Clinical question formulation has been posited to be an important skill for physicians. Sackett’s widely read textbook on evidence-based medicine provides a list of seven reasons for this, although this book acknowledges that there are no controlled trials that demonstrate the relationship between question formation and better clinical care. Two of the reasons given seem especially relevant to family physicians: well-formulated questions help us, as physicians, to focus our learning time on learning needs that are directly relevant to our patients’ clinical needs, and well-formulated questions help us communicate more clearly with our colleagues.

Our earlier study of informal consultation with specialists suggested that these advantages proposed by Sackett might be real. We found that how physicians formulate their clinical questions influences whether a specialist is willing to support a family physician’s effort to personally manage his/her patient’s clinical problem with the informal help of a specialist. We observed that when a primary care physician asked a well-formulated question, the specialist was more likely to answer the question directly and less likely to recommend that the patient be referred or transferred.

Given the purported importance of question formulation, we investigated whether the skills that residents demonstrated in posing real clinical questions improved as they progressed through their 3-year family medicine residency and acquired greater clinical knowledge and experience.

**Methods**

The clinical questions used for this study were those asked by residents in an accredited 3-year family medicine residency using an e-mail-based system. Clinical question formulation was measured on a scale of 0 to 2 by awarding 1 point each for the presence of a proposed intervention and a desired outcome. Changes in question formulation as residents progressed in their training were assessed using cross-sectional and repeated measures. The mean question quality was scored at 1.10, and there was no significant change over the 3 years. Thirteen residents used the e-mail service over their entire 3-year training period. The individual residents showed substantial differences in how well they formulated clinical questions, with their mean question quality scores ranging from 0.38 to 1.45. There was, however, no evidence that the quality of their questions changed as they progressed in their medical training.

**Findings**

The mean question quality was scored at 1.10, and there was no significant change over the 3 years. Thirteen residents used the e-mail service over their entire 3-year training period. The individual residents showed substantial differences in how well they formulated clinical questions, with their mean question quality scores ranging from 0.38 to 1.45. There was, however, no evidence that the quality of their questions changed as they progressed in their medical training.

**Conclusions**

Resident physicians asked moderately well-formulated clinical questions, but question formulation did not improve as they progressed in clinical training. Further training in formulating clinical questions may be helpful.
message. The software did not disclose the identity of the specialty consultant.

Consultation Questions

The informal consultation questions were routed through a dedicated e-mail server, allowing all the messages to be captured for this study. Specialty consultants received the e-mail consults in their usual e-mail inboxes and were able to send back their answers by using the reply function of their usual e-mail client software. During the 5-year study period, primary care physicians who were full-time community family physicians or were faculty members or residents in a family medicine training program made use of this service. The specialty consultants were medical school faculty specifically recruited for this service and participated without payment.

The clinical questions asked by the resident physicians were parsed using the PICO taxonomy. This taxonomy divides questions into four elements: those related to (1) patient’s problem, (2) a proposed intervention, (3) a comparison intervention, or (4) a desired outcome. Two researchers independently analyzed each question for the presence of the four PICO question elements. When the researchers initially disagreed on the presence of PICO elements, the question was discussed until agreement was reached. Only questions that residents asked about specific patients (ie, foreground questions) were used for this analysis. Twenty-one questions about general medical topics (ie, background questions) were excluded from this analysis.

Scoring of Questions

Questions related to the care of specific patients received 1 quality point when they identified an intervention and 1 quality point if they identified a desired outcome. Because all questions used for this analysis identified a patient/problem element, this element was not included in our analyses. The comparison intervention question element was also excluded from our analyses for two reasons. First, according to the PICO taxonomy, this element is not obligatory but should only be present “if relevant.” Second, our previous study found that the intervention and outcome PICO question elements were associated with whether the specialist supported a family physician’s effort to solve a clinical problem, but we did not find a similar association for the comparison intervention question element. Therefore, each of the residents’ clinical questions was given a quality score ranging from 0 to 2. Examples of prototypical clinical questions with quality scores of 0, 1, and 2 are shown in Table 1.

