



Published in final edited form as:

*Arch Intern Med.* 2009 July 27; 169(14): 1326–1332. doi:10.1001/archinternmed.2009.206.

## Agreement Between Drugs-to-Avoid Criteria and Expert Assessments of Problematic Prescribing

Michael A. Steinman, MD<sup>1,2</sup>, Gary E. Rosenthal, MD<sup>3</sup>, C. Seth Landefeld, MD<sup>1,2</sup>, Daniel Bertenthal, MPH<sup>1</sup>, and Peter J. Kaboli, MD<sup>3</sup>

<sup>1</sup> San Francisco VA Medical Center, San Francisco

<sup>2</sup> University of California, San Francisco

<sup>3</sup> The Center for Research in the Implementation of Innovative Strategies in Practice (CRIISP) at the Iowa City VA Healthcare System, Iowa City, IA and the Division of General Internal Medicine, Department of Internal Medicine, University of Iowa

### Abstract

**Background**—Drugs-to-avoid criteria are commonly used to evaluate prescribing quality in elders. However, few studies have evaluated the concordance between these criteria and individualized patient assessments as measures of problem prescribing.

**Methods**—We used data on 256 outpatients from the Iowa City VA Medical Center who were age 65 and older and taking 5 or more medications. After a comprehensive patient interview, a physician/pharmacist study team recommended that certain drugs be discontinued, substituted, or reduced in dose. We evaluated the degree to which drugs considered potentially inappropriate by the drugs-to-avoid criteria of Beers and Zhan were also considered problematic by the study team, and vice versa.

**Results**—In the study cohort, 256 patients were using 3678 medications. The physician/pharmacist team identified 563 drugs (15%) as problematic, the Beers criteria flagged 214 drugs (6%) as potentially inappropriate, and the Zhan criteria flagged 91 drugs (2.5%). Kappa statistics for concordance between drugs-to-avoid criteria and expert assessments were 0.10–0.14, indicating “slight” agreement between these measures. Sixty-one percent of drugs identified as potentially inappropriate by the Beers criteria and 49% of drugs flagged by the Zhan criteria were not judged problematic by the expert reviewers. Correspondence between drugs-to-avoid criteria and expert assessment varied widely across different types of drugs.

**Conclusions**—Drugs-to-avoid criteria have limited power to differentiate between drugs and patients with and without prescribing problems identified on individualized expert review. While these criteria are useful as guides for initial prescribing decisions, they are insufficiently accurate to use as stand-alone measures of prescribing quality.

---

Corresponding author & reprints: Michael Steinman, MD, San Francisco VAMC, Box 181G, 4150 Clement St, San Francisco, CA 94121, mike.steinman@ucsf.edu, (t) 415.221.4810 x3677, (f) 415.750.6641.

Presented at the annual meetings of the Society of General Internal Medicine (Toronto, Canada – April 2007) and the American Geriatrics Society (Seattle, Washington – May 2007). No papers related to this study have been published or are under review elsewhere.

The views expressed in this article are those of the authors and do not necessarily reflect the position or policy of the Department of Veterans Affairs.

**Disclosures:** None

Dr. Steinman had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis

Drs. Landefeld and Rosenthal are Senior Scholars in the VA National Quality Scholars Program. None of these sponsors had any role in the study design, methods, analyses, and interpretation, or in preparation of the manuscript and the decision to submit it for publication.

## Introduction

Drugs-to-avoid criteria are lists of drugs considered potentially inappropriate for elders due to adverse effects, limited effectiveness, or both. These criteria are commonly used as markers of prescribing problems for elders in research and the practice of quality measurement.<sup>1–6</sup> For example, the Centers for Medicare and Medicaid Services (CMS) mandates use of a version of the criteria of Beers et al. in nursing homes,<sup>7</sup> and the National Committee for Quality Assurance uses a version of the criteria of Zhan et al. to compare the quality of U.S. health plans.<sup>3</sup>

Despite the widespread use of drugs-to-avoid criteria, evidence of their validity as markers of prescribing quality for elders is mixed. The most commonly-used criteria were developed by expert panels, and there is substantial disagreement about which drugs should be included on these lists.<sup>1, 8–10</sup> For another marker of validity – the ability of these criteria to predict adverse outcomes – results of observational studies have been inconsistent.<sup>2, 11–15</sup> Interpretation of these outcomes studies is further complicated by difficulty isolating the impact of implicated drugs on clinical outcomes independent of characteristics of patients and concurrent therapies they received. Finally, other work has suggested that drugs-to-avoid criteria medications account for only a small fraction of adverse drug events.<sup>16</sup>

These mixed results highlight the need to better understand the accuracy of drugs-to-avoid criteria as markers of prescribing quality. However, there is little empiric data on the extent to which drugs considered potentially inappropriate by these criteria are in fact inappropriate when reviewed in the context of case histories of actual patients. In this study, we compared two commonly-used drugs-to-avoid criteria with individualized expert assessment of patients' medications in a cohort of over 250 elderly veterans. In doing so, we focused on whether drugs considered inappropriate by the Beers and Zhan criteria were also considered appropriate when evaluated by individualized expert review.

