

BMJ Open Implementing the US Department of Health and Human Services definition of multimorbidity: a comparison between billing codes and medical record review in a population-based sample of persons 40–84 years old

Jennifer L St Sauver ^{1,2}, Alanna M Chamberlain,¹ William V Bobo,³ Cynthia M Boyd,⁴ Lila J Finney Rutten,^{1,2} Debra J Jacobson,⁵ Michaela E McGree,⁵ Brandon R Grossardt,⁵ Walter A Rocca ^{1,6,7}

To cite: St Sauver JL, Chamberlain AM, Bobo WV, *et al.* Implementing the US Department of Health and Human Services definition of multimorbidity: a comparison between billing codes and medical record review in a population-based sample of persons 40–84 years old. *BMJ Open* 2021;**11**:e042870. doi:10.1136/bmjopen-2020-042870

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online ().

Received 16 July 2020
Revised 05 January 2021
Accepted 06 April 2021



© Author(s) (or their employer(s)) 2021. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to

Dr Jennifer L St Sauver;
stsauver.jennifer@mayo.edu

ABSTRACT

Objective To assess the validity of the US Department of Health and Human Services (DHHS) definition of multimorbidity using International Classification of Diseases, ninth edition (ICD-9) codes from administrative data.

Design Cross-sectional comparison of two ICD-9 billing code algorithms to data abstracted from medical records.

Setting Olmsted County, Minnesota, USA.

Participants An age-stratified and sex-stratified random sample of 1509 persons ages 40–84 years old residing in Olmsted County on 31 December 2010.

Study measures Seventeen chronic conditions identified by the US DHHS as important in studies of multimorbidity were identified through medical record review of each participant between 2006 and 2010. ICD-9 administrative billing codes corresponding to the 17 conditions were extracted using the Rochester Epidemiology Project records-linkage system. Persons were classified as having each condition using two algorithms: at least one code or at least two codes separated by more than 30 days. We compared the ICD-9 code algorithms with the diagnoses obtained through medical record review to identify persons with multimorbidity (defined as ≥ 2 , ≥ 3 or ≥ 4 chronic conditions).

Results Use of a single code to define each of the 17 chronic conditions resulted in sensitivity and positive predictive values (PPV) $\geq 70\%$, and in specificity and negative predictive values (NPV) $\geq 70\%$ for identifying multimorbidity in the overall study population. PPV and sensitivity were highest in persons 65–84 years of age, whereas NPV and specificity were highest in persons 40–64 years. The results varied by condition, and by age and sex. The use of at least two codes reduced sensitivity, but increased specificity.

Conclusions The use of a single code to identify each of the 17 chronic conditions may be a simple and valid method to identify persons who meet the DHHS definition of multimorbidity in populations with similar demographic, socioeconomic, and health care characteristics.

Strengths and limitations of this study

- The two diagnostic algorithms were validated against manual medical record review of a representative sample of the general population.
- The records-linkage system includes virtually complete medical record information for code validation, making it unlikely that chronic conditions were missed.
- The study population is from an upper midwest region in the USA, and results may differ in populations with different demographic or socioeconomic characteristics.
- The study population includes data from practices in the upper midwest region of the USA, and results may differ in practices with different approaches to care and to assignment of administrative codes.
- Only International Classification of Diseases, ninth edition (ICD-9) coding algorithms were validated, and further research is needed to specifically validate ICD-10 coding algorithms for multimorbidity.

BACKGROUND

Ageing populations, together with improvements in management of chronic conditions, have resulted in increasing numbers of persons living with multiple chronic conditions (multimorbidity; defined as the presence of two or more chronic conditions).^{1 2} In the USA, 62% of Medicare recipients ages 65–74 years live with multimorbidity, and this frequency increases to 82% in persons older than 85 years.³ Multimorbidity is strongly associated with high healthcare costs and adverse health outcomes, including poor quality of life.^{4–6} Therefore, studying, preventing and

treating multimorbidity have been declared global health priorities.²

Administrative databases are important resources for studies of multimorbidity because they may contain extensive details on many different chronic conditions in large populations. Such databases have been used extensively in a wide range of studies on multimorbidity throughout the world.^{3 7–10} However, there is currently a lack of consensus worldwide regarding how best to define multimorbidity in research studies (number and type of conditions).^{11–16} The prevalence of multimorbidity and the outcomes associated with multimorbidity will vary depending on the definitions that are chosen. In addition, several studies have shown that the International Classification of Diseases (ICD) codes are highly variable in their ability to correctly capture the presence of individual conditions in persons that were assigned these codes.^{17–19} The definition of multimorbidity is dependent on the correct identification of a list of individual chronic conditions. Therefore, if ICD codes do not correctly identify individual conditions, use of ICD codes to identify multimorbidity may also be problematic.

The US Department of Health and Human Services (DHHS) has recommended a list of 20 conditions to be used in studies of multimorbidity because they are ‘chronic, prevalent and are potentially amenable to intervention’.¹ Additionally, Goodman and colleagues have identified the ICD, ninth edition (ICD-9) codes that map to the 20 conditions for use in research studies.¹ We have previously used two codes separated by more than 30 days to identify persons with prevalent chronic conditions to reduce the risk of including false positive diagnoses in our multimorbidity studies.^{20 21} By contrast, other investigators have used a single code to define the conditions included in multimorbidity.²² It is not currently known whether the use of a single code or of multiple codes is best for studies that use the DHHS definition of multimorbidity.^{7 23}

To address this gap, we estimated the positive predictive value (PPV), negative predictive value (NPV), sensitivity and specificity of two algorithms for detecting the conditions included in the DHHS definition of multimorbidity (at least one code or at least two codes) using medical record reviews as the standard for comparison in a sample of persons residing in Olmsted County, Minnesota, USA.

