Research and Applications

Comparing ascertainment of chronic condition status with problem lists versus encounter diagnoses from electronic health records

Robert W. Voss [1], **Teresa D. Schmidt** [1], **Nicole Weiskopf** [2], **Miguel Marino**[3], **David A. Dorr** [2], **Nathalie Huguet**[3], **Nate Warren**[1], **Steele Valenzuela**[3], **Jean O'Malley**[1], and **Ana R. Quiñones** [3,4]

[1]OCHIN, Inc, Portland, Oregon, USA, [2]Department of Medical Informatics and Clinical Epidemiology, Oregon Health & Science University, Portland, Oregon, USA, [3]Department of Family Medicine, Oregon Health & Science University, Portland, Oregon, USA and [4]OHSU-PSU School of Public Health, Oregon Health & Science University, Portland, Oregon, USA

Corresponding Author: Ana R. Quiñones, Department of Family Medicine, Oregon Health & Science University, 3181 S.W. Sam Jackson Park Rd., FM, Portland, OR 97239, USA; quinones@ohsu.edu

ABSTRACT

Objective: To assess and compare electronic health record (EHR) documentation of chronic disease in problem lists and encounter diagnosis records among Community Health Center (CHC) patients.

Materials and methods: We assessed patient EHR data in a large clinical research network during 2012–2019. We included CHCs who provided outpatient, older adult primary care to patients age $\geq$45 years, with $\geq$2 office visits during the study. Our study sample included 1 180 290 patients from 545 CHCs across 22 states. We used diagnosis codes from 39 Chronic Condition Warehouse algorithms to identify chronic conditions from encounter diagnoses only and compared against problem list records. We measured correspondence including agreement, kappa, prevalence index, bias index, and prevalence-adjusted bias-adjusted kappa.

Results: Overlap of encounter diagnosis and problem list ascertainment was 59.4% among chronic conditions identified, with 12.2% of conditions identified only in encounters and 28.4% identified only in problem lists. Rates of coidentification varied by condition from 7.1% to 84.4%. Greatest agreement was found in diabetes (84.4%), HIV (78.1%), and hypertension (74.7%). Sixteen conditions had <50% agreement, including cancers and substance use disorders. Overlap for mental health conditions ranged from 47.4% for anxiety to 59.8% for depression.

Discussion: Agreement between the 2 sources varied substantially. Conditions requiring regular management in primary care settings may have a higher agreement than those diagnosed and treated in specialty care.

Conclusion: Relying on EHR encounter data to identify chronic conditions without reference to patient problem lists may under-capture conditions among CHC patients in the United States.

Key words: multimorbidity, chronic disease ascertainment, concordance, community health centers (CHC), electronic health records (EHR)

BACKGROUND

Chronic medical conditions seriously impact patient quality of life, function, and mortality; their management generates substantial healthcare costs for patients, health systems, and society.[1–3] The United States Department of Health and Human Services (HHS) has identified multimorbidity (≥2 co-occurring chronic conditions) as a pressing, national public health issue. Increasingly, adults contend with multimorbidity, particularly as they age. In this context, the ability to effectively ascertain and monitor chronic disease levels in patient populations is critical for delivering comprehensive patient care, controlling costs, and conducting research to inform care improvement efforts. This need is especially pertinent for patients who receive care at community health centers (CHCs)—clinics who serve uninsured and underinsured patients regardless of their ability to pay—since many CHC patients are from communities that experience high chronic disease burden and have fewer resources to devote to protecting and enhancing their health.[4]

The HHS has developed a framework that identifies chronic conditions with the greatest cumulative impact on population health[5,6] and has disseminated algorithms through the Chronic Condition Data Warehouse (CCW) to identify patients with each condition using administrative claims.[7] These algorithms use International Classification of Disease, Clinical Modification diagnosis codes (9th or 10th editions, hereafter, ICD) to identify patient chronic condition status. The ICD coding system was adapted by the US National Center for Health Statistics for clinical billing purposes[8] and is now routinely used for tracking multimorbidity and mortality in health services research.[9] As organized in the CCW algorithms, ICD codes are aggregated into sets that indicate common condition categories (eg, diabetes, hypertension). CCW algorithms are a validated method to group diagnosis codes to measure the prevalence of common chronic conditions over time, outperforming self-report in some studies.[10–12]

CCW algorithms were developed with Medicare claims data and are endorsed for use with Medicaid claims as well,[13] but administrative data sets have long-known limitations. Claims data usually lag in availability, are limited to billable services,[14–18] can differ substantially from findings in a medical record, and omit uninsured patients, which is a substantial deficit for those studying medically underserved populations. CCW algorithms have additional limitations for populations with limited access to, or utilization of, healthcare services. Their logic includes reference periods (1–3 years, depending on condition) during which 2 claims must occur. This can impede the identification of pre-existing and historical chronic conditions which may be undesirable for some purposes.[11] For example, research and care improvement efforts have employed more flexible case definitions, such as longer periods of "look-back" to identify possible conditions among patients who may have more intermittent access to healthcare.[10–12] Further, claims data generally do not include historical conditions or conditions that are well-managed, limiting the ability to identify cohorts of interest.

Electronic health record (EHR) data provide another source of information regarding patient chronic condition status. CCW algorithms, though developed for claims data, have been previously applied as phenotyping methods to EHR data to ascertain chronic disease status.[19–23] EHR encounter diagnosis data are analogous to administrative claims data when they are billed to an insurer. However, EHR data contain additional discrete fields that can be relevant to identifying individuals with chronic conditions, such as anthropometric (eg, blood pressure) and laboratory measures, problem lists that summarize conditions relevant to a patient's overall health, and free-text clinician notes.[14,15] EHR data are more clinically comprehensive than claims, containing additional services that were not billed[24,25] and including uninsured patients who are absent from claims data. For many research purposes, EHR data are well-suited, more timely, and provide more complete medical information regarding chronic conditions than claims,[26] though EHR data quality for patient identification is an area of ongoing investigation.[27–29]

