Pediatric residents' use of jargon during counseling about newborn genetic screening results

Michael Farrell  
*Medical College of Wisconsin*

Lindsay Deuster  
*Medical College of Wisconsin*

Jodi Donovan  
*Medical College of Wisconsin*

Stephanie Christopher  
*Marquette University, stephanie.christopher@marquette.edu*

Stephanie Christopher was affiliated with the Medical College of Wisconsin at the time of publication.
Pediatric Residents’ Use of Jargon During Counseling About Newborn Genetic Screening Results

Michael Farrell, MD, Lindsay Deuster, BS, Jodi Donovan, BS, Stephanie Christopher, MA

Center for Patient Care and Outcomes Research, Medical College of Wisconsin, Milwaukee, Wisconsin

The authors have indicated they have no financial relationships relevant to this article to disclose.

What’s Known on This Subject
Communication between parents and physicians about newborn screening may be significantly hampered if the physician includes too much medical jargon, especially when the parent’s health literacy is limited.

What This Study Adds
A large number of jargon words and a small number of explanations were found, which suggests that physicians’ counseling about newborn screening may be too complex for some parents.

ABSTRACT

OBJECTIVE. The goal was to investigate pediatric residents’ usage of jargon during discussions about positive newborn screening test results.

METHODS. An explicit-criteria abstraction procedure was used to identify jargon usage and explanations in transcripts of encounters between residents and standardized parents of a fictitious infant found to carry cystic fibrosis or sickle cell hemoglobinopathy. Residents were recruited from a series of educational workshops on how to inform parents about positive newborn screening test results. The time lag from jargon words to explanations was measured by using “statements,” each of which contained 1 subject and 1 predicate.

RESULTS. Duplicate abstraction revealed reliability κ of 0.92. The average number of unique jargon words per transcript was 20; the total jargon count was 72.3 words. There was an average of 7.5 jargon explanations per transcript, but the explained/total jargon ratio was only 0.17. When jargon was explained, the average time lag from the first usage to the explanation was 8.2 statements.

CONCLUSION. The large number of jargon words and the small number of explanations suggest that physicians’ counseling about newborn screening may be too complex for some parents. Pediatrics 2008;122:243–249

The quality of physicians’ communication after newborn screening is a growing area of concern. Routine screening makes early treatment of congenital diseases possible, but screening must be followed by effective communication so that parents understand what the test results mean. High-quality communication services are especially important when abnormal results are false-positive findings or reveal that the infant is a genetic carrier for cystic fibrosis (CF) or sickle cell hemoglobinopathy (SCH). These infants are healthy, but some authors say that such infants have a “nondisease” because many of their parents develop psychosocial complications such as anxiety, depression, stigmatization, or misconceptions about whether the infant has a disease. Depending on the disease, there are 20 to 200 nondisease results for every true-positive result. Therefore, the incidence of nondisease cases is high. Most nondisease results are likely to be presented to parents by the infants’ primary care providers. Unfortunately, communication in primary care about newborn screening has been criticized by screened families and public health officials. Screening programs provide educational materials to help physicians, but we found that most programs lack a follow-up mechanism to monitor psychosocial outcomes. Parents might be helped by a discussion with a genetic counselor, but the supply of counselors is limited enough that most carrier families identified through newborn screening are not able to access their services. Such problems with communication and psychological outcomes are often cited by ethicists and policy experts in arguments against the routine use of genetic and molecular screening technologies. To ensure that newborn screening results in “more good than harm,” we have argued that the screening programs or referral centers should introduce population-scale interventions to assess and to improve parents’ psychological and learning outcomes, as well as the processes of communication in primary care.

This study of jargon usage by pediatric residents after newborn screening is part of our larger effort to develop a communication assessment method that is inexpensive and reliable enough for use as part of routine quality improvement efforts across entire states. “Jargon” is a term for the specialized language of a trade or profession that is unlikely to be easily understood by persons outside the profession. Jargon is most problematic for patients with limited health literacy, but jargon can leave any patient feeling alienated and no better informed than before.
the conversation.22–26 However, there are many situations in which jargon is necessary for communication, such as when there is no common word for an important concept or when it is anticipated that the patient will need to hear or to use the word repeatedly. When jargon must be used, jargon words should be accompanied by an explanation.23–26 If physicians’ discussions about newborn screening include too much jargon or insufficient explanation, then many parents may be confused, emotionally upset, and uninformed about their infants’ health status.