## Methods

### Data source

We used data from the Enhanced Pharmacy Outpatient Clinic (EPOC) trial.<sup>17, 18</sup> This randomized controlled trial evaluated the impact of a specialized medication-review clinic on prescribing and patient outcomes among older veterans in the outpatient clinics of the Iowa City VA Medical Center. Eligibility criteria for participation in the trial included age 65 or older and use of 5 or more medications.

Subjects in the intervention arm were evaluated in a medication review clinic. During the baseline visit, a study pharmacist with expertise in geriatric pharmacotherapy conducted an in-depth interview about the patient's medication use and adverse effects of their medications. The medication list generated from this interview included all drugs, supplements, and herbal preparations currently used by the patient, including prescription and non-prescription drugs from both VA and non-VA sources. Toward the end of the visit, a physician with expertise in prescribing for elders conferred with the pharmacist and patient, after which the physician-pharmacist team generated a consensus list of recommendations which was delivered to the patient's primary care physician. The physician-pharmacist team identified problems using implicit review and then categorized the problems and recommended responses using a process with substantial to excellent inter-rater reliability (kappa 0.64 to 0.85).<sup>19</sup> These recommendations included suggestions about drugs that should be discontinued, substituted with a different drug, or prescribed at different doses, as well as suggestions about initiating new drugs that may benefit the patient. For example, one interview identified that a patient taking a calcium-channel blocker had developed lower extremity edema, and the team

recommended that the drug be stopped. In another case, an older patient taking a highly anticholinergic drug may have reported no adverse effects and good effectiveness of the drug, in which case the team did not recommend a change in therapy. All recommendations were recorded and categorized in the study database and assigned a priority level of high, medium, or low.<sup>19</sup> Of note, the expert reviewers did not explicitly consult the Beers criteria or other such lists in making their recommendations. Nonetheless, as experts in prescribing for elders they were aware of these criteria and likely incorporated their principles into their clinical recommendations.

Other data including basic demographic information, past medical history, and health care utilization were collected from patients at the enrollment interview and through review of medical records.

Of 258 subjects enrolled in the intervention arm, the medication list and expert recommendations were not available for 2. The remaining 256 subjects comprised our study population.

## Measures

We separately evaluated 2 drugs-to-avoid criteria commonly used in research and quality assessment. We used data from the baseline medication list and past medical history to encode each patient's drugs as meeting or not meeting each of the criteria listed below.

**Beers criteria**—The Beers criteria include a list of drugs that are considered inappropriate for all elders (e.g, propoxyphene), drugs that should not to be prescribed above certain doses (e.g., ferrous sulfate > 325mg/day), and drug-disease and drug-drug combinations to avoid (e.g. anticholinergic medications in patients with bladder outflow obstruction). Each criterion on the list is classified as high or low severity. We evaluated all of the criteria specified in the most recent update (2003), including consideration of drug doses, drug-disease interactions, and drug-drug combinations.<sup>20</sup>

**Zhan criteria**—Based on the 1997 version of the Beers criteria, the Zhan criteria focus only on drugs that should generally be avoided in elders, without consideration of drug dosages, drug-disease interactions, or drug-drug combinations. The Zhan criteria categorize drugs into one of three categories: drugs that should always be avoided (e.g., meperidine), drugs that are rarely appropriate (e.g., diazepam), and drugs that are sometimes appropriate but often misused (e.g., amitriptyline).

**Expert review**—We used the study database to identify all drugs from the baseline interview that were recommended by the physician/pharmacist team to be discontinued, substituted with another drug, or prescribed at a lower dose. We considered any of these recommendations to be a prescribing problem as judged by the expert reviewer.

## Analyses

We attempted to match all potentially inappropriate medications (PIMs) identified by the Beers and Zhan criteria with recommendations made by the expert team. In cases where a PIM was recommended to be discontinued, substituted with another drug, or to have its dose lowered, we considered the expert assessment concordant with the drugs-to-avoid criteria.

