Objective: To demonstrate a quantitative abstraction method for Communication Quality Assurance projects to assess physicians’ communication about hidden emotions after newborn genetic screening.

Design: Communication quality indicator analysis.

Setting: Standardized parent encounters performed in practicing physicians’ clinics or during educational workshops for residents.

Participants: Fifty-nine pediatrics residents, 53 pediatricians, and 31 family physicians.

Intervention: Participants were asked to counsel standardized parents about a screening result; counseling was recorded, transcribed, and parsed into statements (each with 1 subject and 1 predicate). Pairs of abstractors independently compared statements with a data dictionary containing explicit-criteria definitions.

Outcome Measures: Four groups of “precautionary empathy” behaviors (assessment of emotion, anticipation/validation of emotion, instruction about emotion, and caution about future emotion), with definitions developed for both “definite” and “partial” instances.

Results: Only 38 of 143 transcripts (26.6%) met definite criteria for at least 1 of the precautionary empathy behaviors. When partial criteria were counted, this number increased to 80 of 143 transcripts (55.9%). The most common type of precautionary empathy was the “instruction about emotion” behavior (eg, “don’t be worried”), which may sometimes be leading or premature.

Conclusions: Precautionary empathy behaviors were rare in this analysis. Further study is needed, but this study should raise concerns about the quality of communication services after newborn screening.


The quality of parent-physician communication after newborn screening is an important area of concern. Communication influences patients’ understanding and psychological outcomes and may be crucial when screening reveals that an infant is a heterozygous “carrier” for cystic fibrosis (CF) or sickle cell hemoglobinopathy (SCH). Carrier infants are healthy, but some say that such infants have a “non-disease” because of the possibility for misconceptions and psychosocial complications. Such problems have been cited in arguments against the routine use of genetic and molecular screening technologies. One of us (M.H.F.) has counter-argued that newborn screening could result in “more good than harm” if psychosocial risks were managed over the same statewide populations served by the screening programs. This article demonstrates the next step in our long-term effort to develop Communication Quality Assurance for population-scale use; it uses the new tool to assess a pilot sample for communication about the potential for emotions after newborn screening.

Emotions are an important topic for communication after newborn screening. Studies after various types of genetic screening have documented that nondisease results can be followed by clinical levels of anxiety and depression, as well as other stress-related problems, such as physical symptoms, feelings of shock, sleep disturbances, reports of infant feeding problems, and uncertainty about the future. In other areas of health care, patients may be more likely to reveal their emotions and discuss them if the physician initiates the topic of emotions, but they may never reveal their emotions if they are not asked. Physicians may not notice their patients’ emotions or may fail to acknowledge them. Failure to address emotions is problematic.

For editorial comment see page 95
because emotions may be distressing or may impede patients’ ability to process information.

For this article, we chose to examine physicians’ communication about emotions that may be present but are not apparent in the patient’s speech, facial expressions, or body language. Addressing potential emotions is important because patients’ facial expressions do not always correlate with their emotional states for cultural reasons or because of a sense of shock.24,25 We developed the term precautionary empathy for these communication behaviors to distinguish between physician behaviors that react to obvious emotions and to avoid confusion with the broader term empathy. Many definitions of empathy have been proposed,36-37 and many scales32-33 to measure empathy exist, but generally speaking, empathy can be defined as “identification with and understanding of another’s situation, feelings, and motives.”34 We also coined the new term precautionary empathy to avoid confusion in the medical education field, where empathy has sometimes been used as a shorthand term for health care providers’ responsiveness to emotions.35-37 A list of the precautionary empathy behaviors is shown in Table 1 and further described in the “Types of Precautionary Empathy Behaviors” subsection in the “Methods” section.

A method to assess physicians’ communication about the potential for emotions informs our larger efforts to improve physician communication quality. This effort, called Communication Quality Assurance, attempts to improve physicians’ communication by providing quantitatively reliable assessment of communication behaviors. Our panel of communication behaviors is not intended to be exhaustively complete,38-46 but the theoretical framework for Communication Quality Assurance holds that each behavior’s usage may incrementally improve the chance of favorable outcomes. Focusing on communication about emotion is a key component of this goal.

