HHS Public Access

Author manuscript

J Am Geriatr Soc. Author manuscript; available in PMC 2017 February 01.

Published in final edited form as:

J Am Geriatr Soc. 2016 January ; 64(1): 22–30. doi:10.1111/jgs.13884.

Predictive validity of the Beers and STOPP Criteria to detect adverse drug events, hospitalizations, and emergency department visits in the United States

Joshua D. Brown, PharmD, MS^{1,3}, Lisa C. Hutchison, PharmD, MPH², Chenghui Li, PhD¹, Jacob T. Painter, PharmD, MBA, PhD¹, and Bradley C. Martin, PharmD, PhD¹

¹Division of Pharmaceutical Evaluation and Policy, College of Pharmacy, University of Arkansas for Medical Sciences

²Department of Pharmacy Practice, College of Pharmacy, University of Arkansas for Medical Sciences

³Institute for Pharmaceutical Outcomes and Policy, University of Kentucky College of Pharmacy

Abstract

Corresponding author: Bradley C. Martin, PharmD, PhD, Professor and Head, Division of Pharmaceutical Evaluation and Policy, College of Pharmacy, University of Arkansas for Medical Sciences, 4301 West Markham St #522, Little Rock, AR 72205, bmartin@uams.edu, P: (501) 603-1992, F: (501) 686-5156. Alternate Contact: Joshua D. Brown, PharmD, MS, Institute for Pharmaceutical Outcomes and Policy, University of Kentucky College of Pharmacy, 292 Pharmacy Building, 789 South Limestone, Lexington, KY 40536, P: (479) 650-8047, F: (501) 686-5156.

Meeting submission: This study was presented as a poster presentation at the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) 17th Annual European Congress, November 8–12, 2014, Amsterdam, the Netherlands.

Conflict of Interest Checklist:

Elements of Financial/Personal Conflicts	JDB		CL		LCH		JTP		BCM	
	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Employment or Affiliation	x			x		x		x		x
Grants/Funds	x			x	x			x	x	
Honoraria		x		x		x		x	x	
Speaker Forum		x		x		x		x		x
Consultant		x	x		x			x		x
Stocks		x		x	x			x		x
Royalties		x		x	x			x		x
Expert Testimony		x		x		x		x		x
Board Member		x		x		x		x		x
Patents		x		x		x		x		x
Personal Relationship						x		x		

Author Contributions: Brown: study concept and design, data analysis and interpretation, preparation and editing of manuscript. Hutchison: study concept and design, data interpretation, editing of manuscript. Li: study design, data analysis and interpretation, editing of manuscript. Painter: study concept and design, data interpretation, editing of manuscript. Martin: study concept and design, data analysis and interpretation, preparation and editing of manuscript.

OBJECTIVES—To compare the predictive validity of the 2003 Beers, 2012 Beers, and STOPP Criteria.

DESIGN—Retrospective cohort.

SETTING—Managed care administrative claims data from 2006 to 2009

PARTICIPANTS—174,275 commercially insured persons 65 and older in the United States.

MEASUREMENTS—Association between adverse drug events, emergency department (ED) visits, and hospitalization outcomes and inappropriate medications using time-varying Cox proportional hazard models. Measures of model discrimination (c-index) and hazard ratios (HR) were calculated to compare unadjusted and adjusted models for associations.

RESULTS—The prevalence of inappropriate prescribing was 34.1%, 32.2%, and 27.6% for the 2012 Beers, 2003 Beers, and the STOPP Criteria. Each criteria modestly discriminated ADEs in unadjusted analyses: STOPP Criteria (HR=2.89 [2.68–3.12]; C-index=0.607), 2012 Beers Criteria (HR=2.51 [2.33–2.70]; C-index=0.603), 2003 Beers Criteria (HR=2.65 [2.46–2.85]; C-index=0.605). Similar results were observed for ED visits and hospitalizations. Adjusted analyses increased the c-indices to between 0.65 and 0.70. The kappa for agreement between criteria was 0.80 for the 2003 and 2012 Beers Criteria, 0.58 for the 2012 Beers and STOPP Criteria, and 0.59 for the 2003 Beers and STOPP Criteria. For the three outcomes, the 2012 Beers Criteria had the highest sensitivity (61.2%–71.2%) and the lowest specificity (41.2%–70.7%) while the STOPP Criteria had the lowest sensitivity (53.8%–64.7%) but the highest specificity (47.8%–78.1%).

CONCLUSIONS—All three criteria were modestly prognostic for ADEs, EDs, and hospitalizations with the STOPP Criteria slightly outperforming both Beers Criteria. With low sensitivity, low specificity, as well as low agreement between the criteria, these criteria can be used in a complementary fashion to enhance sensitivity of detecting ADEs.

Keywords

Beers Criteria; STOPP Criteria; inappropriate prescribing; adverse drug events

INTRODUCTION

A potentially inappropriate medication (PIM) exists when the risk of adverse events due to treatment outweighs the clinical benefit.¹ PIMs are associated with adverse health and economic outcomes,^{2–11} making detection and prevention a primary goal of clinicians, payers, and policymakers. Since its development in 1991,¹² the Beers Criteria has become the most widely used and recognized explicit criteria for the detection of PIMs in older adults.^{8,13,14} The criteria were updated in 1997¹⁵, 2003¹⁶, and again in 2012 by an American Geriatrics Society (AGS) expert panel and includes drugs to always avoid, drugs to use with caution, and drug-disease interactions.¹⁷

The Screening Tool of Older Persons' Prescriptions (STOPP) Criteria is an alternative criteria developed in 2008 by a European consensus group.¹ The STOPP Criteria is organized by physiological system and includes drugs to avoid, drug-drug and drug-disease interactions, and therapeutic duplication to define PIMs. It is purported to be more effective

in a European population where many of the medications considered inappropriate by the Beers Criteria are not available.^{1,18} As a result, the STOPP Criteria has modest overlap with the 2012 Beers Criteria – 55% of the 65 criteria are not found in the 2012 Beers Criteria.¹⁹

The STOPP and 2003 Beers Criteria have been compared in European populations where the STOPP Criteria identified more PIMs and increased the odds of having a serious adverse drug event (ADE) by 85%.^{20–26} A study conducted in Spain compared the 2003 Beers and STOPP Criteria along with the updated 2012 Beers Criteria.²⁷ The PIM prevalence was 24.3%, 35.4%, and 44% for the 2003 Beers, STOPP, and 2012 Beers Criteria and the agreement between the 2012 Beers and STOPP Criteria was 0.35. That study did not compare the criteria on adverse outcomes.

Because of the lack of evidence comparing the Beers Criteria with the STOPP Criteria in a United States (US) population, a comparison of the ability of each criteria to predict relevant clinical outcomes is warranted.¹⁹ Therefore, the current study sought to compare the predictive validity of the 2003 Beers, 2012 Beers, and STOPP Criteria using three outcome measures: 1) ADEs, 2) all cause emergency department (ED) visits, and 3) all cause hospitalizations. Further, the prevalence of PIMs detected with each criteria was investigated as well as measures of agreement between the criteria.

METHODS

Data source

The study sample was selected from a 10% random sample of the proprietary Lifelink Health Plans Claims Database comprised of administrative claims from 80 Managed Care Organizations within the US. The data capture the health claims data of the elderly enrolled in health plans offering employer sponsored coverage and Medicare Advantage plans but do not capture data for persons enrolled in traditional Medicare.