METHODS

Data source

We used the resources of the Rochester Epidemiology Project (REP) medical records-linkage system for this study.²⁴ Details regarding the REP have been described previously.^{24 25} Briefly, the REP has captured virtually complete medical record information for the population residing in Olmsted County since 1966.²⁶ The REP includes all billing codes assigned to a healthcare visit and sent to payers, so the available data are similar to the types of administrative data available in insurance claims

databases. In addition, the full text of the medical records from which these codes are derived is also available, and these records may be reviewed for further details related to chronic diseases.²⁷

Study population

We identified all persons residing in Olmsted County on 31 December 2010 using the REP census resources.²⁴ We then selected an age-stratified and sex-stratified random sample of 1600 persons ages 40–84 years for this validation study. We did not study younger persons because previous work indicates that multimorbidity is most common in persons 40 years of age and older,²⁰ and the sample size needed to validate codes in younger persons was beyond the scope of this study. Therefore, persons were equally distributed across four age and sex strata: women 40–64 years, men 40–64 years, women 65–84 years and men 65–84 years. Ninety-one persons were excluded from the sample because their medical record review indicated that they were not residents of Olmsted County on 31 December 2010, resulting in a final sample size of 1509 persons.

Definition of multimorbidity

As previously described, DHHS has recommended the use of 20 conditions for studies of multimorbidity, and has identified the ICD-9 codes that map to the 20 conditions for use in research studies.¹ For this study, we excluded autism, hepatitis and HIV infection because these conditions are rare in our population, and our sample size was insufficient. The list of the 17 conditions, their acronym or abbreviation, and the corresponding ICD-9 codes used in this study are reported in online supplemental table 1. Persons with two or more of these 17 conditions were classified as having multimorbidity. We also defined multimorbidity as at least three and at least four chronic conditions.

Defining multimorbidity using ICD-9 codes

Using the code sets specified by DHHS, we used the electronic indexes of the REP to extract all ICD-9 codes for the 17 chronic conditions of interest between 1 January 2006 and 31 December 2010 for our study sample (online supplemental table 1). We then created two ICD-9 code-based algorithms to detect each of the 17 conditions within the 5-year time frame: (1) at least one code for a given diagnosis; or (2) two codes for the same diagnosis separated by more than 30 days. Multimorbidity was then defined as the presence of ≥ 2 , ≥ 3 and ≥ 4 of the 17 chronic conditions using each of the two algorithms.

Medical record abstraction

A trained nurse abstractor with extensive medical record review experience in identifying chronic diseases reviewed the medical records of the study sample. She was kept unaware of the definition of multimorbidity based on electronic data extraction. All available records and corresponding medical visit

information for our study sample from 1 January 2006 through 1 December 2010 were reviewed to identify the presence of the 17 chronic conditions. The nurse abstractor recorded the first date on which the person had a ‘definite’ diagnosis, a ‘probable diagnosis plus treatment’ or a ‘history of’ the diagnosis. A ‘history of’ diagnosis was used for a chronic condition that first occurred prior to 1 January 2005. A diagnosis was considered ‘definite’ if a healthcare provider specifically noted that the person had the condition in the medical record. If a patient did not have a diagnosis specifically noted by the provider, but he/she had symptoms or laboratory values described and he/she was treated for the condition, we considered him/her as ‘probable diagnosis plus treatment’. For example, a person with elevated blood pressure readings plus a prescription for an antihypertensive medication was included as ‘probable diagnosis plus treatment’, even if the healthcare provider did not specifically note that the person had ‘hypertension’ in the medical record.

Finally, any diagnoses that were noted in the medical record to be present prior to 1 January 2006 were assigned

a study date of 31 December 2005 and were considered ‘history of’ diagnoses. For example, a person with a medical record note in 2006 that stated ‘patient had a stroke 2 years ago’ was included as ‘history of’ stroke. Our final definition of whether a person had the condition of interest included persons who had a definite, probable plus treatment or history of diagnosis in the 5 years prior to 31 December 2010, as ascertained through medical record review. Persons with ≥ 2 , ≥ 3 and ≥ 4 chronic conditions identified through medical record review were classified as having multimorbidity.

Analysis

We considered the medical record review definition as the reference (gold) standard. We then calculated the PPV, NPV, sensitivity and specificity for each of the 17 conditions and for multimorbidity (≥ 2 , ≥ 3 and ≥ 4 chronic conditions). We conducted the analyses overall and within each age and sex stratum. Overall estimates for the total population ages 40–84 years were weighted using inverse sampling fractions. We measured validity using both sensitivity and specificity, and PPV and NPV. We note

Table 1 Number and per cent of persons identified with definite, probable plus treatment or history of chronic conditions, and of persons with multimorbidity from medical record review (reference or gold standard)

Condition (acronym or abbreviation)	40–64 years		65–84 years		Overall * Both sexes n (%)
	Women n (%)	Men n (%)	Women n (%)	Men n (%)	
Sample size	375 (100)	371 (100)	383 (100)	380 (100)	1509 (100)
Depression (DEP)	154 (41.1)	84 (22.6)	129 (33.7)	65 (17.1)	432 (49.1)
Diabetes (DIA)	93 (24.8)	139 (37.5)	200 (52.2)	231 (60.8)	663 (59.2)
Hyperlipidaemia (LIP)	167 (44.5)	209 (56.3)	280 (73.1)	307 (80.8)	963 (90.5)
Arthritis (ART)	146 (38.9)	130 (35)	273 (71.3)	253 (66.6)	802 (71.8)
Cancer (CAN)	47 (12.5)	30 (8.1)	118 (30.8)	165 (43.4)	360 (26.9)
Asthma (AST)	64 (17.1)	30 (8.1)	56 (14.6)	27 (7.1)	177 (19.8)
Hypertension (HTN)	89 (23.7)	125 (33.7)	247 (64.5)	237 (62.4)	698 (59.3)
Substance abuse (SUB)	22 (5.9)	52 (14)	15 (3.9)	36 (9.5)	125 (14.3)
Cardiac arrhythmias (ARR)	76 (20.3)	50 (13.5)	145 (37.9)	174 (45.8)	445 (36.8)
Coronary artery disease (CAD)	7 (1.9)	29 (7.8)	62 (16.2)	126 (33.2)	224 (15.1)
Congestive heart failure (CHF)	0 (0)	8 (2.2)	21 (5.5)	29 (7.6)	58 (3.8)
Chronic obstructive pulmonary disease (COPD)	43 (11.5)	30 (8.1)	60 (15.7)	64 (16.8)	197 (18.2)
Chronic kidney disease (CKD)	10 (2.7)	18 (4.9)	52 (13.6)	65 (17.1)	145 (10.4)
Osteoporosis (OST)	12 (3.2)	3 (0.8)	101 (26.4)	13 (3.4)	129 (8.7)
Stroke or transient ischaemic attack (STR)	6 (1.6)	9 (2.4)	28 (7.3)	45 (11.8)	88 (6.1)
Schizophrenia and other psychoses (SZO)	3 (0.8)	6 (1.6)	4 (1)	4 (1.1)	17 (1.8)
Dementia and delirium (DEM)	0 (0)	1 (0.3)	12 (3.1)	14 (3.7)	27 (1.5)
Multimorbidity					
≥ 2 conditions	249 (66.4)	237 (63.9)	354 (92.4)	360 (94.7)	1200 (100)
≥ 3 conditions	173 (46.1)	160 (43.1)	315 (82.2)	323 (85)	971 (86.6)
≥ 4 conditions	102 (27.2)	105 (28.3)	266 (69.5)	278 (73.2)	751 (61.3)

*The overall percentages in the total Olmsted County population age 40–84 years reflects the sampling weights.

