

Review of a Randomized Trial Comparing 2 Cerumenolytic Agents

THE STUDY we review by Whatley et al¹ is a randomized, double-blind, controlled interventional trial comparing the relative effectiveness of 2 agents—docusate sodium (Colace), triethanolamine polypeptide (Cerumenex), and saline control—in the removal of cerumen obstructing visualization of the tympanic membrane. The investigators' objective was to determine whether a significant difference existed between these 2 agents with or without irrigation in removing cerumen obstruction.¹ The authors used a convenience sample of 93 pediatric patients with cerumen occlusion

See also page 1177

of the external auditory canal. The subjects were aged 6 months through 5 years and were brought to 1 of 2 sites for treatment: a children's hospital emergency department or a large general pediatric clinic. Patients were eligible if they had complete or partial cerumen obstruction on the basis of clinical examination by 1 of 4 investigators. After subject randomization, a study nurse instilled 1 mL of 1 of the treatment agents or the saline control into the ear canal, leaving it there for 15 minutes. Afterwards, the nurse drained and wiped away the agent before the investigator reexamined the patient. If the ear still appeared obstructed, then 1 or 2 rounds of irrigation with tepid water were performed in a controlled fashion. The authors do not state whether cerumen removal was attempted via curette. The main outcome was the proportion of ears allowing complete visualization of the tympanic membrane after intervention. Complete visualization occurred in 53% of the docusate group, 43% of the triethanolamine group, and 68% of the control group, with no statistically significant difference in outcome between treatment groups ($P=.17$). Therefore, this is a negative results trial. In our review of this article, we focus on 4 topics: framing the question, sampling and randomization, assessment, and interpretation of results.

FRAMING THE QUESTION

Practicing clinicians continually generate research questions regarding patient care. Most of the questions they frame involve some population of interest, an intervention or exposure, and an outcome.² To identify whether the question tested in any research study will answer a question important to the reader, the reader should search for

a study with a hypothesis closely resembling the clinical question of interest. In this study, the investigators' hypothesis is that there is a significant difference in outcome between docusate and triethanolamine with or without irrigation in the removal of cerumen impaction in children. The authors have defined their population, their interventions, and a qualified outcome.

If the clinically framed question were simply which of 2 agents is better at removing stubborn earwax, then we would not be able to form a definitive answer by using this study. By design, this study is set to determine only whether a large (40 percentage point) difference in outcome exists. Smaller differences may exist, but determining them may not have been within the scope of the study. When interpreting results of a clinical trial, the reader must remember that the results generated are usually qualitative information. The reader should not attempt to definitively answer which agent is better on the basis of qualitative information. A more appropriate question would be if there is a difference between the effectiveness of the 2 agents. To interpret results, the reader should adopt a neutral position, called the "null hypothesis." The null hypothesis is the assumption that no differences will exist between intervention groups. The investigators perform the statistical analysis of results to demonstrate whether true differences exist.

The authors report a 40 percentage point difference as their anticipated treatment effect. This estimate was based on a previous trial conducted with adult subjects and without the benefit of a control group.³ Results of that trial demonstrated docusate to be more effective than triethanolamine. The current study, conducted with children and using a control group, allowed the authors to demonstrate that irrigation alone is an effective therapeutic intervention that completely cleared the wax in 68% of the control group. It seems unlikely with such a high clearance rate in the control group alone that the clinical differences between docusate and triethanolamine would be able to exert a further 40 percentage point treatment difference between groups. The authors' demonstration of the effectiveness of saline instillation and irrigation may be an important finding for future studies.

*See the enhanced version at
<http://www.archpediatrics.com>*

SAMPLING AND RANDOMIZATION

Sampling of study participants at the 2 sites occurred only at times when the enrolling physicians were available, a process known as convenience sampling. Convenience sampling is a cost-effective and simple sampling design option, so it is common in clinical research.⁴ The reader should determine whether a convenience sample would represent the broader population that the research question is meant to address. If the convenience sample differs from the broader population, then generalization of study results may be questionable and another sampling design should be used. Although convenience sampling in this study resulted in few nighttime emergency room patients being enrolled, the reader can probably assume that cerumen impaction in young children varies little by time of day. In summary, the authors made reasonable use of convenience sampling.