Data Analysis

Changes in quality scores of the residents’ clinical questions as they progressed through the 3-year residency training program were analyzed using two approaches. The first was cross-sectionally: mean question quality scores were compared after categorizing the question askers as first-, second-, or third-year residents. Because individual residents contributed different numbers of questions to our analyses, the estimates and hypotheses tests for these analyses are based on statistical methods that explicitly accommodate the potential for within-subject clustering. The Cochrane-Mantel-Haenszel test for association was used to compare the quality of the clinical questions across the 3 years of training while adjusting for the clustering effect, thus treating the data as categorical. Generalized Estimating Equations (GEE) were then used to compare the quality of the clinical questions across the 3 years of training, treating the data as continuous in the measure of quality while adjusting for clustering within the individual resident.

For the second approach, we identified the subset residents who asked questions using the ECS during each of their 3 years of training. Repeated measures ANOVA was used to compare the change in question formulation for the individual physicians over his/her 3 years. This second approach permitted us to assess the change in the formulation of questions over time, but, because this analysis involved only 13 of 49 (27%) residents who asked questions using the ECS during each of their 3 years of training. Repeat measures ANOVA was used to compare the change in question formulation for the individual physicians over his/her 3 years. This second approach permitted us to assess the change in the formulation of questions over time, but, because this analysis involved only 13 of 49 (27%) residents who asked questions using the ECS during each of their 3 years of training.

Table 1

Examples of Clinical Questions With Categorization Using the PICO Question Component Taxonomy

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Quality Score</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cannot determine patient’s medication and her condition</td>
<td>0</td>
<td>The patient is already on medication and she is compliant with medications and diet. What can I do with her medication?</td>
</tr>
<tr>
<td>Cannot determine intervention or outcome component</td>
<td>0</td>
<td>“I have a 72-year-old woman who has been admitted to the hospital three times in the last 2 months for decompensated CHF (EF 30% baseline). She is already on enalapril 20 mg bid and furosemide 40 mg bid. BP is 102/70 and without orthostatic symptoms. There is no evidence of ischemia, and she is compliant with medications and diet. What can I do with her medication?”</td>
</tr>
<tr>
<td>Cannot determine patient’s problem and intervention or outcome component</td>
<td>0</td>
<td>“I have a 72-year-old woman who has been admitted to the hospital three times in the last 2 months for decompensated CHF (EF 30% baseline). She is already on enalapril 20 mg bid and furosemide 40 mg bid. BP is 102/70 and without orthostatic symptoms. There is no evidence of ischemia, and she is compliant with medications and diet. Should I add spironolactone to her treatment?”</td>
</tr>
<tr>
<td>Cannot determine intervention or outcome component</td>
<td>1</td>
<td>“I have a 72-year-old woman who has been admitted to the hospital three times in the last 2 months for decompensated CHF (EF 30% baseline). She is already on enalapril 20 mg bid and furosemide 40 mg bid. BP is 102/70 and without orthostatic symptoms. There is no evidence of ischemia, and she is compliant with medications and diet. Should I add spironolactone to her treatment to decrease her frequency of admissions for CHF?”</td>
</tr>
</tbody>
</table>

BP—blood pressure
CHF—congestive heart failure
EF—ejection fraction
bid—twice a day
residents, it was only undertaken as a means of supporting the cross-sectional analysis described above. While the repeated measures ANOVA does not accommodate the within-subject correlation, in these data the within-subject correlation was low (estimated as 4%), suggesting that such correlation would have minimal effect on the results of this analysis when a general linear model procedure is used. We were also interested in whether question formulation was associated with a general measure of physician knowledge. Thus, we studied the Spearman-rank correlation between question quality scores and the American Board of Family Medicine (ABFM) In-training Exam from the third year of training for the 13 residents who used the e-mail service in each of their 3 years of training.

All analyses were performed using Stata® (StataCorp LP, College Station, Tex). All statistical tests were two-tailed, and a P value less than or equal to .05 was considered significant. Approval of the University of Iowa Institutional Review Board was obtained for this study.

Results

During the 5-year study period, 49 family medicine residents (83% of the total number of residents in the program during this time period) used the e-mail informal consultation service to ask a total of 454 questions about specific patients. Individual residents asked a median of six questions each (range=1 to 49). Of the total questions, 110 (24%) were asked by residents in their first year of training, 179 (39%) in their second year, and 165 (36%) in their third year. A total of 197 (43%) of the clinical questions focused on issues related to diagnosis, 165 (36%) on treatment, 45 (10%) on prevention, and 29 (6%) on prognosis. Eighteen (4%) questions had no identifiable clinical task focus. These 18 questions contained a description of a patient’s clinical problem and then asked a general question in the form of “What do you think?” “Any ideas?” or “What would you suggest I do?”

