

From the Editor

Due to the great interest surrounding the publication of two studies by Susan Horn and colleagues in the March issue of The American Journal of Managed Care, the journal has relaxed its policy of substantially shortening letters to the editor. We believe that open discussion of these studies, and the larger ramifications posited by them for healthcare policy, warrant a full and fair exchange of ideas from all sides.

TO THE EDITOR:

The study by Horn et al¹ in the March 1996 issue of *The American Journal of Managed Care* offers a rich opportunity for vested interests to attack systematic efforts designed to improve drug therapy in HMOs and other organized health care settings. While the researchers and the editorialists point out the limitations of the work, many health system managers, abetted by pharmaceutical representatives, will not get beyond the headlines of the media reports on the study. This has unfortunate implications for patient care.

The American Society of Health-System Pharmacists (ASHP) has advocated for many years that a distinction must be made between a *formulary*, as a mere drug list, and a *formulary system*, which embodies the infrastructure necessary to ensure appropriate medication use in an institutional or managed care setting. Elements of that infrastructure include an active pharmacy and therapeutics (P&T) committee, pharmacy support for the P&T committee (including thorough analysis of therapeutic and cost implications of potential committee actions), and assertive use of well-established formulary management techniques (eg, drug use evaluation, therapeutic interchange, and generic product selection). Beginning in 1960, ASHP has developed and regularly updated a number of practice standards on the operation of the formulary system²⁻⁶; we would be pleased to send these documents to any reader of your journal upon request.

Horn et al say nothing about the extent to which the study HMOs applied the wisdom reflected in formulary-system standards in promulgating their so-called formularies. In the absence of that information, one suspects that what the HMOs called formularies were probably more akin to restrictive drug lists than to formulary systems. In this light, then, the Horn study raises questions about the effectiveness of restrictive drug lists in controlling health care utilization, not about the effectiveness of formulary systems.

In addition to the study limitations pointed out by the authors and by Kravitz and Romano,⁷ it should be noted that the research did *not* study patient outcomes, only utilization of health care resources. The implicit assumption that increased utilization means poorer patient outcomes is specious.

The use of average wholesale price as a surrogate for drug product price is a more serious issue than the authors admit. Ironically, this point could help buttress the arguments against restrictive drug lists. Unfortunately, some health systems appear to be driven in their formulary decisions largely by their ability to earn price discounts related to the manufacturer's market share within a therapeutic category. Health professionals should insist that formulary decisions be based primarily on patient care considerations.

The study's method of ranking formulary limitation is a serious flaw. The method assumes that all drugs within a therapeutic category are of equal safety and effectiveness, in sharp contrast with the evidence in the clinical literature.

Many experts believe that a major determinant of the outcome of medication use is the knowledge and ability of the practitioners who manage the therapy. A good formulary system is one of the best tools available to physicians and pharmacists in health systems for bringing objective, scientific information to bear on medication-use decisions. In the interest of good patient care, we must not allow the Horn study to obscure this reality.

Joseph A. Oddis, ScD

American Society of Health-System Pharmacists
Bethesda, MD 20814

1. Horn SD, Sharkey PD, Tracy DM, et al. Intended and unintended consequences of HMO cost-containment strategies: Results from the managed care outcomes project. *Am J Man Care* 1996;2:253-264.
2. ASHP statement on the formulary system. *Am J Hosp Pharm* 1983;40:1384-1385.
3. ASHP statement on the pharmacy and therapeutics committee. *Am J Hosp Pharm* 1992;49:2008-2009.
4. ASHP guidelines on formulary system management. *Am J Hosp Pharm* 1992;49:648-652.
5. ASHP technical assistance bulletin on the evaluation of drugs for formularies. *Am J Hosp Pharm* 198;45:386-387.
6. ASHP technical assistance bulletin on drug formularies. *Am J Hosp Pharm* 1991;48:791-793.
7. Kravitz RL, Romano PS. Managed care cost containment and the law of unintended consequences. *Am J Man Care* 1996;2:323-324.