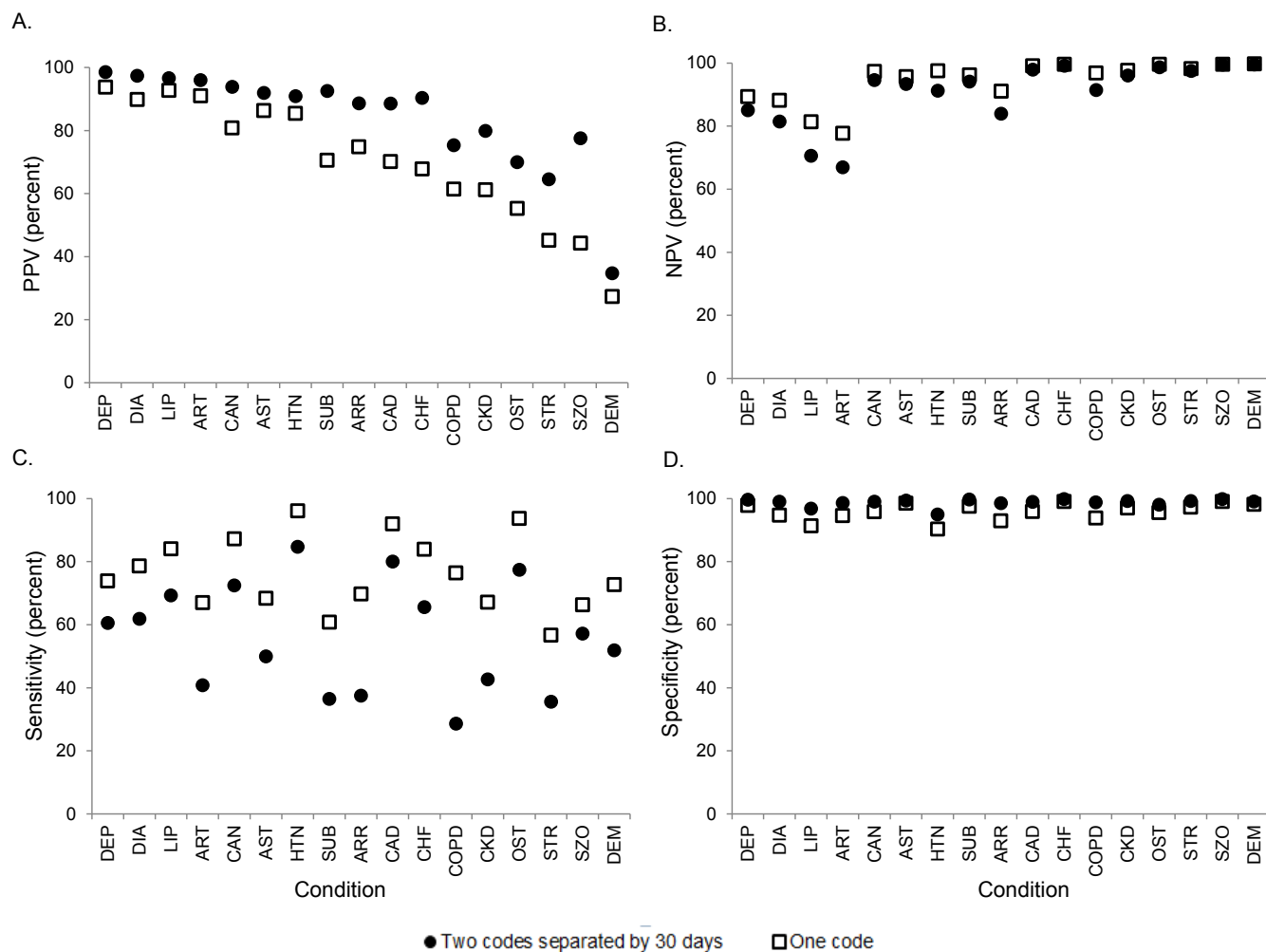


Figure 1 Positive predictive value (PPV; A), negative predictive value (NPV; B), sensitivity (C) and specificity (D) of two International Classification of Diseases, ninth edition (ICD-9) codes separated by >30 days (closed circles) and a single ICD-9 code (open squares) for identifying each of the 17 chronic conditions. ARR, cardiac arrhythmias; AST, asthma; ART, arthritis; CAD, coronary artery disease; CAN, cancer; CHF, congestive heart failure; CKD, chronic kidney disease; COPD, chronic obstructive pulmonary disease; DEM, dementia and delirium; DEP, depression; DIA, diabetes; HTN, hypertension; LIP, hyperlipidaemia; OST, osteoporosis; STR, stroke or transient ischaemic attack; SUB, substance abuse; SZO, schizophrenia and other psychoses.

that PPV and NPV are determined both by sensitivity and specificity of the algorithm and by the frequency of the disease (or of multimorbidity) for which the algorithm is used.²⁸ Therefore, we also explored the performance of the algorithms to detect multimorbidity across age and sex groups with different prevalences of multimorbidity by plotting sensitivity versus PPV and specificity versus NPV. We used the criteria defined by Tonelli and colleagues to determine whether the code-based algorithms had high validity (PPV and sensitivity $\geq 70\%$) or moderate validity (PPV $\geq 70\%$ but sensitivity $< 70\%$).²³

RESULTS

The proportion of persons with each of the 17 chronic conditions of interest based on medical record review is shown in table 1. As expected, the proportion of persons with each of the conditions varied by age and sex. For

example, hypertension was most common in women 65–84 years (65%) and least common in women 40–64 years (24%; table 1). Multimorbidity (≥ 2 conditions) was most common in men 65–84 years of age (95%) and least common in men 40–64 years of age (64%; table 1).

