BRIEF REPORTS

Choice, Numeracy, and Physicians-in-Training Performance: The Case of Medicare Part D

Yaniv Hanoch
University of Plymouth

Talya Miron-Shatz
Princeton University and Ono Academic College

Helen Cole
Mount Sinai School of Medicine

Mary Himmelstein
Princeton University

Alex D. Federman
Mount Sinai School of Medicine

Objective: In this study, we examined the effect of choice-set size and numeracy levels on a physician-in-training's ability to choose appropriate Medicare drug plans. ***Design:*** Medical students and internal medicine residents ($N = 100$) were randomly assigned to 1 of 3 surveys, differing only in the number of plans to be evaluated (3, 10, and 20). After reviewing information about stand-alone Medicare prescription drug plans, participants answered questions about what plan they would advise 2 hypothetical patients to choose on the basis of a brief summary of the relevant concerns of each patient. Participants also completed an 11-item numeracy scale. ***Main outcome measure:*** Ability to answer correctly questions about hypothetical Medicare Part D insurance plans and numeracy levels. ***Results:*** Consistent with our hypotheses, increases in choice sets correlated significantly with fewer correct answers, and higher numeracy levels were associated with more correct answers. Hence, our data further highlight the role of numeracy in financial- and health-related decision making, and also raise concerns about physicians' ability to help patients choose the optimal Part D plan. ***Conclusion:*** Our data indicate that even physicians-in-training perform more poorly when choice size is larger, thus raising concerns about the capacity of physicians-in-training to successfully navigate Medicare Part D and help their patients pick the best drug plan. Our results also illustrate the importance of numeracy in evaluating insurance-related information and the need for enhancing numeracy skills among medical students and physicians.

Keywords: choice, decision making, Medicare Part D, numeracy

The Medicare Prescription Drug, Improvement, and Modernization Act of 2006, the most significant health care policy change in the United States since the inception of Medicare, introduced a prescription drug coverage option (Medicare Part D) for Medicare

beneficiaries. Despite its promise, the program has attracted much criticism, in part because of its complex design. Medicare beneficiaries must choose from more than 50 drug plans in most Part D markets (Kaiser Family Foundation, 2007), as well as a dizzying array of plan features, including variable monthly premiums and drug copays, different drug formularies and formulary structures, and options for gap (or "donut hole") coverage and mail-order pharmacy. It should come as no surprise that a majority of older adults (65 and older) have judged the program too difficult to navigate (Kaiser Family Foundation, 2006).

To cope with such challenges, older adults can obtain information and assistance from family members, friends, Medicare (either by phone or via the Web site), outreach and support groups like state health insurance counseling and assistance programs, media campaigns, and health care providers. Among these resources, health care workers may be particularly important. As prescribers and confidants familiar with their patients' health, financial, and social situations, physicians are often sought for guidance on prescription drug coverage decisions. Indeed, earlier studies on Medicare beneficiaries indicate a widespread expectation that physicians be well informed about the program (Mallya & Bazemore,

2006). More than 30% of beneficiaries would likely enlist their doctor's assistance—second only to Medicare sources—in deciding on Medicare Part D and physicians (Kaiser Family Foundation, 2005), and a majority believe their doctor would spend time helping them choose the best drug plan (Kaiser Family Foundation, 2005). Because physicians are often called on to assist with Part D decisions, it is vital to examine whether they can accurately process the requisite numerical information and appropriately guide their patients.

Earlier research indicates that choice-set size and numeracy skills are key factors affecting an individual's ability to wade through the complexities of health insurance options (Hanoch, Rice, Cummings, & Wood, 2009; Iyengar, Huberman, & Jiang, 2004; Iyengar & Lepper, 2000; Iyengar, Wells, & Schwartz, 2006). Unfortunately, research on physicians' numeracy levels is scant, even though numeracy is essential for practicing evidence-based medicine (Rao, 2008). What little research has been done suggests that physicians may have some difficulty interpreting probability information from mammography screens (Ancker & Kaufman, 2007; Estrada, Barnes, Collins, & Byrd, 1999; Gigerenzer, Gaissmaier, Kurz-Milcke, Schwartz, & Woloshin, 2008). These studies, however, do not indicate how well physicians interpret insurance information—a potentially less familiar topic than, say, mammography and breast cancer risk.