All analyses were performed at the level of the drug (N=3678). In addition, we repeated our analyses at the level of the patient (N=256), whereby any positive result on the Beers criteria, Zhan criteria, or expert review would identify that subject as having “problem prescribing”. As 208 of 256 subjects had at least one drug change recommendation, for this analysis we

considered subjects to have a prescribing problem if expert review yielded a high-priority recommendation (N=166).

We approached our analyses from two complementary perspectives. First, we evaluated the concordance between drugs-to-avoid criteria and individualized expert review using kappa statistics. Kappa statistics provide a measure of agreement between separate ratings of the same construct – in this case, two methods of determining whether a drug was problematic or not problematic - beyond the agreement that would be expected by chance. However, because drugs-to-avoid criteria and individualized expert assessments are designed to measure different aspects of prescribing quality, one would not expect a high kappa even if both evaluations perfectly captured the elements of prescribing quality they attempt to measure. Thus, we employed a second approach whereby we considered the expert assessment a de facto reference standard, and compared the sensitivity and specificity of drugs-to-avoid criteria in comparison to this standard. Such expert assessments are not universally accepted as a criterion standard for defining prescribing quality, in part because reviewers may reasonably disagree in their assessments of a given patient's medications and because there is limited evidence of these reviews' impact on clinical outcomes.<sup>21–23</sup> Nonetheless, their face validity and similarity to a clinical assessment by a thoughtful clinician make them a useful comparison to improve understanding of how drugs-to-avoid criteria perform in a clinically individualized, real-life setting.<sup>21, 24, 25</sup>

This research was approved by the institutional review boards at the University of California, San Francisco and University of Iowa and by the Research and Development Committees of the San Francisco and Iowa City VA Medical Centers. The sponsors had no control over the study question, analyses, or decision to publish this manuscript.

## Results

The study sample comprised 256 patients taking 3678 drugs, including 2425 drugs available through prescription only and 1243 over-the-counter drugs including vitamins, minerals, and herbal preparations. Subjects were predominantly white and male (Table 1). Of the 3678 medications assessed by the expert physician-pharmacist team, 563 (15%) were considered problematic as reflected by a recommendation to discontinue the drug, substitute it with another drug, or reduce the dose (Table 1). The Beers criteria identified 214 of 3678 drugs (5.8%) to be potentially inappropriate, and the Zhan criteria identified 91 drugs (2.5%) as potentially inappropriate. The most common classes of drugs identified by the Beers and Zhan criteria are shown in Table 2.

Figure 1 shows the correspondence between the expert recommendations and the Beers and Zhan criteria. Kappa was 0.14 (95% CI, 0.10 to 0.18) for the Beers criteria and 0.10 (95% CI, 0.07 to 0.14) for the Zhan criteria, indicating “slight” agreement between the drugs-to-avoid criteria and individualized expert review beyond what would be expected by chance. Among 214 drugs meeting one or more of the Beers criteria, 83 (39%) were considered problematic by expert review. Among 563 drugs considered problematic by expert review, 83 (15%) were considered problematic by the Beers criteria. Results for the Zhan criteria followed a generally similar pattern: 46 of 91 drugs (51%) flagged by the Zhan criteria were deemed problematic by expert review, while 46 of 563 drugs (8%) flagged by expert review were deemed problematic by the Zhan criteria.

Expert reviewers cited a variety of reasons for recommending discontinuation, substitution, or dose reduction of the 480 drugs that they but neither the Beers nor Zhan criteria identified as problematic. Among these 480 drugs, 61 (13%) were flagged as causing actual adverse drug reactions, and an additional 111 (23%) were flagged as causing potential adverse drug

reactions. In addition, 138 drugs (29%) had problems relating to indications (e.g., drugs that lacked indications or provided suboptimal treatment for the condition of interest), 105 (22%) had problems with effectiveness (e.g. minimal or no evidence of therapeutic effectiveness), 53 (11%) had problems with inappropriate dose, schedule, or therapeutic duplication, and 12 (3%) had miscellaneous other problems. Among the 83 drugs identified as problematic by both the experts and the Beers or Zhan criteria, 49 (59%) were flagged by experts on the basis of real or potential adverse drug reactions and 22 (27%) were flagged on the basis of lacking indications or providing suboptimal treatment for the condition of interest.

