Construction and Characteristics of the RxRisk-V: A VA-Adapted Pharmacy-Based Case-Mix Instrument

Author(s): Kevin L. Sloan, Anne E. Sales, Chuan-Fen Liu, Paul Fishman, Paul Nichol, Norman T. Suzuki and Nancy D. Sharp

Reviewed work(s):
Source: *Medical Care*, Vol. 41, No. 6 (Jun., 2003), pp. 761-774
Published by: Lippincott Williams & Wilkins
Stable URL: http://www.jstor.org/stable/3768010
Accessed: 13/06/2012 16:07

MEDICAL CARE
Volume 41, Number 6, pp 761–774

Construction and Characteristics of the RxRisk-V

A VA-Adapted Pharmacy-Based Case-mix Instrument

KEVIN L. SLOAN, MD,*† ANNE E. SALES, PHD,*‡ CHUAN-FEN LIU, PHD,*‡ PAUL FISHMAN, PHD,‡¶ PAUL NICHOL, MD,*‖ NORMAN T. SUZUKI, PHARMD,*§ AND NANCY D. SHARP, PHD*‡

BACKGROUND. Assessment of disease burden is the key to many aspects of health care management. Patient diagnoses are commonly used for case-mix assessment. However, issues pertaining to diagnostic data availability and reliability make pharmacy-based strategies attractive. Our goal was to provide a reliable and valid pharmacy-based case-mix classification system for chronic diseases found in the Veterans Health Administration (VHA) population.

OBJECTIVE. To detail the development and category definitions of a VA-adapted version of the RxRisk (formerly the Chronic Disease Score); to describe category prevalence and reliability; to check category criterion validity against ICD-9 diagnoses; and to assess category-specific regression coefficients in concurrent and prospective cost models.

RESEARCH DESIGN. Clinical and pharmacological review followed by cohort analysis of diagnostic, pharmacy, and utilization databases.

SUBJECTS. 126,075 veteran users of VHA services in Washington, Oregon, Idaho, and Alaska.

METHODS. We used Kappa statistics to evaluate RxRisk category reliability and criterion validity, and multivariate regression to estimate concurrent and prospective cost models.

RESULTS. The RxRisk-V classified 70.5% of the VHA Northwest Network 1998 users into an average of 2.61 categories. Of the 45 classes, 33 classes had good-excellent 1-year reliability and 25 classes had good-excellent criterion validity against ICD-9 diagnoses. The RxRisk-V accounts for a distinct proportion of the variance in concurrent (R^2 = 0.18) and prospective cost (R^2 = 0.10) models.

CONCLUSIONS. The RxRisk-V provides a reliable and valid method for administrators to describe and understand better chronic disease burden of their treated populations. Tailoring to the VHA permits assessment of disease burden specific to this population.

Key words: Case-mix; pharmacy; veterans; risk adjustment. (Med Care 2003;41:761–774)

Valid and reliable assessment of disease burden is the key to many aspects of population-based health care management. Comparing systems of care, evaluating cost and quality outcomes, and assessing the effectiveness of medication and procedural interventions all require accounting for

*From the VA Puget Sound Health Care System, Seattle, Washington.

†From the Department of Psychiatry and Behavioral Sciences, ‡ the Department of Health Services, ‖the Department of Medicine, and §the School of Pharmacy, University of Washington, Seattle, Washington.

¶From the Center for Health Studies, Group Health Cooperative, Seattle, Washington.

This research was supported by the Department of Veterans Health Administration, Health Services Research and Development Service, Project #IIR 99001–1.

The views expressed in this report are those of the authors and do not necessarily represent the views of the Department of Veterans Affairs or the Health Services Research and Development Service. This work was presented, in part, at the Academy for Health Services Research and Health Policy 2001 Annual Meeting in Atlanta, Georgia.

Address correspondence and reprint requests to: Kevin L. Sloan, MD, VA Puget Sound Health Care System (116-DDTP), 1660 S. Columbian Way, Seattle, WA 98108. E-mail: Kevin.Sloan@med.va.gov.

differences in case-mix before robust conclusions may be drawn.

The most frequently used case-mix adjustors for large health care systems have been diagnosis-based instruments such as Diagnosis-Related Groups,[1] Diagnostic Cost Groups (DCGs),[2] and Ambulatory Clinical Groups (ACGs).[3] However, not all health care organizations collect diagnoses, and providers vary in their coding practices.[4–9] In addition, anecdotal evidence suggests that physicians may code only one diagnosis even when patients are seen for multiple conditions[10,11] and/or make coding decisions based upon codes most available or frequently used.[12]

Pharmacy-based instruments may address some of the perceived weakness of diagnosis-based strategies.[13–18] Since the advent of large pharmacy benefit management programs, pharmacy data may be more timely, complete, and less expensive to access than comparable diagnostic data. In addition, health care organizations that do not collect diagnostic data may nevertheless be able to access pharmacy data. Even if organizations collect diagnostic data, because clinical pharmacists are required to review prescriptions before dispensing medications, prescription records have a built-in audit that is absent in diagnostic data. The medication fill record may also provide a more complete list of actively treated conditions than coded diagnoses because chronic conditions frequently require continuing prescriptions whether or not the condition is coded as part of a clinic visit. Lastly, given the potency of many drugs, pharmacy data may be more reliable and less subject to "gaming" by providers than diagnosis data because of the greater consequences of manipulating these choices.[14,19]

Among pharmacy-based case-mix instruments, the Chronic Disease Score (CDS) has been the most extensively described.[13,16,18] The CDS was initially developed by Michael Von Korff and colleagues at the Group Health Cooperative (GHC)[13] with the goal that the score should reflect the number of chronic diseases under treatment, the complexity of the treatment regimen, and the likelihood of disease-related morbidity or mortality. An expert panel selected medication classes and assigned weights. Clark[16] modified the original CDS by expanding the list of chronic diseases and by selecting weights based upon regression models. Fishman and colleagues at GHC continue to update the instrument[20,21] and have recently renamed it the RxRisk model to differentiate it from similarly named instruments.

The Veterans Health Administration (VHA) is one of the largest health care systems in the United States. During 2000, approximately 3,700,000 patients made approximately 39 million out-patient visits to more than 700 VHA health care facilities nationwide.[22] The VHA is a good environment for using pharmacy-based case-mix instruments because the very low copayment for drugs in VHA pharmacies ($2 or less per prescription filled) has created substantial financial incentives for patients to obtain their medications from VHA pharmacies. Even with the recent increase to $7 per prescription per month, VHA pharmacy copayment rates remain well below those seen in the private sector, and for many veterans (those treated for service-connected conditions or with an income less than approximately $10,000), prescriptions remain free of charge. Computerized drug dispense records and databases are widely available from individual VHA facilities and in a single national database.