<table>
<thead>
<tr>
<th>Table 1. Precautionary Empathy Behaviors</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Type</strong></td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td>6</td>
</tr>
</tbody>
</table>

**Assessment for Emotion**

The most elementary of the behaviors was for the physician to explicitly ask about emotion. Guidelines suggest that an open-ended question (“How do you feel right now?”) would be more effective at obtaining an accurate answer than a closed-ended question (“Are you worried?”) because the natural answer to the open-ended version would be more detailed than yes or no. Closed-ended questions may be particularly ineffective if a leading syntax is used (“You’re not worried, are you?”).

The Assessment for Emotion behavior may also vary by its timing relative to the emotion. For example, a physician could ask a general question (“Are you worried about all of this?”) or ask whether the patient is worried by the conversation itself (“How is this news making you feel?”). Finally, it may be wise to inquire about emotions from before the conversation (“Were you worried when my nurse asked you to come to my office to talk about this test?”). To facilitate abstraction, we developed criteria for 3 quality indicators: an Assessment of Prior Emotion behavior, an Assessment for Emotion (open-ended), and an Assessment for Emotion (closed-ended).

**Anticipate or Validate Emotion**

In this behavior, a physician reassures the patient that it is normal to have emotions like anxiety or fear or that emotions are nothing to be ashamed about. The patient may then be more likely to reveal emotions to the physician and to discuss them and may even self-reflect and gain a better understanding of his or her feelings. We identified 2 main variants: an “Anticipate Emotion” behavior that indicates the possibility of an emotional reaction (“What I have to say may seem frightening to you at first”) and a “Validate Emotion” behavior in which the physician comments that any current or future emotional response is understandable, common, or normal for people to feel (“Many people become sad when they hear this type of news”). We combined these 2 variants when initial tests found difficulty in discriminating between them.

**Caution About Future Emotion**

This behavior covers remarks about emotions arising after the conversation (“You might get more worried about this after you get home”). We separated this behavior from the Anticipate or Validate Emotion behavior because of differences from the patient perspective. For example, physicians might say, “Here is the number for a sickle cell support group to call if you get to be more worried about this screening result.”
Instruction About Emotion

Physicians sometimes tell patients how to feel (“I wouldn’t be worried about this”). The validity of some instructions may be questionable, have inappropriate timing, or impair trust or openness (eg, “Don’t be sad”). For this analysis, however, we chose to include all instructions to pave the way for future study.

DATA SOURCE

The data dictionary was used to analyze 143 transcripts of conversations between child health providers and standardized parents. To include a variety of communication styles, we developed a heterogeneous sample from 2 data collection efforts. The first 64 transcripts were conversations by pediatric residents who were participating in an educational workshop about newborn screening for CF and SCH. The other 79 transcripts came from a study of counseling about SCH carrier results by physicians who were working outside of academia.

The resident transcripts were collected as part of their curriculum; informed consent was asked for use of tapes in research. Each workshop began with a 10-minute review about newborn screening, CF, SCH, and inheritance. No teaching about communication or emotions was performed during this period. Each resident was taped in 1 SCH trait encounter and 1 likely CF carrier encounter, the order of which was randomly distributed. A handout described the screening result but did not provide any advice about how to communicate with the parent. In the SCH carrier scenarios, residents were told the infant’s screening result showed hemoglobin F, A, and S, a result that had been presented in the review session as definitely indicating that an infant is an SCH carrier. In the likely CF carrier scenarios, residents were told the result showed an elevated immunoreactive trypsinogen, a single F508 mutation, and no multilobulated follow-up screening. Our research group uses the word likely for this result because it is associated with a 5% to 10% chance of having the actual disease due to an undetected second mutation.

All standardized parents were women and chosen to plausibly depict the age and ethnicity of a mother of an infant with CF or SCH. Encounters began with a simple question: “Hello, doctor. I heard you wanted to speak with me, something about my baby’s screening test result?”

The standardized parents were coached to adopt a neutral emotional expression and to avoid any appearance of anxiety or other emotions. These instructions helped our analysis to focus on precautionary empathy rather than on the physicians’ ability to respond to visible emotions. Physicians’ response to visible emotional cues (eg, crying) is a critical skill but is not the subject of this study.