The correspondence between drugs-to-avoid criteria and expert assessment varied across different types of drugs (Table 2). For example, nearly all of the tricyclic antidepressants identified as problematic by the Beers and Zhan criteria were also implicated by the expert assessment. In contrast, there was almost complete lack of overlap in assessments of muscle relaxants. Among 10 cases of cyclobenzaprine use identified by the Beers and Zhan criteria, only 1 was rated problematic by the expert team. However, the expert team recommended changes for 2 of 4 prescriptions of the muscle relaxant baclofen (which is not included in the Beers and Zhan criteria).

Our next analyses focused on results at the level of the subject (Figure 2). Overall, 136 subjects (53%) were taking at least one Beers-criteria drug and 71 subjects (28%) were taking at least one Zhan-criteria drug. Kappa was 0.14 (95% CI, 0.02 to 0.26) for the Beers criteria and 0.17 (95% CI, 0.08 to 0.25) for the Zhan criteria, indicating that most of the observed agreement between drugs-to-avoid criteria and expert review could be attributed to chance alone.

Next, we assessed the performance of drugs-to-avoid criteria in a setting where we designated the expert review as a “gold standard” for detecting prescribing problems (Table 3). In our cohort, 39% of drugs flagged by the Beers criteria and 51% of drugs flagged by the Zhan criteria were considered problematic by expert review (positive predictive value), with positive likelihood ratios of 3.5 and 5.7, respectively. When evaluated at the level of the patient, the ability of the Beers and Zhan criteria to distinguish between patients with and without prescribing problems fell further, with positive likelihood ratios of 1.3 and 2.5, respectively.

We conducted a number of sensitivity analyses in which we varied the thresholds for determining a drug to be problematic, including thresholds for the Beers criteria, the Zhan criteria, and the expert assessments. Most permutations yielded results similar to our main analyses (see Appendix). Finally, since the Beers and Zhan criteria focus principally on systemically-administered allopathic medications, we repeated our analyses after excluding 585 topical preparations, herbal medications, and multivitamins. Results of these analyses were similar to the main analyses.

## Discussion

In this study of elderly veterans, we found substantial discordance between drug quality assessments made by drugs-to-avoid criteria and individualized expert assessments. Half or more of the drugs flagged by the Beers and Zhan criteria were not considered problematic upon individualized, implicit expert review. Moreover, the Beers and Zhan criteria identified only 8–15% of drugs that experts judged to be problematic. Similarly discordant results were observed at the level of the patient, with limited correlation between patients taking drugs-to-avoid medications and those with prescribing problems identified on expert review.

Our finding that drugs-to-avoid criteria detected only a small fraction of prescribing problems found on individualized expert review is not surprising. Drugs-to-avoid criteria are not intended to identify all problematic drugs, but to have high specificity and high positive predictive value – that is, to focus on a limited number of drugs for which consensus indicates that use is often

(or almost always) inappropriate.<sup>1, 14</sup> However, our findings suggest suboptimal accuracy of the Beers and Zhan criteria even for this limited goal. Half or more of the drugs identified as problematic by the Beers and Zhan criteria were not judged as problematic by the expert reviewers. Although the developers of these criteria were careful to note that there may be exceptions to the judgments rendered by their criteria, these exceptions were as or more common than the rule. These findings support the claim, frequently made by physicians, that many of the drugs included in the Beers and Zhan drugs-to-avoid criteria are appropriate in selected circumstances.<sup>3, 8</sup> Of note, there is no single, universally-accepted standard for defining prescribing problems, so we can not definitively conclude that the drugs-to-avoid criteria were incorrect in every instance where they disagreed with individualized expert review. Nonetheless, to the extent that individualized drug review represents a careful, patient-oriented assessment in real-world clinical settings, our findings suggest that drugs-to-avoid criteria have limited ability to distinguish between drugs that do and do not pose a problem for patients.

In addition to their limitations in evaluating individual drugs, our findings suggest limited accuracy of drugs-to-avoid criteria when applied at the level of the patient (defined by the presence or absence of an offending drug on the patient's medication list). Concordance between the Beers criteria and expert review was only slightly above that expected by chance, with the Beers criteria having almost no ability to discriminate between subjects with and without prescribing problems defined by expert review (as reflected by likelihood ratios close to 1). The Zhan criteria had a positive likelihood ratio of 2.5, somewhat better than the Beers criteria but still reflecting weak ability to distinguish between patients with and without prescribing problems identified on expert review.

