framework for R. 2024. Available from: https://CRAN.R-project.org/package=shiny
28. Devleesschauwer B, Torgerson P, Charlier J, Levecke B, Praet N, Roelandt S, et al. prevalence: tools for prevalence assessment studies. 2022. Available from: https://cran.r-project.org/package=prevalence. Cited 2024 Aug 15.
29. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform. 2009;42(2):377–81.
30. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, et al. The REDCap consortium: building an international community of software platform partners. J Biomed Inform. 2019;95:103208.
31. Klassen TP, Dalziel SR, Babl FE, Benito J, Bressan S, Chamberlain J, et al. The pediatric emergency research network (PERN): a decade of global research cooperation in paediatric emergency care. Emerg Med Australas. 2021;33(5):900–10.
32. Hogarth RM. A note on aggregating opinions. Organ Behav Hum Perform. 1978;21(1):40–6.
33. Aregbesola A, Tam CM, Kothari A, Le ML, Ragheb M, Klassen TP. Glucocorticoids for croup in children. Cochrane Database Syst Rev. 2023;(1).
34. Parker CM, Cooper MN. Prednisolone versus dexamethasone for croup: a randomized controlled trial. Pediatrics. 2019;144(3):e20183772.
35. Johnson SR, Tomlinson GA, Hawker GA, Granton JT, Feldman BM. Methods to elicit beliefs for Bayesian priors: a systematic review. J Clin Epidemiol. 2010;63(4):355–69.
36. Rue H, Martino S, Chopin N. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. J R Stat Soc B Stat Methodol. 2009;71(2):319–92.
37. Alshehr M, Almegamsi T, Hammdi A. Efficacy of a small dose of oral dexamethasone in croup. Biomed Res. 2005;16(1):65–72.
38. Fifoot AA, Ting JY. Comparison between single-dose oral prednisolone and oral dexamethasone in the treatment of croup: a randomized, double-blinded clinical trial. Emerg Med Australas. 2007;19(1):51–8.

39. Higgins JPT, Thompson SG, Spiegelhalter DJ. A re-evaluation of random-effects meta-analysis. J R Stat Soc A Stat Soc. 2009;172(1):137–59.
40. Jackson D, Bowden J, Baker R. How does the DerSimonian and Laird procedure for random effects meta-analysis compare with its more efficient but harder to compute counterparts? J Stat Plan Inference. 2010;140(4):961–70.
41. Janis IL. Victims of groupthink: a psychological study of foreign policy decisions and fiascoes. Houghton Mifflin Co; 1972.

Publisher's Note