(continued on page 588)

(continued from page 586)

TO THE EDITOR:

Horn et al's¹ study of managed care patients in six HMOs concludes that restrictive formularies may be counter-productive and actually increase drug and total medical costs. In addition to the concerns of Kravitz & Romano² about the methodologies used, we have serious reservations about the formulary definition methods and the validity of the severity score used in making the case mix adjustment. These concerns are as follows:

Unit of analysis. For every one of the individuals in each of the six HMOs studied, both the cost-containment features (eg, the formulary) and the systems and/or standards of care (eg, number of days wait before an appointment, typical number of office visits for a given medical condition, etc.) are the same. Thus, the results of this study could be attributable to either the cost-containment measures studied or to the systems of care within each of the HMOs. Although Horn et al included variables representing HMO features in their regression equations, the coefficients for these variables are not shown, and the variables do not include two potentially important measures, the HMO's size and its age. The low percentages of variance explained by some of the regression equations would suggest that important variables were in fact not captured in this study.

Non-representative HMOs. As Horn et al acknowledge, the HMOs used in this study are not representative of most HMOs. We would go further and say that they are not necessarily representative of the current state of the art in managed care, which includes a wide variety of environments, practices of all sizes and specialty types, IPAs and PPOs, as well as staff models. To generalize the Horn et al findings from their very small number of staff and group model HMOs to all managed care settings is misleading.

The chicken or the egg? As Kravitz and Romano point out, it is certainly possible that higher-cost HMOs imposed

formularies to control high costs. Because the design is cross-sectional, there is no way to tell whether the higher costs or the formulary "came first." Citing the fact that the formularies measured in this study have been in place for a number of years, Horn et al conclude that the cross sectional design does not obviate their conclusions that "formularies either are causing higher utilization or are ineffective in reducing utilization to the levels of sites with fewer limitations."³ This interpretation, in particular the attribution of causality, is logically flawed and misleading. If an HMO had high utilization and costs in the past and managed to reduce them somewhat using a variety of cost-containment measures, perhaps including a formulary, their costs measured today might still be higher than those of other HMOs, but lower than they would have been had the cost containment measures not been implemented. Such a cost change pattern would go unmeasured by Horn et al's cross-sectional design. A stronger design would have examined changes in costs before and after a more "restrictive" formulary was implemented. Also potentially important is the fact that sites with higher doctor visit copays had less restrictive formularies. Is it then possible that patients with less restrictive formularies were less likely to visit their doctors (because of higher copays) and therefore, were less likely to receive and fill prescriptions?

Apples and oranges. Each patient in the Horn et al study is classified according to the medical condition for which he or she incurred the most office visits. Then all utilization, not just utilization associated with the "primary" disease, is measured. While this approach avoids the methodological pitfalls involved in attributing an office visit or hospitalization to a certain cause or condition, it makes the logical connection between formulary changes and utilization measures even more tenuous. The dependent variables used in the data analyses (eg, number of office visits) are not necessarily attributable to the diseases used to classify the patients for the analyses. It is misleading, for

(continued on page 591)

CPE PROGRAM EVALUATION

Please answer all of the following questions and return this evaluation with your Continuing Education answer sheet. Note that completion of this evaluation is not required in order to receive credit for this course; your participation is completely voluntary. Your responses will in no way affect your score; they will remain anonymous and will be held in strict confidence at all times. Thank you for your assistance.

Please circle your answers:

My pharmacy practice setting is:	Independent	Chain	Hospital	Consultant
The objectives of the lesson were achieved.	Yes	No		
The quality of presentation of the material was:	Excellent	Good	Fair	Poor
The information presented will be useful to me in my practice.	Strongly agree	Mildly agree	Mildly disagree	Strongly disagree

How long did it take you to read the material and respond to the Continuing Education questions: (Please specify the number of hours.)