These results follow a limited body of previous work. In a small study of a homeless geriatric population, a clinical pharmacist recommended drug changes for 60% of Beers criteria drugs identified on medical record review (76% when previously discontinued drugs were excluded).<sup>26</sup> In contrast, another study done in nursing homes identified uneven and generally minimal changes in use of medications from a drugs-to-avoid list after CMS implemented a policy mandating utilization review of patients taking these drugs, suggesting that most such drugs were maintained even after individualized review.<sup>7</sup> Finally, in a previous report from the Enhanced Pharmacy Outpatient Clinic study we found low levels of inter-rater reliability between the Beers criteria and other commonly-used measures of prescribing quality, including the Medication Appropriateness Index and use of  $\geq 9$  medications (one definition of "polypharmacy").<sup>27</sup>

Notwithstanding the problems identified above, the Beers and Zhan criteria are useful when applied in a suitable context. First, these criteria may have utility for identifying prescribing problems in retrospective review of elders' medication lists.<sup>26</sup> This application shows promise insofar as it uses drug-to-avoid criteria to screen drugs for individualized review, rather than using the criteria as the final arbiter of appropriateness.<sup>7</sup> Second, drugs-to-avoid criteria may be particularly valuable when applied at the time of the prescribing decision, for example through prior physician education and/or clinical alerts integrated into electronic prescribing systems.<sup>8</sup> By definition, many of the drugs on these lists have high rates of adverse effects and/or limited efficacy, warranting caution in prescribing. Thus, many of the Beers and Zhan criteria drugs taken by patients in our study may have been suboptimal choices at the time they were initially prescribed even if they later proved to have good efficacy and few side effects for certain patients. For example, a reviewer might caution against beginning elders on diphenhydramine given its high incidence of side effects. However, if a patient with refractory pruritis had been taking diphenhydramine for one year with good symptom control and no side effects, the same reviewer would likely not have recommended the drug be stopped. As a result, the positive predictive value of the Beers and Zhan criteria may be higher when used

prospectively to avoid harmful drugs, rather than retrospectively to evaluate drugs currently in use.

While there may be clinical applications of drugs-to-avoid criteria, these criteria have increasingly been used as quality measures to assess and compare prescribing quality across providers and health systems – and in this process have often been reinterpreted not as “potentially inappropriate medications” but as “definitely inappropriate medications”.<sup>3, 28, 29</sup> Our study demonstrates substantial deficiencies when these criteria are employed for this purpose. In particular, we found that half or more of the quality “problems” identified by the criteria may in fact not have been problems. The ambiguity of quality judgments made by drugs-to-avoid criteria are further amplified when comparing care across physicians or institutions. Given that the appropriateness of these drugs may vary substantially across different clinical settings and that the number of medications a patient receives is strongly linked to the presence of Beers and Zhan criteria drugs, comparisons of prescribing quality using drugs-to-avoid criteria may be particularly challenging when patients’ clinical scenarios, level of illness burden, and medication use vary between institutions or physicians.<sup>3, 8, 18, 30</sup>

Our results should be interpreted in the context of our study design and limitations of our measures. First, subjects were recruited from a single VA medical center, and were taking a minimum of 5 medications. Second, the expert pharmacist reviews are an imperfect measure of prescribing quality, and different experts may give different assessments of prescribing appropriateness. (Of note, although this study did not conduct dual independent ratings of appropriateness for each patient, the ultimate decision about prescribing recommendations were made by consensus by an expert pharmacist and physician, thus limiting the impact of any one rater to influence the results.) Thus, our expert reviews should not be considered a criterion standard of prescribing quality, and further studies are needed to confirm our findings in different care settings and with different expert raters. Third, the recommendations generated by the study’s expert raters reflected the individual clinical circumstances of the patient. Thus, our results should be interpreted as evaluating drugs-to-avoid criteria against real-world clinical situations, rather than against more abstract notions of appropriateness.

Measuring and improving the quality of drug prescribing in older patients is essential for increasing the overall quality of health care for the elderly population. Unfortunately, drugs-to-avoid criteria performed poorly when used as quality measures to assess the current state of a patient’s drug therapy. As a result, use of these tools to judge a physician’s quality of care and to compare performance across providers and health plans may lead to erroneous conclusions. Rather, drug-to-avoid criteria are best used to warn physicians of potential problems prior to prescribing, and as a simple yet insensitive means to identify potentially inappropriate drugs for follow-up with individualized review.

## Acknowledgments

The authors thank Angela Hoth, PharmD, Mitchell Barnett, PharmD, and Sneha Patil, B.A.