In developing a measure of jargon, we considered published work by several investigators who included measures of jargon in their own instruments,26–28 but we became concerned that these measures would be unsuitable for use in future quality improvement projects because they are too labor-intensive, they require highly trained analysts, they are not quantitatively reliable, or they are not transparent enough to be understood by participating physicians. Therefore, we adopted the quality indicator approach that is used by more-traditional versions of quality improvement. Quality indicators are explicitly defined measures that provide a basis for reliably quantifying and comparing the structure, process, or outcome of health care services; each indicator corresponds to a small but important sector within the overall domain of quality.29 A quality indicator technique to quantify jargon usage should help to provide meaningful feedback to physicians about their communication after newborn screening. Once they are established with newborn screening, we expect that communication quality assurance methods will facilitate efforts to improve communication quality in all aspects of health care.

METHODS

Design

This study used an explicit-criteria procedure to abstract transcripts of conversations between pediatric residents and standardized parents of a fictitious infant whose newborn screening test suggested carrier status for CF or SCH. The abstraction procedure was adapted from methods used in medical chart review,30 with a quality improvement-style data dictionary derived from communication guidelines.23–26 Methods were approved by the institutional review boards at Yale University and the Medical College of Wisconsin.

Participants and Data Collection

The study used transcripts obtained during 4 workshops in a prominent pediatrics residency program. The workshops were part of the official curriculum, but residents gave informed consent and were offered a chance to decline the use of their tapes for research. The workshops began with a 10-minute review of newborn screening, CF, SCH, and autosomal recessive genetics. The review avoided any mention of how to discuss newborn screening results with parents. Each resident was then taped in 1 CF carrier encounter and 1 SCH carrier encounter, the order of which was distributed randomly. A handout provided the participant with some contextual information but did not prompt how to inform the parent. In the SCH carrier scenarios, the handout reported a screening result of hemoglobin F, A, and S, a result that had been presented in the review session as definitely indicating that an infant is a SCH carrier. In the likely CF carrier scenarios, the handout reported an elevated screening immunoreactive trypsinogen level and the presence of one ΔF508 mutation, with no multiallele follow-up screening. The review sessions had presented such a result as suggesting that the infant was probably a carrier but still had a 5% to 10% chance of having the disease as the result of an undetected allele.31 The infant’s mother and father were portrayed in the scenario as both being adopted, so that the session could focus on risk communication rather than on taking a family history. Six standardized parents worked on the project; each was female and was chosen to depict plausibly the age and ethnicity of a mother of an infant with CF or SCH. The patients were coached to avoid leading questions, requests for clarification, and any appearance of anxiety or confusion.

Data Handling

The final sample for analysis consisted of 59 transcripts (30 for a SCH carrier infant and 29 for a likely CF carrier infant). The tapes were transcribed verbatim and proofread for accuracy by a board-certified pediatrician. To lessen abstractor bias, all residents’ names were deleted from transcripts.

To guide abstraction and to provide a content-related unit of duration within the transcript, we used a sentence-diagramming procedure to divide transcripts into individual “statements,” each of which either implied or explicitly stated 1 subject and 1 predicate. This approach was selected to correspond closely to the cognitive demands of holding several unfamiliar concepts in mind at the same time, as well as because sentence diagramming was more specific than word counts or time indexes. The statement approach was also simpler for abstractors than the widely used “utterance” approach from the Roter Interaction Analysis System (which looks for individual concepts but also parses speech at 1-second pauses, tonal changes, speaker emphases, and conjunctions).