ICD codes are used in both EHR and claims data. ICD codes are used in claims data to justify a specific service provided. But these same codes may be found in multiple places in an EHR, including encounter diagnoses, problem lists, medical history, and free-text clinical notes. Encounter diagnoses are most analogous to the use of ICD codes in claims data because they are used to indicate relevant conditions at a particular visit. As with claims, misclassification may occur since codes may reflect speculative or confirmed diagnoses. This makes encounter diagnoses the most intuitive location to apply CCW algorithms for identifying patients with chronic disease. Encounter diagnoses may differ from claims records however, as they can include ICD codes not listed on claims because they would not affect reimbursement.[14,15] Workflow differences also create variability in completeness and accuracy of condition entry in encounter records at the condition, provider, clinic, and health system level.[30]

Problem list records are usually recorded in the EHR using diagnostic codes, for example, SNOMED-CT mapped to ICD codes,[31] but unlike encounter-specific diagnoses records they are intended as patient-level summaries of conditions requiring management and follow-up.[30] This structure is similar to the "top sheet" in paper charts and is considered best practice to use,[32,33] but there is variability in adoption at health system, clinic, and provider levels.[27,30] For example, some health systems encourage problem list use more than others, and providers who see patients for specialty, emergency, inpatient, or urgent care may be less able to or invested in maintaining the problem list than a patient's regular or primary care provider.[34] In a problem list, each condition a patient has should be recorded only once, with the date of onset and, when appropriate, the date of resolution. Compared to chart review and patient self-report, problem lists have sometimes been found to have redundant and incomplete information.[28,30] But among chronic conditions indicated by other discrete EHR fields (ie, lab results, prescriptions, and encounter diagnoses), problem lists have been found to be adequately complete and accurate,[35] especially for illnesses of higher severity.[36]

The purpose of this study is to examine EHR records of chronic conditions frequently assessed in geriatric studies of multimorbidity[37] and compare the concordance between encounter diagnoses and problem list sources. Greater understanding of the level of capture is important to population health researchers focused on multimorbidity burden and to clinical and EHR improvement efforts needing to identify patients based on condition status. As CCW algorithms become established as a standard approach to identifying patients with chronic conditions, we aim to contribute knowledge about how the CCW can be applied to EHR data, including a close examination of problem list records as an additional EHR location that may be relevant. This effort supports the HHS Strategic framework.[6]

SIGNIFICANCE

To the best of our knowledge, no prior studies applying CCW algorithms to EHR data have extended these algorithms to leverage data

available in patients' problem lists. Accurate determination of chronic disease using chart review is not feasible for many research contexts. Prior research has shown that combining diagnoses records from different EHR locations can be equivalent to the gold standard of chart review.[35] Combinations of discrete EHR fields with standard code systems (such as ICD) have been found in many studies to be the most useful inputs for algorithms to detect medical conditions, with other EHR elements offering lesser added benefit.[20,38–40] Efforts to assess chronic disease from EHR encounter diagnosis records and patient problem lists could benefit from understanding the level of concordance between these sources.

METHODS

Data

To compare ascertainment of chronic conditions from EHR encounter diagnosis and problem list records, we applied CCW algorithms (https://www2.ccwdata.org/web/guest/condition-categories),[5] to both types of structured fields in EHR data, focusing on 39 chronic conditions emphasized by HHS for their impact on multimorbidity.[37] Using ICD-9-CM and ICD-10-CM codes, we assessed chronic conditions in encounter diagnoses and problem list records, comparing rates of chronic condition ascertainment between the 2 sources.

We used EPIC and Intergy EHR data from the ADVANCE (Accelerating Data Value Across a National Community Health Center Network) clinical research network, which includes outpatient data from >1000 CHC clinics that has been organized into a standardized structure, or common data model.[41] From ADVANCE, we included 679 clinics who provided at least 50 office visits during at least 1 study year (2012–2019) and offered primary care services to at least 50 patients age ≥45 years. Most study-included clinics were Federally Qualified Health Centers (72%) or specialized in primary care (25%), while a minority specialized in public health, women's health, or other specialties but still provided primary care. All clinics were outpatient care facilities serving patients regardless of insurance or ability to pay. Following clinic selection, we identified patients age ≥45 years by the end of 2019 who had no record of death in their medical chart, and who had at least 2 in-person, outpatient ambulatory office visits at a study-included clinic between 2012 and 2019. Our final study sample included 1 180 290 older adult patients from 545 community health centers clinics in 22 states. The study protocol was approved by Oregon Health & Science University's Institutional Review Board (IRB ID# STUDY00020124).

Chronic diseases

We applied the CCW algorithms to patients' EHR encounter diagnoses records. Encounter diagnoses are structured analogously to administrative claims data, so we followed the CCW algorithms requirements for code count and look-back periods, treating visits akin to claims records. For example, CCW identification of asthma requires 2 outpatient claims on different dates within a twelve-month period that list 1 of several ICD-9 or ICD-10 codes for asthma. Once the required confirmation visit was identified, we ascribed the date of the first qualifying visit as the onset date for the condition. We assessed encounter data through 2019, to accommodate the longest, 3-year, look-back period. This allowed the algorithm to assess all conditions through 2016. We required at least 2 office visits for patient inclusion, omitting single-visit patients who were unable to have confirmation visits.

We then used CCW ICD code sets to identify chronic conditions in patients' problem lists. Providers typically record conditions on problem lists only when a diagnosis is confirmed and do not record distinct visits where a patient has a condition. For these reasons, we did not require multiple diagnosis records on the problem list or consider intervals of time between records to ascertain condition status. We took the earliest onset or recorded date for any record that indicated a condition and considered a single record to be sufficient evidence. Since problem list records can "resolve" if a patient no longer has a condition, we retained problem list records that remained unresolved by the end of 2016, to be comparable to our encounter ascertainment.

Demographic characteristics

We identified patient sex and race/ethnicity from the most recent EHR demographic information. Insurance type and household income were determined as of the end of 2016. We assigned patients' race/ethnicity as Hispanic when Hispanic was listed on their medical chart, or when their ethnicity was unknown and their preferred language was Spanish. Non-Hispanic patients were then categorized according to their recorded race, with 21 patients of "other" race reclassified as unknown. Patients missing sex were dropped from the study ($n = 75$) because the rare absence of this information seemed indicative of an incomplete or inaccurate patient record. We created a missing category for all other covariates.