Please send this evaluation, along with your answer sheet and \$5 check payable to University of Kentucky, to:

Homestudy
Continuing Pharmacy Education
University of Kentucky College of Pharmacy
465 E. High Street, Suite 204
Lexington, KY 40507-1935

(continued from page 588)

example, to say that one has analyzed visits for otitis media patients when in reality asthma patients could well be included in the otitis media group, and asthma visits could be included in the otitis media utilization. Thus, each medical condition group contains "apples and oranges," both in terms of patient mix and in terms of services utilized by each patient. It is probably for this reason that the number of office visits associated with each of the medical conditions does not vary appreciably across disease state. We recognize that the Horn et al severity adjustments include all comorbidities, but we have concerns about those adjustments as discussed in the next paragraph.

Severity adjustment concerns. Judging from the data presented in the "Study Design" article, the severity measure for the initial visit appears to have little relationship to utilization measures. For those less than 18 years of age, the first visit Ambulatory Patient Severity (APS) level has no statistically significant relationships to the number of visits, the number of prescriptions, or the number of hospitalizations per 100, although prescription costs are higher for those at APS level 4 in this age group. For those in older age groups, there are a few statistically significant relationships, but all except one of these are at a significance level of only $P < 0.05$ despite the very large sample size of 13,000. Moreover, even statistically significant findings appear to be of limited relevance for utilization analyses. For example, among those age 18 to 64 years, the prescription counts for patients with first visit APS levels of 1, 2, 3 and 4 are, respectively, 21, 20, 20 and 21, a difference which is statistically significant at $P < 0.0001$ but which has no practical import. In addition, among those age 18 to 64 years, hospitalizations per 100 for APS Levels 1 - 4 are, respectively, 8.0, 5.9, 6.7 and 9.8. This inconsistent pattern for hospitalization rates, one of the most important components of medical cost, is observed in all three age groups. Finally, the summed APS measure (severity sum) is in part a function of the number of patient visits, because an office visit must take place for severity to be measured by the APS. Thus, including the severity sum in an equation to predict office visits appears tautological. Notably, the only regression equations with a high percentage of explained variance are those predicting number of office visits.

The importance of this problem for the findings of the formulary analysis is that Horn et al have, in measuring the effects of formularies on utilization, controlled for a severity measure which appears, at least at the bivariate level for the initial visit, to have little or no meaningful relationship to utilization measures. Unfortunately, the authors do not present bivariate measures of the relationship between the severity sum and utilization measures. Similarly, the authors state that the associations between the severity sum and the utilization measures in the regression models are "highly significant and positive", but they do not present the exact value of these coefficients. Although the authors state that the summed severity score is a better

predictor of utilization than the initial visit score for which they present bivariate data, the relationship between the summed score and age is inconsistent. This result is contrary to the well documented association between increasing age and utilization of medical services, as well as the utilization measures of the study population which do increase consistently with age. Moreover, the authors indicate that they used first visit score in the regression models and obtained the same results. Finally, regression equations including not only severity sum but all the other variables, including age and gender, explain very small proportions of the variance in some of the most important cost measures, for example, only 4% to 7% of emergency room visits and only 8% to 12% of hospital admissions. These remarkably low percentages suggest that the severity scale does not do a very good job of predicting utilization.

Faulty assumptions in classifying formulary policies. Simply counting the number of chemical entities available to treat a given disease in a given HMO's formulary is overly simplistic. This method in essence mathematically assigns the same level of value to every drug used to treat a given disease, when in reality some drugs have much more utility than others. The question is not the quantity of drugs covered, but rather the efficacy and cost effectiveness of those drugs that are covered.

Relationship between severity score and standards of care. Judging from the example of the APS scoring methodology provided by the authors, it appears that the severity score is dependent in part upon laboratory results. To have those results available to the reviewer, the physician has to have ordered the laboratory tests. In a system of care which has lower intensity of diagnostic assessment, the patients will have lower scores, not because they are less sick, but because their level of sickness is not as well measured as in a system with a higher intensity of diagnostic assessment. Furthermore, although the authors note that they asked the physicians in the different HMOs to increase their attention to documentation of office visits, it is common for physicians in different systems to document with different levels of accuracy, and it is also common for physicians not to change their practice patterns in response to the requests of researchers.