Development of the Data Dictionary

The data dictionary contained lists for 3 different classes of jargon words (highly specialized, common but confusing, and uncommon) and explicit criterion definitions for 2 levels of explanation (definite and partial). Development of the jargon lists followed a carefully structured combinatorial procedure from corpus linguistics, with 7 steps. In the first step, a corpus document was constructed by merging all 59 transcript files. In the second step, a word frequency analysis was performed with the corpus by using Textanz software (Cro-Code, St Petersburg, Russia) to extract a complete list of words and 2- or 3-word combinations. In the third step, the lemmas (roots) of the words on the frequency list were cross-indexed to remove words included in the Dale-Chall list
of 3000 familiar words, in the 2000-word controlled vocabulary used for the Longman Dictionary of Contemporary English, or in the American Heritage Children’s Dictionary. In the fourth step, the remaining words were cross-indexed against Stedman’s Medical Dictionary, or in the Longman Dictionary of Contemporary English, to identify words for the highly specialized list. All abbreviations were also listed as highly specialized, with the exceptions of “U.S.” and “U.S.A.” Steps 1 through 4 were performed with a spreadsheet, without the need for subjective judgment. In the fifth step, the remaining words were examined collaboratively by the 4 authors, for assignment to either the uncommon list of jargon words, containing words that some people may not recognize, or to the common-but-confusing list, containing words that are common in English but were used for an uncommon concept in the transcripts. For example, the words “sensitive” and “carrier” are moderately common in English but convey different meanings when applied to the fields of statistics or genetics. In the sixth step, the words that had been excluded in the third step were reexamined by the 4 authors, for identification of additional words for the common-but-confusing list. In the final step, a posthoc amendment process allowed abstractors to propose words that they thought were jargon, provided that the word could be ratified by 1 other abstractor. When such a word was identified, an electronic search of previously abstracted transcripts was conducted to verify that the newly designated jargon word had not been missed in previous abstractions.

Another set of criteria was devised to operationalize explanations of jargon. It was expected that the effectiveness of the residents’ explanations would vary widely; therefore, a trichotomous (definite/partial/absent) explicit-criteria quality indicator was developed for abstractors. In this scheme, to be assigned a definite explanation rating, the statement had to not use jargon itself and had to refer to the jargon word or abbreviation directly. Any explanation that used another jargon word to define a referenced jargon word was assigned a partial rating.

Abstraction and Analysis
Abstractors were instructed to read each transcript statement, looking for any type of explanation and comparing all words to the lists in the data dictionary. The abstractors were asked to avoid judgments about whether the residents’ explanations were factual correct. All transcripts were abstracted by 2 authors, for assessment of interabstractor reliability. One third of abstractions were discussed later, to ensure quality control and consistency, following the suggestion made by Feinstein. Interabstractor reliability for jargon words was calculated before merging or consensus, by using Cohen’s method. Discrepancies between abstractors were resolved automatically with a spreadsheet, to avoid subjective judgment.

Statistical analyses and calculations of time lag were performed by using Excel (Microsoft, Redmond, WA) and JMP (SAS Institute, Cary, NC) software. One-way analysis of variance, t tests, and χ^2 tests were used as appropriate for the type of variables being analyzed.

### RESULTS

#### Participant and Interview Characteristics
Participant characteristics were similar to those of the population of the residency program at the time of the study (Table 1). The interviews ranged from 4.1 to 20.75 minutes (mean: 9.8 minutes; SD: 4.2 minutes). On average, interviews lasted longer when the infant was a likely CF carrier than when the infant was a SCH carrier (11 and 8.7 minutes, respectively; P = .03). Transcripts averaged 165.8 statements per transcript (range: 65–401 statements). Interabstractor reliability calculations for abstractions revealed a κ coefficient of 0.92.

#### Jargon Words Included in Counseling
The average number of unique jargon words per transcript was 20.0 (SD: 9.4 words), but many jargon words were used more than once, so that the total jargon count averaged 72.3 words per transcript (SD: 40.5 words). The average number of unique jargon words was greater for likely CF carrier transcripts than for SCH carrier transcripts (23 and 17 unique words, respectively; P = .01). The distributions of total jargon counts for the 2 types of transcripts are depicted in Fig 1. A slight difference in the total counts was not significant (77.8 vs 67.0 total words; P = .31, with a least significant number of 216 transcripts for the observed difference). No significant differences in the amounts of jargon words (unique or total) according to residents’ gender or year in residency were apparent.