Study subjects and design

We used a retrospective cohort study design. Inclusion into the cohort was based on a person being at least 65 years old and having at least 9 months of continuous medical and pharmacy coverage, including a 6 month pre-index period and a minimum 3 months of follow-up between January 1, 2006 and December 31, 2009. The index date was defined as the first day of the seventh month of continuous eligibility. Individuals were followed until the end of continuous enrollment, the end of the study period, or until an outcome event occurred. Because full medical and pharmacy claims data may not be captured, individuals with the payer identified as “Medicaid” were excluded as this group may have additional insurance or incomplete records.

Potentially inappropriate medication exposure

Exposure definitions were individually created according to the 2003¹⁶ and 2012 Beers Criteria¹⁷ and the STOPP Criteria.¹ A composite “All Criteria” exposure measure was also developed based on exposure to any of the three PIM criteria. Therapeutic duplication, present as an over-arching item in the STOPP Criteria, was excluded as this is not unique to

the elderly population and was deliberately excluded from the Beers Criteria.¹⁷ Additionally, dabigatran (2012 Beers Criteria) was not on the market during the time period of this study (2006–2009) and propoxyphene (2003 Beers Criteria) was not included because it is no longer available in the U.S. market. Otherwise, all items from each criteria were included.

Drug-only criteria were mapped using the Medi-Span Generic Product Identifier (GPI, Wolters Kluwer Health, Philadelphia, PA) classification system and the American Hospital Formulary Service Pharmacologic-Therapeutic Classification codes (AHFSCC). These hierarchical coding systems allowed for classification from the drug class, individual medications, formulations (e.g. extended release), or dosing of individual products.

Disease-dependent PIM definitions were based on International Classifications of Disease, 9th Revision, Clinical Modification (ICD-9-CM) codes in conjunction with the GPI and AHFSCC medication codes. As an initial basis for defining disease concepts, the validated Clinical Classification Software (CCS) codes were used to map ICD-9-CM definitions.²⁸ These codes were compared to other validated coding algorithms used by the Center for Medicare and Medicaid Services (CMS)²⁹, Agency for Healthcare Research and Quality (AHRQ)³⁰, and coding algorithms used widely with administrative claims data.^{31,32} Identified codes were included if they were present in at least two of these sources.

For disease states not defined using the above sources, literature searches were performed in MEDLINE using “ICD-9” and “administrative claims” with a description of the disease. Additionally, a manual search of an ICD-9-CM dataset and web pages for ICD-9-CM coding were queried using disease specific terms (<http://www.cms.gov/medicare-coverage-database/staticpages/icd-9-code-lookup.aspx>; <http://icd9.chrisendres.com/>). A review of all code selections was conducted by two clinical pharmacists with experience in administrative claims research and a geriatric pharmacy specialist. (The full details of PIM disease definitions are provided in Supplement 1.)

A time varying approach was used to assess PIM exposure as a monthly binary variable. For drug-only criteria, a subject was only considered exposed to a PIM for the month a medication was dispensed. For PIM definitions based on co-existing disease states, a patient was considered to have that disease in the month of the first inpatient or non-ancillary outpatient claim with a primary or secondary diagnosis for that disease and for all subsequent months of the study.

Outcome variables

ADEs were based on ICD-9-CM codes previously used for surveillance in hospital claims data³³. A similar manual search strategy was performed with the terms “drug-induced”, “adverse effect”, “caused by”, “poisoning”, “drug”, and “allergy” appearing in code descriptions. ADEs were classified based on the subgroups in the original publication³³ with the addition of those identified through the manual search and the removal of ADEs specific to the perinatal period. (ADEs identified in this study, along with the ICD-9-CM codes and rates, are available in Supplement 2.)

All-cause ED visits were defined by procedure and place of service codes and hospitalizations were identified by unique confinement numbers. ADEs, all-cause ED visits, and all-cause hospitalizations were considered separate outcomes; therefore, an individual could experience one or more of the outcomes.

Subject characteristics

Cohort demographics and plan characteristics were determined at the beginning of the post-index period. Age was categorized as 65–74, 75–84, and 85 years and older. Region was classified as South, West, East, and Midwest. Insurance coverage was categorized into five categories based on payer type (Medicare or commercial) and plan type: HMO (health maintenance organization, non-HMO (Preferred Provider Organization, Consumer Directed, Indemnity, Point of Service), or unknown.

Comorbidities were based on the Charlson Comorbidity Index using the ICD-9-CM coding algorithms by Quan *et al.*³² and were assessed during the 6 month pre-index period. Use of long term care was determined during the pre-index period and included the use of skilled nursing facilities, nursing homes, or hospice care. Additionally, prescription utilization was evaluated separately as the total number of prescription fills and refills annualized to the number per 12 months as well as the total number of unique drug classes used during the post-index period.

Data analysis

Baseline variables for the total cohort were compared for those exposed to a PIM from at least one of the three criteria using two-sided Student t-tests for continuous variables and the Chi-square tests for categorical variables. Unadjusted and adjusted Cox proportional hazards models were used to estimate the relationship between PIM exposure and outcomes. In order to preserve the temporal relationship between PIM exposure and outcomes, individuals having an outcome during the pre-index period were excluded in the model assessing the influence of PIMs on that particular outcome but were included in models exploring one of the other two outcomes. Three time-varying models and one time invariant approach were estimated to explore the temporal relationship between PIM exposure and outcome. The primary model assessed PIM exposure in month $t(i)$ and looked for an outcome in month $t(i + 1)$; providing stronger assurances that the exposure preceded the experience of the outcome event. An alternative time-varying model assessed exposures and outcomes within the same month. An additional third time-varying approach used a once-exposed-always-exposed exposure classification where a subject was considered exposed the first month a PIM was detected and all subsequent months.

Dummy variables were created for each covariate. Reference categories were as follows: Age 65–74; Male gender, East region, and Medicare HMO insurance coverage. The Charlson Comorbidity Index diseases were used as individual binary disease states. Prescription utilization variables were considered to possibly exist along the causal pathway and were excluded from the primary analyses. Sensitivity analyses were conducted which considered each prescription utilization variable as a categorical and continuous variable. Additionally, separate models were estimated which stratified the cohort by these measures.

The proportionality assumption was evaluated for each covariate in models by specifying interaction terms between each covariate and log-time – where statistically significant coefficients would indicate a violation of the assumption. Hazard ratios and 95% confidence intervals are reported.

To compare the predictive validity of the criteria, a c-index specifically developed for time-varying models was used.³⁴ The c-index is analogous to the c-statistic often used with logistic regression and ranges between 0.5 and 1 where a value of 0.5 indicates model prediction no better than chance and a value of 1 indicates a model which predicts events perfectly. Concordance or discordance occur when the predictor score for the individual having an event is greater or lesser than individuals not having an event at that time.³⁴

Additionally, Cohen's kappa was calculated to assess the person-level agreement between all possible pairwise PIM criteria. The sensitivity and specificity of each of the PIM criteria were calculated using each of the outcome measures and a composite outcome measure as gold standards. All analyses were conducted using SAS version 9.3 (SAS Institute, Inc., Cary, NC).