Assessment of 50 asthma patients. Although the asthma patient assessment illustrates only one example, which may or may not apply across different disease states, it does provide a process measure linking the independent and dependent variables. The problem with this analysis is that the authors never tell the reader how they selected these fifty cases, nor is any quantitative information presented. If the 50 person sample was selected at random, only 10 persons on average would be included for each HMO. Thus, the statements about the restricted formulary would be based on a very small number of patients.

(continued on page 592)

(continued from page 591)

In summary, the authors have conducted a resource-intensive study in an area of great importance. We conclude that, because of the methodological concerns discussed above, a relationship between a more restrictive formulary and increased healthcare costs is not demonstrated by the findings.

Kathleen Fairman, MA
 Fred Teitelbaum, PhD
 Brenda Motheral, PhD
 Susan M. Barrow, MD
 Michael E. Thomas, PharmD
 Express Scripts, Inc.
 Maryland Heights, MO

1. Horn S, Sharkey P, Tracy D, et al. Intended and unintended consequences of HMO cost containment strategies: Results from the managed care outcomes project. *Am J Man Care* 1996;2:253-263.
2. Kravitz RL, Romano PS. Managed care cost containment and the law of unintended consequences. *Am J Man Care* 1996;2:323-324.
3. Horn SD, Sharkey PD, Gassaway J. Managed care outcomes project: Study design, baseline patient characteristics and outcomes measures. *Am J Man Care* 1996; 2:237-247.

AUTHOR'S REPLY:

We thank these letter writers for the attention they have given to our papers and their thoughtful comments. Regarding the comments of Dr. Oddis, who emphasizes the distinction between a formulary system and formulary list, we pointed out in the text that formulary-management processes varied across study sites, but they all ended up with a formulary list. As we note below in the reply to Fairman et al, we analyzed the concept of formulary limitation in three ways with similar results. Only one definition of formulary limitation was reported in the "Consequences" paper. All three definitions suggest that the number of available drugs per se is a robust predictor of the extent of service utilization. Physicians from the sites reported that it takes an average of 1 hour to address all components needed to obtain exceptions to a formulary, and since they are expected to spend between 5-7 minutes on average with a patient, taxing them with 1 hour of work to obtain an exception is a strong disincentive. It is therefore not surprising that physicians in our study requested exceptions rarely. Study physicians also reported that they are profiled on how often they go outside the formulary, a behavior that is perceived negatively.

Dr. Oddis suggests that health professionals should insist that formulary decisions be based primarily on patient care considerations. We agree, but decisions are not always

made that way today. We recommend that a trio of the patient, physician, and pharmacist determine, as quickly as possible, the treatment that gives the best clinical outcome for the least necessary cost.

The extensive comments of Fairman et al will be considered in turn. In light of their position regarding unit of analysis, we point out that in any multi-center study, patients in the same center face the same "organizational policies and constraints". Yet one often analyzes patients from different centers and uses the patient as the unit of analysis. In our analyses, we adjusted for all site variables we could measure and reported on a combination of variables that resulted in the least significant results for formulary limitation and multisource variables. Coefficients of the most significant site variables—visit co-pay, strict gatekeeper, and formulary limitation—are shown in Table 3 ("Consequences" paper). Another site variable in these equations, for-profit status, was commented on. When other site variables were included (only four site variables can be included in an equation at one time), they increased significance of formulary limitation and multisource variables. We examined the variables HMO size and HMO age and found they were not significant. We would have preferred to have had different groups of patients in each site subject to different levels of formulary restrictions, but that was not possible; every HMO that considered participating in our study had a single formulary. In the real world, one has limitations.

In regard to the HMO's studied not being representative, we would argue that the study HMOs do represent a variety of common HMO cost containment strategies, and that we were studying HMO cost containment strategies, not HMO care in general. The HMOs were not selected randomly because we needed HMOs with sophisticated information systems. Also, since we followed all patient care over the year, we needed HMOs that delivered most of their ambulatory care in a few places.

