When There Is Too Much to Do
How Practicing Physicians Prioritize Among Recommended Interventions

Timothy P. Hofer, MD, MS, Judith K. Zemencuk, MA, Rodney A. Hayward, MD

BACKGROUND: Recent evidence suggests that patients are receiving only 50% of recommended processes of care. It is important to understand physician priorities among recommended interventions and how these priorities are influenced both intentionally as well as unintentionally.

METHODS: A survey was mailed to all primary care physicians (PCPs) from two VA hospital networks (*N* = 289), one of which had participated in a broad, evidence-based guideline development effort 8 to 12 months earlier, and all endocrinologists nationwide in the VA (*N* = 213); response rate, 63% (*n* = 315). Using the method of paired comparisons, we assessed physician priorities among 11 clinical triggers for interventions in the management of an uncomplicated patient with type 2 diabetes.

RESULTS: Both PCPs and specialists consistently identified several high-impact clinical triggers for treatment as the highest priority interventions (hemoglobin A1c = 9.5%, diastolic blood pressure [DBP] = 95 mm Hg, low-density lipoprotein = 145 mg/dl). Several low-impact interventions that are commonly used as performance measures also received relatively high ratings. Treatments that have recently been found to be highly beneficial were often rated as being of low importance (e.g., treating when DBP = 88 mm Hg). Almost 80% of PCPs rated tight glycemic control as more important than tight DBP control, in direct contrast to clinical trial evidence. Specialists' ratings followed the same general pattern, but were more consistent with the epidemiological evidence. The PCPs at the sites that participated in the guideline intervention rated blood pressure control significantly higher.

CONCLUSION: Although several high-priority aspects of diabetes care were clearly identified, there were also notable examples of ratings that were clearly inconsistent with the epidemiological literature. Recommendations based upon more recent evidence were substantially underrated and some guidelines used as performance measures were relatively overrated. These results support the arguments that a more proactive approach is needed to facilitate rapid dissemination of new high-priority findings, and that intervention priority, and not just ease of measurement, should be considered carefully when disseminating guidelines and when selecting performance measures.

KEY WORDS: diabetes; health priorities; decision making; guideline adherence.
J GEN INTERN MED 2004;19:646–653.

Recent studies have documented that barely half of the interventions that are recommended for patients with common conditions are actually provided.[1–6] However, discussions of the "quality gap" sometimes do not acknowledge that for many common chronic diseases, such as coronary artery disease and diabetes, the number of interventions recommended by expert panels and guidelines is proliferating and threatens to overwhelm providers, not to mention potentially greatly surpassing the time and energy that patients are willing to expend on their health care. For example, Yarnall et al. determined that just to carry out the health care maintenance interventions recommended by the U.S. Preventive Services Task Force (USPSTF) for a panel of patients would take a physician close to 8 hours per day of practice, leaving no time for addressing any of the acute complaints or chronic disease management of their patient population.[7] For a routine primary care visit of 10 minutes, the average patient has 15.4 risk factors and 24.5 recommendations based on the USPSTF guidelines.[8] Because many people also have chronic diseases, the competing demands problem is actually much larger, even before considering acute injury and illness.[9–15]

Ultimately, new care models may be needed to significantly improve the number of interventions that it is possible to deliver. However, in the short run, we need to understand, and in some cases modify, the priorities that providers bring to clinical encounters so as to ensure that the most important interventions are not lost amid the blizzard of demands on patients' and providers' time and energy. There have been numerous attempts to develop prioritization schemes for health interventions, usually using economic cost-effectiveness models, as part of population-level resource allocation decisions,[16–23] perhaps most famously the controversial Oregon Basic Health Services Act.[24,25] Despite the controversy that surrounds some of these efforts, Coffield et al. argue that while all of interventions supported by evidence might be desirable, prioritization can allow for efficient step-by-step improvement in population health by devoting quality improvement resources to those interventions that have the greatest impact and value.[16] In the absence of a rational prioritization of interventions, it is argued that provider decisions are based predominantly (and suboptimally) on marketing and tradition, "adding or subtracting at the margin according to some combination of new evidence, new technologies, patient or consumer demands, inertia, and internal politics."[26] While there is an extensive literature (cited in part above) on ways to develop priorities for interventions based on literature review and experts, relatively few studies have attempted to understand how practicing physicians prioritize clinical interventions.[27–29]

Therefore, we designed a study to elicit the relative priority given to 11 diabetes interventions. We selected diabetes

because it is predominantly cared for by primary care physicians (PCPs), is commonly used to profile health plan and provider quality (such as Health Plan Employer Data and Information Set [HEDIS]), and has multiple guideline recommendations that span the spectrum from critically important and strongly evidence based to speculative and probably of low impact. We used the method of paired comparisons, an established method of ranking choices in the psychometric and economics literature, that simplifies prioritization tasks and produces more reliable estimates.[30–34] We sought to evaluate physicians' understanding and prioritization of 11 clinical interventions for specialists and for 2 groups of PCPs (one of which had received a multifaceted intervention trying to increase awareness of treatment priorities).

METHODS

Subject Population

We developed a questionnaire that assessed priorities for diabetes care interventions. The questionnaire was mailed to all PCPs in two VA Integrated Service Networks (VISNs; $n = 289$) and to all endocrinologists in the VA health care system nationwide ($n = 213$ identified in the national VA employment files). Each service network has about 1.5 million annual outpatient visits and consists of up to 7 hospitals and 23 outlying freestanding community clinics. The freestanding clinics are distributed over a wide geographic area. Two reminder postcards followed an initial mailing for those who had not returned the initial questionnaire. The final response rate was 63% ($n = 315$). The survey was conducted in early 2001.