Analyses

We report absolute standardized mean differences (ASMD) to compare demographics of patients who had chronic conditions found in encounter diagnosis records (hereafter referred to as "encounters") versus those having chronic conditions identified in the problem list, excluding patients with no evidence of chronic conditions by either source. ASMDs are measures of effect size, with differences in group means calculated in units of standard deviation, and they can be applied to overlapping sets of study participants,[42,43] as in this study where many patients had conditions identified in both sources. ASMDs can range from 0 to 1, with larger numbers indicating greater disagreement; values greater than 0.1 denote meaningful differences between groups.[43]

We measured encounter and problem list concordance for condition status among all study patients. Because we did not consider either of our sources to be a gold standard of ascertainment, we calculated percent total agreement (PTA), kappa, and prevalence-adjusted bias-adjusted kappa (PABAK) as measures of concordance. Percent total agreement conveys the overall strength of agreement, kappa compares that agreement to chance alone, and PABAK is commonly reported to address bias between raters or prevalence differences.[44] Relying on kappa to assess agreement between conditions with vastly different prevalence can be misleading but PABAK can mask these influences entirely. So we also report bias index (BI) and prevalence index (PI) separately to convey the magnitude of these influences on the agreement measures.[44–47] BI is the difference in rater discordance as a proportion of the total population, with higher values indicating a condition is more likely found in either the problem list or encounter source. PI is the difference in positive versus negative agreement as a proportion of the total population, with high values indicating rare conditions. Finally, we calculated the difference in onset dates between problem list and encounter records to assess onset similarity among conditions that were found in both sources.

We identified patients having all 39 conditions of interest, but only 58 patients had hepatitis-D or hepatitis-E, so we limited analyses of concordance to the remaining 37 conditions. We used SAS version 7.15 for analyses and R version 3.6.0 to generate figures (SAS Institute Inc., Cary, NC, USA, 2017, The R Foundation for Statistical Computing, 2019).

RESULTS

Patients

Out of 1 180 290 eligible CHC patients, 669 144 (56.7%) were female, 456 618 (38.7%) were non-Hispanic white and 383 627 (32.5%) were Hispanic. Most had Medicare/Medicaid as primary payor, 602 787 (51.1%), while 294 819 (25.0%) were uninsured. The largest group by age was 45–54 (494 447, 41.9%), with 268 756 (22.8%) being 65 or older.

There were 493 123 (41.8%) patients with no chronic conditions identified in either source. Patients without chronic conditions tended to be younger; 51.4% were 45–54 compared to 33.6% and 34.6% for the encounter and problem list groups, respectively. Patients without conditions were also more likely to be Hispanic, have private insurance or be uninsured, and have fewer office visits than the encounter or problem list groups (Supplementary material S1).

Patients with conditions

Study sample characteristics for 687 167 patients (58.2% of the population) with 1 or more chronic conditions ascertained by either EHR source are shown in Table 1, comparing patients whose chronic conditions were evidenced in encounters versus problem lists. Among patients with a chronic condition by either source, 589 761 (85.8%) patients had evidence in encounters and 651 521 (94.8%) had evidence in problem lists. Patient characteristics were similar for these highly overlapping groups. Most patients with conditions were aged 55–64 (37.7% and 37.6%, respectively for encounter and problem list groups) and female (57.1% for both groups). The problem list group had slightly more patients with only 1 chronic condition (19.6% vs 16.5%) and 3 or fewer annual office visits (43.3% vs 39.4%) than the encounter group. Household income and primary payor were similar for both groups. ASMDs were below or near 0.1 (chronic condition count just exceeded this threshold at 0.1129), indicating no substantial difference between patients who had evidence of chronic conditions in the 2 sources.

Conditions

We found 2 271 642 chronic conditions among 687 167 patients and most conditions were reflected in both data sources (1 349 824; 59.4%). Additional conditions were identified only in encounters (277 495; 12.2%) or problem lists (644 323; 28.4%). Chronic conditions varied in how often they could be ascertained for the same patient by both sources (Figure 1). The median amount of overlap was 50.4% of patients having evidence of the condition in both places (interquartile range 43.1–57.4%). Conditions with the highest concordance were diabetes, HIV/AIDS, and hypertension, with 84.4%, 78.1%, and 74.7% overlap, respectively. Sixteen conditions had <50% agreement, including all evaluated cancers (breast, colorectal, endometrial, lung, prostate, leukemia, and lymphoma), all evaluated substance use disorders (alcohol, drug, and opioid), asthma, osteoporosis, and most types of hepatitis. Overlap for mental health conditions ranged from 47.4% for anxiety to 59.8% for depression.

Overall, most chronic conditions were identified more often through problem lists than encounters, including atrial fibrillation, schizophrenia, hyperlipidemia, hepatitis, post-traumatic stress disorder (PTSD), depressive disorders, autism, chronic obstructive pulmonary disease (COPD), cancers, and substance use disorders (including alcohol, drug, and opioid). Fewer conditions (heart failure, depression, and Alzheimer's disease) had greater ascertainment using encounter records.

Measures of concordance between encounters and problem lists for each chronic condition among all patients are reported in Table 2. The PABAK estimated values indicated generally strong agreement between problem list and encounter condition identification in the EHR. The percent of total agreement (PTA) was more than 95% for all conditions except anxiety, depression, depressive disorders, hyperlipidemia, hypertension, and arthritis, and PTA was higher than 90% for all conditions except hyperlipidemia. However, PABAK and PTA values increase when condition prevalence is low. High PI values for many diseases indicate when rare disease frequency may be inflating measures of agreement.

Comparisons of the problem list and encounter onset dates for 1 349 824 conditions found in both sources are shown in Figure 2. Condition onset was often on the same day or within a month (54.2%), with 69.9% of dually identified conditions having onset dates within 6 months of one another, and 78.0% conditions having onset dates within 12 months.

DISCUSSION

This study compared patients' chronic disease as ascertained through clinical encounter diagnoses and problem lists in a large EHR database of CHCs across 21 states. We applied CCW algorithms for 37 conditions, including code frequency and look-back period requirements, to ascertain conditions from encounter records and then compared those results to patients' condition status from their problem lists. We found variation across conditions when comparing rates of overlap between these 2 methods of ascertainment, with higher concordance for diabetes, HIV, and hypertension and lower concordance for drug, alcohol, and opioid use disorders, endometrial, breast, and colorectal cancer, stroke, and osteoporosis.