The VHA patient population is known to be different from general patient populations, with distinct socioeconomic characteristics, poorer health status, more medical conditions, and higher medical use.[23,24] Because case-mix adjustment is most effective when tailored to the population,[25,26] we felt that the GHC RxRisk model should be comprehensively reviewed and tested before utilization in VHA patient populations. We also wished to target chronic diseases found at higher rates in the VHA patient population.

The overall goal of our work was to develop a clinically relevant, reliable, and valid pharmacy-based case-mix instrument to improve assessment of disease and disease burden among VHA patients. In this study, we detail the development, category definitions, and psychometric properties of the VHA-adapted RxRisk (RxRisk-V). Specifically, for each RxRisk-V category we (1) assure face validity through careful clinician and pharmacist review; (2) examine reliability by testing year-to-year category stability; (3) check criterion validity by comparing RxRisk categories with their diagnosis-based analogs; and (4) assess discriminant validity by estimating and appraising coefficients for concurrent and prospective cost models. In a companion article by Sales et al,[27] we compare the ability of Rx-Risk-V and other leading risk adjustment approaches to predict total costs of care. Note that the analyses in this manuscript

focus on the single-year user population, whereas the accompanying Sales et al[27] paper examines the 3-year user population.

Methods

RxRisk-V Development

The principles we used when adapting the GHC RxRisk to the VHA population were that categories should be consistent with the GHC RxRisk whenever possible, new categories should be face-valid, and all categories should have at least 300 subjects unless some clinical or economic rationale suggests that a lower threshold is warranted. Subjects for this analysis were all Fiscal Year (FY) 1998 veteran users of VHA Northwest Network services.

First, we reviewed the structure and content of the current RxRisk. The GHC RxRisk software translates National Drug Codes (NDCs)—unique 11-digit numbers assigned by the US Food and Drug Administration that identify the labeler/vendor, product, and trade package size—associated with prescription records into their corresponding RxRisk classifications. Because VHA pharmacy databases do not utilize NDCs, we used the Multum Lexicon,[28] a drug information database that cross-references NDCs, pharmaceutical trade names and generic drug names, to translate NDCs into drug names and routes of administration, which were subsequently mapped onto VHA Product Names and their corresponding VHA Drug Classes. Drug name alone did not provide an adequate translation because certain agents have different indications depending on route of administration (eg, oral doxepin is primarily used as an antidepressant whereas the topical form is used as an antipruritic agent).

After translating the original GHC RxRisk categories into the VHA formulary as outlined above, we rebuilt the RxRisk category definitions using drug and drug class information to facilitate review and updating. These category definitions were assessed by VHA clinician and pharmacist experts to determine whether the categories were best conceptualized at the drug class level (eg, gout), the multiple drug class level (eg, diabetes), or the drug level (eg, liver failure). Some categories were defined using combinations of strategies (eg, anxiety and tension).

We then solicited clinician and pharmacist input for a review of current and candidate new VHA RxRisk-V categories. As part of this process, we reviewed the current VHA formulary for previously unclassified and wholly new drug classes and supplies (in the VHA system, medical supplies are handled as out-patient prescriptions) to see whether their clinical use permitted the definition of a reasonably specific disease state. Disorders of particular interest were those important to the VHA because they are clinical or research priority areas or are expected to have a significant impact on cost.

Category Stability

Once candidate RxRisk-V categories were constructed, we assessed their reliability by comparing FY 1998 (10/1/1997–9/30/1998) and FY 1999 (10/1/1998–9/30/1999) RxRisk-V patient classification on each RxRisk category using the Kappa statistic. We used year-to-year category stability as a proxy for test-retest reliability, reasoning that because chronic illnesses generally receive ongoing treatment, individuals should continue to be classed into the same category from year to year. Subjects for this analysis were all veteran users of the VHA Northwest Network (the eight VHA facilities in the states of Washington, Oregon, Idaho, and Alaska) during FY 1998 who were classified into at least one RxRisk category and had at least one prescription filled during FY 1999.

Our interpretation of the magnitude of Kappa follows Cicchetti,[29] ie, a Kappa of <0.40 indicates a poor level of clinical significance, Kappa between 0.40 and 0.59 indicates a fair level of agreement, Kappa between 0.60 and 0.74 demonstrates a good level of agreement, and Kappa between 0.75 and 1.00 indicates an excellent level of agreement.

Pharmacy versus Diagnostic Classification

We next judged criterion validity by comparing case identification between the 45 RxRisk-V categories and their ICD-9-CM diagnosis-based counterparts. ICD-9 diagnoses were extracted from both in-patient and out-patient diagnosis files. We judged agreement between pharmacy and diagnostic classes using the Kappa statistic. We did not expect a simple correspondence between RxRisk-V and diagnostic categories, particularly for multiindicated drugs (eg, many cardiovascular medications) and those used for a

minority of individuals with a particular diagnosis (eg, pharmacotherapy for individuals with alcohol dependence). Therefore, we also examined how well the RxRisk categories predicted their corresponding diagnoses as well as how well diagnoses predicted RxRisk categories. One would typically utilize positive predictive value and sensitivity for this purpose. However, the values of those statistics depend strongly on the prevalence of the disorder studied, which makes comparisons across disorders problematic when prevalence rates vary by orders of magnitude (as is the case here). We therefore followed Kraemer's methodology[30] and examined the quality indices of positive predictive value and sensitivity:

$K(0,0) = (SP - Q')/Q = (PPV - P)/P'$ measures how well RxRisk-V categories predict the corresponding diagnoses

$K(1,0) = (SE - Q)/Q' = (NPV - P')/P$ measures how well diagnoses predict the corresponding RxRisk-V categories where Q is the prevalence of the RxRisk category; $Q' = 1 - Q$; P is the prevalence of the ICD-9 diagnosis; $P' = 1 - P$; $NPV = (SP^* P')/[SP^* P' + (1 - SE)^*P]$ is negative predictive value; $PPV = (SE^*P)/[SE^*P + (1 - SP)^*P']$ is positive predictive value. These statistics are bounded by 0 and 1, where 0 indicates performance no better than chance and 1 indicates perfect discrimination. Their magnitude may be interpreted in the same manner as the Kappa statistic. Subjects for this analysis were all FY 1998 veteran users of VHA Northwest Network services.