Priority Measurement Method

The paired comparison method has been used in many research disciplines, including health care,[28,35–42] for over 100 years and is particularly useful when the differences between alternatives are not easily quantified by a single objective dimension such as weight, temperature, or size.[30] This technique has been used to examine preferences for political candidates, abstract concepts such as social stability, and sensory inputs such as color or sound quality from hearing aids.

Using this method, respondents were first provided an information scenario that defines a simple and common patient presentation. Then they were presented with a series of pairs of possible clinical triggers for interventions selected from a list of 11. For each pair they were asked to select the clinical trigger that is "more important" to intervene upon (see Fig. 1). We did not define "important" for the physicians, because we wanted them to use their own values and criteria, not one(s) determined by us. The random utility interpretation of this model posits that subjects decide between the two options by determining the underlying utilities of the two options and then selecting

the option with the higher utility.[33] The analysis is designed to recover the strength of the population's preferences from the consistency of their choices (arising from their underlying utilities) versus the random fluctuations introduced by the heterogeneity of the subjects and the difficulty of the decision in choosing between the alternatives.

There are a number of theoretical advantages of the pair comparison method over other rating methods such as category-rating scales (i.e., rating importance on a 1 to 6 scale). The memory demands are less, and it requires a simpler internal conceptual model.[35] The method makes complex judgment tasks easier and frees the judgment process as much as possible from contextual effects caused by the presence of other items.[33] Whereas category-rating scales require subjects to express a preference on an arbitrary numerical scale, something they never do in making decisions in real life, paired comparisons ask whether you think treatment A is more important than treatment B.[38] This is a more natural question than rating treatment A on a 1 to 6 scale. By presenting each intervention multiple times as part of different pairings, the calculated ratings based on multiple measurements are also more reliable than a single category rating.[32] From the results of the pair comparisons the analysis can recover a representation of an interval scale that produces the observed winners of each pairing.

Triggers for Diabetes Interventions

We assembled a series of possible clinical triggers for diabetes interventions (Table 1). By clinical trigger we mean a reason to prescribe a drug, order a test, or perform some other intervention. The triggers can be arranged into 4 groups: hypertension, hyperglycemia, lipid abnormalities, and lapses in screening. Thus, for example, one trigger would be a finding, in the patient described in the scenario in Figure 1, of a systolic blood pressure (SBP) of 150 mm Hg. This trigger represents a possible reason to prescribe a different blood pressure medication.

Measures were selected to give an assortment across the spectrum of degree of importance based upon the strength of the epidemiological literature and degree of benefit in

Table 1. Clinical Triggers for Diabetes Interventions

Clinical Group	Trigger for Intervention
Hypertension	Systolic blood pressure = 150 mm Hg
	Diastolic blood pressure = 95 mm Hg
	Diastolic blood pressure = 80 mm Hg
Hyperglycemia	A1c = 9.5%
	A1c = 8.0%
Lipid abnormalities	LDL = 145 mg/dl
	LDL = 120 mg/dl
	HDL = 30 mg/dl
	Triglycerides = 400 mg/dl
Lapses in screening	Last eye exam = 1 year ago
	Last urine test = 1 year ago

LDL, low-density lipoprotein; HDL, high-density lipoprotein.

Imagine that you have a patient who is an uncomplicated 47 year-old with a two year history of Type 2 diabetes. The patient:

- Has had no complaints or known diabetes complications;
- Is on maximum dose glypizide
- Is on HCTZ 25 mg. qd for hypertension

Assume that the measures listed below (e.g., hemoglobin A1c, LDL cholesterol, etc.) are reliable and accurate reflections of the patient's biological state and are unlikely to improve without further medical intervention (i.e., diet and exercise have been given an adequate trial).

We understand that in practice you may often intervene on both or neither of the items listed below. However, we would like you to record your view of the relative importance of each of the clinical situations listed below with regard to providing high-quality medical care to your patient. There are no "right" answers. We are only interested in your views.

EXAMPLE: Consider ranking the two situations below in terms of your priorities for intervening with medical therapy for the above-described patient.

      ___ Triglycerides = 400 mg/dl vs.
      ___ HgbA1c = 8.0%

If you feel that Triglycerides = 400 mg/dl is more important to intervene upon than HgbA1c = 8.0%, you would put a checkmark on the line for Triglycerides = 400 mg/dl.

If you feel that HgbA1c = 8.0% is more important to intervene upon, put a checkmark in front of HgbA1c = 8.0%

If you are unsure which condition is more important upon which to intervene or feel that they are very close in importance or you would be unlikely to intervene for either please mark both with a checkmark.

FIGURE 1. Information scenario.

preventing major complications. Based upon the epidemiological evidence,[43–48] we *a priori* placed the 11 clinical triggers into 4 general categories (Table 2). The first we have labeled "high priority." These included clear and substantial elevations in blood pressure, LDL-C, and HgbA1c. We felt that anyone who knows the literature well would agree that the interventions for these clinical triggers should be at the top of the heap, because there is good evidence of a large impact on major outcomes for intervening on these triggers (see Table 2).[43,45]

The second group we labeled "overappreciated" consisting of the lapses in annual screening for retinopathy and proteinuria.[45,48] While screening at some interval is of value (especially screening for retinopathy every 2 to 3 years),[48] there is relatively little evidence that annual screening has much benefit, especially for the patient described in this scenario, and even if they are beneficial the impact is certainly much smaller than those for group 1. However, interventions for these indications are easily monitored and part of most health plan "report cards."