We therefore examined in this study whether physicians' numeracy levels affect their ability to evaluate health insurance plans and whether a larger number of plans erodes the capacity to choose appropriately. We predicted that the performance of physicians-in-training would decrease as a function of the number of drug plans presented and that numeracy levels would moderate the effect of a larger choice set.

Method

Participants

Medical students and internal medicine residents ($N = 100$: 40 third-year and 27 fourth-year students, 33 medical residents) at a large teaching hospital in the Northeast were recruited and tested between November 2008 and March 2009. The mean age was 27.2 years (range = 23–39), and 53% were men. There were no significant differences based on age, gender, numeracy, or level of medical training between randomly assigned experimental groups.

Procedure and Materials

The study was approved by the appropriate institutional review board. The instruments and design of the present study were based on an earlier investigation (Hanoch et al., 2009). A member of the research team briefly described the study and distributed surveys, the completion of which required approximately 10 min.

Respondents were randomly given one of three surveys. The surveys shared the same core questions and differed only in the number of plans to be evaluated by each participant. These items, described in greater detail below, compared three, 10, or 20 Medicare plans (see the Appendix). Thirty-four study respondents received the survey that compared three plans, 30 received the survey with 10 plans, and 36 received the survey that included 20 plans. Information about stand-alone Medicare prescription drug plans was obtained from the Medicare Web site (http://www.medicare.gov) and selected to represent a range of plan options.[1] We substituted single letters for the names of Medicare prescription plans to avoid bias associated with name recognition.[2] Respondents were presented with a table including the following information about each drug plan: total estimated annual cost (including both premiums and cost-sharing requirements), annual deductible, availability of drugs by mail order, the number of local pharmacies participating in the plan, and distance in miles from the respondent to the nearest pharmacy. After reviewing this information, respondents answered questions about what plan they would advise two hypothetical patients to choose on the basis of a brief summary of the relevant concerns of each patient.

The main outcome variable was comprehension, and the comprehension items were chosen to test respondents' ability to interpret tabular information, an essential skill for evaluating informa-

---

[1] To increase the ecological validity of the study material, we completed a drug selection process on the official Medicare Web site (using the general search option). After inputting general information (e.g., zip code and age), we entered the names of two commonly prescribed medications (Lipitor and Nexium), which provided us with a list of about 50 drug plans. From the list, we picked 20 with differing total combined annual costs (from the cheapest to the most expensive). Next, to reduce the number of plans to 10, we chose a subset from the 20 plans that offered the lowest total combined annual costs. A similar procedure was used to reduce the number of plans from 10 to three. All individuals who evaluated 20 (10 and 3) drug plans evaluated the same drug plan information. We decided to use annual costs because price is probably important for selection of drug plans. Moreover, given the complexity of the Medicare Part D program, it is impossible to capture the entire spectrum of possible plans or even to characterize the typical plan. For example, a drug plan that offers the lowest total combined annual cost for certain medications might actually be the most expensive one if other medications are required; some drug plans offer coverage during the coverage gap and some plans do not; some offer such coverage only for generic medication and some for both generic and brand names; certain plans charge no monthly premiums, whereas others do; and so forth.

[2] Our decision to use letters instead of company names followed earlier experiments that examined participants' understanding of health care information (Hibbard, Slovic, Peters, Finucane, & Tusler, 2001; Finucane, Mertz, Slovic, & Schmidt, 2005). That is, our study aimed to investigate participants' abilities to compare drug plan information (e.g., identify the drug plan that offers the lowest total annual cost) in relationship to choice size. Thus, although we acknowledge that company name can be an important factor in people's decision, we wished to keep our data free of brand familiarity, as it would not have permitted use of objective measures to evaluate participants' error rates. For example, if a participant chose on the basis of company name (e.g., Blue Cross) a drug plan that does not offer the lowest combined annual cost, it would have been difficult to determine whether the decision represented an error or simply reflected preference or familiarity. Furthermore, insurance companies might offer more than one drug plan to choose from. For instance, Blue Cross's plans may differ on deductible values, coverage during the gap, extent of coverage on brand name drugs, and so forth. Thus, choosing a drug plan on the basis of company name might reduce but not eliminate the complexity of the task. At the same time, we are fully aware that doctors and Medicare beneficiaries may use company name as a cue for drug plan decisions. Evaluating this possibility, however, would require a different study design, one that was beyond the scope of our study. Our current design, then, may have created a more complicated choice environment because it did not allow participants to rely on company name.