The mean quality of all the clinical questions was 1.10, with the mean scores of individual residents ranging from 0 to 2, signifying that residents asked questions that ranged from being unformulated to well-formulated. Additionally, the distribution of the 454 questions by quality scores in each of the 3 years of training were also nearly identical, as shown in Table 2 (P=.63). Approximately 30% of the clinical questions asked by residents in each training year were well formulated. The mean quality scores of the questions asked by first-, second-, and third-year residents were 1.02 (95% confidence interval [CI]: 0.86–1.18), 1.03 (95% CI: 0.90–1.17), and 1.06 (95% CI: 0.92–1.19) respectively, making them nearly identical in all 3 years (P=.98).

The means for each year of training from the GEE analysis are lower than the mean quality score for all the questions, because the GEE analysis takes into account the clustering of clinical questions by individual residents. Residents who asked more questions also tended to ask higher-quality score questions (P=.002) than did residents who only asked a few questions using the e-mail-based system. Thus, we did not detect any differences in the quality of the clinical questions formulated by residents in their first, second, or third years of training when the data are analyzed in a cross-sectional fashion. The change in quality scores between year of training was also minimal. The difference in the question quality scores of second- and first-year residents was .01 (95% CI: -.16–.19) and third- and first-year residents was .04 (95% CI: -.14–.21). These narrow confidence limits suggest that if any real differences existed in question formulation as a function of year of training, then these differences are likely to be very minor.

Thirteen of the residents used the e-mail service during each of their 3 years of training and asked a total of 263 questions. The mean question quality score for this subset of residents was 1.08 (95% CI: 0.93–1.24). These residents showed significant variation in their skills at constructing well-formulated clinical questions (P<.0001), with individual residents generating mean question quality scores ranging from 0.38 to 1.45. Using a repeated measures approach, we did not find evidence that the quality of questions formulated by these individual residents changed as they progressed in their clinical training (P=.36). We also did not detect an association between question formulation skills and a more general measure of medical knowledge.

<table>
<thead>
<tr>
<th>YEAR OF TRAINING</th>
<th>YEAR 1</th>
<th>YEAR 2</th>
<th>YEAR 3</th>
</tr>
</thead>
<tbody>
<tr>
<td># of residents asking questions</td>
<td>30</td>
<td>33</td>
<td>34</td>
</tr>
<tr>
<td>Total # of questions per year</td>
<td>110</td>
<td>179</td>
<td>165</td>
</tr>
<tr>
<td># (%) questions rated as 0</td>
<td>23 (21.9)</td>
<td>37 (21.7)</td>
<td>36 (21.8)</td>
</tr>
<tr>
<td># (%) questions rated as 1</td>
<td>53 (48.2)</td>
<td>87 (48.6)</td>
<td>77 (46.7)</td>
</tr>
<tr>
<td># (%) questions rated as 2</td>
<td>34 (30.9)</td>
<td>55 (30.7)</td>
<td>52 (31.5)</td>
</tr>
<tr>
<td>Mean quality (CI)</td>
<td>1.02 (.86–1.18)</td>
<td>1.03 (.90–1.17)</td>
<td>1.06 (.92–1.19)</td>
</tr>
<tr>
<td>Difference from year 1 (CI)</td>
<td>—</td>
<td>.01 (.16–.19)</td>
<td>.04 (.14–.21)</td>
</tr>
</tbody>
</table>

* The question quality was not significantly different among residents in their first, second, or third years of residency training (P>.05). Each question was awarded 1 point if it contained a proposed intervention and 1 point for a desired outcome.

CI—confidence interval
The rank correlation between the residents’ quality of question scores in their third year of residency and their ABFM In-training Exam scores was low (0.08) and not statistically significant (P=.80).

Discussion

Although we analyzed more than 450 clinical questions, we were not able to detect any change in how well our residents formulated their clinical questions as they progressed through the 3 years of our training program. We were unable to detect a change when all the questions were analyzed cross-sectionally or when only questions from a subset of the residents were analyzed using a repeated measures approach. Our study suggests that residents may not substantially improve in their formulation of clinical questions as they make the transition from medical student to independent clinician.