The third group, consisting of the single indication of a diastolic blood pressure (DBP) of 88, we labeled "recent-evidence" for high-impact; we were anticipating that this measure may be underappreciated, because the evidence

Table 2. Clinical Triggers Categorized by Dimensions Hypothesized to Affect Ranking

Group	Indication	Evidence	Impact	Other
High priority	DBP = 95	Good	Large	
	LDL = 145			
	A1c = 9.5			
Overappreciated	Urine screening	Little	Small	Easily monitored
	Eye screening			
Recent evidence of high-impact	DBP 88	Good	Large	More recent trials
Intermediate importance	A1c 8.0	Good	Small	

DBP, diastolic blood pressure; LDL, low-density lipoprotein.

of a large impact comes from relatively recent clinical trials (studies occurring about 2 years before our study).[49,50] At the time of our survey, the evidence supporting glycemic control to modify microvascular risks in diabetics was almost 10 years old and heavily promoted through national organizations and guidelines, whereas the importance of interventions to prevent macrovascular disease in diabetics, particularly blood pressure control, were less heavily emphasized and a DBP of 88 had until 2 years before our study been considered "within normal limits."[43,45,49,50]

The final group, consisting of mild hyperglycemia, is classified as having good evidence of a small (for later onset diabetes) to moderate (for early onset diabetes) benefit of controlling blood sugar further,[46] and we label it as being of intermediate importance.

Study Design

The pair comparison questions were administered by mailed questionnaire. The questionnaire asked for basic demographic information about the respondents and then presented the information scenario as outlined in Figure 1. The information scenario was the same for every subject with the exception that the age of the patient was randomly assigned to be 47 or 67 across the subjects, with age of onset of diabetes specified as 2 years earlier than the patient's age (i.e., 45 and 65 years old). Interventions related to preventing microvascular outcomes (glycemic control and screening for nephropathy and retinopathy) have much smaller absolute benefit in older onset diabetics.[46,48] For example, the benefit of treating an A1c of 8% is estimated to be over 5 times more beneficial for a 45-year-old patient than for a 65-year-old patient. Thus, if clinician ratings of "importance" consider degree of patient benefit, then clinical triggers related to microvascular disease should be rated lower in older onset diabetics.

Our design included a comparison between different physician groups. Endocrinologists were sampled under the presumption that specialists would be more likely to incorporate new evidence about effective interventions into their treatment priorities.[51] A quasi-experimental design varied the exposure of the PCPs to an educational intervention. The physicians in one of the two VISNs had participated in a broad, evidence-based diabetes guideline development effort 8 to 12 months before the survey was mailed. This intervention included a 2-day conference of clinical opinion leaders from all of the clinical sites in the VISN where the evidence supporting various diabetes interventions was reviewed and a consensus was reached regarding the promulgation of treatment priorities for special attention in the care of patients within the VISN. Subsequently, over a period of several months, an organizer of the conference visited each of the clinical sites in the VISN and presented the treatment priorities and the evidence supporting the recommended interventions in a discussion format with the primary care providers at each site. While presenting precise prioritization among alternative treat-

ments was not a feature of these guidelines, there was a distinct emphasis on prevention of macrovascular disease through diagnosis and treatment of elevated blood pressure and LDL-C.

Analysis

The analysis was carried out using a hierarchical logistic regression method that explicitly models the heterogeneity between respondents to analyze the paired comparison data (see Appendix online, for further details of the analysis. Available at www.jgim.org). The age of the hypothetical patient in the information scenario was introduced as a covariate to test whether the priority of the different interventions expressed by the subjects varied by the age of the patient for whom they were considering the interventions. Results were modeled separately for each of the 3 groups of physicians and specialists and the 2 groups of PCPs. The logistic regression model produces a scale value for each of the triggers relative to the lowest rated trigger. This scale value reflects the mean priority ranking (in terms of the log odds of being selected in preference to the lowest rank trigger) for the physicians in that group. The distance between any two triggers is the log odds of the probability of selecting the higher-ranked trigger over the lower-ranked trigger and the probabilities for the comparison of any two triggers can thus be calculated from the scale values.

RESULTS

Figure 2A shows the relative preference given to the 11 indications for the entire sample of physicians. Interventions for the triggers listed at the top of the scale are preferred over those at the bottom. The distance between two triggers on the y-axis can be interpreted as a probability of selecting one trigger over the other. When choosing between two triggers to intervene upon, a clinical trigger that is 1 point higher will be selected by a provider 73% of the time over the lower clinical trigger, and one that is 2 points higher will be chosen over the other 88% of the time. Two triggers at the same level are 0 units apart and each would be selected over the other with a probability of 50%. For example, in Figure 2A, an A1c of 9.5% is 1.8 points higher than an SBP of 150 mm Hg. Thus, when choosing between the two triggers, the average physician would be expected to choose to intervene first for the elevated A1c 85% of the time over an SBP of 150 mm Hg.