### Support

Supported by the Health Services Research and Development Service, Department of Veterans Affairs through an investigator-initiated research award (SAF98-152; Dr. Rosenthal); Research Career Development awards to Dr. Steinman and to Dr. Kaboli (RCD 03-033-1 and 01-013); a Career Development Award from the National Institute on Aging (1K23AG030999) to Dr. Steinman, the Center for Research in the Implementation of Innovative Strategies in Practice (CRIISP) (HFP 04-149) at the Iowa City VA Medical Center (Drs. Rosenthal and Kaboli); Agency for Healthcare Research and Quality (AHRQ) Centers for Education and Research on Therapeutics cooperative agreement #5 U18 HSO16094; and support from the HSR&D Research Enhancement Award Program at the San Francisco VA Medical Center (Mr. Bertenthal). Additional support was provided by grants from the National Institute on Aging (AG 00912, AG 10418) and the John A. Hartford Foundation, Inc. (Dr. Landefeld).

## References

1. Zhan C, Sangl J, Bierman AS, et al. Potentially inappropriate medication use in the community-dwelling elderly: findings from the 1996 Medical Expenditure Panel Survey. *JAMA* Dec 12;2001 286(22):2823–2829. [PubMed: 11735757]
2. Jano E, Aparasu RR. Healthcare outcomes associated with beers' criteria: a systematic review. *Ann Pharmacother* Mar;2007 41(3):438–447. [PubMed: 17311835]
3. Pugh MJ, Hanlon JT, Zeber JE, Bierman A, Cornell J, Berlowitz DR. Assessing potentially inappropriate prescribing in the elderly Veterans Affairs population using the HEDIS 2006 quality measure. *J Manag Care Pharm* Sep;2006 12(7):537–545. [PubMed: 16981799]
4. Avorn J. Improving drug use in elderly patients: getting to the next level. *JAMA* Dec 12;2001 286(22):2866–2868. [PubMed: 11735764]
5. Steinman MA. Polypharmacy and the balance of medication benefits and risks. *Am J Geriatr Pharmacother* Dec;2007 5(4):314–316. [PubMed: 18179988]
6. Gurwitz JH, Rochon P. Improving the quality of medication use in elderly patients: a not-so-simple prescription. *Arch Intern Med* Aug 12–26;2002 162(15):1670–1672. [PubMed: 12153368]
7. Briesacher B, Limcangco R, Simoni-Wastila L, Doshi J, Gurwitz J. Evaluation of nationally mandated drug use reviews to improve patient safety in nursing homes: a natural experiment. *J Am Geriatr Soc* Jun;2005 53(6):991–996. [PubMed: 15935022]
8. Spinewine A, Schmadier KE, Barber N, et al. Appropriate prescribing in elderly people: how well can it be measured and optimised? *Lancet* Jul 14;2007 370(9582):173–184. [PubMed: 17630041]
9. McLeod PJ, Huang AR, Tamblyn RM, Gayton DC. Defining inappropriate practices in prescribing for elderly people: a national consensus panel. *CMAJ* Feb 1;1997 156(3):385–391. [PubMed: 9033421]
10. Beers MH, Ouslander JG, Rollingher I, Reuben DB, Brooks J, Beck JC. Explicit criteria for determining inappropriate medication use in nursing home residents. *Arch Intern Med* 1991;151(9):1825–1832. [PubMed: 1888249]
11. Fu AZ, Liu GG, Christensen DB. Inappropriate medication use and health outcomes in the elderly. *J Am Geriatr Soc* Nov;2004 52(11):1934–1939. [PubMed: 15507075]
12. Hanlon JT, Fillenbaum GG, Kuchibhatla M, et al. Impact of inappropriate drug use on mortality and functional status in representative community dwelling elders. *Med Care* Feb;2002 40(2):166–176. [PubMed: 11802089]
13. Lau DT, Kasper JD, Potter DE, Lyles A, Bennett RG. Hospitalization and death associated with potentially inappropriate medication prescriptions among elderly nursing home residents. *Arch Intern Med* Jan 10;2005 165(1):68–74. [PubMed: 15642877]
14. Fick DM, Waller JL, Maclean JR, et al. Potentially inappropriate medication use in a Medicare managed care population: association with higher drug costs and utilization. *J Manag Care Pharm* 2001;7(5):407–413.
15. Budnitz DS, Pollock DA, Weidenbach KN, Mendelsohn AB, Schroeder TJ, Annett JL. National surveillance of emergency department visits for outpatient adverse drug events. *JAMA* Oct 18;2006 296(15):1858–1866. [PubMed: 17047216]
16. Budnitz DS, Shehab N, Kegler SR, Richards CL. Medication use leading to emergency department visits for adverse drug events in older adults. *Ann Intern Med* Dec 4;2007 147(11):755–765. [PubMed: 18056659]
17. Kaboli PJ, Hoth A, Carter B, et al. The VA Enhanced Pharmacy Outpatient Clinic (EPOC) Study: A Randomized Controlled Pharmacist-Physician Intervention Trial [abstract]. *J Gen Intern Med* 2004;19(Suppl 1):114.
18. Steinman MA, Landefeld CS, Rosenthal GE, Berthenthal D, Sen S, Kaboli PJ. Polypharmacy and prescribing quality in older people. *J Am Geriatr Soc* Oct;2006 54(10):1516–1523. [PubMed: 17038068]
19. Hoth AB, Carter BL, Ness J, et al. Development and reliability testing of the clinical pharmacist recommendation taxonomy. *Pharmacotherapy* May;2007 27(5):639–646. [PubMed: 17461698]
20. Fick DM, Cooper JW, Wade WE, Waller JL, Maclean JR, Beers MH. Updating the Beers criteria for potentially inappropriate medication use in older adults: results of a US consensus panel of experts. *Arch Intern Med* Dec 8–22;2003 163(22):2716–2724. [PubMed: 14662625]