We considered a "before and after formulary implementation" design, but felt it would have been inappropriate since the drugs available pre- and post-formulary over a period of up to 12 years would have been different. Moreover, pre- vs. post-longitudinal studies can be flawed by unknown confounders of other changes in the health care system. We chose to refine the issue of "formulary vs no formulary" by looking at a range of formulary limitations. Thus, we used level of formulary limitation, not just existence of a formulary, as an independent variable. As Figures 1 and 2 in our "Consequences" paper show, utilization of resources generally increases as the level of formulary limitation increases. This relationship between degree of formulary limitation and service utilization indicates a systematic association between these two variables. In summary, we addressed the question "When one looks across multiple managed care organizations at a year's worth of actual data on the care of thousands of typical

(continued on page 595)

(continued from page 592)

patients treated by their regular doctors, is there any correlation between the amount of health care services used and cost-containment efforts by the HMO?" Thus, we selected a prospective observational design, rather than controlling variability. Such an approach, used frequently in health economics, permits study of complex, real-world systems.

We recognized, as did Fairman et al, that visit co-pay could influence utilization. Therefore, we included it in our regression equations. Our results show that even after controlling for visit co-pay in the regression analyses, formulary limitation was positive and significant; our findings do not change as a result of visit co-pay amount.

Fairman et al. consider our classifications of patients and their healthcare utilizations to be an "apples and oranges" comparison. This is not correct. We examined disease-specific utilization with the outcome variables of disease-group drug count and cost (Table 3, "Consequences" paper). It is difficult to determine how to group patients with multiple diseases over a year's period of time. We performed a great many analyses that were not included in the paper in order to assess how sensitive our results were to how we grouped patients. For instance, we repeated all analyses grouping patients by the first visit study disease as well as separately grouping each patient into each study group for which they had that study disease coded for any visit throughout the year. Hence, patients could be placed into more than one study disease group in the latter analysis. Since we analyzed disease groups separately, this did not effect independence of patients. In all analyses, formulary limitation was always positively correlated with service utilization. No matter how patients were grouped, as long as they had the study disease and their prescribing was subject to limitations on drugs to treat that disease, we found greater health care utilization associated with greater formulary limitation. Although visits for any purpose are included, one would expect that, with all the "noise" this causes, the variable formulary limitation would not end up being significant. Hence, the fact that we found formulary limitation to be significantly associated with utilization suggests the effects of formulary limitation must be very large.

In regard to the concerns of Fairman et al regarding severity adjustments, we never thought that the first visit severity would be a good predictor of annual resource use, and we are not surprised that it isn't. Indeed, we are unaware of any study in which information on a single visit reliably predicts resource use for the rest of a year. However, as Table 3 ("Study Design" paper) shows, first visit severity predicts severity for the year very well. Also, for patients over 65 years, drug count, drug cost, and number of office visits are monotonically increasing with increasing first visit severity levels (Table 2, "Study Design" paper). Only severity criteria at level 2 or higher are included in the calculation of summed APS. However, if patients have more abnormal signs and symptoms, one would expect

them to visit the doctor more frequently. Hence, higher severity should result in more visits, but more visits do not drive higher severity. This is exactly what was found in the regression analyses, with highest R² values for predicting number of visits.

Further, the severity sum was always highly significant ($P < 0.0001$) and positive, and was often the most significant variable in predicting resource use. Because this was an expected finding, ie, sicker patients use more health care resources, the reviewers suggested we not show these coefficients. These data are available by request.

For our data, severity sum is not monotone increasing with increasing age group, but this is expected. First, it is well known that younger as well as older patients are sicker, ie, that the relationship is U-shaped, not monotone. Second, there is a small difference in average severity sum scores: for the 0-17 years group the sum is 55, compared to a sum of 50 for the 18-64 years group. Third, this difference is expected since younger patients are more likely to be in the asthma study group, which has the highest average severity of all study diseases (Table 2, "Study Design" paper). Thus, contrary to the assertion by Fairman et al, the results are as expected when one interprets the data carefully.