The relative position of the clinical triggers for the entire sample of physicians (Fig. 2A) is consistent with some of our hypotheses about the general groupings of the indicators. The more severe levels of hypertension, hyperlipidemia, and hyperglycemia are at the top of the scale. An HgbA1c of 8.0% is in the middle of the scale. The remaining clinical triggers are clustered at the bottom, with an HDL of 30 having the lowest rank. Further, the DBP of 88 mm Hg is in the lower part of the scale, especially when looking at the PCP comparison group (Fig. 2B), supporting
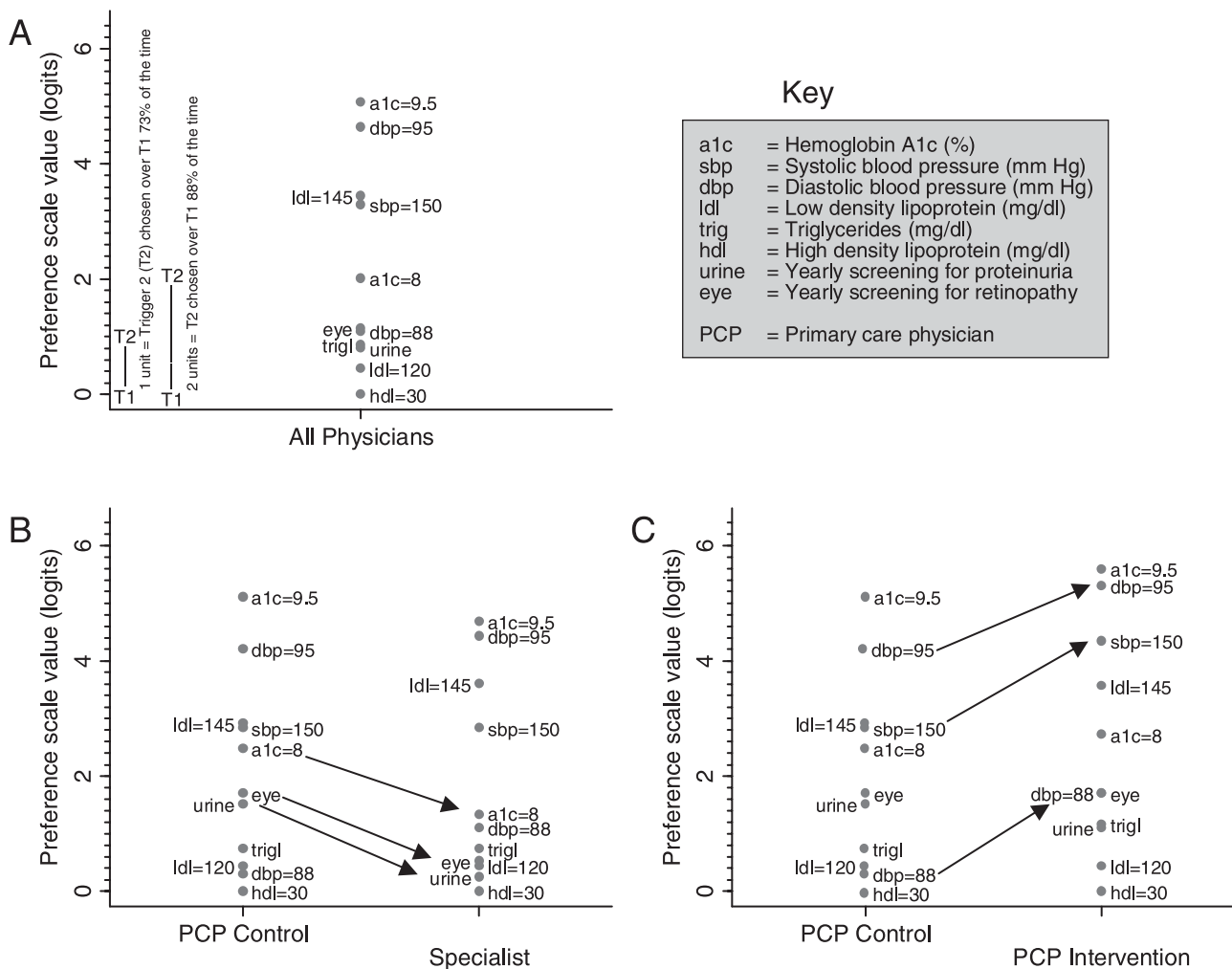
FIGURE 2. Prioritization of diabetes interventions by physician group. The higher the clinical trigger is on the y-axis, the higher its priority rating. A 1-unit higher rating denotes a 73% greater probability of being selected, and a 2-unit higher rating denotes an 88% higher probability. For example, in Figure 2A, a diastolic blood pressure of 95 mm Hg is 1.8 units higher than an A1c of 8.0% and thus 85% of the time PCPs would choose to intervene for the elevated blood pressure as compared to the Hgb A1c of 8. Figures 2B and 2C show the results broken out for endocrinologist (specialists), a group of generalist physicians who received no specific intervention (control PCPs), and a group of primary care physicians who received an intervention designed to increase attention to interventions directed at preventing macrovascular complications (intervention PCPs). Relative to the model for All Physicians, the model that estimates separate priorities for the 3 physician groups results in improved fit (*P* < .001). The arrows indicate the individual comparisons that were statistically significant.

the hypothesis that this "recent evidence" trigger would be underappreciated.

Figures 2B and 2C also show the ratings for the endocrinologist and "intervention" PCPs. There is clear evidence of significant heterogeneity in the rating scales between the 3 physician groups (likelihood ratio test $\chi^2 = 50$; degrees of freedom = 20; *P* < .001). First, there is a greater spread in the scale for the PCPs than for the specialists, particularly for the intervention PCPs. In general, the specialists seem to classify the triggers into two groups, the more severe levels of hypertension, hyperlipidemia, and hyperglycemia, and then everything else.

In terms of specific differences, the specialists rate the lapses in screening lower than the "control" PCPs (*P* < .002 for both; arrows in Fig. 2B). Interestingly, they also rate an intervention for an Hgb A1c of 8.0% much lower than the control PCPs (*P* = .01). The intervention PCPs are notable for their significantly higher ratings given to systolic hypertension (*P* = .003) and both levels of diastolic blood pressure (DBP = 88, *P* = .01; DBP = 95, *P* = .05; arrows in Fig. 2C).