RESULTS

A total of 538,532 individuals were 65 or older during the time period January 1, 2006 and December 31, 2009. Applying eligibility inclusion criteria, 257,206 had at least 9 months of continuous medical insurance enrollment and 175,696 also had 9 months of continuous pharmacy benefit enrollment. Combined, 175,581 had at least 9 months continuous enrollment with both medical and pharmacy benefit during the study period. An additional 1,306 (<1%) individuals were excluded having "Medicaid" identified as the payer type. The final cohort consisted of 174,275 individuals representing 32.4% of the original elderly sample (Supplement 3). The mean follow up time of the cohort was slightly over 2 years (24.9 months, Median 27.0 months, IQR 12–39 months) and the cohort contributed a combined 361,621 person-years. Baseline cohort demographics by PIM exposure are presented in Table 1.

Over the entire post-index period, 72,493 (41.6%) of the cohort were exposed to at least one of the criteria and 19.7% were exposed to all three. Exposure to at least one PIM from the 2012 Beers Criteria was 34.1%, the 2003 Beers Criteria 32.2%, and the STOPP Criteria 27.6%. Person-level agreement between each of the PIM criteria, measured by Cohen's kappa, was "good" between the 2012 and 2003 Beers Criteria ($\kappa = 0.80$, Table 2), and "moderate" between the STOPP Criteria and the 2012 and 2003 Beers Criteria ($\kappa = 0.58$ and 0.59).³⁵

The top 5 individual PIMs for each criteria included many of the same medication groups but differed in prevalence because of different definitions of inappropriateness. A "use with caution" criteria which included SSRIs, SNRIs, antipsychotics, and other medications associated with syndrome of inappropriate anti-diuretic hormone (SIADH) was the most prevalent PIM for the 2012 Beers Criteria (16.2% of cohort). This was followed by benzodiazepines (11.3%), skeletal muscle relaxants (6.6%), non-benzodiazepine hypnotics

(5.8%), and NSAIDs (5.4%). The top five 2003 Beers Criteria PIMs included anticholinergics and first generation antihistamines (19.4%), SSRIs (“with caution”, 10.5%), benzodiazepines (11.2%), muscle relaxants and antispasmodics (7.4%), and long-term NSAID use (5.1%). The most prevalent STOPP Criteria PIMs included NSAIDs (16.2%), opioids (4.8%), beta-blockers (4.7%), corticosteroids (3.8%), and first generation antihistamines (3.8%). (Complete PIM exposure prevalence is available in Supplements 4–6.)

A total of 1,911 individuals experienced a post-index ADE (67 ADEs per 10,000 person-years, 1.12% of the total cohort) after excluding 3,558 people who had pre-index adverse events. Additionally, 24,614 individuals were excluded who had a pre-index ED visit and an additional 29,864 had a post-index event (140 ED visits per 1,000 person-years, 17.1% of the total cohort). Post-index hospitalizations occurred for 16,444 persons (67 hospitalizations per 1,000 person-years, 9.4% of the total cohort) with 22,190 individuals excluded with hospitalizations occurring in the pre-index period. The associations of demographic and health-related characteristics with each outcome are shown in Supplement 7.

PIM exposure was strongly associated with all study outcomes in both unadjusted and adjusted models (Table 3). In the primary unadjusted model, PIM exposure was associated with a 2 to 3 fold increased risk across all outcomes for the 2003 Beers, 2012 Beers, and STOPP Criteria. The associations were similar across the three outcome measures. A stronger relationship between PIM exposure with all three of the criteria and each of the three outcomes (HRs: 3.67 – 5.30) was observed in the time varying models that assessed exposure and outcome in the same month. The time dependent once exposed always exposed model found more modest associations between all the PIM criteria (HRs: 1.30 – 1.76), however all remained significant. The hazard ratios for the STOPP Criteria in the primary time varying model trended higher than those for either of the Beers Criteria.

For the primary unadjusted model, the c-indices were similar for each of the criteria for each of the outcomes and indicated modest levels of discrimination with c-indices between 0.58 and 0.61 (Table 4). When the models included the pre-index covariates, the levels of discrimination increased to 0.65 to 0.70 and were similar across the criteria for each of the outcomes. The model that assessed PIM exposure and outcome in the same month had the highest measures of discrimination than the other models. Inclusion of prescription utilization measures as covariates increased the discrimination of the models less than 1% and stratification had no significant effect. The sensitivity and specificity of the 2012 Beers, 2003 Beers, and STOPP Criteria for the separate composite outcomes are shown in Table 5.

DISCUSSION

In studies using the Beers Criteria, PIM rates of 40–50% are common and have ranged as low as 12%^{7,25,36,37} while rates for the STOPP Criteria have ranged from 13–70%.^{20,21,23,38,39} Our study found 41.6% of the cohort to be exposed to at least one of the criteria. The 2003 and 2012 Beers and STOPP Criteria identified PIMs in 32.2% and 34.1%, and 27.6% of the cohort. These rates are similar to a study in Spain comparing the three

criteria in an ambulatory population.²⁷ Differences between the Beers and STOPP Criteria when they consider similar drug classes are due to differences in the definition of inappropriateness according to each criteria.

We found that exposure to a PIM from any criteria was associated with an increased risk of ADEs, ED visits, and hospitalizations. Individuals with exposure to PIM from the STOPP Criteria had slightly higher risks than either of the Beers Criteria. Despite the slightly higher risk associations for the STOPP Criteria compared to the Beers Criteria, there were only marginal differences in discrimination between the criteria. The 2012 Beers Criteria performed better in terms of sensitivity across all outcomes but was less specific while the STOPP Criteria was less sensitive but more specific.

For the Beers Criteria, the slightly lower performance appears to be a result of higher exposures resulting in more false-positives weakening the association with outcomes. The STOPP Criteria detected only 53% of individuals having any outcome while the 2012 Beers Criteria detected an additional 7% of individuals having each outcome. When the combined “All Criteria” exposure was considered, sensitivity increased for ADEs, ED visits, and hospitalizations. Overall, the composite exposure measure had a sensitivity of 71.4% and specificity of 67.4% for the composite outcome. Therefore, future updates of the Beers Criteria should consider evidence-based refinement of the criteria to include more drug classes that are predictive of serious adverse outcomes.^{40,41}

The AGS has adapted the Beers Criteria into a mobile application and a pocket guide for the practicing clinician who they acknowledge as the target audience.¹⁷ The Beers Criteria were also intended for and have been widely used by researchers, pharmacy benefit managers, and policy-makers – greatly broadening the impact of the Beers Criteria over the last twenty years.^{42,43} For example, the criteria have been used by the CMS and the National Committee for Quality Assurance (NCQA) as quality indicators in long-term care and ambulatory settings. There have even been cases of “misuse” of the criteria to deny coverage of medications⁴⁴ leading to an AGS letter to insurers specifying the appropriate uses of the criteria which do not include formulary decision making.⁴⁵

The STOPP/START Criteria were updated in 2014 further expanding and specializing these criteria to detect inappropriate prescribing in older adults.⁴⁶ Additionally, the 2015 update to the Beers Criteria were available for public comment until May 5th, 2015 on the AGS’s website and will be made available soon for use. The updates to these criteria were not available at the time this study was conducted. Future work should further investigate the predictive validity of these criteria as they are updated.