21. Hanlon JT, Lindblad CI, Gray SL. Can clinical pharmacy services have a positive impact on drug-related problems and health outcomes in community-based older adults? *Am J Geriatr Pharmacother* Mar;2004 2(1):3–13. [PubMed: 15555474]
22. Hofer TP, Bernstein SJ, DeMonner S, Hayward RA. Discussion between reviewers does not improve reliability of peer review of hospital quality. *Med Care* Feb;2000 38(2):152–161. [PubMed: 10659689]
23. Barber N, Bradley C, Barry C, Stevenson F, Britten N, Jenkins L. Measuring the appropriateness of prescribing in primary care: are current measures complete? *J Clin Pharm Ther* Dec;2005 30(6):533–539. [PubMed: 16336285]
24. Hanlon JT, Schmader KE, Ruby CM, Weinberger M. Suboptimal prescribing in older inpatients and outpatients. *J Am Geriatr Soc* 2001;49(2):200–209. [PubMed: 11207875]
25. Viktil KK, Blix HS. The impact of clinical pharmacists on drug-related problems and clinical outcomes. *Basic Clin Pharmacol Toxicol* Mar;2008 102(3):275–280. [PubMed: 18248511]
26. Spiker EC, Emptage RE, Giannamore MR, Pedersen CA. Potential adverse drug events in an indigent and homeless geriatric population. *Ann Pharmacother* Oct;2001 35(10):1166–1172. [PubMed: 11675839]
27. Steinman MA, Rosenthal GE, Landefeld CS, Bertenthal D, Sen S, Kaboli PJ. Conflicts and concordance between measures of medication prescribing quality. *Med Care* Jan;2007 45(1):95–99. [PubMed: 17279026]
28. Barnett MJ, Perry PJ, Langstaff JD, Kaboli PJ. Comparison of rates of potentially inappropriate medication use according to the Zhan criteria for VA versus private sector medicare HMOs. *J Manag Care Pharm* Jun;2006 12(5):362–370. [PubMed: 16792442]
29. Crownover BK, Unwin BK. Implementation of the Beers criteria: sticks and stones--or throw me a bone. *J Manag Care Pharm* Jun;2005 11(5):416–417. [PubMed: 15934801]
30. Aparasu RR, Mort JR. Inappropriate prescribing for the elderly: Beers criteria-based review. *Ann Pharmacother* Mar;2000 34(3):338–346. [PubMed: 10917382]

		Expert Review		
		Problem	No Problem	
Beers Criteria	Problem	83	131	214
	No Problem	480	2984	3464
		563	3115	3678

  

		Expert review		
		Problem	No Problem	
Zhan Criteria	Problem	46	45	91
	No Problem	517	3070	3587
		563	3115	3678

**Figure 1.**  
 Concordance of drugs-to-avoid criteria with individualized expert review: drug level data

		Expert Review					Expert Review		
		Problem	No Problem				Problem	No Problem	
Beers Criteria	Problem	97	39	136	Zhan Criteria	Problem	58	13	71
	No Problem	69	51	120		No Problem	108	77	185
		166	90	256			166	90	256