The 5 most frequent jargon words included in counseling about SCH carrier and likely CF carrier screening results are listed in Table 2. Five of the words were from the highly specialized list of jargon words and 4 were from the common-but-confusing list.

#### Explanations of Jargon
Definite criteria for a jargon explanation were found in 51 (86.4%) of 59 transcripts. There was an average of 7.5 explanations per transcript (SD: 4.0 explanations), with no apparent difference between SCH carrier and likely CF carrier transcripts. The 5 most frequently explained jargon words for the 2 types of screening results are listed in Table 3.

The average explained/total jargon ratio was 0.17 (SD: 0.13). In other words, residents failed to explain...
83% of their jargon, on average. There was no apparent difference in the explained/total jargon ratios in the SCH carrier and likely CF carrier transcripts.

Time Lag Between Jargon Words and Explanations
When jargon words were explained, it was common for several concepts to be introduced between the first usage of the jargon word and its explanation. The average time lag from the first usage of the word to its explanation was 8.2 statements (SD: 9.3 statements). The lag was greater in the likely CF carrier transcripts than in the SCH carrier transcripts (9.5 and 7.0 statements, respectively), but there was inadequate power to ascertain whether this was a significant difference. There were no significant differences according to residents’ gender or year in residency.

DISCUSSION
Jargon is a key barrier to effective communication, especially when the topic is as complicated as the implications of positive newborn screening test results. In this study, we examined residents’ use of jargon during counseling about newborn screening results and found that jargon words were common, explanations were rare, and explanations often lagged far behind the first usage of the jargon word. There were no apparent differences between the 2 screening results with respect to total jargon counts or explanations, although the higher unique jargon word count and longer duration for the likely CF carrier transcripts led to speculation that the residents might have had more difficulty explaining this result.

Our results raise questions about whether many parents will be able to understand health care providers’ explanations of these types of screening results. This is a troubling possibility, because much newborn screening policy-making has focused on “affected” infants with diseases instead of infants who are affected by the screening process itself. It has been helpful for us to use the “nondisease” term for these infants, because the term suggests that parents’ learning of the screening result can lead to an actual condition with its own symptoms, risks, and need for treatment. We are particularly concerned about the challenge that jargon
and nondisease results present for parents with limited health literacy, who often are unfamiliar with medical terminology.

To put our results in perspective, it may help to envision a hypothetical group of 59 parents coming to their infant’s physician to hear about the results of the newborn screening test. In this scenario, the average physician’s answer would have included 72 potentially confusing words, ~83% of which would not have been explained. If these parents felt uncomfortable asking questions about the jargon, then they might have developed psychosocial complications and had difficulty making informed decisions or adhering to a recommended treatment plan. Indeed, if jargon usage left the parents uninformed, anxious, and alienated, then it could be said that the counseling was more harmful than beneficial.

The finding that 86% of the residents explained ≥1 jargon word suggests that many may already be aware of the potential for patients to misunderstand technical language. This awareness is important to recognize when interventions to reduce jargon usage are being designed, because informational interventions such as guideline dissemination may have trouble improving the behavior of physicians who are already aware of a problem. Additional research should investigate whether physicians tend to overestimate their patients’ vocabularies for medical terminology or whether there are skill barriers or some other reason why jargon is not adequately explained.

A more-subtle problem is presented by our finding of a mean time lag of 8.4 statements between the first instance of the jargon word and its explanation. There are no previously published data about the effect of a time lag, but research can proceed now that a reliable method for quantifying it is available. For now, physicians should be guided by the cognitive psychology literature, which suggests that explanations should closely accompany jargon because delays can increase the patient’s cognitive workload.16-41

This study was limited by its small sample size, but we see some limitations as strengths from a quality improvement perspective. Qualitative methods would have provided a richer description of conversations, but qualitative methods have limited reliability and would be prohibitively expensive for use in quality improvement. The use of standardized patients instead of real patients avoided logistic, privacy, and consent difficulties that would make quality improvement activities difficult. Simulation is useful because a sense of observation prompts physicians to perform as well as they can; the resulting competence data suggest a likely ceiling for the physician’s processes of communication, because competence is necessary but not sufficient for real-world performance.42 Simulation also allows an equal-footing comparison that would be impossible with real patients because of variations across actual patients. Data from residents may not be generalizable to other residency programs or to clinicians already in practice, but we saw residents as being ideal for this demonstration project because many are near the peak of their content knowledge about genetics and newborn screening.