Concurrent and Future Cost Prediction

Lastly, we examined each category's ability to predict total cost of VHA care, including both in-patient and out-patient care, at the patient level. Complete cost models, including split-half validation and comparison with diagnosis-based instruments, are presented in Sales et al.[27] In this article, we present the parameter estimates for each category of the new RxRisk-V instrument to demonstrate the performance of each category in explaining variance in concurrent and future costs. All parameter estimates show the marginal cost for the classification category. Categories are more credible if their parameter estimates correspond to known utilization patterns and magnitudes.

Both concurrent and prospective models were estimated using 12-month periods of pharmacy and cost data with the patient-year as the unit of analysis. To better compare our results with other models previously described, we followed the convention in the risk adjustment literature by estimating parameters using weighted linear regression.[3,14,16,31] For concurrent models, we used weighted least squares regression to regress FY 1998 costs on FY 1998 explanatory variables—the RxRisk-V and age/gender categories. Prospective models were estimated using FY 1998 explanatory variables to predict FY 1999 costs. Subjects for the concurrent analysis were all FY 1998 veteran users in seven of the eight facilities in the VHA Northwest Network. (One facility was excluded because of site-specific cost data issues.) Subjects for the prospective analysis were all FY 1998 veteran users in seven of the eight facilities in the VHA Northwest Network who were alive at the beginning of FY 1999.

Data Sources

Noncost data were extracted from the VHA Northwest Network Data Warehouse, a relational database containing data from the clinical information systems of each of the eight Northwest Network medical facilities. We extracted patient demographics, out-patient and in-patient diagnoses, and out-patient pharmacy records. Drugs may be identified through several fields in the pharmacy record including VHA National Product Name and site-specific drug name.

Cost data were obtained from the Decision Support System (DSS) from each local site. DSS is a cost accounting system that allows VHA managers and researchers to determine the cost of specific patient care encounters using Relative Value Units (RVUs).[32] Hospital managers in the private sector generally consider RVU-based systems such as DSS to be the most accurate approach to costing currently available.[33,34]

Results

RxRisk-V Development

The RxRisk-V category names, descriptions, and prevalence rates are shown in Table 1. Although most of the original RxRisk categories continued to appear well grounded based on pharmacolog-

ical mechanisms and current clinical indications of the included drugs, our clinical investigators felt that the definitions of the cardiovascular RxRisk categories: cardiac disease, coronary/peripheral vascular and hypertension, should be made more homogeneous. As a result, we eliminated cardiac disease and coronary/peripheral vascular while introducing five additional categories: anticoagulation, antiplatelet agents, arrhythmias, congestive heart failure (CHF)/hypertension, and ischemic heart disease (IHD)/angina.

Two of the original RxRisk categories contained fewer than 300 patients: hyperkalemia (71 patients) and tuberculosis (173 patients). Tuberculosis also did not correspond well to parallel ICD-9 diagnoses and neither category was particularly stable over time (see below). However, because the medications used to define tuberculosis have few other clinical indications, we felt that the poor correspondence with diagnostic data was most likely related to diagnostic coding issues (see Examples below). In addition, because both conditions are treated in a time-limited fashion, we would not expect high levels of year-year stability. We therefore elected to include both of these classes in the RxRisk-V model, in part for consistency with the original GHC instrument.

Examining previously unclassified drugs, we constructed alcohol dependence, smoking cessation, benign prostatic hypertrophy, psoriasis, allergies, and malnutrition as indicated in Table 1. Newer drug classes provided the basis for osteoporosis/Pagets disease, dementia, hepatitis C, and migraine. Our review of supplies on the VHA formulary suggested several additional potential markers for high-cost individuals including ostomy, neurogenic bladder, and urinary incontinence.

With two exceptions, all new RxRisk-V categories included at least 300 patients. Only 242 individuals classed into dementia; however, we elected to maintain that category because (1) the drugs were highly specific to the diagnosis; (2) the drugs were likely to be prescribed in increasing numbers in future years; and (3) aging of the VHA population makes dementia a clinical and cost priority. Only eight individuals classed into hepatitis C. However, given the high priority in the VHA of Hepatitis C treatment and research[35] as well as more recent data that prescription rates are accelerating rapidly (eg, 103 individuals classed into hepatitis C in 2000), we elected to maintain this category for future research.

Category Stability

This analysis included 76,772 veterans who were classified into at least one RxRisk-V category in FY 1998 and had at least one prescription fill during FY 1999. Subjects had a mean age of 59.1 years (SD 14.2), and 94.1% were male. Eligible subjects classified into an average of 3.86 (SD 2.40) RxRisk-V categories in FY 1998 and 4.19 (SD 2.62) RxRisk-V categories in 1999.

Of the 45 RxRisk-V categories, 33 had Kappas indicating a good to excellent level of consistency ($K > 0.6$) with HIV (Kappa = 0.92), diabetes (Kappa = 0.91), and hypothyroidism (Kappa = 0.90) all having Kappas ≥ 0.90. Only tuberculosis (Kappa = 0.36), smoking cessation (Kappa = 0.27), hyperkalemia (Kappa = 0.24), and hepatitis C (Kappa = 0.10) had Kappas <0.40; all are conditions treated in time-limited fashions. Table 1 provides category-by-category details.

Pharmacy versus Diagnostic Classification

This analysis included 126,075 veterans who were FY 1998 system users. These individuals had a mean age of 56.9 years (SD 15.2) and 93.9% were male. Subjects averaged 2.61 (SD 2.64) RxRisk-V categories per individual, with 70.5% classified into at least one RxRisk-V category. The ICD-9 diagnostic definitions we utilized for this work are in Table 2. Table 3 details the relationship between RxRisk-V and ICD-9 diagnostic classification including simple concordance (Kappa), how well RxRisk predicts diagnostic category [K(0,0)], and how well diagnosis predicts RxRisk category [K(1,0)]. Detailed results of the comparison are presented in Appendix 1.

Risk categories with Kappas >0.6 (indicating good to excellent concordance between pharmacy and ICD-9 strategies) included diabetes, HIV, glaucoma, hypothyroidism, reactive airway disease, Parkinsons disease, psychotic illness, hyperlipidemia, and IHD/angina. The 14 categories with Kappas between 0.4 and 0.6 (indicating fair agreement) were gout, depression, transplant, gastric acid disorder, CHF/hypertension, IHD/hypertension, pain/inflammation, anticoagulation, benign prostatic hypertrophy, migraine, osteoporosis/Pagets, neurogenic bladder, bipolar disorder, arrhythmias, and neurogenic bladder. All 22 remaining categories had Kappas <0.4.

A low Kappa does not mean that there is no important or interpretable relationship between

RxRisk category and diagnosis. CHF/hypertension and alcohol dependence both had Kappas <0.4. However, a diagnosis of CHF very strongly predicted an RxRisk classification of CHF/hypertension [K(1,0)=0.82] whereas an RxRisk classification of alcohol dependence very strongly predicted the presence of an alcohol use disorder diagnosis [K(0,0) = 0.90].