tion provided by Medicare. Four dichotomous comprehension variables indicated whether a respondent correctly or incorrectly identified the health plans that (a) minimized the total annual cost of prescription drugs, (b) minimized annual costs so long as drugs could be obtained by mail order, (c) included the most pharmacies, and (d) did not offer mail-order prescriptions but included the closest pharmacy. A continuous variable, total comprehension, was then derived from responses to each dichotomous comprehension variable: Respondents received 1 point for each correct answer (4 points maximum).

Finally, participants completed an 11-item numeracy scale (Lipkus, Samsa, & Rimer, 2001) assessing basic understanding of probabilities given that such knowledge is required for health insurance decision making. Each item on the scale was coded as correct or incorrect, and a score from 0 to 11 was assigned by tabulating correct responses.

Results

Numeracy

Numeracy scores ranged from 5 to 11 ($M = 10.43$, $SD = 1.00$), with 64% of our sample scoring 11. Therefore, using a median split of numeracy (Peters et al., 2006) would have been futile; we treated numeracy as a continuous variable instead.

Comprehension

Over two thirds of the sample (68%) chose the appropriate Medicare prescription drug plan across all situations. Questions involving annual cost generated the most errors: 77% chose the correct plan that minimized total annual cost, and 86% chose the appropriate plan that both minimized annual cost and allowed mail-order prescriptions (see Table 1 and Figure 1).

Between-groups differences in comprehension based on number of plans. A detailed description of results from Fisher's exact tests on comprehension are shown in Table 1. Respondents who compared either 10 or 20 plans had more trouble identifying the plan that allowed mail-order prescriptions and minimized the total annual cost (Fisher's exact $p < .00$) than those who analyzed only three plans. Similarly, respondents who evaluated just three plans did better in choosing the plan that minimized total annual cost than those who dealt with 10 or 20, with the difference marginally significant (Fisher's exact $p < .10$). Respondents assigned to only three plans also scored significantly higher on total comprehension than did respondents assigned to 10 or 20 plans, $F(2, 99) = 4.69$, $p = .01$.[3]

Predictors of comprehension. Predicting total comprehension by number of plans, numeracy, class year, and age yielded a significant model, $F(4, 93) = 34.43$, $p = .003$, with an $R^2 = .17$. Number of plans ($B = -.28$, $SE = .09$, $p = .00$) and numeracy ($B = .21$, $SE = .08$, $p = .01$) emerged as the only significant predictors. Neither level of training (third-year medical student, fourth-year medical student, or resident) nor age contributed significantly to the model, $\Delta R^2 = .02$, $F(2, 89) = 0.94$, ns.

Discussion

Can physicians-in-training successfully navigate the new Medicare Part D program? In line with an earlier study (Hanoch et al.,

2009), our results show that even physicians-in-training perform more poorly when choice size is larger. That is, even highly educated and numerate individuals found our tasks complicated. Indeed, respondents who compared 10 or 20 drug plans (as opposed to just three) were significantly less likely to identify the plan that minimized the total annual costs or the plan that both minimized the annual costs and allowed mail order. These results remain significant when we examine the respondents' ability to answer not just one of the questions but all four factual questions correctly. Our findings support earlier indications that most (92%) physicians find Medicare Part D too complicated (Kaiser Family Foundation, 2006). The program's complexities make it difficult to identify which drugs are covered by plan formularies so that physicians sometimes prescribe drugs that fall outside their patients' formularies (Tseng, Mangione, Brook, Keeler, & Dudely, 2007). The American Academy of Family Physicians' educational campaign has foreshadowed this possibility (Mallya & Bazemore, 2006), and our findings raise further concerns over physicians' ability to help older adults choose a drug plan.