Single conditions

The algorithm requiring at least two codes separated by more than 30 days had consistently higher PPVs than the one code algorithm for identifying individual conditions (figure 1A, closed circles; online supplemental table 2). NPVs were higher for the one code algorithm, ranging from 78% for arthritis to 99% for dementia and delirium (figure 1B, open squares; online supplemental table 2). The one code algorithm was consistently more sensitive for detecting each of the 17 conditions of interest compared with two codes separated by more than 30 days (figure 1C, open squares; online supplemental table

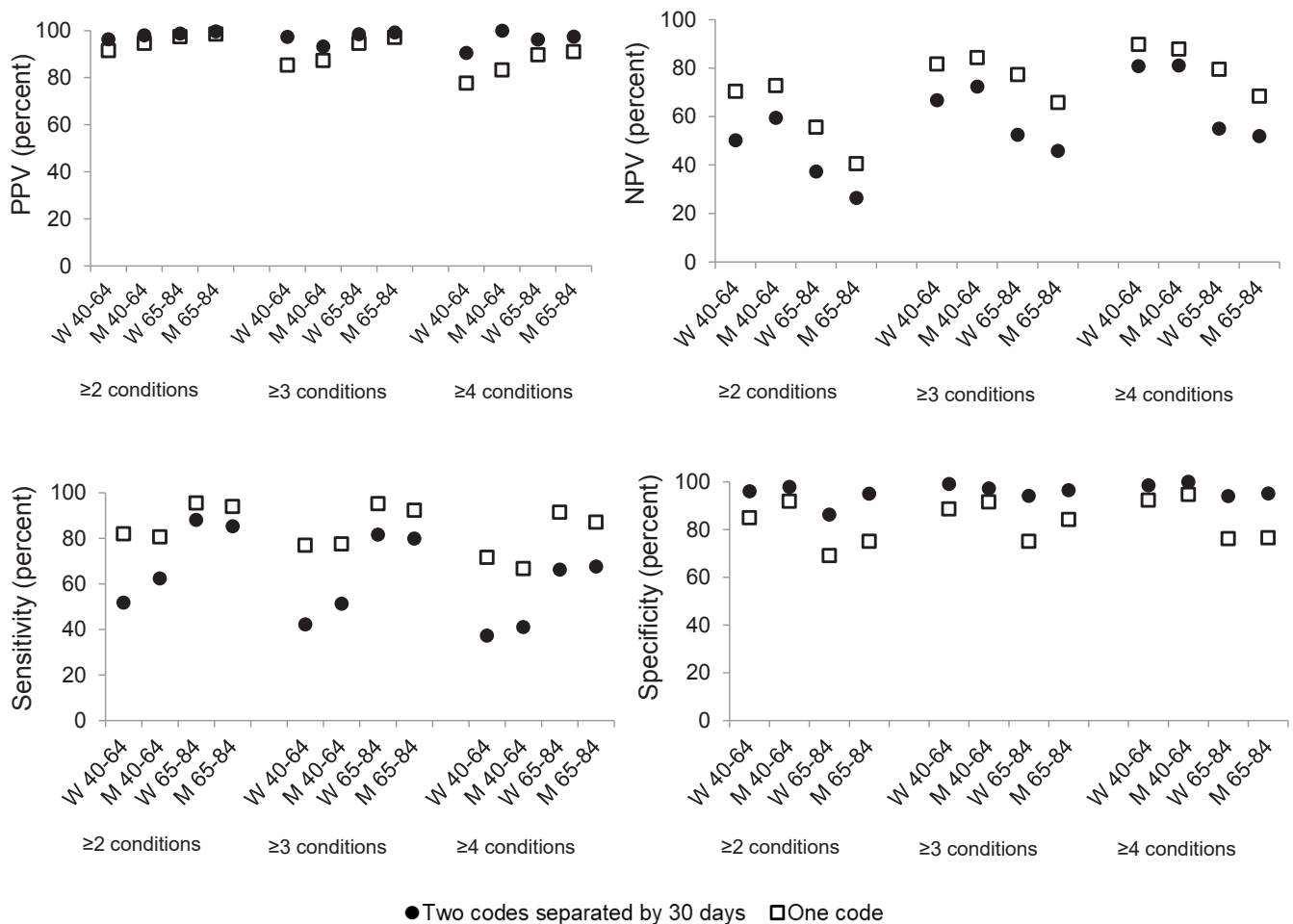


Figure 2 Positive predictive value (PPV; A), negative predictive value (NPV; B), sensitivity (C) and specificity (D) of two International Classification of Diseases, ninth edition (ICD-9) codes separated by >30 days or a single ICD-9 code for identifying the presence of multimorbidity. M, men; W, women.

3). The sensitivity of a single code ranged from 57% for stroke or transient ischaemic attack to 96% for hypertension. Finally, two codes separated by more than 30 days were more specific than a single code (figure 1D, closed circles; online supplemental table 3). However, specificity was generally high for all conditions (>90%) with use of a single code.

Multimorbidity

PPVs and specificity for the detection of multimorbidity were highest when two codes separated by more than 30 days were used (figure 2 and online supplemental table 4). By contrast, the sensitivity for detection of multimorbidity was highest when a single code definition was used (86% for presence of ≥ 2 conditions and 79% for ≥ 4 conditions; figure 2 and online supplemental table 4). Similarly, NPVs were highest when a single code was used (70% for ≥ 2 conditions and 87% for ≥ 4 conditions).

Overall, the use of one code to detect the individual conditions had high validity (PPV and sensitivity $\geq 70\%$) for multimorbidity defined as ≥ 2 or ≥ 3 conditions for all

age and sex strata (figure 3A,B, open symbols). For multimorbidity defined as ≥ 4 conditions, the use of a single code had high validity for all age and sex strata except men 40–64 years (figure 3C). The use of a single code had moderate validity (PPV $\geq 70\%$, but sensitivity <70%) for identifying ≥ 4 conditions in men 40–64 years (figure 3C, open blue circle). The use of single codes had NPVs and specificity $\geq 70\%$ in the overall cohort for ≥ 2 , ≥ 3 and ≥ 4 conditions (figure 4A–C, open black triangles). However, results varied by age and sex. For example, when defining multimorbidity as ≥ 2 conditions, both specificity and NPV were <70% for women 65–84 years (figure 4A, red open square). These results indicate that use of a single code may incorrectly classify older persons as having multimorbidity defined as ≥ 2 conditions when they do not actually have multiple chronic conditions. However, both specificity and NPV improve when multimorbidity is defined as ≥ 3 or ≥ 4 conditions (figure 4B,C, open circles and squares).