More detailed examination of categories with intermediate Kappas revealed different patterns of overlap. Neither epilepsy or ostomy had a particularly high Kappa (0.24 and 0.34, respectively), but their high K(1,0) scores (0.88 and 0.84, respectively) indicated that, although the RxRisk category is a poor predictor of having a corresponding same-year ICD-9 diagnosis, if the ICD diagnosis is present, in most cases a corresponding prescription is found. This situation could arise either through "under-coding" of ICD diagnoses (most likely for the Ostomy case) or if the drug class has other indications (eg, chronic pain or bipolar disorder for some anticonvulsants).

The reverse holds for dementia. Although the ICD-9 diagnosis is a poor predictor of having an RxRisk dementia prescription, the presence of a prescription strongly predicts the corresponding ICD-9 diagnosis. This situation could arise if medications not mapped to that RxRisk category are also used to treat the disorder (eg, divalproex and carbamazepine used to treat bipolar disorder) or if a substantial number of individuals receive the diagnosis but do not receive drug treatment (most likely the reason for the dementia results).

The correspondence between drug (or supply) and coded diagnosis was uniformly poor (low Kappa, K(0,0), and K(1,0)) for only two of the new categories: urinary incontinence and malnutrition. Much of the poor fit could be caused by underreporting of diagnosis: these conditions are frequently seen in the clinical context of multiple and more severe other medical conditions and the year-to-year stability is reasonable for both pharmacy classes (urinary incontinence: Kappa = 0.63; malnutrition: Kappa = 0.54), indicating ongoing treatment. Lastly, these categories significantly predicted both concurrent and prospective cost (see below and Table 1). Consequently, we included both categories in the RxRisk-V but continue to investigate their utility.

Concurrent and Future Cost Prediction

The concurrent analysis included 121,067 veterans seen in seven of eight facilities of the Northwest network during FY 1998, while the prospective analysis included 117,936 individuals from the first group who were alive at the beginning of FY 1999. At the beginning of FY 1998, individuals had a mean age of 56.8 years (SD 15.3) and 93.8% were male. Subjects averaged 2.60 (SD 2.64) RxRisk-V categories per individual, with 70.1% classified into at least one RxRisk-V category. Table 1 provides category-by-category details.

The concurrent model had an $R^2 = 0.179$ ($F = 470$; $P < 0.0001$). The coefficients or parameter estimates indicate marginal costs of each RxRisk-V category. Categories with estimated coefficients >$10,000 included end stage renal disease, hyperkalemia, transplant, malnutrition, and tuberculosis. The prospective model had an $R^2 = 0.106$ ($F = 250$; $P < 0.0001$). Categories with estimated coefficients >$10,000 included end stage renal disease and hyperkalemia. In general, the estimated coefficients for both concurrent and prospective models appeared to be clinically rational, with the categories at highest cost being either intrinsically expensive (eg, end stage renal disease) or typically found in concert with other high-cost disorders (eg, malnutrition). Disorders usually requiring fewer resources to treat (eg, gout, psoriasis, hypertension) had lower concurrent and prospective estimated cost coefficients.

Substantially negative coefficients, such as the estimated concurrent cost coefficient of −$1,458 for inflammatory bowel syndrome (IBS), were unexpected. However, further analysis revealed that individuals in the IBS category averaged 1.72 (95% CI 1.52–1.91) RxRisk-V categories more than those outside the class, excluding IBS itself. In addition, three categories with some of the highest marginal costs, malnutrition, transplant, and ostomy, were among the conditions substantially over-represented ($P < 0.001$) among individuals with IBS. A similar analysis for hyperlipidemia (estimated concurrent cost coefficient of −$650) revealed that individuals in that RxRisk category were older (63.0 years vs. 55.9 years; $t = -72.8$; $P < 0.001$) and were more highly comorbid, with 2.27 (95% CI 2.23–2.31) additional RxRisk-V categories, excluding hyperlipidemia. It is therefore possible that the negative coefficients for IBS and hyperlipidemia served to offset the multiple additional comorbidities seen in these individuals.

TABLE 1. RxRisk-V Category Descriptions, Prevalence, Stability, and Cost Coefficient Estimates

RxRisk-V class	Drug classes/drugs	Prevalence N	Prevalence (%)	Stability†	Concurrent Cost′	Prospective Cost″
Total patients		126,075		76,772	121,067	117,936
Alcohol dependence	disulfiram, naltrexone	1,115	0.9%	0.54	6,397***	3,147***
Allergies	antihistamines (except hydroxyzine and diphenhhydramine), nasal anti-inflammatories	10,903	8.6%	0.58	−64	564***
Anticoagulation	anticoagulants	4,826	3.8%	0.80	3,709***	2,178***
Antiplatelet agents	antiplatelet agents	1,102	0.9%	0.55	4,275***	1,685***
Anxiety and tension	anxiolytics (benzodiazepine and barbiturate)	10,598	8.4%	0.65	2,357***	1,268***
Arrhythmias	antiarrhythmics, digoxin	6,380	5.1%	0.84	1,837***	1,584***
Benign prostatic hypertrophy	alpha blockers	9,346	7.4%	0.73	153	416**
Bipolar disorder	lithium	1,526	1.2%	0.78	843***	1,477***
CHF/hypertension	loop diuretics, ACE, and angiotensin II inhibitors	26,064	20.7%	0.81	625***	996***
Dementia	donepezil, tacrine	245	0.2%	0.71	−1,188	−509
Depression	antidepressants	24,023	19.1%	0.71	1,699***	1,561***
Diabetes	insulins, oral hypoglycemics	12,956	10.3%	0.91	1,550***	2,189***
End stage renal disease	alpha erythopoetin, calcifediol, calcitriol, sevelamer	420	0.3%	0.65	17,325***	17,083***
Epilepsy	anticonvulsants	6,582	5.2%	0.79	2,698***	1,750***
Gastric acid disorder	H_2 blockers, proton pump inhibitors	27,966	22.2%	0.72	1,897***	1,408***
Glaucoma	topical (ophthalmic) antiglaucoma agents	3,053	2.4%	0.85	530**	1,543***
Gout	antigout agents	3,650	2.9%	0.82	965***	600**
Hepatitis C	interferon/ribavirin combinations	4	0.0%	0.10	(N/A)	(N/A)
HIV	anti-HIV antivirals	361	0.3%	0.92	5,831***	5,470***
Hyperkalemia	sodium polystyrene sulfonate	71	0.1%	0.24	11,083***	12,536***
Hyperlipidemia	antilipemic agents	16,726	13.3%	0.79	−650***	−360**
Hypertension	thiazides, potassium-sparing agents, combination antihypertensives, other antihypertensives (eg, clonidine, hydralazine)	13,075	10.4%	0.71	647***	1,128***
Hypothyroidism	thyroid replacements	4,942	3.9%	0.90	339*	649**
IHD/angina	nitrates	13,866	11.0%	0.73	1,949***	1,545***
IHD/hypertension	beta blockers, calcium channel blockers	28,059	22.3%	0.80	777***	638***
Inflammatory bowel syndrome	IBS-specific drugs, rectal anti-inflammatories	854	0.7%	0.77	−1,458***	−1,532**
Liver failure	lactulose	649	0.5%	0.48	9,631***	8,441***
Malignancies	antineoplastic agents (all oral and system agents but excluding topicals)	1,705	1.4%	0.72	3,519***	4,795***
Malnutrition	enteral nutritional supplements	1,111	0.9%	0.54	12,149***	5,461***
Migraine	antimigraine medications (ergots, triptans, methysergide)	1,249	1.0%	0.67	−734**	−202
Neurogenic bladder	urinary catheters (supplies)	1,026	0.8%	0.67	5,088***	7,679***
Osteoporosis/Paget's	alendronate, etidronate	643	0.5%	0.68	738	1,427**
Ostomy	colostomy and urostomy supplies	1,464	1.2%	0.74	6,780***	4,944***
Pain	opiate-containing pain medications	23,285	18.5%	0.44	4,023***	2,436***
Pain/Inflammation	nonsteroidal anti-inflammatory drugs	34,591	27.4%	0.57	152*	271**
Pancreatic insufficiency	pancreatic exocrine enzyme replacements	417	0.3%	0.72	5,253***	1,908**