These results also highlight the fact that physicians-in-training can occasionally fail simple data interpretation tasks, illustrating the important role of numeracy in evaluating insurance-related information.[4] Our data, therefore, not only fit well with earlier studies' findings, but also extend the field to a new domain. The results support the growing body of evidence showing that numeracy plays an important role in medical students' ability to interpret medical data (Sheridan & Pignone, 2002), physicians' capacity to practice evidence-based medicine (Rao, 2008), and clinicians' ability to communicate risk information to patients (Gigerenzer et al., 2008; Nelson, Reyna, Fagerlin, Lipkus, & Peters, 2008). Our data also corroborate earlier indications

---

[3] It is possible that better performance with fewer plans is driven by chance, as participants had a higher probability of guessing correctly when faced with three plans compared with 10 or 20 plans. To examine this issue, we analyzed the data according to the following equation: p(observed) $= p$(guess)p(correct|guessed) $+ (1 - p$(guess))p(correct|didn't guess). When p(correct|didn't guess) is set in the two equations (e.g., three plans and 10 plans) to equal each other, we can determine the value of p(guess). By setting the p(correct|didn't guess) to equal each other, we can ensure that difference in performance is due to differences in the likelihood of guessing and conditional on guessing performing better when the odds of guessing correctly are to the participant's advantage. Consider our finding that participants correctly choose the lowest cost plan 88% of the time when they had only three plans to choose from compared with 77% when facing 10 plans. If the 11% difference between three- and 10-plan options could be explained by facing more advantageous odds of guessing, then the probability of being correct conditional on not guessing exceeds 100%: $[0.88 - (0.11/0.33)]/(1 - 0.33) = 1.15$. Hence, our results indicate no evidence that the difference in performance between the three- and 10-plan options could be explained by the increased probability of guessing correctly given that a participant guessed. Rather, equivalent values of p(correct|didn't guess) can only be obtained if participants are much more likely to guess when facing more options. For example, if participants guess 1% of the time when choosing from among three plans and 15% of the time when choosing from among 10 plans, then the p(correct|didn't guess) in both treatments is equal to 89%.

[4] Because we did not include certain other measures (e.g., processing speed and working memory), it is difficult to conclude whether numeracy alone can explain our results; other factors could have contributed to our findings.

Table 1
Between-Groups Differences in Comprehension Based on Number of Plans

Objective	3 plans	10 plans	20 plans	Total	Fisher's exact p	ϕ
1. Plan that minimizes total annual cost					.10*	.21*
% choosing incorrectly	4	7	12	23		
% choosing correctly	30	23	24	77		
Total	34	30	36	100		
2. Plan that provides access to the most pharmacies					.33	.19
% choosing incorrectly	0	0	2	2		
% choosing correctly	34	30	34	98		
Total	34	30	36	100		
3. Plan that provides access to the closest pharmacy, without a mail-order option					.15	.21*
% choosing incorrectly	1	5	2	8		
% choosing correctly	33	25	34	92		
Total	34	30	36	100		
4. Minimize total annual cost, and allows mail-order prescriptions					.00**	.29**
% choosing incorrectly	0	7	7	14		
% choosing correctly	34	23	29	86		
Total	34	30	36	100		

* $p < .10$.   ** $p < .05$.

that a range of health care professionals (physicians, nurses, doctorate faculty, and medical students) have trouble answering numeracy questions (Estrada et al., 1999) and that higher numeracy levels correlate positively with correct interpretation of medical data (Sheridan & Pignone, 2002). Thus, we concur with previous studies arguing for the need to provide physicians with greater numeracy education. Finally, our results extend existing work into the domain of health insurance, demonstrating that numeracy levels are also associated with the ability to process health insurance information. These findings highlight the need for enhancing numeracy skills among medical students and physicians (Davis & Wolf, 2004).

Nonetheless, our study has a number of limitations. First, our sample is one of convenience and may not accurately represent the general physician population. Because our sample consisted of medical students and residents, it is possible that their knowledge and experience with Medicare do not reflect the ability of more experienced physicians. In addition, their decisions did not directly affect real patients, only imaginary ones. Thus, participants who had to review 10 or 20 insurance plans might have been less motivated to

Figure 1. Comprehension by number of plans reviewed (bars represent standard error).

carefully review the information and hence made more mistakes. Would practicing physicians, who have more experience with Medicare and are willing to devote the necessary time and effort, be better at finding the best insurance plan for their patients? Further studies are needed to address this important question. Finally, despite our efforts to provide a decision task that mimics the actual Medicare drug benefit, it was not possible to render them identical.