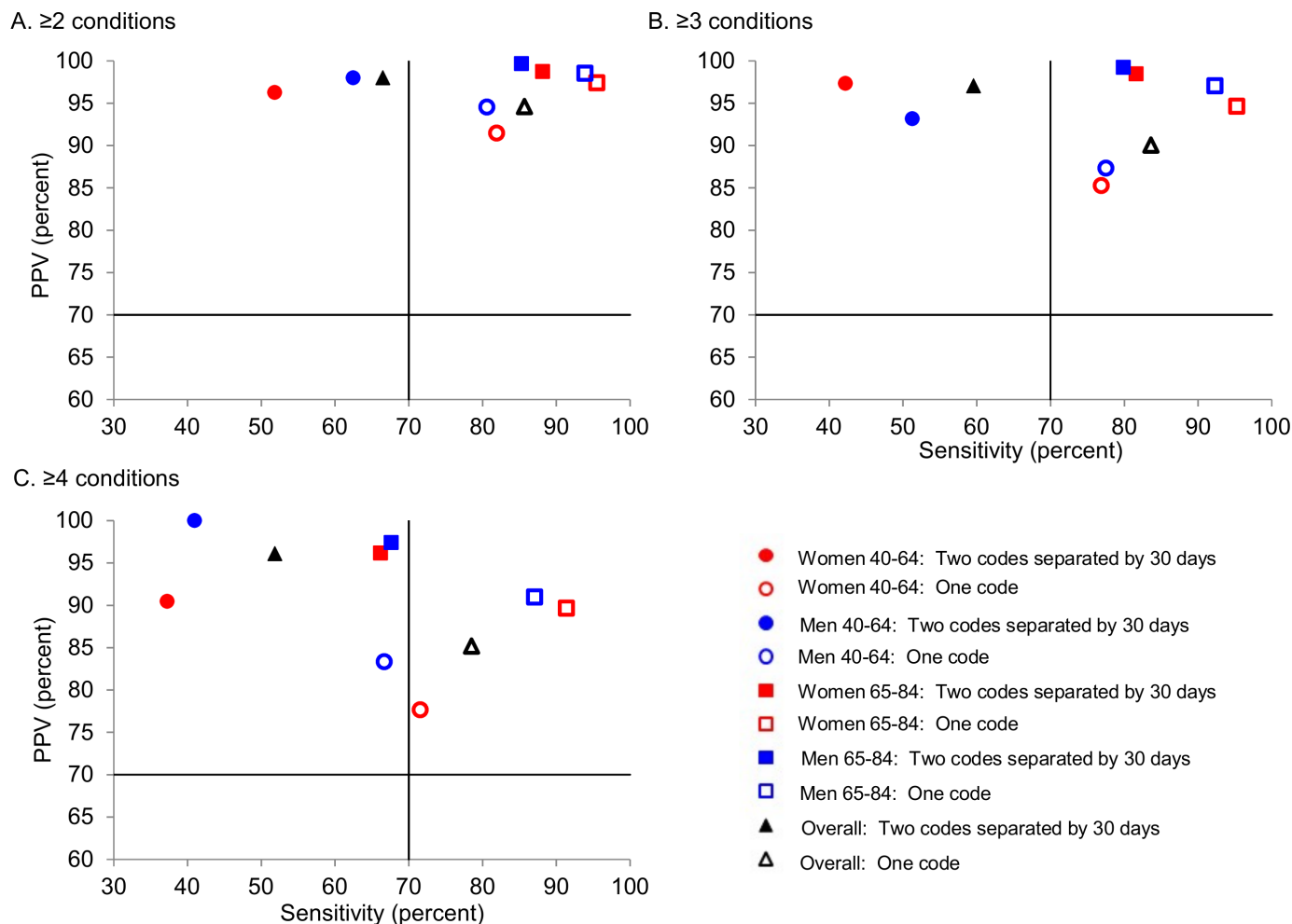


Figure 3 Positive predictive value (PPV) and sensitivity of two codes separated by >30 days compared with a single code definition for identifying persons with ≥ 2 (A), ≥ 3 (B) or ≥ 4 (C) conditions by age and sex. Shapes in the upper right quadrant indicate that PPV and sensitivity are both $\geq 70\%$.

DISCUSSION

We studied the PPV, NPV, sensitivity and specificity of using two ICD-9 code-based algorithms for identifying the 17 conditions included in our definition of multimorbidity. We found that the use of a single code resulted in high validity (sensitivity and PPV $\geq 70\%$) for defining multimorbidity, regardless of whether multimorbidity was defined as the presence of ≥ 2 , ≥ 3 or ≥ 4 conditions. In addition, use of a single code resulted in high validity in the opposite direction (NPV and specificity $\geq 70\%$) for the overall cohort, but these results varied by age and sex.

Our data are consistent with previous studies that have shown high variability in the ability of ICD-based algorithms to accurately identify the presence of a single chronic condition.¹⁷⁻¹⁹ Specificity and NPV tended to be high for all 17 conditions studied, regardless of whether a single code or two codes separated by more than 30 days was used. Using a single code was consistently more sensitive than two codes separated by more than 30 days across all age and sex strata. Conversely, two codes separated by more than 30 days had consistently higher PPVs compared with one code. However, none of our code-based algorithms were able to identify dementia and

delirium, or stroke and transient ischaemic attack with even moderate validity (PPV $\geq 70\%$, sensitivity $< 70\%$). Delirium and stroke are acute events, and both are known to be consistently undercoded.²⁹⁻³¹

By contrast, we found that the use of a single code to identify the individual conditions had high validity (PPV $\geq 70\%$, sensitivity $> 70\%$) for identifying persons with multimorbidity, regardless of whether multimorbidity was defined as ≥ 2 or ≥ 3 conditions. These results were consistent across all age and sex strata. For multimorbidity defined as ≥ 4 conditions, a single code definition had high validity for persons 65-84 years, and for women 40-64 years. The single code definition had moderate validity (PPV $\geq 70\%$ but sensitivity $< 70\%$) for men 40-64 years. These results indicate that the use of a single code for identification of the chronic conditions included in multimorbidity was the best of our two algorithms in persons 40-84 years of age.