(*Continued*)

TABLE 1. (Continued)

RxRisk-V class	Drug classes/drugs	Prevalence N	Prevalence (%)	Stability†	Concurrent Prospective Cost′	Concurrent Prospective Cost″
Parkinson disease	antiparkinsonian agents	928	0.7%	0.77	468	729
Psoriasis	systemic and topical antipsoriatics	1,541	1.2%	0.66	1,019***	991**
Psychotic illness	antipsychotics	5,193	4.1%	0.80	5,517***	3,853***
Reactive airway disease	inhaled bronchodilators, leukotriene inhibitors	16,326	12.9%	0.74	1,311***	1,423***
Smoking cessation	nicotine, Zyban	2,474	2.0%	0.27	2,724***	1,761***
Steroid-responsive conditions	glucocorticoids (steroids)	5,850	4.6%	0.49	2,973***	2,873***
Transplant	immune suppressants	441	0.3%	0.84	10,392***	4,060***
Tuberculosis	anti-tubercular agents	173	0.1%	0.36	11,548**	6,545***
Urinary incontinence	diapers and pads (supplies)	1,312	1.0%	0.63	6,407***	5,362***
Male 0–34	(Reference Category)	8,249	6.5%		(0)	(0)
Male 35–44		14,872	11.8%		163	592**
Male 45–54		31,668	25.1%		139	676***
Male 55–59		9,897	7.9%		146	698**
Male 60–64		10,193	8.1%		288*	1,176***
Male 65–69		12,456	9.9%		240	931***
Male 70–74		12,710	10.1%		273	1,207***
Male 75–79		11,896	9.4%		421**	1,293***
Male 80+		6,440	5.1%		1,062***	1,922***
Female 0–34		1,913	1.5%		17	39
Female 35–44		2,182	1.7%		−93	455
Female 45–64		2,244	1.8%		152	1,177***
Female 65+		1,355	1.1%		88	592
Intercept					363***	341*

†Kappa. ***$P < 0.01$; **$0.05 < P < 0.01$; *$0.10 < P < 0.05$; N/A indicates not included in regression models.
′Concurrent cost indicates FY98 RxRisk–V categories predicting FY98 total pharmacy cost.
″Prospective cost indicates FY98 RxRisk–V categories predicting FY99 total pharmacy cost.

As a sensitivity analysis, we also recalculated the coefficients for the concurrent and prospective cost models after excluding categories with negative coefficients. We found no substantive difference in the coefficients calculated.

Examples

Based on these results, we feel that pharmacy- and diagnosis-based case-detection strategies are best viewed as complementary rather than alternate strategies for defining patient case-mix. Pharmacy- and diagnosis-based strategies may correspond in a number of ways:

Diabetes is an example of a condition in which pharmacy- and diagnosis-based case-detection strategies highly overlap, as can be seen by the high Kappa, K(0,0) and K(1,0). Clinically, this makes sense, because diabetes is the sole clinical indication for the prescription of oral hypoglycemics and insulin. Some individuals with diabetes are controlled by diet alone, which likely accounts for the observation that medication case-detection strategies are more specific than sensitive [K(1,0)>K(0,0)] when compared with criterion ICD-9 diagnoses. Additionally, treatment for diabetes is chronic, reflected in high year-to-year stability with a Kappa of 0.90. For conditions similar to diabetes, pharmacy- and diagnosis-based case-mix instruments might be viewed as alternative strategies for capturing the same (or very similar) patient populations.

Alcohol dependence is an example of a condition in which pharmacy-based case-detection strategies find a subset of the population defined by ICD-9 diagnoses. Because these medications are rarely used outside the specialist treatment setting, one could hypothesize that individuals

TABLE 2. Diagnostic Categories with ICD-9 Code Mappings

Diagnostic Category	ICD-9 Codes
Alcohol use disorders	303*, 305*
Allergies	477*
Anxiety disorders	300*, 300.2*
Arrhythmias (all)	427*
Atrial fibrillation/flutter	427.3*
Benign prostatic hypertrophy (BPH)	600
Bipolar disorder	296.00–296.16, 296.40–296.81, 296.89
Malignancies	140–208.91
Cerebrovascular disease (CVD)	433–438.9
Congestive heart failure (CHF)	428*
Chronic renal failure	585, E879.1, V45.1, V56.*
Dementia	290.40, 290.0–290.10, 331.0–331.1
Depression	296.20–296.36, 311
Diabetes	250*
Glaucoma	365.10–365.9
Gout	274*
Hepatic coma	572.2
Hepatitis C	070.41, 070.44, 070.51, 070.54
HIV	042–044.9, V08
Hyperlipedemia	272–272.4
Hyperpotassemia	276.7
Hypertension	401–405.99
Hypothyroidism	243–244.9
Ischemic heart disease (IHD)	410–414.9
Ostomy	V44.2–V44.3, V55.2–V55.3
Malnutrition	260–263.9
Inflammatory bowel syndrome	555–556.9, 564.1
Migraine	346*
Pain (musculoskeletal)	714–729.9
Pancreatitis–chronic	577.1
Parkinson disease	332
Psoriasis	696–696.1
Psychotic disorders	295*, 296.9–298.9
Quadraparesis	344–344.1
Reactive airway disease	492*, 493*, 496*
Tobacco use disorder	305.1*
Transplant	996.8*, V42.0–V42.1, V42.6–V42.9, V43.2
Tuberculosis	01*
Urinary incontinence	788.3*