One of the notable limitations of this study is the outcome measures selected. We used a narrow set of ICD-9-CM codes specific to drug-induced syndromes to define an adverse drug event, some of which are based on supplementary E-codes. These codes were based on previous work which measured the performance of these codes as an ADE surveillance system. They found that the codes had an overall sensitivity and specificity of 55% and 97% for ADEs causing hospital admission and positive predictive value greater than 70%.³³ Though these codes may have only detected half of all adverse drug events in that study, the

codes can be expected to detect true ADEs. Conversely, the all-cause hospitalizations and ED visits are not specific to ADEs and may have higher sensitivity detecting serious ADEs but will be less specific. For example, up to 31% of hospitalizations may be medication related⁴⁷ leaving two thirds that are not. This should be considered when interpreting our findings, particularly when we report the sensitivity and specificity which should not be interpreted in the conventional fashion as measures of diagnostic or screening accuracy for a verified outcome.

Non-prescription medications, such as aspirin or NSAIDs, and prescriptions not processed through claims were not present in the data. For example, inappropriate use of proton pump inhibitors based on the STOPP Criteria has been found to be highly prevalent and its underrepresentation in our data may bias the associations between the STOPP Criteria and the outcomes toward the null. The absence of medications considered by all three criteria sets would have a similar but balanced effect. Similarly, disease-dependent PIM definitions may suffer from missing data due to undercoding.^{48,49} Thus, our findings are likely conservative as more individuals are likely to be exposed to PIMs than were observed in this study.

We excluded the therapeutic duplication criteria from the STOPP Criteria PIM definition because this item has been specifically mentioned for exclusion from the Beers Criteria as it is not a problem unique to the elderly.⁴¹ While therapeutic duplication has been reported to have a prevalence of nearly 5%, it has not been associated with ADEs in published studies.^{5,38} Our exclusion of this item may have decreased the exposure prevalence and the association of the STOPP Criteria with the outcomes.

The most prevalent PIM from the 2012 Beers Criteria considered “use with caution” medications because of the risk of SIADH. Based on the original wording of the 2012 Beers Criteria, this criterion did not require an individual to have previous episodes of SIADH, compared to the STOPP Criteria which did require a previous diagnosis of hyponatremia or SIADH. Thus, all individuals exposed to these commonly used medications, including selective serotonin and norepinephrine reuptake inhibitors, were considered exposed to the Beers Criteria.

This study was strengthened by considering the temporal relationship of exposure and outcomes with a time-varying approach. This allowed for the observation of the initial period of PIM exposure when adverse events may be more apt to occur.⁵⁰ This method also allows individuals to move to and from exposed and unexposed status taking into account changes, additions, and discontinuations of therapy. Our primary model in which exposure was assessed in one month and outcomes assessed in the following month strongly preserves the temporal relationship where exposure precedes outcomes. Though the month-to-month model provided stronger associations between exposure and outcome, reverse causality may explain the stronger association. On the other hand, this approach would not capture associations to PIM and ADEs that have latency periods beyond one month which would bias the findings towards the null.

The administrative claims capture the healthcare utilization of members enrolled in commercial coverage and Medicare Advantage plans and would be expected to be generalizable to that population. Individuals covered under traditional Medicare are not included. This population may differ from the general Medicare population by demographic characteristics such as income status, education, and health behaviors which could not be compared in the current study.

CONCLUSIONS

This was the first study to compare the predictive validity of the updated Beers Criteria to the STOPP Criteria in a population of older adults as well as the first application of the full Beers Criteria including drug-disease items in the US. Our study showed low agreement and no significant differences between the two iterations of the Beers Criteria and the STOPP Criteria in the level of discrimination for ADEs, ED visits, and hospitalizations, though each was moderately prognostic of these outcomes. Future evidence-guided updates of these widely used tools should identify medications and medication classes that may increase the predictive ability of the criteria. These criteria can be used in a complementary fashion to enhance sensitivity of detecting ADEs to decrease adverse drug events in older patients.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding: The project described was supported by the Translational Research Institute (TRI), grant UL1TR000039 through the NIH National Center for Research Resources and National Center for Advancing Translational Sciences.

BCM was paid by the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) to teach courses in retrospective database analysis. This study was unrelated to that course content and ISPOR had no affiliation or review of the submitted work. BCM received a grant (NIH Grant # 1UL1RR029884) which supported acquisition of the data used in this study. CL is a consultant for eMaxHealth Systems on unrelated studies. LCH received a grant from MedEdPortal/Josiah Macy Foundation on interprofessional education development and served as a consultant for the Arkansas Foundation for Medical Care drug safety quality improvement projects. LCH has stock in Cardinal Health and CareFusion and has received royalties from the American Society of Health System Pharmacists for a pharmacy textbook which are unrelated to this work.

JDB is now the University of Kentucky, Humana, Pfizer Doctoral Fellow at the Institute for Pharmaceutical Outcomes and Policy in Lexington, KY. This work was completed before taking this new position and the aforementioned companies had no involvement in the concept, design, interpretation, or drafting of this manuscript.

We do not believe these are potential conflicts of interest, but report them in the interest of full disclosure.

Sponsor's Role: This project was supported by the UAMS Translational Research Institute (NIH Grant # 1UL1RR029884) which supported acquisition of the data. The sponsor had no other role in this study.

References

1. Gallagher P, Ryan C, Byrne S, Kennedy J, O'Mahony D. STOPP (screening tool of older person's prescriptions) and START (screening tool to alert doctors to right treatment). consensus validation. *Int J Clin Pharmacol Ther*. 2008; 46(2):72–83. [PubMed: 18218287]