Problem lists and encounters each have strengths and limitations for chronic condition ascertainment. Both sources can be incomplete: encounter diagnoses will not consistently reflect chronic conditions that are not treated at a given encounter, and problem lists can be unmaintained or underused. Data entry burden on providers can affect both sources. Without a gold standard, we cannot determine "correctness" in this study but given these shortcomings, the use of both sources may be warranted.

We observed higher rates of ascertainment in problem lists than encounters across nearly all conditions except heart failure, depression, Alzheimer's and related dementia, and chronic kidney disease. The problem list was particularly useful for identifying patients with endometrial, breast, and colorectal cancer, drug use disorders, stroke, and osteoporosis, more than doubling the number of patients who were identified with each of those conditions. These chronic diseases may not be treated as regularly in primary care settings, resulting in referrals to specialty care after an initial encounter. While specialist providers themselves may be less likely to utilize the problem list,[34] primary care providers may document them while managing follow-up care. This may be the only indication of some conditions in the EHR, making them difficult to detect with CCW algorithms in encounter records.

Table 1 Characteristics of patients having at least 1 chronic condition in EHR sources

Characteristics	Conditions by encounter diagnosis		Conditions by problem list		ASMD
	N = 589 761		N = 651 521		
	n	%	*n*	%	
Age					0.0256
45–54	198 199	33.6	225 312	34.6	
55–64	222 120	37.7	245 075	37.6	
65+	169 442	28.7	181 134	27.8	
Sex					0.0014
Female	336 438	57.1	372 134	57.1	
Race/ethnicity					<0.0001
Hispanic	176 803	30.0	196 182	30.1	
Non-Hispanic Asian	18 832	3.2	20 821	3.2	
Non-Hispanic Black	107 161	18.2	119 319	18.3	
Non-Hispanic American Indian/Alaska Native	2528	0.4	2711	0.4	
Non-Hispanic Native Hawaiian/Other Pacific Islander	2160	0.4	2333	0.4	
Non-Hispanic White	250 456	42.5	273 784	42.0	
Non-Hispanic multiple	2372	0.4	2599	0.4	
Unknown (race or ethnicity)	29 449	5.0	33 772	5.2	
Household income					<0.0001
<138% FPL	405 532	68.8	447 058	68.6	
≥138% FPL	91 101	15.5	100 280	15.4	
Missing	93 128	15.8	104 183	16.0	
Insurance payor					0.0421
Medicare/Medicaid	341 088	57.8	364 904	56.0	
Private	94 726	16.1	109 541	16.8	
Other	31 474	5.3	34 130	5.2	
Uninsured	122 297	20.7	142 728	21.9	
Missing	176	≤0.1	218	≤0.1	
Chronic condition count					0.1129
0	0	0.0	0	0.0	
1	97 218	16.5	127 588	19.6	
2–3	236 654	40.1	260 105	39.9	
4+	255 889	43.4	263 828	40.5	
Office visits per year					0.0815
≤3	232 380	39.4	282 221	43.3	
4–5	193 076	32.7	202 256	31.0	
>5	164 305	27.9	167 044	25.6	

ASMD: absolute standardized mean difference, (values greater than 0.1 denote meaningful differences between comparison groups); FPL: Household income was calculated as a percentage of the Federal Poverty Level.

Note: Clinics were present in 22 US states: Alaska, California, Florida, Georgia, Hawaii, Indiana, Kansas, Massachusetts, Maryland, Minnesota, Missouri, Montana, North Carolina, New Mexico, Nevada, Ohio, Oregon, Rhode Island, South Carolina, Texas, Washington, and Wisconsin.

Problem list data may also be helpful for identifying conditions that are asymptomatic, under control, in remission, or predate the first visit recorded in the EHR, as encounter diagnoses may not be available in those circumstances. And they may provide more complete information on younger patients, those with less regular access to care, or infrequent or intermittent healthcare utilization. Single-visit patients (omitted from our study) are a clear example of this; they cannot meet the CCW requirements for confirmation visits. Problem lists might be particularly advantageous for targeted patient care efforts (eg, registries or interventions), where timely identification could assist in better disease management.[30] Particularly in cases when relevant care may be referred out to specialists, problem lists may offer timely identification of patients in need of appropriate follow-up care.[28]

There are circumstances under which problem lists may be less relevant for identifying chronic disease. The utility of problem list data hinges on them being up to date, regularly used by providers, and adopted by healthcare systems who support their use. When problem lists are not maintained, they may suffer from incomplete or redundant information,[30,48] causing doubt in their accuracy. For direct comparisons to Medicare or Medicaid populations, or when studying populations with frequent healthcare utilization, researchers might prefer to apply CCW algorithm logic as closely as possible. When detailed accuracy on disease status is of high importance, researchers may prefer to triangulate encounter diagnoses with laboratory test results, medication orders, referrals, or other locations in the EHR,[19,20] or conduct chart review. However, when identifying all patients with any indication of a condition is the primary goal, the use of problem lists and/or modifying algorithm parameters might be a better approach.[11]