Limitations notwithstanding, we believe that our results are conservative by nature and that results obtained with practicing physicians would be similar. As noted earlier, the decision environment faced by respondents was far simpler than real-world situations: There was a maximum of 20 plans, rather than the 50-plus Part D plans available on the market. Moreover, complex elements of Part D, such as drug formularies and "donut holes," were not part of the study. Perhaps most important, respondents did not have to navigate through the Medicare Web site and input detailed information about their patients' prescriptions. Rather, our simplified version of the tasks is likely to have mitigated the real-world complexities. These difficulties are probably best exemplified by findings highlighting practicing physicians' troubles in prescribing drugs that appear on their patients' plan formularies (Tseng et al., 2007).

A growing number of studies have come to challenge the merit of offering more choice, arguing that a greater choice set increases task complexity. Our findings support this notion and, thus, raise concerns about the capacity of physicians-in-training to successfully navigate Medicare Part D and help their patients pick the best drug plan. Given the financial and medical ramifications involved, policymakers must recognize these difficulties and their potential harms. Regardless of the driving factor, Medicare beneficiaries would benefit from a simplified program and fewer insurance plans (Cummings, Rice, & Hanoch, 2009).
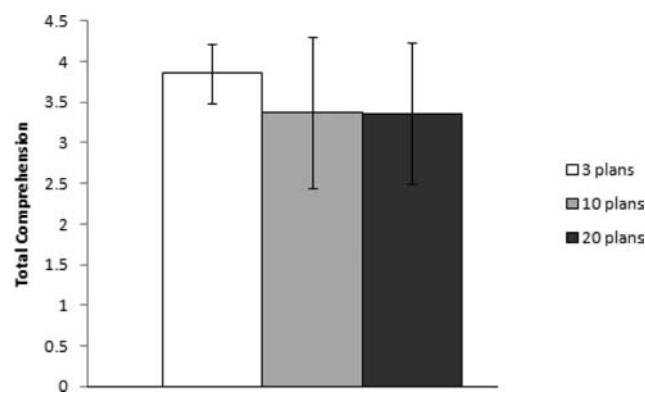
References

Ancker, J. S., & Kaufman, D. (2007). Rethinking health numeracy: A multidisciplinary literature review. *Journal of the American Medical Informatics Association, 14,* 713–721.

Cummings, J. R., Rice, T., & Hanoch, Y. (2009). Who thinks that Part D is too complicated? Survey results on the Medicare prescription drug benefit. *Medical Care Research and Review, 66,* 97–115.

Davis, T. C., & Wolf, M. S. (2004). Health literacy: Implications for family medicine. *Family Medicine, 36,* 595–598.

Estrada, C., Barnes, V., Collins, C., & Byrd, J. C. (1999). Health literacy and numeracy. *Journal of the American Medical Association, 282,* 527.

Finucane, M. L., Mertz, C. K., Slovic, P., & Schmidt, E. S. (2005). Task complexity and older adults' decision-making competence. *Journal of Psychology and Aging, 20,* 71–84.

Gigerenzer, G., Gaissmaier, W., Kurz-Milcke, E., Schwartz, L. M., & Woloshin, S. (2008). Helping doctors and patients make sense of health statistics. *Psychological Science in the Public Interest, 8,* 53–96.

Hanoch, Y., Rice, T., Cummings, J., & Wood, S. (2009). How much choice is too much? The case of the Medicare prescription drug benefit. *Health Service Research, 44,* 1157–1168.

Hibbard, J. H., Slovic, P., Peters, E., Finucane, M. L., & Tusler, M. (2001). Is the informed-choice policy approach appropriate for Medicare beneficiaries? *Health Affairs, 20,* 199–203.

Iyengar, S., Huberman, G., & Jiang, W. (2004). How much choice is too much? Contributions to 401(k) retirement plans. In O. S. Mitchell & S. Utkus (Eds.), *Pension design and structure: New lessons from behavioral finance* (pp. 83–95). Oxford, England: Oxford University Press.

Iyengar, S. S., & Lepper, M. R. (2000). When choice is demotivating: Can one desire too much of a good thing? *Journal of Personality and Social Psychology, 76,* 995–1006.

Iyengar, S. S., Wells, R. E., & Schwartz, B. (2006). Doing better but feeling worse: Looking for the "best" job undermines satisfaction. *Psychological Science, 17,* 143–150.