2. Laroche ML, Charmes JP, Nouaille Y, Picard N, Merle L. Is inappropriate medication use a major cause of adverse drug reactions in the elderly? *Br J Clin Pharmacol*. 2007; 63(2):177–186. DOI: 10.1111/j.1365-2125.2006.02831.x [PubMed: 17166186]
3. Spinewine A, Schmader KE, Barber N, et al. Appropriate prescribing in elderly people: How well can it be measured and optimised? *Lancet*. 2007; 370(9582):173–184. DOI: 10.1016/S0140-6736(07)61091-5 [PubMed: 17630041]
4. Fillenbaum GG, Hanlon JT, Landerman LR, et al. Impact of inappropriate drug use on health services utilization among representative older community-dwelling residents. *Am J Geriatr Pharmacother*. 2004; 2(2):92–101. [PubMed: 15555485]
5. Cahir C, Fahey T, Teeling M, Teljeur C, Feely J, Bennett K. Potentially inappropriate prescribing and cost outcomes for older people: A national population study. *Br J Clin Pharmacol*. 2010; 69(5): 543–552. DOI: 10.1111/j.1365-2125.2010.03628.x [PubMed: 20573091]
6. Chiatti C, Bustacchini S, Furneri G, et al. The economic burden of inappropriate drug prescribing, lack of adherence and compliance, adverse drug events in older people: A systematic review. *Drug Saf*. 2012; 35(Suppl 1):73–87. DOI: 10.1007/BF03319105 [PubMed: 23446788]
7. Perri M 3rd, Menon AM, Deshpande AD, et al. Adverse outcomes associated with inappropriate drug use in nursing homes. *Ann Pharmacother*. 2005; 39(3):405–411. aph.1E230 [pii]. [PubMed: 15671088]
8. Corsonello A, Onder G, Abbatecola AM, Guffanti EE, Gareri P, Lattanzio F. Explicit criteria for potentially inappropriate medications to reduce the risk of adverse drug reactions in elderly people: From beers to STOPP/START criteria. *Drug Saf*. 2012; 35(Suppl 1):21–28. DOI: 10.1007/BF03319100 [PubMed: 23446783]
9. Stockl KM, Le L, Zhang S, Harada AS. Clinical and economic outcomes associated with potentially inappropriate prescribing in the elderly. *Am J Manag Care*. 2010; 16(1):e1–10. 12554 [pii]. [PubMed: 20059286]
10. Jano E, Aparasu RR. Healthcare outcomes associated with beers' criteria: A systematic review. *Ann Pharmacother*. 2007; 41(3):438–447. aph.1H473 [pii]. [PubMed: 17311835]
11. Price SD, Holman CD, Sanfilippo FM, Emery JD. Association between potentially inappropriate medications from the beers criteria and the risk of unplanned hospitalization in elderly patients. *Ann Pharmacother*. 2014; 48(1):6–16. DOI: 10.1177/1060028013504904 [PubMed: 24396090]
12. Beers MH, Ouslander JG, Rollingher I, Reuben DB, Brooks J, Beck JC. Explicit criteria for determining inappropriate medication use in nursing home residents. UCLA division of geriatric medicine. *Arch Intern Med*. 1991; 151(9):1825–1832. [PubMed: 1888249]
13. Dimitrow MS, Airaksinen MS, Kivela SL, Lyles A, Leikola SN. Comparison of prescribing criteria to evaluate the appropriateness of drug treatment in individuals aged 65 and older: A systematic review. *J Am Geriatr Soc*. 2011; 59(8):1521–1530. DOI: 10.1111/j.1532-5415.2011.03497.x [PubMed: 21797829]
14. Luo R, Scullin C, Mullan AM, Scott MG, McElnay JC. Comparison of tools for the assessment of inappropriate prescribing in hospitalized older people. *J Eval Clin Pract*. 2012; 18(6):1196–1202. DOI: 10.1111/j.1365-2753.2011.01758.x [PubMed: 21883716]
15. Beers MH. Explicit criteria for determining potentially inappropriate medication use by the elderly. an update. *Arch Intern Med*. 1997; 157(14):1531–1536. [PubMed: 9236554]
16. 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*. 2003; 163(22):2716–2724. DOI: 10.1001/archinte.163.22.2716 [PubMed: 14662625]
17. The American Geriatrics Society 2012 Beers Criteria Update Expert Panel. American geriatrics society updated beers criteria for potentially inappropriate medication use in older adults. *J Am Geriatr Soc*. 2012; 60(4):616–631. DOI: 10.1111/j.1532-5415.2012.03923.x [PubMed: 22376048]
18. O'Mahony D, Gallagher PF. Inappropriate prescribing in the older population: Need for new criteria. *Age Ageing*. 2008; 37(2):138–141. DOI: 10.1093/ageing/afm189 [PubMed: 18349010]
19. Dalleur O, Boland B, Spinewine A. 2012 updated beers criteria: Greater applicability to europe? *J Am Geriatr Soc*. 2012; 60(11):2188–9. author reply 2189–90. DOI: 10.1111/j.1532-5415.2012.04219.x [PubMed: 23148435]

20. O'Sullivan DP, O'Mahony D, Parsons C, et al. A prevalence study of potentially inappropriate prescribing in Irish long-term care residents. *Drugs Aging*. 2013; 30(1):39–49. DOI: 10.1007/s40266-012-0039-7 [PubMed: 23229766]
21. Gallagher PF, O'Connor MN, O'Mahony D. Prevention of potentially inappropriate prescribing for elderly patients: A randomized controlled trial using STOPP/START criteria. *Clin Pharmacol Ther*. 2011; 89(6):845–854. DOI: 10.1038/clpt.2011.44 [PubMed: 21508941]
22. Hamilton H, Gallagher P, Ryan C, Byrne S, O'Mahony D. Potentially inappropriate medications defined by STOPP criteria and the risk of adverse drug events in older hospitalized patients. *Arch Intern Med*. 2011; 171(11):1013–1019. DOI: 10.1001/archinternmed.2011.215 [PubMed: 21670370]
23. Ryan C, O'Mahony D, Kennedy J, Weedle P, Byrne S. Potentially inappropriate prescribing in an Irish elderly population in primary care. *Br J Clin Pharmacol*. 2009; 68(6):936–947. DOI: 10.1111/j.1365-2125.2009.03531.x [PubMed: 20002089]
24. O'Sullivan D, O'Mahony D, Byrne S. Inappropriate prescribing in Irish nursing home residents. *Eur Geriatr Med*. 2010; 1S(S84)
25. Hill-Taylor B, Sketris I, Hayden J, Byrne S, O'Sullivan D, Christie R. Application of the STOPP/START criteria: A systematic review of the prevalence of potentially inappropriate prescribing in older adults, and evidence of clinical, humanistic and economic impact. *J Clin Pharm Ther*. 2013; doi: 10.1111/jcpt.12059
26. Gallagher P, O'Mahony D. STOPP (screening tool of older persons' potentially inappropriate prescriptions): Application to acutely ill elderly patients and comparison with Beers' criteria. *Age Ageing*. 2008; 37(6):673–679. DOI: 10.1093/ageing/afn197 [PubMed: 18829684]
27. Blanco-Reina E, Ariza-Zafra G, Ocana-Riola R, Leon-Ortiz M. 2012 American Geriatrics Society Beers criteria: Enhanced applicability for detecting potentially inappropriate medications in European older adults? A comparison with the screening tool of older person's potentially inappropriate prescriptions. *J Am Geriatr Soc*. 2014; doi: 10.1111/jgs.12891
28. Agency for Healthcare Research and Quality, AHRQ. [Accessed 6/1, 2013] Clinical classifications software (CCS) for ICD-9-CM: Healthcare cost and utilization project (HCUP) tools & software. <http://www.hcup-us.ahrq.gov/toolssoftware/ccs/ccs.jsp#pubs>
29. Bucaaneer, A General Dynamics Company, under contract with CMS. [Accessed 05/02, 2014] Chronic conditions data warehouse, condition categories. <https://www.ccwdata.org/web/guest/condition-categories>
30. Elixhauser, A., Steiner, C., Kuzik, D. [Accessed 05/01, 2014] Comorbidity software documentation. 2004. HCUP methods series report #2004-1. <http://www.hcup-us.ahrq.gov/toolssoftware/comorbidity/comorbidity.jsp>. Updated 2004
31. Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. *Med Care*. 1998; 36(1):8–27. [PubMed: 9431328]
32. Quan H, Sundararajan V, Halfon P, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. *Med Care*. 2005; 43(11):1130–1139. [PubMed: 16224307]
33. Hougland, P., Nebeker, J., Pickard, S., et al. Using ICD-9-CM codes in hospital claims data to detect adverse events in patient safety surveillance. In: Henriksen, K., Battles, J.B., Keyes, M.A., Grady, M.L., editors. *Advances in patient safety: New directions and alternative approaches* (vol. 1: Assessment). Rockville (MD): 2008.
34. Kremers, W. Concordance for survival time data: Fixed and time-dependent covariates and possible ties in predictor and time. Department of Health Sciences Research, Mayo Clinic; Rochester, MN: 2007. (Technical Report #80)
35. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977; 33(1):159–174. [PubMed: 843571]
36. Lau DT, Kasper JD, Potter DE, Lyles A. Potentially inappropriate medication prescriptions among elderly nursing home residents: Their scope and associated resident and facility characteristics. *Health Serv Res*. 2004; 39(5):1257–1276. DOI: 10.1111/j.1475-6773.2004.00289.x [PubMed: 15333108]