Similarly, a single code had high validity for excluding multimorbidity in the overall cohort (NPV $\geq 70\%$, specificity $> 70\%$); however, results varied within age and sex strata. For example, when multimorbidity was defined as ≥ 2 conditions, NPV and specificity were moderate or low

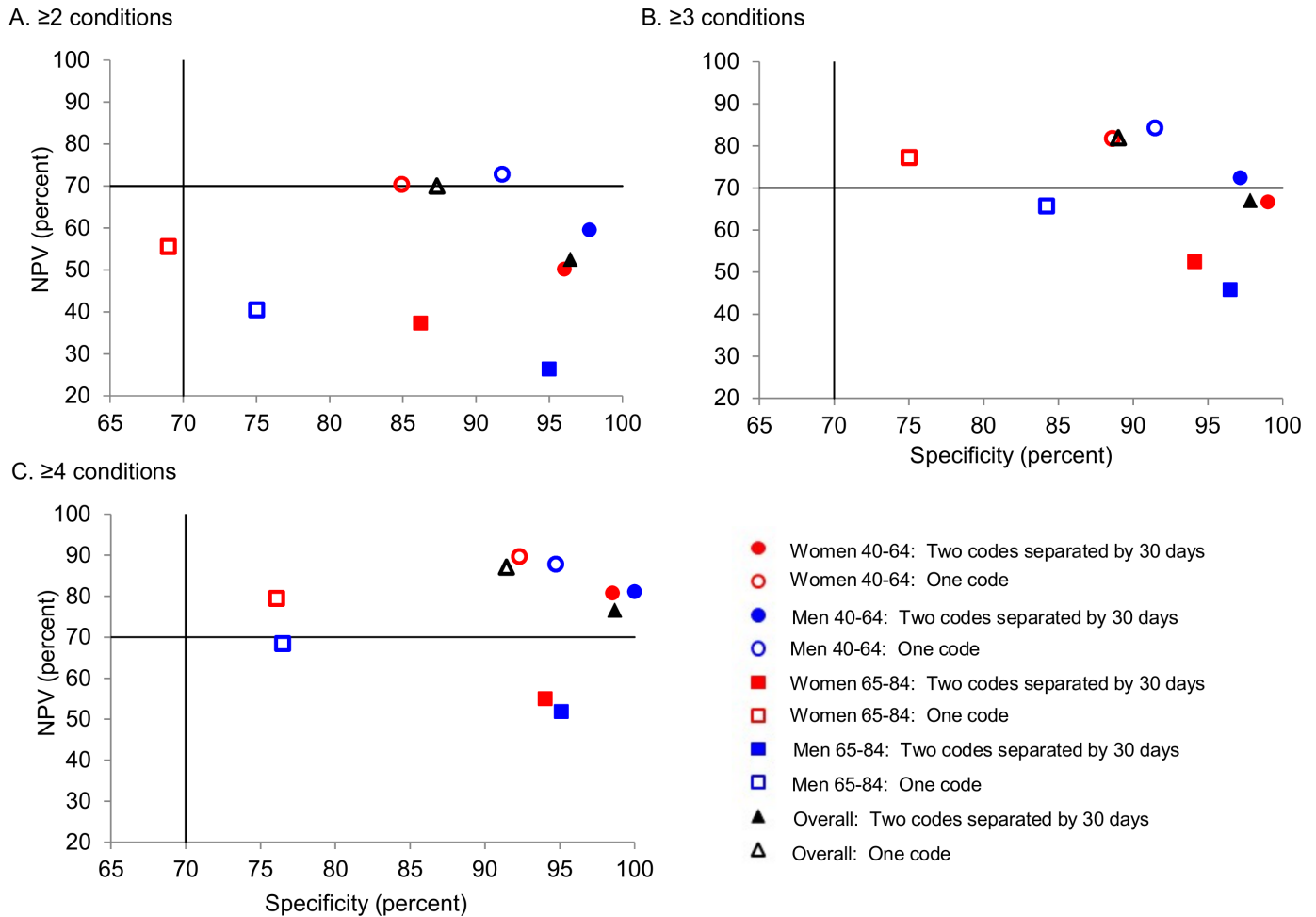


Figure 4 Negative predictive value (NPV) and specificity of two codes separated by >30 days compared with a single code definition for identifying persons with ≥ 2 (A), ≥ 3 (B) or ≥ 4 (C) conditions by age and sex. Shapes in the upper right quadrant indicate that NPV and specificity are both $\geq 70\%$.

for persons 65–84 years. These results suggest that the use of a single code algorithm for multimorbidity may incorrectly identify some older persons as having multimorbidity who do not have ≥ 2 conditions.

We note that the identification of persons with multimorbidity will depend on whether a person has two or more conditions from a defined list of possible conditions. Unfortunately, there is no universally accepted list of conditions to consider in studies of multimorbidity, and definitions of multimorbidity vary widely from study to study.^{11–16} There is therefore an urgent need to identify a minimum acceptable list of conditions to consider in studies of multimorbidity. In particular, it is important to understand whether persons who are considered multimorbid by one set of chronic conditions are also considered multimorbid using a different set of chronic conditions. Although we only considered 17 of the 20 conditions currently defined by the US DHHS for studies of multimorbidity, our study conclusions did not differ when we considered a smaller list of 10 chronic conditions defined by Drye and colleagues.³² Overall, a smaller proportion of persons were identified as having multimorbidity (two or more of the 10 conditions). However,

a single code still yielded the highest NPV (one code: 80.9%; two codes: 67.4%) and sensitivity (one code: 76.7%; two codes: 47.6%). Two codes were again the best for optimising PPV (one code: 87.5%; two codes: 96.8%) and specificity (one code: 90.0%; two codes: 98.6%).