After deciding on a sample, proper randomization of study participants is essential in a clinical trial because it provides the basis for testing the statistical significance of any difference in outcome between study groups.⁴ If faults occur in randomization, any difference between groups after the intervention could have been caused by differences in baseline characteristics rather than as a result of the intervention. In this study, the authors report that a hospital pharmacist randomized subjects by using a computerized random-number program and consecutively numbered opaque envelopes. We can infer that all envelopes were generated at 1 site, or central randomization.

Central randomization by someone who does not have contact with study participants or investigators prevents investigators from influencing, often unconsciously, which study intervention a patient receives. Knowing the details of how subjects were randomized becomes important in interpreting the results of the current study because the distribution of who received which drug was not evenly balanced—subjects in the emergency department were disproportionately treated with triethanolamine. Unfortunately, it is not clear how the numbered envelopes were assigned to the 2 sites. Were the envelopes distributed as a group to each site, with each site picking an envelope as each subject enrolled? Or did each site contact the central randomization person, and then the central person chose the envelope? Proper randomization technique does not eliminate misdistribution due to chance variation. The authors attributed the disproportionate number of emergency department participants who received triethanolamine to chance. However, elaboration by the authors on the particulars of the randomization process would support their contention that differences in the study groups' allocation occurred because of chance alone.

A method of ensuring balance in numbers by site and within treatment arms is block randomization. Block randomization ensures that, within a certain number of subjects, typically some multiple of the number of treatment arms, the allocation of patients to treatment arms is equal.⁴ This technique ensures that there are not more subjects receiving a certain treatment at 1 site and that the allocations to treatment during the study are equal.

For example, block randomization would prevent having one treatment disproportionately assigned early in the study and another later in the study.

Blinding, another essential component of a clinical trial, protects against confounding or cointervention by the investigators and the participants.⁴ Cointervention occurs when an investigator treats patients differently because of knowledge of which intervention they received.⁴ The color and texture differences of the cerumenolytics and saline made absolute blinding impossible in this study. To compensate, the nurses attempted to passively drain the ear prior to the second assessment. The authors' approach was reasonable and was perhaps the best option given the study design. When blinding is essential to outcome measurement, one should always measure whether the blinding method was effective. Measurement of the effectiveness of blinding can be as simple as asking the assessors whether they knew which treatments the patients received.

ASSESSMENT

It is often difficult for researchers to determine whether subjects in a study meet study eligibility or whether they have a particular condition or outcome of interest. In this particular study, the eligibility of each patient was dependent on whether the tympanic membrane was completely or partially obstructed by cerumen. When multiple examiners participate in a trial and their determination of whether a subject has a condition is subjective (in this case, their interpretation of external auditory canal occlusion), how well the multiple raters agree becomes important to interpreting the findings. Agreement among observers, especially for common outcomes, is often because of chance or expected agreement. To correct for chance agreement, κ or weighted κ values are calculated to quantify nonrandom agreement among observers, investigators, or measurements.⁵

INTERPRETATION OF RESULTS

Although it is sometimes difficult for the reader to assess the validity of research study results, the challenge can be more difficult with negative results studies. In a previous Evidence-Based Journal Club article in the ARCHIVES in which the authors reviewed a study evaluating the postpartum interview and the factors affecting patients' learning and satisfaction,⁶ the authors reviewed the concepts of power and sample size in the context of how they help the reader assess the effectiveness of an intervention. The number of subjects and the variation among them determine whether a study has sufficient power to show a difference between 2 groups or among more groups. Whatley et al¹ determined the sample size needed to have 80% power to detect a 40 percentage point difference between the treatment groups. The choice of 80% power implies that even with a 40 percentage point difference between treatment groups, there is a 20% chance that this difference would not be detected because the authors are sampling from a larger population. It may be the case that there truly is a 40 percentage point difference in the larger population, but that degree of varia-

tion may not be present in the sample the investigators evaluated (ie, you would accept the null hypothesis of no difference when a difference actually existed).