37. Levy HB, Marcus EL, Christen C. Beyond the beers criteria: A comparative overview of explicit criteria. *Ann Pharmacother*. 2010; 44(12):1968–1975. DOI: 10.1345/aph.1P426 [PubMed: 21081709]
38. Vishwas HN, Harugeri A, Parthasarathi G, Ramesh M. Potentially inappropriate medication use in indian elderly: Comparison of beers' criteria and screening tool of older persons' potentially inappropriate prescriptions. *Geriatr Gerontol Int*. 2012; 12(3):506–514. DOI: 10.1111/j.1447-0594.2011.00806.x [PubMed: 22239067]
39. Hamilton H, Gallagher P, Ryan C, Byrne S, O'Mahony D. Potentially inappropriate medications defined by STOPP criteria and the risk of adverse drug events in older hospitalized patients. *Arch Intern Med*. 2011; 171(11):1013–1019. DOI: 10.1001/archinternmed.2011.215 [PubMed: 21670370]
40. Curtain CM, Bindoff IK, Westbury JL, Peterson GM. A comparison of prescribing criteria when applied to older community-based patients. *Drugs Aging*. 2013; 30(11):935–943. DOI: 10.1007/s40266-013-0116-6 [PubMed: 24002743]
41. Marcum ZA, Hanlon JT. Commentary on the new american geriatric society beers criteria for potentially inappropriate medication use in older adults. *Am J Geriatr Pharmacother*. 2012; 10(2): 151–159. DOI: 10.1016/j.amjopharm.2012.03.002 [PubMed: 22483163]
42. Fick DM, Semla TP. 2012 american geriatrics society beers criteria: New year, new criteria, new perspective. *J Am Geriatr Soc*. 2012; 60(4):614–615. DOI: 10.1111/j.1532-5415.2012.03922.x [PubMed: 22376003]
43. Resnick B, Pacala JT. 2012 beers criteria. *J Am Geriatr Soc*. 2012; 60(4):612–613. DOI: 10.1111/j.1532-5415.2012.03921.x [PubMed: 22375952]
44. Berger MS. Misuse of beers criteria. *J Am Geriatr Soc*. 2014; 62(7):1411. doi: 10.1111/jgs.12911
45. McCormick WC. American geriatrics society response to letter to the editor from marc S. berger "misuse of beers criteria" july 2014. *J Am Geriatr Soc*. 2014; 62(12):2466–2467. DOI: 10.1111/jgs.13231 [PubMed: 25516058]
46. O'Mahony D, O'Sullivan D, Byrne S, O'Connor MN, Ryan C, Gallagher P. STOPP/START criteria for potentially inappropriate prescribing in older people: Version 2. *Age Ageing*. 2015; 44(2):213–218. DOI: 10.1093/ageing/afu145 [PubMed: 25324330]
47. Salvi F, Marchetti A, D'Angelo F, Boemi M, Lattanzio F, Cherubini A. Adverse drug events as a cause of hospitalization in older adults. *Drug Saf*. 2012; 35(Suppl 1):29–45. DOI: 10.1007/BF03319101 [PubMed: 23446784]
48. Zhan C, Miller MR. Administrative data based patient safety research: A critical review. *Qual Saf Health Care*. 2003; 12(Suppl 2):ii58–63. [PubMed: 14645897]
49. Schneeweiss S, Avorn J. A review of uses of health care utilization databases for epidemiologic research on therapeutics. *J Clin Epidemiol*. 2005; 58(4):323–337. S0895-4356(04)00298-7 [pii]. [PubMed: 15862718]
50. Dedhiya SD, Hancock E, Craig BA, Doebbeling CC, Thomas J 3rd. Incident use and outcomes associated with potentially inappropriate medication use in older adults. *Am J Geriatr Pharmacother*. 2010; 8(6):562–570. DOI: 10.1016/S1543-5946(10)80005-4 [PubMed: 21356505]
51. Medina-Ramon M, Goldberg R, Melly S, Mittleman MA, Schwartz J. Residential exposure to traffic-related air pollution and survival after heart failure. *Environ Health Perspect*. 2008; 116(4): 481–485. DOI: 10.1289/ehp.10918 [PubMed: 18414630]
52. Oliphant SS, Wang L, Bunker CH, Lowder JL. Trends in stress urinary incontinence inpatient procedures in the united states, 1979–2004. *Am J Obstet Gynecol*. 2009; 200(5):521.e1–521.e6. DOI: 10.1016/j.ajog.2009.01.007 [PubMed: 19375571]
53. Anger JT, Saigal CS, Madison R, Joyce G, Litwin MS. Urologic Diseases of America Project. Increasing costs of urinary incontinence among female medicare beneficiaries. *J Urol*. 2006; 176(1):247–51. discussion 251. DOI: 10.1016/S0022-5347(06)00588-X [PubMed: 16753411]
54. Wallace KL, Riedel AA, Joseph-Ridge N, Wortmann R. Increasing prevalence of gout and hyperuricemia over 10 years among older adults in a managed care population. *J Rheumatol*. 2004; 31(8):1582–1587. 0315162X-31-1582 [pii]. [PubMed: 15290739]