**Figure 2.**  
Concordance of drugs-to-avoid criteria with individualized expert review: patient level data

**Table 1**

## Characteristics of study subjects and drugs

Characteristics of study subjects	N (%), mean (SD) (N=256)
Age (years)	73.5 (+/-5.1)
Male sex	249 (97%)
White race	252 (98%)
<b>Clinical conditions</b>	
Hypertension	213 (83%)
Ischemic heart disease	138 (54%)
Diabetes mellitus	105 (41%)
Chronic obstructive pulmonary disease	91 (36%)
Depression	47 (18%)
<b>Characteristics of drugs taken by study subjects</b>	<b>(N=3678)</b>
<b>Number of drugs per subject*</b>	<b>14.4 (+/- 5.0)</b>
Prescription-only "regular" drugs	8.7 (+/- 3.8)
Prescription-only other drugs	0.8 (+/- 1.1)
Non-prescription "regular" drugs	2.3 (+/- 1.4)
Non-prescription other drugs	2.6 (+/- 2.1)
<b>Problem drugs per expert review<sup>†</sup></b>	<b>563 (15.3%)</b>
<u><b>By priority</b></u>	
High priority	308 (8.4%)
Intermediate priority	190 (5.2%)
Low priority	65 (1.8%)
<u><b>By recommendation</b></u>	
Discontinue drug	379 (10.3%)
Substitute with another drug	105 (2.9%)
Reduce dose of drug	79 (2.2%)
<b>Problem drugs per Beers criteria</b>	<b>214 (5.8%)</b>
"High severity"	156 (4.2%)
"Low severity"	58 (1.6%)
<b>Problem drugs per Zhan criteria</b>	<b>91 (2.5%)</b>
"Always avoid"	1 (<0.1%)
"Rarely appropriate"	16 (0.4%)
"Sometimes appropriate but often misused"	74 (2.0%)

\* "Regular" drugs = systemically active (e.g., oral) drugs other than vitamins, minerals, and herbal preparations

Other drugs = topical drugs, vitamins, minerals, and herbal preparations

*Non-prescription drugs* = drugs available over-the-counter, regardless of whether or not the drug was recommended or ordered by a physician

†“Problem drugs per expert review” includes drugs recommended to be discontinued, substituted, or reduced in dose.

**Table 2**

Most common drugs identified by Beers and Zhan criteria, and expert assessment of those drugs

	Beers criteria (Total N = 214)		Zhan criteria (Total N = 91)	
	Number of drugs meeting criteria	N (%) of drugs that met criteria that were deemed problematic by expert team	Number of drugs meeting criteria	N (%) of drugs that met criteria that were deemed problematic by expert team
Antihistamines	42	21 (50%)	42	21 (50%)
Tricyclic antidepressants	18	15 (83%)	15	14 (93%)
Benzodiazepines	15	5 (33%)	2	1 (50%)
Digoxin *	15	10 (67%)	N/A	N/A
Urinary antispasmodics	12	6 (50%)	10	5 (50%)
Pseudoephedrine	12	6 (50%)	N/A	N/A
Skeletal muscle relaxants	10	1 (10%)	10	1 (10%)
Ticlopidine and clopidogrel	6	1 (17%)	4	1 (25%)
Propoxyphene	4	1 (25%)	4	1 (25%)

\* Digoxin is rated problematic by the Beers criteria if dosed at >0.125 mg/day. Expert assessment included both recommendations for drug discontinuation and dose reduction.

Note: Discrepancies between number of problem drugs identified using the Beers and Zhan criteria reflect different criteria for each drug class. For example, the Beers criteria recommend against a variety of benzodiazepines in different settings, while the Zhan criteria cover only 3 specific benzodiazepines. Cells reading "N/A" indicate drugs not covered by the Zhan criteria.

**Table 3**  
Test characteristics of drugs-to-avoid criteria compared with individualized expert review

	Sensitivity	Specificity	Positive Predictive Value	Negative Predictive Value	Likelihood Ratio (+)	Likelihood Ratio (-)
Beers Criteria	15%	96%	39%	86%	3.5	0.9
Zhan Criteria	8%	99%	51%	86%	5.7	0.9
<b>For Individual Drugs (N=3678 drugs) *</b>						
Beers Criteria	58%	57%	71%	43%	1.3	0.7
Zhan Criteria	35%	86%	82%	42%	2.5	0.8
<b>For Subjects Taking Any Beers or Zhan Criteria Drugs (N=256 subjects) *</b>						

\* In this table, individualized expert review is the reference standard against which drugs-to-avoid criteria are evaluated. Test characteristics for individual drugs correspond to the data in Figure 1, and test characteristics for subjects taking or more Beers or Zhan criteria drugs correspond to the data in Figure 2.