Depending on the study objectives, investigators may wish to optimise some parameters over others. For example, when it is most important to identify all possible cases of multimorbidity, optimisation of sensitivity is appropriate (eg, prevalence, incidence or time trends studies). By contrast, in studies where more precise identification of multimorbidity is needed, optimisation of PPV would take priority (eg, case–control or cohort studies). Our study provides data to allow investigators to choose the implementation of the DHHS multimorbidity definition that best suits their study needs. Our study indicates that a single code algorithm for the multimorbidity definition proposed by Drye and colleagues would also be adequate.³²

Strengths of this study included our ability to validate the DHHS code sets against manual medical record review of a representative sample of the general population. In addition, the REP records-linkage system



provided access to virtually complete medical records for this population, and it is unlikely that we missed any of the 17 chronic conditions considered. As expected, the gold standard estimates of multimorbidity in this population were higher than what we have previously observed using only billing codes.²⁰ We would expect the number of conditions identified by a chart review of clinical notes to be higher than the number of conditions identified through billing data. Existing conditions may be briefly mentioned in a clinical note, but are not always assigned a billing code because they are not the focus of the clinical visit. For example, hyperlipidaemia that is controlled through statin use may be mentioned briefly in a clinical note, but may not receive a code for a visit that is scheduled to manage an acute episode of arrhythmia.

Limitations of our study include the fact that we validated only ICD-9 coding algorithms against medical record reviews, and further research is needed to specifically validate ICD-10 coding algorithms for multimorbidity. However, Quan and colleagues have found that ICD-9 and ICD-10 algorithms identified a similar proportion of persons with the comorbidities included in the Charlson and Elixhauser indexes using single codes from comparable code sets. These findings suggest that similar algorithms may work across both coding systems.¹⁷

Second, we included 'history of' conditions in our gold standard definition because we expected that most of the conditions considered in this study were unlikely to resolve. However, including 'history of' a condition may result in over-identification of prevalent cases, particularly for conditions that may become asymptomatic (eg, a previous stroke or cancer episode). To address this question, we examined the number of 'history of' chronic conditions that were not accompanied by a current treatment specific for that condition in our study population. Although the number of cases that might be excluded varied depending on the condition, for 15 of the conditions, <5% of the cases would have been excluded. Only history of hyperlipidaemia (7.6%) and history of arthritis (9.3%) had slightly more cases that did not include a current treatment for these conditions. Because the ICD-9 billing codes are less likely to capture historic events that have become asymptomatic, we expect that the exclusion of these cases from our gold standard definition would result in better sensitivity and PPV values. Therefore, our current results represent conservative estimates.

The data used in our study are from a limited number of healthcare providers in a single county of southern Minnesota, USA. Variability of coding practices in other institutions, in other regions of the USA, and in other countries may result in variation in the PPV, NPV, sensitivity and specificity of these DHHS code algorithms for identifying multimorbidity. Finally, our study population only included persons 40–84 years of age, and we observed variability in our measures by age and sex. PPV and sensitivity were highest in persons 65–84 years of age, whereas NPV and specificity were highest in persons 40–64 years. Because all of the conditions included in the

DHHS definition of multimorbidity are less prevalent in persons younger than 40 years, we expect that PPV and sensitivity would be lower, but that NPV and specificity would be higher.

We also note that women and older persons are more likely to visit their doctors, and more frequent visits may result in better documentation of health conditions in the medical record, and in a higher likelihood of receiving a billing code for any given condition. In our study population, both men and women 65–84 years had a median of 49 healthcare visits between 2006 and 2010 (IQR 28, 83 for men; IQR 32, 84 for women). By contrast, women 40–64 years had a median of 37 visits (IQR: 21, 57), and men had a median of 20 visits (IQR: 10, 35) in the same time frame. These differences in healthcare utilisation may account for some of the differences in validity that we observed across age and sex. In particular, sensitivity of codes was best for older persons. Therefore, our results may not apply to younger populations, or to populations with different socioeconomic characteristics. Social determinants of health may affect the use of healthcare services, the likelihood of diagnosis and the assignment of billing codes. In addition, differences across countries may limit the generalisability of these data to populations with other healthcare systems (eg, USA vs Canada).

CONCLUSIONS

The ability of a single ICD-9 code algorithm to accurately identify 17 chronic conditions in the general population varied by condition, and by age and sex. However, the use of a single ICD-9 code for each condition had high validity (high sensitivity, PPV, specificity and NPV) for identifying persons with multimorbidity in this population. Therefore, the use of a single code algorithm may represent the simplest way to identify persons with multimorbidity in studies that use the multimorbidity definition proposed by DHHS, and in populations with similar demographic, socioeconomic, and health care characteristics.

Author affiliations

¹Division of Epidemiology, Department of Quantitative Health Sciences, Mayo Clinic, Rochester, Minnesota, USA

²The Robert D. and Patricia E. Kern Center for the Science of Health Care Delivery, Mayo Clinic, Rochester, Minnesota, USA

³Department of Psychiatry and Psychology, Mayo Clinic, Jacksonville, Florida, USA

⁴Division of Geriatric Medicine and Gerontology, Department of Medicine, Johns Hopkins University, Baltimore, Maryland, USA

⁵Division of Clinical Trials and Biostatistics, Department of Quantitative Health Sciences, Mayo Clinic, Rochester, Minnesota, USA

⁶Department of Neurology, Mayo Clinic, Rochester, Minnesota, USA

⁷Women's Health Research Center, Mayo Clinic, Rochester, MN, USA

Acknowledgements The authors thank Ms Connie Fortner for reviewing the medical records of the study participants, and Ms Kristi Klinger for manuscript formatting and preparation.

Contributors JLS, AMC, WVB, CMB, LJFR, DJJ, BRG and WAR participated in the conception of the study. JLS, AMC, DJJ and WAR designed the study and supervised data collection. DJJ and MEM cleaned the data, and performed all study analyses. JLS, AMC, WVB, CMB, LJFR, DJJ, MEM, BRG and WAR reviewed the study results and participated in the data interpretation process. JLS drafted the manuscript and

AMC, WV, CMB, LJFR, DJJ, MEM, BRG and WAR critically revised the manuscript. All authors read and approved the final draft.

Funding This study was supported by grants from the National Institute on Aging (R01 AG034676 and R01 AG052425).

Competing interests None declared.

Patient consent for publication Obtained.

Ethics approval This study was approved by the Mayo Clinic and Olmsted Medical Center Institutional Review Boards.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request.