Finally, the age of onset of diabetes (45 years old vs 65 years old) did not change the magnitude of any of the relative priorities given to the triggers by the 3 physician groups. This finding is quite notable given that younger patients will, on average, get dramatically more benefit from tight glycemic control.[46]

CONCLUSION

In this study, we found that several high-priority aspects of diabetes care were clearly and consistently identified by all 3 physician groups, including the more severe abnormalities of hyperglycemia, hyperlipidemia, and hypertension. However, many of their other priority selections were highly inconsistent with the epidemiological evidence. In particular, all 3 physician groups prioritized interventions for an Hgb A1c of 8.0 higher than an intervention for a DBP of 88 mm Hg. This is very difficult to justify when comparing the moderate impact on microvascular outcomes achieved by intervening on an Hgb A1c of 8.0 compared to the large impact on cardiovascular outcomes and death found for the aggressive treatment of hypertension intervention.[43,45,46,49,50,52] This disturbing result illustrates how slowly critically important clinical evidence can diffuse to even specialist physicians, as our study was conducted over 2 years after publication of the relevant clinically trials.[49,50] In addition, based on the substantial benefits found with gemfibrozil treatment of low HDL in the VA HDL Intervention Trial study, some would argue that the trigger of an HDL = 30 mg/dl should be ranked higher.[53] Even for specialists we may need to find a way to disseminate new information faster.

The relatively high priority that the PCPs gave to the eye and urine screening tests suggests another concerning phenomenon. The widespread practice of profiling providers using easily monitored, but often low-priority, aspects of clinical care may be distorting primary care provider priorities relative to those of specialists and the evidence in the epidemiological literature.[45] For example, 70% of the time, the control PCPs chose to intervene for a lapse in urine screening for proteinuria over a DBP of 88 mm Hg. There are some theoretical reasons for annual proteinuria screening, but no evidence-based argument for it substantially impacting patient outcomes.[43,54] In contrast, the Hypertension Optimal Treatment and UK Prospective Diabetes Study Group studies provided evidence that using 3 to 4 antihypertensive medications with a target DBP less than 80 mm Hg reduced cardiovascular mortality by 30% or more and also decreased visual loss and reduced cardiovascular events.[45,49] Organizations such as the National Committee for Quality Assurance (NCQA) need to consider that selecting easily measured, but low-priority, interventions as performance measures may distort clinician understanding of treatment priorities.

A group of PCPs who underwent an educational intervention 12 months prior to the survey showed priorities that differed significantly from the control PCP group. The intervention appeared to result in a de-emphasis of the lapses in screening (although intervention PCPs still showed signs of possibly being influenced by performance measurement) and an increased emphasis on blood pressure triggers, which were one of the major topics of the educational intervention. The survey materials did not refer to the educational intervention and this, along with 12-month period of time between the intervention and the survey, suggest that this observation was not merely a transient effect of the intervention. However, the quasi-experimental design of the study does not allow us to draw firm conclusions about the causal relationship between the intervention and the observed differences in priorities between the 2 primary care groups.

We found no evidence that any of the physician groups gave different rankings based on the age of onset of diabetes of the patient. This is striking given that younger onset diabetes is associated with dramatically greater benefit from interventions that decrease microvascular complications (glycemic control and screening for nephropathy and retinopathy).[46,48,55] Reducing a hemoglobin A1c from 9% to 7% is over 4 times more likely to prevent blindness (in terms of absolute risk reduction) in a patient with onset of diabetes at age 45 versus a patient with onset at age 65. In contrast, the absolute benefits of blood pressure reduction are substantial regardless of age of onset.[45] Clearly one of the challenges of disseminating information about treatment priorities is to communicate the importance of incorporating individual patient risk and benefits into treatment priorities and not just "treat the number."

Finally, this study demonstrates a method that can be used to elicit physician treatment priorities. Assessing physician treatment priorities could allow us to evaluate whether physicians have a general understanding of the clinical literature, especially regarding new findings, and could help identify important areas for educational interventions. Given that less than half of recommended interventions are actually provided,[1-6] the priority given to different interventions is important. We would hope and expect that treatment priorities that a physician brings to an encounter would be modified by the preferences and particular circumstances of each individual patient. Yet we propose that these treatment priorities, elicited for a relatively generic patient scenario such as we presented, still form the base from which individualized decisions about interventions are made and should have a significant impact on the rate at which different interventions are employed.

This method could also be used to help prioritize guideline recommendations. Formal cost-effectiveness analyses have sometimes been used in an attempt to identify clinical priorities;[16,23] however, conducting explicit cost-effectiveness analyses is time consuming and sometimes is not feasible (because of deficiencies in the available evidence). Further, when the results of cost-effectiveness analyses vary substantially with sensitivity analyses, there is no established way to simultaneously convey the importance of the estimated absolute risk reduction (if true) and our confidence that the estimate is true. In these cases, the best available standard may be for an expert panel to review the evidence and then rate treatment priorities using the type of method used in our study. Subsequently, the most critical high-consensus priorities in the clinical guidelines could be better understood and emphasized during dissemination efforts and performance measurement activities.

In conclusion, while showing that some clinical triggers used as performance measures were relatively overrated, our study also showed that even critically important triggers based on more recent evidence can be substantially underrated by both generalists and specialists. We also showed some preliminary evidence that a cooperative guideline development effort improved PCPs' awareness of the importance of several underrated diabetes interventions. It is clear that we need a more proactive approach to facilitate rapid dissemination of new high-priority findings. But it is also important to consider that selecting performance measures based on availability of data or ease of measurement may distort provider priorities in ways that are counterproductive.