Future research is warranted to explore why heart failure, depression, chronic kidney disease, and Alzheimer's disease and related dementias were identified proportionately less often by problem lists than by ascertainment methods relying on encounter diagnoses. In the case of depression, this may be attributable to a mapping difference between clinically similar concepts, as providers
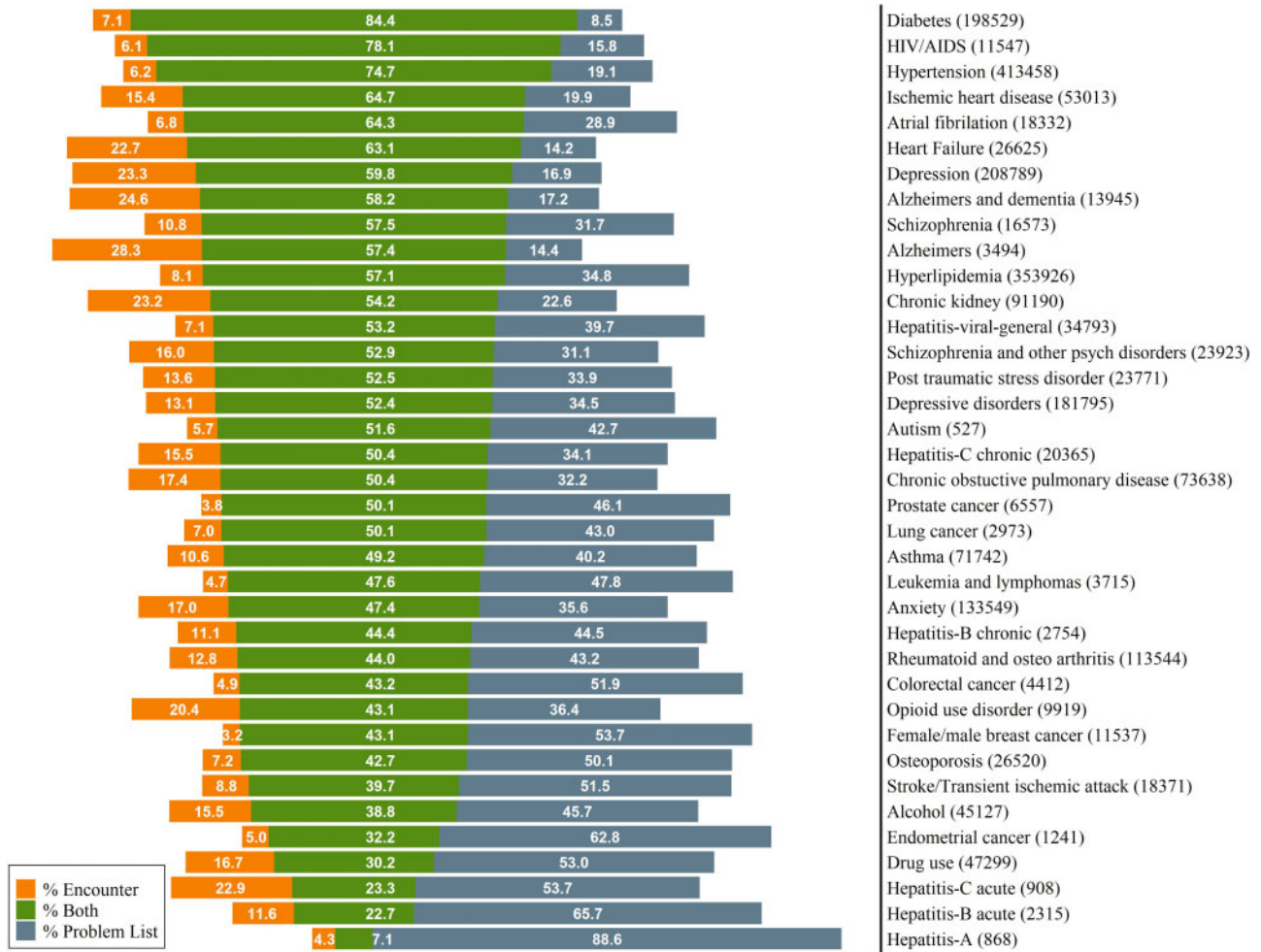
Figure 1. Overlap of chronic condition ascertainment from EHR problem list and encounter diagnosis records.

may be more inclined to list depressive disorders on a problem list but may only need to list (acute) depression on an encounter for billing purposes. For Alzheimer's, heart failure, or chronic kidney disease, there may be barriers to screening and diagnosis in the primary care setting which cause delays in problem list documentation.[49,50] However, we note that all cancers and arthritis were found to be more commonly captured by problem lists than encounter diagnoses, despite often requiring specialty equipment to diagnose. It is unclear, then, why these few conditions were less likely to be noted in patients' problem list records than in patterns of encounter diagnosis that were detected by CCW algorithms.

Our study has several strengths. First, our data come from community-based healthcare clinics in various US regions, containing a large clinical population of adults with a high chronic disease burden. Second, our study evaluates medical records from many uninsured patients, who are frequently missing from health services studies, especially claims-based research. When available, EHR records can fill in this reporting gap and aid in care improvement, clinical practice, and public health efforts for many under-resourced people. Third, this study contributes to the rapidly evolving literature on multimorbidity ascertainment. While there is currently no consensus on the best approach for defining or ascertaining multimorbidity, recent guidance highlights consideration of data availability and oper-

ationalization, given fit with the intended and stated purpose of study.[51] Expanding EHR applications of CCW algorithms to include data from patient problem list records is a strategy that aligns with these considerations, increasing capture of the chronic conditions most frequently used to operationalize multimorbidity.[6]

This study also has limitations. Both problem list and encounter diagnosis records can be incomplete or inaccurate in different ways, depending on the clinic, healthcare system, and provider factors.[27] We did not conduct chart review to validate conditions, potentially omitting information available from other EHR fields or narrative text. Additionally, we applied CCW algorithms to encounter diagnoses on patients younger than 65, and to EHR data sources, despite CCW algorithms being developed and validated for Medicare patient populations using administrative claims data. However, CMS now endorses applications of the CCW ascertainment methods to Medicaid populations,[13] and application to EHR data is increasingly common in other studies.[19–22] Lastly, our study used outpatient CHC data from 2 EHR systems (Epic and Intergy) that may not be generalizable to others.

Future work should emphasize validation of EHR encounter and problem list diagnoses to assess the accuracy and refine EHR ascertainment methods when using a gold standard is not practical.[52,53] This is especially important as EHR accuracy and com-

Table 2. Measures of concordance for chronic conditions as ascertained from EHR sources