55. Asche CV, Joish VN, Camacho F, Drake CL. The direct costs of untreated comorbid insomnia in a managed care population with major depressive disorder. *Curr Med Res Opin.* 2010; 26(8):1843–1853. DOI: 10.1185/03007995.2010.488037 [PubMed: 20515321]
56. Finkelstein MM, Jerrett M. A study of the relationships between parkinson's disease and markers of traffic-derived and environmental manganese air pollution in two canadian cities. *Environ Res.* 2007; 104(3):420–432. S0013-9351(07)00054-0 [pii]. [PubMed: 17445792]
57. Oliveria SA, Liperoti R, L'italien G, et al. Adverse events among nursing home residents with alzheimer's disease and psychosis. *Pharmacoepidemiol Drug Saf.* 2006; 15(11):763–774. DOI: 10.1002/pds.1274 [PubMed: 16823890]
58. Williams C, Simon TD, Riva-Cambrin J, Bratton SL. Hyponatremia with intracranial malignant tumor resection in children. *J Neurosurg Pediatr.* 2012; 9(5):524–529. DOI: 10.3171/2012.1.PEDS11465 [PubMed: 22546031]
59. Movig KL, Leufkens HG, Lenderink AW, Egberts AC. Validity of hospital discharge international classification of diseases (ICD) codes for identifying patients with hyponatremia. *J Clin Epidemiol.* 2003; 56(6):530–535. S0895435603000064 [pii]. [PubMed: 12873647]
60. Khan N, Abbas AM, Almukhtar RM, Khan A. Prevalence and predictors of low bone mineral density in males with ulcerative colitis. *J Clin Endocrinol Metab.* 2013; 98(6):2368–2375. DOI: 10.1210/jc.2013-1332 [PubMed: 23596137]
61. Anderson FA Jr, Wheeler HB, Goldberg RJ, et al. A population-based perspective of the hospital incidence and case-fatality rates of deep vein thrombosis and pulmonary embolism. the worcester DVT study. *Arch Intern Med.* 1991; 151(5):933–938. [PubMed: 2025141]
62. Zhan C, Battles J, Chiang YP, Hunt D. The validity of ICD-9-CM codes in identifying postoperative deep vein thrombosis and pulmonary embolism. *Jt Comm J Qual Patient Saf.* 2007; 33(6):326–331. [PubMed: 17566542]
63. Shorr RI, Ray WA, Daugherty JR, Griffin MR. Incidence and risk factors for serious hypoglycemia in older persons using insulin or sulfonylureas. *Arch Intern Med.* 1997; 157(15):1681–1686. [PubMed: 9250229]
64. Ginde AA, Blanc PG, Lieberman RM, Camargo CA Jr. Validation of ICD-9-CM coding algorithm for improved identification of hypoglycemia visits. *BMC Endocr Disord.* 2008; 8:4-6823-8-4.doi: 10.1186/1472-6823-8-4 [PubMed: 18380903]
65. Wang YR, Fisher RS, Parkman HP. Gastroparesis-related hospitalizations in the united states: Trends, characteristics, and outcomes, 1995–2004. *Am J Gastroenterol.* 2008; 103(2):313–322. DOI: 10.1111/j.1572-0241.2007.01658.x [PubMed: 18047541]

Table 1

Baseline Characteristics for Cohort and Those Exposed to the 2012 Beers, 2003 Beers, and STOPP Criteria

Characteristics	Total Cohort N=174,275 No. (%)	2012 Beers N=59,426 No. (%)	2003 Beers N=56,144 No. (%)	STOPP N=48,121 No. (%)
Age (years) *				
65–74	128,306 (73.6)	37,150 (62.5)	36,603 (65.2)	30,951 (64.3)
75–84	34,637 (19.9)	17,098 (28.8)	14,991 (26.7)	13,098 (27.2)
85 and older	11,332 (6.5)	5,178 (8.7)	4,550 (8.1)	4,072 (8.5)
Female *	94,588 (54.3)	34,779 (58.5)	33,997 (60.6)	27,809 (57.8)
Insurance Type *				
Medicare HMO	22,570 (13.0)	11,071 (18.6)	9,907 (17.7)	8,630 (17.9)
Medicare non-HMO	24,992 (14.3)	8,995 (15.1)	8,357 (14.9)	7,248 (15.1)
Commercial HMO	20,432 (11.7)	5,449 (9.2)	5,311 (9.5)	4,755 (9.9)
Commercial non-HMO	96,412 (55.3)	31,116 (52.4)	29,868 (53.2)	25,214 (52.4)
Unspecified	9,869 (5.7)	2,795 (4.7)	2,701 (4.8)	2,274 (4.7)
Region				
East	35,987 (20.7)	11,739 (19.8)	11,333 (20.2)	10,233 (21.3)
Midwest	57,514 (33.0)	20,388 (34.3)	19,219 (34.2)	16,688 (34.7)
South	43,528 (25.0)	14,294 (24.1)	13,393 (23.9)	11,508 (23.9)
West	37,246 (21.4)	13,005 (21.9)	12,199 (21.7)	9,692 (20.1)
Charlson Co-morbidity Index (Pre-index)				
Mean (SD) *	1.3 (1.7)	1.6 (1.9)	1.6 (1.8)	1.8 (2.0)
0–1	117,690 (67.5)	35,013 (58.9)	33,803 (60.2)	27,111 (56.3)
2–3	39,736 (22.8)	16,513 (27.8)	15,265 (27.2)	13,744 (28.6)
3+	16,849 (9.7)	7,900 (13.3)	7,076 (12.6)	7,266 (15.1)
Prescription utilization				
Total prescription fills per 12 months				
Mean (SD) *	12.1 (36.5)	20.5 (17.7)	20.0 (16.3)	19.4 (16.3)
Unique drug classes				
Mean (SD) *	4.4 (5.3)	8.8 (5.5)	8.7 (5.5)	8.8 (5.7)
Long term care	3,682 (2.1)	2,287 (3.9)	2,088 (3.7)	2,126 (4.42)
Follow-up time (months)				
Mean (SD) *	24.9 (13.2)	29.1 (11.8)	29.4 (11.6)	30.1 (11.2)
Median	27.0	36.0	36.0	36.0
IQR	12–39	18–39	19–39	12–39

* p<0.01 for comparison between “Any Exposure” and Total Cohort.

Significant differences were not observed between criteria

Abbreviations: HMO (health maintenance organization); SD (standard deviation); IQR (inter-quartile range)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2

Inappropriate prescribing criteria person-level concordance and agreement

Exposure to inappropriate prescribing	No. (% of cohort) N=174,275
Any exposure to criteria	72,493 (41.6)
Exposed to more than one criteria	58,915 (32.7)
Exposed to all criteria	34,283 (19.7)
Concordance Between Criteria	No. (% Agree)
2012 Beers*2003 Beers	50,182 (84.4)
2012 Beers*STOPP	38,006 (64.0)
2003 Beers*2012 Beers	50,182 (89.4)
2003 Beers*STOPP	37,293 (66.4)
STOPP*2012 Beers	38,006 (79.0)
STOPP*2003 Beers	37,293 (77.5)
2012 Beers*All Criteria	59,426 (82.0)
2003 Beers*All Criteria	56,144 (77.4)
STOPP*All Criteria	48,121 (66.4)
Agreement Between Criteria	Cohen's Kappa
2012 Beers*2003 Beers	0.80
2012 Beers*STOPP	0.58
2003 Beers*STOPP	0.59
2012 Beers*All Criteria	0.84
2003 Beers*All Criteria	0.80
STOPP*All Criteria	0.70

Table 3

Adjusted and Unadjusted Hazards ratios for the 2012 Beers, 2003 Beers, and STOPP Criteria for time varying and non-time varying models.