Kaiser Family Foundation. (2005). *The Medicare drug benefit: Beneficiary perspectives just before implementation.* Retrieved from http://www.kff.org/kaiserpolls/upload/The-Medicare-Drug-Benefit-Beneficiary-Perspectives-Just-Before-Implementation-Chartpack.pdf

Kaiser Family Foundation. (2006). *National survey of pharmacists* and *National survey of physicians.* Retrieved from http://www.kff.org/kaiserpolls/upload/7556.pdf

Kaiser Family Foundation. (2007). *Medicare Part D plan characteristics, by state, 2008: Fact sheet.* Retrieved from http://www.kff.org/medicare/upload/7426_04.pdf

Lipkus, I. M., Samsa, G., & Rimer, B. K. (2001). General performance on a numeracy scale among highly educated samples. *Medical Decision Making, 21,* 37–44.

Mallya, G., & Bazemore, A. (2006). Medicare Part D: Practical and policy implications for family physicians. *American Family Physician, 73,* 395–396.

Nelson, W., Reyna, V. F., Fagerlin, A., Lipkus, I., & Peters, E. (2008). Clinical implications of numeracy: Theory and practice. *Annals of Behavioral Medicine, 35,* 261–274.

Peters, E., Vastfjall, D., Slovic, P., Mertz, C. K., Mazzocco, K., & Dickert, S. (2006). Numeracy and decision making. *Psychological Science, 17,* 407–413.

Rao, G. (2008). Physician numeracy: Essential skills for practicing evidence-based medicine. *Family Medicine, 40,* 354–358.

Sheridan, S. L., & Pignone, M. (2002). Numeracy and the medical student's ability to interpret data. *Effective Clinical Practice, 5,* 35–40.

Tseng, C. W., Mangione, C. M., Brook, R. B., Keeler, E., & Dudely, A. (2007). Identifying widely covered drugs and drug coverage variation among Medicare Part D formularies. *Journal of the American Medical Association, 297,* 2596–2602.

Appendix

Study Material

We start by presenting you with a hypothetical situation. Please read the following carefully and answer the questions that follow using only the information you gained by reading the example.

Imagine that one of your friends, whom we'll call Bill, has asked you to help him in choosing a Medicare prescription drug plan. He has made it clear that he is only concerned about saving money, and therefore would like to enroll in the plan that minimizes his total combined annual cost for prescription drugs and the insurance for those drugs. We will provide you with a card containing information about several insurance companies covering prescription drugs. After you review this information in Card A, you will be asked several questions. Please choose the answer which you think is the best one.

We will also provide you with a card containing the definitions of various terms used throughout this experiment that can help you in answering the questions.

If you have any questions during the study, do not hesitate to ask a member of the project team.

Please answer the following questions using Card A:

1. Given Bill's desire to minimize total annual cost of prescription drugs, which one of the drug plans listed on Card A would you recommend that he choose? Please write down the Company Letter.

Company Letter _____

2. Looking at the information in Card A, which company has the largest number of pharmacies? Please write down the Company Letter.

Company Letter _____

3. Which of the companies listed does NOT offer mail order but has the closest pharmacy location? Please write down the Company Letter.

Company Letter _____

Card A

Company letter	Total combined annual cost	Mail order	Number of pharmacies	Distance to closest pharmacy (in miles)	Annual deductible	Monthly cost share
A	$4,860	Yes	14	3	$0	$140
B	$5,230	No	13	6	$0	$160
C	$4,530	Yes	20	4	$250	$250
D	$7,020	No	16	10	$250	$420
E	$5,140	Yes	20	8	$100	$210
F	$4,890	No	17	5	$0	$220
G	$4,830	No	16	9	$250	$270
H	$6,740	Yes	14	8	$0	$380
I	$4,870	Yes	18	2	$0	$200
J	$6,250	Yes	18	1	$0	$310
K	$4,410	No	12	2	$250	$240
L	$4,880	No	15	9	$250	$130
M	$4,950	No	15	6	$0	$230
N	$6,870	Yes	19	7	$100	$310
O	$4,990	No	21	8	$100	$210
P	$5,250	Yes	16	1	$100	$220
Q	$4,950	Yes	12	3	$0	$230
R	$4,800	No	16	6	$50	$170
S	$4,830	No	13	15	$0	$180
T	$6,720	Yes	15	7	$0	$420