Chronic condition	Encounter diagnosis	Problem list	Kappa	PTA	|PI|	|BI|	PABAK
Alcohol use disorder	24 501	38 145	0.55	0.98	0.95	0.01	0.95
Alzheimer's disease	2992	2506	0.73	>0.99	>0.99	<0.01	>0.99
Alzheimer's disease and related disorders or senile dementia	11 546	10 513	0.73	>0.99	0.98	<0.01	>0.99
Anxiety disorders	86 031	110 781	0.61	0.94	0.83	0.02	0.88
Arthritis—rheumatoid or osteoarthritis	64 454	99 047	0.58	0.95	0.86	0.03	0.89
Asthma	42 899	64 117	0.64	0.97	0.91	0.02	0.94
Atrial fibrillation	13 032	17 086	0.78	>0.99	0.97	<0.01	0.99
Autism spectrum disorders	302	497	0.68	>0.99	>0.99	<0.01	>0.99
Cancer—breast (female/male)	5343	11 170	0.60	>0.99	0.99	<0.01	0.99
Cancer—colorectal	2122	4196	0.60	>0.99	>0.99	<0.01	>0.99
cancer—endometrial	462	1179	0.49	>0.99	>0.99	<0.01	>0.99
cancer—leukemia and lymphoma	1940	3542	0.64	>0.99	>0.99	<0.01	>0.99
Cancer—lung	1696	2765	0.67	>0.99	>0.99	<0.01	>0.99
Cancer—prostate	3532	6310	0.66	>0.99	0.98	<0.01	0.99
Chronic kidney disease	70 553	70 072	0.68	0.96	0.88	<0.01	0.93
Chronic obstructive pulmonary disease and bronchiectasis	49 916	60 853	0.65	0.97	0.91	<0.01	0.94
Depression	173 539	160 115	0.71	0.93	0.72	0.01	0.86
Depressive disorders	119 035	158 011	0.65	0.93	0.77	0.03	0.85
Diabetes	181 634	184 457	0.90	0.97	0.69	<0.01	0.95
Drug use disorders	22 208	39 381	0.45	0.97	0.95	0.01	0.94
HIV/AIDS	9725	10 839	0.88	>0.99	0.98	<0.01	>0.99
Heart failure	22 853	20 578	0.77	>0.99	0.96	<0.01	0.98
Hepatitis—general viral	20 981	32 322	0.69	0.99	0.95	<0.01	0.97
Hepatitis A	99	831	0.13	>0.99	>0.99	<0.01	>0.99
Hepatitis B—acute or unspecified	793	2047	0.37	>0.99	>0.99	<0.01	>0.99
Hepatitis B—chronic	1528	2449	0.61	>0.99	>0.99	<0.01	>0.99
Hepatitis C—acute	420	700	0.38	>0.99	>0.99	<0.01	>0.99
Hepatitis C—chronic	13 422	17 215	0.67	>0.99	0.97	<0.01	0.98
Hyperlipidemia	230 816	325 144	0.65	0.87	0.53	0.08	0.74
Hypertension	334 506	387 671	0.79	0.91	0.39	0.05	0.82
Ischemic heart disease	42 461	44 863	0.78	0.98	0.93	<0.01	0.97
Opioid use disorder	6307	7891	0.60	>0.99	0.99	<0.01	>0.99
Osteoporosis	13 235	24 619	0.59	0.99	0.97	<0.01	0.97
Post-traumatic stress disorder	15 711	20 546	0.68	>0.99	0.97	<0.01	0.98
Schizophrenia	11 320	14 786	0.73	>0.99	0.98	<0.01	0.99
Schizophrenia and other psychiatric disorders	16 472	20 103	0.69	>0.99	0.97	<0.01	0.98
Stroke/transient ischemic attack	8910	16 753	0.56	>0.99	0.98	<0.01	0.98

PTA: percentage of total agreement; PI: prevalence index; BI: bias index; PABAK: prevalence-adjusted bias-adjusted kappa; HIV/AIDS: human immunodeficiency virus/acquired immunodeficiency syndrome.

Note: Prevalence measures were calculated using entire patient population as denominator, except endometrial and prostate cancers used only female and male patients, respectively. Bias Index is the difference in rater discordance, as a proportion of the total population. High values indicate 1 rater is more likely to be positive than the other (condition is more likely to be found in either problem list or encounter source). Prevalence Index is the difference in positive versus negative agreement, as proportion of total population. High values indicate either a very rare or very common condition. Disease prevalence greatly affects these concordance measures. For example, since acute hepatitis-B is infrequently found in our data, most patients have no indications in both problem list and encounter sources, resulting in very high PI, PTA, and PABAK scores, even though the kappa is low. Conversely, since hypertension is very common in our data, a greater proportion of the study sample could be misclassified and PTA is lower, the PI is small, and the PABAK is lower. There is much less "prevalence correction" from kappa to PABAK in conditions with higher prevalence.

pleteness are a rapidly evolving landscape, critical to the provision of timely patient care, control of healthcare costs, and maximization of provider effectiveness. The clinical process of adding conditions to encounters or problem lists is often overly tedious, which can discourage and disincentivize providers, leading to the inaccurate and incomplete recording of conditions and undermining the EHR's full potential. Healthcare systems and EHR vendors should reduce this burden to profoundly improve patient care and provider satisfaction.[48] A patient and provider-centered approach toward reducing data entry burden will only improve accuracy and completeness of EHR data, increasing the utility of EHR data for patient care and downstream data uses.

CONCLUSION

This study highlights important considerations for the ascertainment of chronic disease in EHR data. This is critical for the HHS strategy to measure and monitor chronic disease and multimorbidity, particularly among low-income, uninsured, or underinsured patients. Comparing encounter diagnosis and problem list records, we found varying levels of concordance in chronic condition status, but problem list data routinely provided higher levels of capture than encounter diagnoses alone. Our findings suggest that applications of CCW algorithms to EHR data, depending on study aims, may benefit from incorporating problem list data to identify more patients with chronic disease. Greater investment in the usability and mainte-
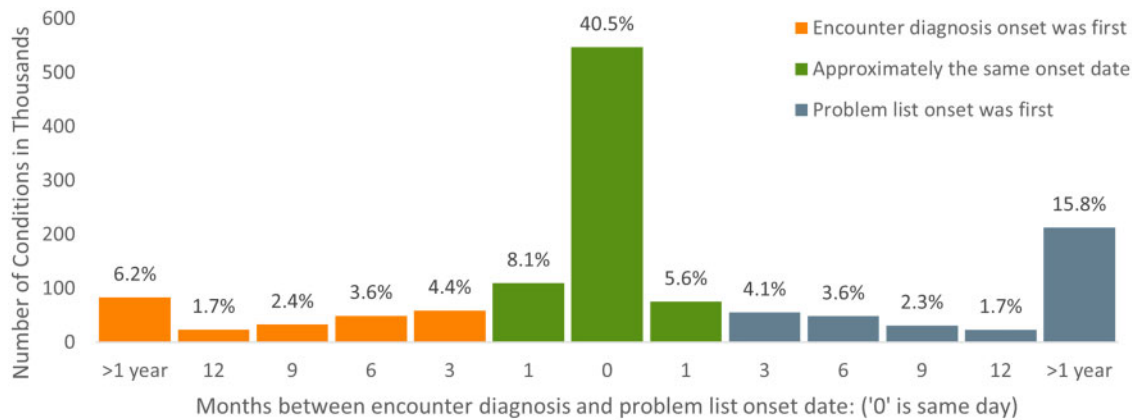
Figure 2. Distribution of months between problem list and encounter diagnosis onset dates among conditions ascertained by both EHR sources.

nance of high-quality problem lists may further improve this valuable tool for clinical management and reporting of chronic disease throughout the United States.