REFERENCES

1. McGlynn EA, Asch SM, Adams J, et al. The quality of health care delivered to adults in the United States. N Engl J Med. 2003;348:2635–45.
2. Kessler RC, Berglund P, Demler O, et al. The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). JAMA. 2003;289:3095–105.
3. Cabana MD, Rand CS, Becher OJ, Rubin HR. Reasons for pediatrician nonadherence to asthma guidelines. Arch Pediatr Adolesc Med. 2001;155:1057–62.
4. Ruffin MT, Gorenflo DW, Woodman B. Predictors of screening for breast, cervical, colorectal, and prostatic cancer among community-based primary care practices. J Am Board Fam Pract. 2000;13:1–10.
5. Anderson LM, May DS. Has the use of cervical, breast, and colorectal cancer screening increased in the United States? Am J Public Health. 1995;85:840–2.
6. Mottur-Pilson C, Snow V, Bartlett K. Physician explanations for failing to comply with "best practices." Eff Clin Pract. 2001;4:207–13.
7. Yarnall KS, Pollak KI, Ostbye T, Krause KM, Michener JL. Primary care: is there enough time for prevention? Am J Public Health. 2003;93:635–41.
8. Medder JD, Kahn NB Jr, Susman JL. Risk factors and recommendations for 230 adult primary care patients, based on U.S. Preventive Services Task Force guidelines. Am J Prev Med. 1992;8:150–3.
9. Chernof BA, Sherman SE, Lanto AB, Lee ML, Yano EM, Rubenstein LV. Health habit counseling amidst competing demands: effects of patient health habits and visit characteristics. Med Care. 1999;37:738–47.
10. Rost K, Nutting P, Smith J, Coyne JC, Cooper-Patrick L, Rubenstein L. The role of competing demands in the treatment provided primary care patients with major depression. Arch Fam Med. 2000;9:150–4.
11. Stange KC, Zyzanski SJ, Jaen CR, et al. Illuminating the "black box." A description of 4454 patient visits to 138 family physicians. J Fam Pract. 1998;46:377–89.
12. Nutting PA, Rost K, Smith J, Werner JJ, Elliot C. Competing demands from physical problems: effect on initiating and completing depression care over 6 months. Arch Fam Med. 2000;9:1059–64.
13. Nutting PA, Baier M, Werner JJ, Cutter G, Conry C, Stewart L. Competing demands in the office visit: what influences mammography recommendations? J Am Board Fam Pract. 2001;14:352–61.
14. Kiefe CI, Funkhouser E, Fouad MN, May DS. Chronic disease as a barrier to breast and cervical cancer screening. J Gen Intern Med. 1998;13:357–65.
15. Jaen CR, Stange KC, Nutting PA. Competing demands of primary care: a model for the delivery of clinical preventive services. J Fam Pract. 1994;38:166–71.
16. Coffield AB, Maciosek MV, McGinnis JM, et al. Priorities among recommended clinical preventive services. Am J Prev Med. 2001;21:1–9.
17. Maciosek MV, Coffield AB, McGinnis JM, et al. Methods for priority setting among clinical preventive services. Am J Prev Med. 2001;21:10–9.
18. Granata AV, Hillman AL. Competing practice guidelines: using cost-effectiveness analysis to make optimal decisions. Ann Intern Med. 1998;128:56–63.
19. Hadorn DC, Holmes AC. The New Zealand priority criteria project. Part 1: overview. BMJ. 1997;314:131–4.
20. Torgerson DJ, Gosden T. Priority setting in health care: should we ask the tax payer? BMJ. 2000;320:1679.
21. Anonymous. Assessing the effectiveness of disease and injury prevention programs: costs and consequences. MMWR Recomm Rep. 1995;44(RR-10):1–10.
22. Hersh AL, Black WC, Tosteson AN. Estimating the population impact of an intervention: a decision-analytic approach. Stat Methods Med Res. 1999;8:311–30.
23. Pinkerton SD, Johnson-Masotti AP, Holtgrave DR, Farnham PG. Using cost-effectiveness league tables to compare interventions to prevent sexual transmission of HIV. AIDS. 2001;15:917–28.
24. Klevit HD, Bates AC, Castanares T, Kirk EP, Sipes-Metzler PR, Wopat R. Prioritization of health care services. A progress report by the Oregon Health Services Commission. Arch Intern Med. 1991;151:912–6.
25. Eddy DM. Clinical decision making: from theory to practice. What's going on in Oregon? JAMA. 1991;266:417–20.
26. Lawrence DM. Priorities among recommended clinical preventive services. Am J Prev Med. 2001;21:66–7.
27. Stange KC, Fedirko T, Zyzanski SJ, Jaen CR. How do family physicians prioritize delivery of multiple preventive services? J Fam Pract. 1994;38:231–7.
28. Dunham C, Mattern WD, McGaghie WC. Preferences of nephrologists among end-stage renal disease treatment options. Am J Nephrol. 1985;5:470–5.
29. Rakow T. Differences in belief about likely outcomes account for differences in doctors' treatment preferences: but what accounts for the differences in belief? Qual Health Care. 2001;10(suppl 1):i44–i9.
30. Bock RD, Thurstone LL, Jones LV. The Measurement and Prediction of Judgment and Choice. San Francisco, Calif: Holden-Day; 1968.
31. Bradley RA. Science, statistics, and paired comparisons. Biometrics. 1976;32:213–39.
32. Peterson GL, Brown TC. Economic valuation by the method of paired comparison, with emphasis on evaluation of the transitivity axiom. Land Econ. 1998;74:240–62.
33. Bockenholt U. Hierarchical modeling of paired comparison data. Psychol Methods. 2001;6:49–66.
34. Rosenberger RS, Peterson GL, Loomis JB. Applying a method of paired comparisons to measure economic values for multiple goods sets. J Agric Appl Econ. 2002;34:215–29.
35. Eisenberg LS, Dirks DD, Gornbein JA. Subjective judgments of speech clarity measured by paired comparisons and category rating. Ear Hear. 1997;18:294–306.
36. Fletcher RH, O'Malley MS, Earp JA, et al. Patients' priorities for medical care. Med Care. 1983;21:234–42.
37. Gosden T, Bowler I, Sutton M. How do general practitioners choose their practice? Preferences for practice and job characteristics. J Health Serv Res Policy. 2000;5:208–13.
38. Hadorn DC, Hays RD, Uebersax J, Hauber T. Improving task comprehension in the measurement of health state preferences. A trial of informational cartoon figures and a paired-comparison task. J Clin Epidemiol. 1992;45:233–43.
39. McKenna SP, Hunt SM, McEwen J. Weighing the seriousness of perceived health problems using Thurstone's method of paired comparisons. Int J Epidemiol. 1981;10:93–7.