Data sets used and analysed during the current study are available from the corresponding author, with the enactment of appropriate data use agreements that comply with HIPAA regulations.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Jennifer L St Sauver <http://orcid.org/0000-0002-9789-8544>

Walter A Rocca <http://orcid.org/0000-0002-1832-7664>

REFERENCES

- Goodman RA, Posner SF, Huang ES, *et al*. Defining and measuring chronic conditions: imperatives for research, policy, program, and practice. *Prev Chronic Dis* 2013;10:E66.
- Academy of Medical Sciences. *Multimorbidity: a priority for global health research*. London, UK: Academy of Medical Sciences, 2018: 1–127.
- Salive ME. Multimorbidity in older adults. *Epidemiol Rev* 2013;35:75–83.
- Wolff JL, Starfield B, Anderson G. Prevalence, expenditures, and complications of multiple chronic conditions in the elderly. *Arch Intern Med* 2002;162:2269–76.
- 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 Suppl 3:391–5.
- Chen H-Y, Baumgardner DJ, Rice JP. Health-Related quality of life among adults with multiple chronic conditions in the United States, behavioral risk factor surveillance system, 2007. *Prev Chronic Dis* 2011;8:A09.
- Barnett K, Mercer SW, Norbury M, *et al*. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. *Lancet* 2012;380:37–43.
- Frolich A, Ghith N, Schiøtz M, *et al*. Multimorbidity, healthcare utilization and socioeconomic status: a register-based study in Denmark. *PLoS One* 2019;14:e0214183.
- Lujic S, Simpson JM, Zwar N, *et al*. Multimorbidity in Australia: comparing estimates derived using administrative data sources and survey data. *PLoS One* 2017;12:e0183817.
- Nicholson K, Terry AL, Fortin M, *et al*. Prevalence, characteristics, and patterns of patients with multimorbidity in primary care: a retrospective cohort analysis in Canada. *Br J Gen Pract* 2019;69:e647–56.
- Diederichs C, Berger K, Bartels DB. The measurement of multiple chronic diseases—a systematic review on existing multimorbidity indices. *J Gerontol A Biol Sci Med Sci* 2011;66:301–11.
- Fortin M, Hudon C, Dubois M-F, *et al*. Comparative assessment of three different indices of multimorbidity for studies on health-related quality of life. *Health Qual Life Outcomes* 2005;3:74.
- Willadsen TG, Bebe A, Køster-Rasmussen R, *et al*. The role of diseases, risk factors and symptoms in the definition of multimorbidity - a systematic review. *Scand J Prim Health Care* 2016;34:112–21.
- Le Reste JY, Nabbe P, Manceau B, *et al*. The European general practice research network presents a comprehensive definition of multimorbidity in family medicine and long term care, following a systematic review of relevant literature. *J Am Med Dir Assoc* 2013;14:319–25.
- Pati S, Swain S, Hussain MA, *et al*. Prevalence and outcomes of multimorbidity in South Asia: a systematic review. *BMJ Open* 2015;5:e007235.
- Xu X, Mishra GD, Jones M. Evidence on multimorbidity from definition to intervention: an overview of systematic reviews. *Ageing Res Rev* 2017;37:53–68.
- Quan H, Li B, Saunders LD, *et al*. Assessing validity of ICD-9-CM and ICD-10 administrative data in recording clinical conditions in a unique dually coded database. *Health Serv Res* 2008;43:1424–41.
- Wilkinson T, Ly A, Schnier C, *et al*. Identifying dementia cases with routinely collected health data: A systematic review. *Alzheimers Dement* 2018;14:1038–51.
- Jiang J, Southern D, Beck CA, *et al*. Validity of Canadian discharge Abstract data for hypertension and diabetes from 2002 to 2013. *CMAJ Open* 2016;4:E646–53.
- Rocca WA, Boyd CM, Grossardt BR, *et al*. Prevalence of multimorbidity in a geographically defined American population: patterns by age, sex, and race/ethnicity. *Mayo Clin Proc* 2014;89:1336–49.
- St Sauver JL, Boyd CM, Grossardt BR, *et al*. Risk of developing multimorbidity across all ages in an historical cohort study: differences by sex and ethnicity. *BMJ Open* 2015;5:e006413.
- Lochner KA, Cox CS. Prevalence of multiple chronic conditions among Medicare beneficiaries, United States, 2010. *Prev Chronic Dis* 2013;10:E61.
- Tonelli M, Wiebe N, Fortin M, *et al*. Methods for identifying 30 chronic conditions: application to administrative data. *BMC Med Inform Decis Mak* 2015;15:31.
- St Sauver JL, Grossardt BR, Yawn BP, *et al*. Use of a medical records linkage system to enumerate a dynamic population over time: the Rochester epidemiology project. *Am J Epidemiol* 2011;173:1059–68.
- St Sauver JL, Grossardt BR, Yawn BP, *et al*. Data resource profile: the Rochester epidemiology project (Rep) medical records-linkage system. *Int J Epidemiol* 2012;41:1614–24.
- Rocca WA, Yawn BP, St Sauver JL, *et al*. History of the Rochester epidemiology project: half a century of medical records linkage in a US population. *Mayo Clin Proc* 2012;87:1202–13.
- Rocca WA, Grossardt BR, Brue SM, *et al*. Data resource profile: expansion of the Rochester epidemiology project medical records-linkage system (E-REP). *Int J Epidemiol* 2018;47:368–368j.
- Porta M. *A dictionary of epidemiology*. 6th edn. New York: Oxford University Press, 2014.
- Inouye SK, Leo-Summers L, Zhang Y, *et al*. A chart-based method for identification of delirium: validation compared with interviewer ratings using the confusion assessment method. *J Am Geriatr Soc* 2005;53:312–8.
- Benesch C, Witter DM, Wilder AL, *et al*. Inaccuracy of the International classification of diseases (ICD-9-CM) in identifying the diagnosis of ischemic cerebrovascular disease. *Neurology* 1997;49:660–4.
- Leibson CL, Naessens JM, Brown RD, *et al*. Accuracy of hospital discharge Abstracts for identifying stroke. *Stroke* 1994;25:2348–55.
- Drye EE, Altaf FK, Lipska KJ, *et al*. Defining multiple chronic conditions for quality measurement. *Med Care* 2018;56:193–201.