For transcripts from graduated physicians, the counseling task always concerned an SCH carrier result. Names and contact information were obtained from a search of the American Medical Association Masterfile for Pediatricians in Connecticut and Wisconsin and Family Physicians in Wisconsin. Search parameters included self-report of working in direct patient care and graduation from residency at least 2 years before. Participants were offered $20.

The standardized parent encounters followed procedures of the resident project, except that physicians were randomized to counsel either by telephone or in person at their offices. Recordings were transcribed, proofread, and stripped of identifying information. To facilitate the abstraction procedures, we used a sentence diagramming technique to parse transcripts into individual “statements,” each with 1 subject and 1 predicate.

ABSTRACTION

Our procedure for abstracting transcripts is derived from quality improvement methods used for explicit-criteria review of medical records,62 with a data dictionary containing explicit definitions and examples. Abstractors read transcripts statement-by-statement, comparing each statement with the data dictionary.

Abstraction used a trichotomous scale, in which the abstractors rated the statement “definite” if it met full criteria for the communication behavior described in the data dictionary or “absent” if there was no reference to emotion. As with our previous studies,63–65 we used “partial” ratings to represent statements that may have been an attempt to discuss emotion, such as leading syntax, assessment questions without a pause for the patient to answer, or hinting at emotion with a vague term that fails to mention an actual emotion (“so what’s going on in your mind?”).

Each transcript was abstracted duplicatively by 2 of the abstractors, and discrepancies were automatically resolved by a spreadsheet algorithm to final results. The algorithm resulted in a final status of “definite” if the individual abstractors had entered definite or if one had entered definite and the other entered partial. The algorithm resulted in a final status of “partial” for all other combinations of abstractor entries except if both statements had been left blank or deliberately entered as “absent.”

One-third of the duplicate abstractions were discussed further to ensure quality and consistency, following the suggestion by Feinstein.63

STATISTICAL ANALYSIS

Interabstractor reliability was calculated using a weighted adaptation of the Cohen method for the trichotomous (definite/partial/absent) schema. Data were analyzed using the χ2 test for grouped categorical responses and the Wilcoxon rank sum test for nonparametric continuous responses to categorical variables.

We also ran a series of regression models to adjust for overlapping effects between hypothesized predictor variables. The rates of the individual precautionary empathy behaviors were low enough, however, that all of the individual factors lost their significance even in 2-parameter logistic models.

RESULTS

Characteristics of the composite sample are shown in Table 2. Data for the residents were similar to data for the residency program at the time of the study. Data for the graduated physicians were similar to the Masterfile data for physicians who declined participation.

PHYSICIANS’ USE OF PRECAUTIONARY EMPATHY BEHAVIORS

Use of precautionary empathy behaviors was rare. Definite criteria were identified for 58 instances of the precautionary empathy behaviors. Partial criteria were identified for an additional 119 instances. Abstractors agreed on more than 99% of the instances, but the behaviors were rare enough in the sample that the overall Cohen κ was only 0.73.

Only 38 of 143 transcripts (26.6%) met definite criteria for at least 1 precautionary empathy behavior (Figure 1). When abstractors’ partial-criteria ratings were
also counted, the number of transcripts increased by 42 to 80 of 143 (55.9%) (Figure 2).

Transcripts were more likely to include at least 1 precautionary empathy behavior if the infant had likely CF carrier status (86.2% vs 48.3% for SCH carrier infants; χ² test; P = .001). Precautionary empathy was also more common if the physician were a resident (74.6% vs 42.9% for graduates; χ² test; P = .001) or were female (74.4% vs 48.0% for males; χ² test; P = .004). We tested whether effects would persist after adjustment for each other, but the numbers of instances were insufficient for 2-parameter regression.

**INDIVIDUAL TYPES OF PRECAUTIONARY EMPATHY**

**Assessment for Emotion**

Abstractors identified definite criteria in only 1 transcript for open-ended Assessment for Emotion. No transcripts met definite criteria for closed-ended Assessment for Emotion. When partial-criteria abstractions were also counted, open-ended Assessments for Emotion were found in 8 other transcripts (6.3% total). These were more common in likely CF carrier transcripts (17.2% vs 3.6% for SCH carrier infants; χ² test; P = .006).