AUTHOR CONTRIBUTIONS

RV, TS, and AQ contributed to the literature search, study design, data interpretation, drafting and editing the manuscript. RV and TS conducted the data management and analysis. NH, NW (Weiskopf), and MM contributed to the data interpretation, writing, and editing of the manuscript. DD, JO, SV, and NW (Warren) contributed to data interpretation, and editing of the manuscript. All authors approved the final version of the manuscript. RV had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

SUPPLEMENTARY MATERIAL

Supplementary material is available at *Journal of the American Medical Informatics Association Journal* online.

CONFLICT OF INTEREST STATEMENT

Dr Weiskopf reports personal fees from Merck, outside the submitted work. Remaining authors have no conflicts to declare.

DATA AVAILABILITY

Raw data underlying this article were generated from multiple health systems across institutions in the ADVANCE Network; restrictions apply to the availability and re-release of data under organizational agreements.

Prior Presentations. A preliminary version of these analyses was presented at the American Medical Informatics Virtual Informatics Summit, March 2021.

REFERENCES

1. Lochner KA, Goodman RA, Posner S, *et al.* Multiple chronic conditions among Medicare beneficiaries: state-level variations in prevalence, utilization, and cost, 2011. *Medicare Medicaid Res Rev* 2013; 3: 3.
2. Vogeli C, Shields AE, Lee TA, *et al.* Multiple chronic conditions: prevalence, health consequences, and implications for quality, care management, and costs. *J Gen Intern Med* 2007; 22 (S3): 391–5.
3. Goodman RA, Ling SM, Briss PA, *et al. Multimorbidity Patterns in the United States: Implications for Research and Clinical Practice.* J Gerontol: Series A 2016; 71 (2): 215–20.
4. Shin P, Alvarez C, Sharac J, *et al.* A Profile of Community Health Center Patients: Implications for Policy. KFF. 2013. www.kff.org/medicaid/issue-brief/a-profile-of-community-health-center-patients-implications-for-policy. Accessed October 10, 2021.
5. CMS Chronic Conditions Data Warehouse. CCW Chronic Condition Algorithms, 2019. www2.ccwdata.org/documents/10280/19139421/ccw-chronic-condition-algorithms.pdf. Accessed August 18, 2020.
6. *Multiple Chronic Conditions - A Strategic Framework: Optimum Health and Quality of Life for Individuals with Multiple Chronic Conditions.* Washington, DC: US Department of Health and Human Services; 2010. www.hhs.gov/sites/default/files/ash/initiatives/mcc/mcc_framework.pdf. Accessed October 4, 2021.
7. CMS Chronic Conditions Data Warehouse. Technical Guidance Documentation. 2019. www2.ccwdata.org/web/guest/technical-guidance-documentation. Accessed October 4, 2021.
8. Wu P, Gifford A, Meng X, *et al.* Mapping ICD-10 and ICD-10-CM codes to phecodes: workflow development and initial evaluation. *JMIR Med Inform* 2019; 7 (4): e14325.
9. Kirby JC, Speltz P, Rasmussen L, *et al.* PheKB: a catalog and workflow for creating electronic phenotype algorithms for transportability. *J Am Med Inform Assoc* 2016; 23 (6): 1046–52.
10. Rector TS, Wickstrom SL, Shah M, *et al.* Specificity and sensitivity of claims-based algorithms for identifying members of Medicare+ Choice health plans that have chronic medical conditions. *Health Serv Res* 2004; 39 (6 Pt 1): 1839–58.

11. Gorina Y, Kramarow EA. Identifying chronic conditions in Medicare claims data: evaluating the Chronic Condition Data Warehouse algorithm. *Health Serv Res* 2011; 46 (5): 1610–27.

12. Clair PS, Gaudette E, Zhao H, *et al.* Using self-reports or claims to assess disease prevalence: it's complicated. *Med Care* 2017; 55 (8): 782–8.

13. CMS Chronic Condition Data Warehouse. CCDW: condition categories. 2016. www.ccwdata.org/web/guest/condition-categories. Accessed October 4, 2021.

14. Kharrazi H, Chi W, Chang HY, *et al.* Comparing population-based risk-stratification model performance using demographic, diagnosis and medication data extracted from outpatient electronic health records versus administrative claims. *Med Care* 2017; 55 (8): 789–96.

15. Wilson J, Bock A. The benefit of using both claims data and electronic medical record data in health care analysis. *Optum Insight.* 2012. https://www.optum.com/content/dam/optum/resources/whitePapers/Benefits-of-using-both-claims-and-EMR-data-in-HC-analysis-WhitePaper-ACS.pdf (Accessed 7 February 2022).

16. Klabunde CN, Warren JL, Legler JM. Assessing comorbidity using claims data: an overview. *Med Care* 2002; 40 (Supplement): IV–IV35.

17. Fisher ES, Whaley FS, Krushat WM, *et al.* The accuracy of Medicare's hospital claims data: progress has been made, but problems remain. *Am J Public Health* 1992; 82 (2): 243–8.

18. Kieszak SM, Flanders WD, Kosinski AS, *et al.* A comparison of the Charlson comorbidity index derived from medical record data and administrative billing data. *J Clin Epidemiol* 1999; 52 (2): 137–42.

19. Richesson RL, Rusincovitch SA, Wixted D, *et al.* A comparison of phenotype definitions for diabetes mellitus. *J Am Med Inform Assoc* 2013; 20 (e2): e319– 326.

20. Upadhyaya SG, Murphree DH, Ngufor CG, *et al.* Automated diabetes case identification using electronic health record data at a tertiary care facility. *Mayo Clin Proc* 2017; 1 (1): 100–10.

21. Xu J, Wang F, Xu Z, *et al.* Data-driven discovery of probable Alzheimer's disease and related dementia subphenotypes using electronic health records. *Learn Health Syst* 2020; 4 (4): 10246.