Criteria	Unadjusted models (exposure only)			Adjusted models ^a		
	Time-varying monthly lag (Primary Model) ^b					
	ADEs	Emergency	Hospitalization	ADEs	Emergency	Hospitalization
2012 Beers	2.51 (2.33–2.70)	2.21 (2.16–2.25)	2.25 (2.20–2.30)	2.17 (2.01–2.34)	2.00 (1.96–2.04)	2.03 (1.98–2.07)
2003 Beers	2.65 (2.46–2.85)	2.29 (2.25–2.34)	2.31 (2.26–2.37)	2.33 (2.16–2.52)	2.14 (2.10–2.19)	2.16 (2.11–2.21)
STOPP	2.89 (2.68–3.12)	2.66 (2.60–2.72)	2.80 (2.74–2.87)	2.43 (2.24–2.63)	2.38 (2.32–2.43)	2.46 (2.40–2.52)
Time-varying month to month ^c						
	ADEs	Emergency	Hospitalization	ADEs	Emergency	Hospitalization
2012 Beers	4.33 (4.11–4.56)	4.38 (4.31–4.44)	4.27 (4.20–4.34)	3.67 (3.48–3.87)	3.93 (3.87–3.99)	3.75 (3.68–3.81)
2003 Beers	5.01 (4.75–5.28)	4.89 (4.81–4.97)	4.76 (4.68–4.84)	4.30 (4.08–4.54)	4.51 (4.44–4.58)	4.32 (4.25–4.40)
STOPP	5.21 (4.91–5.52)	5.18 (5.09–5.28)	5.30 (5.20–5.41)	4.18 (3.92–4.44)	4.52 (4.43–4.60)	4.47 (4.38–4.56)
Time-dependent once exposed, always exposed ^d						
	ADEs	Emergency	Hospitalization	ADEs	Emergency	Hospitalization
2012 Beers	1.71 (1.57–1.87)	1.45 (1.42–1.48)	1.46 (1.42–1.49)	1.43 (1.31–1.56)	1.32 (1.29–1.35)	1.30 (1.26–1.33)
2003 Beers	1.66 (1.53–1.81)	1.39 (1.36–1.42)	1.38 (1.35–1.42)	1.45 (1.33–1.58)	1.32 (1.29–1.35)	1.30 (1.26–1.33)
STOPP	1.76 (1.62–1.91)	1.50 (1.46–1.53)	1.54 (1.51–1.58)	1.47 (1.35–1.60)	1.37 (1.34–1.40)	1.38 (1.34–1.42)
Ever exposure ^e						
	ADEs	Emergency	Hospitalization	ADEs	Emergency	Hospitalization
2012 Beers	3.06 (2.77–3.37)	2.34 (2.28–2.39)	2.58 (2.51–2.65)	2.60 (2.35–2.88)	2.08 (2.03–2.13)	2.27 (2.21–2.34)
2003 Beers	2.83 (2.57–3.12)	2.18 (2.13–2.23)	2.33 (2.27–2.39)	2.49 (2.25–2.74)	2.01 (1.97–2.06)	2.15 (2.09–2.21)
STOPP	3.11 (2.83–3.42)	2.44 (2.38–2.49)	2.71 (2.64–2.78)	2.64 (2.39–2.91)	2.18 (2.13–2.23)	2.38 (2.32–2.45)

^a Covariates included: Age, gender, insurance status, region, long term care, and Charlson comorbidities

^b Outcome events associated with time-varying exposure in the preceding month (i.e. March outcome associated with February exposure

^c Outcome events associated with time-varying exposure in the same month

^d Once exposed to a criteria, always exposed whether or not exposure status changes

^e Exposed at any point during the post-index follow up period

Sample size for final model excluding pre-index events: ADEs (170,717); ED visits (147,661); Hospitalizations (152,085)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 4

C-indices and 95% confidence intervals for 2003 Beers, 2012 Beers, and STOPP Criteria for the time varying and non-time varying models

Unadjusted Model (exposure only)				Adjusted Model ^{d*}	
Time-varying monthly lag (Primary Model) ^b					
Criteria	ADEs	Emergency	Hospitalization	ADE	Hospitalization
2012 Beers	0.603 (0.597–0.609)	0.585 (0.583–0.587)	0.590 (0.588–0.592)	0.688 (0.677–0.700)	0.652 (0.649–0.655)
2003 Beers	0.605 (0.599–0.611)	0.585 (0.583–0.587)	0.588 (0.586–0.590)	0.695 (0.684–0.706)	0.653 (0.650–0.656)
STOPP	0.607 (0.601–0.614)	0.590 (0.588–0.592)	0.599 (0.597–0.601)	0.695 (0.685–0.706)	0.661 (0.658–0.664)
Time-varying month to month ^c					
ADEs	Emergency	Hospitalization	ADE	Emergency	Hospitalization
2012 Beers	0.642 (0.639–0.645)	0.635 (0.634–0.636)	0.636 (0.634–0.638)	0.733 (0.723–0.744)	0.709 (0.706–0.712)
2003 Beers	0.646 (0.643–0.650)	0.635 (0.634–0.636)	0.637 (0.636–0.638)	0.741 (0.730–0.751)	0.708 (0.705–0.711)
STOPP	0.642 (0.638–0.647)	0.626 (0.625–0.628)	0.634 (0.633–0.634)	0.741 (0.730–0.752)	0.707 (0.704–0.710)
Time-varying once exposed, always exposed					
ADEs	Emergency	Hospitalization	ADE	Emergency	Hospitalization
2012 Beers	0.566 (0.557–0.574)	0.548 (0.546–0.551)	0.551 (0.549–0.554)	0.666 (0.654–0.679)	0.628 (0.624–0.631)
2003 Beers	0.563 (0.554–0.571)	0.542 (0.540–0.545)	0.544 (0.541–0.546)	0.667 (0.655–0.680)	0.626 (0.622–0.629)
STOPP	0.567 (0.559–0.574)	0.548 (0.546–0.551)	0.554 (0.552–0.557)	0.670 (0.658–0.682)	0.630 (0.627–0.634)
Ever exposure ^e					
ADEs	Emergency	Hospitalization	ADE	Emergency	Hospitalization
2012 Beers	0.566 (0.557–0.574)	0.548 (0.546–0.551)	0.551 (0.549–0.554)	0.666 (0.654–0.679)	0.628 (0.624–0.631)
2003 Beers	0.563 (0.554–0.571)	0.542 (0.540–0.545)	0.544 (0.541–0.546)	0.667 (0.655–0.680)	0.626 (0.622–0.629)
STOPP	0.636 (0.624–0.647)	0.599 (0.596–0.603)	0.612 (0.608–0.615)	0.713 (0.701–0.725)	0.659 (0.656–0.663)

* Covariate only model c-indices: ADE 0.664 (0.651–0.676); Emergency 0.606 (0.603–0.610); Hospitalization 0.647 (0.644–0.651)

^a Covariates included: Age, gender, insurance status, region, long term care, and Charlson comorbidities

^b Outcome events associated with time-varying exposure in the preceding month (i.e. March outcome associated with February exposure)

^c Outcome events associated with time-varying exposure in the same month

^d Once exposed to a criteria, always exposed whether or not exposure status changes

Exposed at any point during the post-index follow up period
Sample size for final model excluding pre-index events: ADEs (170,717); ED visits (147,661); Hospitalizations (152,085)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 5

Sensitivity and specificity of PIM criteria predicting study outcomes

Criteria	Sensitivity (%)	Specificity (%)
2012 Beers		
ADEs	71.2	41.2
Emergency Visits	61.2	70.7
Hospitalizations	64.3	69.0
Composite Outcome	60.6	73.9
2003 Beers		
ADEs	67.7	42.8
Emergency Visits	57.8	72.2
Hospitalizations	60.3	70.4
Composite Outcome	57.3	75.4
STOPP		
ADEs	64.7	47.8
Emergency Visits	53.8	78.1
Hospitalizations	57.6	76.3
Composite Outcome	53.4	80.2
All Criteria exposure		
ADEs	79.8	30.1
Emergency Visits	71.8	63.2
Hospitalizations	74.8	61.4
Composite Outcome	71.4	67.4

Study supplement to facilitate review only, not intended for publication or for inclusion in the word or table count.