If, on the other hand, there is a true difference between treatment groups but the difference is less than 40 percentage points, the actual power of the study would be less than 80%. With a smaller true difference, the investigators would need more patients enrolled to still have 80% power. In addition to evaluating the statistically significant differences in a study, the reader must determine how big of a difference between groups would seem clinically relevant. For example, if the true difference in the effectiveness of 2 drugs was 33 percentage points, would a reader decide that a drug is better if it worked that much better than another drug?

CONCLUSIONS

This study is a well-conducted clinical intervention trial comparing 2 agents commonly used to help remove cerumen. The issue is particularly pertinent to the pediatric population and practicing physicians. The investigators used appropriate methods of sampling, but the trial had notable limitations in how randomization and blinding were conducted. The study results demonstrate that, when combined with irrigation, there is no statistically significant difference in outcome between docusate sodium and triethanolamine polypeptide in treating impacted cerumen in the ambulatory setting. The study results also demonstrate that irrigation is an effective method of cerumen removal. There was a trend toward finding that docusate sodium or triethanolamine polypeptide impeded irrigation (53% and 43% cleared vs 68% for saline). The difference in clearance rates between triethanolamine polypeptide and saline is 58% and 25 percentage points but is not statistically significant.

Because this was a negative results study, there are limitations to answering the framed question of which agent is better at removing wax. First, the fact that the patients receiving saline had the best clearance rate suggests that it should not have been the control but perhaps could be considered another active intervention. The authors concede that fact. Perhaps an irrigation-alone arm of the study

would have helped clarify the findings. Second, the study is powered to detect large outcome differences and therefore is susceptible to a type II error, meaning that there is a difference in outcomes but there are too few patients to detect it. Third, another point important to clinicians who may use these medications is that use of only a single dose of the agents is evaluated in the study. This study cannot show whether either agent may be useful when used for a longer time, but this was not what the authors attempted to evaluate. Future studies in which this clinical research question is evaluated should include more patients and consider that saline irrigation alone might actually be the superior approach and not the control method, and they might also seek to answer whether these medications are useful when used for longer periods.

Henry M. Lemon, MD
Jennifer Mazyck-Brown, PharmD
Tara A. Cancellaro, MD
Paul M. Darden, MD
William T. Basco, Jr, MD
Charleston, SC

Corresponding author: William T. Basco, Jr, MD, Department of Pediatrics, Medical University of South Carolina, 326 Calhoun St, PO Box 250106, Charleston, SC 29425 (e-mail: bascob@musc.edu).

REFERENCES

1. Whatley VN, Dodds CL, Paul RI. Randomized clinical trial of docusate, triethanolamine polypeptide, and irrigation in cerumen removal in children. *Arch Pediatr Adolesc Med.* 2003;157:1177-1180.
2. Guyatt G, Rennie D, eds. *Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice.* Chicago, Ill: American Medical Association; 2002: 16.
3. Singer AJ, Sauris E, Viccellio AW. Ceruminolytic effects of docusate sodium: a randomized, controlled trial. *Ann Emerg Med.* 2000;36:228-232.
4. Hulley SB, Cummings SR, eds. *Designing Clinical Research: An Epidemiologic Approach.* Baltimore, Md: Williams & Wilkins; 1988.
5. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33:159-174.
6. Mabry IR, Richmond T, Bialostozky A, Rushton J. Interpreting negative results: postpartum interview position not associated with improved outcomes. *Arch Pediatr Adolesc Med.* 2003;157:333-335.