22. Xu Z, Wang F, Adekkanattu P, *et al.* Subphenotyping depression using machine learning and electronic health records. *Learn Health Sys* 2020; 4 (4): 10241.

23. He Z, Bian J, Carretta HJ, *et al.* Prevalence of multiple chronic conditions among older adults in Florida and the United States: comparative analysis of the OneFlorida data trust and national inpatient sample. *J Med Internet Res* 2018; 20 (4): e137.

24. Schulz WL, Young PH, Coppi A, *et al.* Temporal relationship of computed and structured diagnoses in electronic health record data. *BMC Medical Inform Decis Mak* 2021; 21: 1–9.

25. Bailey SR, Heintzman JD, Marino M, *et al.* Measuring preventive care delivery: comparing rates across three data sources. *Am J Prev Med* 2016; 51 (5): 752–61.

26. Shephard E, Stapley S, Hamilton W. The use of electronic databases in primary care research. *Fam Pract* 2011; 28 (4): 352–4.

27. Singer A, Yakubovich S, Kroeker AL, *et al.* Data quality of electronic medical records in Manitoba: do problem lists accurately reflect chronic disease billing diagnoses? *J Am Med Inform Assoc* 2016; 23 (6): 1107–12.

28. Weiskopf NG, Cohen AM, Hannan J, *et al.* Towards augmenting structured EHR data: a comparison of manual chart review and patient self-report. *AMIA Annu Symp Proc* 2019; 2019: 903–12.

29. Heintzman J, Bailey SR, Hoopes MJ, *et al.* Agreement of Medicaid claims and electronic health records for assessing preventive care quality among adults. *J Am Med Inform Assoc* 2014; 21 (4): 720–4.

30. Wright A, McCoy AB, Hickman TT, *et al.* Problem list completeness in electronic health records: a multi-site study and assessment of success factors. *Int J Med Inform* 2015; 84 (10): 784–90.

31. Duarte J, Castro S, Santos M, *et al.* Improving quality of electronic health records with SNOMED. *Proc Technol* 2014; 16: 1342–50.

32. Hartung DM, Hunt J, Siemienczuk J, *et al.* Clinical implications of an accurate problem list on heart failure treatment. *J Gen Intern Med* 2005; 20 (2): 143–7.

33. Klappe ES, de Keizer NF, Cornet R. Factors influencing problem list use in electronic health records—application of the unified theory of acceptance and use of technology. *Appl Clin Inform* 2020; 11 (3): 415–26.

34. Wright A, Feblowitz MS, Maloney FL, *et al.* Use of an electronic problem list by primary care providers and specialists. *J Gen Intern Med* 2012; 27 (8): 968–73.

35. Martin S, Wagner J, Lupulescu-Mann N, *et al.* Comparison of EHR-based diagnosis documentation locations to a gold standard for risk stratification in patients with multiple chronic conditions. *Appl Clin Inform* 2017; 8 (3): 794–809.

36. Wang EC, Wright A. Characterizing outpatient problem list completeness and duplications in the electronic health record. *J Am Med Inform Assoc* 2020; 27 (8): 1190–7.

37. Goodman R, Posner SF, Huang ES, *et al.* Defining and measuring chronic conditions: imperatives for research, policy, program, and practice. *Prev Chronic Dis* 2013; 10 (4): 1–16.

38. McBrien KA, Souri S, Symonds NE, *et al.* Identification of validated case definitions for medical conditions used in primary care electronic medical record databases: a systematic review. *J Am Med Inform Assoc* 2018; 25 (11): 1567–78.

39. Wei MY, Luster JE, Chan CL, *et al.* Comprehensive review of ICD-9 code accuracies to measure multimorbidity in administrative data. *BMC Health Serv Res* 2020; 20 (1): 1–11.

40. Hansen ML, Gunn PW, Kaelber DC. Underdiagnosis of hypertension in children and adolescents. *JAMA* 2007; 298 (8): 874–9.

41. Data. PCORnet. 2017. https://pcornet.org/data. Accessed January 7, 2022.

42. Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Stat Med* 2009; 28 (25): 3083–107.

43. Austin PC. Using the standardized difference to compare the prevalence of a binary variable between two groups in observational research. *Commun Stat* 2009; 38 (6): 1228–34.

44. Wolinsky FD, Jones MP, Ullrich F, *et al.* The concordance of survey reports and Medicare claims in a nationally representative longitudinal cohort of older adults. *Med Care* 2014; 52 (5): 462–8.

45. Byrt T, Bishop J, Carlin JB. Bias, prevalence and kappa. *J Clin Epidemiol* 1993; 46 (5): 423–9.

46. Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Phys Ther* 2005; 85 (3): 257–68.

47. Chen G, Faris P, Hemmelgarn B, *et al.* Measuring agreement of administrative data with chart data using prevalence unadjusted and adjusted kappa. *BMC Med Res Methodol* 2009; 9 (1): 1–8.

48. Hodge CM, Narus SP. Electronic problem lists: a thematic analysis of a systematic literature review to identify aspects critical to success. *J Am Med Inform Assoc* 2018; 25 (5): 603–13.

49. Aranda MP, Kremer IN, Hinton L, *et al.* Impact of dementia: Health disparities, population trends, care interventions, and economic costs. *J Am Geriatr Soc* 2021; 69 (7): 1774–83.

50. Babulal GM, Quiroz YT, Albensi BC, *et al.* International Society to Advance Alzheimer's Research and Treatment, Alzheimer's Association. Perspectives on ethnic and racial disparities in Alzheimer's disease and related dementias: update and areas of immediate need. *Alzheimers Dement* 2019; 15 (2): 292–312.

51. Suls J, Bayliss EA, Berry J, *et al.* Measuring multimorbidity: selecting the right instrument for the purpose and the data source. *Med Care* 2021; 59 (8): 743–56.

52. Nissen F, Quint JK, Morales DR, *et al.* How to validate a diagnosis recorded in electronic health records. *Breathe (Sheff)* 2019; 15 (1): 64–8.

53. Reitsma JB, Rutjes AW, Khan KS, *et al.* A review of solutions for diagnostic accuracy studies with an imperfect or missing reference standard. *J Clin Epidemiol* 2009; 62 (8): 797–806.