* indicates "wildcard", ie, matches any numbers.

found using the RxRisk-V will be more costly than those with diagnosis alone because of costs generated by specialist care as well as the level of severity needed to generate specialist referral. Partially supporting this notion are the positive and substantial concurrent and prospective cost estimates ($6,397 and $3,147, respectively). For alcohol dependence and similar conditions, pharmacy-based case-mix

TABLE 3. Comparison of RxRisk-V and ICD-9 Diagnostic Classification

K(1,0): / K(0,0):	**K(1,0) is ≥0.6** *Chart diagnosis is a good-excellent predictor of prescription written within this RxRisk category*	**K(1,0) is 0.4–0.6** *Chart diagnosis is a fair predictor of prescription written within this RxRisk category*	**K(1,0) is <0.4** *Chart diagnosis is a poor predictor of prescription written within this RxRisk category*
K(0,0) is ≥0.6 *RxRisk category is a good-excellent predictor of chart diagnosis*	**Diabetes** **HIV** **Glaucoma** **Hypothyroidism** **Reactive airway disease** **Parkinson disease** **Psychotic illness** **Hyperlipedemia**	**IHD/Angina** *CHF/hypertension–hypertension* *IHD/hypertension–hypertension*	*Osteoporosis/Pagets* *Bipolar disorder* Hypertension Smoking cessation Dementia Malignancies Alcohol dependence
K(0,0) is 0.4–0.6 *RxRisk category is a fair predictor of chart diagnosis*	*Gout* *Depression* *Transplant*	*Pain/inflammation* *Anticoagulation* *Migraine* *Arrhythmias*	Inflammatory bowel syndrome Pain Steroid-responsive conditions End stage renal disease Hyperkalemia Antiplatelet agents Hepatitis C
K(0,0) is <0.4 *RxRisk category is a poor predictor of chart diagnosis*	*Gastric Acid Disorder* Ostomy Allergies CHF/hypertension–CHF Epilepsy	*Benign prostatic hypertrophy* *Neurogenic bladder* *IHD/hypertension–IHD* Psoriasis Pancreatic insufficiency Anxiety and tension Liver failure	Tuberculosis Urinary incontinence Malnutrition

Bold indicates Kappa ≥0.6 or good–excellent agreement; Italics indicate Kappa 0.4–0.6 or fair agreement; K(0,0): the quality of the specificity; K(1,0): the quality of the sensitivity

instruments are likely detecting a subset of the diagnostic population whose particular characteristics would bear further scrutiny.

In contrast, epilepsy is an example of a condition where diagnosis-based case-detection strategies find a subset of a larger medication-defined population. Because clinical practice guidelines suggest that individuals should be seizure-free for 2 to 5 years before attempting to discontinue anticonvulsants,[36] the clinical process reflected here might be that many individuals with epilepsy have their medications "routinely" renewed with no diagnosis coded. If this is the case, then individuals found using the RxRisk-V would be a more complete accounting of individuals with epilepsy than those found using diagnosis alone. For conditions similar to epilepsy, pharmacy-based case-detection algorithms might more completely capture the total population with a given condition and, therefore, provide a better index of the true prevalence of the condition in the population examined.

Lastly, in the case of tuberculosis, pharmacy- and diagnosis-based case-detection strategies find largely different, albeit quite costly, individuals. In general, this disagreement between pharmacy and diagnostic data may occur when a particular drug class cannot be consistently associated with a

specific diagnosis or when the treated condition is not well captured in the diagnosis files. In the case of tuberculosis, because infected individuals are treated for 6 to 12 months, the diagnosis may not be carried forward even when the treatment continues. In addition, antitubercular therapy may be instituted as preventive therapy[37]; these individuals would be on the drug but not carry a diagnosis. For conditions similar to tuberculosis, close study of the correspondence between pharmacy- and diagnosis-based case-detection algorithms would be required to fully understand the interplay.

Discussion

We have successfully adapted the RxRisk to the VHA population and demonstrated that RxRisk-V classes are stable over time, valid against ICD-9 criterion diagnoses, and possess rational concurrent and prospective marginal cost coefficients. For individuals seen in the VHA Northwest Network during Fiscal Year 1998, the RxRisk-V classifies 70.5% into an average of 2.61 categories.

Because our focus was on *chronic* disease, we were encouraged that more than 2/3 of the RxRisk-V categories had year-to-year Kappa coefficients in the good-excellent range. This stability broadens the utility of the instrument for case-mix assessment. For example, one might use the RxRisk-V to predict which individuals will require ongoing treatment for certain conditions to track these patients in appropriate disease management programs. Our findings also indicate that pharmacy- and diagnosis-based strategies for disease identification have a complex relationship, which would benefit from closer examination and analysis tailored to the specific research or management application.

Of the 44 RxRisk-V categories (excluding hepatitis C), 39 had statistically significant coefficients in the concurrent cost model and 41 had statistically significant cost coefficients in the prospective model. Categories without significant coefficients (eg, dementia) might still be important for noncost case-mix applications and, therefore, we have elected to keep them in the overall model. Also noteworthy is that some categories with the lowest year-to-year stability and correspondence to ICD-9 diagnoses (eg, tuberculosis, malnutrition, and hyperkalemia) had among the highest marginal cost coefficients, whereas some categories with high year-to-year stability and correspondence to ICD-9 diagnoses (eg, Parkinson disease and hypothyroidism) had insignificant or small marginal cost coefficients. Results such as these indicate the importance of assessing new case-mix instruments using a number of methods.

One limitation is that this study used a VHA population from one region of the United States. The VHA population is known to be sicker than the general patient population,[23,38] and the VHA Northwest Network population is younger than the rest of the VHA population; in addition, the small area variations literature indicates that patient treatment can differ markedly from region to region. Therefore, generalization to other patient populations requires additional study.