40. Prieto L, Alonso J. Exploring health preferences in sociodemographic and health related groups through the paired comparison of the items of the Nottingham health profile. J Epidemiol Community Health. 2000;54:537–43.

41. Revicki DA, Shakespeare A, Kind P. Preferences for schizophrenia-related health states: a comparison of patients, caregivers and psychiatrists. Int Clin Psychopharmacol. 1996;11:101–8.

42. Coleman EA, Pennypacker H. Measuring breast self-examination proficiency. A scoring system developed from a paired comparison study. Cancer Nurs. 1991;14:211–7.

43. Vijan S, Stevens DL, Herman WH, Funnell MM, Standiford CJ. Screening, prevention, counseling, and treatment for the complications of type II diabetes mellitus. Putting evidence into practice. J Gen Intern Med. 1997;12:567–80.

44. Hayward RA. Diabetes care priorities: preventing cardiovascular complications. VA Practice Matters (volume 5, no. 1). May 2000; Available at: http://www.appc1.va.gov/resdev/prt/pm_v5_n1.pdf. Accessed September 29, 2003.

45. Vijan S, Hayward RA. Treatment of hypertension in type 2 diabetes mellitus: blood pressure goals, choice of agents, and setting priorities in diabetes care. Ann Intern Med. 2003;138:593–602.

46. Vijan S, Hofer TP, Hayward RA. Estimated benefits of glycemic control in microvascular complications in type 2 diabetes. Ann Intern Med. 1997;127:788–95.

47. VA Office of Quality and Performance. Diabetes mellitus—clinical practice guideline. November 2, 2002; Available at: http://www.oqp.med.va.gov/cpg/DM/DM_base.htm. Accessed September 29, 2003.

48. Vijan S, Hofer TP, Hayward RA. Cost-utility analysis of screening intervals for diabetic retinopathy in patients with type 2 diabetes mellitus. JAMA. 2000;283:889–96.

49. Hansson L, Zanchetti A, Carruthers SG, et al. Effects of intensive blood-pressure lowering and low-dose aspirin in patients with hypertension: principal results of the Hypertension Optimal Treatment (HOT) randomised trial. HOT Study Group. Lancet. 1998;351:1755–62.

50. UK Prospective Diabetes Study Group. Tight blood pressure control and risk of macrovascular and microvascular complications in type 2 diabetes: UKPDS 38. BMJ. 1998;317:703–13.

51. Ayanian JZ, Hauptman PJ, Guadagnoli E, Antman EM, Pashos CL, McNeil BJ. Knowledge and practices of generalist and specialist physicians regarding drug therapy for acute myocardial infarction. N Engl J Med. 1994;331:1136–42.

52. UK Prospective Diabetes Study (UKPDS) Group. Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33). Lancet. 1998;352:837–53.

53. Robins SJ, Collins D, Wittes JT, et al. Relation of gemfibrozil treatment and lipid levels with major coronary events: VA-HIT: a randomized controlled trial. JAMA. 2001;285:1585–91.

54. Berg M, Meulen RT, van den Burg M. Guidelines for appropriate care: the importance of empirical normative analysis. Health Care Anal. 2001;9:77–99.

55. Cost-effectiveness of intensive glycemic control, intensified hypertension control, and serum cholesterol level reduction for type 2 diabetes. JAMA. 2002;287:2542–51.

Appendix

Paired comparison method

This method is a two level extension of Luces' paired comparison method known as the hierarchical paired comparison model.[33] In this model the mean evaluation of the importance of an intervention for a population of subjects is estimated independently of the random effects due to the between subject variation. The data is organized so that each of the 28 pairs submitted to a subject represent an observation clustered within subject. The dependent variable is dichotomous reflecting which member of the pair was chosen. The data are analyzed with a hierarchical logistic regression model that can accommodate the incomplete pair design and accounts for the heterogeneity of the physician raters. A design matrix was coded so that we estimated ten parameters representing the mean evaluation of 10 out of the 11 triggers for interventions relative to an omitted 11th trigger. The mean evaluation of a trigger is estimated on a logit scale and represents the likelihood of choosing that trigger for an intervention relative to the omitted 11th trigger. As the trigger for a low HDL was the least likely to be chosen over any of the other triggers in the simple tables of probabilities, we used that trigger as the omitted reference category, against which the likelihood of choosing the other triggers was estimated.

With 11 clinical triggers there are a total of 110 possible pairings (n*[n−1]). In order to reduce respondent burden we randomly varied the order of presenting any two triggers across subjects. This reduced the number of possible pairs in half to 55. By employing an incomplete pair design we further reduced the number of pairs presented by half again, randomly selecting 28 pairs to present to each subject. Based on feedback from pre-testing this was an acceptable number of items to complete taking 5–7 minutes.