While there is admittedly limited objective evidence about the importance of question formulation, there are strong arguments to support its importance related to self-directed learning. To practice medicine, an experienced physician uses about 2 million pieces of information. To remain competent, every year many of these pieces need to be replaced as new information is produced. Formal medical education is clearly effective in allowing novices to rapidly acquire medical knowledge. But, despite the popularity of using formal didactic lectures to transmit new information, this venue has little influence on the practices of physicians. The other hand, formal interventions that compel physicians to be active participants in the learning process are effective. This can occur when the core learning activity of the intervention is solving authentic clinical problems. Primary care physicians engage in this type of active learning on their own when they turn to specialists for assistance in solving clinical problems at the edge of their zones of proficiency. Primary care physicians have identified this type of interaction, frequently undertaken through informal or “curbside” consultations, as being one of the most common reasons for changing their approach to practice. Learning through expert guidance is not unique to medicine and has been shown to be effective in many different fields.

The alternative to directly solving the clinical problem with the assistance of a colleague is to refer the patient to the colleague for this second physician to deal with. Our earlier study suggests that how a primary care physician formulates a clinical question in part determines whether a specialist colleague will support the efforts of the primary care physician to solve the clinical problem instead of asking to manage the patient through formal consultation or transfer of care. Prior research would predict that the second approach does not provide the same active learning environment because it does not provide the primary care physician with direct experience in solving the problem. Corroborating these findings, primary care physicians often report that the formal consultation path does not meet their ongoing learning needs.

Because question formulation skills did not improve as physicians progressed through our family medicine training program, many of our residents may have left the program without the skills needed to maximize their self-directed learning. At this point, the association between question formulation skill and future competency is conjecture, although it has been noted that graduates of a medical school (McMaster University) that uses problem-based learning (PBL) tended to be more successful in updating their knowledge than did graduates of a more-traditional, lecture-based school (University of Toronto). It is possible that those at the PBL-based school received greater training in question formulation, especially because it is the school where faculty members developed the PICO taxonomy and focused on the importance of question formulation early on.

Limitations

Our study has a number of additional limitations that must be appreciated by the educator before drawing conclusions and designing an intervention. The first is that we had only a small number of residents who used the ECS in each of their 3 years, and this gives us limited power to detect differences in question formulation. We did find that among these 13 residents there was a detectable difference in the quality of their questions, although we were unable to detect differences within individual residents as they progressed in their training. However, the small number of residents and questions means that our study has limited power.

Additionally, this research was completed at only one training program, and it is possible that family medicine residents training in other programs acquire greater skill in formulating their questions as they progress through their training.

Lastly, the questions analyzed for this study were asked using an e-mail-based informal consultation service, and the results may not apply to questions asked in other situations. But, it is also possible that the questions asked using this e-mail service are actually better formulated than those typically transmitted verbally.

Conclusions

There are studies suggesting that targeted intervention can enhance question formulation skills in the area of searching the literature for answers to clinical questions. Whether a focused intervention would enhance the question formulation skills of a physician within the venue of informal consultation is, to our knowledge, unstudied. Although our resident physi-
Physicians did not receive interactive training on question formulation, they attended several lectures about question formulation using the PICO taxonomy within the residency curriculum during the study period. Additionally, our data were also collected at a time when the PICO taxonomy was being widely disseminated through books, journal articles, and Web sites.22-24 Thus, it is likely that without more-intensive interventions, residents do not improve in their formulation of questions related to patient care.

In conclusion, our family medicine residents demonstrated wide variation in how well they formulated their clinical questions, and we found that the questions from individual resident physicians did not significantly improve as they advanced in their clinical training. Whether providing instruction on question formulation may help family physicians ask well-formulated questions when they seek out the help of a specialty colleague deserves further study, as does the relationship between question formulation skills and maintenance of medical competency.

Corresponding Author: Address correspondence to Dr Bergus, University of Iowa, Department of Family Medicine, Iowa City, IA 52245. 319-384-7566. Fax: 319-384-7822. george-bergus@uiowa.edu.

REFERENCES