Another limitation is that we were unable to ascertain medical service or pharmacy utilization of other health care systems. VHA patients are known to receive services from other health care providers.[39–41] Unfortunately, to date there are no studies that examine dual system utilization in the VHA population as a whole. Furthermore, although anecdotal reports indicate that a subset of patients seek VHA care primarily to have prescriptions obtained elsewhere rewritten by a VHA providers (because VHA pharmacies may not fill prescriptions written by non-VHA providers), there are no data currently available that examine the scope of this "inflow" of prescriptions. Outside medical service or pharmacy utilization would produce incomplete case ascertainment, downwardly biasing the Kappa statistics that examine category stability and that compare RxRisk-V categories with their ICD-9-CM diagnosis-based counterparts. In addition, coefficients for both concurrent and future cost prediction would be similarly biased. However, if patients consistently obtain their medications from the same source (ie, VHA or non-VHA pharmacies), then dual system utilization might not materially affect the year-to-year stability of RxRisk-V category assignments.

An additional issue is that approximately 37,000 of the 121,000 FY 1998 veterans seen were not classified by the RxRisk-V. Of these, 30,710 (82.6%) had no prescriptions filled in the year and thus would be unclassifiable by any pharmacy-based instrument; the others had fills of non-RxRisk-V drugs such as antibiotics. Although these individuals would be assigned to age and sex categories when using the RxRisk-V in a case-mix adjustment application, lack of drug information clearly places limits on what any pharmacy-based instrument could learn about this subpopulation.

Lastly, new medications are continually being introduced, requiring ongoing monitoring and updating of any pharmacy-based instrument. This problem is somewhat reduced (although not eliminated) by coding category definitions at the drug *class* level, because new agents in a previously utilized class may be added with minimal additional review.

Our goal was to tailor an instrument that has been used successfully in general health care populations to a specialized patient population segment—veterans using the VHA system. The result of this process is a reliable and valid instrument for administrators to describe and understand better the chronic disease burden of their clinic and medical center populations. We have attempted to report category definitions and psychometric properties with sufficient detail so that these findings may be validated in other populations as pharmacy-based case-mix instruments receive increasing attention and researchers explore population-specific case-mix and risk adjustment instruments.

References

1. **Fetter R.** DRG Refinement with Diagnosis Specific Cormorbidity: A Synthesis of Current Approaches to Patient Classification. New Haven: Yale University Press; 1989.

2. **Ash A, Porell F, Gruenberg L, et al.** Adjusting Medicare capitation payments using prior hospitalization data. Health Care Financ Rev 1989;10:17–29.

3. **Weiner JP, Starfield BH, Steinwachs DM, et al.** Development and application of a population-oriented measure of ambulatory care case-mix. Med Care 1991;29:452–472.

4. **Dunn D, Rosenblatt A, Taira D.** A Comparative Analysis of Methods of Health Risk Assessment. SOA Monograph M HB96–1. Schaumberg, IL: Society of Actuaries; 1995.

5. **Medicare Payment Advisory Commission.** Report to the Congress: Medicare Payment Policy. Washington, DC; 1998.

6. **Corn RF.** The sensitivity of prospective hospital reimbursement to errors in patient data. Inquiry 1981;18:351–360.

7. **Demlo LK, Campbell PM.** Improving hospital discharge data: lessons from the National Hospital Discharge Survey. Med Care 1981;19:1030–1040.

8. **Institute of Medicine.** Reliability of Hospital Discharge Abstracts. Washington, DC: National Academy of Sciences; 1977.

9. **Institute of Medicine.** Reliability of Hospital Discharge Abstracts. Washington, DC: National Academy of Sciences; 1980.

10. **Kashner TM.** Agreement between administrative files and written medical records: a case of the Department of Veterans Affairs. Med Care 1998; 36:1324–1336.

11. **Szeto H, Goldstein M.** Accuracy of Computer Identified Diagnoses in a VA General Medicine Clinic 1999. Presented at the 17th VA Health Service Research and Development Service Annual Meeting, Washington, DC; February 24–26,1999.

12. **Hughes JS, Ash AS.** Reliability of Risk-Adjustment Models. In: Iezzoni L, ed. Risk Adjustment for Measuring Healthcare Outcomes. 2nd ed. Chicago: Health Administration Press; 1997:365–390.

13. **Von Korff M, Wagner EH, Saunders K.** A chronic disease score from automated pharmacy data. J Clin Epidemiol 1992;45:197–203.

14. **Johnson RE, Hornbrook MC, Nichols GA.** Replicating the chronic disease score (CDS) from automated pharmacy data. J Clin Epidemiol 1994;47:1191–1199.

15. **Roblin DW.** Patient case mix measurement using outpatient drug dispense data. Manag Care Q 1994;2:38–47.

16. **Clark DO, Von Korff M, Saunders K, et al.** A chronic disease score with empirically derived weights. Med Care 1995;33:783–795.

17. **Roblin DW.** Physician profiling using outpatient pharmacy data as a source for case mix measurement and risk adjustment. J Ambul Care Manage 1998;21:68–84.

18. **Lamers LM.** Pharmacy costs groups: a risk-adjuster for capitation payments based on the use of prescribed drugs. Med Care 1999;37:824–830.

19. **Hornbrook MC, Goodman MJ.** Health plan case mix: definition, measurement, and use. Adv Health Econ Health Serv Res 1991;12:111–148.

20. **Fishman PA, Shay DK.** Development and estimation of a pediatric chronic disease score using automated pharmacy data. Med Care 1999;37:874–883.

21. **Fishman PA, Goodman M, Hornbrook M, et al.** Risk adjustment using automated pharmacy data: the RxRisk model. Med Care 2003;41:84-99.

22. **United States General Accounting Office.** VA Health Care: More National Action Needed to Reduce Waiting Times, but Some Clinics Have Made Progress. Report GAO-01–953. Washington, D.C.: Government Printing Office; 2001.

23. **Kazis LE, Miller DR, Clark J, et al.** Health-related quality of life in patients served by the Department of Veterans Affairs: results from the Veterans Health Study. Arch Intern Med 1998; 158:626–632.

24. **Agha Z, Lofgren R, VanRuiswyk J, et al.** Are patients at Veterans Affairs medical centers sicker? A comparative analysis of health status and medical resource use. Arch Intern Med 2000;160:3252–3257.

25. **Wray NP, Hollingsworth JC, Peterson NJ, et al.** Case-mix adjustment using administrative databases: a paradigm to guide future research. Med Care Res Rev 1997;54:326–356.

26. **DeLong ER, Peterson ED, DeLong DM, et al.** Comparing risk-adjustment methods for provider profiling. Stat Med 1997;16:2645–2664.

27. **Sales AE, Liu CF, Sloan KL, et al.** Predicting costs of care using a pharmacy-based measure: risk adjustment in a veteran population. Med Care 2003;41:753–760.

28. **Multum Lexicon.** Multum Information Services; 2001.