Moreover, when first visit severity was used in regression equations, effects of formulary limitation remained positive and became more significant. This is logical if limiting drugs could result in patients getting sicker or staying sicker longer. The severity sum variable picked this up as a sicker patient, not differentiating that it was drug limitation that caused extra sickness, and took that variation out from the resource use dependent variables, leaving less variation to be explained by a formulary limitation variable. When we used first visit severity, an extra amount of health care utilization from increased severity sum due to drug limitations was left to be explained by a formulary limitation variable, and hence, it was more significant. Our statement that we found the same results when we used either first visit severity or severity sum, means that the positive direction was the same (but significance of formulary limitation was greater with first visit severity).

A final matter in regards to severity concerns the lower explained variance for the dependent variables emergency department visits and hospital admissions. This lower variance is to be expected, for the following reasons. First, the lowest explained variance is for ED visits. Each of the HMOs had 24-hour urgent care centers, so there were very few ED visits (urgent care visits were counted with office visits). In fact, the number of ED visits was less than half the number of hospitalizations, which would not be expected if urgent care facilities were not present. Thus, severity sum and the other variables, although statistically significant, did not explain much variation in the few ED visits. Hospitalizations were also infrequent, although more frequent than ED visits, and there was greater ex-

(continued on page 596)

(continued from page 595)

planatory power using the study variables. For neither of these two outcomes was there very much variation to be explained. For the other outcome variables, there was much greater variability. In all cases, the models had highly statistically significant variables and were statistically significant overall.

Fairman et al point out the difficulties in classifying formulary policies. While some drugs may have more utility than others, this property is difficult to quantify. We would welcome suggestions from Fairman et al. of useful ways to describe formulary limitations. We tried several other definitions, including limitations at the drug category level within each disease group, and weighting each drug by its national market share. For every choice we tried, study results came out in the same direction: greater limitations were always associated with higher health care utilization.

While Fairman et al may disagree with the utility of the Ambulatory Patient Severity (APS), we stand by this methodology. The APS system takes into account all of a patient's diagnoses and/or problems presented at a visit, as well as degree of abnormality of the signs and symptoms. Hence, APS uses data far beyond those available to claims data-based severity systems that include only a limited number of diagnosis codes. The severity score in APS depends on laboratory results in a minimal way, with only common labs included. In the case cited by Fairman et al.

there are 26 APS criteria for patients with respiratory illness, only 2 are lab tests (both common tests: white blood cell count and blood gases), and one is based on x-ray findings. Even if these tests are not performed, there are 23 other criteria available to determine the patient's level of severity.

Lastly, the 50 asthma cases had either many visits and high severity sum per annum, or fewer visits and lower severity sum, so we have information that permits us to look for possible associations with greater utilization. In the future we want to assemble medical specialty teams to help determine, based on Managed Care Outcomes Project data, which treatments are associated with better outcomes, controlling for patient severity differences.

Each year, providers are increasingly challenged to control costs and account for resource utilization while maintaining quality patient care. This study provides new empirical data for debating these issues, and we welcome the debate that has begun. We invite the originators of these letters to the editor to conduct their own studies, using study designs that correct for all flaws they believe were present in our study, and report their results.

Susan D. Horn, PhD
Institute for Clinical Outcomes Research
Salt Lake City, UT

Instructions for Letters to the Editor of AJMC

Letters to the Editor of *The American Journal of Managed Care* will be considered for publication if they conform to the following guidelines:

- The letter is typewritten and double-spaced.
- The full name, address, and telephone number(s) of the corresponding author(s) are included.
- Any financial relationship with a product, service, or equipment company mentioned in the letter (or article referenced) that could pose a conflict of interest must be disclosed.

The American Journal of Managed Care reserves the right to edit the text of letters (for space limitations). Although the Editor cannot acknowledge receipt of all letters sent, notification will be given if and when letters are to be published in *The American Journal of Managed Care*. Submission of a letter to *The American Journal of Managed Care* constitutes permission for American Medical Publishing, LLC, to use it in any form.

Letters can be sent via mail or fax to:

Letters to the Editor
The American Journal of Managed Care
American Medical Publishing, LLC
322-D Englishtown Road
Old Bridge, NJ 08857
FAX: (908) 723-0172