Twelve transcripts (8.4%) met partial criteria for at least 1 closed-ended Assessment for Emotion. Transcripts were more likely to include a closed-ended Assessment for Emotion if the physician were female (16.3% vs 5.0% for males; χ² test; P = .03) or a resident (15.3% vs 3.6% for graduates; χ² test; P = .01). Eleven transcripts (7.7%) met partial criteria for at least 1 “Assess for Prior Emotion” behavior.

Abstractors agreed on 93.5% of the open-ended assessments, 90.7% of the closed-ended assessments, and 95.1% of the Assessments of Prior Emotion. All of these were so rare that the κ coefficients were not able to exceed 0.5.

**Anticipate or Validate Emotion**

Definite criteria for this anticipation or validation of emotion were identified in 8 of 143 transcripts (3.1%). Partial criteria were identified in 22 transcripts (20.9% total). Transcripts were more likely to Anticipate or Validate Emotion if the physician were female (32.6% vs 16.0% for males; χ² test; P = .004), a resident (34.0% vs 11.9% for graduates; χ² test; P = .004), or younger (Wilcoxon; P = .05). Physicians were less likely to Anticipate or Validate Emotion if the encounter were by telephone (2.9% vs 26.9% for in-person encounters), and more likely if the infant were a likely CF carrier (51.7% vs 13.2% for SCH carrier infants). Abstractors disagreed on 10% of the Anticipate/Validate Behaviors but there were more instances, so the interabstractor reliability was better than for the other behaviors (κ = 0.67).

**Caution About Future Emotion**

Abstractors failed to identify any transcripts with definite criteria for this behavior. When partial criteria were counted, 8 of 143 transcripts (3.6%) were identified with this behavior. Female physicians were more likely than males to Caution About Future Emotion (11.6% vs 3.0%; χ² test; P = .04).

Instruction About Emotion was the most common type of precautionary empathy found in this analysis. The ab-

---

**Table 2. Characteristics of 143 Participants**

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>No. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex</strong></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>73 (51.0)</td>
</tr>
<tr>
<td>Female</td>
<td>70 (49.0)</td>
</tr>
<tr>
<td><strong>Age, y</strong></td>
<td></td>
</tr>
<tr>
<td>21-40</td>
<td>83 (58.0)</td>
</tr>
<tr>
<td>41-60</td>
<td>40 (27.9)</td>
</tr>
<tr>
<td>61-80</td>
<td>16 (11.2)</td>
</tr>
<tr>
<td><strong>Specialty</strong></td>
<td></td>
</tr>
<tr>
<td>Family medicine</td>
<td>31 (21.7)</td>
</tr>
<tr>
<td>Pediatrics</td>
<td>112 (78.3)</td>
</tr>
<tr>
<td><strong>Postgraduate year</strong></td>
<td></td>
</tr>
<tr>
<td>Second</td>
<td>39 (27.3)</td>
</tr>
<tr>
<td>Third or fourth</td>
<td>20 (14.0)</td>
</tr>
<tr>
<td>&gt;2 y out of residency</td>
<td>84 (58.7)</td>
</tr>
</tbody>
</table>

---

*Four participants declined to provide their age.
**Fourth-year pediatric residents were from the combined medicine-pediatrics program.*
Abstractor disagreements were more common for the Instruc-
tions About Emotion behavior than for the other be-
haviors (17.0% disagreement; \( \kappa = 0.56 \)).

COMMEN

Communicating about the possibility for emotions re-
quires different empathic behaviors than responding to
apparent emotional cues. This article demonstrates our
communication quality indicator method for measur-
ing these precautionary empathy behaviors. This method
would be an important component in a communication
quality assessment toolkit because unaddressed emo-
tions may be distressing or impair understanding. The
results raise concern about an apparent problem with com-
munication after “nondisease” carrier newborn screen-
ing results because it would seem that many of our phy-
sicians were unaware that emotional responses may not
be apparent. The sample size is small enough that it is
difficult to draw firmer conclusions about associated fac-
tors, but if our findings are generalizable, then many phy-
sicians fail to recognize parents’ emotional problems af-
ter newborn screening.