While the model takes into account the random differences between physician raters it does not directly estimate these variances for each trigger. With some assumptions allowing the use of identification constraints in a transformation of the estimated covariance matrix it is possible to recover a covariance matrix of the 11 triggers directly.[33] This matrix was examined but the covariance pattern did not suggest any structure that contributed further to the understanding of the observed results. The models were estimated using Laplace approximations of the maximum likelihood as for binary outcomes in HLM 5.0, there is some evidence that these estimates are more accurate than the quasi-likelihood estimation and more computationally efficient than the simulation estimations used in MlwiN.[33]

While Figure 2 contains all of the information in the model, it can be difficult to appreciate the relative differences in distance between the 55 possible pairs of triggers for each physician group and the impact of those differences on the probability of choosing one clinical trigger over another. The entire pattern can be visualized on a more natural scale by transforming the relative distances back to a probability scale (Fig. 3).

Figure 3. Presentation of scale relationships on a probability scale

In this figure, the ordering of the triggers on the x-axis is set to that of the specialists, with their priorities used as the standard against which to compare the priorities of the primary care physicians. Thus on the x-axis of Figure 3a. the triggers on the left are those with the lowest priority, meaning that they are least likely to be chosen over the triggers on the y-axis when considering a pair. This is reflected with the preponderance of the topographic colors being blues. Conversely, the higher priority triggers on the right are highly likely to be chosen over any of the other triggers as is shown by the topographic colors of red and orange.

The endocrinologists' graph reveals the grouping of the triggers into two groups, a high priority group versus the others. The dividing line is clearly shown by the 3-4 topographic lines between an Hgb A1c of 8.0 and the diastolic blood pressure of 150 mm Hg. There is more ordering evident within the high priority group than within the low priority group (illustrated by the greater density of topographic lines in the matrix in the upper right corner comparing the high priority group to itself than the lower left corner, where the lower priority group is compared to itself.)

The graph of the Primary Care Physicians priorities most obviously illustrates the "hot spots" along the ridge that follows the vertical lines for the lapses in screening triggers (urine and eye). There is evidence of these triggers being lifted in priority relative to the specialists over the whole set of triggers but it is most pronounced when they are compared to a diastolic blood pressure of 88 or an LDL of 120. The white square that appears in the upper right quadrant indicates that the PCPs find the LDL of 145 and systolic BP of 150 of equal priority whereas the specialists would prefer to treat the hyperlipidemia first 75% of the time.

The Primary Care Intervention group displays a considerable attenuation of the preference given to the lapses in screening shown by the PCP comparison group although echoes of this effect are still evident along the vertical lines extending up from the urine and eye triggers. However, the blood pressure triggers are substantially raised in priority for both the diastolic of 88 and the systolic blood
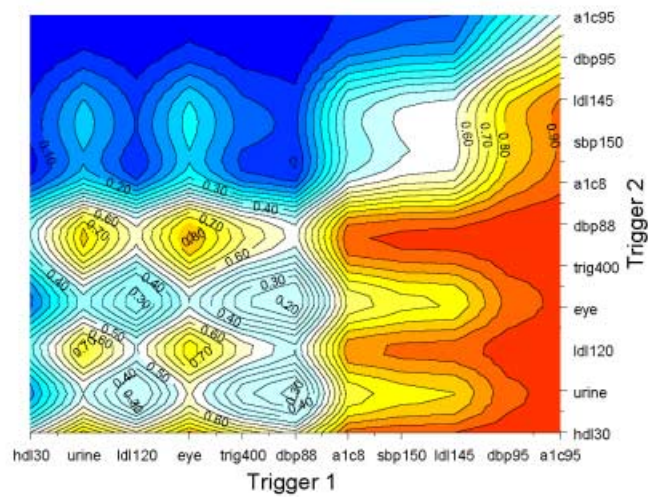
pressure of 150 as shown by the greater preponderance of yellow and red colors along the vertical lines from these triggers. The probability of choosing these blood pressure related triggers relative to the other triggers is higher than for the specialists in many cases.

Figure 3. Probability of choosing to intervene on Trigger 1 over Trigger 2

a. Endocrinologists
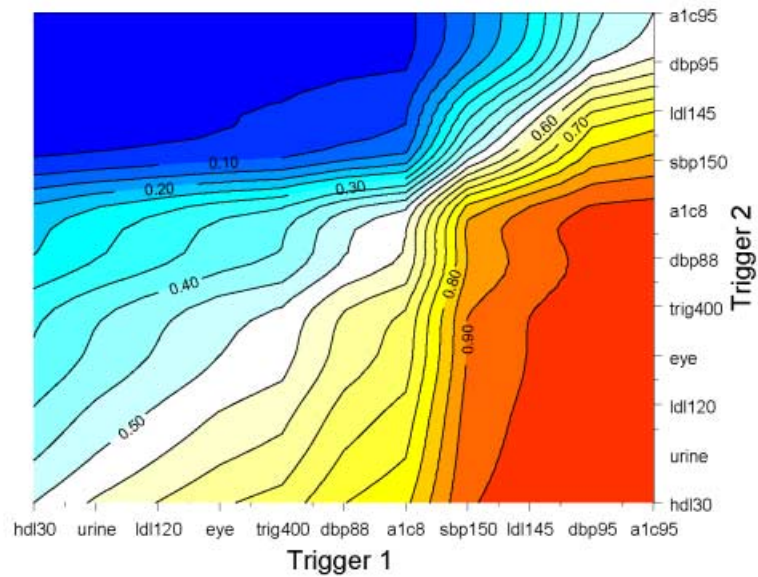
b. Primary Care Physicians

c. Primary Care Intervention