Another methodologic challenge to this project has been the difficulty of populating our data dictionary’s 3 jargon word lists. Our method met the required feasibility criteria for quality improvement, but a more-accurate list might have been derived with cross-sectional surveys in which healthy people were asked to define various words or to use them in a sentence. Such surveys would be similar to those used to construct the Dale-Chall list of common words32 to construct his list of common words, but it is unclear how long the results of those surveys would be relevant, given the variance of patient vocabularies and the increased use of the Internet. The chief advantage of our automated combinatorial approach is that it is simple enough to be replicated and tailored as needed for each new clinical topic or patient population. For now, the best practice for clinicians may be to ask regularly about the patient’s understanding. We address such assessment-of-understanding communication behaviors in another set of quality indicator articles.17,43 Future studies of jargon and assessments of patient understanding can determine whether there is a dose-response relationship between the amount of jargon included or explained and communication outcomes such as patient comprehension, satisfaction, and decision-making.

We are incorporating the jargon quality indicators into our population-scale “communication quality assurance” approach to assessing and improving the processes and outcomes of communication in health care. We designed the methods to meet key demands of quality improvement, such as quantitative reliability, transparency, fairness, and ability to be implemented on a lean budget. By comparison, many projects that included a jargon-related item among Likert-type interview rating scales used labels ranging in detail from simple adjectives (eg, good/fair/poor) to a paragraph of definition for each point on the scale. These types of flexible-choice scales are not known for high inter-rater reliability and may be subject to bias resulting from raters’ personal preferences for communication style or raters’ opinions about whether any given word is jargon. If communication is to be included in quality improvement or pay-for-performance schemes, then physicians will probably demand that the assessment methods be transparent enough for them to understand exactly why they received (for example) a “fair” rating instead of a “good” rating. These challenges may be met through the use of explicit-criteria methods that we have been developing for use in communication quality assurance efforts.16-20 We anticipate the finding of similar problems with jargon in many clinical settings besides newborn screening. With an assessment method in hand, targeted interventions can be developed to reduce physicians’ use of jargon in counseling and to increase clinicians’ use of jargon explanations.

**CONCLUSIONS**

The large number of jargon words and the small number of explanations suggest that communication quality dur-
ing counseling about newborn screening tests often is suboptimal. When excessive jargon is used and not explained, patients may not understand the screening tests enough to participate fully in the decision-making process. Furthermore, patients may not be adequately prepared for positive newborn screening results if they do not fully understand the nature of these tests before they are administered. Increased jargon explanations and decreased time lag between jargon use and explanation should enhance communication quality and have the potential to increase patient participation in and satisfaction with care.

ACKNOWLEDGMENTS
Ms Deuster was supported by a training grant from the National Heart, Lung, and Blood Institute (grant T35-HL72483-24). Dr Farrell was supported in part by grant HL72483-24). Dr Farrell was supported in part by grant National Heart, Lung, and Blood Institute (grant T35-MP000004). Ms Deuster was supported by a training grant from the National Heart, Lung, and Blood Institute (grant T35-HL72483-24).

REFERENCES
27. Aspegren K, Lönnberg-Madsen P. Which basic communication skills in medicine are learnt spontaneously and which need to be taught and trained? Med Teach. 2005;27(6):539–543
41. Miller GA. The magical number seven, plus or minus two: some limits on our capacity for processing information. Psychol Rev. 1994;101(2):343–352