Our analysis has several potential limitations. We used
standardized parents to evaluate physicians on a level play-
ing field, but physicians’ awareness that the “parent” is
not truly harboring emotions may have reduced the us-
age of precautionary empathy. On the other hand, the phy-
sicians’ sense of being evaluated should have led to
their best behavior.

Several other factors could have accounted for the rar-
ity of precautionary empathy. Physicians may lack ex-
perience with genetics or may not realize that parents can
become distressed about nondisease results. Some phy-
sicians may not be attuned to the possibility for emo-
tions to be hidden, having been taught to respond to
feigned emotions in standardized parents. Reliability may
be improved in future efforts by increasing the explicit-
ness of the data dictionary, by increasing abstractor train-
ing, or by focusing on fewer behaviors.

Some of these questions may be addressed by our on-
going, statewide study of counseling by primary care pro-
viders of actual infants with nondisease newborn screen-
ing results.55-56 This effort to assess communication quality
may allow us to examine both the usage of communica-
tion behaviors and the direct impact of those behaviors on
parental outcomes. The fact that the community phy-
sicians did not fare any better than the residents under-
lines the need for Communication Quality Assurance ef-
forts on a population scale.

Our quality indicator methods are intended to allow
existing quality improvement professionals to work on
physicians’ communication rather than to assume that
physicians will continue to use skills learned during train-
ing.36,37,39 Addressing communication problems in health
care does not need to interfere with the biomedical im-
portance of screening or change medical educators’ tradi-
tional focus on emotions—but such problems may need
to be addressed if population-scale newborn screening is
to result in more good than harm.

Accepted for Publication: June 28, 2011.
Correspondence: Michael H. Farrell, MD, Center for Pa-
tient Care and Outcomes Research, 8701 Watertown
Plank Rd, Milwaukee, WI 53226-0509 (mfarrell@mcw.
ed).

Author Contributions: Dr Farrell had full access to all the
data in the study and takes responsibility for the in-
tegrity of the data and the accuracy of the data analysis.

Study concept and design: Farrell and Christopher.
Acquisi-
tion of data: Farrell. Analysis and interpretation of data:
Farrell, Speiser, Deuster, and Christopher. Drafting of the
manuscript: Farrell, Speiser, and Christopher. Critical re-
vision of the manuscript for important intellectual content:
Farrell, Deuster, and Christopher. Statistical analysis: Far-
rell. Obtained funding: Farrell and Christopher. Admin-
istrative, technical, and material support: Farrell and Speiser.
Study supervision: Farrell.

Financial Disclosure: None reported.

Funding/Support: This analysis was completed while Drs
Speiser and Deuster were medical students at the Medi-
cal College of Wisconsin. The project was funded in part
by National Institutes of Health grants K01-HL072530
and R01-HL086691 and by a training grant (T35-
HL72483-24) from the National Heart, Lung and Blood
Institute (Dr Deuster).

REFERENCES

1. Roter DL, Hall JA, Kern DE, Barker LR, Cole KA, Roca RP. Improving physicians’
interviewing skills and reducing patients’ emotional distress: a randomized clini-
2. Hampton ML, Anderson J, Lavizzo BS, Bergmen AB. Sickle cell disease 
screening programs for sickle cell disease, Tay-Sachs disease, neural tube de-
oting Safe and Effective Genetic Testing in the United States: Final Report of the
Task Force on Genetic Testing. Baltimore, MD: Johns Hopkins University Press;
7. Mischler EH, Wilford BS, Fost N, et al. Cystic fibrosis newborn screening: im-
8. Tluczek A, Kosick RL, Farrell PM, Rock MJ. Psychosocial risk associated with
newborn screening for cystic fibrosis: parents’ experience while awaiting the sweat-
results when screening for cystic fibrosis. Pediatr Pulmonol Suppl. 1991:7-
29-37.
10. Hewlett J, Waisbren SE. A review of the psychosocial effects of false-positive
results on parents and current communication practices in newborn screening. 
11. Tluczek A, Mischler EH, Farrell PM, et al. Parents’ knowledge of neonatal screen-
42. Deuster L, Donovan J, Christopher SA, Farrell MH. Quantity of jargon used by internal medicine residents in their counseling about mammography and prostate cancer screening. Poster presented at: Society of General Internal Medicine 30th Annual Meeting; April 25-28, 2007; Toronto, Canada.