29. **Cicchetti DV.** Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychol Assess 1994;6:284–290.

30. **Kraemer H.** Evaluating Medical Tests: Objective and Quantitative Guidelines. Newbury Park: Sage Publications; 1992.

31. **Ellis RP, Ash A.** Refinements to the Diagnostic Cost Group (DCG) model. Inquiry 1995;32:418–429.

32. **Barnett PC, Rodgers JH.** Use of the Decision Support System for VA Cost-Effectiveness Research. MDRC Working Meeting on Cost Studies; 1998.

33. **Ashby J.** The Accuracy of Cost Measures Derived from Medicare Cost Report Data. Intramural Report I-93–01. Washington, DC: Prospective Payment Assessment Commission; 1993.

34. **Shwartz M, Young DW, Siegrist R.** The ratio of costs to charges: how good a basis for estimating costs? Inquiry 1995;32:476–481.

35. **United States General Accounting Office.** Standards and Accountability Could Improve Hepatitis C Screening and Testing Performance. Report GAO-01–807T. Washington, D.C.: Government Printing Office; 2001.

36. **American Academy of Neurology.** Practice parameter: a guideline for discontinuing antiepileptic drugs in seizure-free patients–summary statement. Neurology 1996;47:600–602.

37. **Bass JB Jr., Farer LS, Hopewell PC, et al.** Treatment of tuberculosis and tuberculosis infection in adults and children. Am J Respir Crit Care Med 1994;149:1359–1374.

38. **Kazis LE, Ren XS, Lee A, et al.** Health status in VA patients: results from the Veterans Health Study. Am J Med Qual 1999;14:28–38.

39. **Passman L, Garcia R, Campbell L, et al.** Elderly veterans receiving care at a Veterans Affairs Medical Center while enrolled in Medicare-financed HMOs. Is the taxpayer paying twice? J Gen Intern Med 1997;12:247–249.

40. **Wright S, Petersen L, Lamkin R, et al.** Increasing use of Medicare services by veterans with acute myocardial infarction. Med Care 1999;37:529–537.

41. **Wright SM, Petersen LA.** Enrollment in the Medicare program among veterans turning age 65. Paper presented at the 18th VA Health Service Resesearch and Development Service Annual Meeting, Washington, DC; March 24, 2000.

APPENDIX 1. Details of Comparison Between RxRisk-V and ICD-9 Diagnostic Categories: K(0,0) is the Quality of the Specificity; K(1,0) is the Quality of the Sensitivity

RxRisk-V Category	ICD-9 Diagnosis Category	Kappa*	K(0,0)	K(1,0)
Alcohol dependence	Alcohol use disorders	0.14	**0.90**	0.08
Allergies	Allergies	0.25	0.15	**0.72**
Anticoagulation	Atrial fibrillation/flutter	*0.45*	*0.42*	*0.48*
Antiplatelet agents	Cerebrovascular disease	0.15	*0.49*	0.09
Anxiety and tension	Anxiety disorders	0.25	0.18	*0.41*
Arrhythmias	Arrhythmias	*0.41*	*0.41*	*0.41*
Benign prostatic hypertrophy	Benign prostatic hypertrophy	*0.44*	0.38	*0.52*
Bipolar disorder	Bipolar disorder	*0.42*	**0.69**	0.31
CHF/hypertension	Congestive heart failure (CHF)	0.25	0.15	**0.82**
	Hypertension	*0.51*	**0.64**	*0.42*
Dementia	Dementia	0.23	**0.71**	0.14
Depression	Depression	0.52	*0.41*	**0.71**
Diabetes	Diabetes	**0.83**	**0.95**	**0.74**
End stage renal disease	Chronic renal failure	0.22	*0.40*	0.15
Epilepsy	Epilepsy	0.24	0.14	**0.88**
Gastric acid disorder	Gastric acid disorders	*0.51*	0.37	**0.82**
Glaucoma	Glaucoma	**0.74**	**0.70**	**0.78**
Gout	Gout	*0.55*	*0.50*	**0.62**
Hepatitis C	Hepatitis C	0.00	*0.49*	0.00
HIV	HIV	**0.78**	**0.94**	**0.67**
Hyperkalemia	Hyperpotassemia	0.17	*0.44*	0.10
Hyperlipidemia	Hyperlipidemia	**0.64**	**0.67**	**0.60**
Hypertension	Hypertension	0.32	**0.71**	0.20
Hypothyroidism	Hypothyroidism	**0.73**	**0.63**	**0.88**
IHD/Angina	Ischemic heart disease (IHD)	**0.62**	**0.72**	*0.54*
IHD/Hypertension	Hypertension–Ischemic Heart Disease (IHD)	0.50	**0.60**	*0.43*
		0.40	0.31	*0.54*
Inflammatory bowel syndrome	Inflammatory bowel syndrome	0.31	*0.43*	0.25
Liver failure	Hepatic coma	0.21	0.13	**0.60**
Malignancies	Malignancies	0.17	**0.61**	0.10
Malnutrition	Malnutrition	0.11	0.09	0.13
Migraine	Migraine	*0.44*	*0.45*	*0.43*
Neurogenic bladder	Neurogenic bladder	*0.40*	0.35	*0.47*
	Quadraparesis	*0.43*	0.37	*0.52*
Osteoporosis/pagets	Osteoporosis/Pagets	*0.44*	**0.62**	0.34
Ostomy	Ostomy	0.34	0.21	**0.84**
Pain	Pain (musculoskeletal)	0.29	*0.45*	0.21
Pain/inflammation	Pain (musculoskeletal)	*0.45*	*0.51*	*0.40*
Pancreatic insufficiency	Pancreatitis–chronic	0.37	0.28	*0.54*
Parkinson disease	Parkinson disease	**0.64**	**0.60**	**0.69**
Psoriasis	Psoriasis	0.39	0.37	*0.40*
Psychotic illness	Psychotic disorders	**0.64**	**0.62**	**0.66**
Reactive airway disease	Reactive airway disease	**0.66**	**0.62**	**0.71**
Smoking cessation	Tobacco use disorder	0.27	**0.68**	0.17
Steroid-responsive conditions	Reactive airway disease	0.23	*0.42*	0.16
Transplant	Transplant	*0.51*	*0.45*	**0.61**
Tuberculosis	Tuberculosis	0.22	0.27	0.19
Urinary incontinence	Urinary incontinence	0.20	0.20	0.20

*Bold type indicates Kappa ≥0.6 or good-excellent agreement; Italic type indicates Kappa 0.4–0.6 or fair agreement; K(0,0): the Quality of the Specificity; K(1,0): the quality of the Sensitivity