COMPREHENSIVE PEDIATRIC HOSPITAL MEDICINE*

Edition: 1st
Editors: Zaoutis, Lisa B., Chiang, Vincent W., Ballantine, Allison, Landrigan, Christopher P., Mirkinson, Laura J.
Publisher: Elsevier
Imprint: Mosby
Reviewer: Barry Starr, MD (Ochsner Clinic Foundation)
Description: The book covers a broad range of topics related to the care of hospitalized infants, children, and adolescents from the perspective that pediatric hospital medicine is a new and changing field.
Purpose: The purpose is to provide a valuable resource to those practitioners involved with the hospital care of pediatric patients. There are few books that deal exclusively with pediatric hospital medicine and the authors have tried to cover most topics related to hospitalized pediatric patients. The book approaches this subject not only from a medical standpoint, but also from the perspectives of care delivery, ethics, and quality of care. The authors succeed in meeting their objectives.
Audience: General practitioners or hospitalists who care for hospitalized newborns, children, and adolescents are the intended audience. The focus is on the general practitioner, not the subspecialist or intensivist.
Features: The book is divided into four parts covering general issues of inpatient pediatric medicine, common presenting signs and symptoms, a systems approach, and procedures. Medical subjects are well represented by system. Chapters include sections on admission/discharge criteria, consultations, suggested readings, references, and future research/advances. This format is useful because it gives readers not only useful clinical information on pathophysiology and presenting symptoms, but also recommendations for admission/discharge and a look at future therapeutic modalities or practices. Specific, nonmedical care topics related to hospitalists, including organization, research, funding, quality of care, and billing are also covered. The rich tables and illustrations are easy to comprehend.
Assessment: This book does an excellent job of presenting a wide range of medical conditions affecting hospitalized infants, children, and adolescents. It will be a valuable tool for practitioners caring for hospitalized pediatric patients as well as for those attempting to establish or improve a pediatric hospitalist practice. It compares well to others in the field (eg, Perkin et al., Pediatric Hospital Medicine: Textbook of Inpatient Management, 2nd edition (Lippincott Williams & Wilkins, 2008)), while providing the same type of information in a similar format. Unique to this book, and which support its purchase, are some topics not generally covered as well as features such as admission/discharge criteria and administrative aspects of a hospitalist practice.

Pediatric Residents' Use of Jargon During Counseling About Newborn Genetic Screening Results

Michael Farrell, Lindsay Deuster, Jodi Donovan and Stephanie Christopher

*Pediatrics* 2008;122;243

DOI: 10.1542/peds.2007-2160

<table>
<thead>
<tr>
<th>Updated Information &amp; Services</th>
<th>including high resolution figures, can be found at: <a href="http://pediatrics.aappublications.org/content/122/2/243.full.html">http://pediatrics.aappublications.org/content/122/2/243.full.html</a></th>
</tr>
</thead>
<tbody>
<tr>
<td>References</td>
<td>This article cites 27 articles, 6 of which can be accessed free at: <a href="http://pediatrics.aappublications.org/content/122/2/243.full.html#ref-list-1">http://pediatrics.aappublications.org/content/122/2/243.full.html#ref-list-1</a></td>
</tr>
<tr>
<td>Citations</td>
<td>This article has been cited by 7 HighWire-hosted articles: <a href="http://pediatrics.aappublications.org/content/122/2/243.full.html#related-urls">http://pediatrics.aappublications.org/content/122/2/243.full.html#related-urls</a></td>
</tr>
<tr>
<td>Subspecialty Collections</td>
<td>This article, along with others on similar topics, appears in the following collection(s): Administration/Practice Management <a href="http://pediatrics.aappublications.org/cgi/collection/administration/practice_management_sub">http://pediatrics.aappublications.org/cgi/collection/administration/practice_management_sub</a> Interpersonal &amp; Communication Skills <a href="http://pediatrics.aappublications.org/cgi/collection/interpersonal_communication_skills_sub">http://pediatrics.aappublications.org/cgi/collection/interpersonal_communication_skills_sub</a></td>
</tr>
<tr>
<td>Permissions &amp; Licensing</td>
<td>Information about reproducing this article in parts (figures, tables) or in its entirety can be found online at: <a href="http://pediatrics.aappublications.org/site/misc/Permissions.xhtml">http://pediatrics.aappublications.org/site/misc/Permissions.xhtml</a></td>
</tr>
<tr>
<td>Reprints</td>
<td>Information about ordering reprints can be found online: <a href="http://pediatrics.aappublications.org/site/misc/reprints.xhtml">http://pediatrics.aappublications.org/site/misc/reprints.xhtml</a></td>
</tr>
</tbody>
</table>